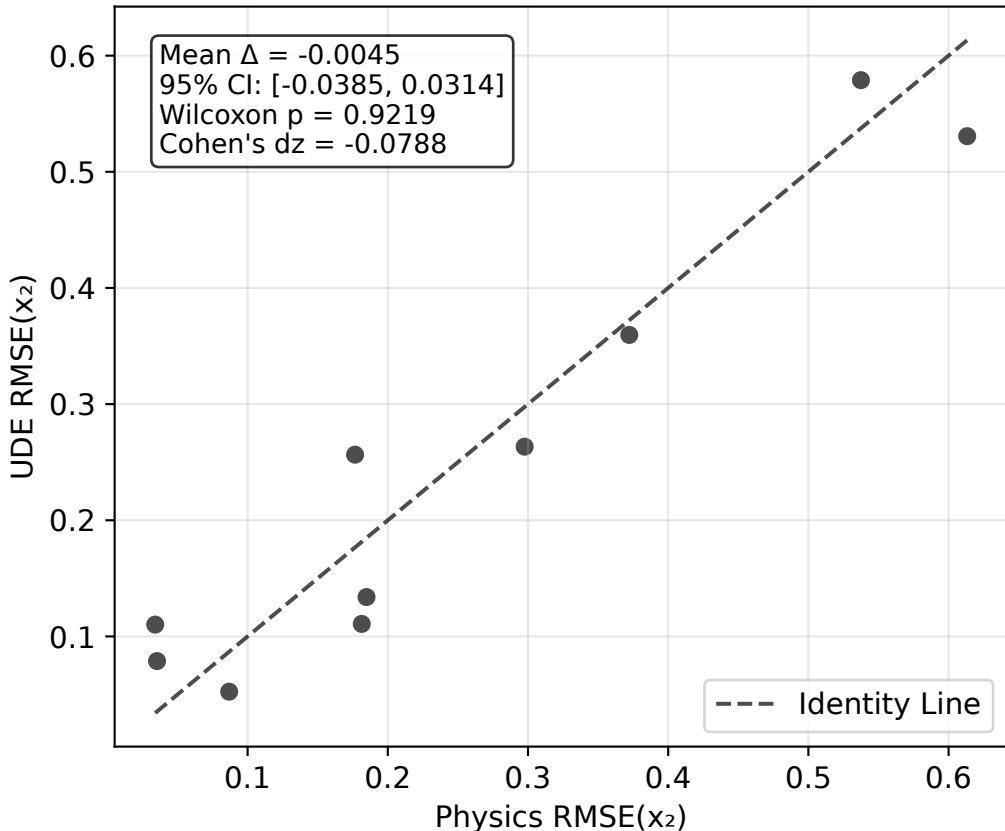
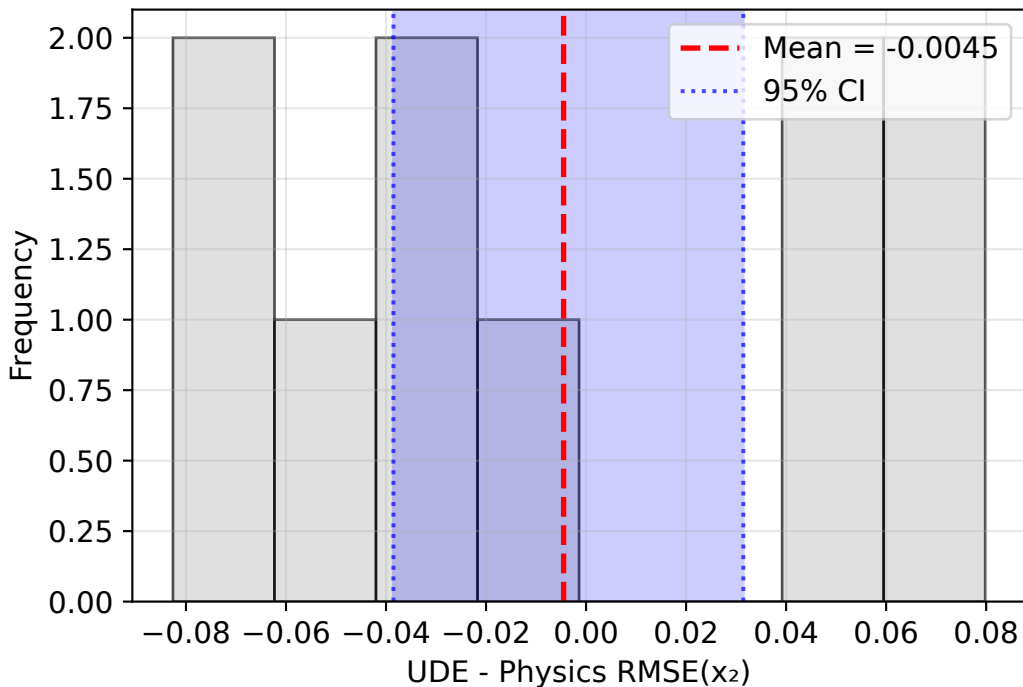


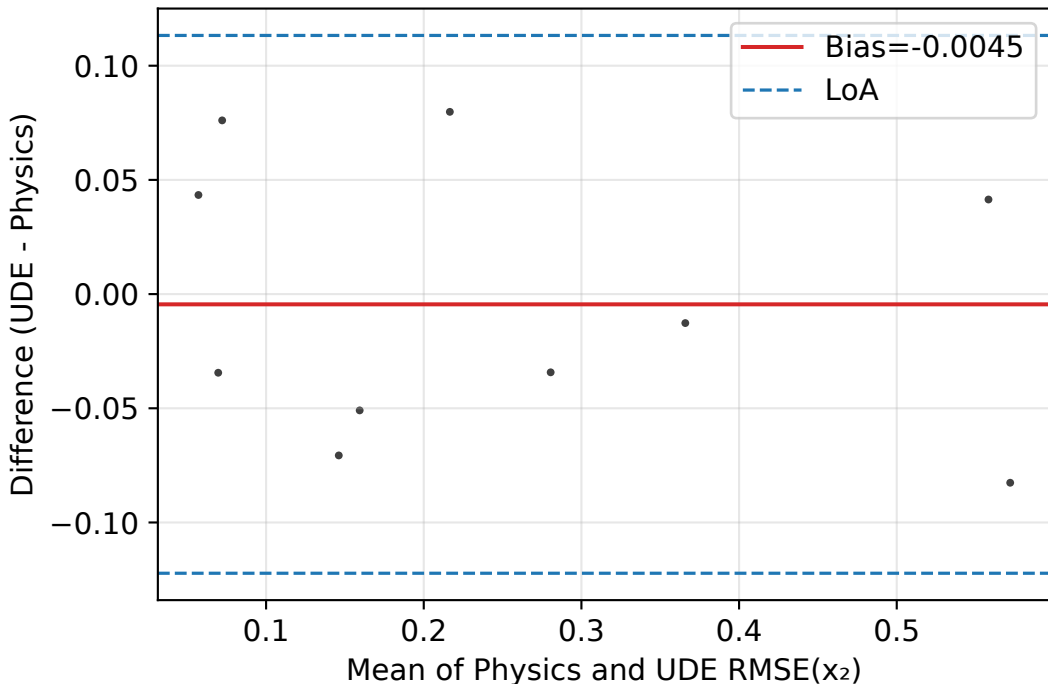
# UDE vs Physics Performance Comparison



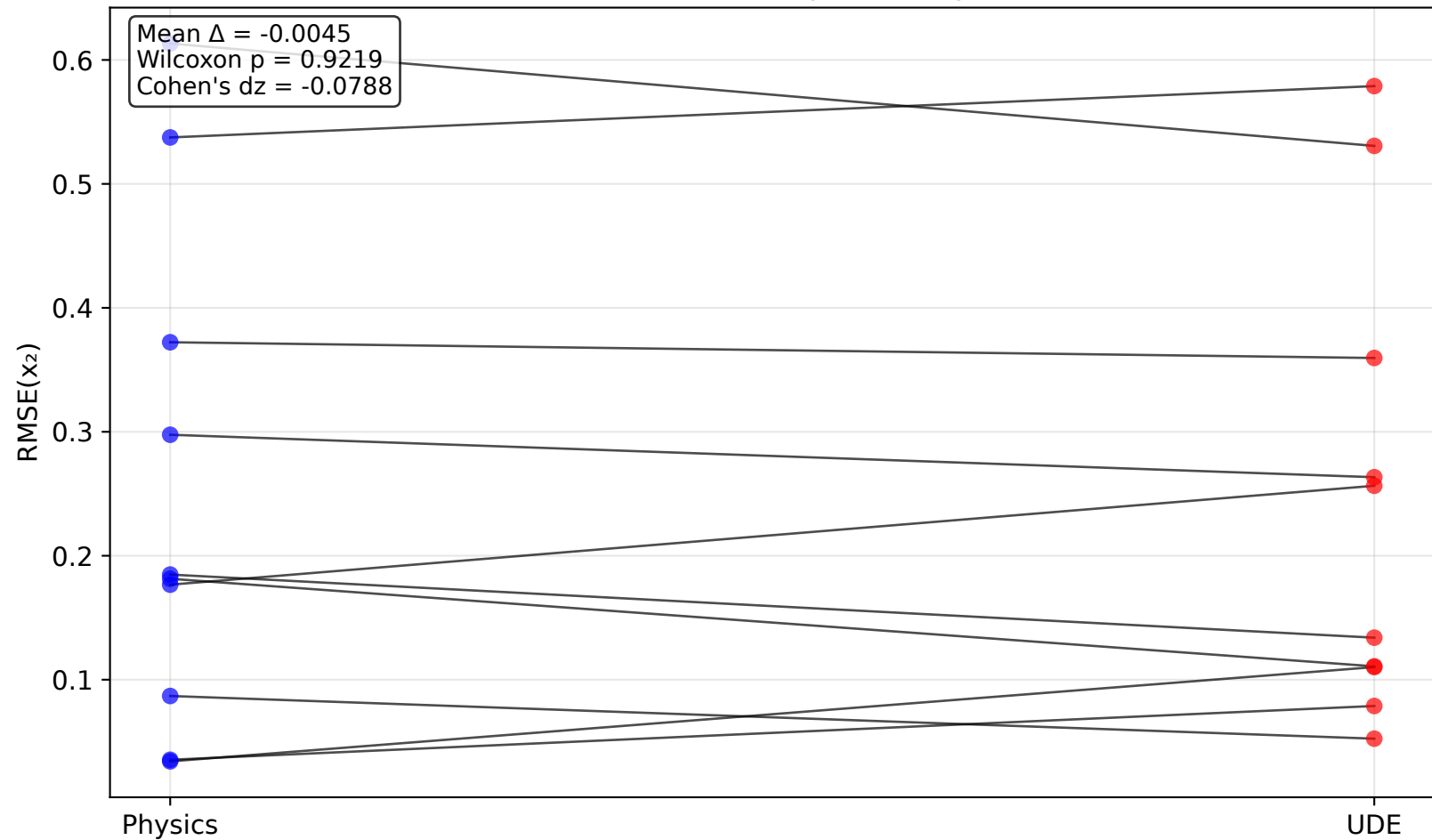
# Distribution of Performance Differences



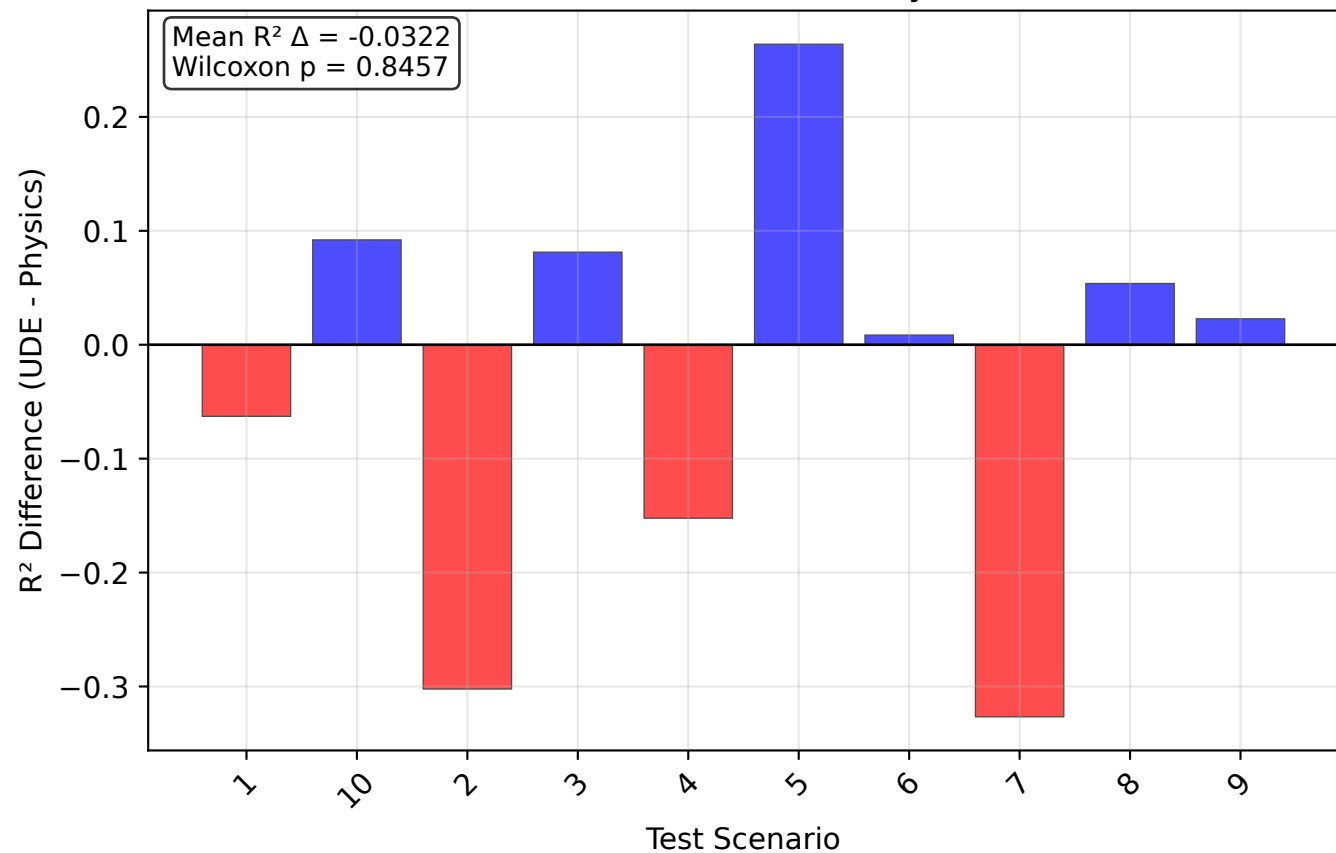
## Bland-Altman (RMSE $x_2$ )



