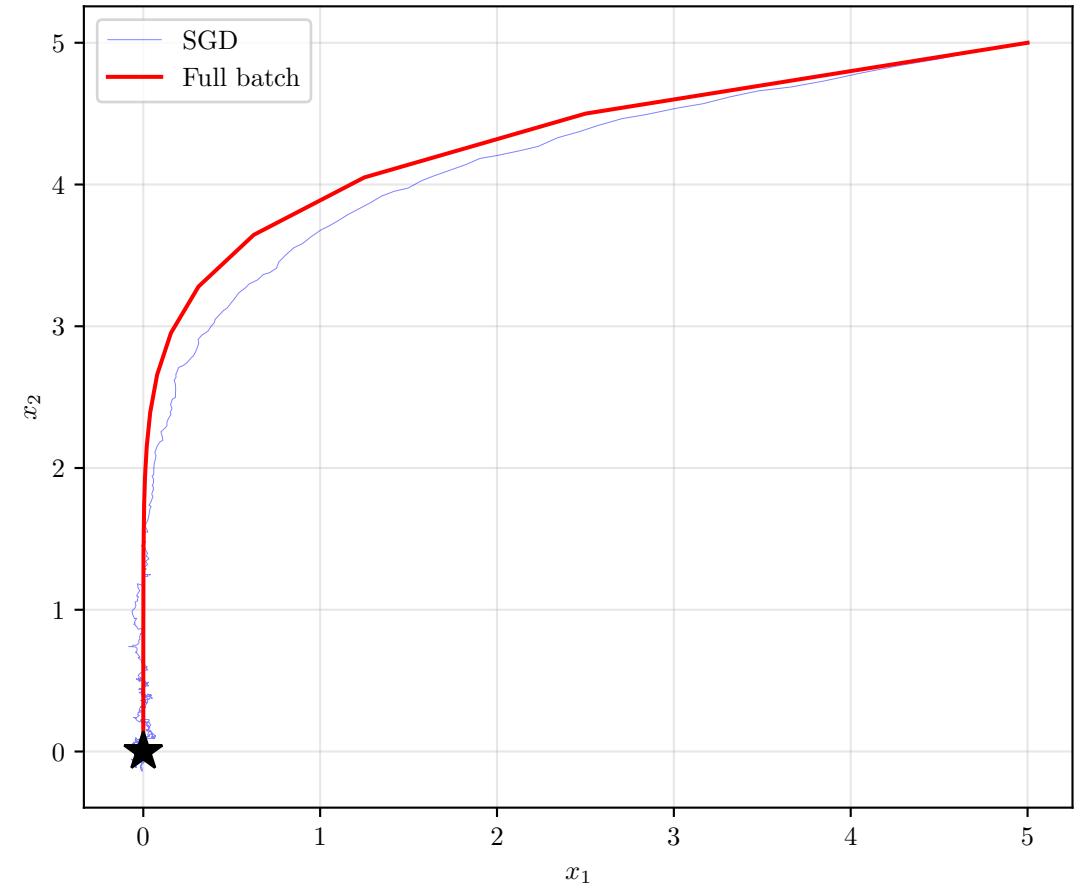
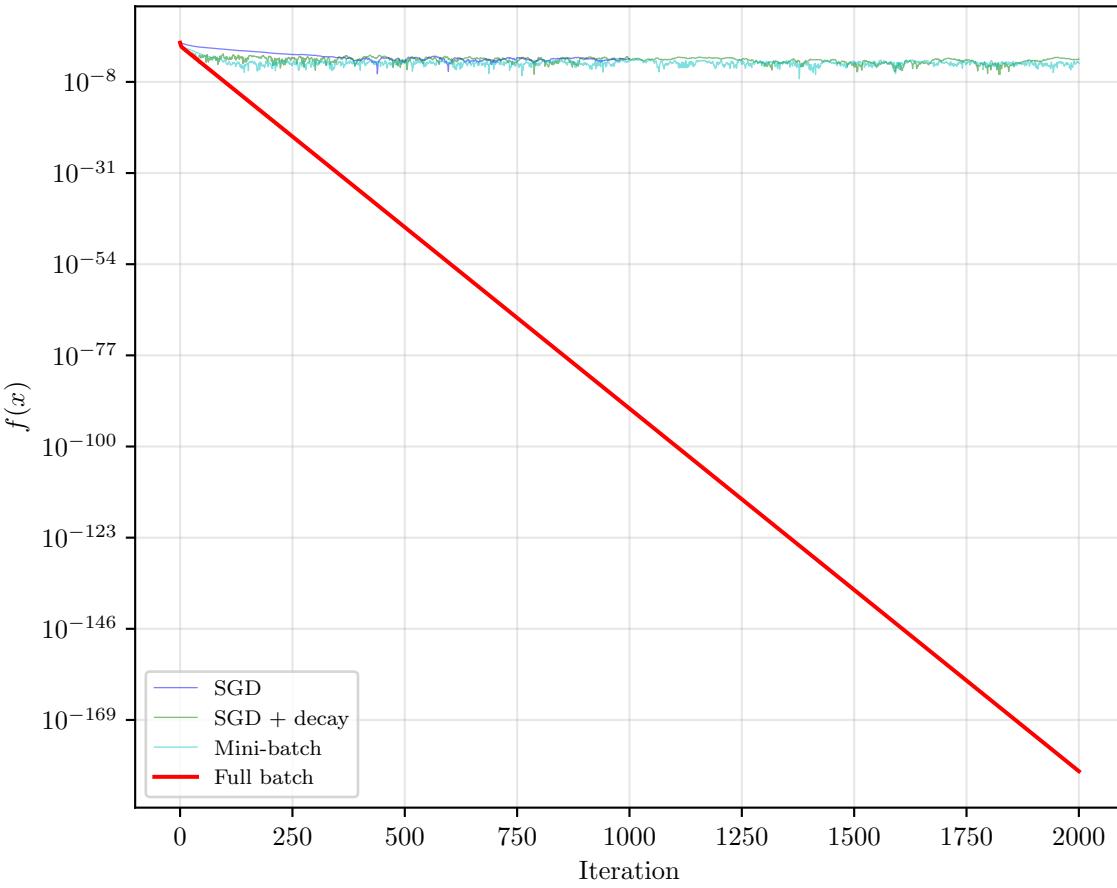


