Xintong Li
xintong.li1@uwaterloo.ca

Dept. Statistics and Actuarial Science
University of Waterloo

Efficient Machine Learning Approaches for Fast Risk Evaluation of VAs

Supervised by Prof. Tony Wirjanto and Prof. Mingbin Feng

Thesis Defense, University of Waterloo

Outline

Nested Simulation Procedures

Nested simulation procedures are necessary for **complex** financial derivatives and insurance products.

$$\rho(L) = \rho(L(X)), \quad L(X) = \mathbb{E}\left[Y|X=x\right]|_{x=X}$$

Involves two levels of Monte Carlo simulations:

- Outer level: generates underlying risk factors (outer scenarios), $X_i \sim F_X$
- Inner level: generates scenario-wise samples of portfolio losses (inner replications), $Y_{ij} \sim F_{Y|X_i}$

Computationally expensive due to its nested structure.

Common Risk Measures

- Smooth h, e.g., quadratic tracking error

$$\rho(L) = \mathbb{E}\left[(L-b)^2\right]$$

- hockey-stick h: mean excess loss

$$\rho(L) = \mathbb{E}\left[L \cdot \mathbb{1}_{\{L \geq u\}}\right]$$

- indicator h: probability of large loss

$$\rho(L) = \mathbb{E}\left[\mathbb{1}_{\{L \geq u\}}\right]$$

- Value at Risk (VaR)

$$\rho_\alpha(L) = Q_\alpha(L) = \inf\{u : \mathbb{P}(L \leq u) \geq \alpha\}$$

- Conditional Value at Risk (CVaR) [1]

$$\rho_\alpha(L) = \mathbb{E}\left[L | L \geq Q_\alpha(L)\right]$$

[1] Note: If $Q_\alpha(L)$ falls in a probability mass, $\rho(L) = \frac{(\beta-\alpha)Q_\alpha(L)+(1-\beta)\mathbb{E}[L|L\geq Q_\alpha(L)]}{1-\alpha}$.

Standard Nested Simulation

$$\hat{L}_{N,i} = \frac{1}{N}\sum_{j=1}^{N} Y_{ij}; \quad Y_{ij} \sim F_{Y|X_i}$$

- Uses inner sample mean to estimate $L(X_i)$.
- Proposed by Gordy and Juneja (2010); finds optimal growth order of M and N.
- Zhang et al. (2021) estimate the optimal M and N using a bootstrap method.
- Computationally expensive and potentially **wasteful** use of budget.

Other Nested Simulation Procedures

Subsequent works focus on improving the efficiency of nested simulation:

- Regression-based (Broadie et al., 2015)
- Kernel smoothing (Hong et al., 2017)
- Likelihood ratio (Feng and Song, 2020)
- Kernel ridge regression (Zhang et al., 2022)

Key ideas:

- Pool inner replications from different outer scenarios
- Use metamodeling techniques to approximate the inner simulation model

Metamodeling Approach

In this thesis, we focus on procedures that use **supervised learning metamodels** to approximate the inner simulation model.

- Treat the inner simulation as a black-box function
- Approximate $L(\cdot)$ with $\hat{L}^{\mathsf{SL}}_{M,N}(\cdot)$
- Train with a set of feature-label pairs generated from the standard procedure:

$$\{(X_i, \hat{L}_{N,i}) | i = 1, \ldots, M, j = 1, \ldots, N\}$$

- Use trained metamodel to make predictions for all $X \in \mathcal{X}$

There are **computational costs** associated with pooling inner replications.