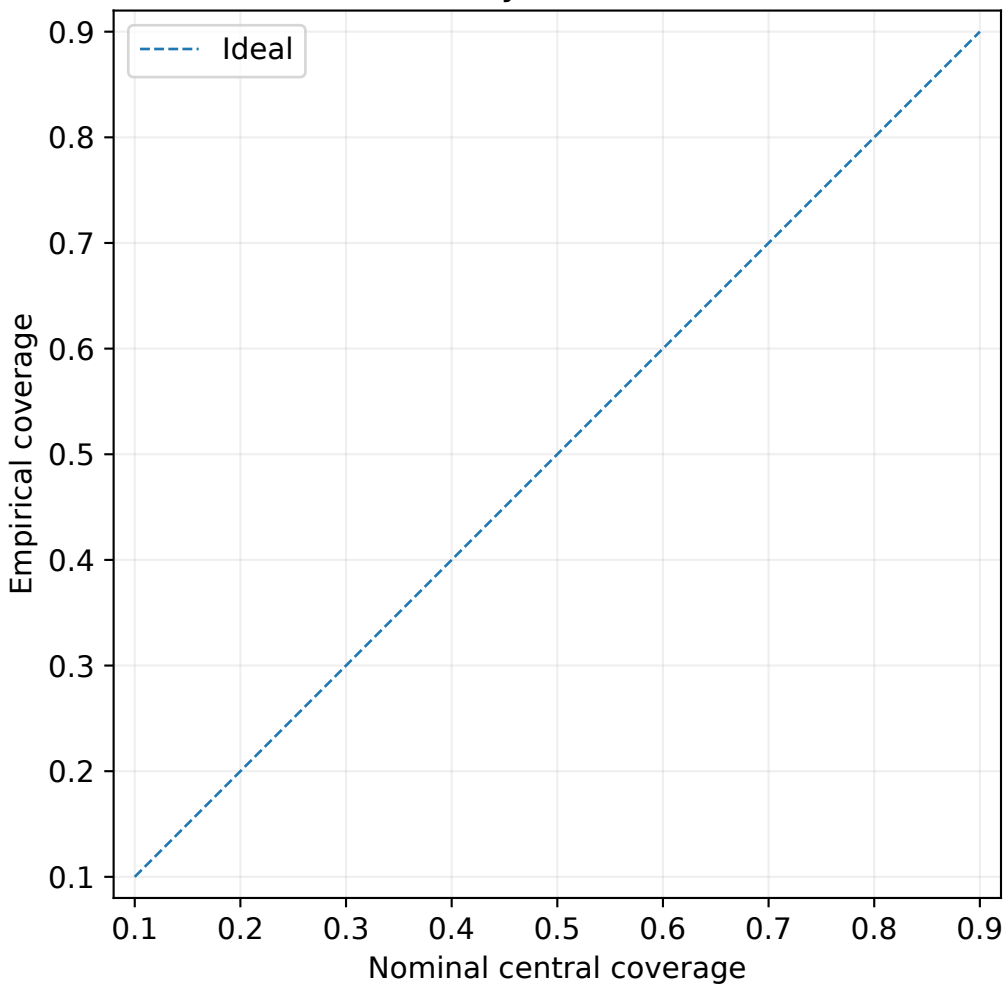
# Paired Performance Comparison by Scenario



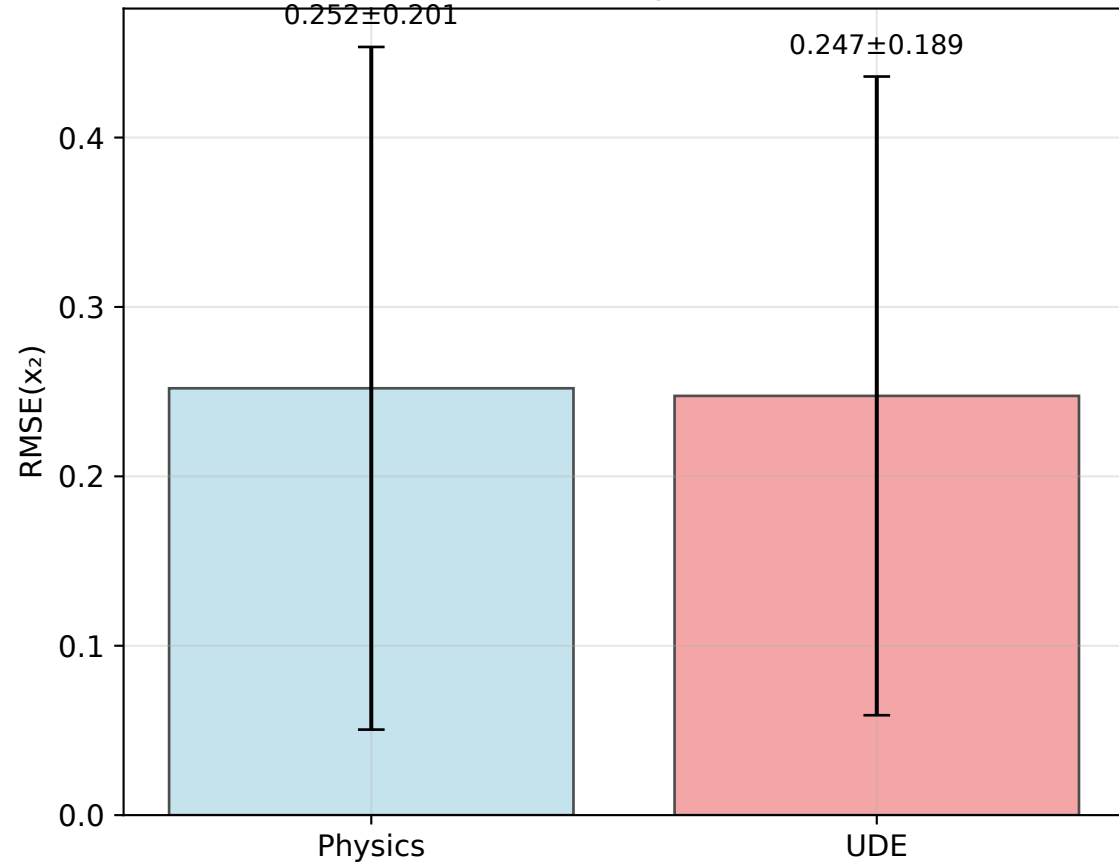
## R<sup>2</sup> Performance Differences by Scenario



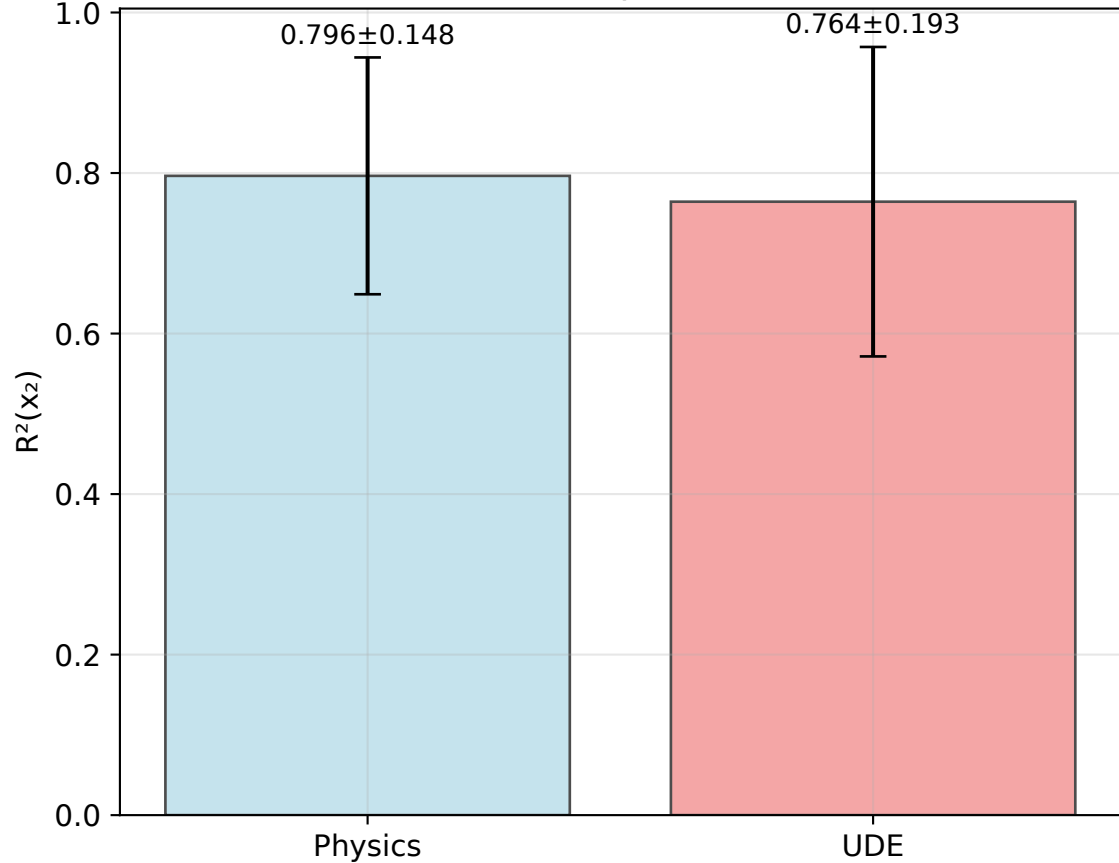
# Schematic — Reliability (could not decode artifact)



