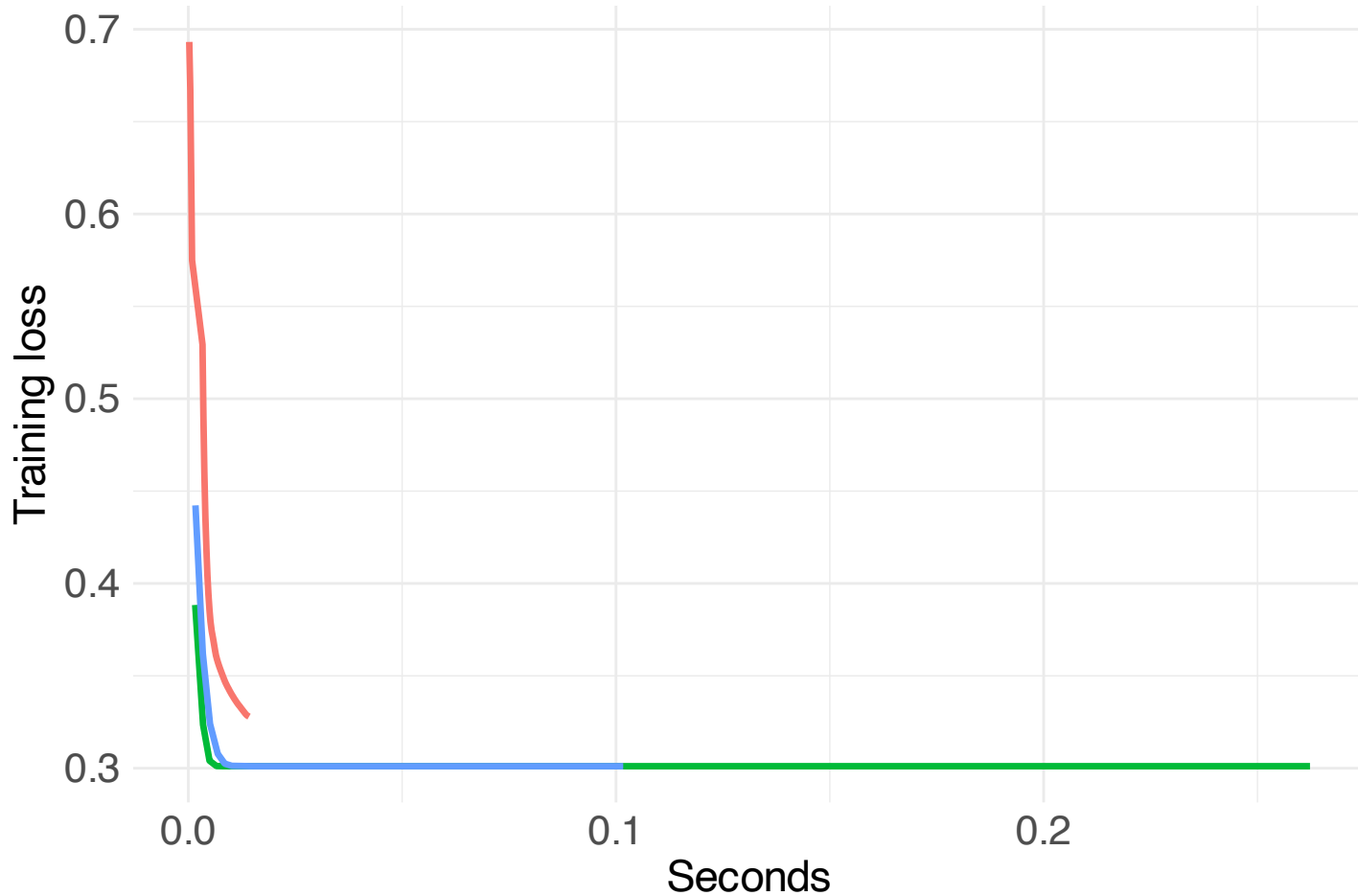
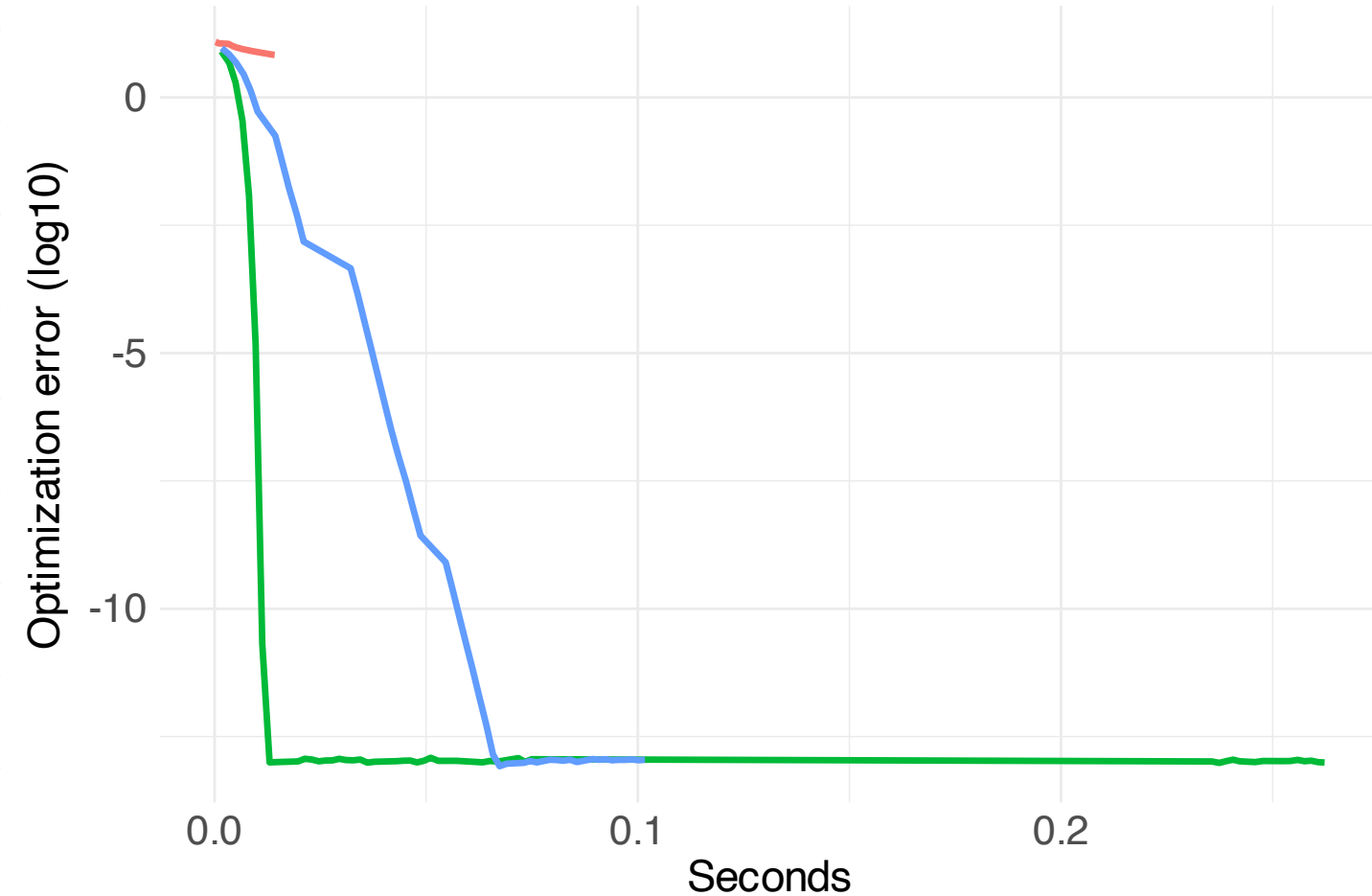


Training loss



Parameter optimization error



method — GD+mom — Newton-Raphson — Newton-Raphson+step size