Problem Statement

Minimize mean squared error (MSE) of the estimator subject to total simulation budget:

$$\min_{M,N} \quad \mathbb{E}\left[(\hat{\rho}_{M,N} - \rho)^2\right]$$
$$\text{subject to } M \cdot N = \Gamma$$

Interested in convergence order as $\Gamma \to \infty$

Asymptotic Convergence Rates of Different Procedures

Procedures	**Smooth** h	**Hockey-Stick** h	**Indicator** h
Standard Procedure	$\mathcal{O}(\Gamma^{-2/3})$	$\mathcal{O}(\Gamma^{-2/3})$	$\mathcal{O}(\Gamma^{-2/3})$
Regression	$\mathcal{O}(\Gamma^{-1})$	$\mathcal{O}(\Gamma^{-1+\delta})$	No Result
Kernel Smoothing	$\mathcal{O}(\Gamma^{-\min(1,4/(d+2))})$		
Kernel Ridge Regression	$\mathcal{O}(\Gamma^{-1})$		
Likelihood Ratio	$\mathcal{O}(\Gamma^{-1})$		

- We show the asymptotic convergence rates of the standard procedure for smooth and hockey-stick h.
- Only kernel smoothing depends on the asset dimension d.

Key Theoretical Results

Observations:

- Most literature focuses on the MSE of $\hat{\rho}$.
- Wang et al. (2022) analyze convergence of absolute error in probabilistic order.

Contribution: bridging the gap between MSE and absolute error convergence.

- Convergence in MSE:

$$\mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2\right] = \mathcal{O}\left(\Gamma^{-\xi}\right)$$

- Convergence in Probabilistic Order:

$$|\hat{\rho}_\Gamma - \rho| = \mathcal{O}_{\mathbb{P}}(\Gamma^{-\xi})$$

Key Theoretical Results

Theorem

If $\hat{\rho}_\Gamma$ converges in MSE to ρ in order ξ, then $\hat{\rho}_\Gamma$ converges in probabilistic order to ρ in order $\frac{\xi}{2}$.

- First result to draw connection between MSE and probabilistic order convergence.
- Applicable to any nested simulation procedure.
- Convergence in MSE implies convergence in probabilistic order.

Experiment Design

We compare 5 nested simulation procedures

- Standard nested simulation
- Regression-based
- Kernel smoothing
- Likelihood ratio
- Kernel ridge regression

And their empirical convergence stable across different:

- Risk measures
- Option types
- Asset dimensions
- Asset models (GBM vs. Heston)
- Regression bases (only for the regression-based procedure)

Finite-Sample Performance

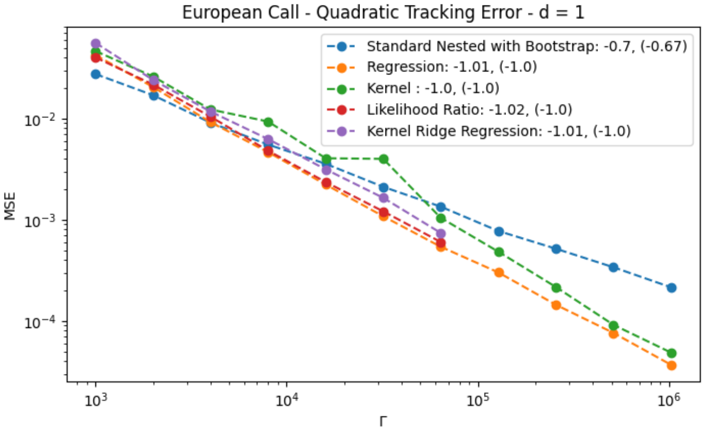
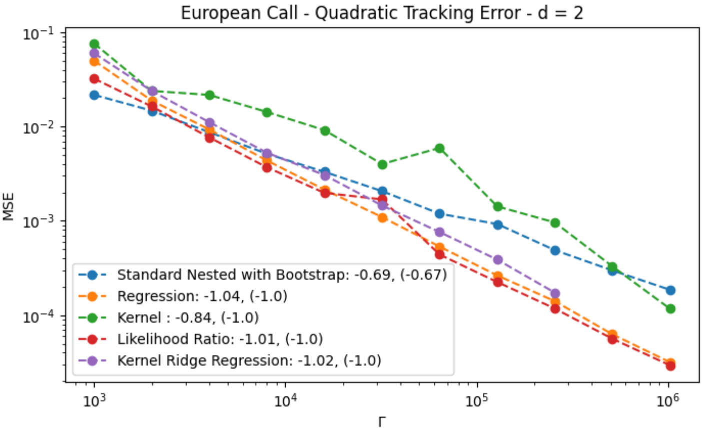


