Unified Survey-Belief Propagation Scheme for Satisfiability

A. PELIZZOLA[(1)], M. PRETTI[(1)], and R. ZECCHINA[(2)]

(1) Politecnico di Torino, (2) Abdus Salam I.C.T.P.

Outline

→ The random satisfiability problem

→ The Bethe approximation

→ Belief Propagation

→ From Belief Propagation to Warning Propagation

→ From Warning Propagation to Survey Propagation

→ From a double-loop algorithm to “damped” Belief Propagation

→ Damped Survey Propagation

→ Numerical experiments

→ Conclusions

(Random) K-satisfiability

→ The K-sat problem deals with N boolean variables x_i, $i \in \{1,...,N\}$ and M "clauses" $a \in \{1,...,M\}$, which must be verified simultaneously.

→ Each clause is built as the OR function of K (randomly chosen) variables, which can be (randomly) negated or not.

→ For $K \geq 3$ the problem is NP-complete.

→ In physical terms, a boolean variable x_i can be mapped onto an Ising spin $\sigma_i = +1, -1$ if x_i =TRUE,FALSE.

→ For each variable x_i appearing in clause a, one introduces a "coupling" $J_{a \to i} = +1, -1$ if the variable appears as $\overline{x_i}, x_i$, so that $\sigma_i = -J_{a \to i}$ satisfies the clause.

→ Given a spin configuration, we can define an "energy" function, proportional to the number of violated clauses

$$E = \sum_a E_a(\sigma_a) \propto \sum_a \prod_{i \in a} \delta(\sigma_i - J_{a \to i})$$

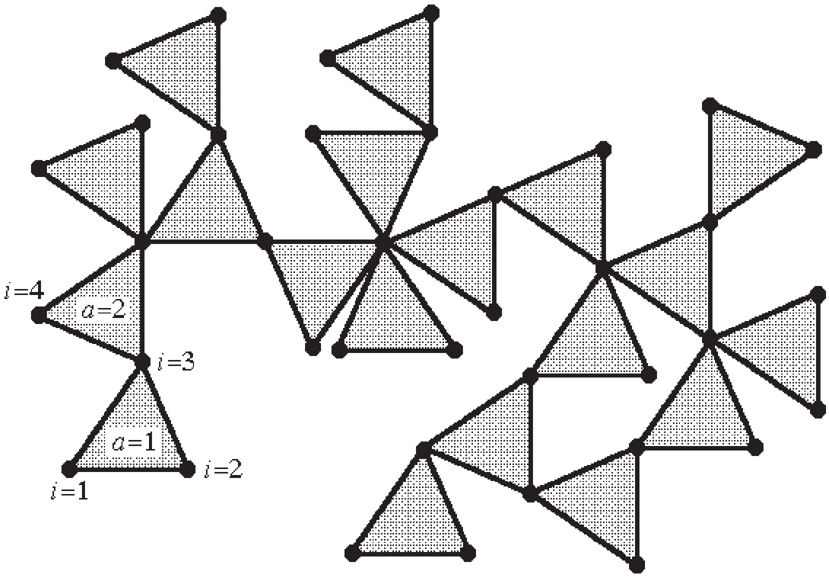
where $\sigma_a = \{\sigma_i | i \in a\}$. Satisfying assignments correspond to ground states of a random lattice model with K site interactions.

→ In the "thermodynamic" limit $N \to \infty$, for $K = 3$, there exists a sat-unsat phase transition around $\alpha \equiv M/N \approx 4.26$.

→ Formulae in the vicinity of the transition are extremely hard to solve, due to clustering phenomenon.

The Bethe approximation

The physical system we have defined has a hypergraph structure, which, for $K = 3$, looks like the following picture



For a random formula, the graph is locally treelike (loop size is $\mathcal{O}(\ln N)$ for $N \to \infty$), therefore a Bethe approximation is expected to be satisfactory. The free energy reads

$$\beta F = \sum_a \sum_{\sigma_a} p_a(\sigma_a) \left[\beta E_a(\sigma_a) + \ln p_a(\sigma_a)\right] - \sum_i B_i \sum_{\sigma_i} p_i(\sigma_i) \ln p_i(\sigma_i),$$

where $B_i = C_i - 1$, C_i being the connectivity of node i.

Minimization of this free energy with a message-passing scheme gives rise to the algorithm known as Belief Propagation.

