

Bengio *et al.* (2014) showed that if the autoencoder  $p(\mathbf{x} \mid \tilde{\mathbf{x}})$  forms a consistent estimator of the corresponding true conditional distribution, then the stationary distribution of the above Markov chain forms a consistent estimator (albeit an implicit one) of the data generating distribution of  $\mathbf{x}$ .

### 20.11.2 Clamping and Conditional Sampling

Similarly to Boltzmann machines, denoising autoencoders and their generalizations (such as GSNs, described below) can be used to sample from a conditional distribution  $p(\mathbf{x}_f \mid \mathbf{x}_o)$ , simply by clamping the *observed* units  $\mathbf{x}_f$  and only resampling the *free* units  $\mathbf{x}_o$  given  $\mathbf{x}_f$  and the sampled latent variables (if any). For example, MP-DBMs can be interpreted as a form of denoising autoencoder, and are able to sample missing inputs. GSNs later generalized some of the ideas present in MP-DBMs to perform the same operation (Bengio *et al.*, 2014). Alain *et al.* (2015) identified a missing condition from Proposition 1 of Bengio *et al.* (2014), which is that the transition operator (defined by the stochastic mapping going from one state of the chain to the next) should satisfy a property called **detailed balance**, which specifies that a Markov Chain at equilibrium will remain in equilibrium whether the transition operator is run in forward or reverse.

An experiment in clamping half of the pixels (the right part of the image) and running the Markov chain on the other half is shown in figure 20.12.