#### Stefano Sarao Mannelli

in collaboration with

Florent Krzakala, Pierfrancesco Urbani and Lenka Zdeborová

ICML 2019

# Descent Algorithms and Local Minima in Spiked Matrix-Tensor Models





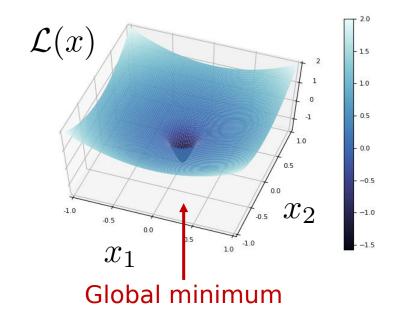
$$Y_{ij} = \frac{1}{\sqrt{N}} x_i^* x_j^* + \xi_{ij} \underset{\sim}{N}(0, \Delta_2)$$

$$T_{ijk} = \sqrt{\frac{(p-1)!}{N^{p-1}}} x_i^* x_j^* x_k^* + \xi_{ijk}$$

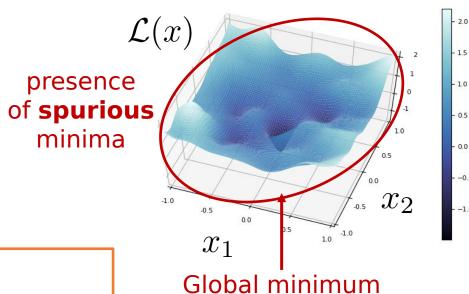
$$Y_{ij} = \frac{1}{\sqrt{N}} x_i^* x_j^* + \xi_{ij}$$
  $x^* = ?$   $\hat{x} = \arg\min_{x} \mathcal{L}(x)$  minus log-likelihood  $T_{ijk} = \sqrt{\frac{(p-1)!}{Np-1}} x_i^* x_j^* x_k^* + \xi_{ijk}$ 

$$Y_{ij} = \frac{1}{\sqrt{N}} x_i^* x_j^* + \xi_{ij}$$

$$T_{ijk} = \sqrt{\frac{(p-1)!}{N^{p-1}}} x_i^* x_j^* x_k^* + \xi_{ijk}$$



$$Y_{ij} = \frac{1}{\sqrt{N}} x_i^* x_j^* + \xi_{ij}$$



$$T_{ijk} = \sqrt{\frac{(p-1)!}{N^{p-1}}} x_i^* x_j^* x_k^* + \xi_{ijk}$$

#### **Gradient Flow Analysis**

#### Analytical analysis of gradient flow dynamics.

$$\frac{\partial}{\partial t}C(t,t') = 2R(t',t) - \mu(t)C(t,t') + Q'(m(t))m(t') +$$

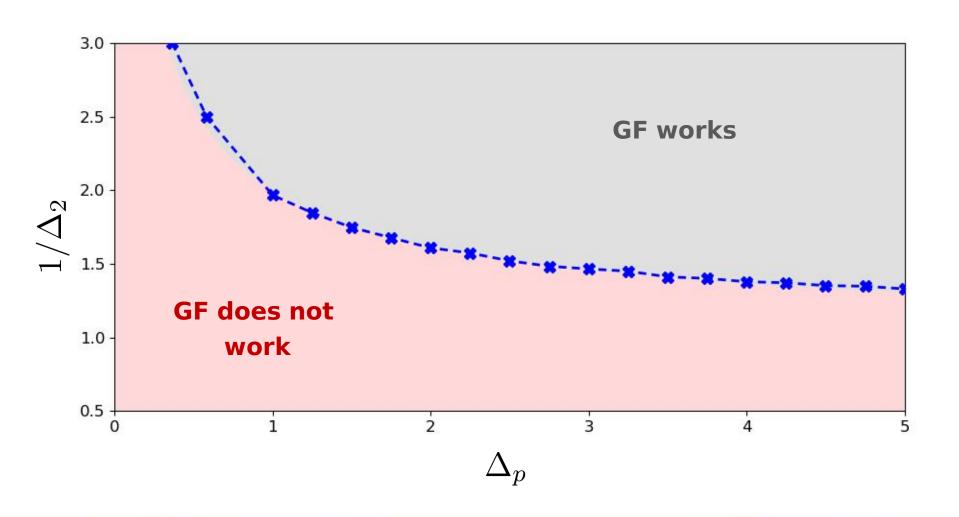
$$+ \int_0^t dt'' R(t,t'')Q''(C(t,t''))C(t',t'') + \int_0^{t'} dt'' R(t',t'')Q'(C(t,t'')),$$