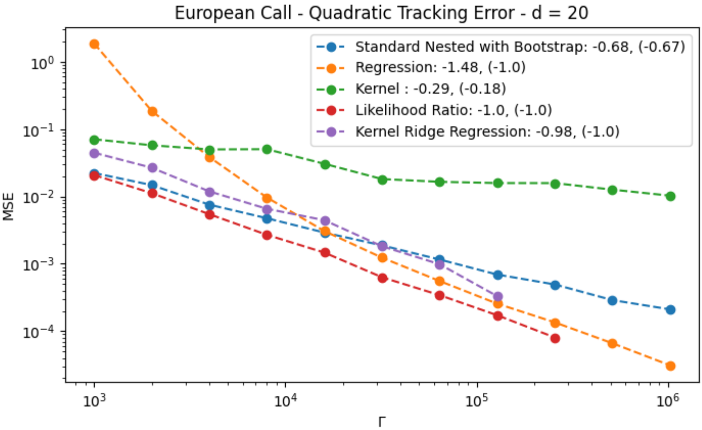
Figure: Empirical convergence rates of different procedures for the base case

Sensitivity to Asset Dimension



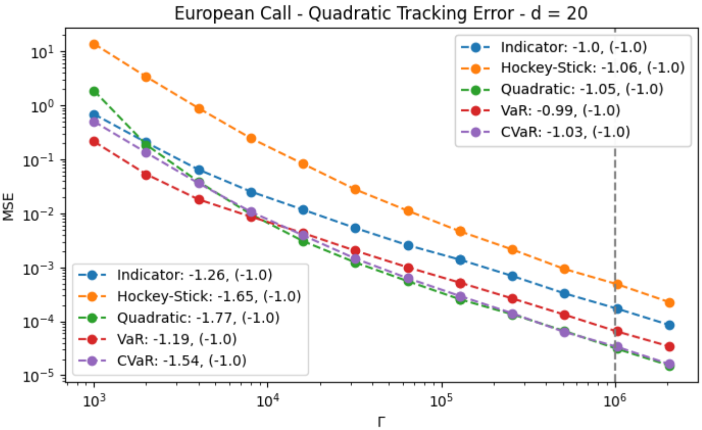
- Standard, KRR, and likelihood ratio procedures are dimension-independent
- Kernel smoothing and regression show sensitivity to dimension, but in different ways

Sensitivity to Asset Dimension



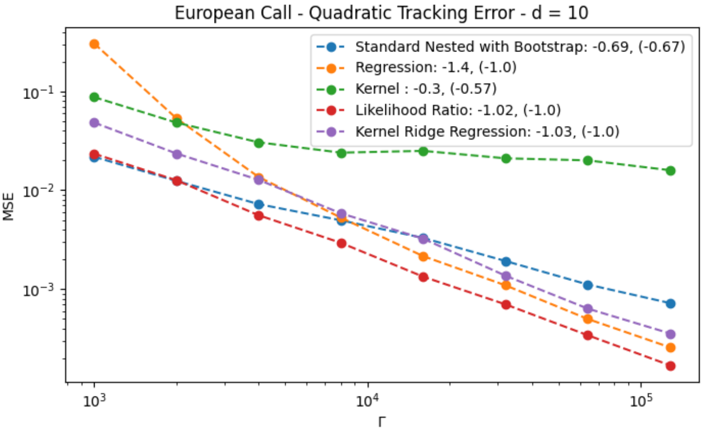
- Standard, KRR, and likelihood ratio procedures are dimension-independent
- Kernel smoothing and regression show sensitivity to dimension, but in different ways

