

Glassy Dynamics in Physics & Beyond

(Landscape & Glassy Dynamics)

- Dynamical Mean-Field Theory of Glassy dynamics
Dynamical mean-field equation via the dynamic cavity method
- How & Why dynamics slow down in rough high-dimensional landscape?
Analysis of the Langevin dynamics for the p-spin spherical model
- Rough-landscape with a signal (“crystal hunting”)
Dynamics for the spiked-tensor model and generalization

Rough Landscape and Generalized Spin-Glasses

$$\sum_{\langle i,j \rangle} J_{ij} s_i s_j \xrightarrow[\text{More complex random interactions}]{\text{More complex random interactions}} \sum_{\langle i_1, \dots, i_p \rangle} J_{i_1, \dots, i_p} s_{i_1} s_{i_2} \dots s_{i_p}$$

$$\overline{J_{i_1, \dots, i_p}^2} = \frac{p!}{2N^{p-1}}$$

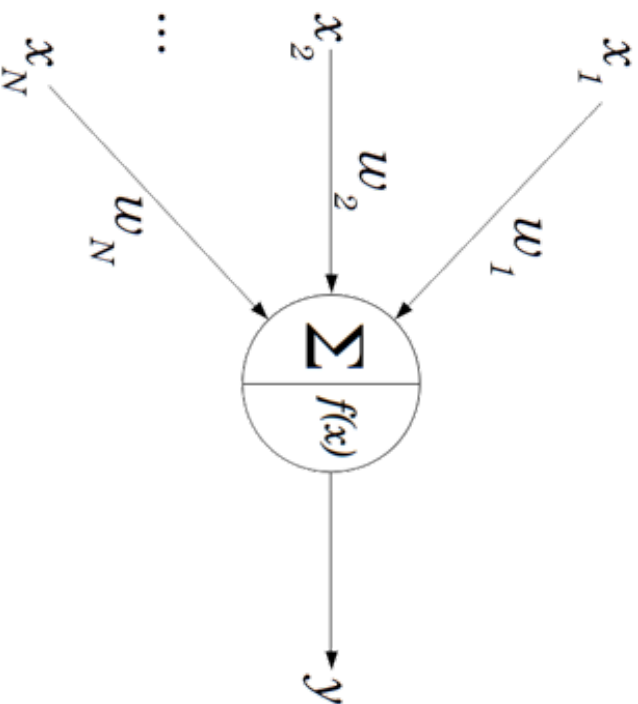
$$\frac{ds_i(t)}{dt} = -\frac{\partial V}{\partial s_i} + \frac{1}{(p-1)!} \sum_{i_2 \neq \dots i_p} J_{i, i_2, \dots, i_p} s_{i_2} \dots s_{i_p} + \xi_i(t)$$

DMFT:

$$\frac{ds(t)}{dt} = -\frac{\partial V}{\partial s} + \frac{p(p-1)}{2} \int_0^t R(t, t') C^{p-2}(t, t') s(t') dt' + \xi(t)$$

$$\langle \xi(t) \xi(t') \rangle = 2T \delta(t - t') + \frac{p}{2} C^{p-1}(t, t')$$

Perceptron



$$E(\{w_i\}) = \sum_{\mu=1}^M f(x_\mu \cdot w)$$

$$w^2 = N \quad M = \alpha N$$

- DMFT for Langevin dynamics
Agoritsas, Biroli, Urbani, Zamponi 2018
- SGD for online learning
Mace, Coolen '97
- DMFT for Glauber dynamics
Horner '92

Non-Conservative Dynamics (No Landscape)

Mean-field models of recurrent neural networks, large-interacting eco-systems and sheared glasses

$$\frac{dh_i}{dt} = -h_i + \sum_{l \neq i}^N J_{il} S_l = -h_i + \sum_{l \neq i}^N J_{il} \tanh(gh_l)$$

$$\overline{J_{ij}^2} = \frac{1}{N} \quad \overline{J_{ji} J_{ji}} = 0$$

$$\text{DMFT:} \quad \frac{dh}{dt} = -h + \xi(t) \quad \langle \xi(t) \xi(t') \rangle = \langle \tanh(gh(t)) \tanh(gh(t')) \rangle$$

Transition to **chaos** by increasing g
(Crisanti, Sompolinsky '87 ,...; Kadmon, Sompolinsky PRX 2015)

How & Why Dynamics slow down in rough high -D Landscapes?

