Go back 1 page, to ①:

What is the coeff. of x_k^2 in ① ?

$$H(x_1, \dots, x_{2^n}) = n + \sum_{j=1}^{2^n} \sum_{i=1}^{\infty} a_{ij} x_j^i$$

Need: a_{2k}

$$\frac{\partial^2 H}{\partial x_k^2} = \left[2 a_{2k} + \text{terms involving } x_k \right]$$

$$\frac{1}{2} \cdot \left. \frac{\partial^2 H}{\partial x_k^2} \right|_{x_k = 0} = a_{2k}$$

$H(x_1, \dots, x_m)$, $\sum x_i \le 1$

$u = \left(\frac{1}{m+1}, \dots \frac{1}{m+1} \right)$

entropy of $\underline{x}$, expanded arou

$$d \|X\|_2^2 \le \underbrace{\log(\;)}_{\cancel{H(u)}} - \underbrace{H(\underline{X} - u)} \le c \cdot \|X\|_2^2$$

$$1 - \frac{(x - 1/2)^2}{\ln 2} \dots \le 1 - \frac{(x}{\ln}$$

$$\overset{?}{\geq} 1 - H(x - 1/2) \geq \frac{(x - 1/2)^2}{\ln 2}$$

9/8/10 <u>TAYLOR EXPANSIONS</u> FOR <u>h(x)</u>:

$$h(x) = -x \log x - (1-x) \log (1-x)$$

$$= 1 - \frac{2}{\ln 2}\left(x - \frac{1}{2}\right)^2 - \frac{4}{3 \ln 2}\left(x - \frac{1}{2}\right)^4 - \frac{32}{15 \ln 2}\left(x\right.$$

$$+ o\left(\left(x - \frac{1}{2}\right)^8 \right)$$

$$1 - h(x) = \frac{2}{\ln 2}\left(x - \frac{1}{2}\right)^2 + \frac{4}{3\ln 2}\left(x - \frac{1}{2}\right)^4 + \frac{32}{15\ln 2}\left(x - \frac{1}{2}\right)^6 + o\left(\left(x - \frac{1}{2}\right)^8\right)$$

$$\geq \frac{2}{\ln 2}\left(x - \frac{1}{2}\right)^2$$

Claim $\quad 1 - h(x) \leq c\left(x - \frac{1}{2}\right)^2$

$$\frac{\partial h(x)}{\partial x} = \cancel{-x\left(\frac{1}{x}\right)} - \log x \left(\frac{-1}{x^2}\right) - \cancel{(1-x)\frac{1}{(1-x)}(-1)}$$

$$- \log(1-x)\,(-1)$$

$$= \frac{\log x}{x^2} + \log(1-x)$$

$$\frac{\partial^2 h(x)}{\partial x^2} = \log x \,\frac{(-2)}{x^3} + \frac{1}{x^3} + \frac{1}{1-x}\,(-1)$$

$$= -\frac{2}{x^3}\log x + \frac{1}{x^3} - \frac{1}{1-x}$$

$$\frac{\partial^3 h(x)}{\partial x^3} = -\frac{2}{x^3}\,\frac{1}{x} - 2\log x\,\frac{(-3)}{x^4} - \frac{3}{x^4} + \frac{1}{(1-x)^2}\,($$

22/09/10 THOUGHTS ON SPIN-GLASS WORKS:

- These works give us upper bounds on entropy / MI

$$I(x;z) \leq \_\_\_\_$$

which in-turn give us lower bounds on

$$\Pr\{I(x;z) < n\,\mathrm{Cap}(\mathcal{C})\} \geq \_\_\_\_ \quad \begin{cases} 0 \\ 1 \end{cases}$$

Can we use these upper bounds to show the other side of our conjecture?

How sharp is the threshold obtained from these

24|09|10 GAUSSIAN WEIGHTS FOR HOPFIELD N/Ws :

