Input 1: differentiable simulator or variational approximation $g(\boldsymbol{\theta})$.
Input 2: initial parameter values $\boldsymbol{\theta}_s$.
Input 3: parameter of interest $\omega_0 = \theta_k$.
Output: learned summary statistic $\boldsymbol{s}(D; \boldsymbol{\phi})$.

for $i = 1$ to N **do**
    Sample a representative mini-batch G_s from $g(\boldsymbol{\theta}_s)$.
    Compute differentiable summary statistic $\hat{\boldsymbol{s}}(G_s; \boldsymbol{\phi})$.
    Construct Asimov likelihood $\mathcal{L}_A(\boldsymbol{\theta}, \boldsymbol{\phi})$.
    Get information matrix inverse $I(\boldsymbol{\theta})^{-1} = \boldsymbol{H}_{\boldsymbol{\theta}}^{-1}(\log \mathcal{L}_A(\boldsymbol{\theta}, \boldsymbol{\phi}))$.
    Obtain loss $U = I_{kk}^{-1}(\boldsymbol{\theta}_s)$.
    Update network parameters $\boldsymbol{\phi} \rightarrow \mathrm{SGD}(\nabla_{\boldsymbol{\phi}} U)$.
end for