Belief Propagation (BP)

Probability distributions for single nodes and clauses can be written as a function of messages

$$
\begin{aligned}
p_i(\sigma_i) &\propto \prod_{a \in i} m_{a\to i}(\sigma_i), \\
p_a(\sigma_a) &\propto e^{-\beta E_a(\sigma_a)} \prod_{i \in a} m_{i\to a}(\sigma_i),
\end{aligned}
$$

where messages must satisfy certain recursion relations

$$
m_{i\to a}(\sigma_i) = \prod_{b \in i \backslash a} m_{b\to i}(\sigma_i)
$$

$$
m_{a\to i}(\sigma_i) \propto \sum_{\sigma_{a\backslash i}} e^{-\beta E_a(\sigma_a)} \prod_{j \in a \backslash i} m_{j\to a}(\sigma_j)
$$

BP for satisfiability

For satisfiability we can write

$$
E_a(\sigma_a) = 2 \prod_{i \in a} \delta(\sigma_i - J_{a\to i}),
$$

whence

$$
\lim_{\beta\to\infty} e^{-\beta E_a(\sigma_a)} = 1 - \prod_{i \in a} \delta(\sigma_i - J_{a\to i}).
$$

Assuming that node-to-clause messages are normalized to unit

$$
\sum_{\sigma=\pm 1} m_{j\to a}(\sigma) = 1,
$$

the second recursion relation reads

$$
m_{a\to i}(\sigma) \propto 1 - \delta(\sigma - J_{a\to i}) \prod_{j \in a \backslash i} m_{j\to a}(J_{a\to j}).
$$

Warning Propagation (WP)

- Let us rewrite BP equations in terms of fields.

For Ising variables it is possible to write

$$m_{a\to i}(\sigma_i) = \exp\beta(u_{a\to i}\sigma_i + \text{const.}) \quad (1)$$

where $u_{a\to i}$ are usually denoted as cavity biases. The constant can be chosen to be zero, due to freedom in message normalization. From the first recursion relation one obtains

$$m_{i\to a}(\sigma_i) = \exp\beta h_{i\to a}\sigma_i \quad (2)$$

where

$$h_{i\to a} = \sum_{b\in i\backslash a} u_{b\to i}$$

are usually denoted as cavity fields. The latter equation represents the first recursion relation for Ising fields. The second relation can be obtained making use of expressions (1) and (2) and taking the linear combination $\frac{1}{2}\sum_{\sigma_i=\pm 1}\sigma_i(\cdot)$:

$$u_{a\to i} = \frac{1}{2}\sum_{\sigma_i}\sigma_i\beta^{-1}\ln\sum_{\sigma_{a\backslash i}}\exp\beta\left(\sum_{j\in a\backslash i} h_{j\to a}\sigma_j - E_a(\sigma_a)\right).$$

- Let us now take the limit $\beta\to\infty$

$$u_{a\to i} = \frac{1}{2}\sum_{\sigma_i}\sigma_i \max_{\sigma_{a\backslash i}}\left\{\sum_{j\in a\backslash i} h_{j\to a}\sigma_j - E_a(\sigma_a)\right\}$$

- Let us notice that, in case fields are $\mathcal{O}(\beta^{-1})$ (evanescent), there is some loss of information.

WP for satisfiability

Let us evaluate the second recursion relation, considering the sum over $\sigma_i = \pm J_{a\to i}$:

$$u_{a\to i} = \frac{1}{2} J_{a\to i} \left(\max_{\sigma_{a\setminus i}} \left\{ \sum_{j\in a\setminus i} h_{j\to a}\sigma_j - 2 \prod_{j\in a\setminus i} \delta(\sigma_j - J_{a\to j}) \right\} - \max_{\sigma_{a\setminus i}} \left\{ \sum_{j\in a\setminus i} h_{j\to a}\sigma_j \right\} \right) .$$

To maximize the second term, we can choose

$$\begin{aligned} \sigma_j &= \operatorname{sgn} h_{j\to a} && \text{if } h_{j\to a} \neq 0 \\ \sigma_j &= \text{don't care} && \text{if } h_{j\to a} = 0 \end{aligned}$$

The same choice is ok also for the first term, with the ansatz that cavity fields are integer (thanks to the prefactor 2), but

$$\sigma_j = -J_{a\to j} \qquad \text{if } h_{j\to a} = 0$$

As a consequence

$$u_{a\to i} = \frac{1}{2} J_{a\to i} \left(\sum_{j\in a\setminus i} |h_{j\to a}| - 2 \prod_{j\in a\setminus i} \delta(\operatorname{sgn} h_{j\to a} - J_{a\to j}) - \sum_{j\in a\setminus i} |h_{j\to a}| \right) ,$$

whence

$$\boxed{u_{a\to i} = -J_{a\to i} \prod_{j\in a\setminus i} \delta(\operatorname{sgn} h_{j\to a} - J_{a\to j})}$$

