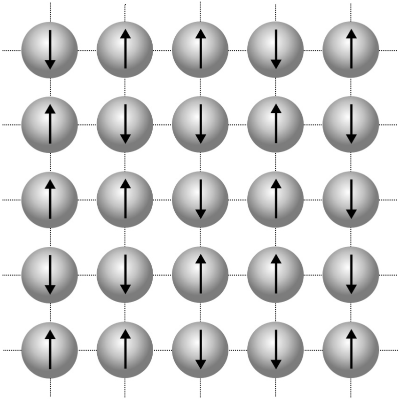
Machine Learning reading group 4/24/18

An exact mapping between the Variational Renormalization Group and Deep Learning

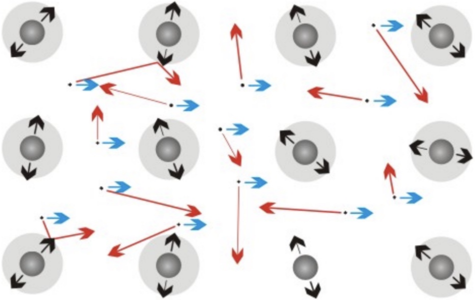
Mehta and Schwab, *arXiv* 2014

Ising model of interacting spins



$$H(\mathbf{v}) = \sum_i K_i v_i + \sum_{ij} K_{ij} v_i v_j \quad \text{for } v_i = \pm 1$$

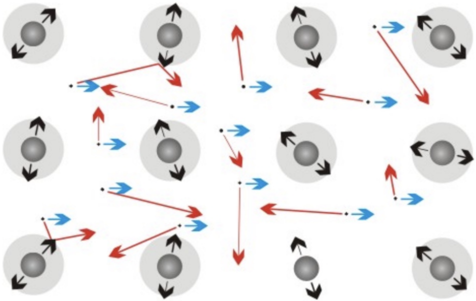
Spins under thermal motion



$$P(\mathbf{v}) = \frac{e^{-H(\mathbf{v})}}{Z}$$

where $Z = \sum_{\mathbf{v}'} e^{-H(\mathbf{v}')}$

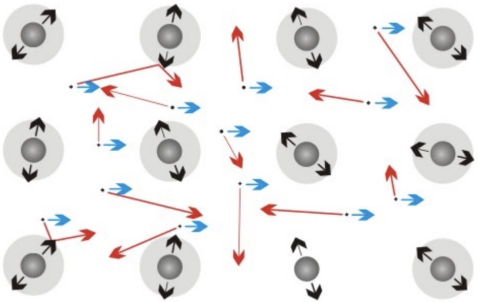
Thermal spin distribution



$$P(\uparrow / \downarrow |_i) = P(v_i = \pm 1)$$

$$= \sum_{v_{j \neq i} = \pm 1} \frac{e^{-H(\mathbf{v})}}{Z} \bigg|_{v_i = \pm 1}$$

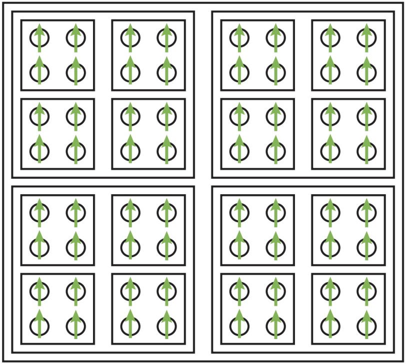
Definition of *free energy*



$$F = -\log Z$$

F measures disequilibrium (e.g. polarization).

Renormalized spins

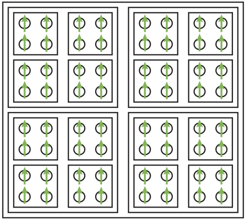


- parameters λ

- $\mathbf{v} \longrightarrow \mathbf{h}$
- $K \longrightarrow \tilde{K}$

- $H \longrightarrow H_\lambda^{RG}$
- $F^{\mathbf{v}} \longrightarrow F_\lambda^{\mathbf{h}}$

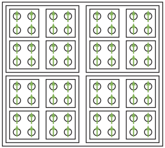
Calculating effective spins $\mathbf{h}$



$\mathbf{h}$ is an average spin per block:

$$h_i = \frac{1}{A} \sum_j v_j$$

Calculating effective coupling strengths $\tilde{K}$



Define T_λ as (negative) interaction energy between $\mathbf{v}$ and $\mathbf{h}$, and integrate out $\mathbf{v}$:

