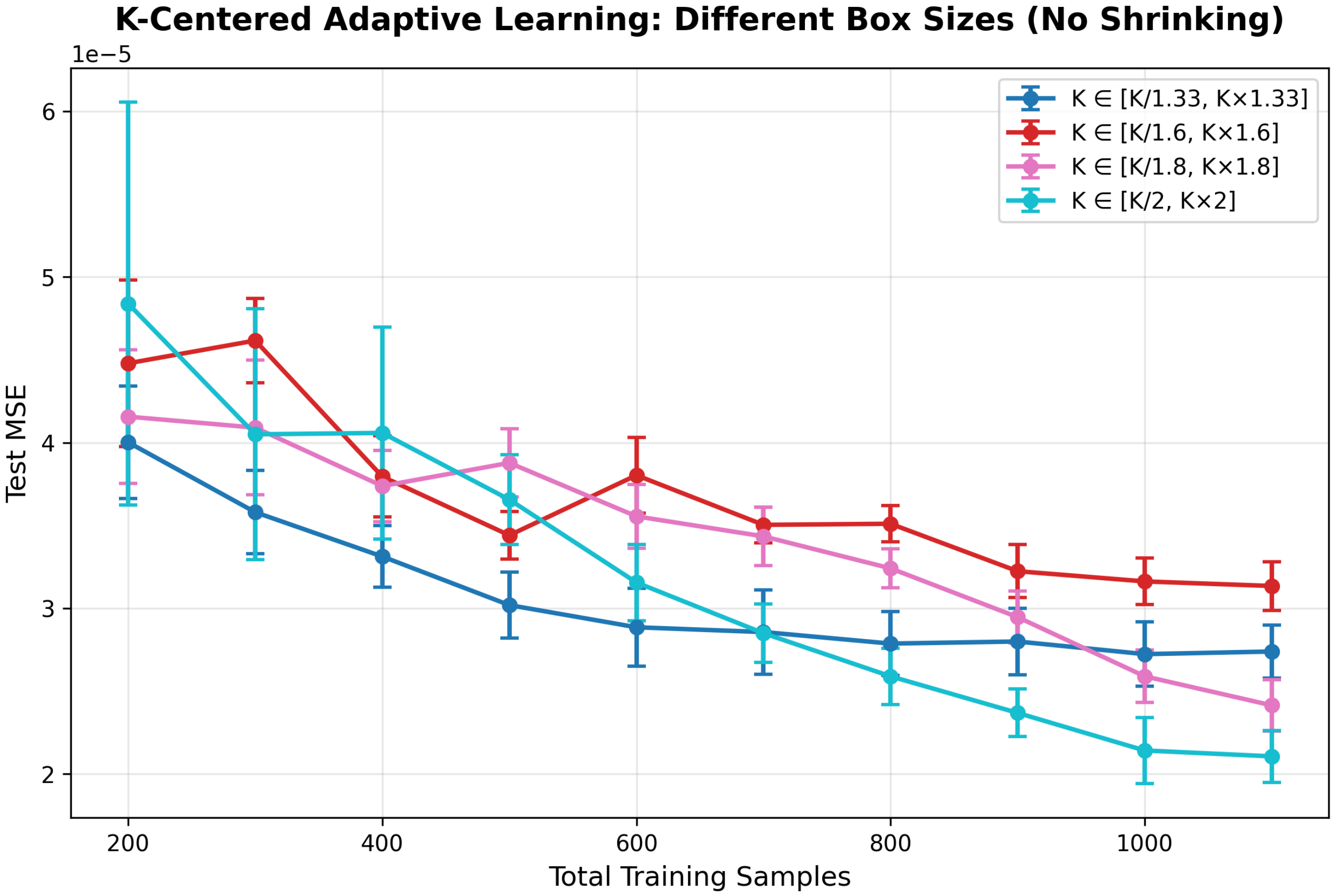
Adaptive Learning Strategy - Morris Sampling

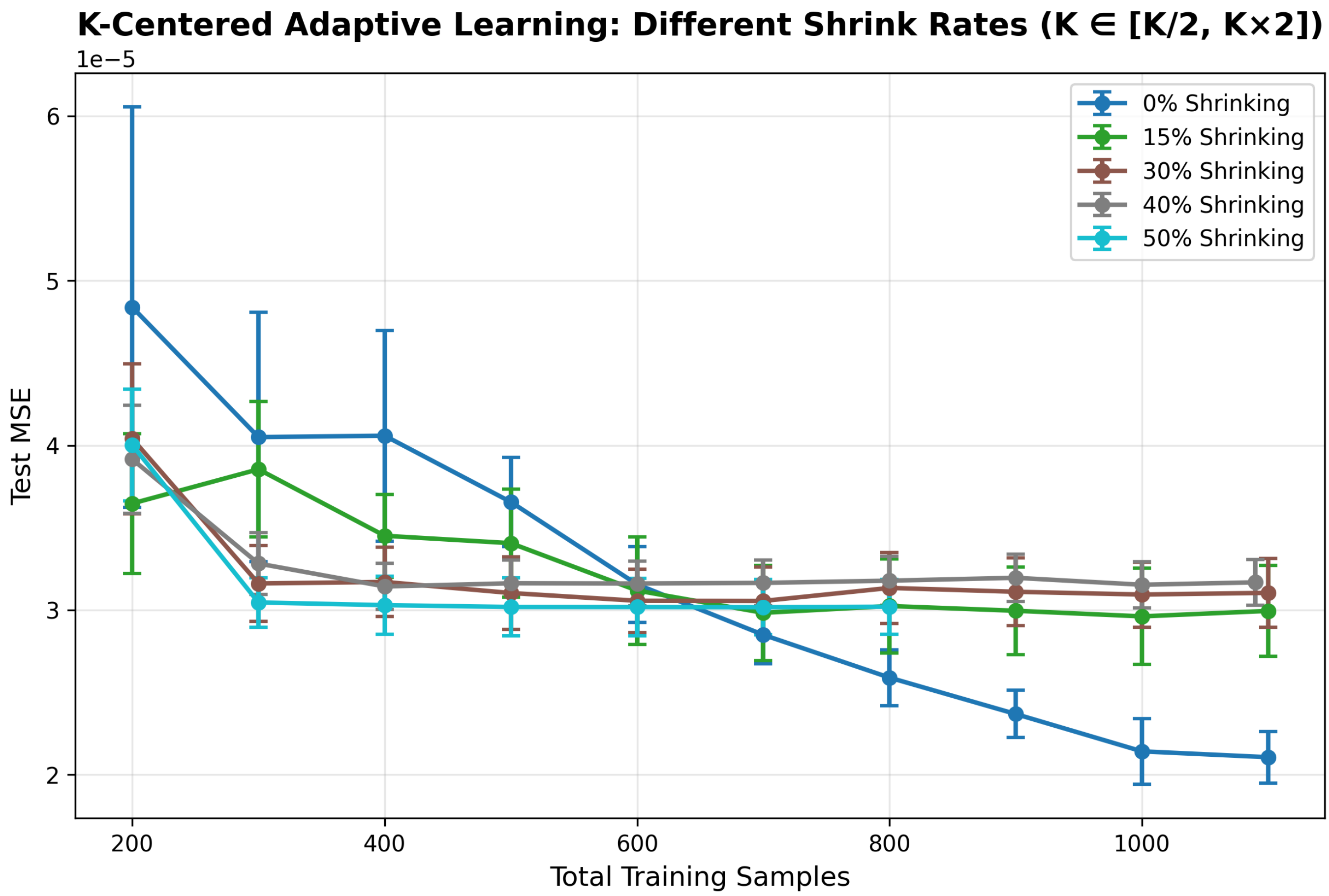
22_10_25

Previous Meeting Results: K-Centered Adaptive Learning

Box Sizes Comparison (No Shrinking)



Shrink Rates Comparison (K ∈ [K/2, K×2])



Two Topics

- Chemestry distribution
- LoKI error

1. Absolute Density and Rate Coefficient Distribution

The following graphs show the Absolute Density and Rate Coefficient distributions for the different sampling strategies.

Introduction

Setup file:

- setup_O2_simple.in

Species:

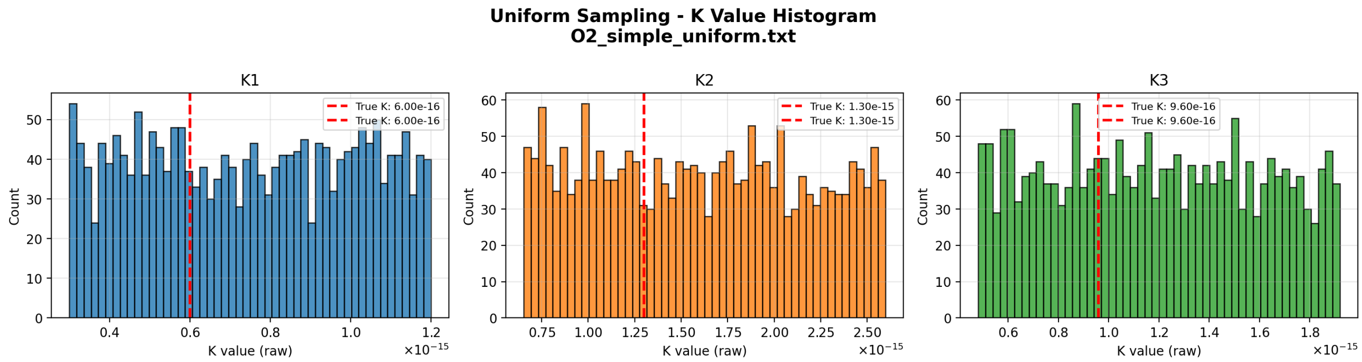
- Species 1: O2(X)
- Species 2: O2(a)
- Species 3: O(3P)