These equations are consistent with the ansatz of integer fields. In particular, $u_{a\to i} = 0, \pm 1$ while $h_{i\to a} = 0, \pm 1, \pm 2, \ldots$

Survey Propagation (SP) for satisfiability

In WP equations, only the sign of cavity fields is relevant, therefore let us define new variables ("cavity signs")

$$s_{i\to a} \doteq \operatorname{sgn} h_{i\to a} = 0, \pm 1.$$

The equations become

$$\begin{aligned} s_{i\to a} &= \operatorname{sgn} \sum_{b\in i\setminus a} u_{b\to i} \\ u_{a\to i} &= -J_{a\to i} \prod_{j\in a\setminus i} \delta(s_{j\to a} - J_{a\to j}). \end{aligned}$$

• Let us build up SP as a statistics over field values for sat configurations, according to the work of Braunstein, Mezard, and Zecchina [`cond-mat:0212002`], but defining probability "pseudo-distributions"

$$Q_{i\to a}(s) \doteq \mathbb{P}\{s_{i\to a} = s \vee 0\},$$

where $s = 0, \pm 1$, related to ordinary probability distributions

$$P_{i\to a}(s) \doteq \mathbb{P}\{s_{i\to a} = s\}$$

by the following simple relations

$$\begin{aligned} Q_{i\to a}(0) &= P_{i\to a}(0) \\ Q_{i\to a}(\sigma) &= P_{i\to a}(\sigma) + P_{i\to a}(0), \end{aligned}$$

where $\sigma = \pm 1$. In a completely analogous way, let us define distributions and pseudo-distributions for cavity biases

$$\begin{aligned} P_{a\to i}(s) &\doteq \mathbb{P}\{u_{a\to i} = s\} \\ Q_{a\to i}(s) &\doteq \mathbb{P}\{u_{a\to i} = s \vee 0\}. \end{aligned}$$

Let us consider the first recursion relation. Assuming that all $u_{b\to i}$ are statistically independent, and excluding unsat configurations (the ones in which $u_{b\to i}$ have different signs), it is possible to write

$$Q_{i\to a}(s) \quad \propto \quad \prod_{b\in i\backslash a} Q_{b\to i}(s)$$

Proportionality (not equality) is related to the fact that contradictory configurations are forced to be excluded.

As far as the second recursion relation is concerned, let us observe that we can only have $u_{a\to i} = 0, -J_{a\to i}$. As a consequence

$$\begin{aligned} Q_{a\to i}(-J_{a\to i}) &= 1 \\ Q_{a\to i}(\ \ J_{a\to i}) &= Q_{a\to i}(0) = P_{a\to i}(0), \end{aligned}$$

that is

$$Q_{a\to i}(s) \quad = \quad 1 - [\delta(s) + \delta(s - J_{a\to i})][1 - P_{a\to i}(0)],$$

where $P_{a\to i}(0)$ is the probability that at least one $s_{j\to a}$, for $j \in a\backslash i$, is different from $J_{a\to j}$. Still assuming statistical independence, we have

$$P_{a\to i}(0) \quad = \quad 1 - \prod_{j\in a\backslash i} P_{j\to a}(J_{a\to j}).$$