P-spin spherical model

$$E = - \sum_{\langle i_1, \dots, i_p \rangle} J_{i_1, \dots, i_p} s_{i_1} s_{i_2} \dots s_{i_p} \quad \sum_{i=1}^N s_i^2 = N$$

$$\frac{ds_i(t)}{dt} = - \frac{\partial E}{\partial s_i} + \xi_i(t)$$

How & Why Dynamics slow down in rough high -D Landscapes?

P-spin spherical model

$$E = - \sum_{\langle i_1, \dots, i_p \rangle} J_{i_1, \dots, i_p} s_{i_1} s_{i_2} \dots s_{i_p}$$
$$\sum_{i=1}^N s_i^2 = N$$

$$\frac{ds_i(t)}{dt} = - \frac{\partial E}{\partial s_i} + \xi_i(t) - \lambda(t) s_i$$
$$\lambda(t) = -p \frac{E(t)}{N} + T$$

How & Why Dynamics slow down in rough high -D Landscapes?

P-spin spherical model

$$E = - \sum_{\langle i_1, \dots, i_p \rangle} J_{i_1, \dots, i_p} s_{i_1} s_{i_2} \dots s_{i_p} \quad \sum_{i=1}^N s_i^2 = N$$

$$\frac{ds_i(t)}{dt} = - \frac{\partial E}{\partial s_i} + \xi_i(t) - \lambda(t) s_i$$

$$\lambda(t) = -p \frac{E(t)}{N} + T$$

$$\frac{ds_i(t)}{dt} = -\lambda(t) s_i + \frac{1}{(p-1)!} \sum_{i_2 \neq \dots i_p} J_{i, i_2, \dots, i_p} s_{i_2} \dots s_{i_p} + \xi_i(t)$$

How & Why Dynamics slow down in rough high -D Landscapes?

P-spin spherical model

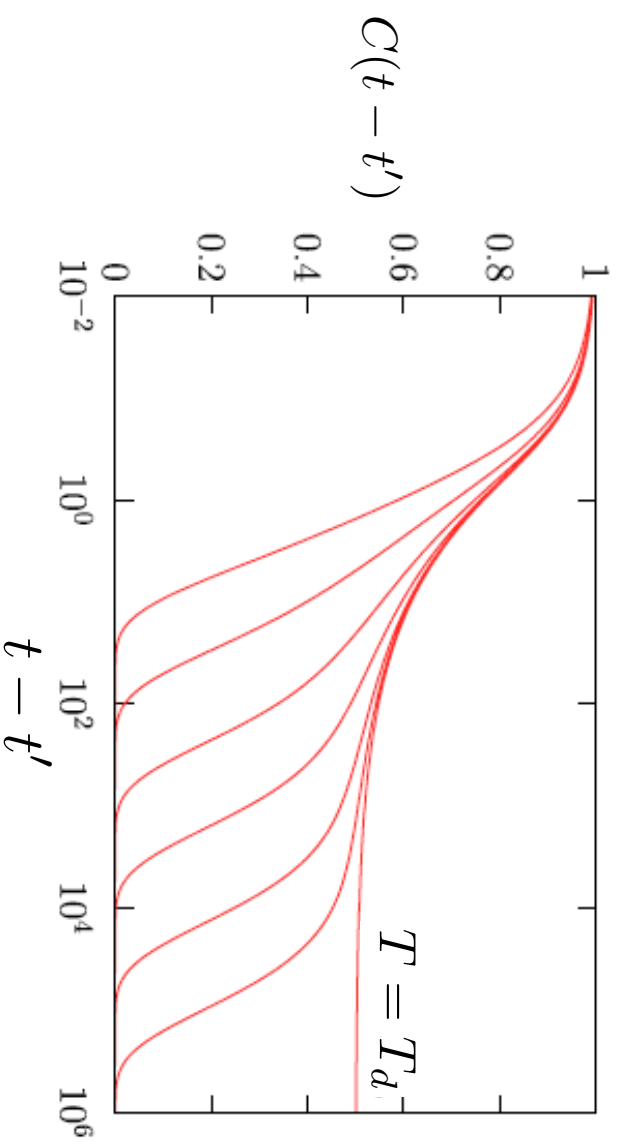
$$\frac{ds_i(t)}{dt} = -\lambda(t)s_i + \frac{1}{(p-1)!} \sum_{i_2 \neq \dots i_p} J_{i,i_2,\dots,i_p} s_{i_2} \dots s_{i_p} + \xi_i(t)$$

