1/4

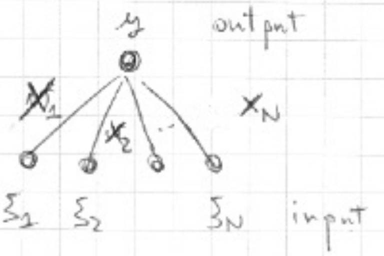
(A.) PERCEPTRON TUTORIAL

BASED ON
SciPost Phys. 2, 019 (2017)
[check SciPost if you don't know it]

(won't talk about its full history)

(1) Definitions

(a) Perceptron: simplest model of a neuron
(Rosenblatt 1957) also a simple binary classifier



$$y = \mathrm{sgn}(\vec{x} \cdot \vec{\xi})$$

Many questions about perceptron: learning etc.

(b) Today, very simple questions: how many random patterns can I store?

[Derrida-Gardner 1988]

$$y^\mu = \mathrm{sgn}(\vec{x} \cdot \vec{\xi}^\mu) \iff y^\mu \, \vec{x} \cdot \vec{\xi}^\mu \geq 0$$

$\mu = 1 \dots M = \alpha N$ random patterns.

Choose: $\xi_i^\mu \sim \mathcal{N}(0, \frac{1}{N})$ independently

$y^\mu = \pm 1 \quad p = \frac{1}{2} \Rightarrow$ then, y^μ is irrelevant, neglect it

$\vec{x} \in \mathbb{R}^N$ but note that if $\vec{x}$ sol $\Rightarrow \lambda \vec{x}$ sol. $\forall \lambda \in \mathbb{R}^+$

If I want a finite volume of solutions, need to normalize $\vec{x}$

$$\vec{x} \cdot \vec{x} = \sum_i x_i^2 = N$$

Finally we generalize slightly the constraint: $\vec{x} \cdot \vec{\xi}^\mu \geq \sigma \geq 0$

Finally, the problem is:

Can I find $\vec{x} \in \mathbb{R}^N$, $\sum_i x_i^2 = N$, such that

$$\vec{x} \cdot \vec{\xi}^\mu \geq \sigma \quad \forall \mu = 1 \dots M = \alpha N, \quad \xi_i^\mu \sim \mathcal{N}(0, \frac{1}{N})$$

(c) Stat. mech. formulation

$$R^\mu(\vec{x}) = \vec{x}\cdot\vec{\xi}^\mu \quad \geq \sigma$$

$$H(\vec{x}) = \sum_{\mu=1}^{M} v\left(R^\mu(\vec{x}) - \sigma\right)$$

$$\begin{cases} v \geq 0 & h \leq 0 \\ v = 0 & h \geq 0 \end{cases}$$

$H(\vec{x}) = 0 \iff \vec{x}$ is a solution

$H(\vec{x}) > 0 \iff \vec{x}$ is not a solution

Introduce a temperature and compute the partition function.

$$Z = \underbrace{\int d^N\!x\, \delta\left(\sum x_i^2 - N\right)}_{D\vec{x}} e^{-\beta H(\vec{x})} = \int D\vec{x} \prod_\mu e^{-\beta v(R^\mu(\vec{x}) - \sigma)}$$

$$= \int D\vec{x} \prod_\mu \int dr^\mu\, e^{-\beta v(r^\mu - \sigma)}\, \delta\left(r^\mu - \vec{x}\cdot\vec{\xi}^\mu\right)$$

$F = -T \log Z$

$$\propto \int D\vec{x} \prod_\mu \int dr^\mu\, dk^\mu\, e^{-\beta v(r^\mu - \sigma) + i k^\mu (r^\mu - \vec{x}\cdot\vec{\xi}^\mu)}$$

(2) Replica method

(a) General calculation.

$$F = -T\,\overline{\log Z} = -T \lim_{n\to 0} \partial_n \overline{Z^n} = -T \lim_{n\to 0} \partial_n \log \overline{Z^n} \quad \left[\,\overline{\cdot}\ \text{average over } \vec{\xi}^\mu\right]$$

$$\overline{Z^n} = \int D\vec{x}_a \int dr_a^\mu\, dk_a^\mu\, e^{-\beta \sum_{a,\mu} v(r_a^\mu - \sigma)}\, \overline{e^{\sum_{a,\mu} i k_a^\mu (r_a^\mu - \vec{x}_a\cdot\vec{\xi}^\mu)}}$$