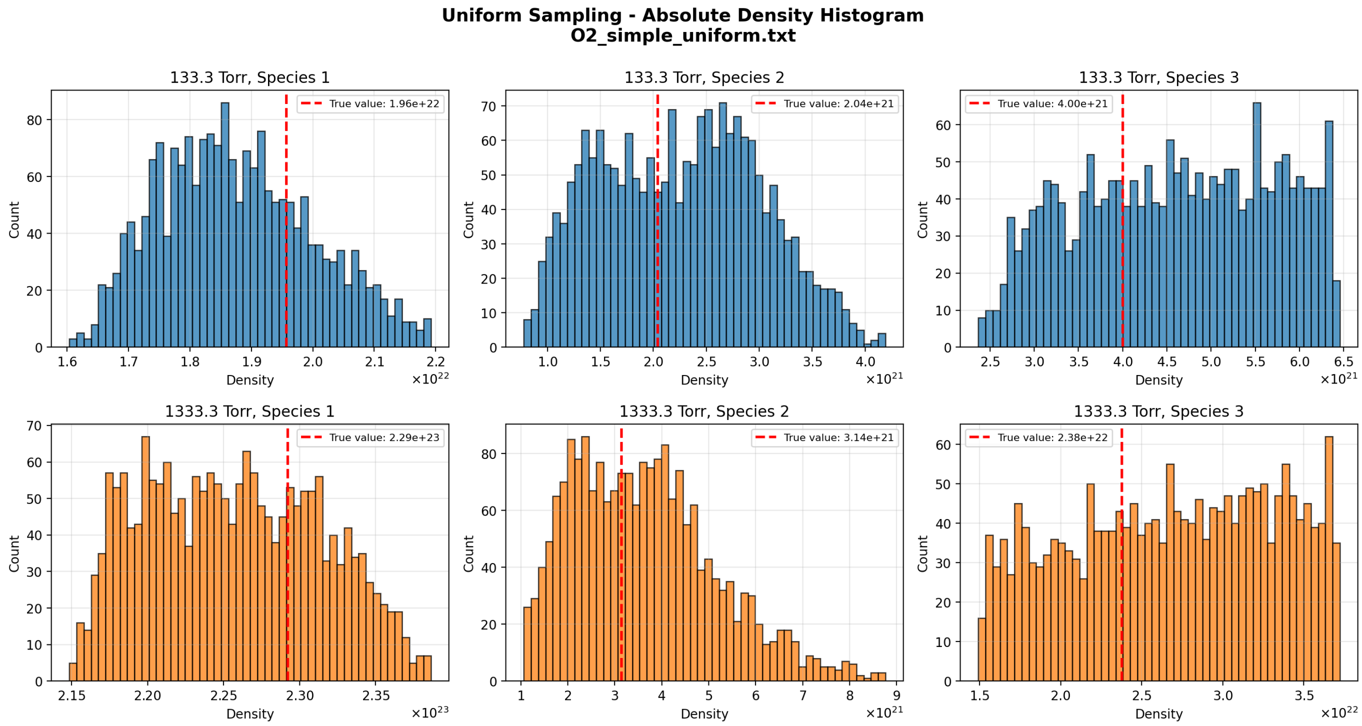
Sampling Strategies - With K between [0.5 * K_true, 2 * K_true]:

- Uniform Sampling - Training Dataset
- Uniform Sampling - Test Dataset for Marcelo's work
- Log-Uniform Sampling
- Uniform Latin Hypercube Sampling
- Log-Uniform Latin Hypercube Sampling
- Morris Method (Continuous) Sampling

