**Supplementary Note B**

**Algorithms for implementing stochastic gradient decent**

Algorithm is a modified version of Algorithms in (4).

Let be the number of steps to apply to the discriminator.

**for** number of training iterations **do**

**for** steps **do**

1. Sampling minibatch of noise vectors from prior .
2. Sampling minibatch of samples of feature, observed outcomes, treatment and treatment assignment indicator data from data generating distributions.
3. Update the parameters in discriminator by descending its stochastic gradient:

. (B1)

**end if**

1. Sampling minibatch of noise vectors from prior .
2. Update the parameters in generator by descending its stochastic gradient:

. (B2)

**end if**