Gaussian integral.
$$\overline{e^{-\sum_{a\mu} i k_a^\mu \vec{x}_a\cdot\vec{\xi}^\mu}} = e^{-\frac{1}{2}\sum_{ab,\mu} k_a^\mu k_b^\mu Q_{ab}} \qquad Q_{ab} = \frac{1}{N}\vec{x}_a\cdot\vec{x}_b$$

$$= \int D\vec{x}_a \int dr_a^\mu\, dk_a^\mu\, e^{-\beta\sum_{a\mu} v(r_a^\mu - \sigma) + \sum_{a\mu} i k_a^\mu r_a^\mu - \frac{1}{2}\sum_{ab,\mu} k_a^\mu k_b^\mu Q_{ab}}$$

$$= \int D\vec{x}_a \left[\int dr_a\, dk_a\, e^{-\beta\sum_a v(r_a - \sigma) + \sum_a i k_a r_a - \frac{1}{2}\sum_{ab} Q_{ab} k_a k_b}\right]^M$$

$$\propto \int D\vec{x}_a\, I(\hat{Q})^M \qquad I(\hat{Q}) = \int dr_a\, \frac{e^{-\frac{1}{2}\sum_{ab} Q_{ab}^{-1} r_a r_b - \beta\sum_a v(r_a - \sigma)}}{\sqrt{2\pi \det Q_{ab}}}$$

Rotationally invariant integral. (generalized spherical coordinates)

$$\int d\vec{x}_1 \dots d\vec{x}_n\, F(\{\vec{x}_i\}) = \int d\hat{q}\;\; C_{n+1,N}\, e^{\frac{N-(n+1)}{2}\log\det\hat{q}}\; F(\hat{q})$$

Neglect proportionality constant [integral over q_{ab} with $a \leq b$]

$$\overline{Z^n} \propto \int d\hat{q}\, e^{\frac{N}{2}\log\det\hat{q}} \prod_a \delta(q_{aa}-1)\;\; I(\hat{q})^M$$

$$= \int d\hat{q}\, e^{N A(\hat{q})} \quad \leftarrow q_{ab} \text{ with } a<b,\; q_{aa}=1$$

$$\frac{1}{N}\log\overline{Z^n} \xrightarrow[N\to\infty]{} \max_{\hat{q}} A(\hat{q}) \qquad \boxed{A(\hat{q}) = \frac{1}{2}\log\det\hat{q} + \alpha\log I(\hat{q})}$$

$$f = \frac{F}{N} = -T \lim_{n\to 0} \partial_n\left[\max_{\hat{q}} A(\hat{q})\right]$$

Recipe: fix n; find $\max_{\hat{q}} A(\hat{q})$; analytically continue to $n\to 0$.

Difficult!

(b) Replica symmetric solution.

Replica permutation symmetry : $\hat{q} = \begin{pmatrix} 1 & & q \\ & \ddots & \\ q & & 1\end{pmatrix} = (1-q)\mathbb{1} + q\, v v^T$

$v = (1,\dots,1)$ vector of all ones

Then:

eigenvector	eigenvalue
v	$1+(n-1)q$
$v_\perp$	$1-q$

$$\Rightarrow \det q = (1-q)^{n-1}(1+(n-1)q)$$

$$\hat{q}^{-1} = \frac{1}{1-q}\mathbb{1} + \frac{q}{(1-q)[1+(n-1)q]}\, v v^T$$

$$a_{RS}(q) = \lim_{n\to 0}\partial_n\left[\frac{1}{2}(n-1)\log(1-q) + \frac{1}{2}\log[1+(n-1)q]\right] + \alpha\lim_{n\to 0}\partial_n \log I(\hat{q})$$

$$= \frac{1}{2}\log(1-q) + \frac{1}{2}\frac{q}{1-q} + \alpha\lim_{n\to 0}\partial_n\log I(\hat{q})$$

A slightly tedious computation gives

$$\lim_{n\to 0} \partial_n \log I(\hat q) = \int dh \, \frac{e^{-\frac{h^2}{2q}}}{\sqrt{2\pi q}} \log\left[\int dz \, \frac{e^{-\frac{z^2}{2(1-q)}}}{\sqrt{2\pi(1-q)}} \, e^{-\beta v(h-\sigma-z)}\right]$$

$$= \int dh \, \gamma_q(h) \log\left[\int \gamma_{1-q}(z) \, e^{-\beta v(h-\sigma-z)} dz\right]$$

$$a(q) = \frac{1}{2}\log(1-q) + \frac{1}{2}\frac{q}{1-q} + \alpha \int dh \, \gamma_q(h) \log\left[\int dz \, \gamma_{1-q}(z) \, e^{-\beta v(h-\sigma-z)}\right]$$

$$f = -T\, a(q^*) \qquad q^* \text{ such that } \frac{\partial a}{\partial q} = 0$$

[note: inversion of the max and the min]

Exercise: check at $\alpha = 0$

(3) The SAT-UNSAT transition.

