The Replica Method in Wireless Communications

— National University of Singapore —

March 2006

Ralf R. Müller

Professor for Wireless Networks
Norwegian University of Science & Technology
mueller@iet.ntnu.no

The Replica Method and Multistage Detection in Wireless Communications

— National University of Singapore —

March 2006

Ralf R. Müller

Professor for Wireless Networks
Norwegian University of Science & Technology
mueller@iet.ntnu.no

Part I:

The Replica Method

© *Ralf R. Müller 2005-2006*

Thermodynamics vs. Wireless Communications

Since Boltzmann, physicists have studied the behavior of systems with interactions of many particles.

Particular correspondences can be drawn between spin glasses (amorph magnetic materials, e.g. the magnetic surface of a hard disk drive) and communication systems. The binary nature of the bits corresponds to the quantum-mechanical constraints of electron spins to $\pm\frac{1}{2}$.

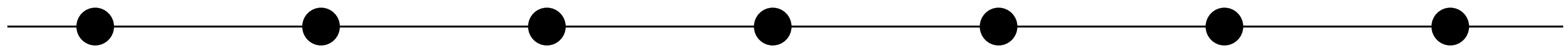
Spin glass theory is both very rich and very complicated. Various methods have been proposed by physicists to analyze them:

- The replica method
- The TAP approach
- The cavity method
- Gauge Theory
- ...

This course will be restricted in scope to the replica method.