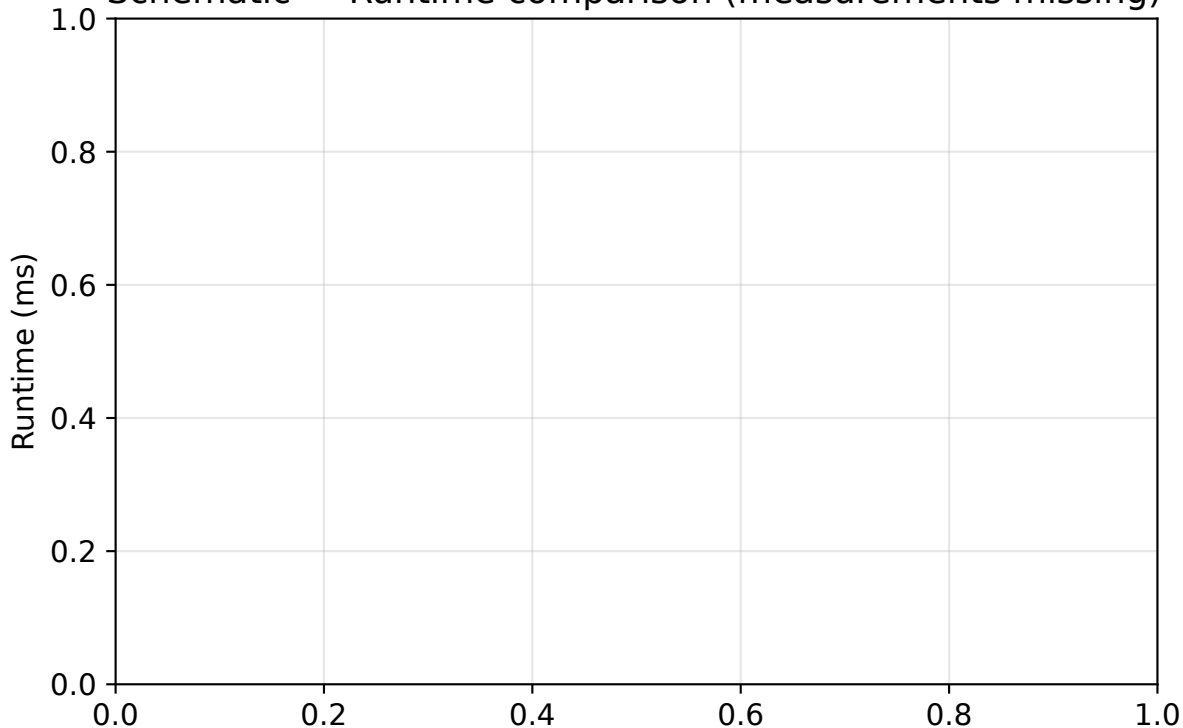
### RMSE Comparison



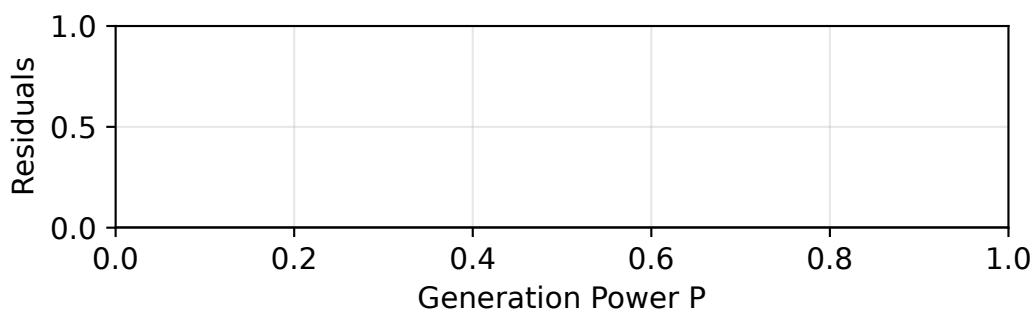
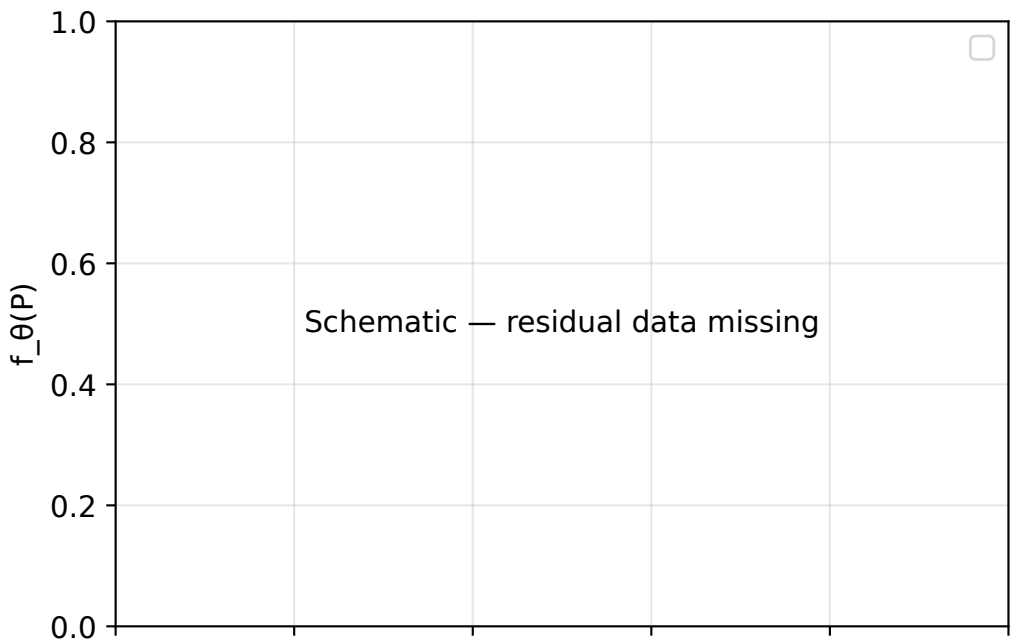
### R<sup>2</sup> Comparison

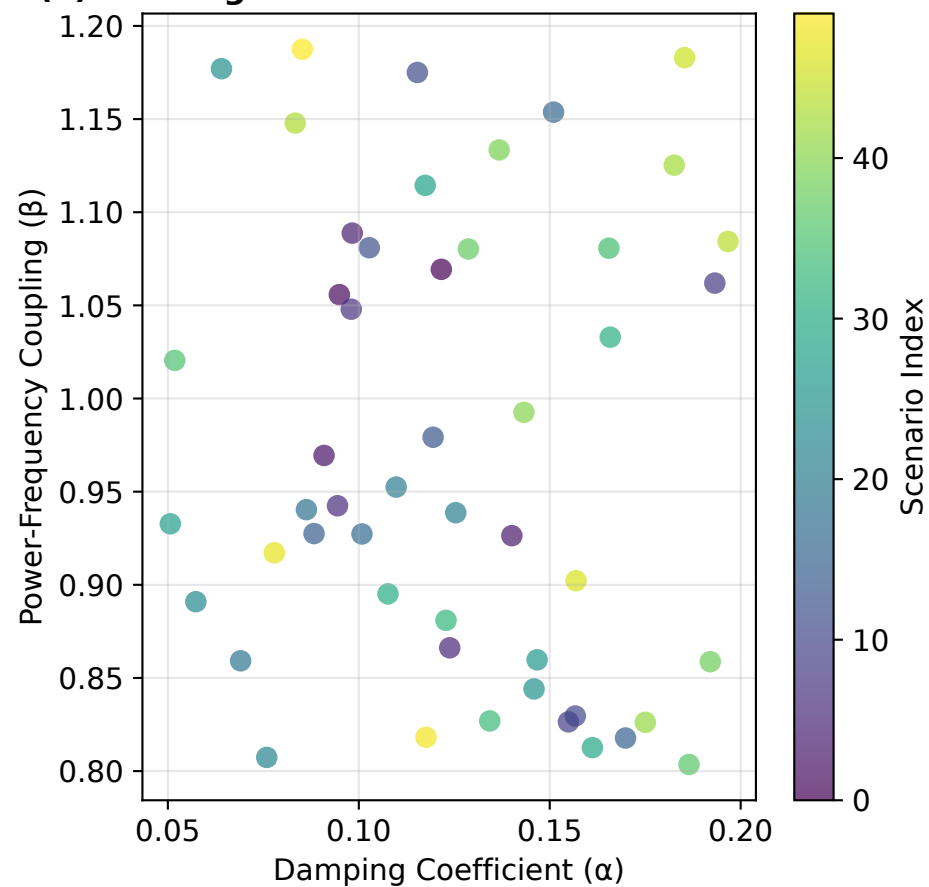
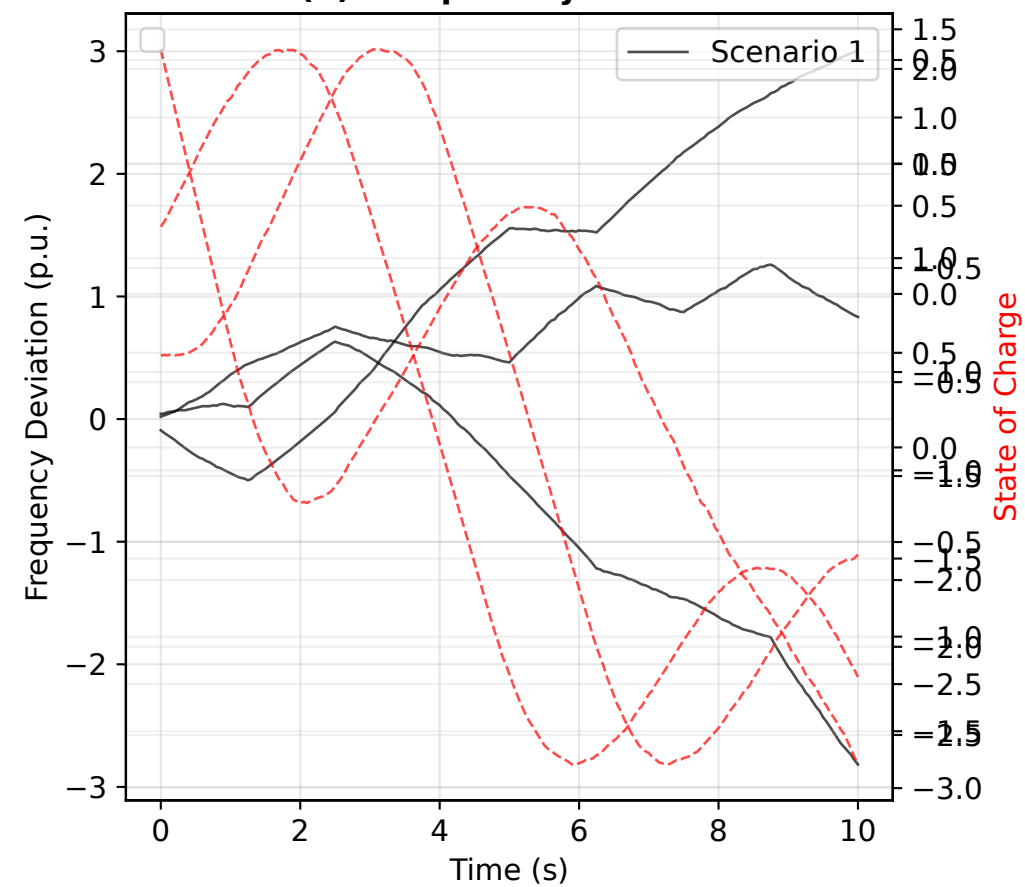
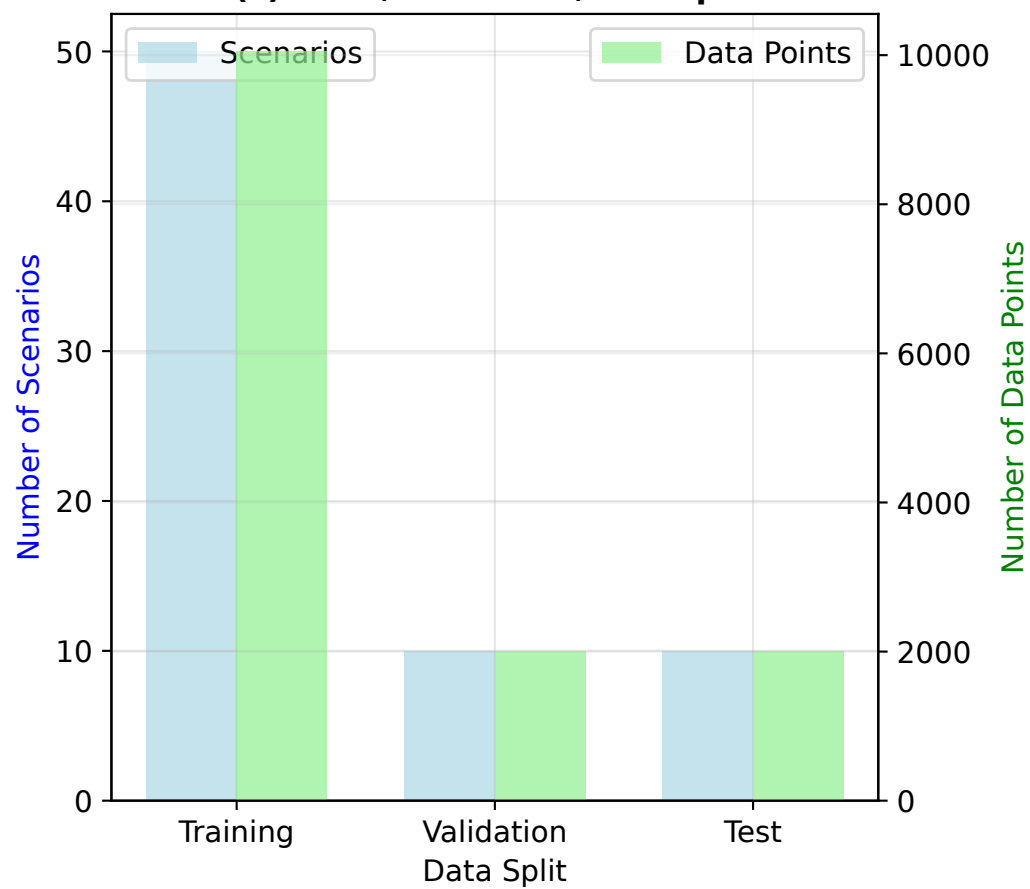
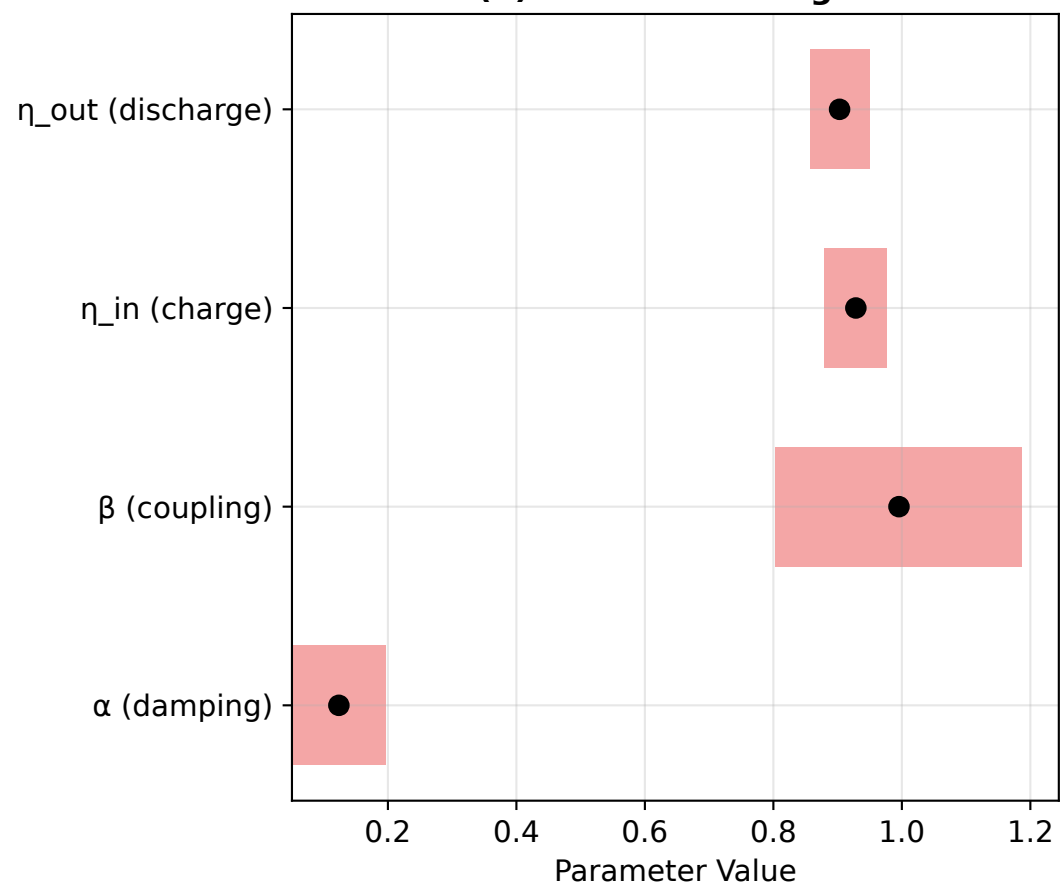