We want to study what happens for $T \to 0$

(a) SAT phase.

q finite for $T \to 0$

$a(q)$ finite $\Rightarrow$ $f = -TS$, $S = a(q^*)$

energy $= 0$, entropy finite

$$\Theta(x) = \frac{1}{2}\left(1 + \mathrm{erf}(x)\right)$$

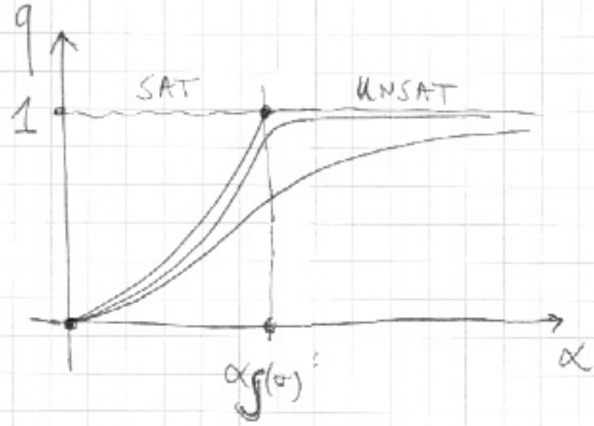
$$T = 0 \Rightarrow a(q) = \frac{1}{2}\log(1-q) + \frac{1}{2}\frac{q}{1-q} + \alpha \int dh \, \gamma_q(h) \log \Theta\left(\frac{h-\sigma}{\sqrt{2(1-q)}}\right)$$

$$\frac{\partial a}{\partial q} = 0 \Rightarrow \frac{1}{2}\frac{q}{(1-q)^2} + \alpha \frac{d}{dq} \int dh \, \gamma_q(h) \log \Theta\left(\frac{h-\sigma}{\sqrt{2(1-q)}}\right) = 0$$

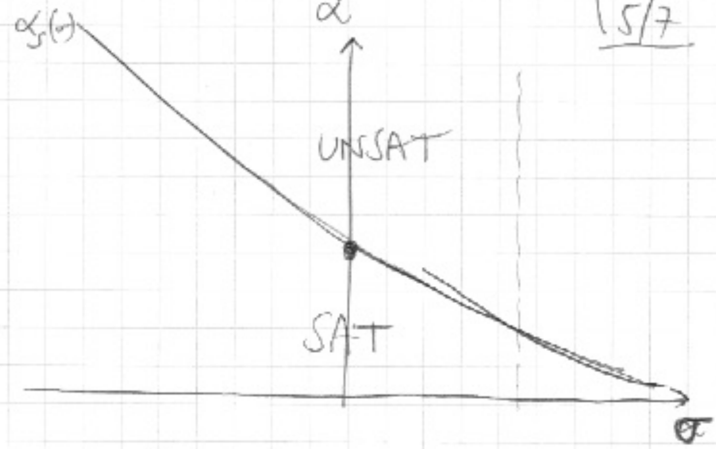
When we approach jamming, $q \to 1$ $\quad \Theta\left(\frac{h-\sigma}{\sqrt{2(1-q)}}\right) \sim \begin{cases} 1 & h > \sigma \\ e^{-\frac{(h-\sigma)^2}{2(1-q)}} & h < \sigma \end{cases}$

$$\frac{1}{2}\frac{1}{(1-q)^2} + \alpha \frac{d}{dq}\left[-\frac{1}{2(1-q)} \int_{-\infty}^{\sigma} dh \, \gamma_1(h) \, (h-\sigma)^2\right] = 0$$

$$1 - \alpha \int_{-\infty}^{\sigma} \mathcal{D}h \, (h-\sigma)^2 = 0 \Rightarrow \boxed{\alpha_J(\sigma) = 1 \Big/ \int_{-\infty}^{\sigma} \mathcal{D}h \, (h-\sigma)^2}$$



