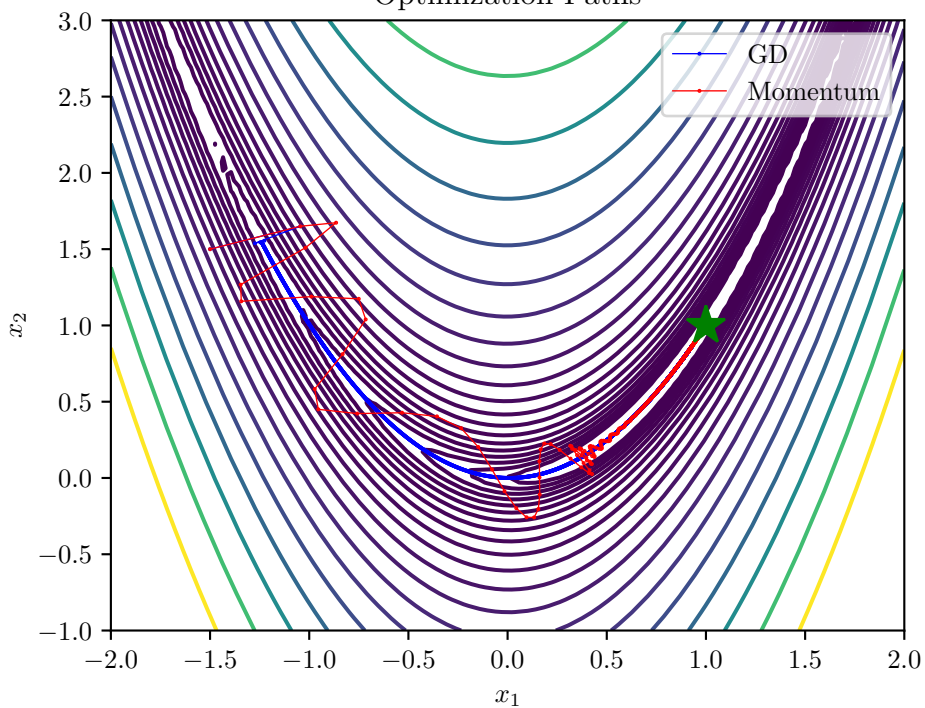
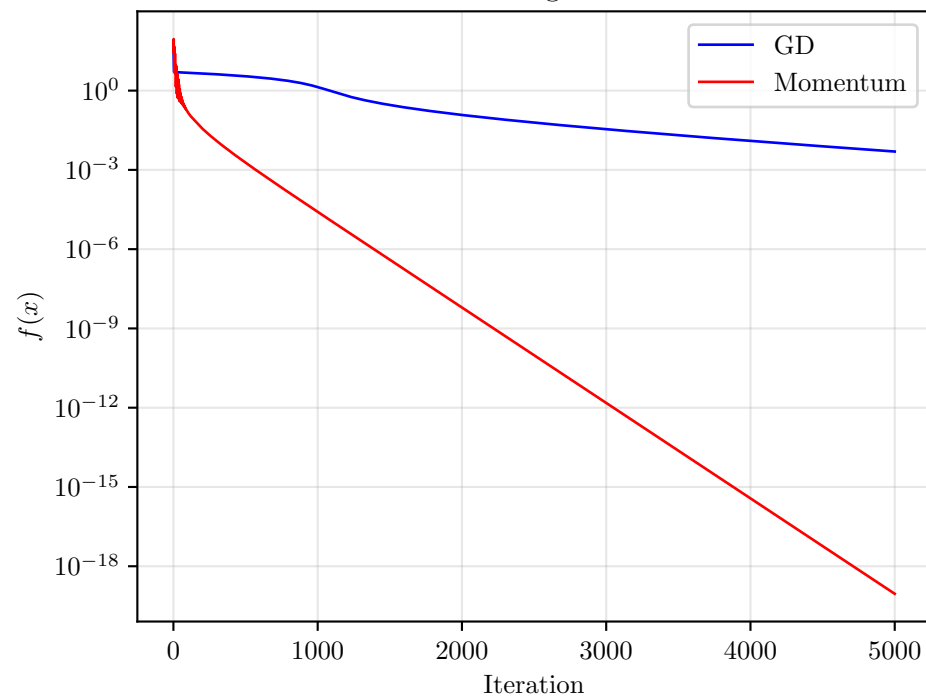


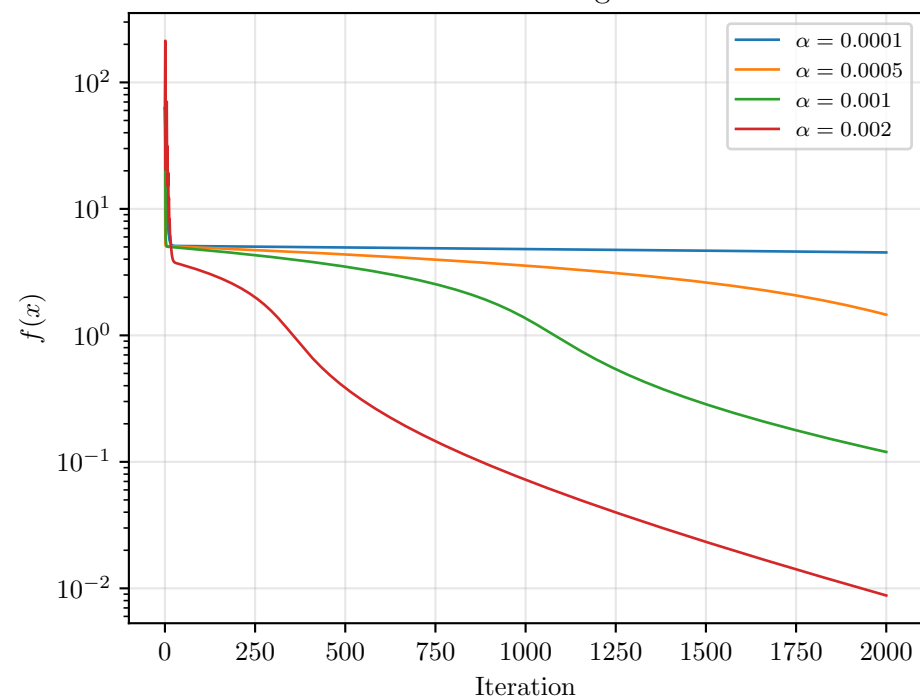
Optimization Paths



Convergence



Effect of Learning Rate



1D Function and Gradient

