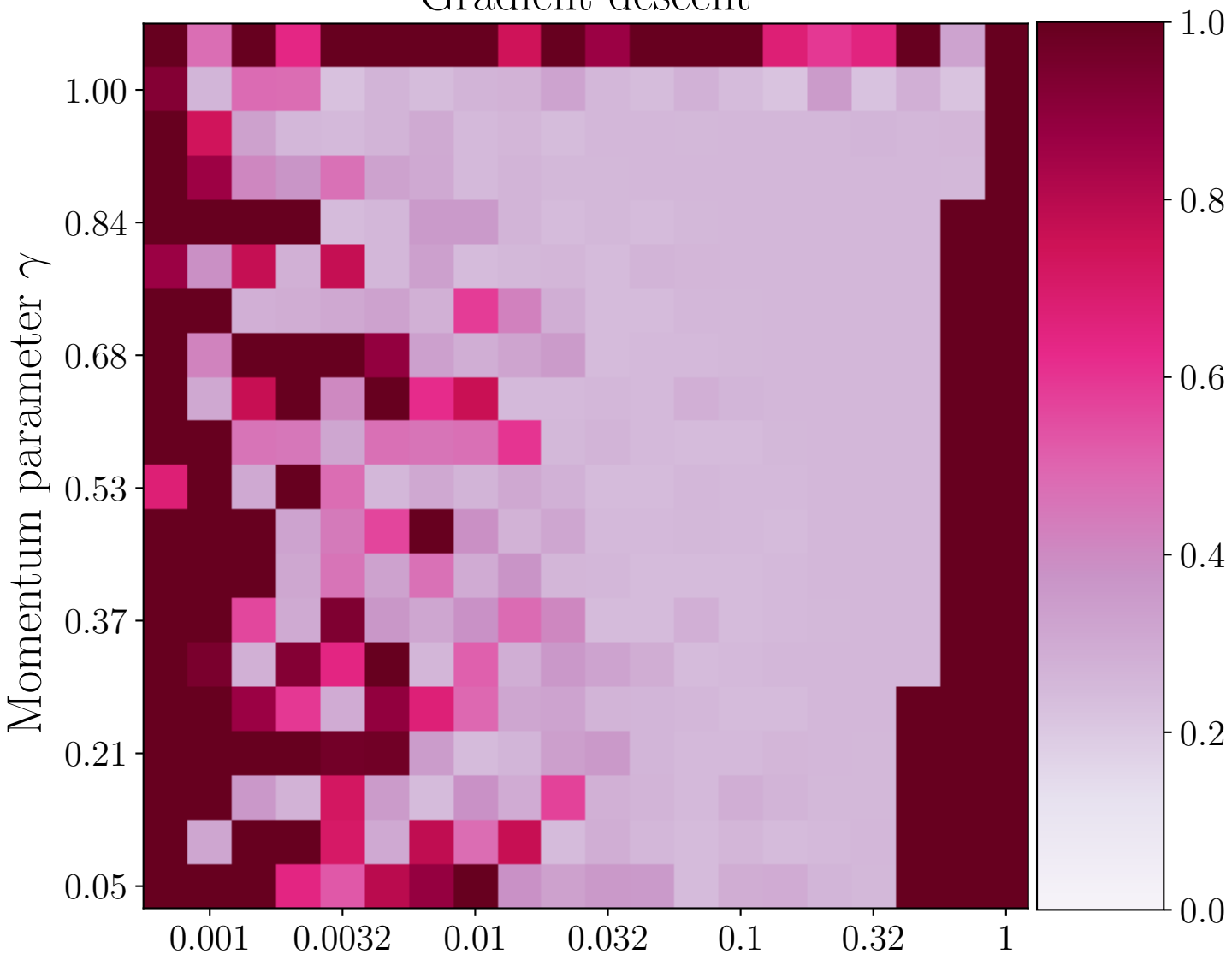
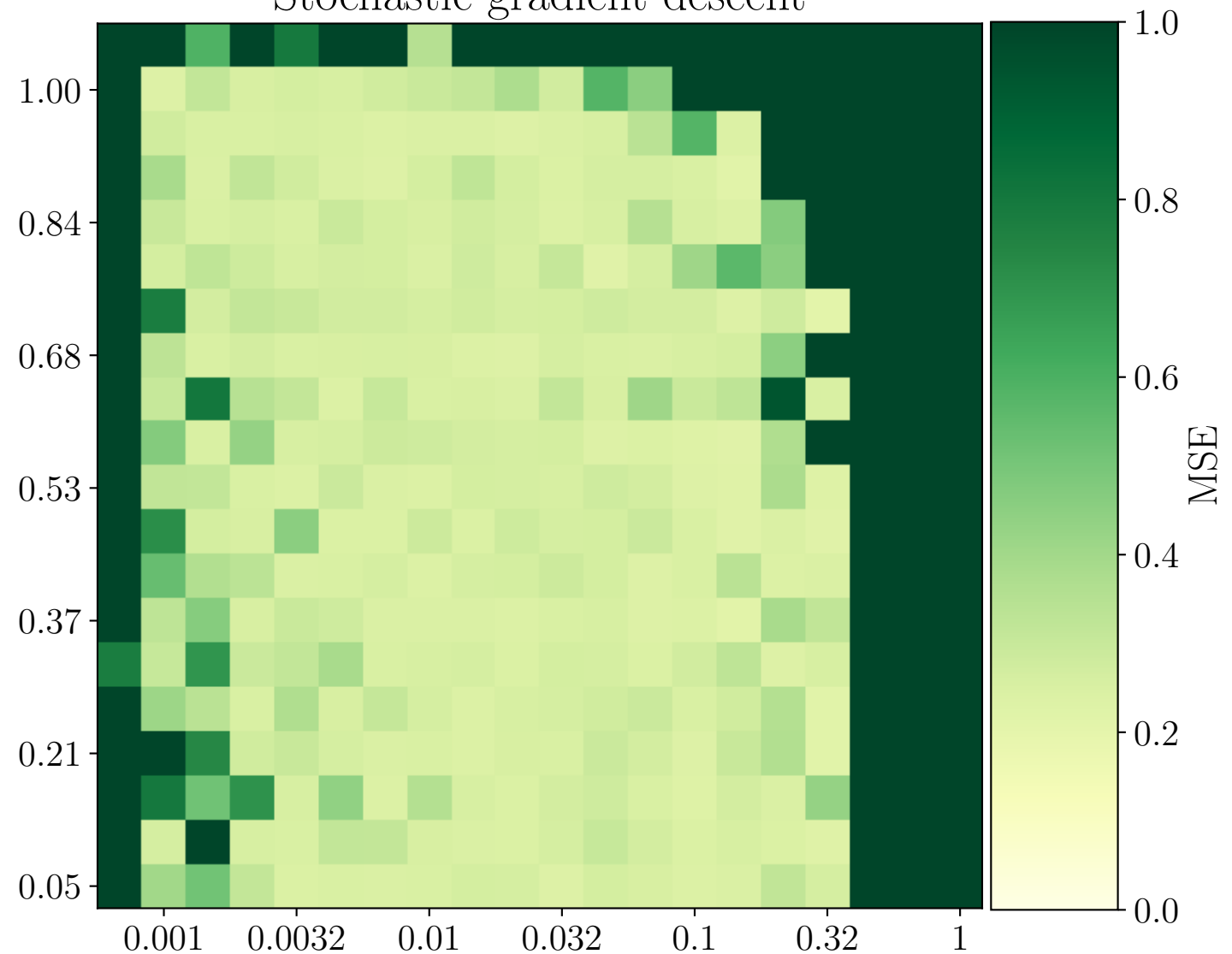


Gradient descent



Stochastic gradient descent

Learning rate  $\eta$