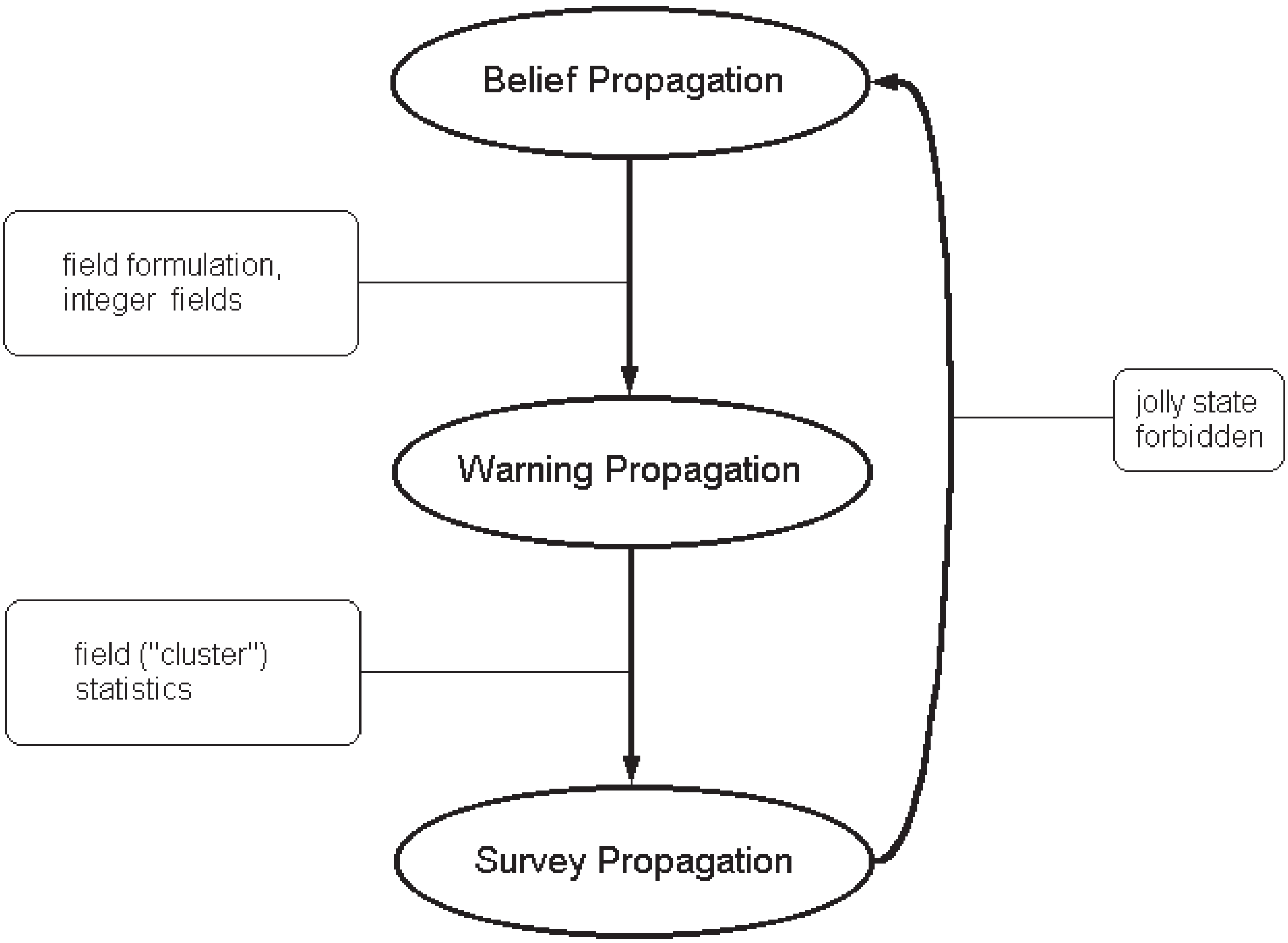
Finally, going back to pseudo-distributions, we obtain

$$Q_{a\to i}(s) = 1 - [\delta(s) + \delta(s - J_{a\to i})] \prod_{j\in b\backslash i} [Q_{j\to a}(J_{a\to j}) - Q_{j\to a}(0)]$$

$\rightarrow$ SP recursion relations degenerate into BP ones, if the $s = 0$ ("jolly") state is forbidden.

Summary

This is what we have done so far:



From double loop algorithm to "damped" BP

BP is efficient (single loop), but does not converge for $\alpha > 3.86$, where a full replica symmetry breaking (RSB) occurs.

$\rightarrow$ We have performed minimization with the algorithm proposed by Heskes, Albers, and Kappen (HAK), a double loop (less efficient) algorithm, which is nevertheless guaranteed to converge [*UAI-2003 proceedings pp. 313-320 (2003)*].

At each main loop iteration, the HAK algorithm minimizes (by message passing) a convex upperbound to the free energy

$$\begin{aligned} \beta\bar{F} &= \sum_a \sum_{\sigma_a} p_a(\sigma_a)\left[\beta\tilde{E}_a(\sigma_a) + \ln p_a(\sigma_a)\right] \\ &\quad - \sum_i \tilde{B}_i \sum_{s_i} p_i(\sigma_i)\ln p_i(\sigma_i), \end{aligned}$$

where $\tilde{B}_i$ are suitably chosen, and

$$\beta\tilde{E}_a(\sigma_a) = \beta E_a(\sigma_a) - \sum_{i\in a}\frac{B_i - \tilde{B}_i}{C_i}\ln\bar{p}_i(\sigma_i).$$

• Notice that $\bar{p}_i(\sigma_i)$ are fixed during minimization of the upperbound (inner loop).