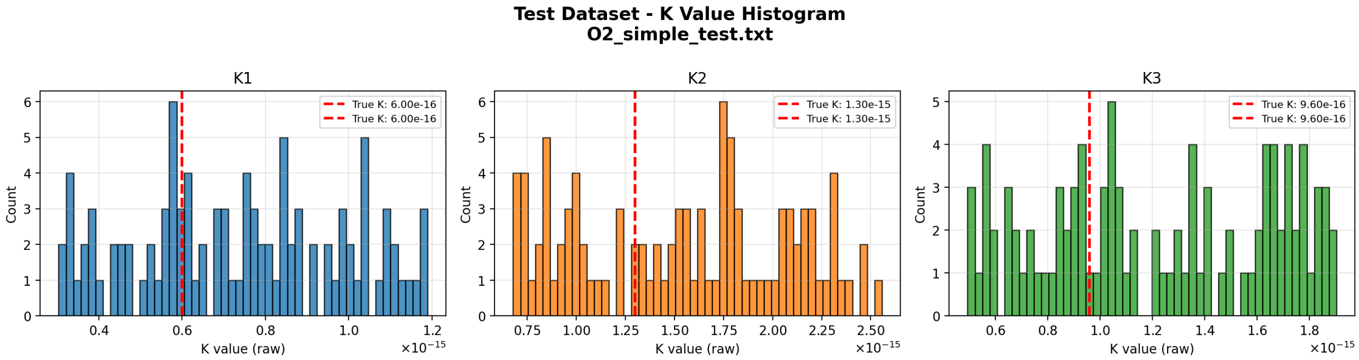
Distribution Plots by Sampling Strategy

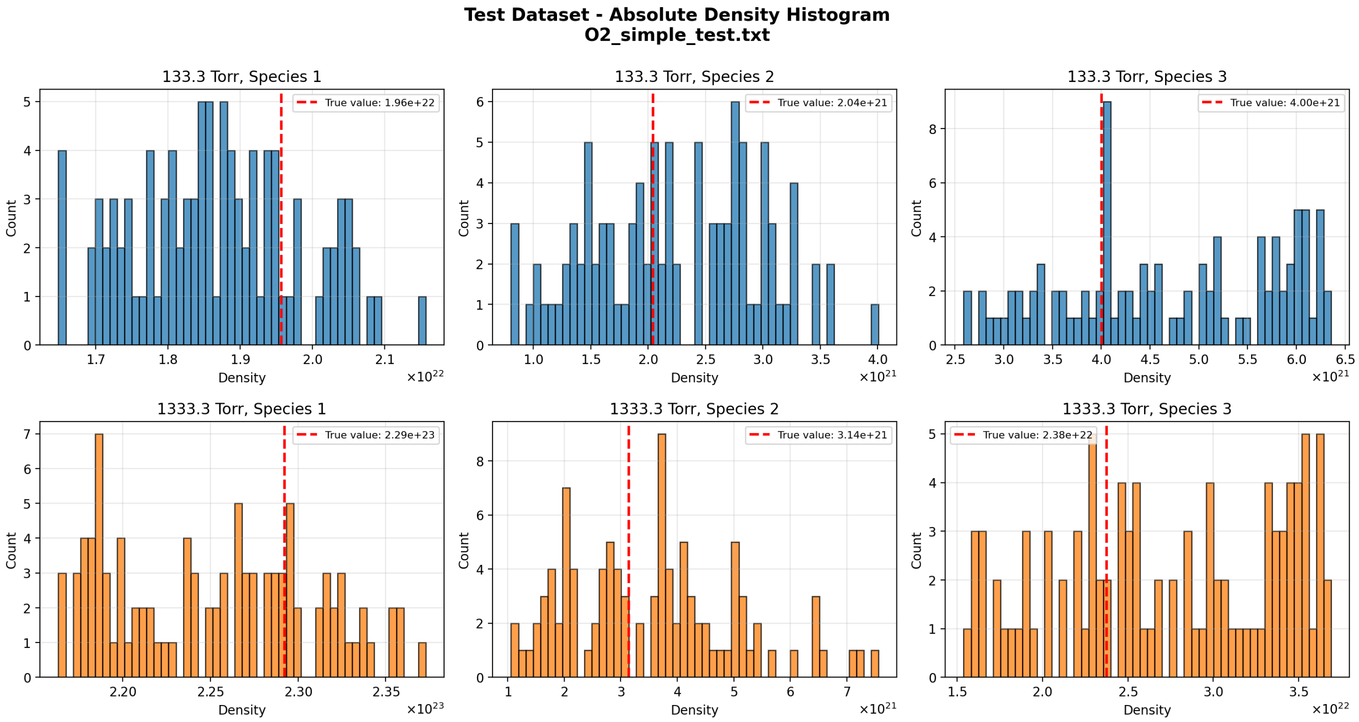
Uniform Sampling - Training Dataset



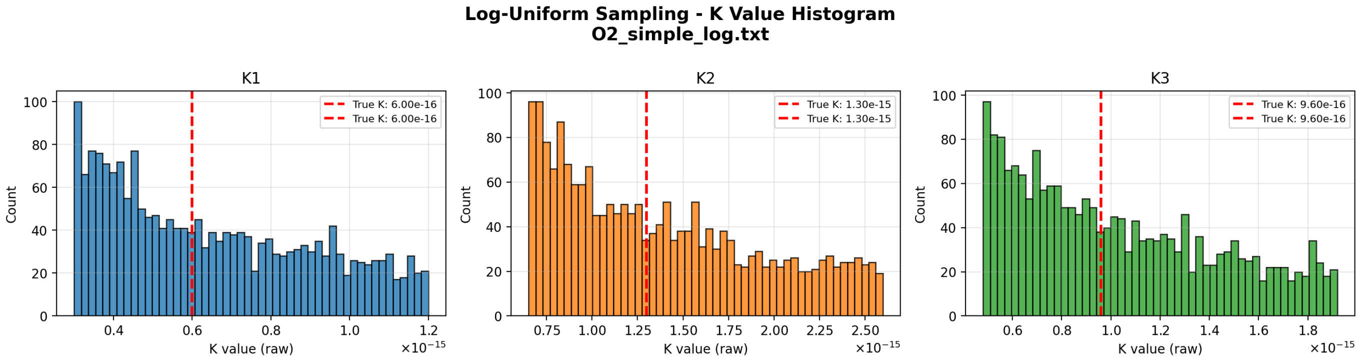


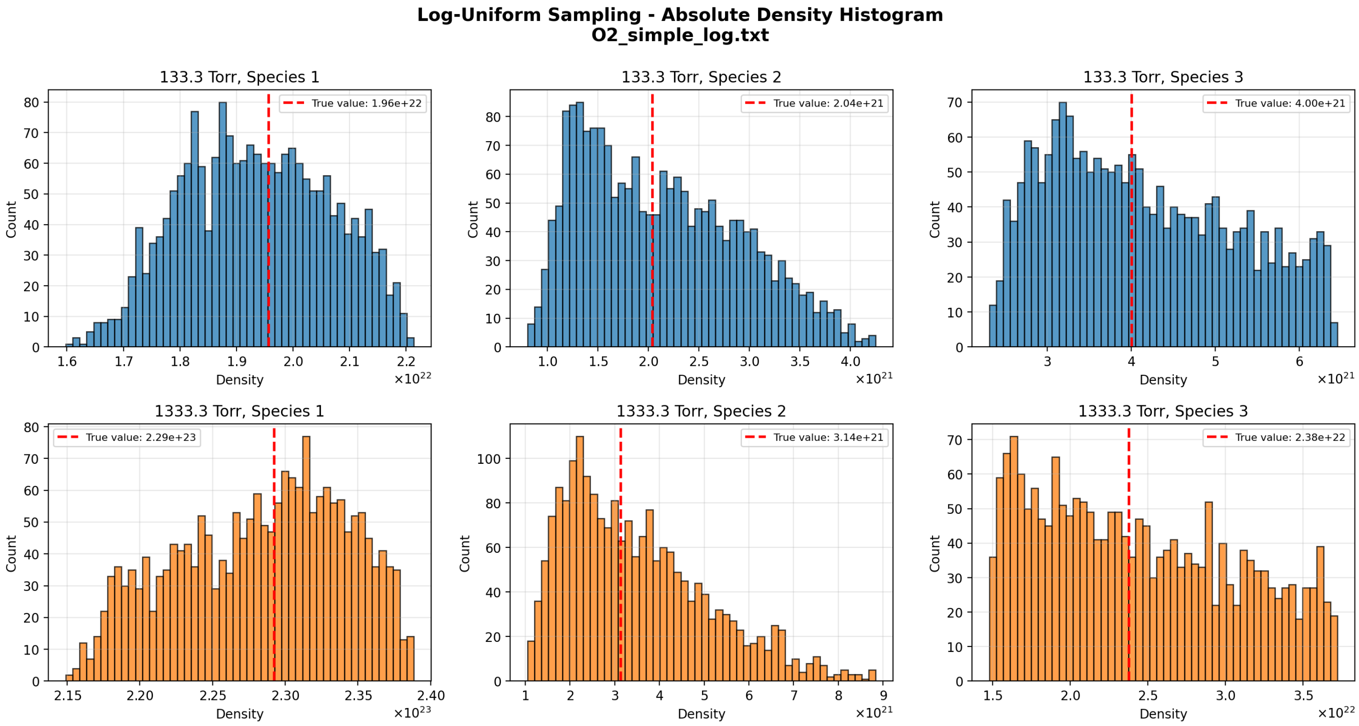
Uniform Sampling - Test Dataset



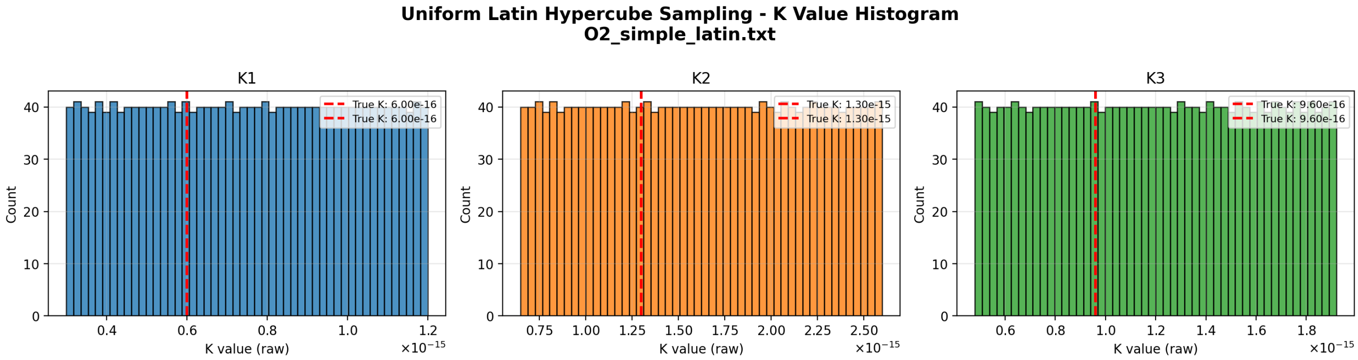


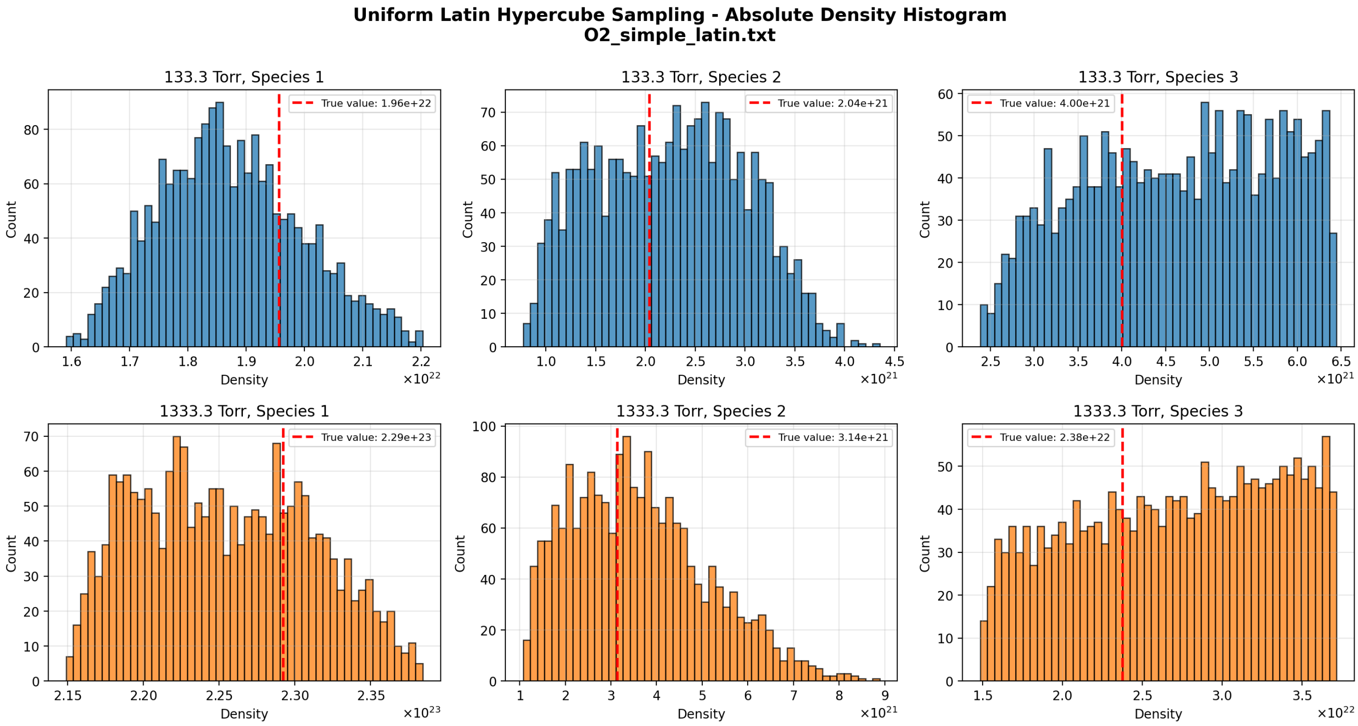
Log-Uniform Sampling



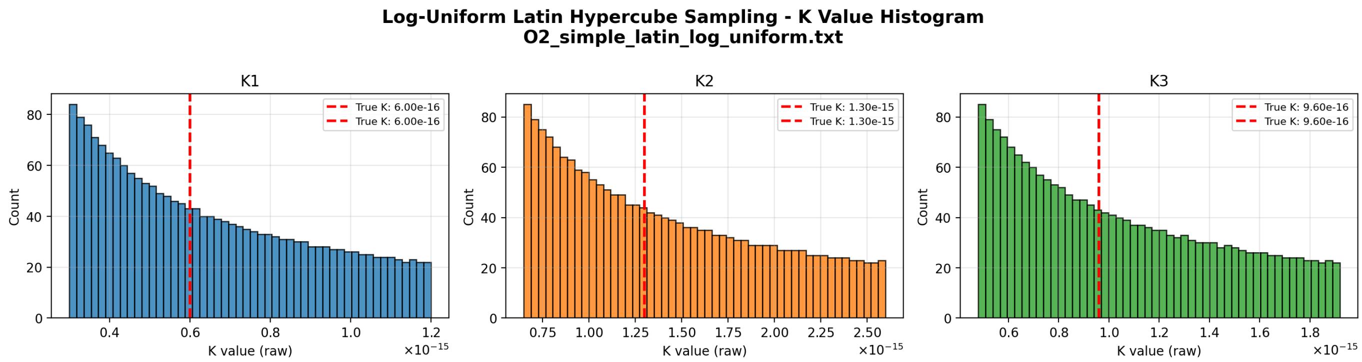


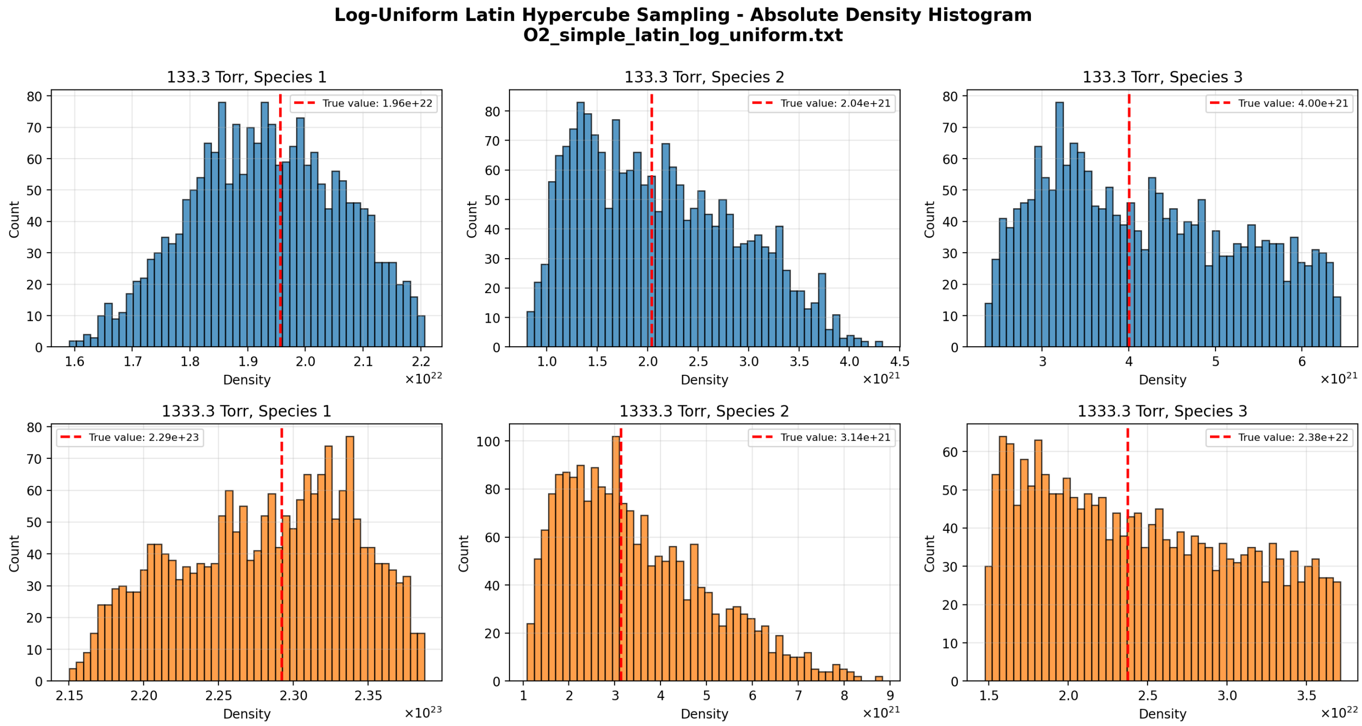
