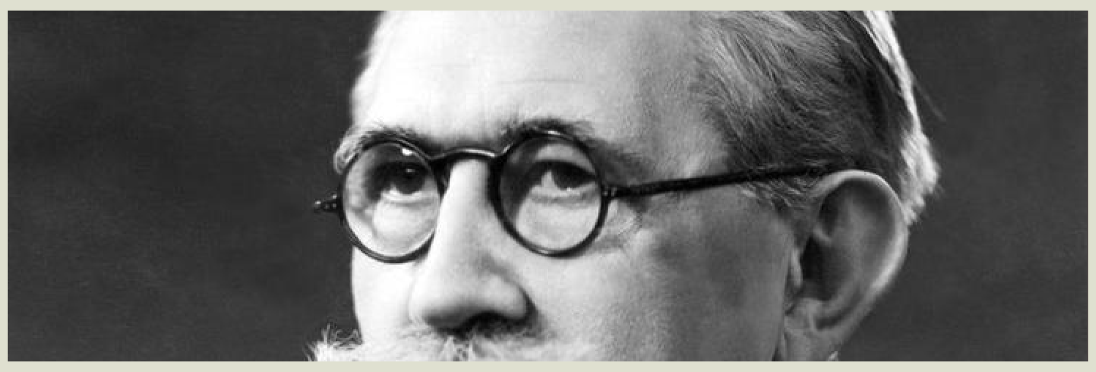
IPhT SACLAY

Stefano Sarao■, Giulio Biroli■●, ChiaraCammarota◆, Florent Krzakala●, Pierfrancesco Urbani■, Lenka Zdeborová■

■ Institut de Physique Théorique (IPhT), CEA, Université Paris-Saclay
◆ Laboratoire de Physique Statistique, Ecole Normale Supérieure, Paris
● King's College London, Department of Mathematics

cea

Performance of Langevin dynamics in high dimensional inference



INTRODUCTION

Does Langevin dynamics improve the boundaries of detection? We considered a simple detection problem and we studied the performance of different algorithms based on a generalization of cavity method called Approximate Message Passing (AMP)[1] and on Langevin dynamics[2,3]. We consider Langevin dynamics because on the inference prospective it mimics the commonly used Stochastic Gradient Descent algorithms, on the physics side it represents the natural relaxation of the systems.

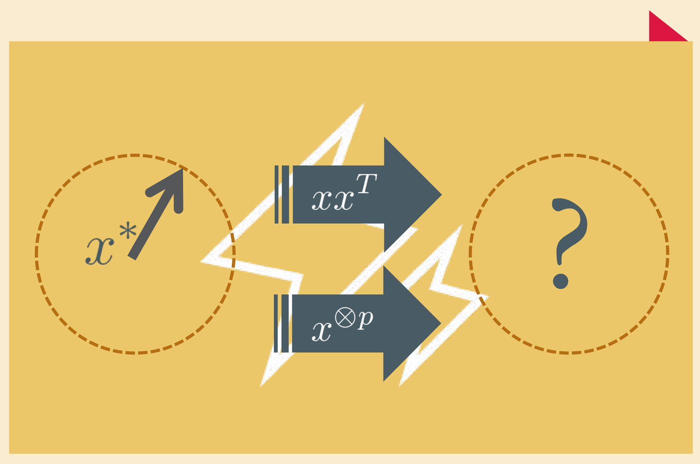
Spoiler Alert: it doesn't at all! Naive quenching and annealing protocols get much worse inference boundaries with respect to AMP. Using the standard replica machinery we investigated the complexity of the region where Langevin gets stuck and we found a good estimation of the threshold line[4].