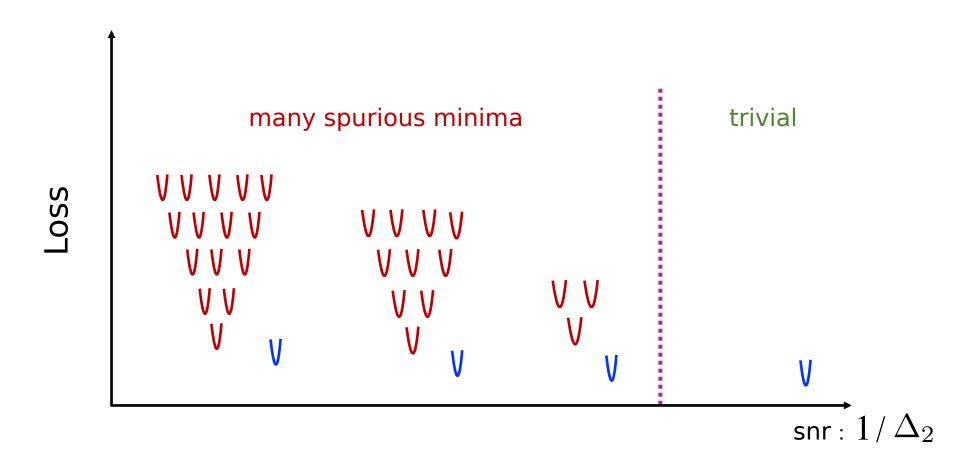
$$\frac{\partial}{\partial t}R(t,t') = \delta(t-t') - \mu(t)R(t,t') + \int_{t'}^t dt'' R(t,t'')Q''(C(t,t''))R(t'',t'),$$

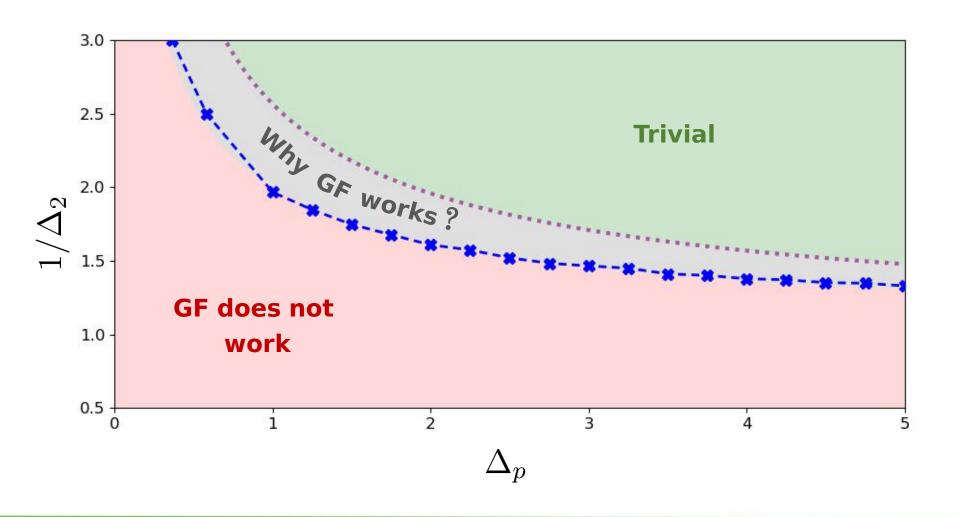
$$\frac{\partial}{\partial t}m(t) = -\mu(t)m(t) + Q'(m(t)) + \int_0^t dt'' R(t,t'')m(t'')Q(C(t,t'')),$$

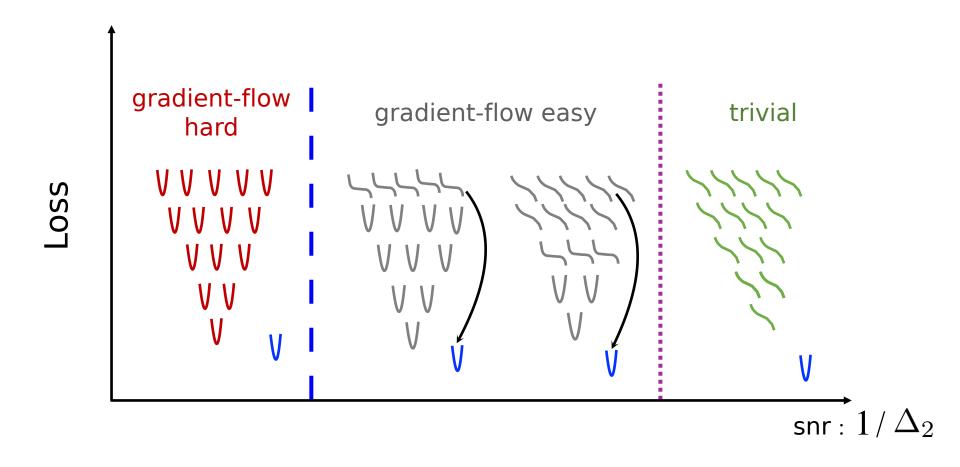
with 
$$Q(x) = x^2/(2\Delta_2) + x^p/(p\Delta_p)$$
.

$$C(t,t') = \frac{1}{N} \sum_{i=1}^{N} \left\langle x_i(t) x_i(t') \right\rangle, \ R(t,t') = \frac{1}{N} \sum_{i=1}^{N} \left\langle \frac{\delta x_i(t)}{\delta \eta_i(t')} \right\rangle, \ m(t) = \frac{1}{N} \sum_{i=1}^{N} \left\langle x_i(t) x_i^* \right\rangle.$$









\*joint work with Giulio Biroli and Chiara Cammarota