Uniform Latin Hypercube Sampling



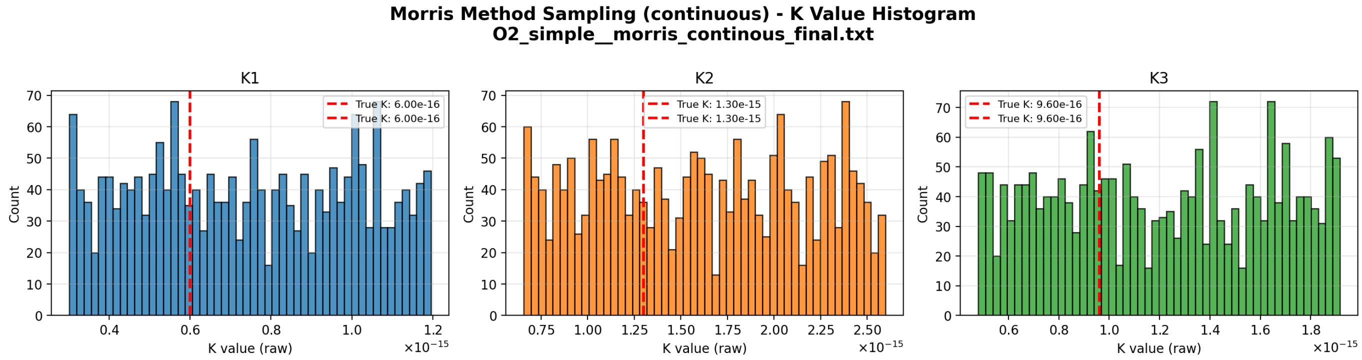


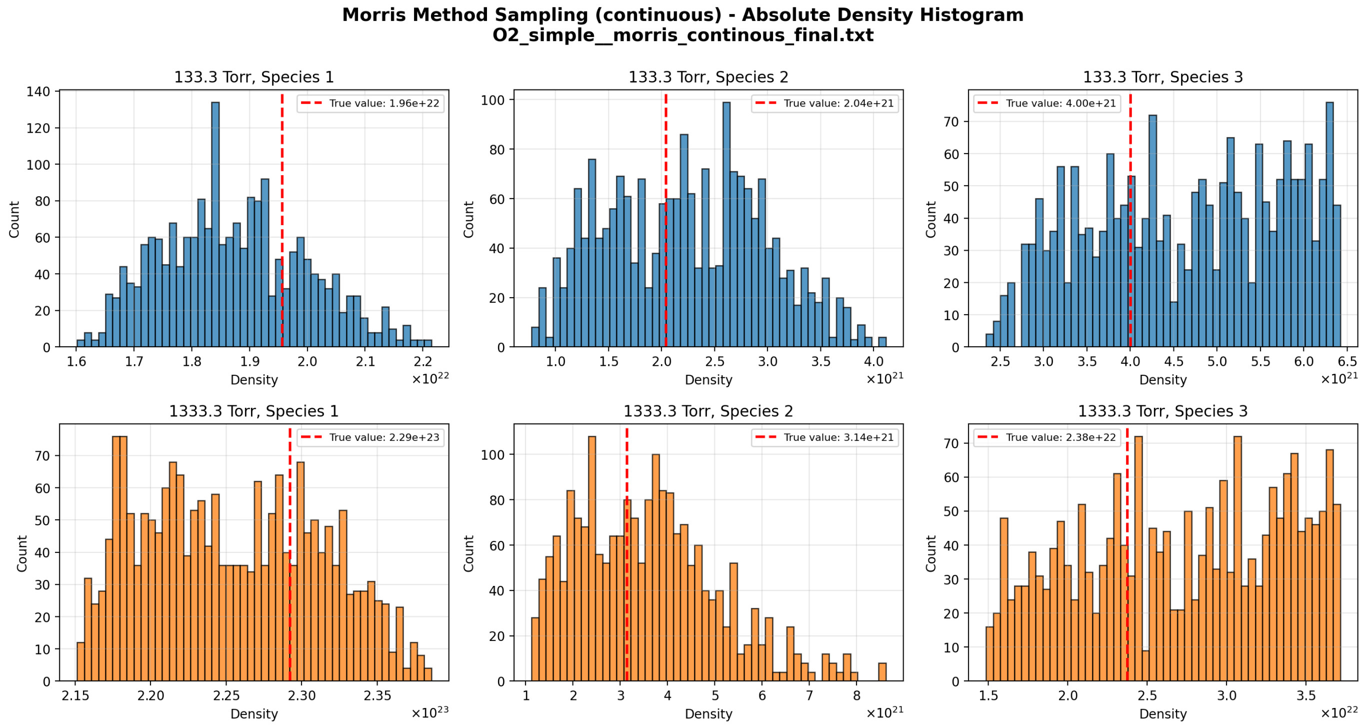
Log-Uniform Latin Hypercube Sampling





Morris Method Sampling



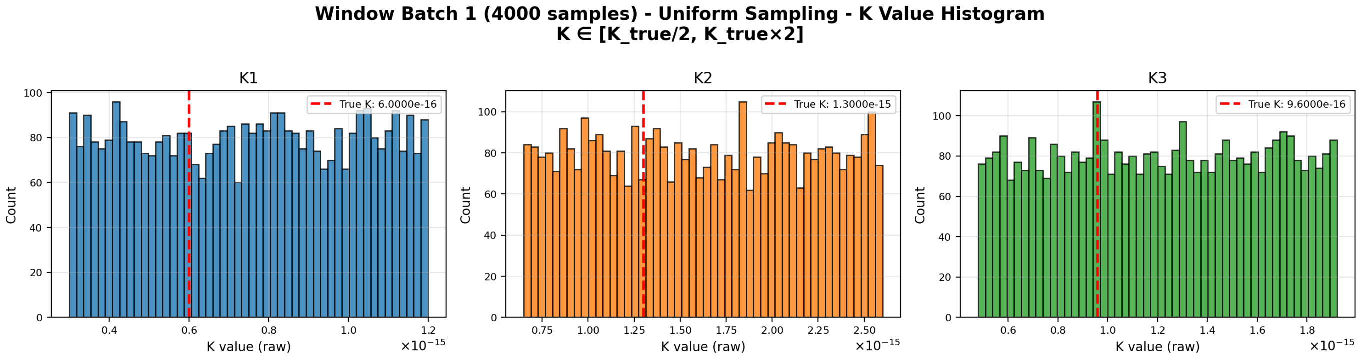


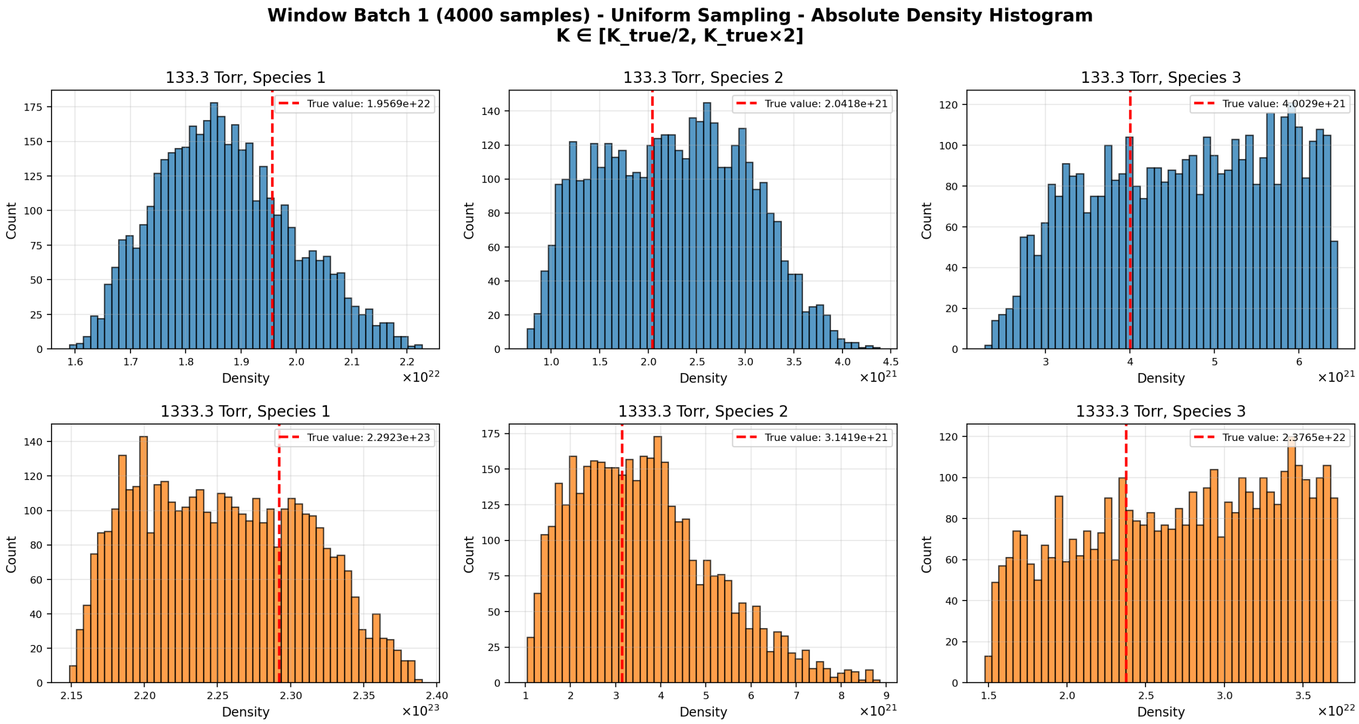
Uniform Sampling - With Varying K boundaries:

Here we show distributions coming only from Uniform Sampling. However, the allowed values for K (K_range) is now changing:

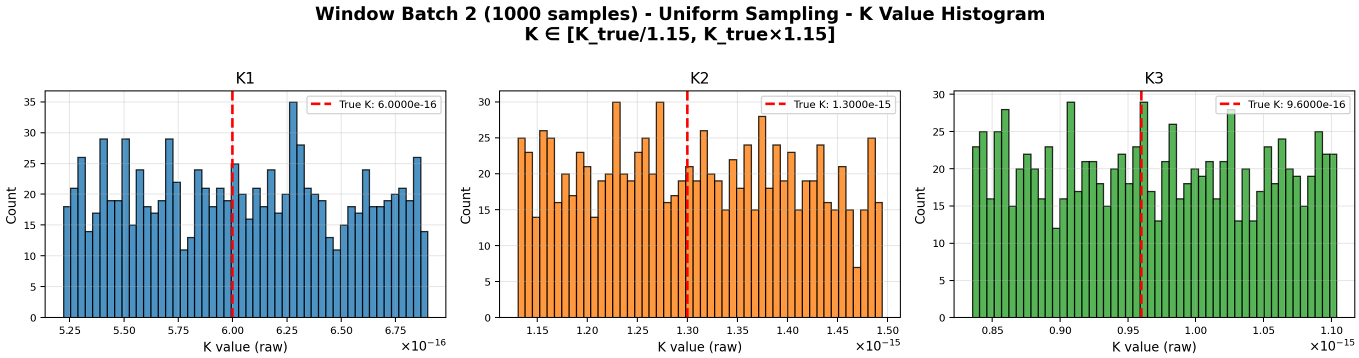