DMFT:

$$\begin{aligned} \frac{ds(t)}{dt} &= -\lambda(t)s(t) + \frac{p(p-1)}{2} \int_0^t R(t,t') C^{p-2}(t,t') s(t') dt' + \xi(t) \\ \langle \xi(t) \xi(t') \rangle &= 2T \delta(t-t') + \frac{p}{2} C^{p-1}(t,t') \end{aligned}$$

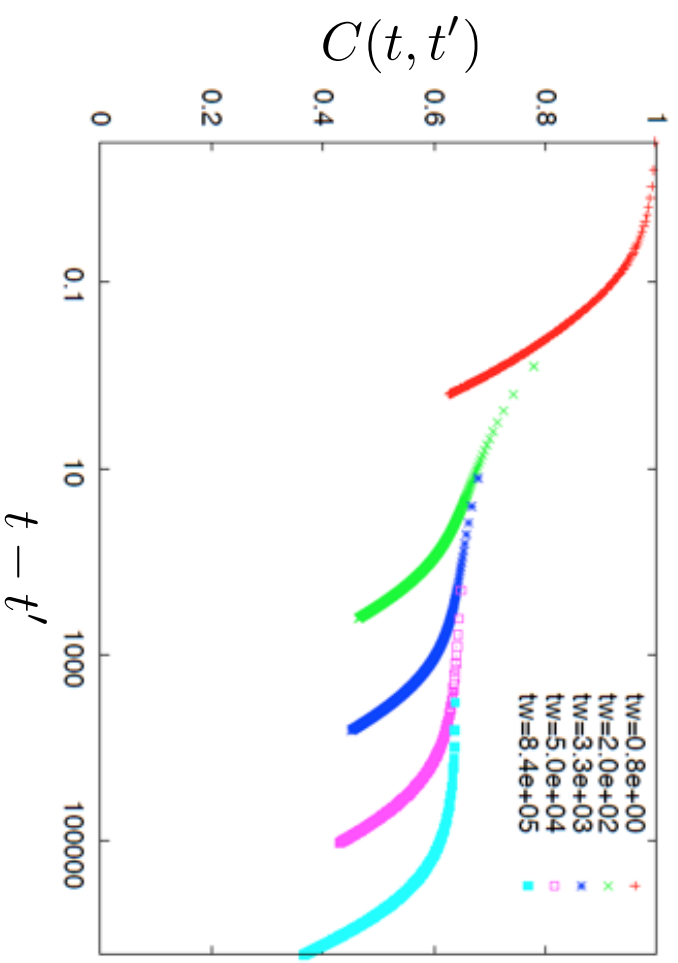
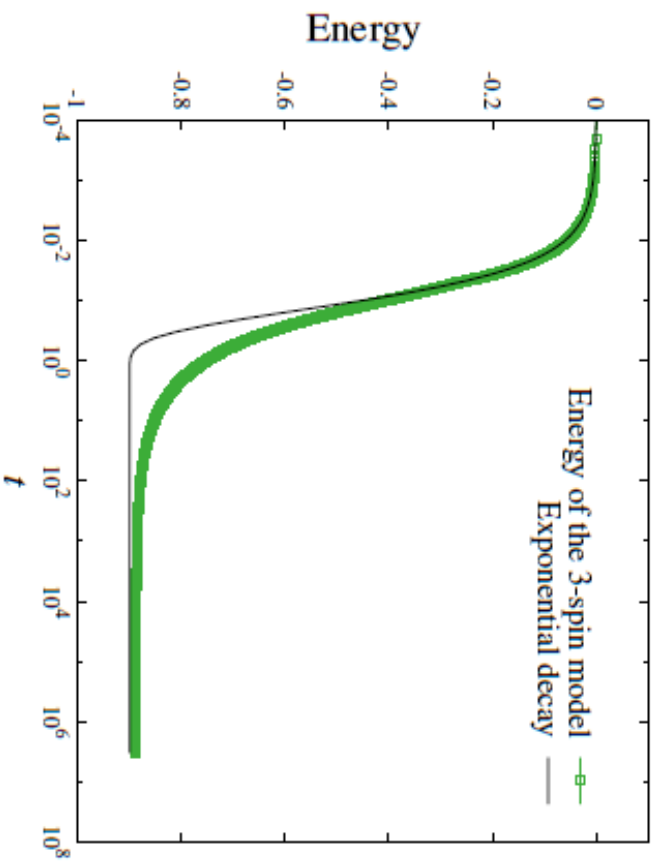
How & Why Dynamics slow down in rough high -D Landscapes?