## Schematic — Runtime comparison (measurements missing)

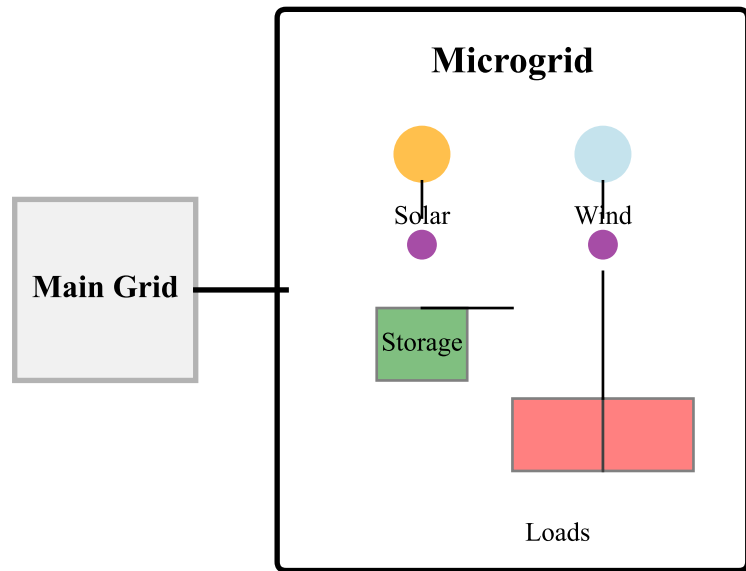




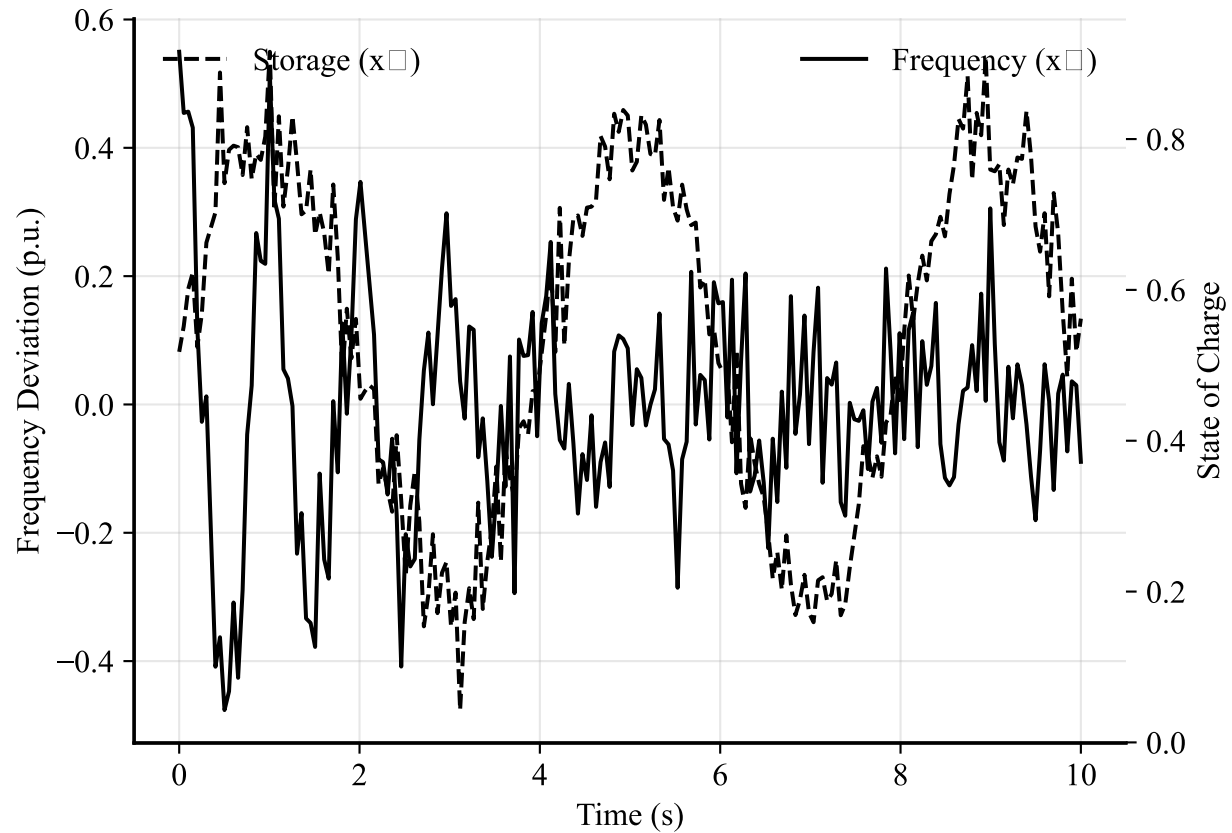


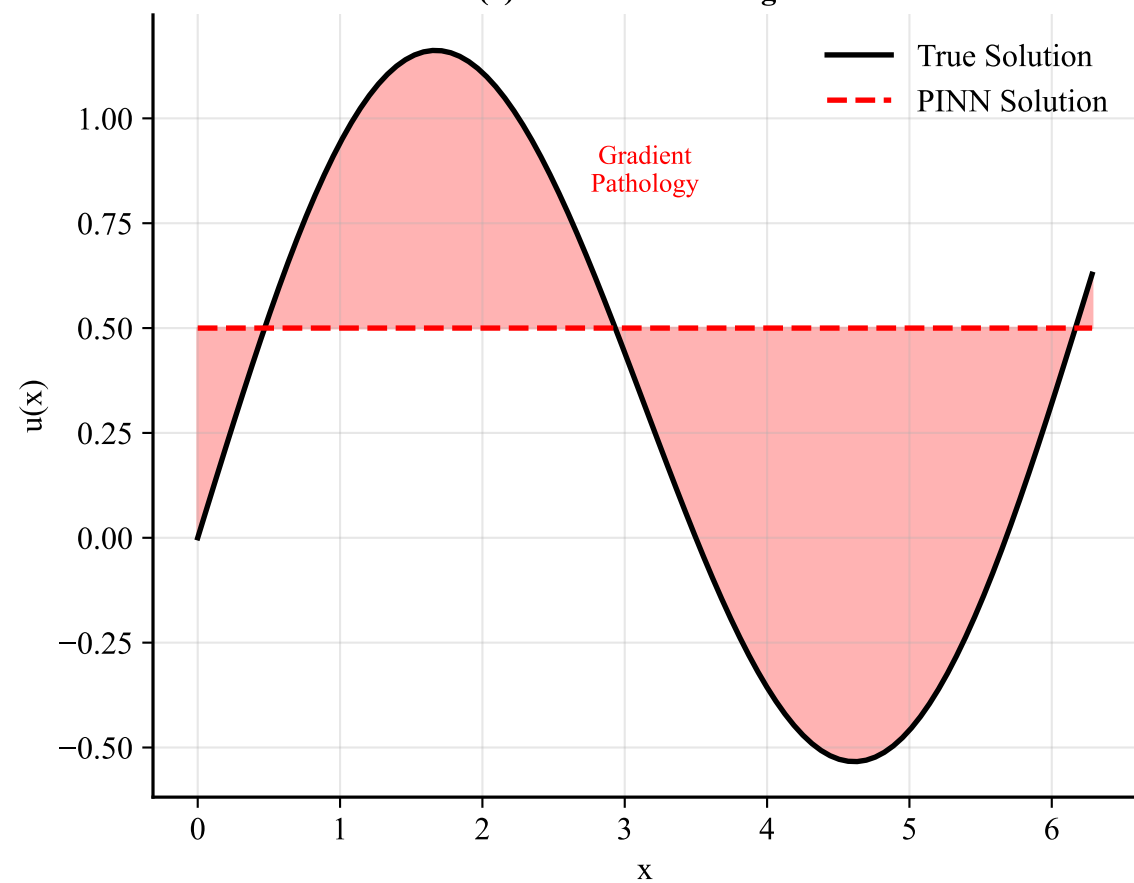
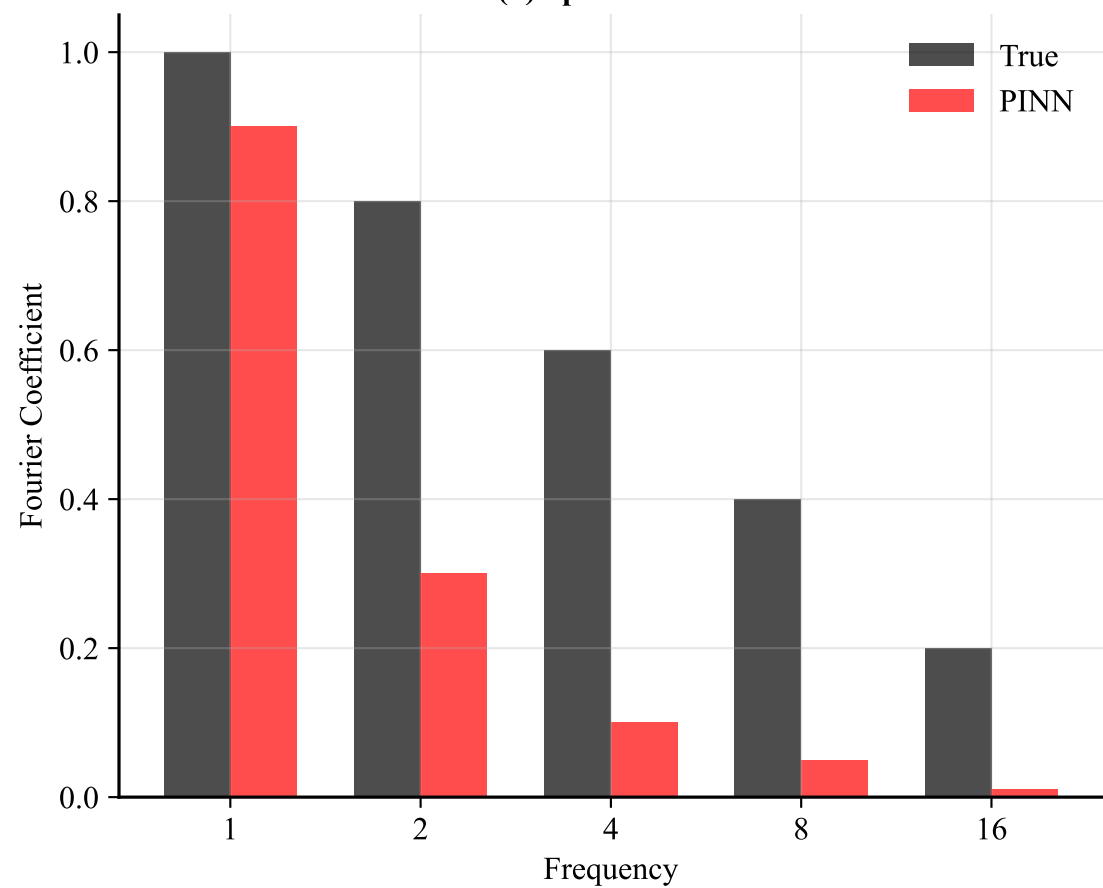
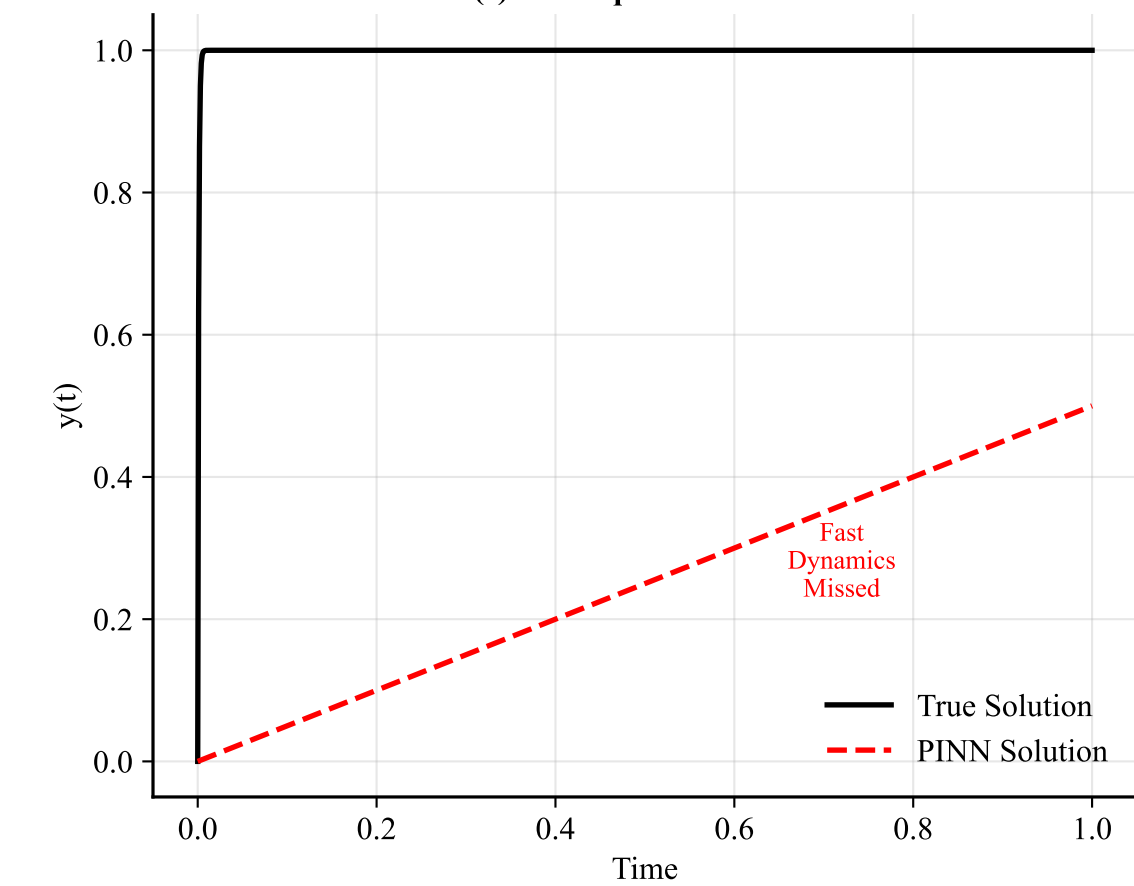
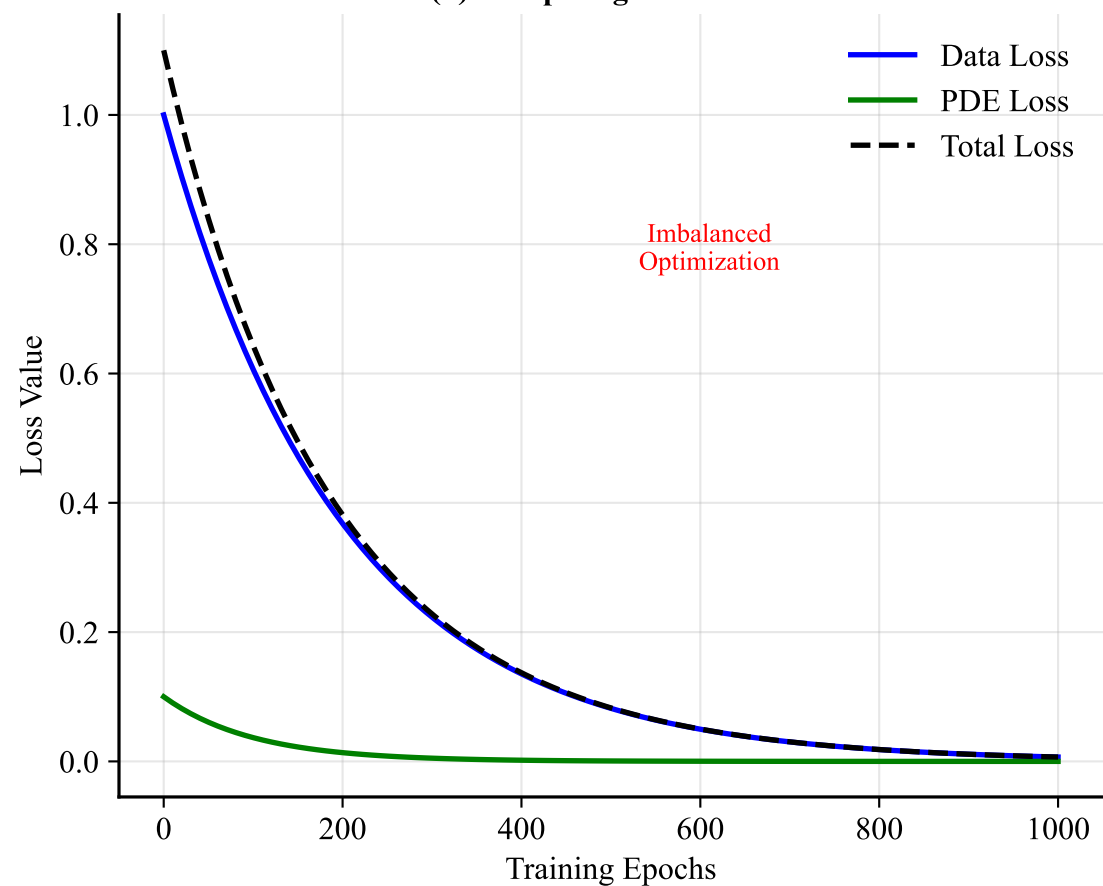
**(a) Training Scenarios Parameter Distribution****(b) Sample Trajectories****(c) Train/Validation/Test Split****(d) Parameter Ranges**