- **Batch 1** (4000 samples) - K ∈ [K_true/2, K_true×2]
- **Batch 2** (1000 samples) - K ∈ [K_true/1.15, K_true×1.15]
- **Batch 3** (2500 samples) - K ∈ [K_true/1.15, K_true×1.15]
- **Batch 4** (2000 samples) - K ∈ [K_true/1.005, K_true×1.005]
- **Batch 5** (1500 samples) - K ∈ [K_true/1.0005, K_true×1.0005]
- **Batch 6** (2000 samples) - K ∈ [K_true/1.00005, K_true×1.00005]

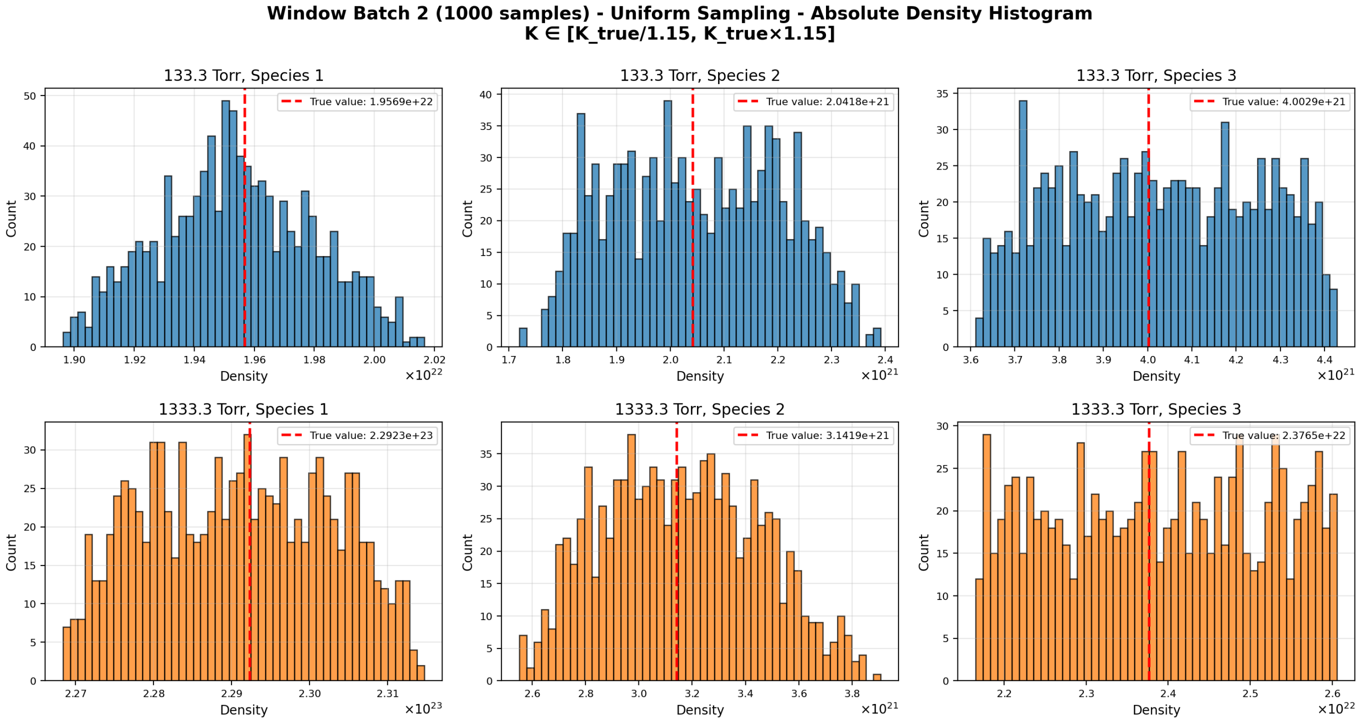
Batch 1 - K ∈ [K_true/2, K_true×2]



