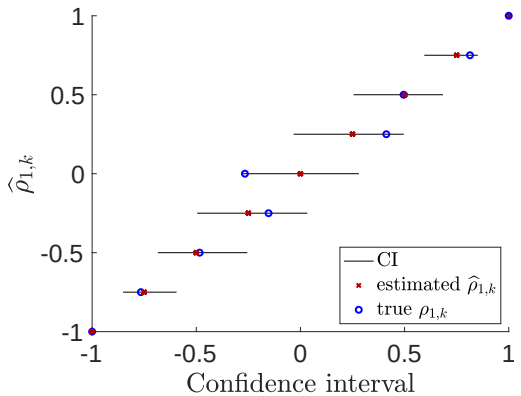
Multi-fidelity Monte Carlo

Dr. Jiaxing Liang (Rice University)
Dr. Matthias Heinkenschloss (Rice University)

May 26, 2025

Comparison between $\widehat{\rho}_{1,k}$ and $\rho_{1,k}$



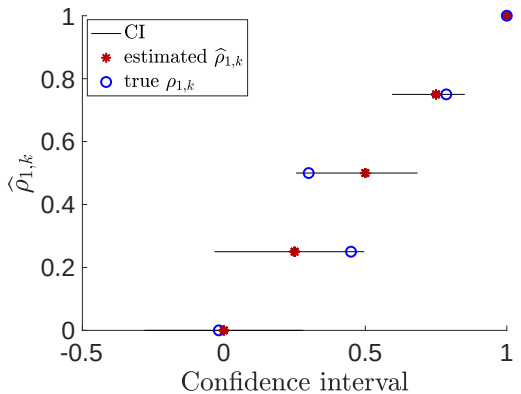
- Confidence intervals are generated with pilot sample size $Q = 10$ and $z_{\alpha/2} = 1.96$
- $\epsilon = 10^{-2}$, $\sigma_1 = 10^{-1}$,

Cost per sample	[100, 10, 1, 1000, 1, 1, 1, 1, 100]	
	true $\rho_{1,k}$	estimate $\widehat{\rho}_{1,k}$
$\rho_{1,k}$	[-1, -0.650, -0.464, -0.402, -0.057, 0.038, 0.268, 0.836, 1]	[-1,-0.75,-0.5,-0.25,0,0.25,0.5,0.75,1]
Model selection (model index)	[1,8]	[1,2,3]
Sample size before ceil	[34.73, 528.74]	[58.75, 157.02, 444.11]
Variance before ceil	1e-4	1e-4
Sampling cost before ceil	4.00e+03	7.89e+03
Sample size after ceil	[35, 529]	[59, 158, 445]
Variance after ceil	9.9331e-05	9.9549e-05
Sampling cost after ceil	4029	7925

Observation

- **Observation:**
  - Sampling cost of $\widehat{\rho}_{1,k}$ is **higher** than that of $\rho_{1,k}$. This is due to the fact that different sets of models are selected and the difference between the true $\rho_{1,k}$ and the estimated $\widehat{\rho}_{1,k}$.
  - Overlap of confidence intervals, $\widehat{\rho}_{1,k}$ has monotonic trend, whereas the true $\rho_{1,k}$ does not have this behavior. This mismatch influences model selection and cost estimation.
  - The variance constraint is satisfied exactly when using real-valued sample sizes. When rounding up to the nearest integer, the resulting variance is slightly below the target tolerance.
  - Rounding sample sizes to integers has minimal effect on both variance and cost. In contrast, poor approximation of $\rho_{1,k}$ can lead to significant increases in sampling cost (e.g., from $\sim$4k to $\sim$8k).

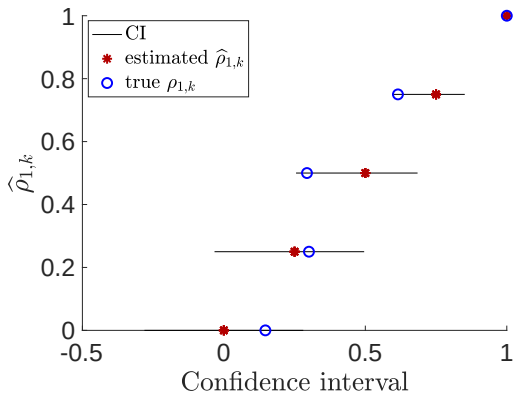
Comparison between $\widehat{\rho}_{1,k}$ and $\rho_{1,k}$



Cost per sample	[10, 1, 1, 1000, 100]	
	true $\rho_{1,k}$	estimate $\widehat{\rho}_{1,k}$
$\rho_{1,k}$	[-1.7700e-02, 4.4940e-01, 2.9981e-01, 7.8605e-01, 1]	[0,0.25,0.5,0.75,1]
Model selection (model index)	[5,3]	[5,2]
Sample size before ceil	[83.82, 421.66]	[79.33, 458.01]
Variance before ceil	1e-4	1e-4
Sampling cost before ceil	8.8035e+03	8.3910e+03
Sample size after ceil	[84, 422]	[80, 459]
Variance after ceil	9.9790e-05	9.9197e-05
Sampling cost after ceil	8822	8459

- **Observation:**
  - Sampling cost of $\widehat{\rho}_{1,k}$ is **smaller** than that of $\rho_{1,k}$. Different models are selected.

Comparison between $\widehat{\rho}_{1,k}$ and $\rho_{1,k}$



Cost per sample	[10, 1, 1, 1000, 100]	
	true $\rho_{1,k}$	estimate $\widehat{\rho}_{1,k}$
$\rho_{1,k}$	[1.4674e-01, 3.0076e-01, 2.9375e-01, 6.1525e-01, 1]	[0,0.25,0.5,0.75,1]
Model selection (model index)	[5,4]	[5,4]
Sample size before ceil	[67.00, 522.88]	[48.71, 552.33]
Variance before ceil	1e-4	1e-4
Sampling cost before ceil	7.2226e+03	5.4234e+03
Sample size after ceil	[67, 523]	[49, 553]
Variance after ceil	9.9994e-05	9.9458e-05
Sampling cost after ceil	7223	5453

- **Observation:**
  - Sampling cost of $\widehat{\rho}_{1,k}$ is **smaller** than that of $\rho_{1,k}$, even though same models are selected, the difference is due to the discrepancy in correlation coefficients.

Numerical results

Thank You |¬_¬|