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$$

p patterns
N neurons

If p is large, then $w_{ij} \to \mathcal{N}(\mu, \sigma^2)$
Suppose that ξ_i^μ and ξ_j^μ are picked iid Bernoulli(½
Then for $i \neq j$,

$$\mathbb{E}[w_{ij}] = \frac{1}{N} \sum_{m=1}^{p} \mathbb{E}[\xi_i^m] \, \mathbb{E}[\xi_j^m]$$

$$= \frac{p}{4N}$$

$$\mathbb{E}[w_{ij}^2] = \frac{1}{N^2} \sum_{m=1}^{p} \sum_{n=1}^{p} \mathbb{E}[\xi_i^m \xi_j^m \xi_i^n \xi_j^n]$$

$$= \frac{1}{N^2} \sum_{m=1}^{p} \mathbb{E}[(\xi_i^m)^2 (\xi_j^m)^2]$$

$$+ \frac{1}{N^2} \sum_{\substack{m,n=1 \\ m \neq n}}^{p} \mathbb{E}[\xi_i^m] \, \mathbb{E}[\xi_j^m] \, \mathbb{E}[\xi_i^n] \, \mathbb{E}[\xi_j^n$$

$$\mathbb{E}[(\xi_i^m)^2] = \frac{1}{2}(0) + \frac{1}{2}(1)^2 = \frac{1}{2}$$

$$\mathbb{E}[w_{ij}^2] = \frac{p}{4N^2} + \frac{p}{16N^2} = \frac{5p}{16N^2}$$

QN. Storing codewords instead of random patterns will increase correlation among weights (?). Will an inc in correlation increase the capacity of the hopfield

THOUGHT: Assume that the patterns stored are codewo corresponding to a (random?) linear code. WOLOG assume that all-zero pattern is fed as I/P. What is the condition for it to be stable?

→ In order for us to be able to store 2^k patterns "reliably" through this linear coding scheme, what is the 'n' that is required?

<u>Linear coding framework</u>:

$$p = 2^k, \quad N = n$$

Let v^l denote the output of a neuron at the l iteration. Assume that

$$\Pr\{ v^l$$

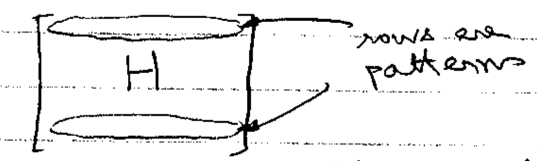
<u>THINGS TO WORK ON</u>:

① Can we compute $\mathbb{E}[w_{ij}]$ & $\mathbb{E}[w_{ij}^2]$ for familie of linear codes, whose codewords are used as patter. In order to do this, we need to compute the correlat between codeword bits of a linear code (asymptotically?) <u>One method</u>: use a systematic code, so that first k codeword bits are uncorrelated among each other. Then t to bound correlation with other bits combinatorially.

② Assume $N(\mu, \sigma^2)$ for the weights. Since we memorize codewords of a linear code, the all (-1) vector is a pattern. Assume that the Hopfield N/W is fed with the all (-1) pattern, with d errors. Can we write down the density evolution equations, given this initial conditi Try to determine the rate of the code for convergence the correct pattern. Can this Hopfield decoder correct er all the way till dmin of the code ?

IMPACT OF CORRELATION ON CAPACITY

Assume we use a Hadamard matrix to obta
our patterns.

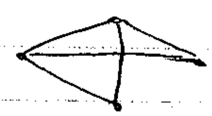


$H \in \{+1, -1\}$

Columns of H are orthogonal

$$\Rightarrow \quad w_{ij} = 0 \quad \forall \; i \neq j$$

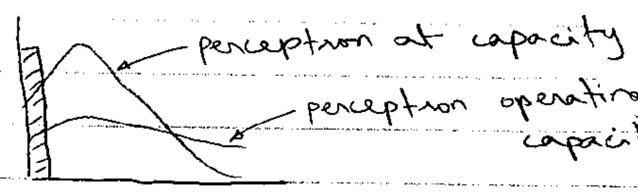
This tells us that if there are no correla
between pattern bits, recalling patterns is
possible



would like some strong & weak connections for recallin patterns, not all strong

Evidence that (pattern) correlation increases capacity:

By increasing correlation between ξ_i^m & ξ_j^m
$\mathbb{E}[w_{ij}]$ decreases, so does $\sigma^2[w_{ij}]$ (check th



$\alpha = $

This observation is consistent with results ob
in the perceptron literature
(Optimal storage & synaptic weights - Brunel,

CONTOUR PLOTS IN MAPLE:

To generate the contour plots, maple uses equal spacing on the z-axis (assume a 2-D fn. $f(x,y)$, wh values are plotted on the z-axis)

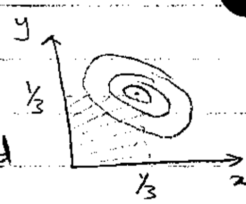
29/09/2010 TAYLOR SERIES FOR 3-D ENTROPY FUNCTION:

$$h(x,y) = -x\log_2 x - y\log_2 y - (1-x-y)\log_2(1-x-y)$$

Need a bound

$$2 - h(x,y) \le \_\_\_\_$$

From the Taylor series expansion of $2-h(x,y)$, c drop -ve terms to get a bound.