For fixed σ



Phase diagram at $T=0$

ⓑ <u>UNSAT phase</u> $\quad q \to 1$ for $T \to 0$, $\quad q = 1 - \chi T$ (harmonic expansion)

Choose a potential $\quad v(h) = \frac{1}{2} h^2 \,\theta(-h)$

$$\log \int dz\, \gamma_{1-q}(z)\, e^{-\beta v(h-\sigma-z)} \sim \begin{cases} -\dfrac{\beta (h-\sigma)^2}{2(1+\chi)} & h<\sigma \\ 0 & h>\sigma \end{cases}$$

$$f = -T\, a_{RS}(q) \xrightarrow[T\to 0]{} -\frac{1}{2\chi} + \frac{\alpha}{2}\int_{-\infty}^{\sigma} \mathcal{D}h\, \frac{(h-\sigma)^2}{1+\chi} = -\frac{1}{2\chi} + \frac{\alpha}{\alpha_J(\sigma)}\frac{1}{2(\chi+1)} = e(\chi)$$

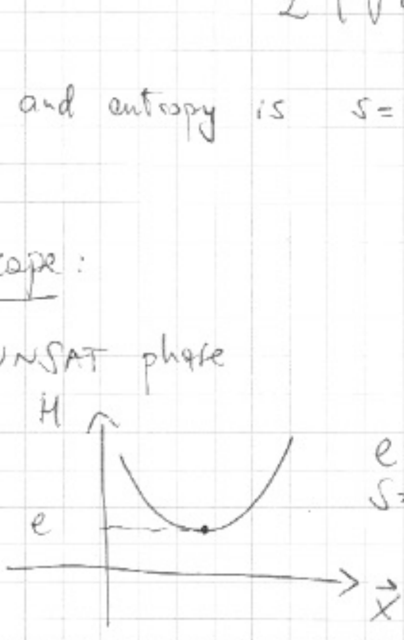
$$\frac{\partial f}{\partial \chi} = 0 \Rightarrow \frac{1}{2\chi^2} - \frac{\alpha}{\alpha_J(\sigma)}\frac{1}{2(\chi+1)^2} = 0 \Rightarrow \boxed{\left(1+\frac{1}{\chi}\right)^2 = \frac{\alpha}{\alpha_J(\sigma)}}$$

$$\Rightarrow e = \frac{1}{2}\left(\sqrt{\frac{\alpha}{\alpha_J(\sigma)}} - 1\right)^2 \propto \left[\alpha - \alpha_J(\sigma)\right]^2$$

and entropy is $S = -\infty$ everywhere in the SAT phase.

ⓒ <u>Landscape</u>:

UNSAT phase



$e > 0$
$S = -\infty$

SAT phase

$e = 0$
S finite

Summary $\boxed{\text{For } \sigma > 0}$

- Simplest RS calculation of a SAT-UNSAT transition.
- Analytic expression for $\alpha_J(\sigma)$: $\begin{cases} \alpha > \alpha_J(\sigma) & \text{no solution} \\ \alpha < \alpha_J(\sigma) & \text{finite entropy of solutions} \end{cases}$

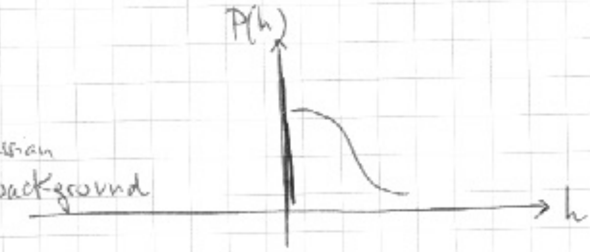
- Very simple landscape : always convex
- Energy $\sim \delta\alpha^2$ as in unjamming of soft spheres
- Entropy $\xrightarrow[\delta\alpha \to 0^-]{\delta\alpha \to 0^+} -\infty$ as in jamming of hard spheres.

But.......