Fast Convergence of Regression-based Procedure



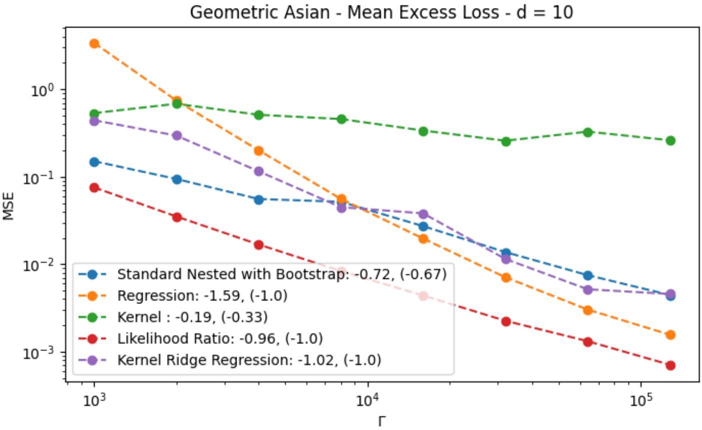
- Higher initial convergence rate
- Stabilizes to match asymptotic rate at higher budgets
- Consistent across different asset dimensions

Sensitivity to Option Type



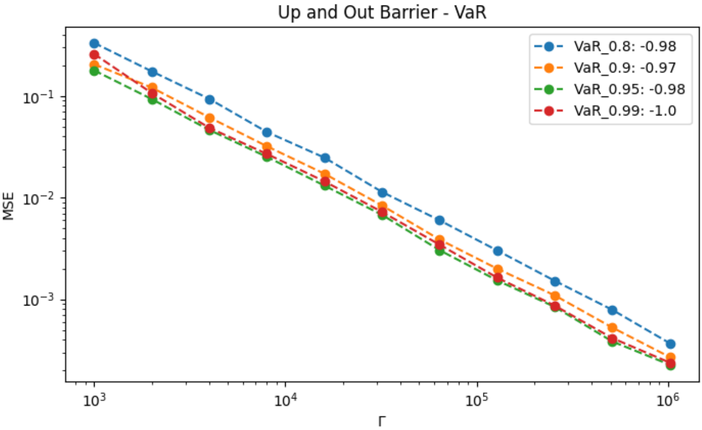
- Similar convergence patterns across different option types
- Regression and kernel smoothing show higher empirical rates for barrier options

Sensitivity to Risk Measure



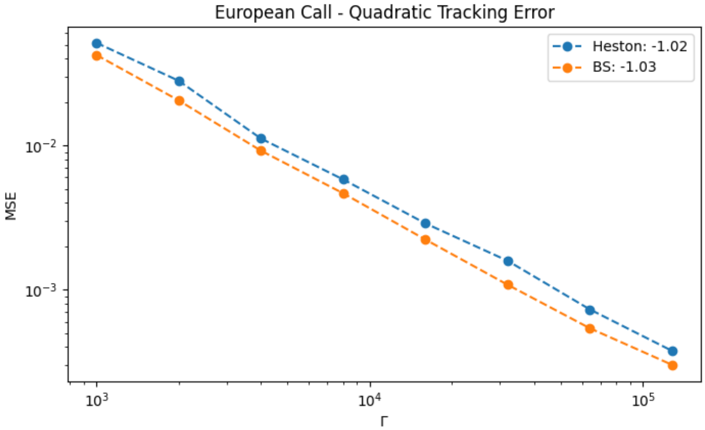
- Convergence behavior consistent across different risk measures
- Regression-based method shows highest empirical convergence rates

Sensitivity to VaR/CVaR Level



- Regression-based method not sensitive to VaR/CVaR level
- Consistent performance across different levels

Sensitivity to Asset Model



- Regression-based method insensitive to asset model (GBM vs. Heston)
- Consistent performance across different asset models

Computational Complexity

There are **computational costs** associated with pooling inner replications.