**(a) Microgrid Architecture**

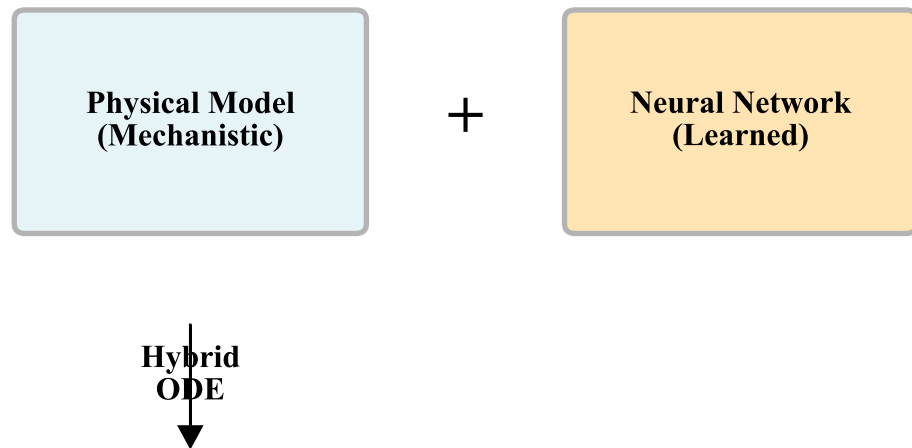


**(b) System Dynamics**

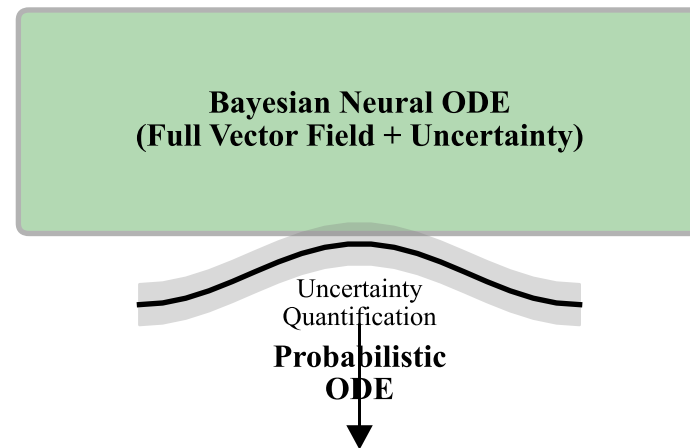


**(a) Gradient Pathologies****(b) Spectral Bias****(c) Stiff Equation Failure****(d) Competing Loss Terms**

**(a) Universal Differential Equation (UDE)**



**(b) Bayesian Neural ODE (BNODE)**



## (a) Two-State Microgrid Model

### Variables:

$x_1$ : Storage state-of-charge

$$\frac{dx_1}{dt} = \eta_{in} \cdot u \cdot \mathbb{1}_{\{u > 0\}} - \frac{1}{\eta_{out}} \cdot u \cdot \mathbb{1}_{\{u < 0\}} - d(t)$$

$x_2$ : Frequency/power deviation

*Storage Dynamics (Conservation Law)*

$$\frac{dx_2}{dt} = -\alpha x_2 + \beta \cdot P_{gen} - \beta \cdot P_{load} + \gamma x_1$$

*Frequency Dynamics (Droop Control)*

$u$ : Control input

$P_{gen}, P_{load}$ : Generation/load power

## (b) UDE Modification

$\beta \cdot P_{gen}$  (Original)

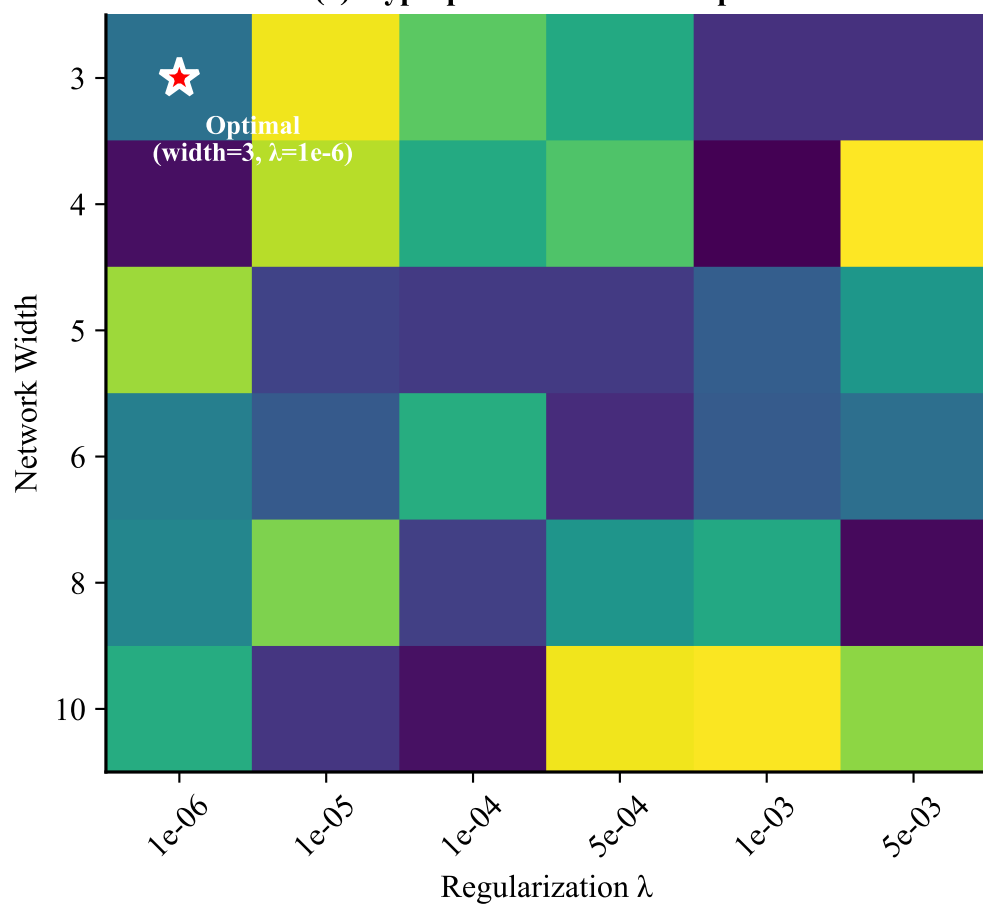


$f_{\theta}(P_{gen})$  (Neural Residual)

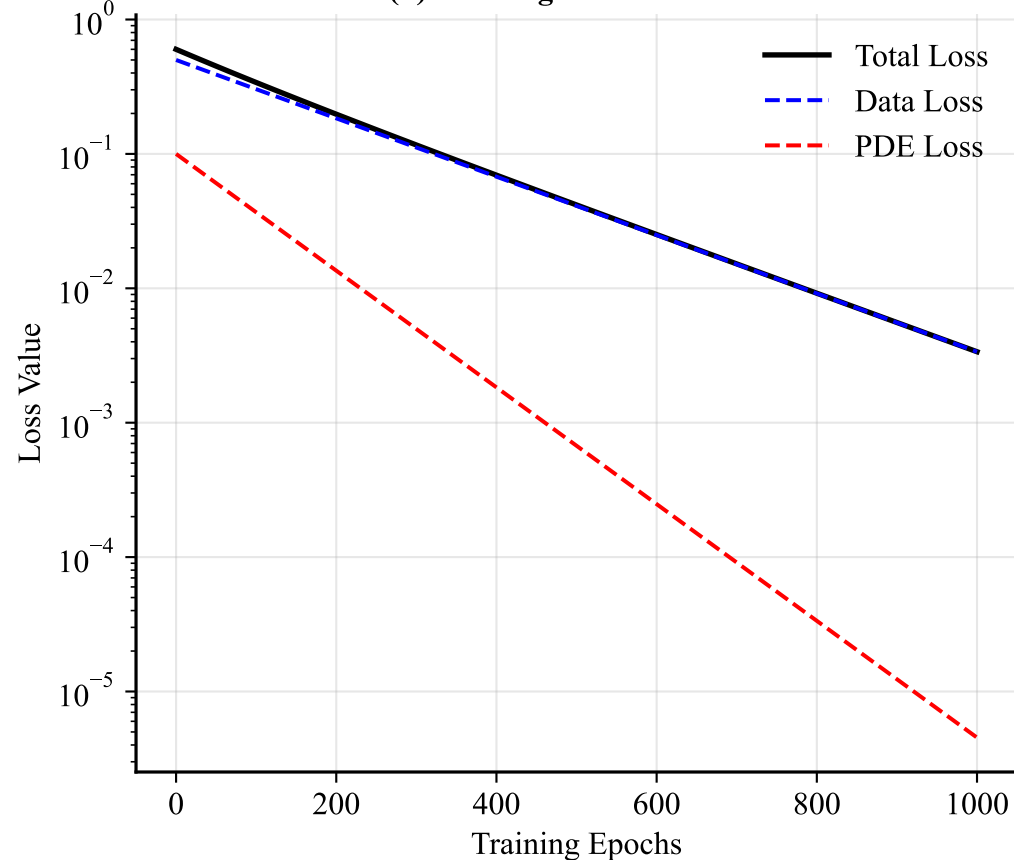
$$f_{\theta}(P) = \sum_{i=1}^3 w_i \tanh(W_{i1}P + b_i)$$