Experimentally, it turns out that (at least for satisfiability) the double-loop procedure converges even for a single inner loop iteration, giving rise to a numerical routine similar to ordinary BP with "damping" coefficients $D_i = (B_i - \tilde{B}_i)/(1 + B_i - \tilde{B}_i)$.

The first update rule of ordinary BP can be rewritten as

$$\begin{aligned} p_i(\sigma) &:= \prod_{b\in i} m_{b\to i}(\sigma) \\ m_{i\to a}(\sigma) &:= \frac{p_i(\sigma)}{m_{a\to i}(\sigma)} \end{aligned}$$

Damped BP replaces the former assignment by a geometric average

$$p_i(\sigma) \quad := \quad p_i(\sigma)^{D_i} \prod_{b \in i} m_{b \to i}(\sigma)^{1-D_i}$$

Damped BP is found to converge even for site independent damping coefficients $D < D_i$, increasing speed, but reducing stability. For $D_i = 0$ we get back ordinary BP.

Also a linear average is found to work

$$p_i(\sigma) \quad := \quad D p_i(\sigma) + (1 - D) \prod_{b \in i} m_{b \to i}(\sigma)$$

Damped BP, applied to subsequently reduced (decimated) formulae, finds the correct ground state up to $\alpha \approx 4.15$, where the hard (1-RSB) region begins.

Survey Propagation overcomes this problem [M. Mézard and R. Zecchina, Phys. Rev. E **66**, 056126 (2002)].

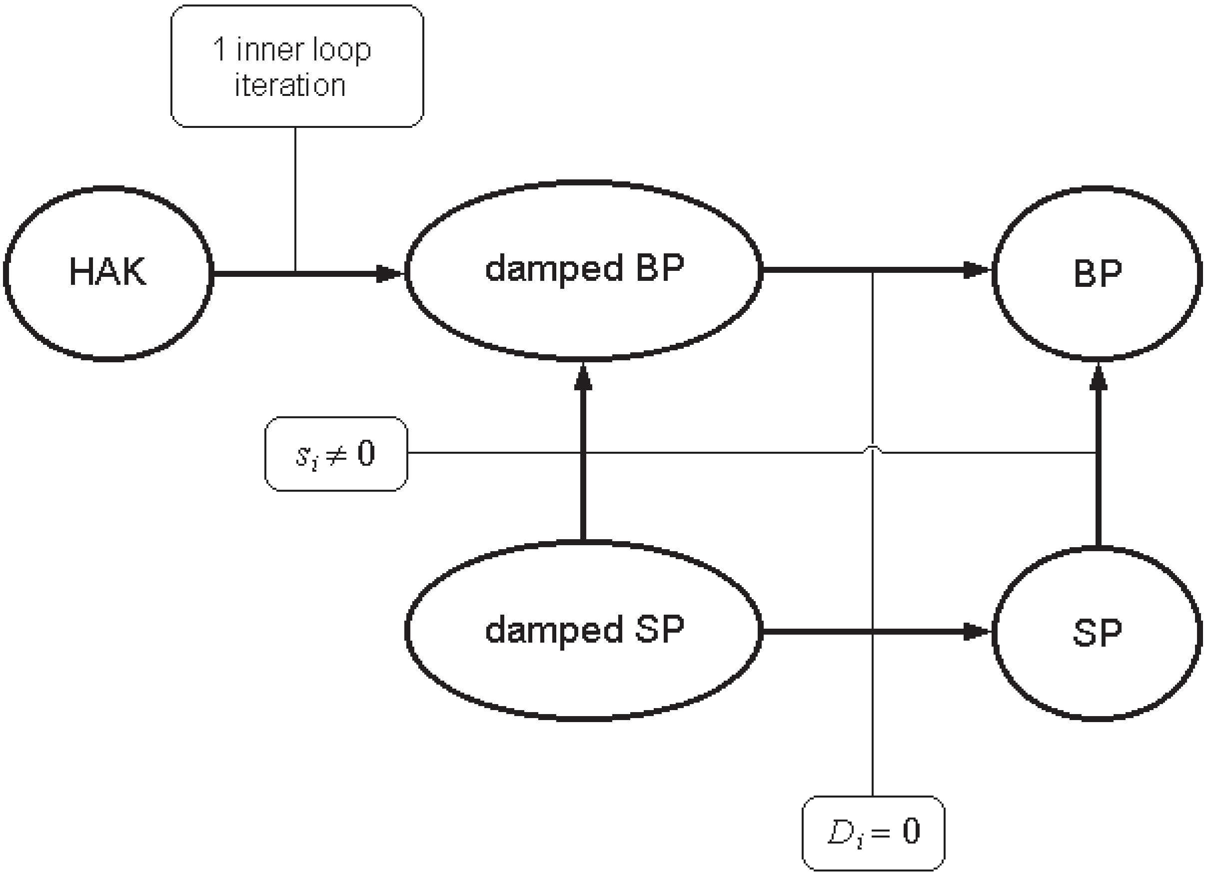
• Formally, SP introduces a third (jolly) state $s_i = 0$, describing spins that are not frozen.