(d) Isostaticity

we can compute the probability distribution of $h = R_\mu(\vec{x}) - \sigma$

at $\alpha = \alpha_J(\sigma)$

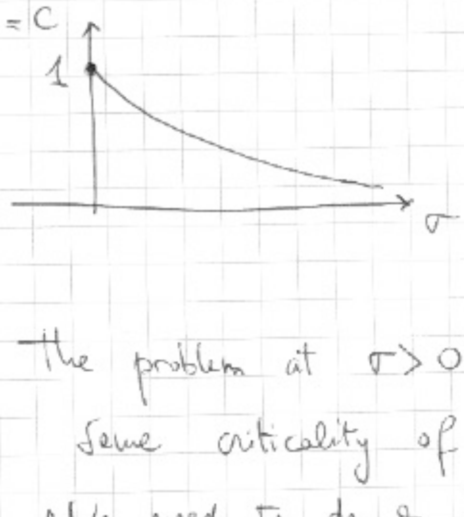


$P(h) = z\,\delta(h) + \underset{\text{background}}{\text{gaussian}}$

$$z = \int_{-\infty}^{0} \frac{dh}{\sqrt{2\pi}} e^{-\frac{(h+\sigma)^2}{2}} = \Theta\left(\frac{\sigma}{\sqrt{2}}\right)$$

Isostaticity : $Mz = N \Rightarrow \alpha z = 1$

$\alpha z = C$

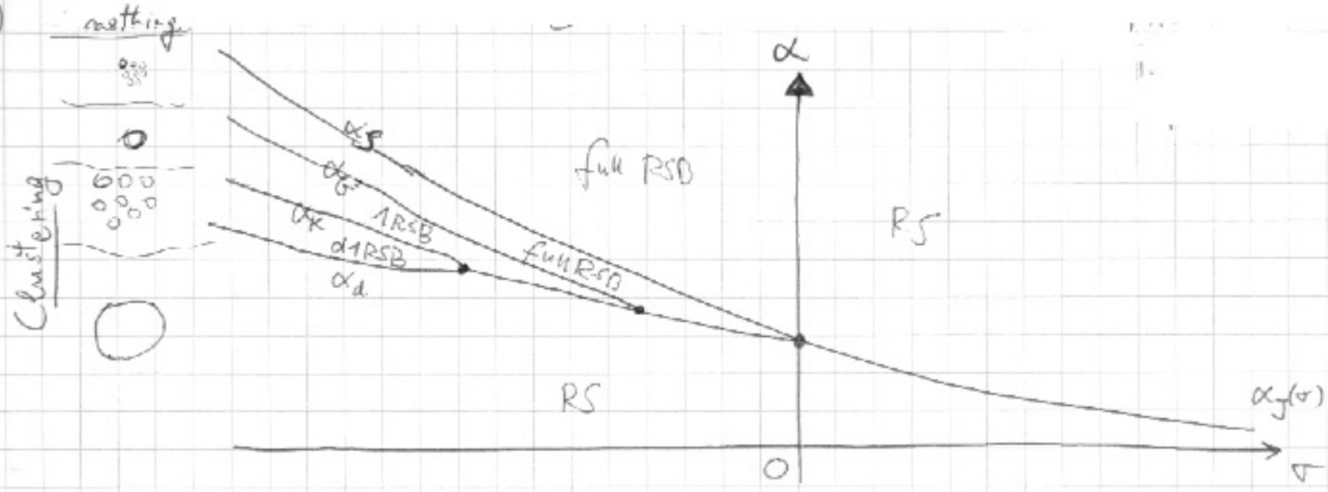


only isostatic at $\sigma = 0$!

The problem at $\sigma > 0$ does not have the same criticality of jamming (not isostatic)
We need to do a bit more work.

(4) The non convex perceptron and RSB

(a)



$$A(\hat{q}) = \frac{1}{2}\log\det\hat{q} + \alpha\log I(\hat{q}) \qquad q_{ab} = \langle \frac{1}{N}\,\vec{X}_a\cdot\vec{X}_b \rangle$$

$$f = -T\lim_{n\to 0}\max_{\hat{q}} A(\hat{q})$$

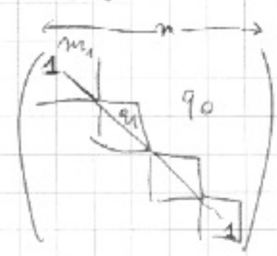
RS: $\hat{q} = (1-q)\mathbb{1} + q\,vv^T$

(closed algebra)

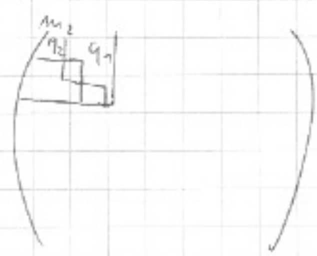
Not always stable: spontaneous breaking of replica symmetry.

