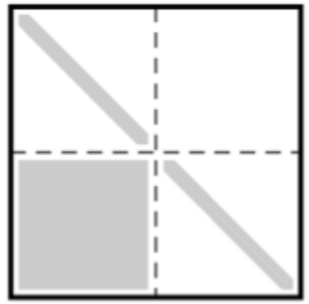
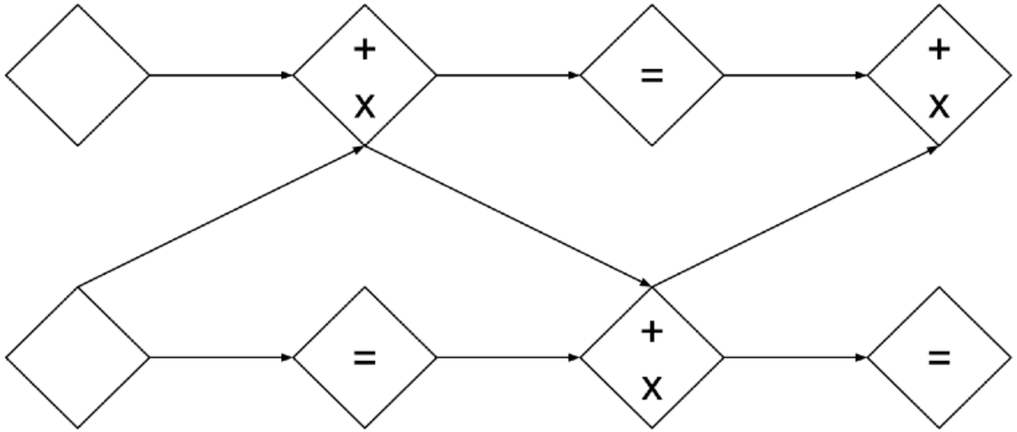
Mixing layer

We need all areas of vector to “talk to each other”

$$\begin{bmatrix} \mathbb{I} & 0 \\ \dfrac{\partial z_2}{\partial x_1} & diag(\exp[s(x_1)]) \end{bmatrix}$$



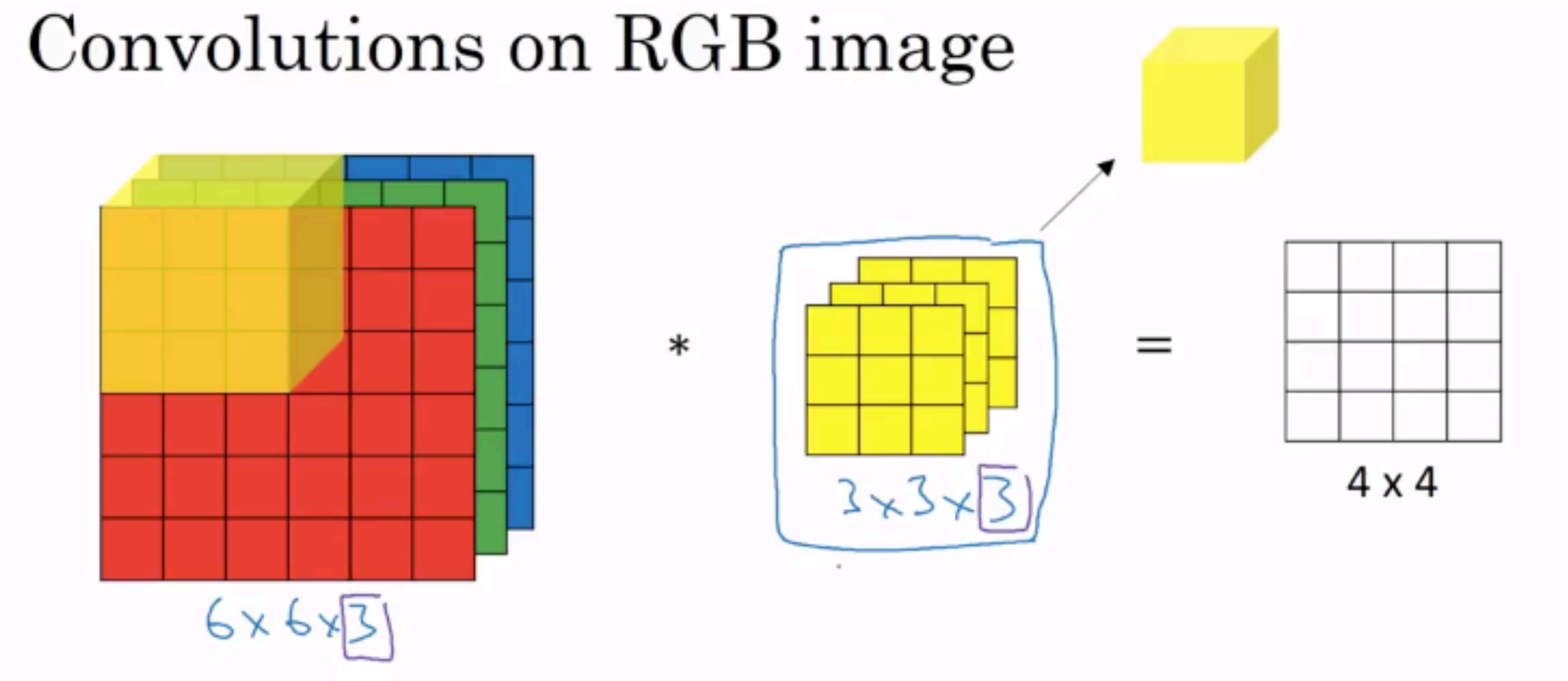
Naively, we can flip coupling layers:



Dinh, Laurent, Jascha Sohl-Dickstein, and Samy Bengio. "Density estimation using real nvp." *arXiv preprint arXiv:1605.08803* (2016).

Mixing layer - 1x1 Convolutions

A mixing operation is permuting information so lets generalize that:



Kingma, Durk P., and Prafulla Dhariwal. "Glow: Generative flow with invertible 1x1 convolutions." *Advances in neural information processing systems* 31 (2018).