MODEL

tensor+matrix factorization OR planted 2+p spherical spin model

Inference problem. Given a canonical Gaussian signal, we construct a matrix, Y, and a tensor, T, that we observe through two Gaussian channels, with variances Δ_2 and Δ_p. *What are the variances of the channels that allow to reconstruct the signal?* Using Bayesian statistics we write the posterior as:

$$P(X|Y,T) \propto \prod_i e^{-\frac{1}{2}x_i^2} \prod_{i<j} e^{-\frac{1}{2\Delta_2}\left(Y_{ij}-\frac{x_i x_j}{\sqrt{N}}\right)^2} \prod_{i_1<\cdots<i_p} e^{-\frac{1}{2\Delta_p}\left(T_{i_1\dots i_p}-\frac{\sqrt{(p-1)!}}{N^{(p-1)/2}}x_{i_1}\dots x_{i_p}\right)^2}$$



Physical problem. After half-line computation you can notice that the problem is the old (and well known) 2+p spherical spin problem with a preferential configuration that acts like a ferromagnetic bias. The interactions are Gaussian with variances Δ_2 and Δ_p.

$$\beta\mathcal{H} = -\frac{1}{\Delta_2\sqrt{N}}\sum_{i<j}\xi_{ij}x_i x_j - \frac{\sqrt{(p-1)!}}{\Delta_p N^{\frac{p-1}{2}}}\sum_{i_1<\cdots<i_p}\xi_{i_1\dots i_p}x_{i_1}\dots x_{i_p} - \frac{N}{2\Delta_2}\left(\frac{1}{N}\sum_i x_i x_i^*\right)^2 - \frac{N}{p\Delta_p}\left(\frac{1}{N}\sum_i x_i x_i^*\right)^p$$

AMP & REPLICA

AMP state evolution and Parisi-scheme have been used to explore the static phase diagram of the problem. AMP is morally equivalent to look for a Replica Symmetric solution in an algorithmic way.
According to replica scheme, we write the partition function as

$$\overline{Z_x^n} \propto \int \prod_{ab} dQ_{ab} e^{NnxS(Q)} \simeq \lim\sup_Q e^{NnxS(Q)}$$

where Q is a generalized Parisi-matrix in the 1RSB scheme

$$Q = \begin{pmatrix} 1 & m\cdots m \\ m & \\ \vdots & \tilde{Q} \\ m & \end{pmatrix}$$

and S is

$$S(Q) = \frac{1}{x}\left[\frac{1}{2}\log\det Q + \frac{1}{2p\Delta_p}\sum_{ab}Q_{ab}^p + \frac{1}{4\Delta_2}\sum_{ab}Q_{ab}^2\right]$$

with x the Parisi-parameter.
The extremum is taken using Saddle Point approximation in the thermodynamic limit. The fixed point of the Saddle Point equations contains relevant information concerning the final magnetization and the presence of metastable states.

DYNAMICS

We consider a Langevinian evolution of the particles in our system. The ith particle will evolve as

$$\dot{x}_i(t) = -\mu(t)x_i(t) - \frac{\partial\mathcal{H}}{\partial x_i} - \eta_i(t)$$

where μ is a Lagrange multiplier enforcing spherical constrain and η represent thermal noise. Δ_2 was chosen as temperature in order to sample the posterior probability, thus

$$\langle\eta_i(t)\eta_j(t')\rangle = 2\Delta_2\delta_{ij}\delta(t-t')$$

using standard techniques we can derive the average dynamical evolution of some observable of our system, forming a closed system of differential equations