{Castellani-Cavagna
FZ arXiv: 1008.4844

1RSB matrices



2RSB



etc.

⊛ Closed algebra at each step

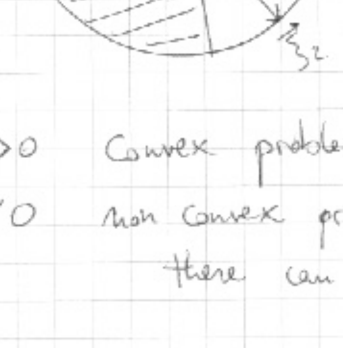
⊛ Can be analytically continued at $n\to 0$

⊛ $\overline{P(q)} = m_1\,\delta(q-q_0) + (1-m_1)\,\delta(q-q_1)$ at 1RSB level.

(b)

$\sigma > 0$ Convex, no RSB

$\sigma < 0$ non-convex, RSB



Geometric interpretation: $\sigma > 0$ Convex problem, no RSB

$\sigma < 0$ non convex problem, there can be RSB

Summary of ~~last~~ lecture 1

$R_\mu(\vec{x}) \qquad \phi_i^\mu \sim \mathcal{N}(0, 1/N)$

⊛ Perceptron: (as a random CSP) $\quad \exists \vec{x}$ s.t. $\vec{x}\cdot\vec{x} = N$ and $\overset{\|}{\vec{x}\cdot\vec{\phi}^\mu} \geq \sigma \quad \forall \mu = 1 \ldots M$

$N, M \to \infty \qquad \alpha = \frac{M}{N}$ $\qquad$ Control parameters (α, σ)

⊛ Hamiltonian $\qquad H(\vec{x}) = \sum_{\mu=1}^{M} v(R_\mu(\vec{x}) - \sigma)$ $\qquad \begin{cases} v(h) > 0, & h < 0 \\ v(h) = 0, & h > 0 \end{cases}$

$$Z = \int \mathcal{D}\vec{x}\, e^{-\beta H(\vec{x})} \qquad f = -\frac{T}{N}\overline{\log Z}$$

⊛ Replica calculation (general)

$$f = -T \lim_{n \to 0} \partial_n \Big[\max_{\hat{q}: \begin{pmatrix} q_{aa} = 1 \\ q_{ab} = q_{ba} \end{pmatrix}} A(\hat{q}) \Big] \qquad A(\hat{q}) = \frac{1}{2}\log\det\hat{q} + \alpha \log I(\hat{q})$$

$$I(\hat{q}) = \int dr_a \frac{e^{-\frac{1}{2}\sum_{ab} q^{-1}_{ab} r_a r_b}}{\sqrt{(2\pi)^n \det(q^{-1})}}\, e^{-\beta \sum_a v(r_a - \sigma)}$$

⊛ Replica symmetric $\quad \boxed{q_{ab} = q \quad \forall a \neq b}$

$$f = -T \min_q a_{RS}(q) \qquad a_{RS}(q) = \frac{1}{2}\log(1-q) + \frac{1}{2}\frac{q}{1-q} + \alpha \int dh\, \gamma_q(h) \log \int dz\, \gamma_{1-q}(z)\, e^{-\beta v(h+z\cdots}$$

$$q \in [0,1]$$

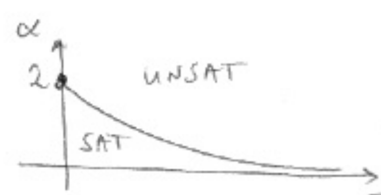
⊛ SAT phase: $\quad T \to 0 \qquad q < 1 \qquad a_{RS}(q) = \frac{1}{2}\log(1-q) + \frac{1}{2}\frac{q}{1-q} + \alpha \int dh\, \gamma_q(h) \log \Theta\left(\frac{h-\sigma}{\sqrt{2(1-q)}}\right)$

$e = 0$

$S = a_{RS}(q^*)$ finite $\qquad$ Indep. of potential $v(h)$

⊛ Approaching α_s from SAT phase: $\quad q \to 1$

$$\alpha_s(\sigma) = 1 \Big/ \int_{-\infty}^{\sigma} dh\, \gamma_1(h)\,(h-\sigma)^2$$



Restart from here

goo.gl/LXsCoH