Single hidden layer, 3 units, 9 parameters

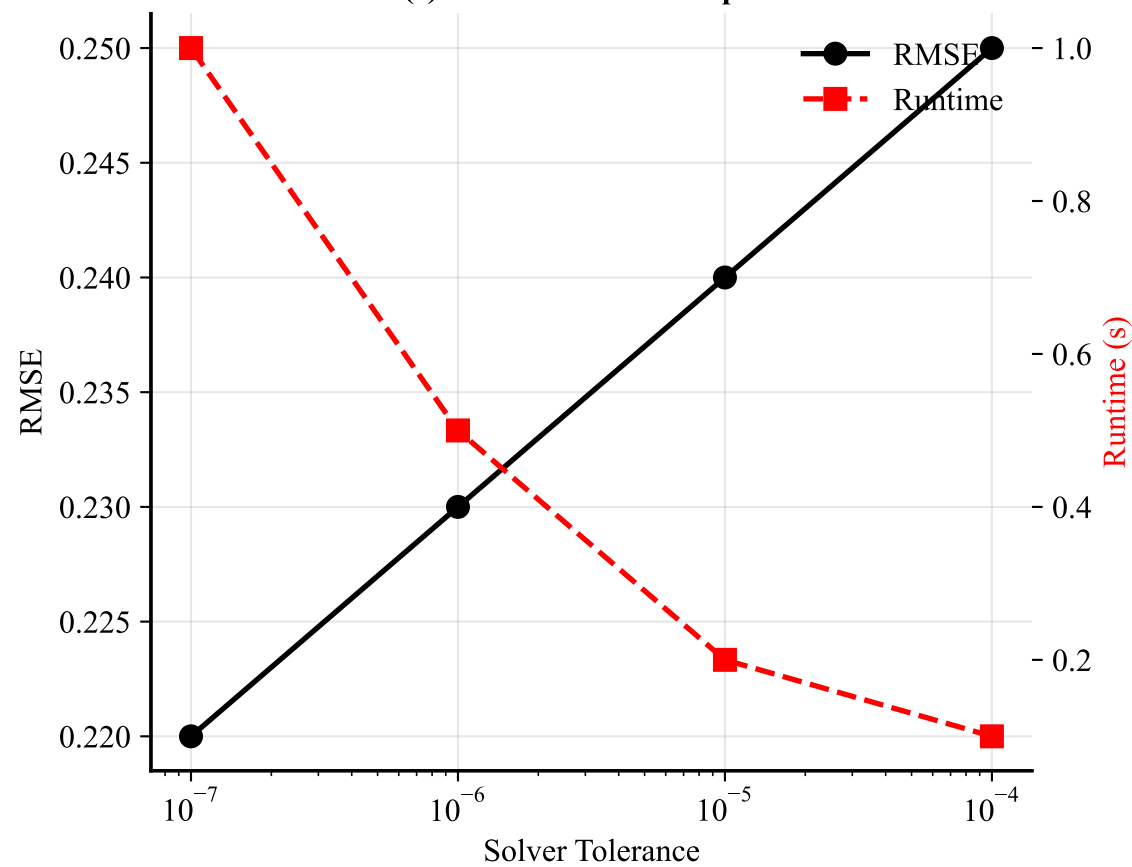
(a) Hyperparameter Search Space



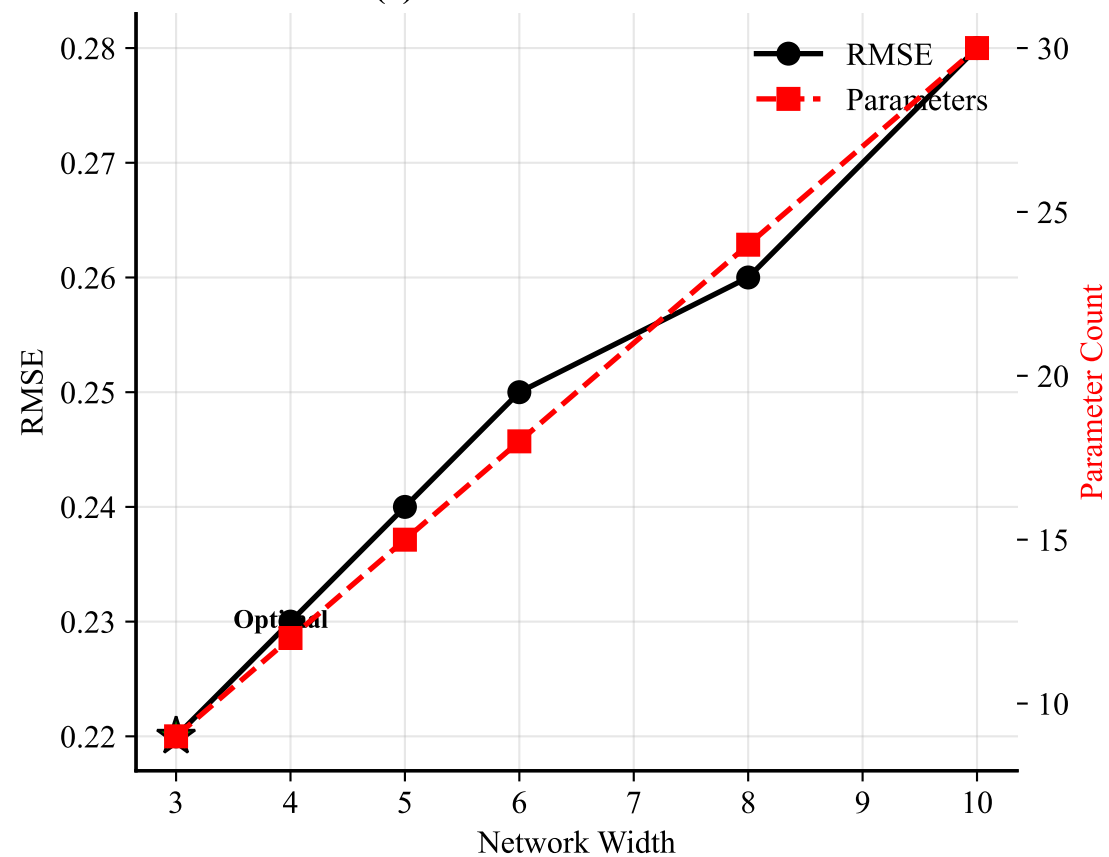
(b) Training Loss Evolution

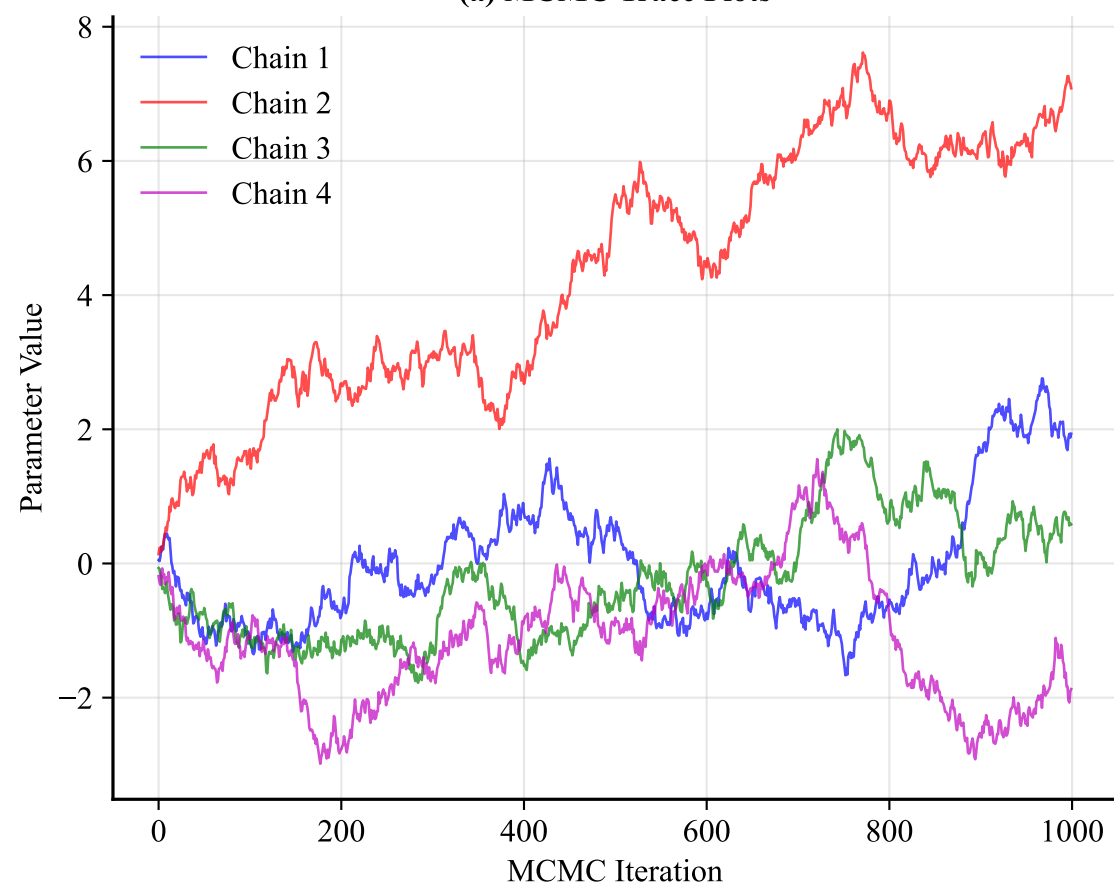
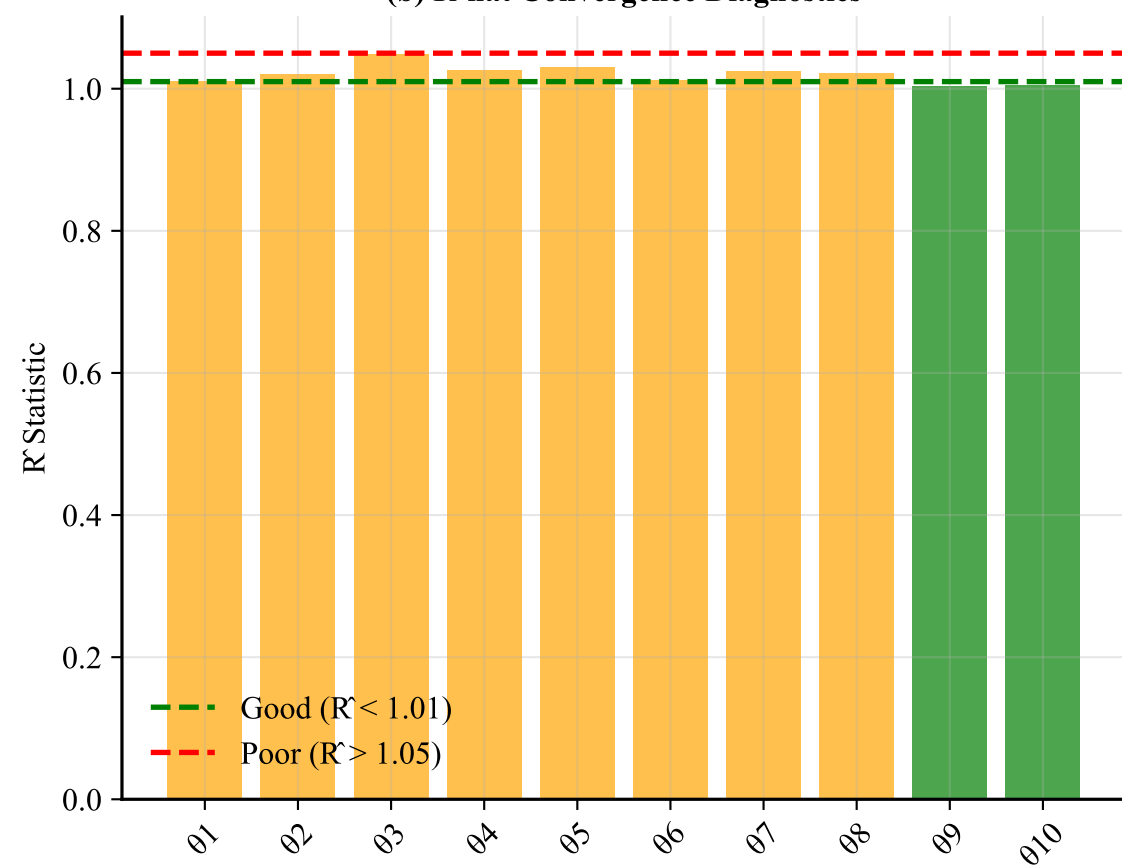
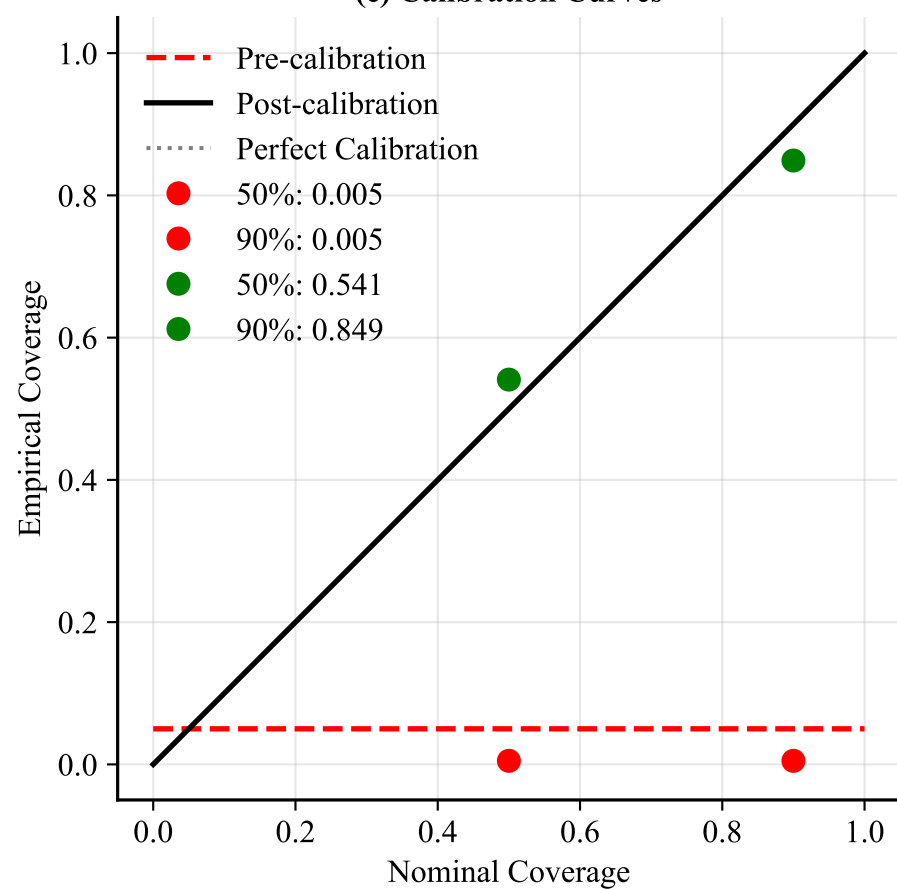
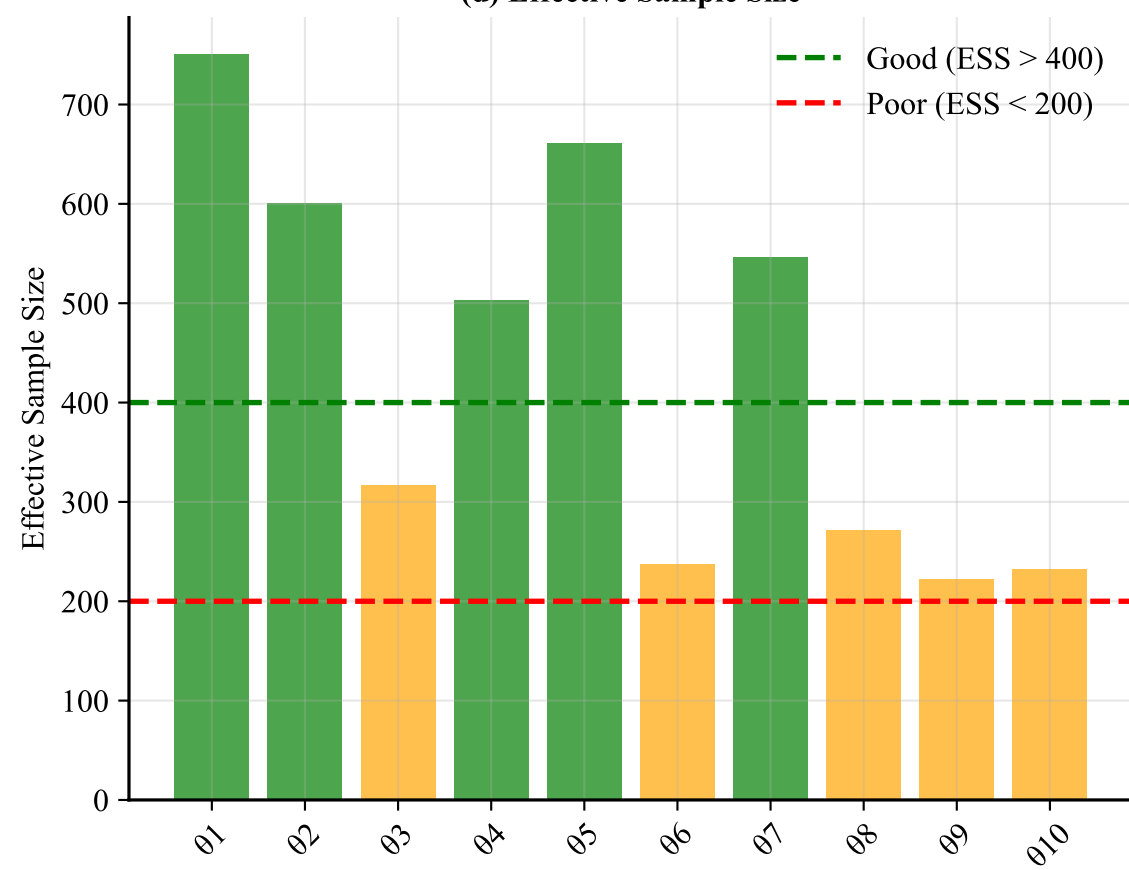