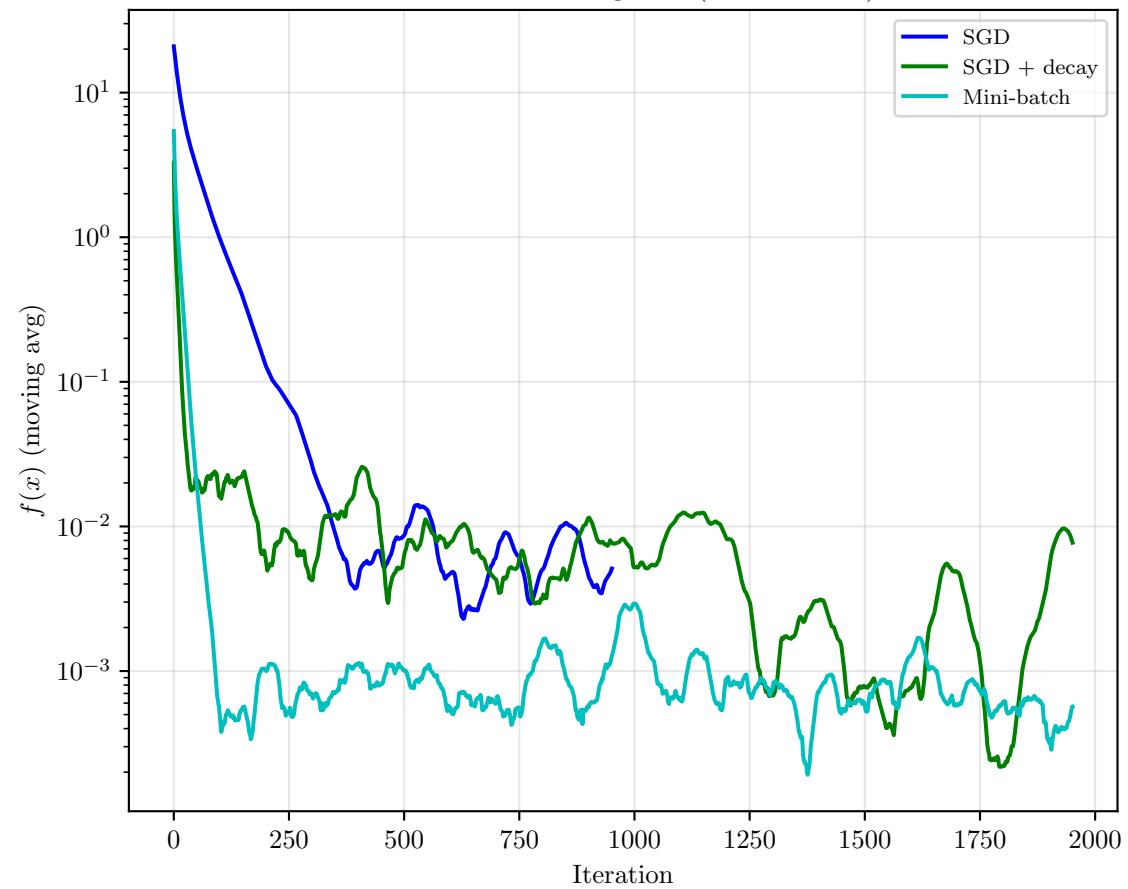
SGD vs Full Batch GD



Stochastic Optimization Convergence



Smoothed Convergence (window=50)



Variance Reduction with Batch Size

