TD 5: Random XORSAT problems

I) Bounds on α_{sat}

We recall from TD2 that for Z a random var. in $\mathbb{N}$:

$$\frac{\mathbb{E}(Z)^2}{\mathbb{E}(Z^2)} \leq \mathbb{P}(Z>0) \leq \mathbb{E}(Z) \quad (*)$$

1.) $\mathbb{E}(Z) = ?$, $Z = \sum_{\underline{x}} \mathbb{1}(\underline{x} \text{ satisfies } F)$

↑ indicator function

$$\mathbb{1}(\text{"A"}) = \begin{cases} 1 \text{ if A true} \\ 0 \text{ otherwise} \end{cases}$$

We rewrite $Z = \sum_{\underline{x}} \prod_{a=1}^{M} \mathbb{1}\left(x_{i_a^1} + x_{i_a^2} + \dots + x_{i_a^k} = y_a \bmod 2\right)$

$$\Rightarrow \mathbb{E}(Z) = \sum_{\underline{x}} \prod_{a=1}^{M} \mathbb{E}\left(\mathbb{1}\left(x_{i_a^1} + x_{i_a^2} + \dots + x_{i_a^k} = y_a \bmod 2\right)\right)$$

since the M eqs. are drawn independently

Using that:

$$\mathbb{E}\left(\mathbb{1}\left(x_{i_1^a} + x_{i_2^a} + \dots + x_{i_L^a} = y_a \bmod 2\right)\right) = \mathbb{P}\left(y_a = x_{i_1^a} + x_{i_2^a} + \dots + x_{i_L^a} \bmod 2\right)$$

$$= \frac{1}{2} \quad \text{since } y_a \text{ is independent of } \underline{x}$$

One finds that: $\mathbb{E}(Z) = \sum_{\underline{x}} \prod_{a=1}^{M} \frac{1}{2} = \sum_{\underline{x}} \left(\frac{1}{2}\right)^M$

$$\Rightarrow \boxed{\mathbb{E}(Z) = \frac{2^N}{2^M} = 2^{N - \alpha N}}$$

Suppose that $\alpha > 1$, then $\mathbb{E}(Z) \xrightarrow[N\to\infty]{} 0$

and from the property (*) [right inequality]

using $P_{sat}(\alpha, N) = \mathbb{P}(Z > 0) \leq \mathbb{E}(Z)$

one obtains $P_{sat}(\alpha > 1, N) \xrightarrow[N\to\infty]{} 0$

and therefore $\boxed{\alpha_{sat} \leq 1}$.

2.) We now want to compute $\mathbb{E}(Z^2)$ to exploit the left inequality of (⋆).

We write Z^2 as:

$$Z^2 = \sum_{\underline{x},\underline{x}'} \prod_{a=1}^{M} \mathbb{1}\left(x_{i_a^1} + \dots + x_{i_a^h} = x'_{i_a^1} + \dots + x'_{i_a^h} = y_a\right)$$

where the equalities '=' hold modulo 2.

Taking the average yields:

$$\mathbb{E}(Z^2) = \sum_{\underline{x},\underline{x}'} \prod_{a=1}^{M} \mathbb{P}\left(x_{i_a^1} + \dots + x_{i_a^h} = x'_{i_a^1} + \dots + x'_{i_a^h} = y_a\right)$$

recall that $y_a = 0,1$ independently of $\underline{x}, \underline{x}'$

Therefore

$$\mathbb{P}\left(x_{i_a^1} + \dots + x_{i_a^h} = x'_{i_a^1} + \dots + x'_{i_a^h} = y_a\right) = \frac{1}{2}\underbrace{\mathbb{P}\left(x_{i_a^1} + \dots + x_{i_a^h} = x'_{i_a^1} + \dots + x'_{i_a^h}\right)}$$

$$= \frac{1}{2} \quad P(\underline{x}, \underline{x}')$$

which is independent of 'a'

Hence: $\boxed{\mathbb{E}(Z^2) = \sum_{\underline{x}, \underline{x}'} \frac{1}{2^N} p(\underline{x}, \underline{x}')^M}$

Follow the hint and write the double sum as:

$$\sum_{\underline{x}, \underline{x}'} = \sum_{\underline{x}} \sum_{D=0}^{N} \sum_{\underline{x}' \,|\, d(\underline{x}, \underline{x}') = D}$$

where the "Hamming distance" $d(\underline{x}, \underline{x}')$ is defined as follows:

Suppose without loss of generality that $\underline{x} = 0$,

i.e. $\underline{x} = 0, 0, 0, \ldots, 0 \,|\, 0, 0 \ldots, 0$

ensemble of indices $\mathcal{D}$ | ensemble $\overline{\mathcal{D}}$

Then $\underline{x}' = \underbrace{1, 1, 1, \ldots, 1}_{D \text{ distinct 'spins' 'variables'}} \,|\, \underbrace{0, 0, \ldots, 0}_{N-D \text{ common 'spins' 'variables'}}$

then $d(\underline{x}, \underline{x}') = D = \text{card}(\mathcal{D})$.

Defining: $P_D(\underline{x}, \underline{x}') = \mathbb{P}\left(x_{i_1} + \cdots + x_{i_k} = x'_{i_1} + \cdots + x'_{i_k} \text{ & } d(\underline{x}, \underline{x}') = D\right)$

on can rewrite:

$$\mathbb{E}(Z^2) = \frac{1}{2^N} \sum_{D=0}^{N} \sum_{\underline{x}} \sum_{\underline{x}' \, | \, d(\underline{x}, \underline{x}') = D} \left[P_D(\underline{x}, \underline{x}') \right]^M$$

We now compute $P_D(\underline{x}, \underline{x}')$:

To proceed we denote by:

P_ℓ = proba. to have a k-uplets $\{i_1, i_2, \ldots, i_k\}$ which has exactly ℓ indices that fall into the D distinct variables that separate $\underline{x}'$ from $\underline{x}$.

One has:

$$P_\ell = \frac{1}{\binom{N}{k}} \underbrace{\binom{D}{\ell}\binom{N-D}{k-\ell}}_{\#\, k\text{-uplets with } \begin{cases} \ell \text{ indices} \in D \\ k-\ell \text{ "} \in \overline{D} \end{cases}}$$

Let us now assume that $\{i^1, \dots, i^h\}$ has exactly ℓ indices that fall in the ensemble $\mathcal{D}$. One has:

$$x_{i^1} + x_{i^2} + \dots + x_{i^h} = x_{i^1,\mathcal{D}} + \dots + x_{i^\ell,\mathcal{D}} + x_{i^1,\bar{\mathcal{D}}} + \dots + x_{i^{k-\ell},\bar{\mathcal{D}}}$$

$$x'_{i^1} + x'_{i^2} + \dots + x'_{i^h} = \underbrace{x'_{i^1,\mathcal{D}} + \dots + x'_{i^\ell,\mathcal{D}}}_{\text{here } x_{i,\mathcal{D}} = x'_{i,\mathcal{D}} + 1} + \underbrace{x'_{i^1,\bar{\mathcal{D}}} + \dots + x'_{i^{k-\ell},\bar{\mathcal{D}}}}_{\text{here } x_{i,\bar{\mathcal{D}}} = x'_{i,\bar{\mathcal{D}}}}$$

Hence: if ℓ is even : $x_{i^1} + \dots + x_{i^h} = x'_{i^1} + \dots + x'_{i^h} \bmod 2$

if ℓ is odd : $x_{i^1} + \dots + x_{i^h} \neq x'_{i^1} + \dots + x'_{i^h} \bmod 2$

and therefore:

$$P_{\mathcal{D}}(\underline{x}, \underline{x}') = \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{h} P_\ell$$

$$\Rightarrow \boxed{P_{\mathcal{D}}(\underline{x}, \underline{x}') = \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{h} \frac{1}{\binom{N}{h}} \binom{D}{\ell} \binom{N-D}{h-\ell}}$$

We obtain :

$$\mathbb{E}(Z^2) = \sum_{D=0}^{N} \sum_{\underline{x}} \sum_{\underline{x}' \,|\, d(\underline{x},\underline{x}')=D} \left[\frac{1}{2} \sum_{\substack{l=0 \\ l \text{ even}}}^{k} \frac{1}{\binom{N}{l}} \binom{D}{l} \binom{N-D}{k-l} \right]^M$$

Nber of terms in this double sum $= 2^N \binom{N}{D}$

(2^N: choose $\underline{x}$; $\binom{N}{D}$: D diff. indices among N)