• We have reformulated SP in such a way that BP is reobtained simply by forbidding the jolly state.

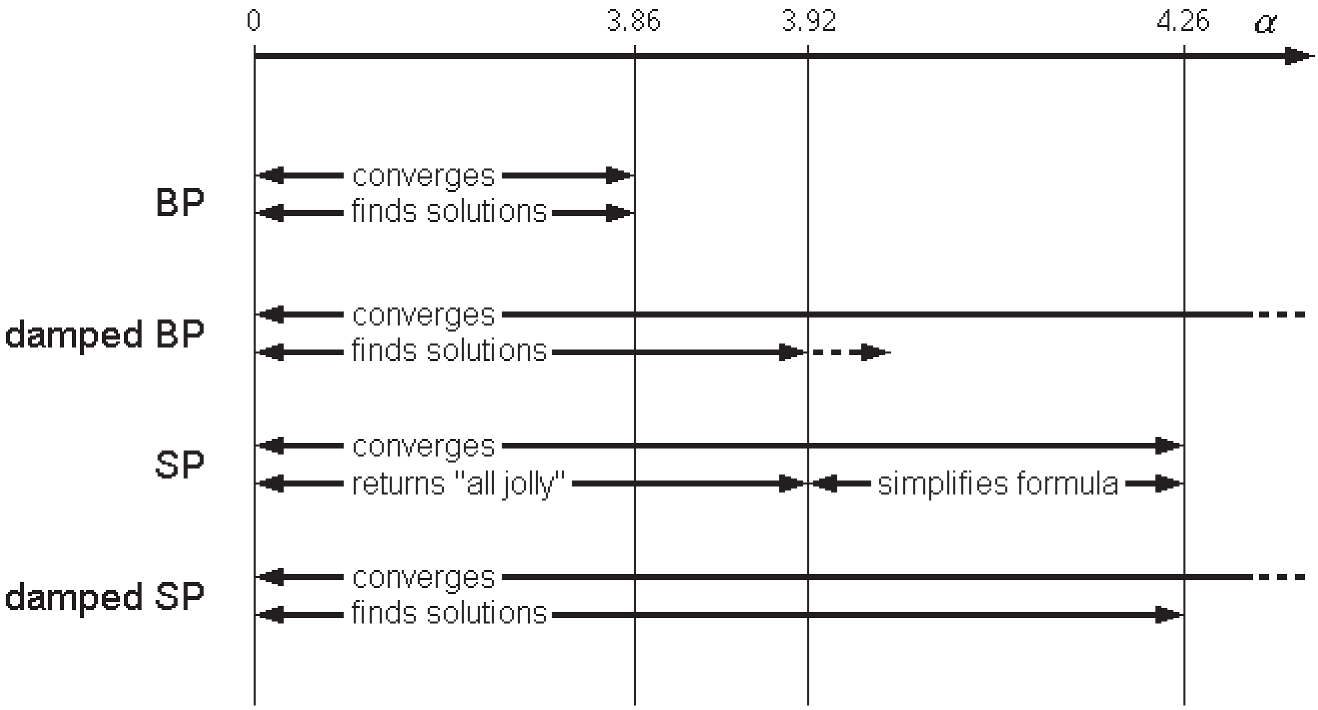
$\longrightarrow$ The damping trick can be straightforwardly extended to SP, with some advantages:

- robustness is improved in particular cases (small instances close to the sat-unsat transition);

- a unique routine is able to solve hard formulae and "easy" subformulae obtained by decimation.