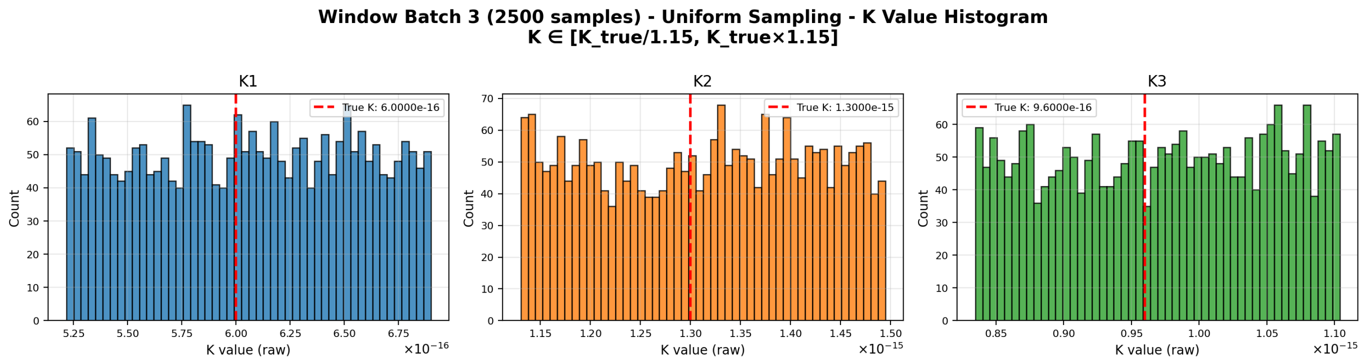


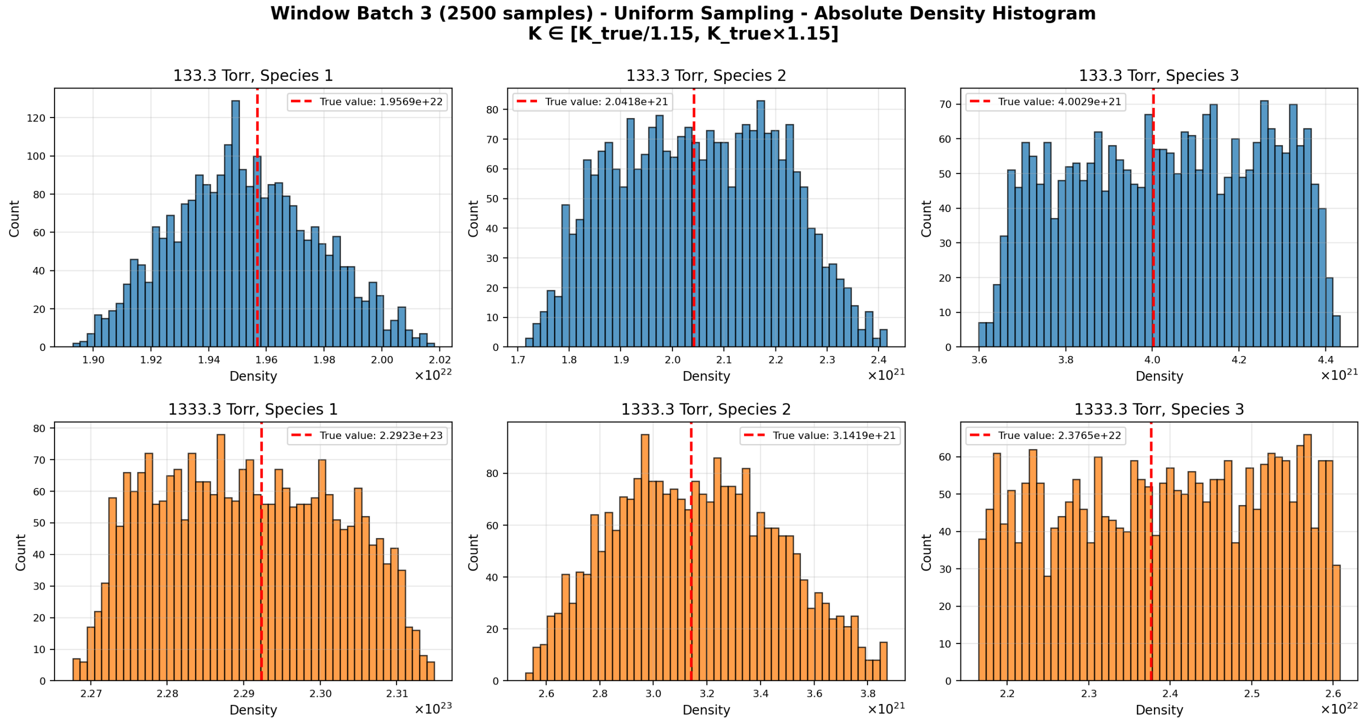
Batch 2 - K ∈ [K_true/1.15, K_true×1.15]



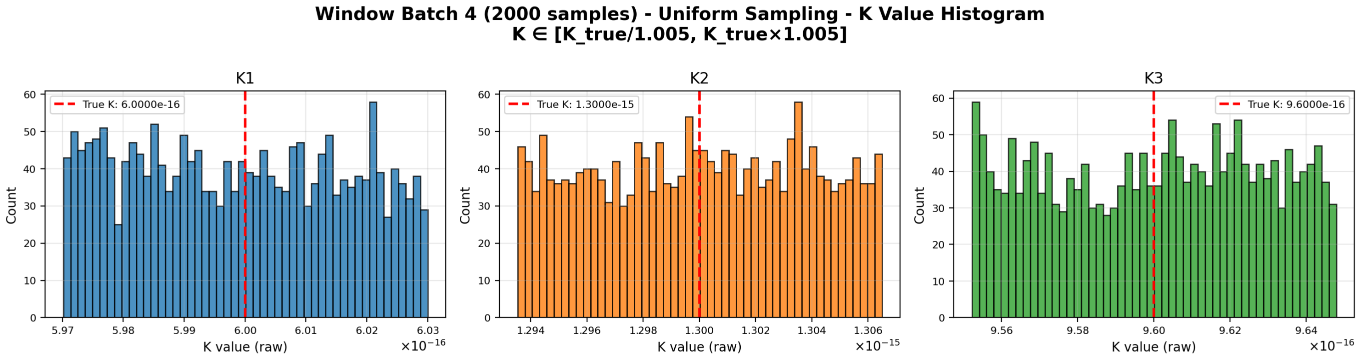


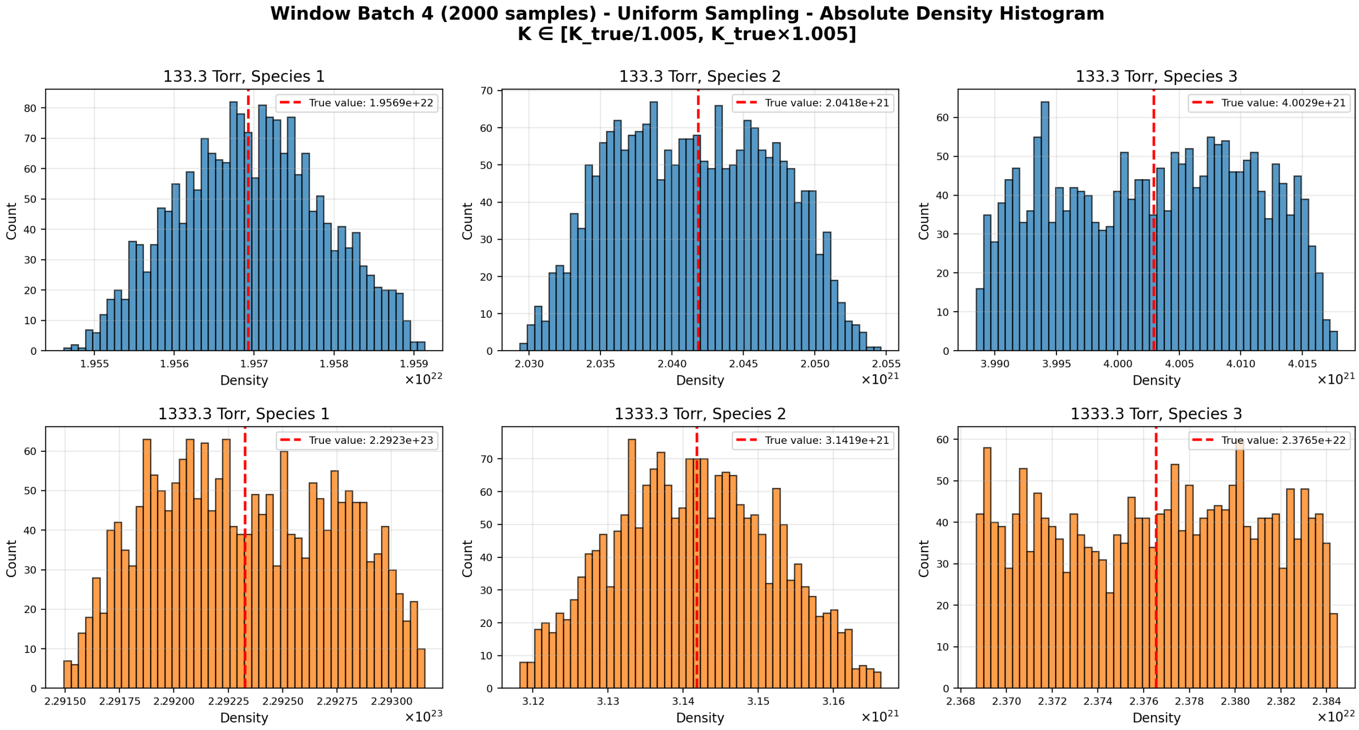
Batch 3 - K ∈ [K_true/1.15, K_true×1.15]