The final formula reads:

$$\boxed{\mathbb{E}(Z^2) = 2^N \sum_{D=0}^{N} \binom{N}{D} \left[\frac{1}{2} \sum_{\substack{l=0 \\ l \text{ even}}}^{k} \frac{1}{\binom{N}{l}} \binom{D}{l} \binom{N-D}{k-l} \right]^M}$$

3.) In the large N limit we assume that the sum over D in $\mathbb{E}(Z^2)$ is dominated by $D = \alpha(N)$. We thus set $D = \alpha N$, $\alpha \in [0,1]$. In this limit one can use Stirling's formula:

$$n! \underset{n \to \infty}{\sim} \sqrt{2\pi n}\, e^{n \ln n - n}$$

To obtain:

a): $\binom{N}{D} = \binom{N}{dN} \underset{N\to\infty}{\sim} e^{N\left[-d\ln d - (1-d)\ln(1-d)\right]} \times \frac{1}{\sqrt{2\pi N d(1-d)}}$

(keeping d fixed)

b): $\frac{\binom{D}{\ell}\binom{N-D}{k-\ell}}{\binom{N}{k}} \sim \binom{k}{\ell} d^{\ell}(1-d)^{k-\ell}$

We then rewrite the $\sum_{\substack{\ell=0 \\ \text{even}}}^{k}$ in Eq. (4) of the sheet as:

$$\sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{k} \frac{1}{\binom{N}{k}} \binom{D}{\ell}\binom{N-D}{k-\ell} \simeq \sum_{\ell=0}^{k} \frac{1+(-1)^{\ell}}{2} \binom{k}{\ell} d^{\ell}(1-d)^{k-\ell}$$

$$= \frac{1}{2}\left[\left(d+(1-d)\right)^{k} + \left(-d+(1-d)\right)^{k} \right]$$

$$= \frac{1}{2}\left[1 + (1-2d)^{k} \right]$$

$$\rightarrow \mathbb{E}(Z^2) \simeq 2^N \sum_{D=0}^{N} \frac{1}{\sqrt{2\pi N d(1-d)}} e^{N(-d\ln d - (1-d)\ln(1-d))} \left[\frac{1+(1-2d)^k}{4}\right]^N$$

$$\simeq \frac{2^N}{4^{\alpha N}} \sum_{D=0}^{N} e^{N\left[-d\ln d - (1-d)\ln(1-d) + \alpha \ln\left(1+(1-2d)^k\right)\right]}$$

at exponential order

where we have set $M = \alpha N$.

The sum over D can then be evaluated by a saddle point method:

$$\mathbb{E}(Z^2) \sim \frac{2^N}{4^{\alpha N}} e^{N \sup_{d\in[0,1]} \left[\underbrace{-d\ln d - (1-d)\ln(1-d) + \alpha \ln\left(1+(1-2d)^k\right)}_{\varphi(\alpha, d)}\right]}$$

On the other hand we have seen that:

$$\mathbb{E}(Z) = \frac{2^N}{2^M} \Rightarrow \mathbb{E}(Z)^2 = \frac{2^{2N}}{4^{\alpha N}}$$

Hence:

$$\frac{\mathbb{E}(Z)^2}{\mathbb{E}(Z^2)} = 2^N e^{-N \sup_{d\in[0,1]} \varphi(\alpha, d)}$$

$$= e^{N\ln 2 + N \inf_{d\in[0,1]} (-\varphi(\alpha, d))}$$

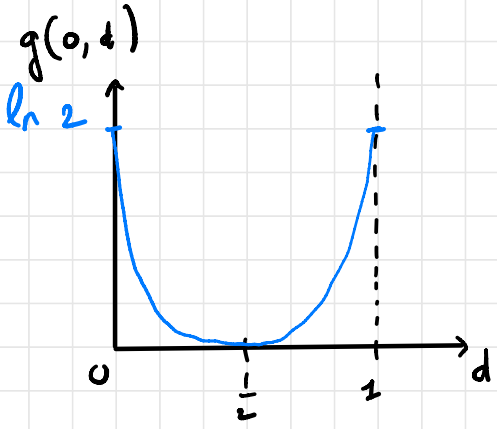
and finally: $\lim_{N\to\infty} \frac{1}{N} \ln \frac{\mathbb{E}(Z)^2}{\mathbb{E}(Z^2)} = \ln 2 + \inf_{d\in[0,1]} -\varphi(\alpha,d)$