P-spin spherical model: EQUILIBRIUM



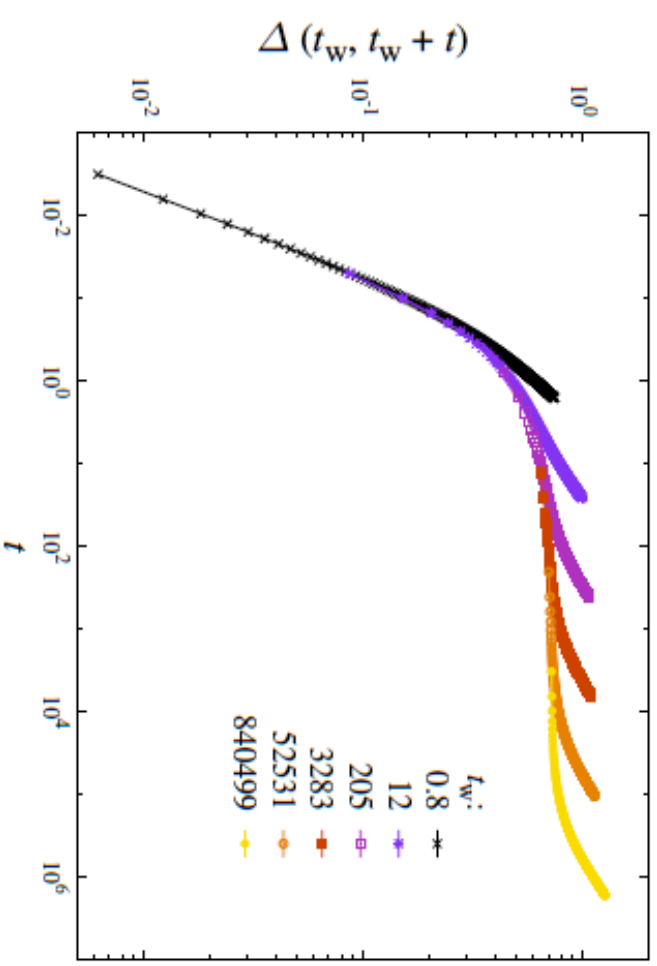
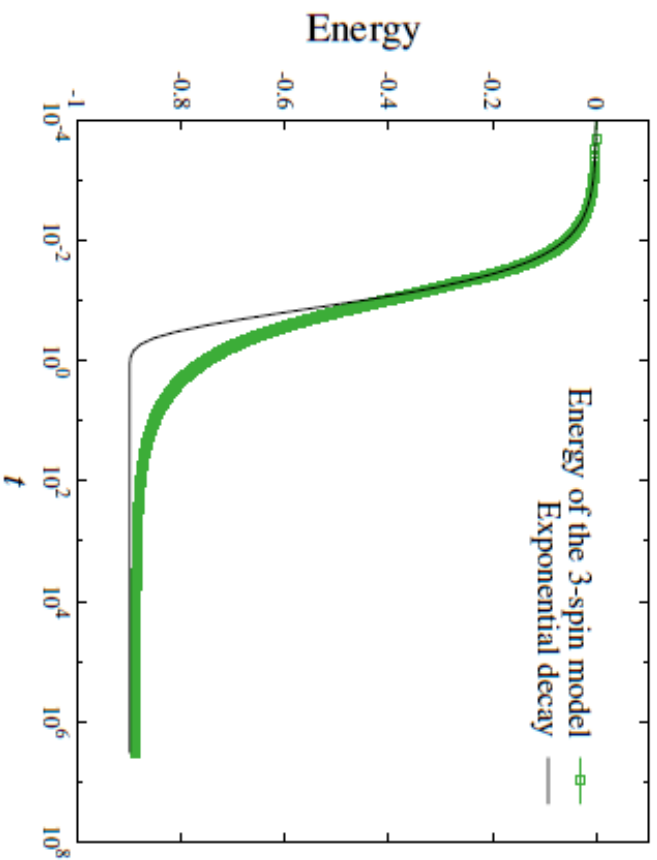
How & Why Dynamics slow down in rough high -D Landscapes?