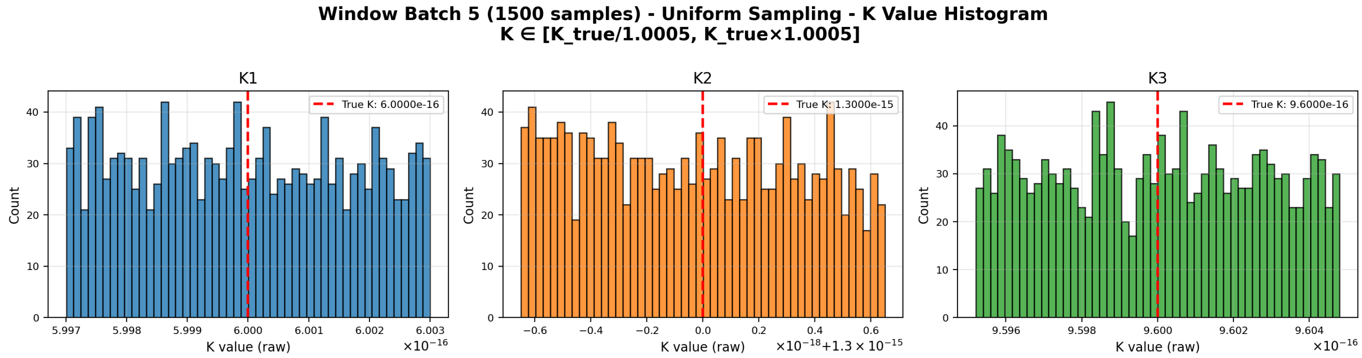


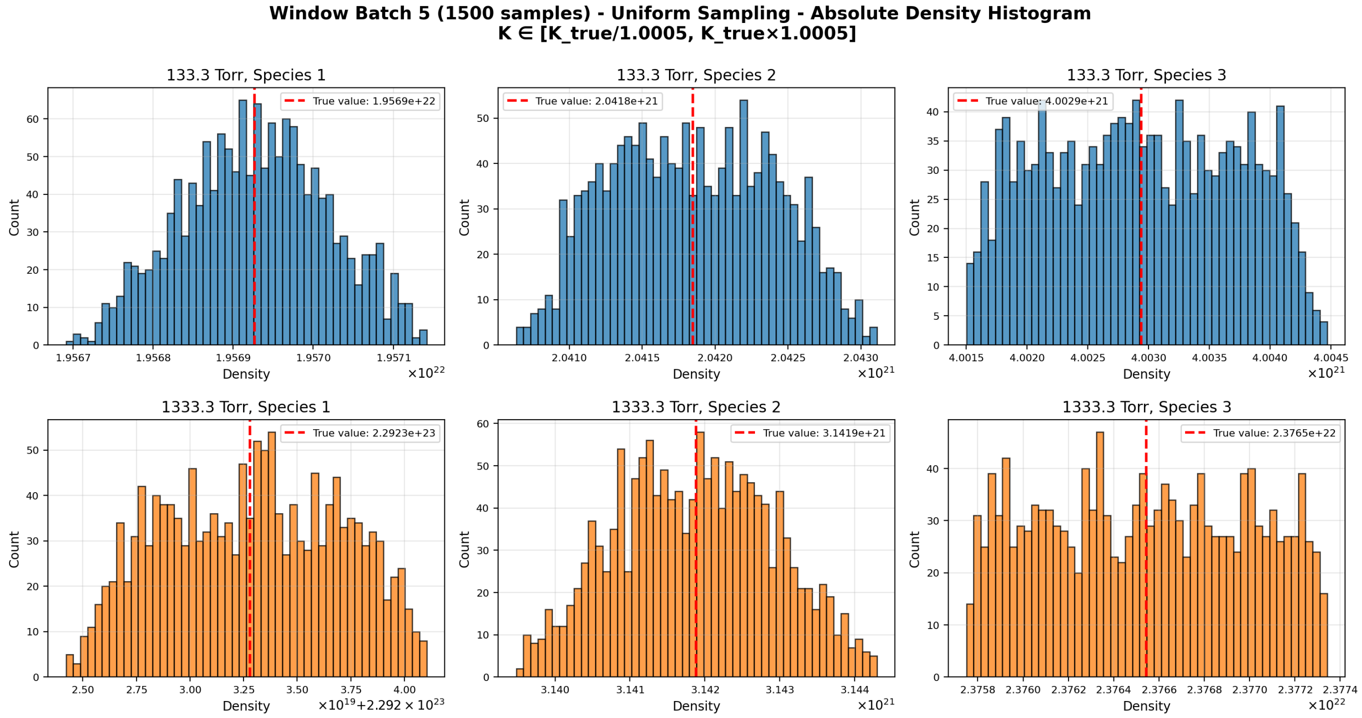
Batch 4 - K ∈ [K_true/1.005, K_true×1.005]



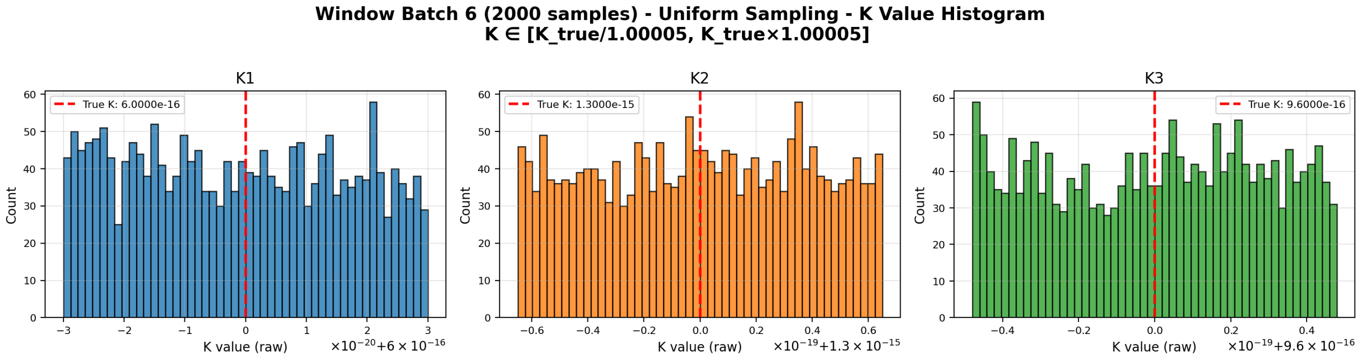


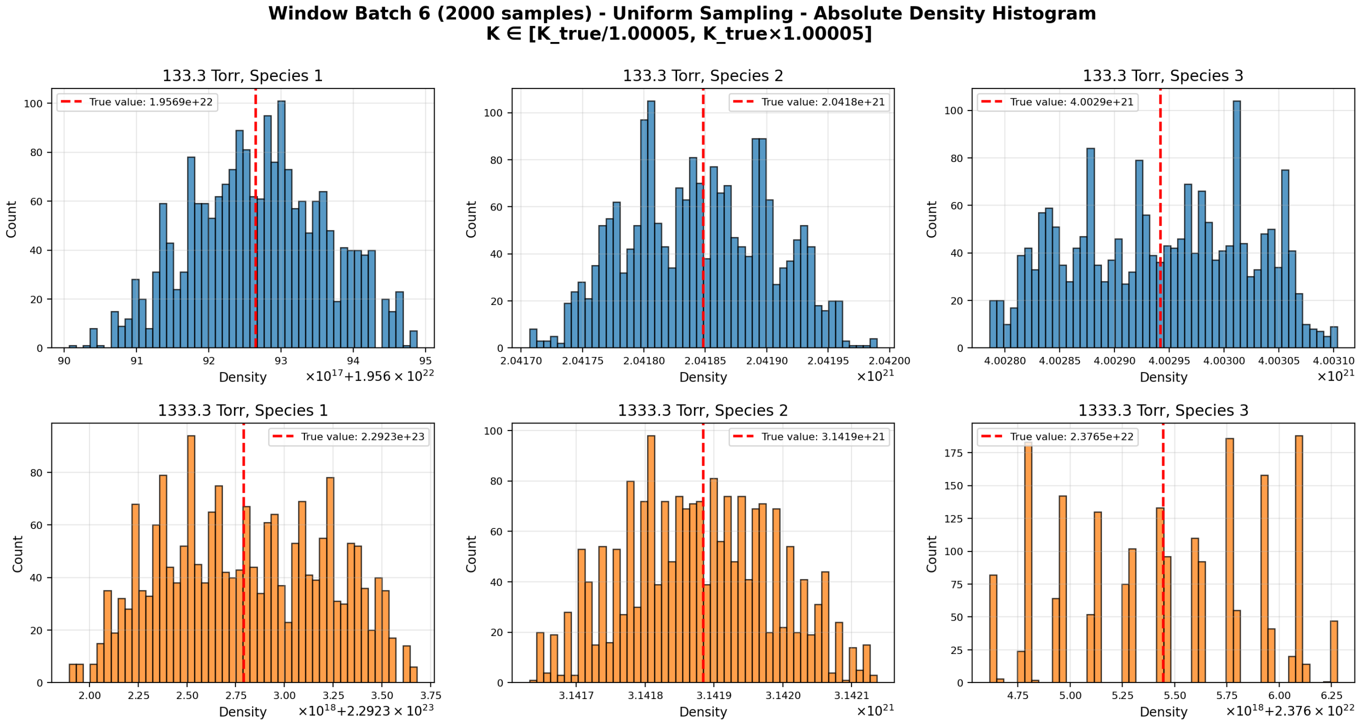
Batch 5 - K ∈ [K_true/1.0005, K_true×1.0005]





Batch 6 - K ∈ [K_true/1.00005, K_true×1.00005]





2. Loki Error Calculation

From the setup file, I find the following parameters:

iterationSchemes: pressureRelError: 1e-3 pressureMaxIterations: 800 neutralityRelError: 1e-2 neutralityMaxIterations: 100 globalRelError: 1e-3 globalMaxIterations: 200 timeIntegrationConf: odeSolver: ode15s steadyStateTime: 1e3 postDischargeTime: 0 odeSetParameters: % optional parameters that can be sent to the odeSolver % RelTol: 1e-7 % AbsTol: 1e-10 % MaxStep: 0.1

Errors:

- pressureRelError: 1e-3
- neutralityRelError: 1e-2
- globalRelError: 1e-3
- RelTol: 1e-3
- AbsTol: 1e-6

For solving the ODE:

- odeSolver: ode15s
- Default Values: -RelTol: 1e-3 -AbsTol: 1e-6
- Source: https://www.mathworks.com/matlabcentral/answers/1819175-stiff-differential-equation-solver-euler

Issue: When i try to decrease the Relative Error i have two issues:

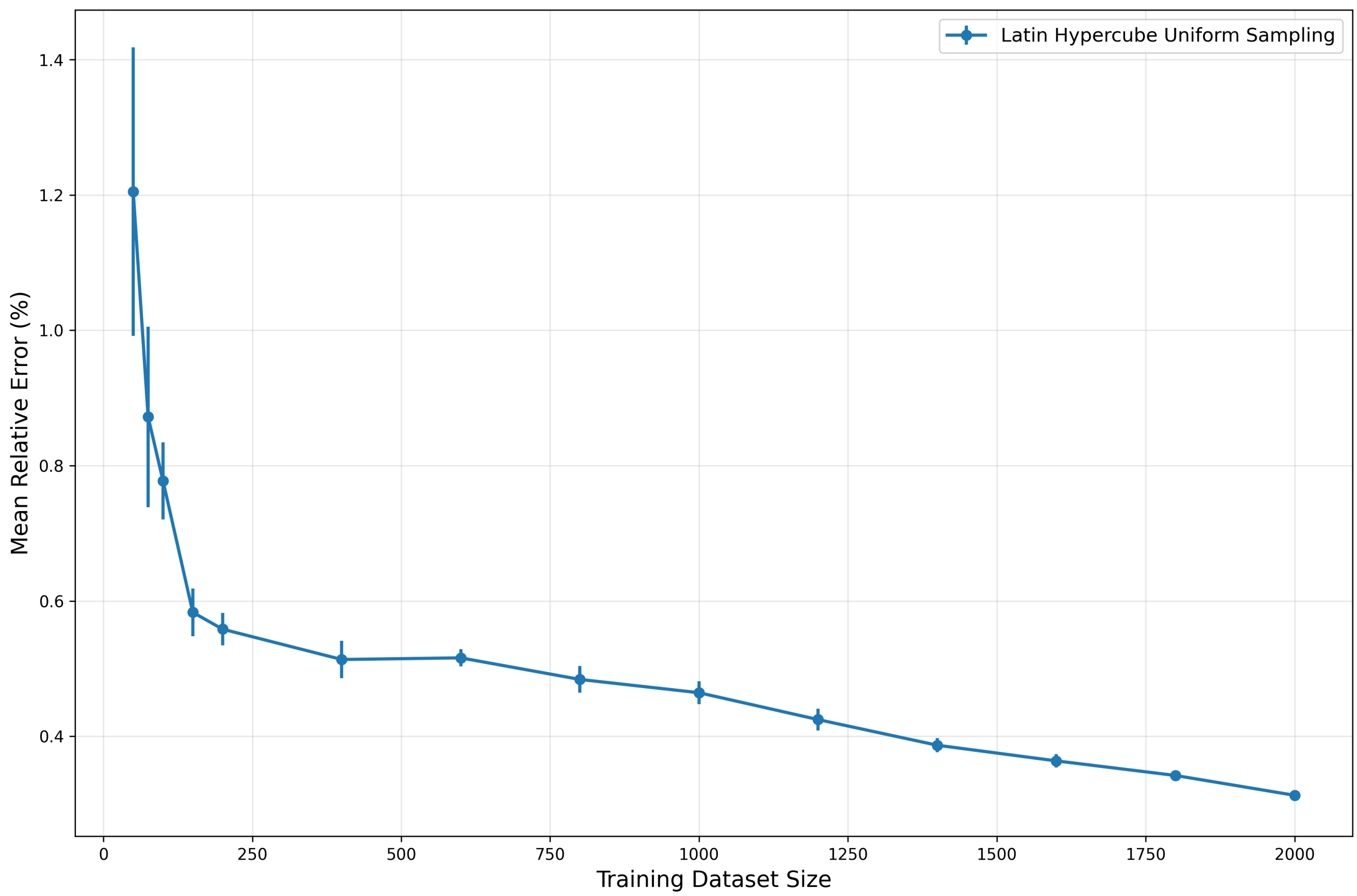
- "Error using odearguments (line 126) RelTol must be a positive scalar." - This one i could fix by changing the LoKI matlab code to cast the parameter correctly.
- I decreased the RelTol to a very small value (1e-12) but I get the same output.
    - Only changing - pressureRelError: 1e-3 changes the output.

Should we address the inverse problem as a regression?

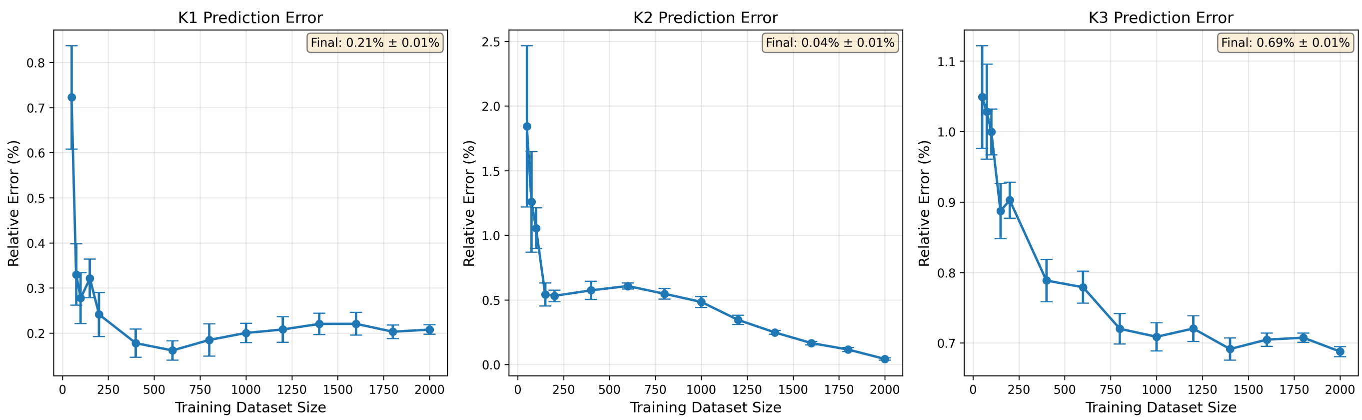
4. Sample Efficiency Analysis

Error Bars - 10 Seeds

Full MSE Comparison

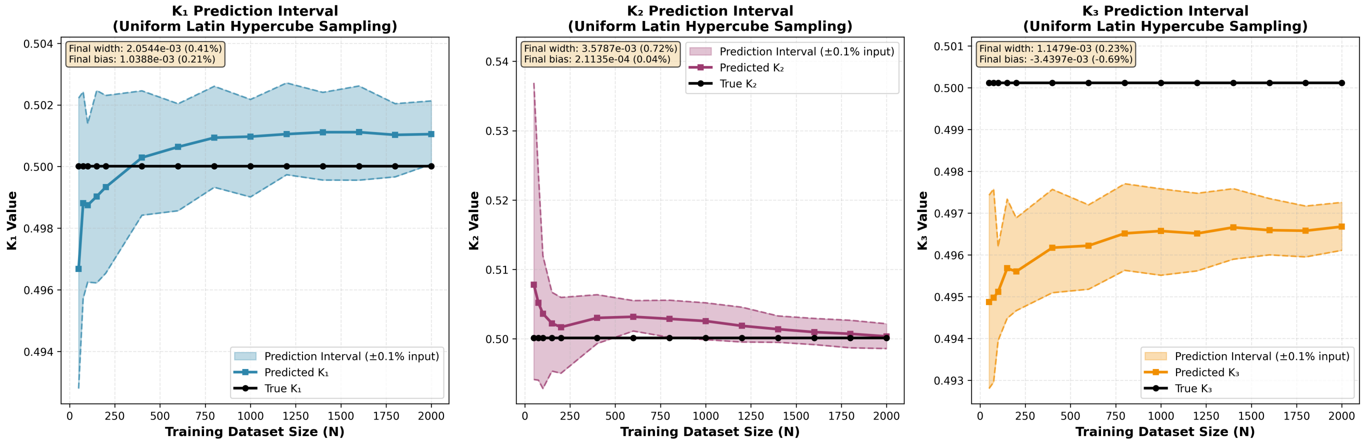


Error Per Output - Uniform Sampling



4. Perturbation Analysis

Testing model robustness with ±0.1% input perturbations across different training dataset sizes (N).



Perturbation Results Table

N	K_1 Interval	True K_1	Contained?	K_2 Interval	True K_2	Contained?	K_3 Interval	True K_3	Contained?
50	[0.49280, 0.50222]	0.50001	✓	[0.49413, 0.53688]	0.50014	✓	[0.51053, 0.53639]	0.50012	✗
75	[0.49572, 0.50243]	0.50001	✓	[0.49404, 0.52384]	0.50014	✓	[0.50795, 0.52584]	0.50012	✗
100	[0.49625, 0.50140]	0.50001	✓	[0.49284, 0.51197]	0.50014	✓	[0.50729, 0.51777]	0.50012	✗
150	[0.49622, 0.50247]	0.50001	✓	[0.49534, 0.50670]	0.50014	✓	[0.50645, 0.51546]	0.50012	✗
200	[0.49654, 0.50231]	0.50001	✓	[0.49504, 0.50596]	0.50014	✓	[0.49948, 0.51482]	0.50012	✓
400	[0.49842, 0.50246]	0.50001	✓	[0.49931, 0.50635]	0.50014	✓	[0.50111, 0.50947]	0.50012	✗
600	[0.49856, 0.50204]	0.50001	✓	[0.50112, 0.50551]	0.50014	✗	[0.50129, 0.50848]	0.50012	✗
800	[0.49933, 0.50261]	0.50001	✓	[0.50021, 0.50556]	0.50014	✗	[0.50176, 0.50903]	0.50012	✗
1000	[0.49901, 0.50218]	0.50001	✓	[0.49988, 0.50518]	0.50014	✓	[0.50070, 0.50756]	0.50012	✗
1200	[0.49973, 0.50271]	0.50001	✓	[0.49953, 0.50456]	0.50014	✓	[0.50118, 0.50692]	0.50012	✗
1400	[0.49956, 0.50241]	0.50001	✓	[0.49949, 0.50329]	0.50014	✓	[0.50152, 0.50701]	0.50012	✗
1600	[0.49956, 0.50261]	0.50001	✓	[0.49916, 0.50293]	0.50014	✓	[0.50257, 0.50759]	0.50012	✗
1800	[0.49966, 0.50204]	0.50001	✓	[0.49870, 0.50267]	0.50014	✓	[0.50407, 0.50724]	0.50012	✗

N	K_1 Interval	True K_1	Contained?	K_2 Interval	True K_2	Contained?	K_3 Interval	True K_3	Contained?
2000	[0.50008, 0.50213]	0.50001	✗	[0.49859, 0.50217]	0.50014	✓	[0.50503, 0.50603]	0.50012	✗

Key Observations:

- ✓ indicates true value is contained within prediction interval
- ✗ indicates true value falls outside prediction interval
- K_1 shows best containment (13/14 cases)
- K_2 shows good containment (11/14 cases)
- K_3 shows poor containment (2/14 cases) - systematic bias toward overestimation
- Prediction intervals narrow as N increases, showing improved model confidence