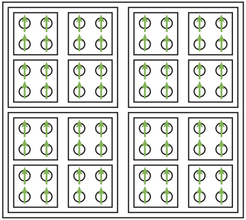
$$e^{-H^{RG}(\mathbf{h})} = \sum_{\mathbf{v}|\mathbf{h}} e^{T_\lambda(\mathbf{v},\mathbf{h})} \cdot e^{-H(\mathbf{v})}$$

This implies a renormalized free energy:

$$Z^{RG} = \sum_{\mathbf{h}} e^{-H^{RG}(\mathbf{h})}$$

$$F_\lambda = -\log Z^{RG}$$

Objective of renormalization process

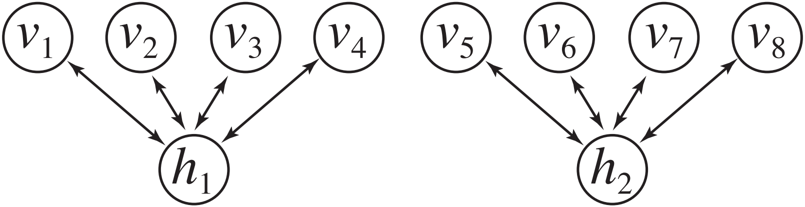


Adjust λ so that free energies match:

$$F_{\lambda}^{\mathbf{h}} = F^{\mathbf{v}}$$

Equivalently, we want $e^{T_\lambda} = 1$ for all states.

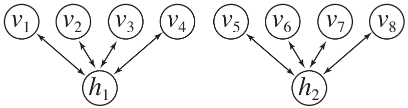
RBM activity depends on an energy function



$$p_\lambda(\mathbf{v}, \mathbf{h}) = \frac{e^{-E_\lambda(\mathbf{v},\mathbf{h})}}{Z_\lambda}$$

E is typically 2nd order in $\mathbf{v}$, $\mathbf{h}$

Training objective



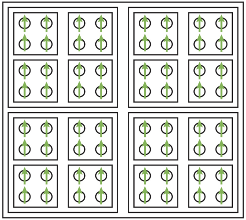
Adjust λ to reproduce input statistics:

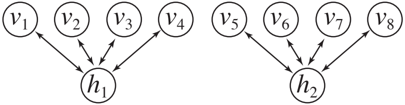
$$0 = \text{KL divergence}\,(p(\mathbf{v}), p_\lambda(\mathbf{v}))$$

$$= \sum_{\mathbf{v}} p(\mathbf{v}) \log \frac{p(\mathbf{v})}{p_\lambda(\mathbf{v})}$$

$$\longrightarrow p(\mathbf{v}) = p_\lambda(\mathbf{v})$$

Correspondence

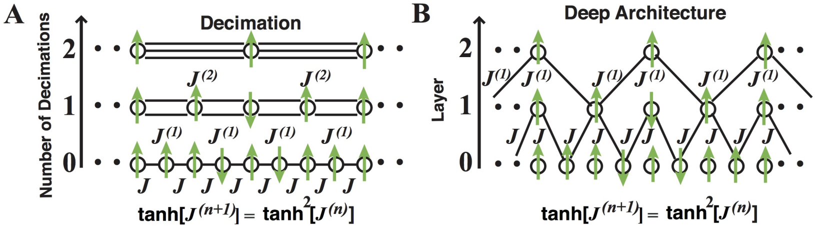




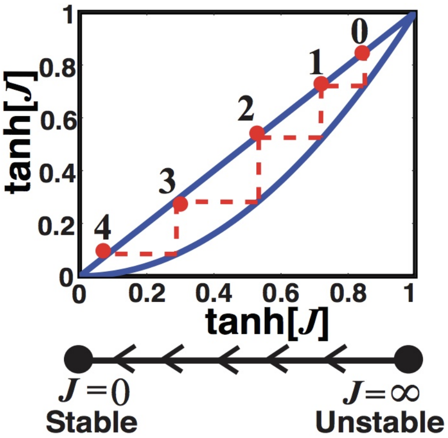
$$T - H \quad \longleftrightarrow \quad -E$$

$$F = F_\lambda \quad \longleftrightarrow \quad \frac{e^{-E(\mathbf{v},\mathbf{h})}}{Z} = \frac{e^{-E_\lambda(\mathbf{v},\mathbf{h})}}{Z_\lambda}$$

Renormalizing by ‘decimation’ in 1D



Renormalized 1D coupling strengths



Renormalization in 2D