Our chief insight is that the $d$-separation properties of the directed graphical model (DGM) structure of the forward model should be used to produce an encoder factorization that is faithful to the posterior, in the sense that it does not impose independencies not present in the true posterior.

Auto-encoding variational inference requires the design of an encoder for learning either the model and/or the posterior.

Typically, the structure of the encoder is formed in an ad hoc way by simply reversing the edges in the DGM of the generative model.

We introduce a novel algorithm that given the factorization for a generative model, produces a faithful factorization for the posterior that is optimal in a certain technical sense.

A faithful encoder outperforms an ad hoc one at every capacity level.

After 10,000 epochs, reaches 86.5 nats, equivalent to results from an unfaithful encoder with 200 latent units and a multi-sample objective with 50 importance weighted samples (Maddison et al., 2016; Jang et al., 2016).

The unfaithful encoder is unable to learn the true posterior.

A fully connected MADE encoder (Germain et al., 2015) and an encoder based on a novel MADE variant for trees (both faithful, the later minimal) are able to learn the posterior.

The minimally faithful encoder learns quicker than the fully connected one. Why?

A preprint arXiv:1312.6114v10