Convergence regions

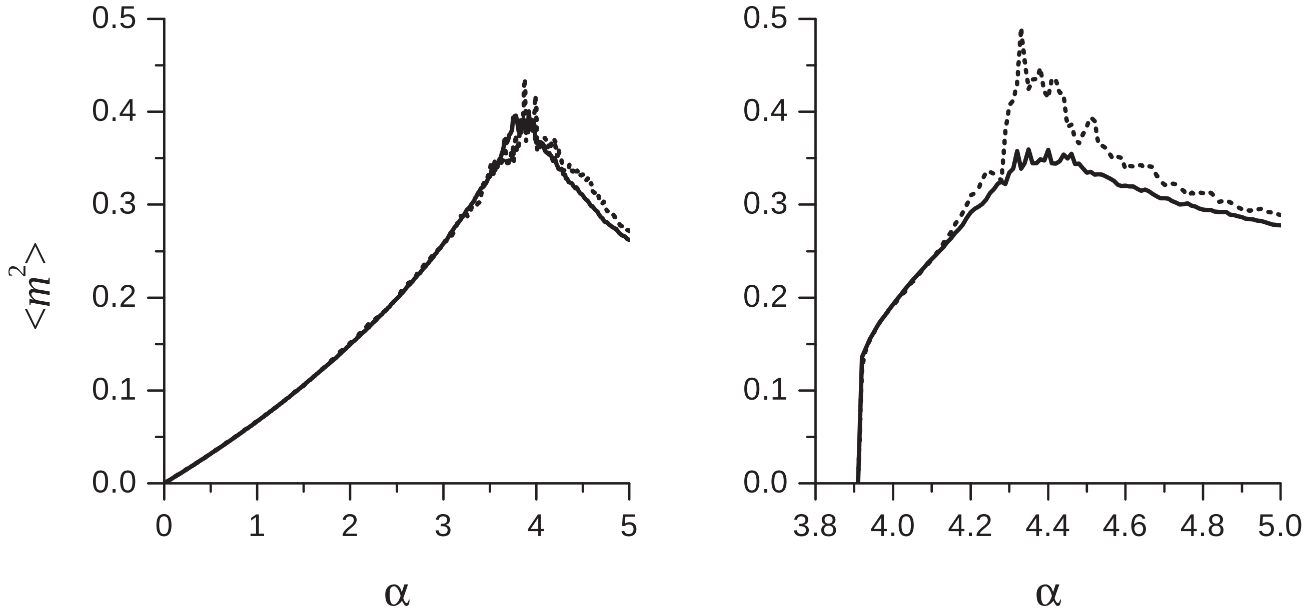


Experiments on random instances

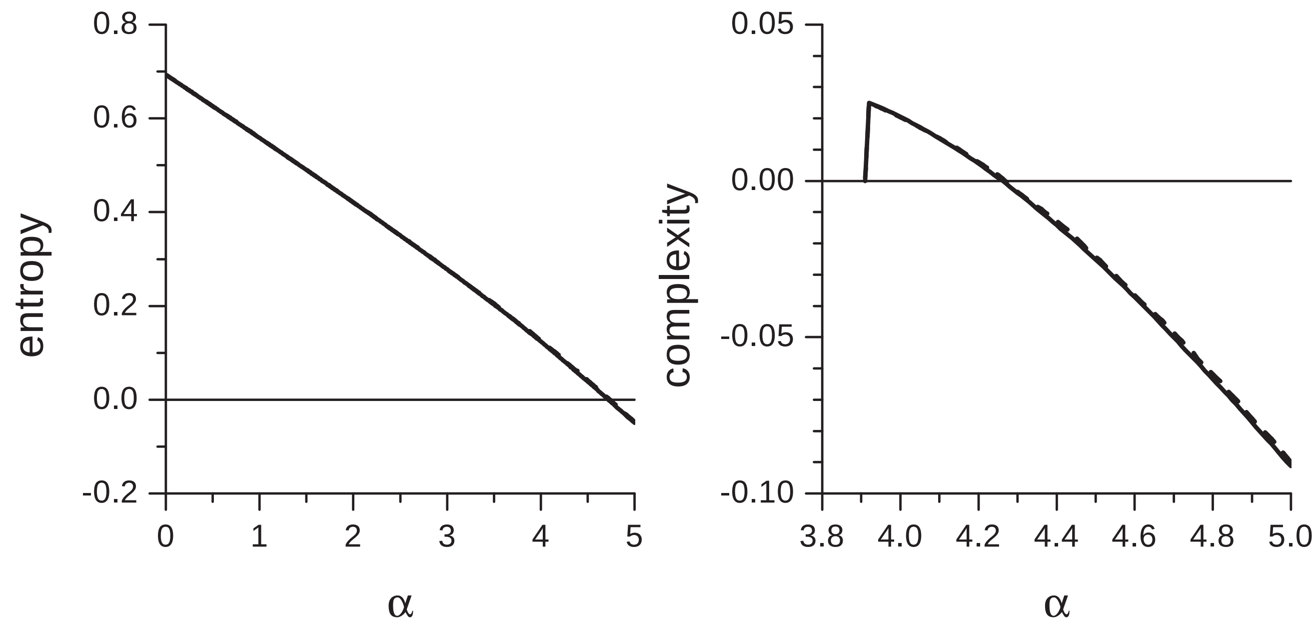
Let us compare the results of damped BP (left) and damped SP (right) for random 3-SAT instances with $N = 10000$ (dashed lines) and $N = 100000$ (solid lines).