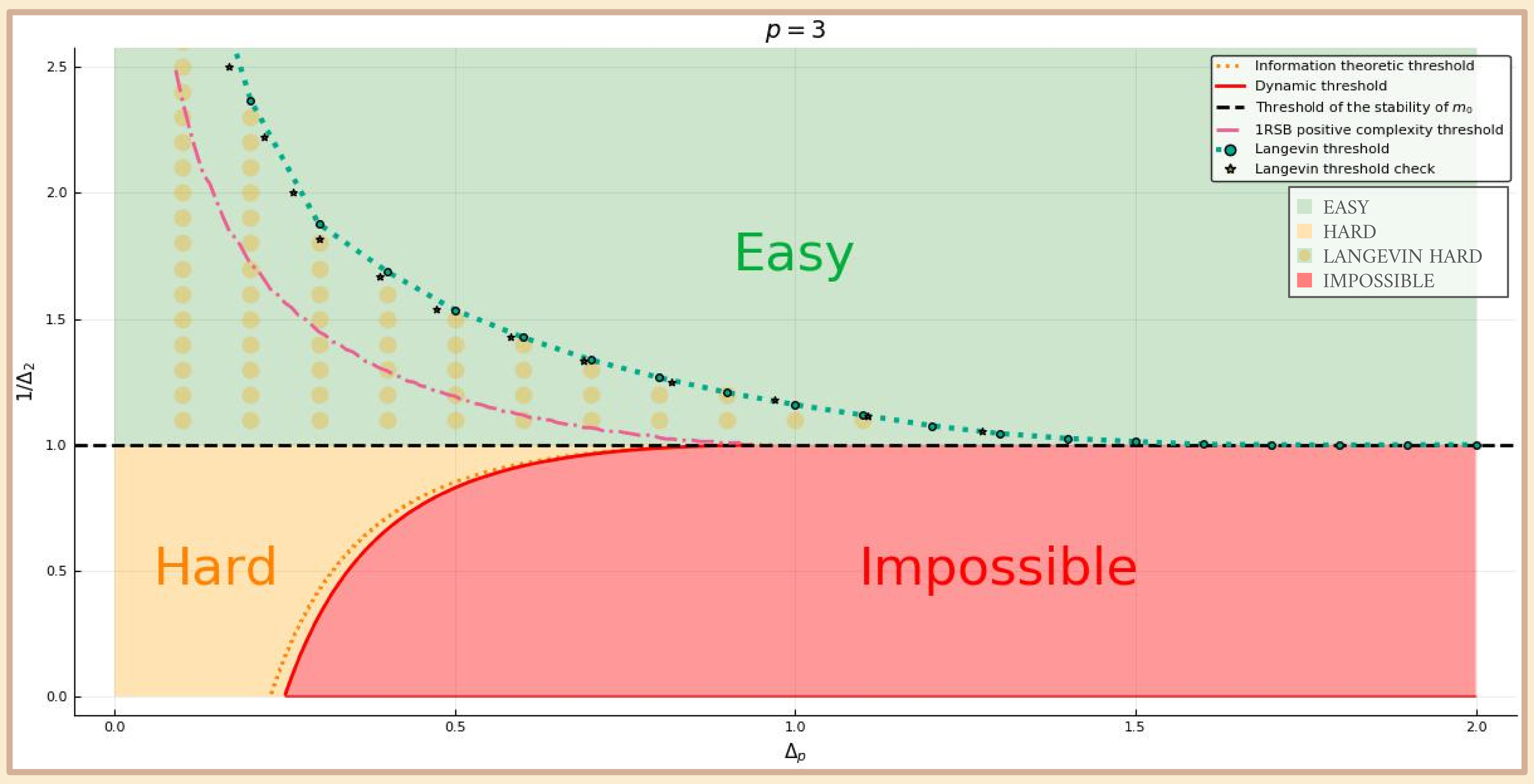
$$(\partial_t+\mu(t))C(t,t') = r_2\overline{C}(t')f_2'(\overline{C}(t)) + r_p\overline{C}(t')f_p'(\overline{C}(t)) +$$
$$+\Delta_2\int_0^{t'} dt'' f_2''(t,t'')R(t',t'') + \Delta_2\int_0^{t} dt'' f_2''(t,t'')R(t,t'')C(t',t'')+$$
$$+\frac{2}{p}\frac{\Delta_2^2}{\Delta_p}\int_0^{t'} dt'' f_p'(t,t'')R(t',t'') + \frac{2}{p}\frac{\Delta_2^2}{\Delta_p}\int_0^{t} dt'' f_p''(t,t'')R(t,t'')C(t',t'')$$
$$(\partial_t+\mu(t))R(t,t') = \Delta_2\int_{t'}^{t} dt'' f_2''(t,t'')R(t,t'')R(t'',t')+$$
$$+\frac{2}{p}\frac{\Delta_2^2}{\Delta_p}\int_{t'}^{t} dt'' f_p''(t,t'')R(t,t'')R(t'',t')$$
$$\mu(t) = \Delta_2 + r_2\overline{C}(t)f_2'(\overline{C}(t)) + r_p\overline{C}(t)f_p'(\overline{C}(t))+$$
$$+2\Delta_2\int_0^t dt'' f_2'(t,t'')R(t,t'') + 2\frac{\Delta_2^2}{\Delta_p}\int_0^t dt'' f_p'(t,t'')R(t,t'')$$
$$(\partial_t+\mu(t))\overline{C}(t) = r_2 f_2'(\overline{C}(t)) + r_p f_p'(\overline{C}(t))+$$
$$+\Delta_2\int_0^t dt'' f_2''(t,t'')R(t,t'')\overline{C}(t'') + \frac{2}{p}\frac{\Delta_2^2}{\Delta_p}\int_0^t dt'' f_p''(t,t'')R(t,t'')\overline{C}(t'')$$