- Standard procedure: cost of estimating the optimal M and N
- Regression: most efficient among metamodel-based procedures
- Kernel smoothing: costly distance calculations and cross-validation
- Likelihood ratio: No training, but costly weight calculations
- KRR: even more expensive than kernel smoothing

Total Computation Time

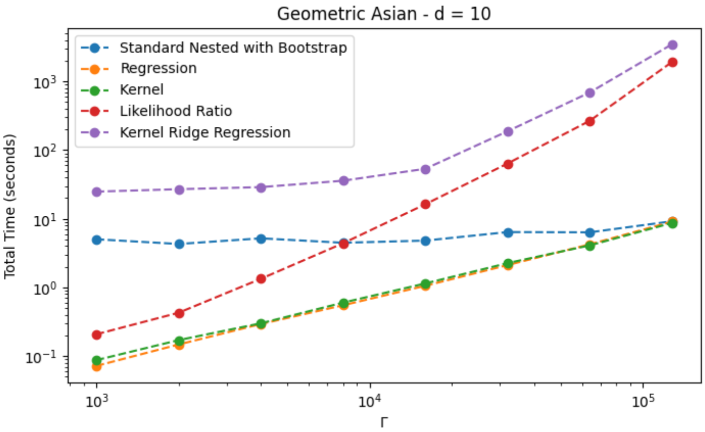


Figure: Total computation time for different procedures

Cost of Hyperparameter Tuning

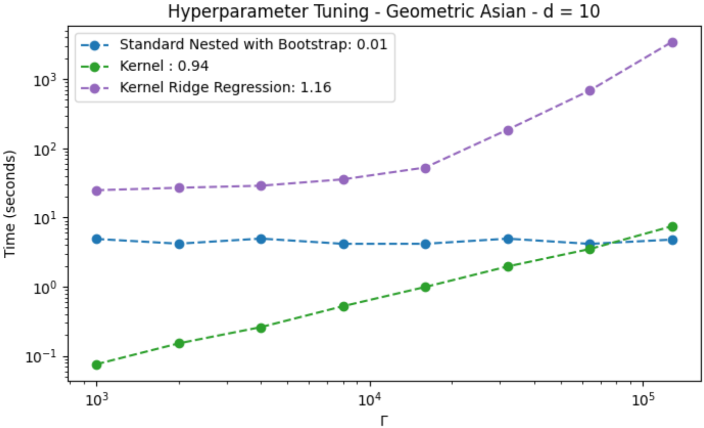


Figure: Cost of hyperparameter tuning for different procedures

Cost of Model Fitting and Validation

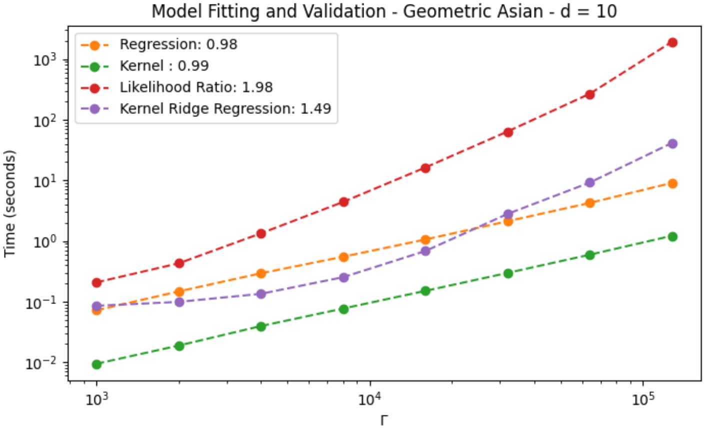


Figure: Cost of model fitting and validation for different procedures

Conclusion

Regression-based nested simulation procedure:

- Most robust and stable for limited budgets
- Efficient to implement
- Fast empirical convergence for option portfolios

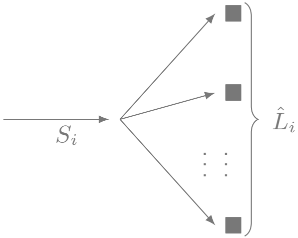
For high-dimensional or complex payoffs:

- Difficult to find a good regression basis
- Neural network-based procedures may be more suitable

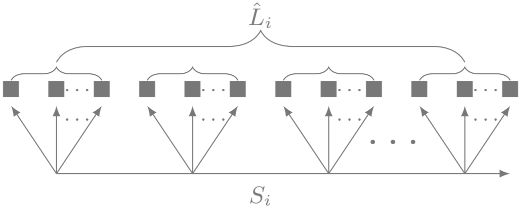
Next project: examining performance of metamodel-based simulation procedures for variable annuities

From Options to Variable Annuities

Variable annuities (VAs) poses a challenge for nested simulation due to its **high-dimensional** and **complex payoff structure**.



1 Outer Path for Options



1 Outer Path for VAs

- Need to reconstruct a metamodeling-based nested simulation procedure

Nested Simulation for Risk Management of VAs

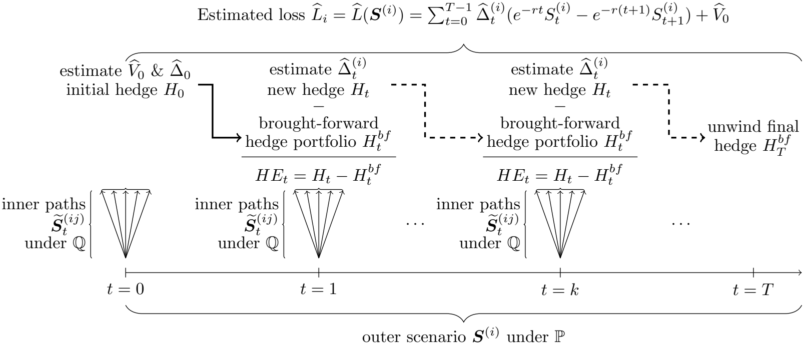


Figure: Illustration of nested simulation that estimates the P&L for one outer scenario

Standard Nested Simulation for VAs

Standard nested simulation for VAs is similar to the one for options.

- Generate M outer scenarios
- For each outer scenario,
  - Perform N inner simulations
  - Estimate hedging loss L_i with $\hat{L}_i$
- Use estimated losses to calculate tail risk measures (e.g., 95%-CVaR)

Main challenge:

- Computational budget is limited.
- Accuracy of the estimation depends on both M and N.

Two-Stage LSTM-based Nested Simulation for VAs

References

Broadie, M., Du, Y., and Moallemi, C. C. (2015). Risk estimation via regression. *Operations Research*, 63(5):1077–1097.

Feng, M. and Song, E. (2020). Optimal nested simulation experiment design via likelihood ratio method. *arXiv preprint arXiv:2008.13087*.

Gordy, M. B. and Juneja, S. (2010). Nested simulation in portfolio risk measurement. *Management Science*, 56(10):1833–1848.

Hong, J. L., Juneja, S., and Liu, G. (2017). Kernel smoothing for nested estimation with application to portfolio risk measurement. *Operations Research*, 65(3):657–673.

Wang, W., Wang, Y., and Zhang, X. (2022). Smooth nested simulation: bridging cubic and square root convergence rates in high dimensions. *arXiv preprint arXiv:2201.02958*.

Zhang, K., Feng, M., Liu, G., and Wang, S. (2022). Sample recycling for nested simulation with application in portfolio risk measurement. *arXiv preprint arXiv:2203.15929*.

Zhang, K., Liu, G., and Wang, S. (2021). Bootstrap-based budget allocation for nested simulation. *Operations Research*.