Notes on RG flow through decimation of undetected degrees of freedom

Serena Bradde and William Bialek

Gaussian-like models have been used to describe system that are not embedded in a dimensional space like neural or gene interaction networks. In absence of a space structure, locality can not justify the fact that we neglect three-body or higher order interaction terms. Still the "effective" joint probability distribution with two-body interaction term seems to capture in several cases the basic property of the systems. There is indeed a ratio behind the use of effective models to describe more complex system. Here we want to give an intuition behind this assumption and give a criterion to understand why we can neglect higher order terms that may be present at the molecular level. Using mean field models, the paradigm of non local systems, we build up a theory that shows that in absence of full knowledge when part of the system is unobserved, learning higher order terms is not only unfeasible but also useless given that a two-body interaction hamiltonian is efficient in describing the system distribution.

I. PARTIAL INFORMATION

Let us start assuming that we have a system of N spins whose energy in mean-field regime is just a function of the average magnetization, $e(\{\sigma\}) = e(m)$ with $m = 1/N \sum_i \sigma_i$. The system is described by the Boltzman distribution with temperature $1/\beta$. Let us assume that we do not observe to the full system but only a fraction $(1 - x)$ We thus divide the degrees of freedom in two sets: $\mathcal{O}$ the set of observed degrees of freedom σ_i with $i = 1, \ldots, N(1-x)$, and its complement, $\overline{\mathcal{O}}$, that is hidden or unobserved, τ_j with $j = 1, \ldots, xN$. However, they interact with each other and with the rest of the system and in order to get the distribution of the observed system we need to integrate them out

$$P(\{\sigma\}) = \frac{1}{Z} \sum_{\tau_j} e^{-\beta N e((1-x)m_\sigma + x m_\tau)} \tag{1}$$

where $\sigma_i \in \mathcal{O}$ while $\tau_j \in \overline{\mathcal{O}}$ and the magnetization are respectively, $m_\sigma = 1/N(1-x) \sum_{i \in \mathcal{O}} \sigma_i$ while $m_\tau = 1/Nx \sum_{\alpha \in \overline{\mathcal{O}}} \tau_\alpha$. What happens to the observed system when we integrate over the unobserved degrees of freedom? In other terms what is the role in hidden nodes in the effective interactions of observed variables? In the following we will answer this question drawing parallel from the Renormalization group analysis.

II. SMALL X EXPANSION

When the fraction of unobserved nodes is small, we can perform analytic calculation of the mean field "effective" couplings for a general form of the free energy. In this case by assuming additivity of the entropic part we get at leading order in x that

$$P(\{\sigma\}) = \frac{1}{Z} \int dm \, e^{-N(\beta e((1-x)m_\sigma + xm) - xs(m))} = \frac{1}{Z} e^{-N\beta e(m_\sigma)} \int dm \, e^{Nxs(m) - Nx(m-m_\sigma)e'(m_\sigma) + O(x^2)} \tag{2}$$

where $e'(x) = \frac{de(m)}{dm}\big|_x$. The integral over m can be performed in the limit $N \to \infty$ by saddle point approximation such that

$$\int dm \, e^{Nx(s(m) - e(m_\sigma)(m - m_\sigma))} \sim e^{Nx(s(m^*) - (m^* - m_\sigma)e'(m_\sigma))} \quad \text{where} \quad s'(m^*) = e'(m_\sigma)$$

This means that the new system is characterized by

$$P(\{\sigma\}) = \frac{1}{Z} e^{-N(1-x)\beta \mathcal{E}(m_\sigma)} \tag{3}$$

where the new energy density is $\mathcal{E}(m_\sigma) = e(m_\sigma) + x\left[e(m_\sigma) - m_\sigma e'(m_\sigma) - \log \cosh e'(m_\sigma)\right]$. This last equation is obtained by substituting into the definition of $s(m)$, the saddle point solution m^* and properly taking into account the renormalization of degrees of freedom $N \to N(1-x)$. This means that the change in the energy function due to the integration of internal degrees of freedom

$$\frac{d\mathcal{E}}{dx} = e(m) - me'(m) - \log \cosh e'(m) \tag{4}$$

A.   Ising model

In order to obtain the normalization of the Ising model, we need to expand this equation and compute it in terms of the couplings. For example, let us assume that

$$e(m) = -\frac{J}{2}m^2 - \frac{u}{4}m^4 \,.$$

Let us specify that the limits we are interested in, i.e. $x \to 1$ and $N(1 - x) \to \infty$. In order to have a well behaving distribution in the limit $N \to \infty$, we have to impose the field scales as $m \to m(1 - x)^\beta$ with $\beta = 1/2$ whenever $J \neq 1$. When the coupling $J \neq 1$, it turns out that J is an adimensional parameter while the forth order term get a dimension, it goes to zero as $u \sim O(1 - x)$. If we define $u' = u/(1 - x)$ then using equation (**??**) we get that expanding for small value of m the couplings are renormalized

$$\frac{dJ}{dx} = J - 1 \qquad \frac{du'}{dx} = 3(J - 1) + J\left(u' - \frac{J^3}{3}\right) - u' \tag{5}$$