with t>t', C(t,t') and R(t,t') the correlation and the response function, C(t) i the magnetization and $f_k(x) = x^k/2$.

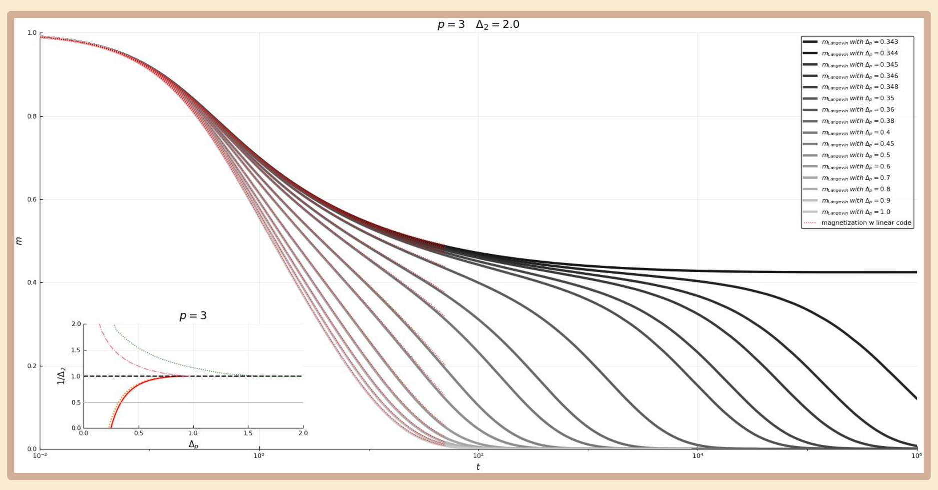
RESULTS

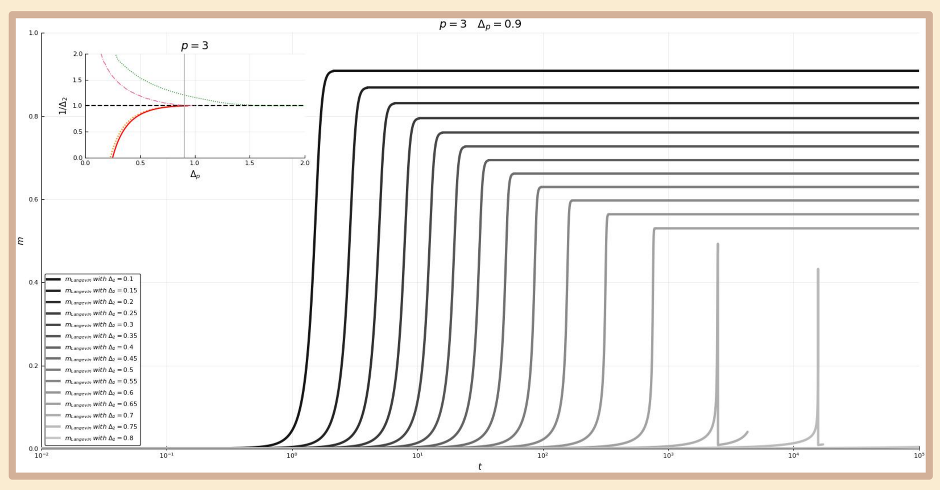
The static phase diagram shows three region easy, impossible and hard that representing the possibility of finding the ferromagnet in the problem. The hard phase is characterized by competing minima separated by exponentially high barriers.

Integrating numerically Langevin equations[5], we compared this picture with the one portraited by the static. Surprisingly Langevin's hard region is much larger then AMP's. The boundary were obtained by fitting the relaxation time with a power law and extrapolating the critical Δs.