$$= \inf_{d\in[0,1]} g(\alpha,d)$$

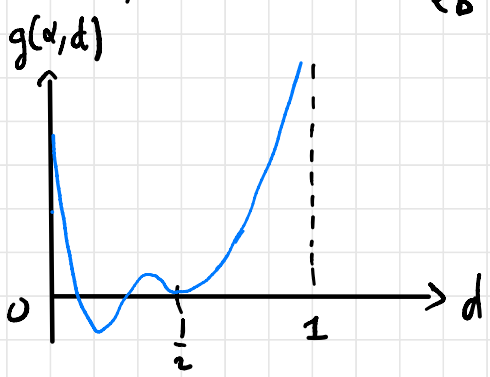
where $g(\alpha,d) = \ln 2 - \varphi(\alpha,d)$

$$= \ln 2 + d\ln d + (1-d)\ln(1-d) - \alpha \ln\left(1+(1-2d)^k\right)$$

4.) Let us analyse $g(\alpha,d)$ as a function of d, increasing α



by increasing α, $g(\alpha,d)$ develops a secondary minimum where, for $\alpha > \alpha_{eb}$:



Therefore, for $\alpha < \alpha_{lb}$ one has $\inf_{\alpha \in [0,1]} q(\alpha, d) = 0$

while for $\alpha > \alpha_{lb}$ one has $\inf_{\alpha \in [0,1]} q(\alpha, d) < 0$.

Here in the case $\alpha > \alpha_{lb}$ the left inequality in (*) gives:

$$\mathbb{P}(Z > 0) \geq \frac{\mathbb{E}(Z)^2}{\mathbb{E}(Z^2)} \xrightarrow[N \to \infty]{} 0$$

and we do not get useful information.

However, in the case $\alpha < \alpha_{lb}$ one needs to work out all the prefactors in the saddle-point computation (see R. Monasson, arXiv 0704.2536, App. C) and one obtains:

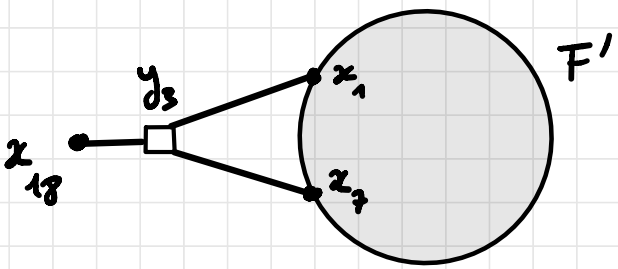
$$\frac{\mathbb{E}(Z)^2}{\mathbb{E}(Z^2)} \xrightarrow[N \to \infty]{} 1, \text{ hence } \mathbb{P}(Z > 0) \xrightarrow[N \to \infty]{} 1.$$

Therefore we deduce that $\boxed{\alpha_{sat} \geq \alpha_{lb}}$.

Rk: For $k = 3$, a numerical evaluation gives $\alpha_{lb} \simeq 0,889\ldots$

II| Leaf removal.

1.) Suppose that the system F contains a leaf, for instance:



$$F = F' \cup \{x_{18} + x_1 + x_7 = y_3\}$$

Let us show that F satisfiable $\Leftrightarrow$ F' satisfiable.

- It is clear that F' unsat. $\Rightarrow$ F unsat.
  hence F' sat. $\Leftarrow$ F sat.

- Suppose now that F' is satisfiable.
Then one can take one sol. and find

$$x_{18} = y_3 - x_1 - x_7 = y_3 + x_1 + x_7$$

(modulo 2)

since this is the only constraint implying x_{18}.