$\rightarrow$ Convergence is easily obtained up to $\alpha = 5$ and beyond.

Average squared magnetization

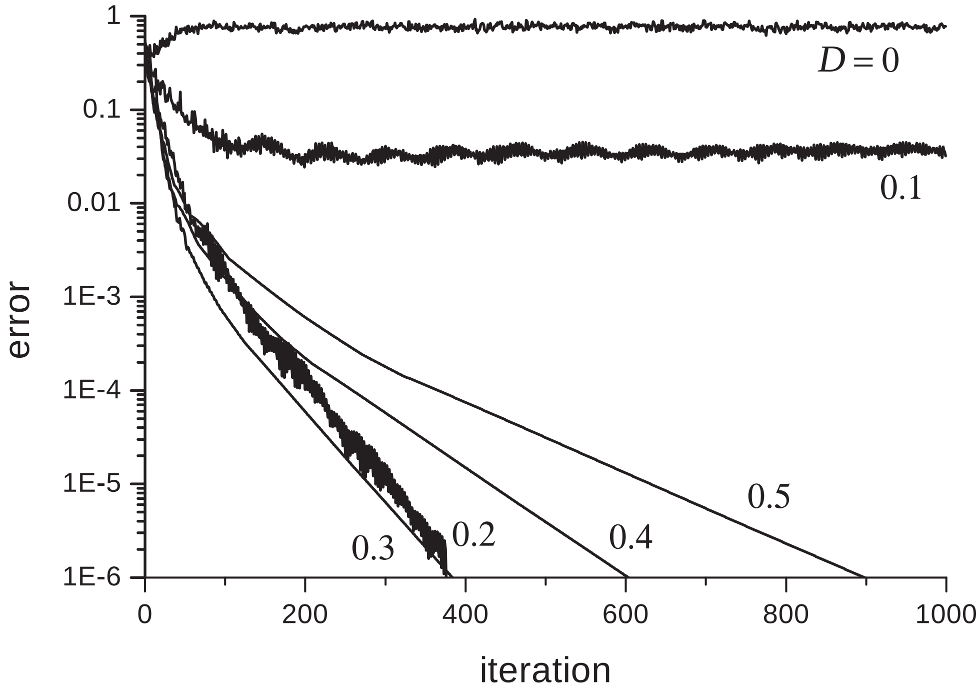


Entropy (BP) and complexity (SP)



Effects of damping

Let us analyze the effect of different damping coefficients on a random instance with $N = 10000$ and $\alpha = 3.90$.



In this region:

→ Ordinary BP ($D = 0$) does not converge, whereas damped BP converges for some damping coefficient $D > 0$;

→ Satisfying assignments can be found by decimation and repeated damped BP runs.

Conclusions

We have proposed a modified (damped) message-passing procedure for the satisfiability of random boolean formulae.

Such idea is based on a double-loop method, recently proposed for the minimization of Bethe and Kikuchi free energies.

The method can be also extended to the framework of Survey Propagation.

We obtain a unified message-passing scheme, with improved convergence properties.

- → As BP, it can be used in the full RSB region, where ordinary BP does not converge.

- → As SP, it can be used in the "hard" region near the sat-unsat transition. If the jolly state is forbidden, it also allows to solve subformulae generated by "survey inspired decimation".