

MSE convergence in SGD with momentum