Assume first that both $x, y < 1/3$.
In this case, can drop all terms with odd degree from the Taylor series in entropy_taylor_2D.mw

Define $x' = x - 1/3$, $y' = y - 1/3$.

$$3x'y' + 9x'^3y' + \frac{27}{2}x'^2y'^2 + 9x'y'^3 + \frac{243}{5}x'^5y'$$

$$+ \frac{243}{2}x'^4y'^2 + 162x'^3y'^3 + \frac{243}{2}x'^2y'^4 + \frac{243}{5}x'y'^5 + \cdots$$

30/09/2010

HOPFIELD N/Ws - MATRIX UPDATES?

$$w_{ij} = \frac{1}{N}\sum_{m=1}^{p} \xi_i^m \xi_j^m, \quad i,j = 1,\dots,N, \; i \neq j$$

Let $v_i^{(t)}$ denote the state of the neuron at time t.
Update rule:

$$v_i^{(t+1)} = f\left(\sum_{j=1}^{N} w_{ij} v_j^{(t)}\right) \qquad (\text{Assume } w_i = w_j)$$

$$\begin{bmatrix} v_1^{(t+1)} \\ v_2^{(t+1)} \\ \vdots \\ v_N^{(t+1)} \end{bmatrix} = f\left(\begin{bmatrix} 0 & w_{12} & \cdots & w_{1N} \\ w_{21} & 0 & & \vdots \\ \vdots & & \ddots & \\ w_{N1} & \cdots & & 0 \end{bmatrix} \begin{bmatrix} v_1^{(t)} \\ v_2^{(t)} \\ \vdots \\ v_N^{(t)} \end{bmatrix} \right)$$

Assume that $w_{ij} \sim N(\mu, \sigma^2)$. Further, WOLOG that $v_1^{(t)}, \ldots, v_\ell^{(t)} = 1$, $v_{\ell+1}^{(t)}, v_{\ell+2}^{(t)}, \ldots, v_N^{(t)} = -1$.

$$\Pr\{v_1^{(t+1)} = 1\} = \Pr\{\underbrace{W_{12} + W_{13} + \cdots + w_{1\ell} - w_{1\,\ell+1} - \cdots - w_N}_{N(\mu(d-1-(N-d)),\ (N-1)\sigma^2) \ = N((2d-N-1)\mu,\ (N-1)\sigma^2)}$$

$$= \Pr\{N(0, (N-1)\sigma^2) > \mu(N-d-(d-1))\}$$

$$= \Pr\left\{N(0,1) > \frac{\mu(N-2d+1)}{\sqrt{N-1}\,\sigma}\right\}$$

$$= Q\left(\frac{\mu(N-2d+1)}{\sqrt{N-1}\,\sigma}\right) \qquad \{\text{Assume that } \mu(N+1-2d) > 0$$

w_{ij}s are correlated Gaussian RVs. Can we est[imate] the correlation between these RVs?

Patterns are codewords of a linear code (say LDG[M] with each col. [having] d ones)

$$\overbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 1 \\ & & \vdots & \\ 1 & 1 & \cdots & 1 \end{bmatrix}_{2^k \times k}}^{\triangleq A} \underset{k \times n}{G} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_{2^k} \end{bmatrix}_{2^k \times n} \triangleq \Xi$$

$$W = \frac{1}{N}(2\Xi - \underline{1})^T(2\Xi - \underline{1}), \quad \text{where } \underline{1} \rightarrow \text{all ones matri[x]} \ (2^k \times n \text{ in this [case]})$$

$$\underset{\substack{\uparrow \\ \text{real ops.}}}{A^T A} = \begin{bmatrix} 2^{k-1} & 2^{k-2} & 2^{k-2} \\ 2^{k-2} & 2^{k-1} & 2^{k-2} \\ & \vdots & \\ 2^{k-2} & 2^{k-2} & 2^{k-1} \end{bmatrix}_{k \times k} = 2^{k-2}\left\{ \underset{k}{I} + \begin{bmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \\ & \vdots & \\ 1 & \cdots & 1 \end{bmatrix}_{k \times k} \right\}$$