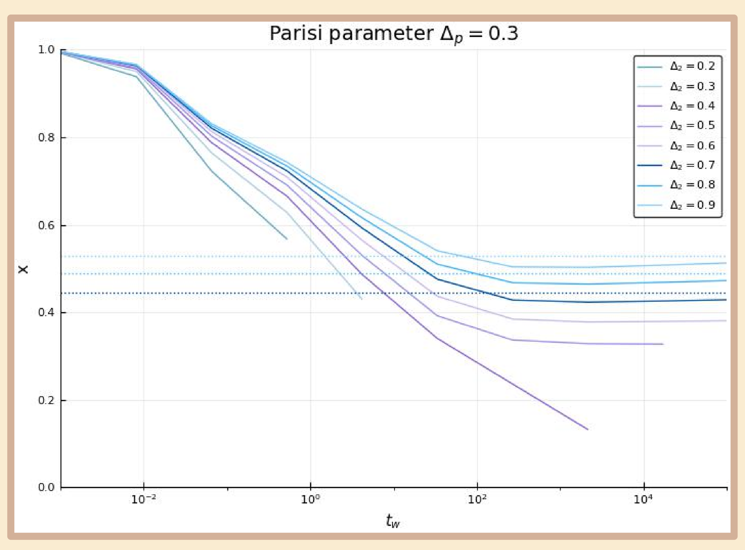
Extrapolation of the Langevin Hard line





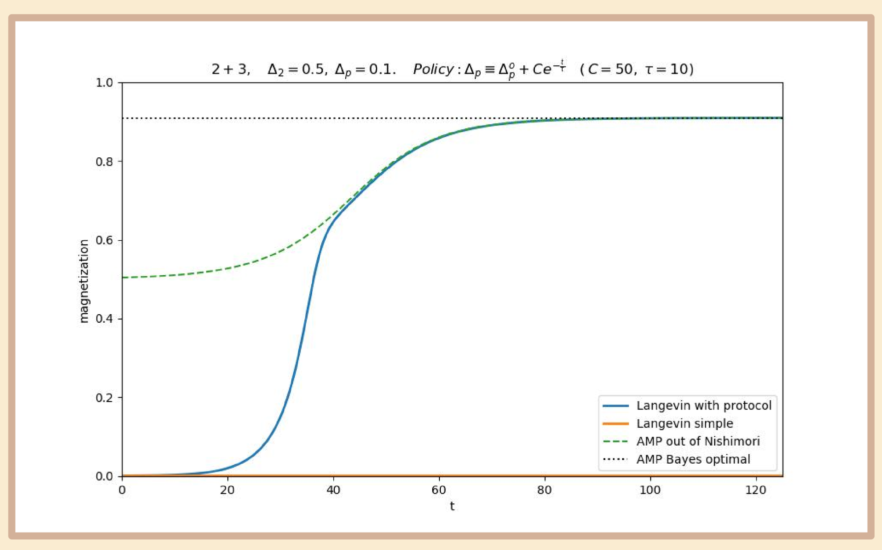
Langevin Hard region interpretation

The presence of minima that, although don't dominate the action, block the dynamics has been highlighted in 1RSB study. We compare the Parisi parameter in the two appraoches.



Cracking the Langevin Hard region

Using specific annealing protocols we showed that Langevin dynamics can enter in the hard region. However, this protocols are problem-specific, limiting their practical use.



References:

1. Lesieur T., Krzakala F, Zdeborová L. *Constrained low-rank matrix estimation: Phase transitions, approximate message passing and applications.* Journal of Stat. Mech. 2017;
2. Franz S, Mézard M. *Off-equilibrium glassy dynamics: a simple case.* EPL 1994;
3. Cugliandolo LF, Kurchan J. *Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model.* PRL. 1993;
4. Antenucci F, Franz S, Urbani P, Zdeborová L. *On the glassy nature of the hard phase in inference problems.* arXiv:1805.05857;
5. Kim B, Latz A. *The dynamics of the spherical p-spin model: From microscopic to asymptotic.* EPL 2001;

cnrs · PIFDE · erc European Research Council · PALM · SIMONS FOUNDATION