(c) Solver Tolerance Impact

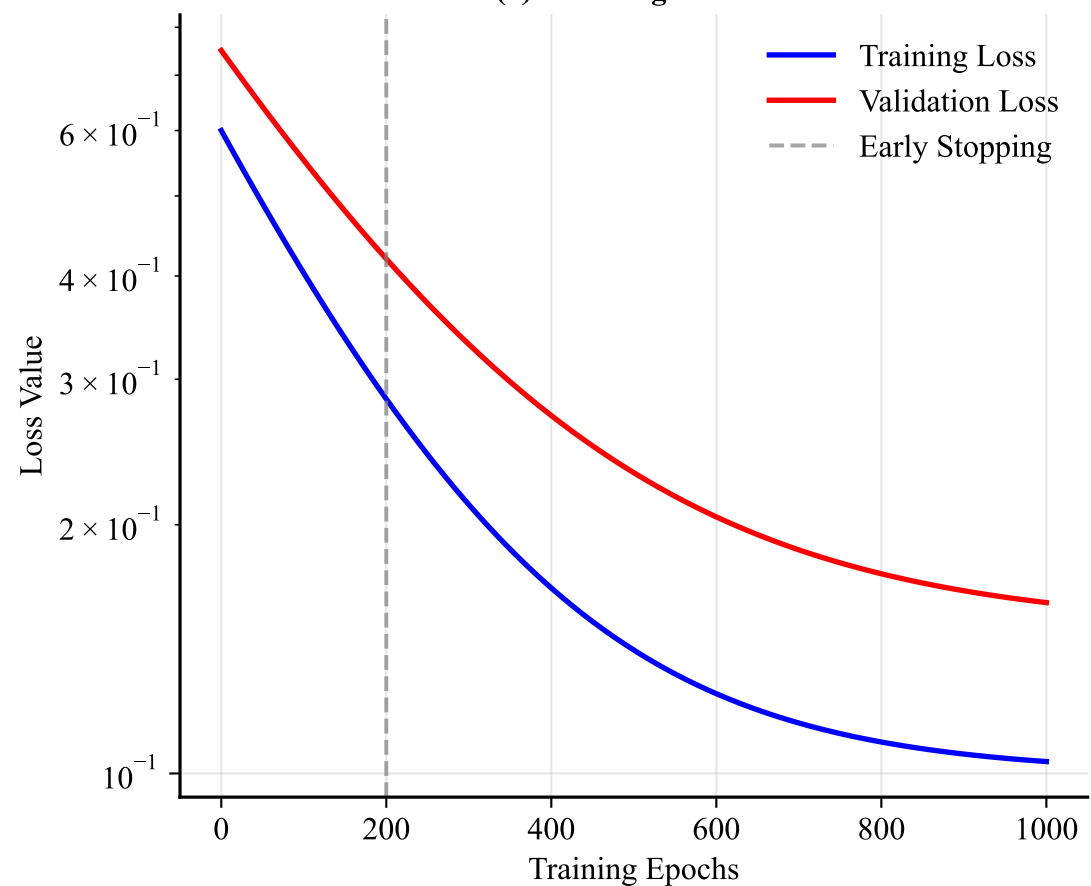
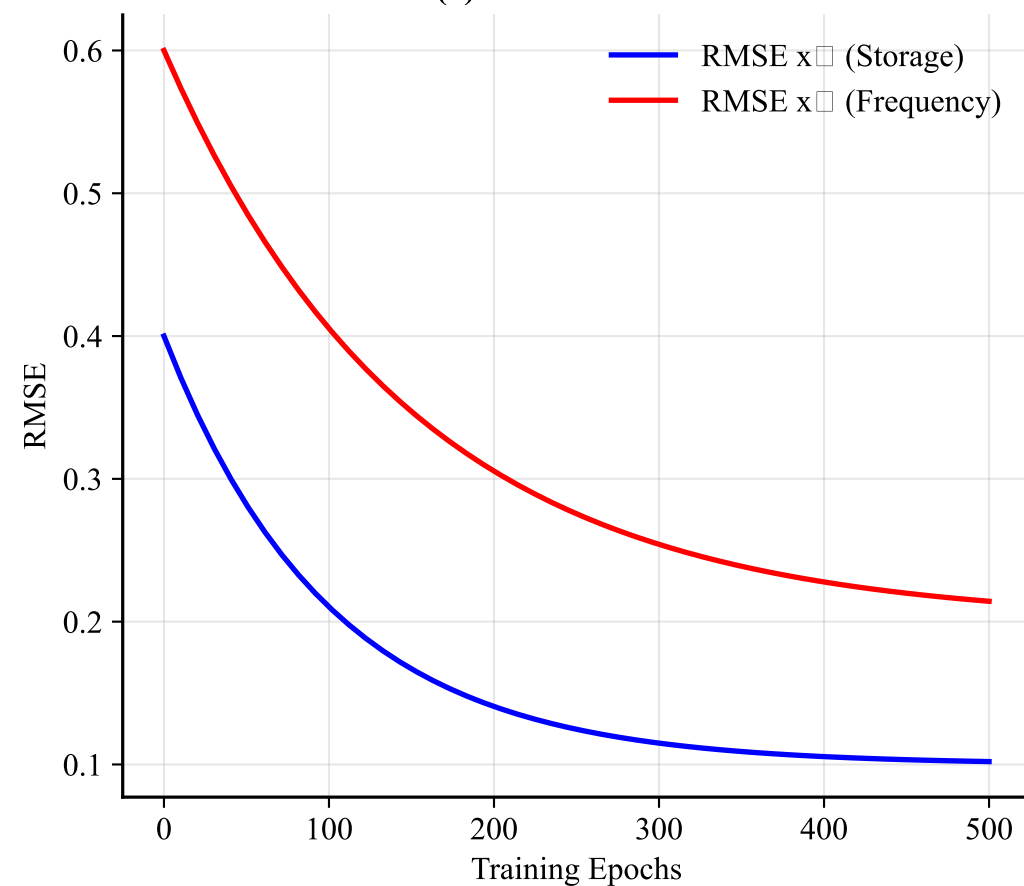
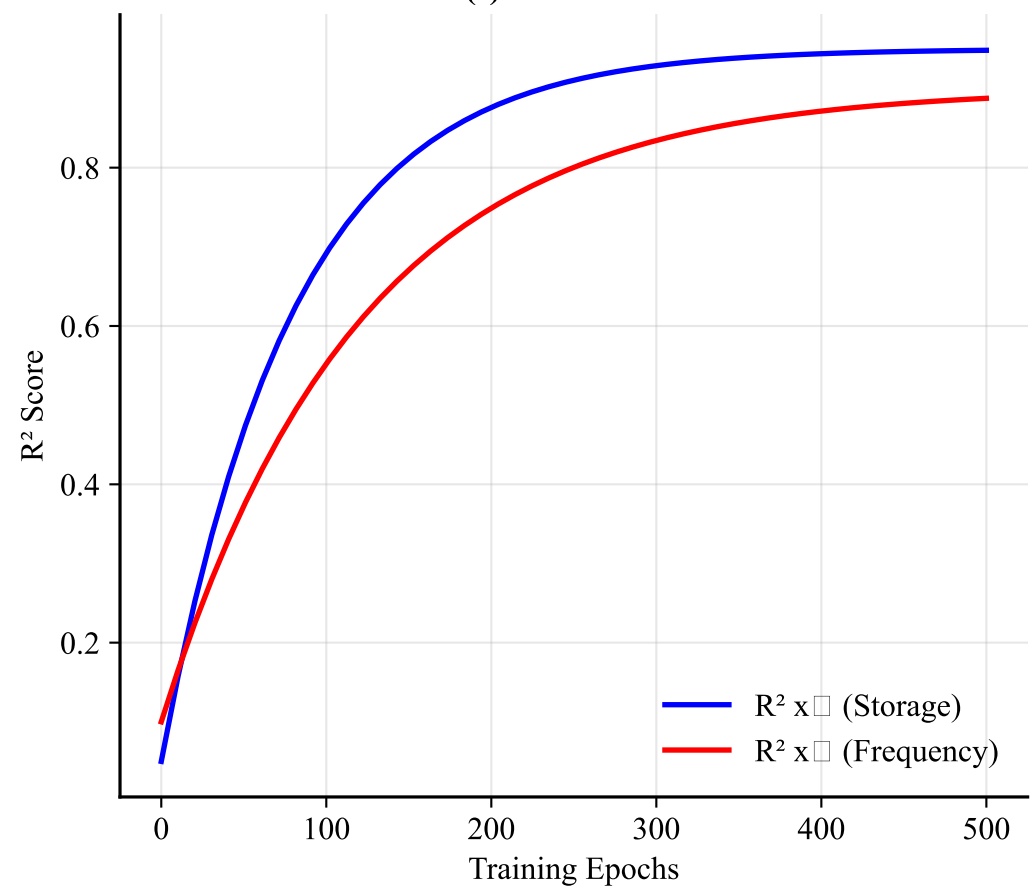
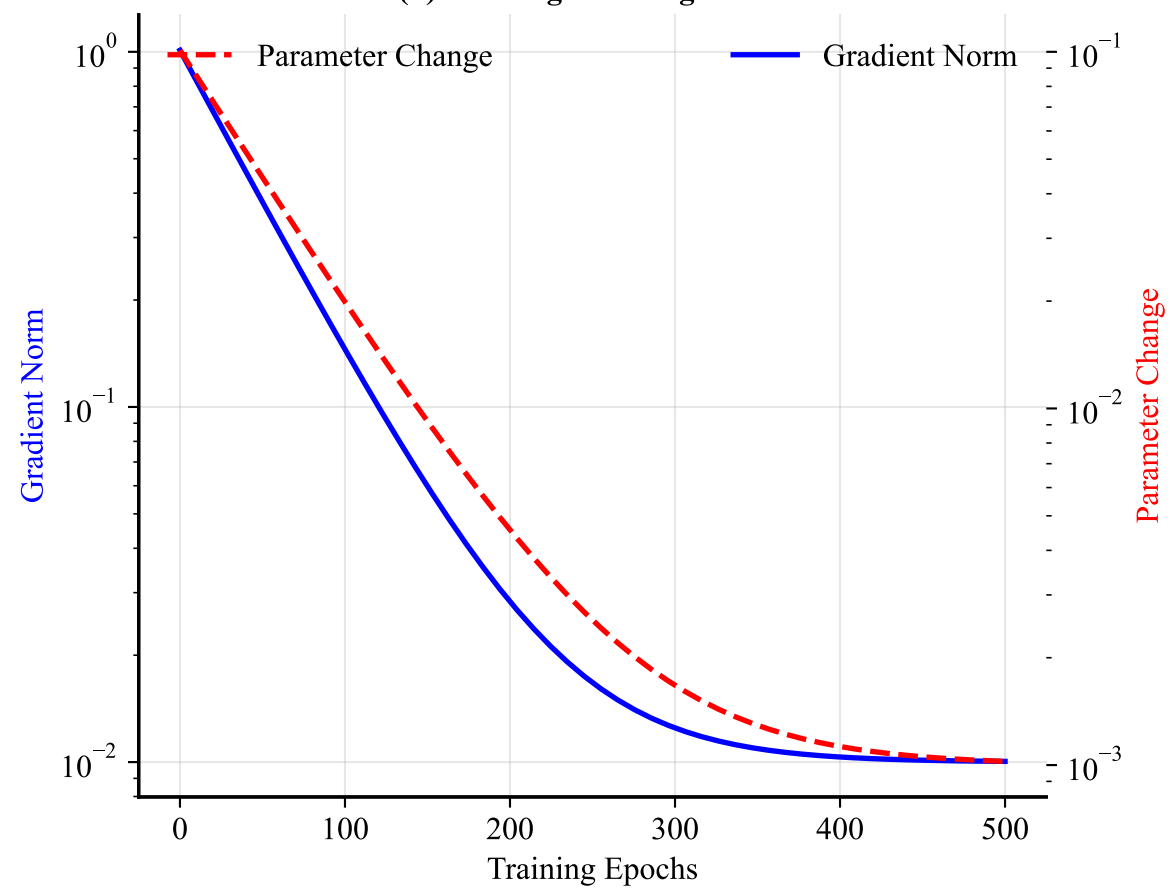