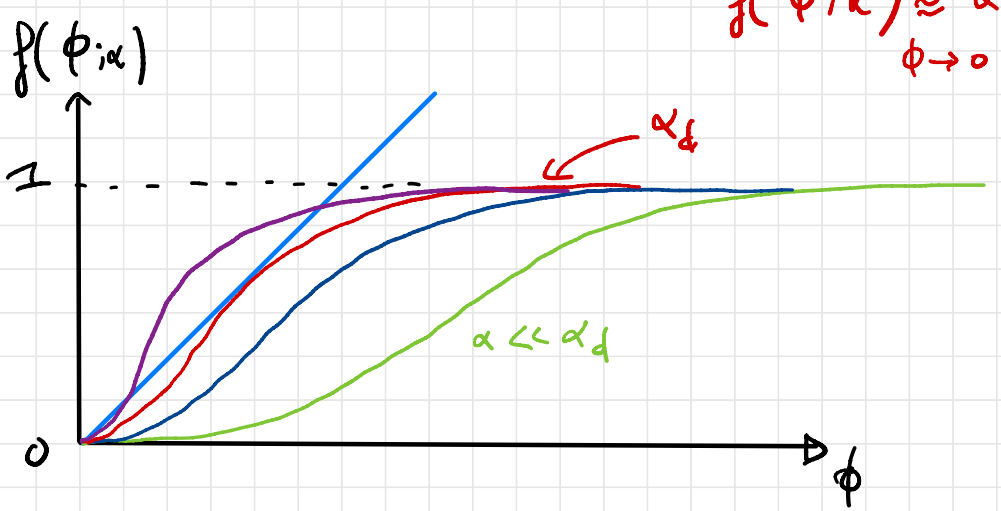
2.) One arrives at the relation:

$$f_{core} = 1 - e^{-\alpha k \phi^{k-1}} - \alpha k \phi^{k-1} e^{-\alpha k \phi^{k-1}}$$

where ϕ is the largest sol. of $\phi = 1 - e^{-\alpha k \phi^{k-1}}$

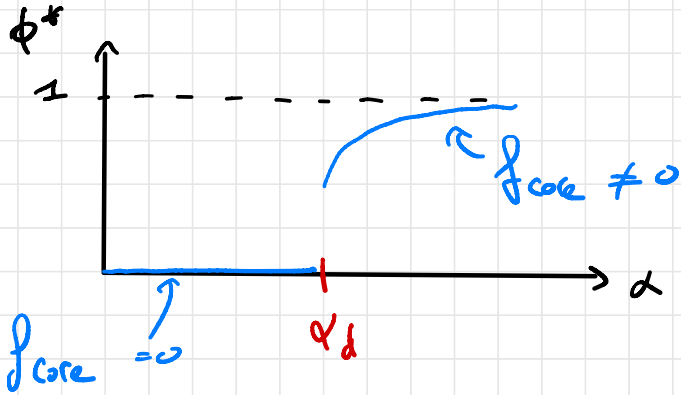
Rk: in the case $k=2$, the eq. for ϕ is reminiscent of the eq. found for the fraction of sites in the giant component of the Erdös-Rényi random graph with the correspondance $2\alpha = c$.

Graphical solution: $\phi = \underbrace{1 - e^{-\alpha k \phi^{k-1}}}_{f(\phi;\alpha) \underset{\phi \to 0}{\approx} \alpha k \phi^{k-1}}$

