Convergence of the validation loss as a function of the number of steps