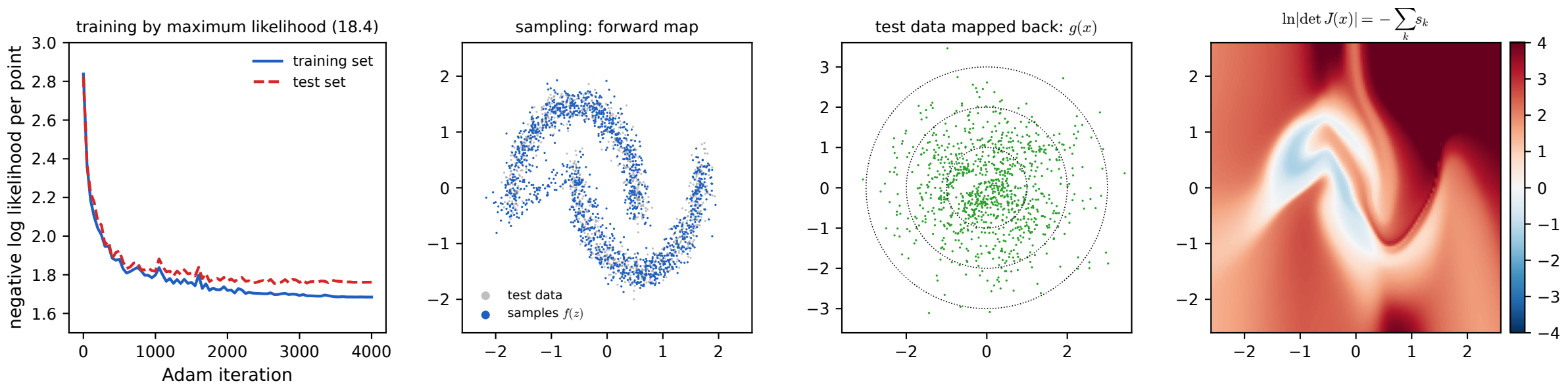

training by maximum likelihood (18.4)
negative log likelihood per point
Adam iteration
training set
test set
sampling: forward map
test data
samples $f(z)$
test data mapped back: $g(x)$
$\ln|\det J(x)| = -\sum_k s_k$