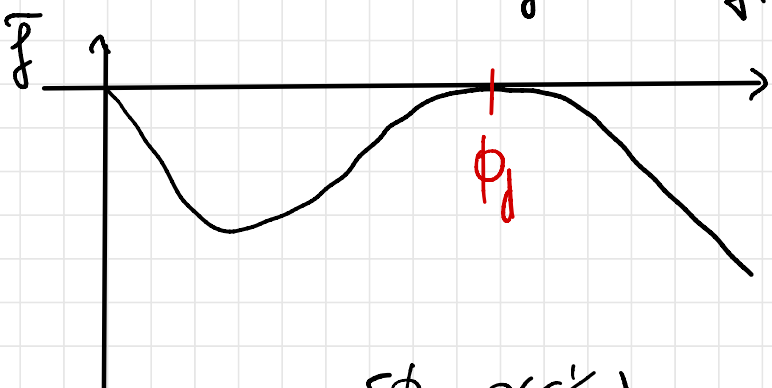


=> ϕ^* the largest sol. of $\phi = f(\phi;\alpha)$ is a discontinuous

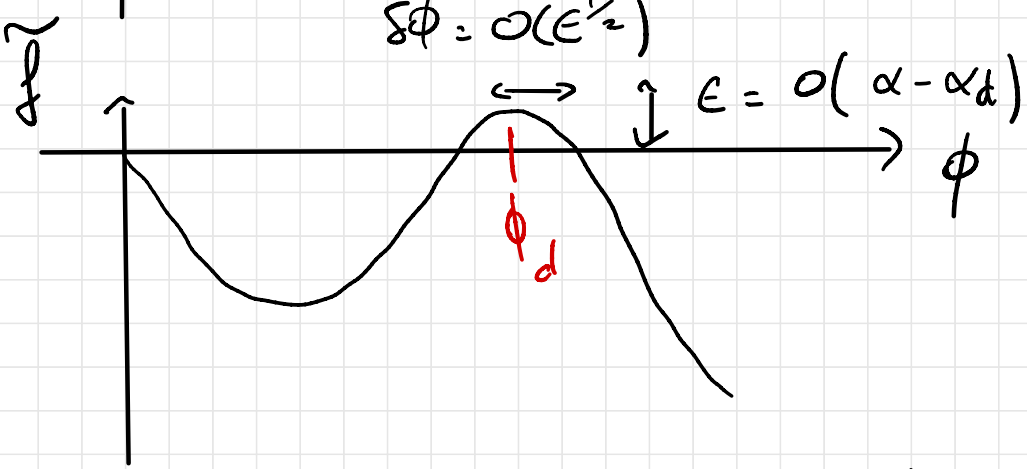
function of α:



Let us introduce the function $\tilde{f}(\phi, \alpha) = f(\phi, \alpha) - \phi$



$\alpha = \alpha_d$



$$\Rightarrow \delta\phi = O\left((\alpha - \alpha_d)^{1/2}\right).$$

More formally we see (geometrically) that:

$$\tilde{f}(\phi_d, \alpha_d) = 0 \quad \text{and} \quad \frac{\partial \tilde{f}}{\partial \phi_d}(\phi_d, \alpha_d) = 0$$

We want to solve:

$$\tilde{f}(\phi, \alpha) = 0 \quad , \quad \alpha = \alpha_d + \delta\alpha$$
$$\phi = \phi_d + \delta\phi$$

one has: $\tilde{f}(\phi, \alpha) = \tilde{f}(\phi_d + \delta\phi, \alpha_d + \delta\alpha) = 0$

$$\Leftrightarrow \underbrace{\tilde{f}(\phi_d, \alpha_d)}_{=0} + \delta\phi \underbrace{\frac{\partial \tilde{f}}{\partial \phi}(\phi_\alpha, \alpha_d)}_{=0} + \delta\alpha \underbrace{\frac{\partial \tilde{f}}{\partial \alpha}(\phi_\alpha, \alpha_d)}_{\neq 0}$$

$$+ \frac{1}{2}(\delta\phi)^2 \underbrace{\frac{\partial^2 \tilde{f}}{\partial \phi}(\phi_\alpha, \alpha_d)}_{\neq 0} + O(\delta\phi\delta\alpha, (\delta\alpha)^2)$$

$$= 0$$

$$\Rightarrow \boxed{\delta\phi = o((\delta\alpha)^{1/2})}$$

Rk: one checks a posteriori that

$$\mathcal{O}(\delta\phi\,\delta\alpha) = \mathcal{O}((\delta\phi)^3) = o((\delta\phi)^4)$$

$$o((\delta\alpha)^4) = \mathcal{O}((\delta\phi)^4) = o((\delta\phi)^2)$$

3.) Note that:

$$\mathbb{E}[Z_{core}] = 2^{N_{core} - M_{core}} = 2^{N_{core}(1-\alpha_{core})}$$

↑ # of solutions of the reduced formula (same computation as in I.1).

For $\alpha > \alpha_*$, $\alpha_{core} > 1 \Rightarrow \mathbb{E}(Z_{core}) \to 0$

$\Rightarrow \mathbb{P}(Z_{core} > 0) \to 0$ as $N \to \infty$

and therefore for $\alpha > \alpha_*$ the core is unsat

with proba. 1 and the full formula is thus also unsat with proba. 1,

$$\Rightarrow \quad \boxed{\alpha_{sat} \leq \alpha_*} \; .$$

Rk: it turns out that the left inequality of (*) applied to the core yields a tight bound $\Rightarrow \alpha_{sat} = \alpha_*$.

For $k=3$: $\alpha_{sat} = \alpha_* = 0,9318\ldots$