(d) Architecture Performance

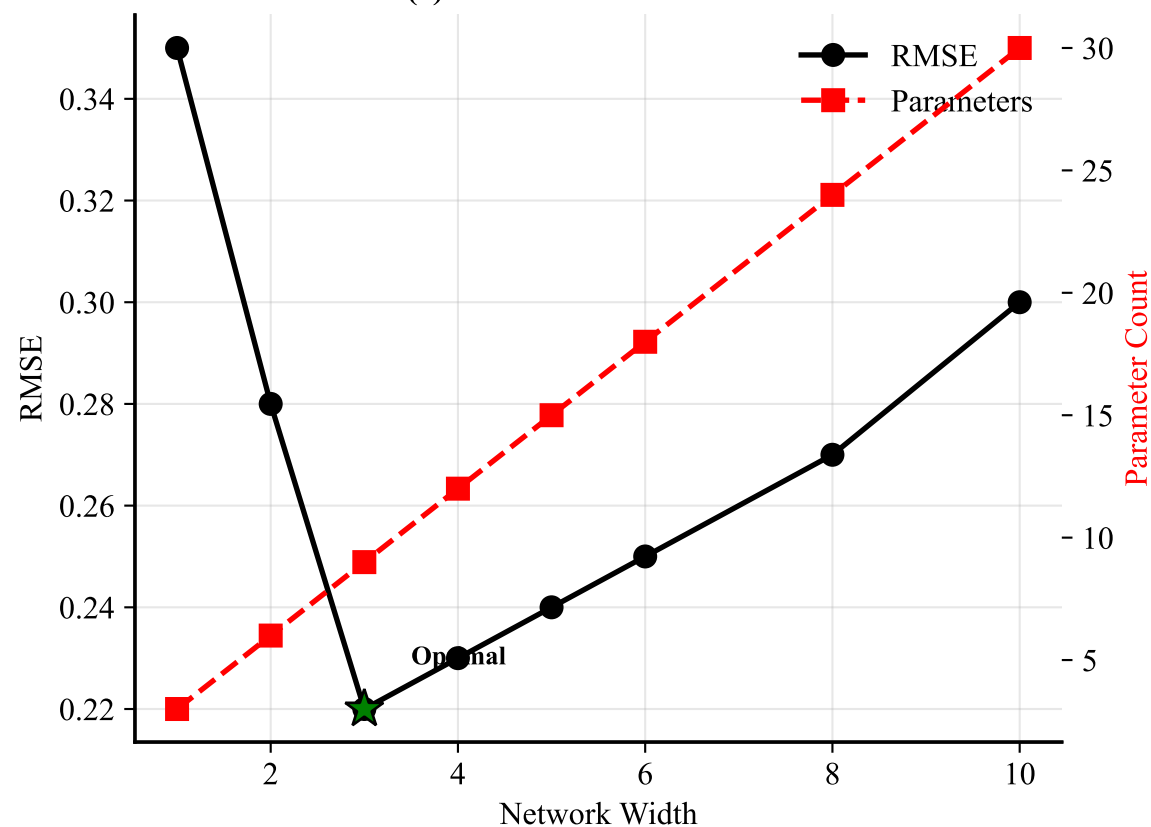


**(a) MCMC Trace Plots****(b) R-hat Convergence Diagnostics****(c) Calibration Curves****(d) Effective Sample Size**

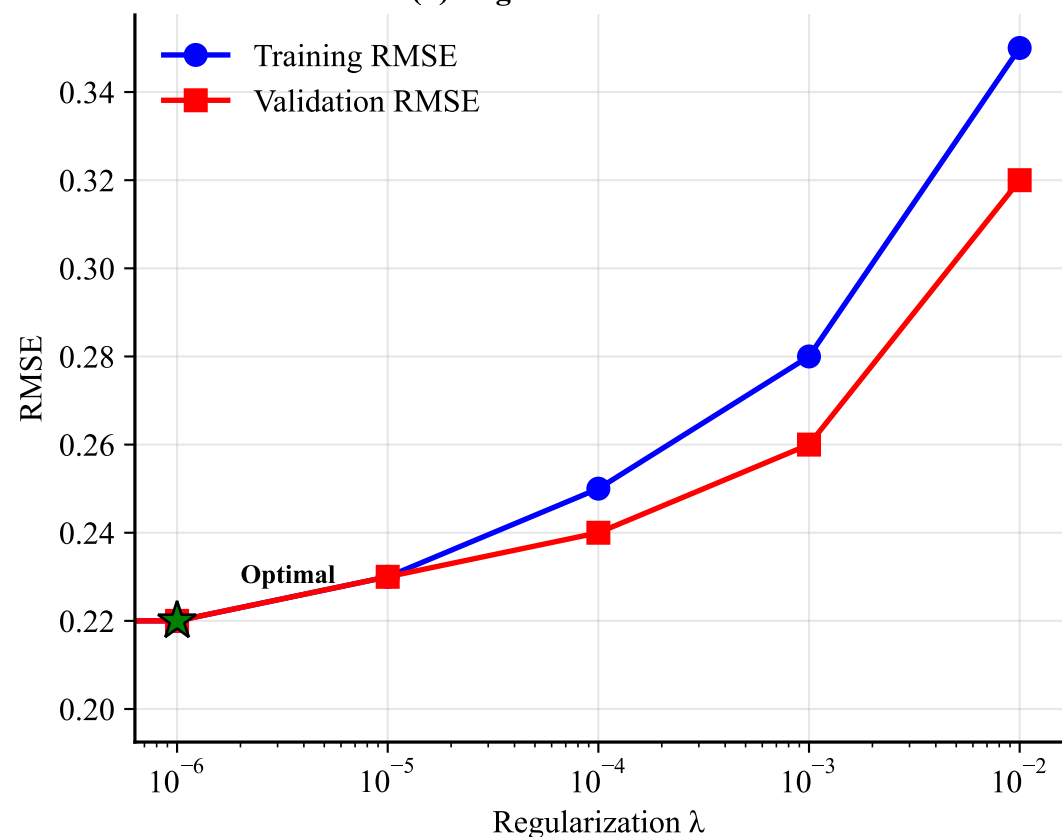


**(a) Learning Curves****(b) RMSE Evolution****(c) R<sup>2</sup> Evolution****(d) Convergence Diagnostics**

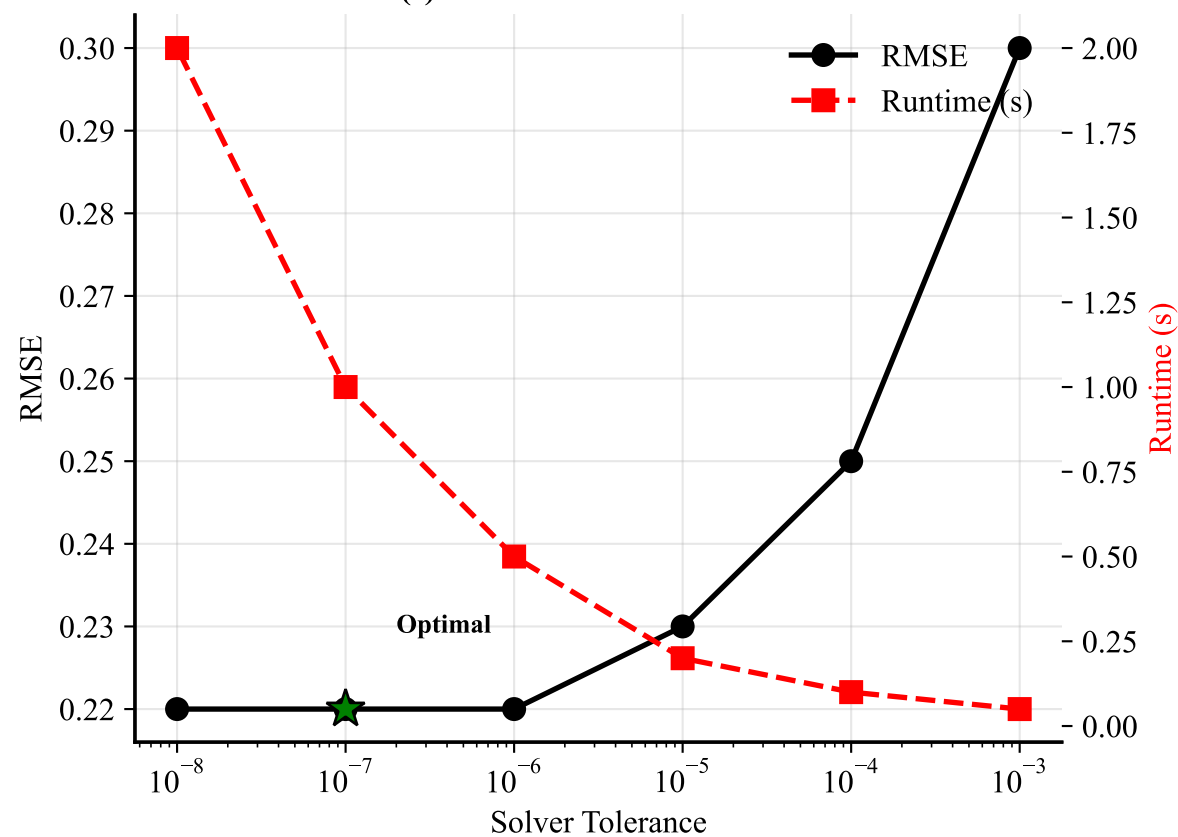
(a) Network Width Ablation



(b) Regularization Ablation



(c) Solver Tolerance Ablation



(d) Loss Weighting Ablation