P-spin spherical model: AGING



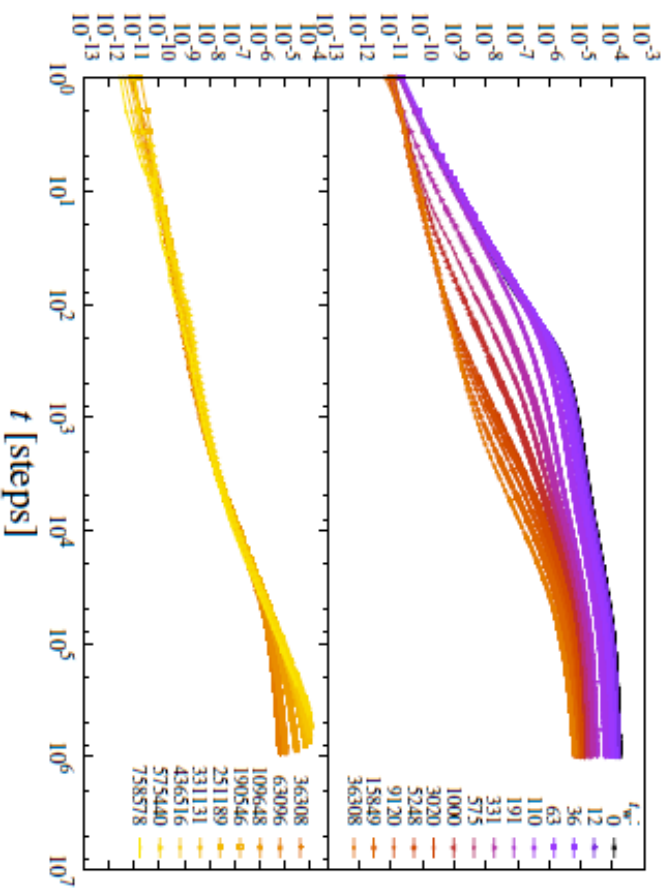
How & Why Dynamics slow down in rough high -D Landscapes?

P-spin spherical model: AGING

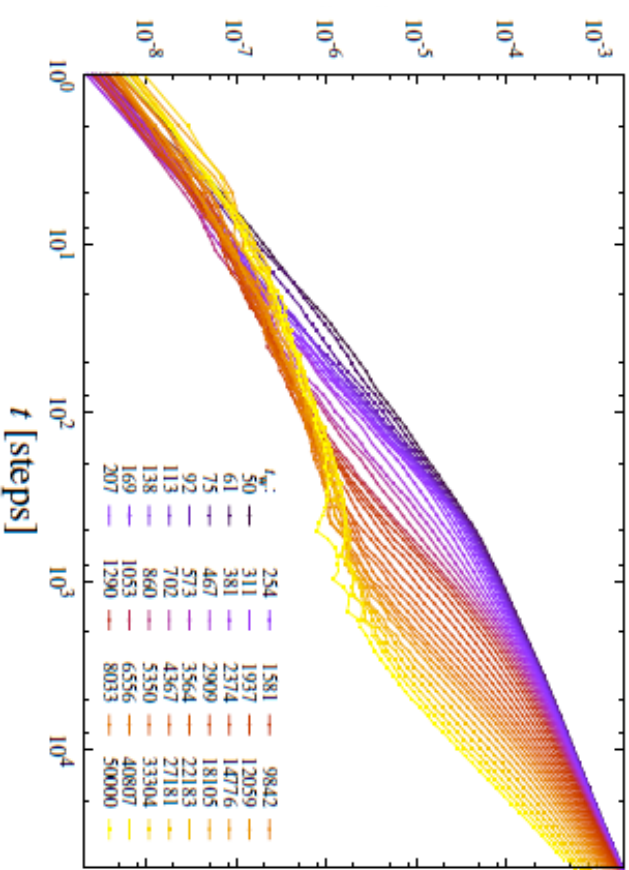


$$\Delta(t, t') = \frac{1}{N} \sum_i \langle (s_i(t) - s_i(t'))^2 \rangle = 2 - 2C(t, t')$$

Correlation-Functions in DNNs



(b) Fully Connected on MNIST, $B = 128$, $\alpha = 0.01$.



(b) Mean square displacement of the under-parametrized model.

Over-parametrized

Under-parametrized

M. Baity-Jesi, L. Sagun, M. Geiger, S. Spigler, G. Ben Arous, C. Cammarota, Y. LeCun, M. Wyart, G. Biroli (2018)
arXiv1803.06969 & ICML 2018

Explanation in terms of changes in the landscape: L. Sagun
More on the transition: M. Wyart