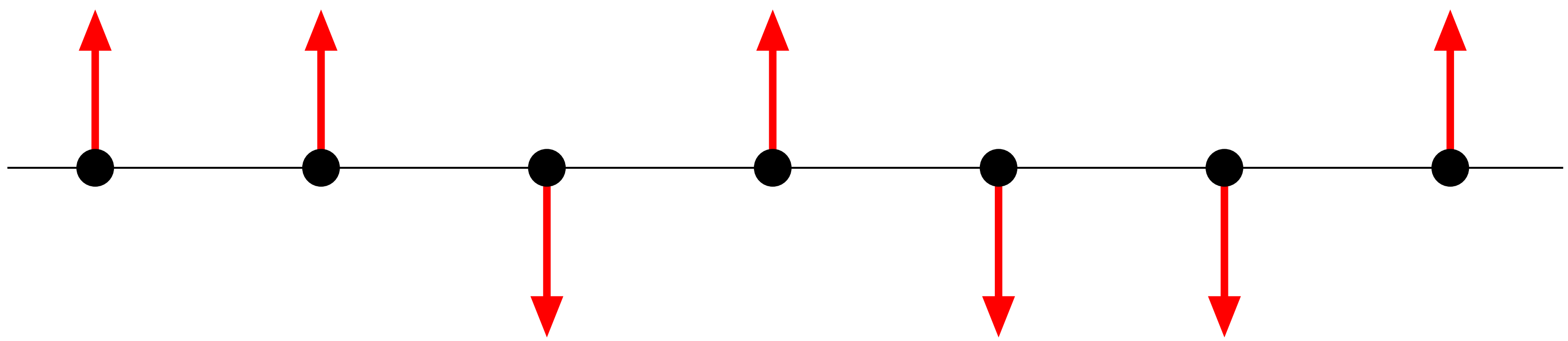
 © Ralf R. Müller 2005-2006

Spin Glasses

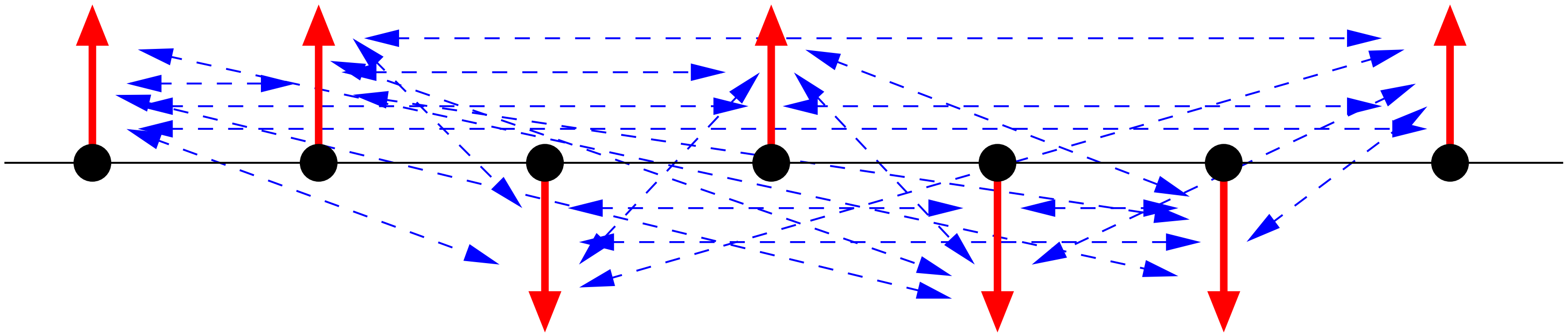


© *Ralf R. Müller 2005-2006*

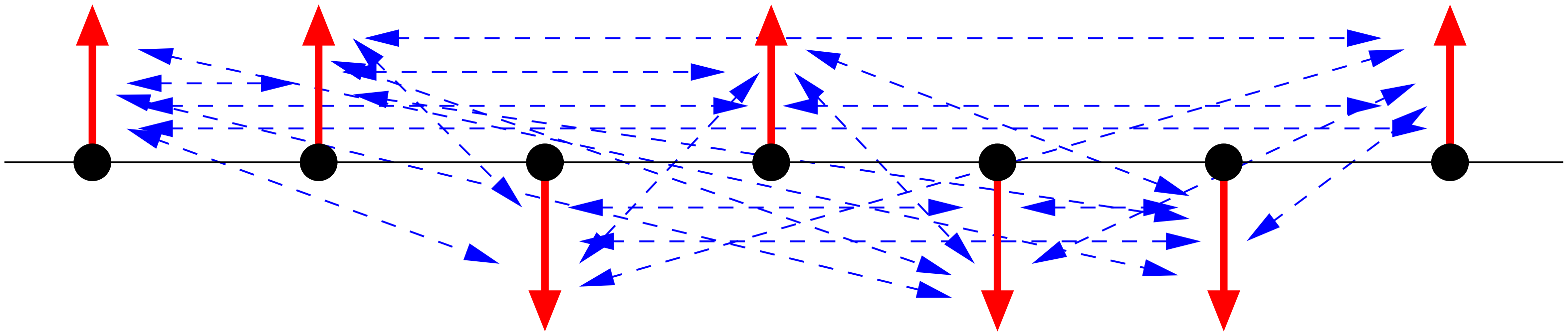
Spin Glasses



Spin Glasses



Spin Glasses

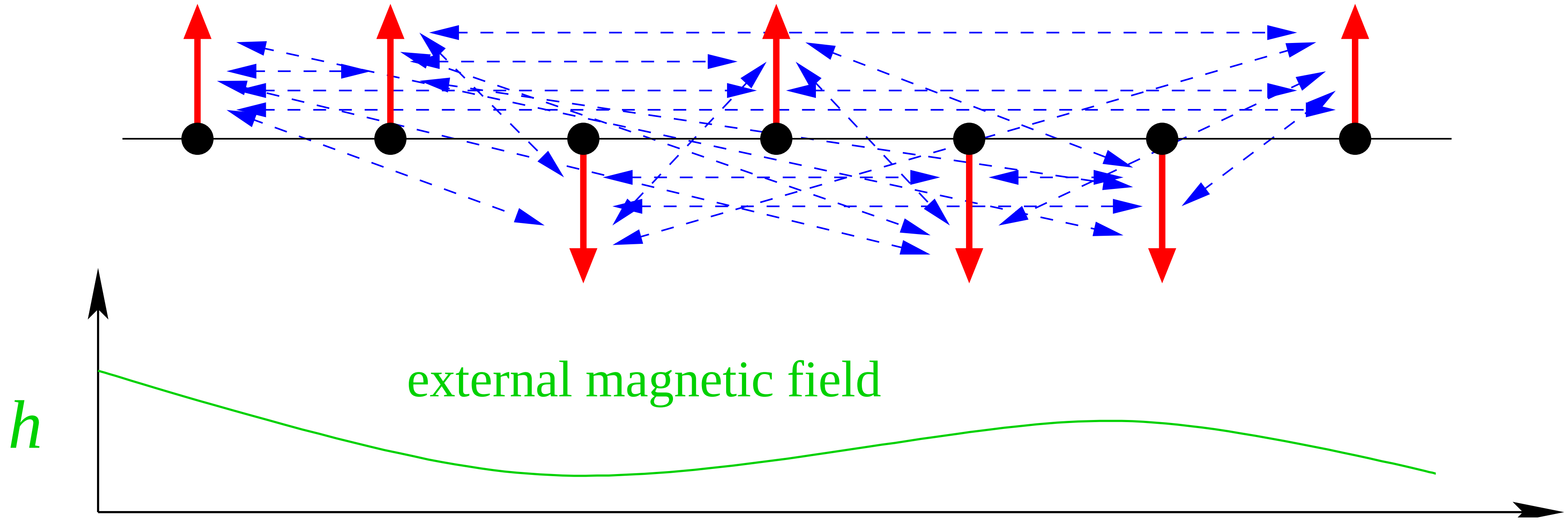


Energy function (Hamiltonian):

$$-\sum_i \sum_{j<i} r_{ij} x_i x_j$$

© *Ralf R. Müller 2005-2006*

Spin Glasses



Energy function (Hamiltonian):

$$-\sum_i \sum_{j<i} r_{ij} x_i x_j - \sum_i h_i x_i$$

Optimal Detection of Vector Channel

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{n}$$

© Ralf R. Müller 2005-2006

Optimal Detection of Vector Channel

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{n}$$

Best estimate for transmitted data:

$$\hat{\boldsymbol{x}} = \operatorname*{argmin}_{\boldsymbol{x} \in \{\pm 1\}^K} ||\boldsymbol{y} - \boldsymbol{S}\boldsymbol{x}||$$

© Ralf R. Müller 2005-2006

Optimal Detection of Vector Channel

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{n}$$

Best estimate for transmitted data:

$$\begin{aligned}\hat{\boldsymbol{x}} &= \operatorname*{argmin}_{\boldsymbol{x}\in\{\pm 1\}^K} ||\boldsymbol{y} - \boldsymbol{S}\boldsymbol{x}|| \\ &= \operatorname*{argmin}_{\boldsymbol{x}\in\{\pm 1\}^K} -\frac{1}{2}\boldsymbol{x}^\dagger \boldsymbol{R}\boldsymbol{x} - \boldsymbol{h}^\dagger \boldsymbol{x} + y^\dagger y\end{aligned}$$

with

$$\begin{aligned}\boldsymbol{R} &= -2\boldsymbol{S}^\dagger \boldsymbol{S} \\ \boldsymbol{h} &= 2\boldsymbol{S}^\dagger \boldsymbol{y}\end{aligned}$$

© Ralf R. Müller 2005-2006

Optimal Detection of Vector Channel

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{n}$$

Best estimate for transmitted data:

$$\begin{aligned}\hat{\boldsymbol{x}} &= \operatorname*{argmin}_{\boldsymbol{x}\in\{\pm 1\}^K} ||\boldsymbol{y} - \boldsymbol{S}\boldsymbol{x}|| \\ &= \operatorname*{argmin}_{\boldsymbol{x}\in\{\pm 1\}^K} -\frac{1}{2}\boldsymbol{x}^\dagger \boldsymbol{R}\boldsymbol{x} - \boldsymbol{h}^\dagger \boldsymbol{x} \qquad \text{with} \qquad \begin{aligned}\boldsymbol{R} &= -2\boldsymbol{S}^\dagger \boldsymbol{S} \\ \boldsymbol{h} &= 2\boldsymbol{S}^\dagger \boldsymbol{y}\end{aligned} \\ &= \operatorname*{argmin}_{\boldsymbol{x}\in\{\pm 1\}^K} -\sum_i \sum_{j<i} r_{ij} x_i x_j - \sum_i h_i x_i - \frac{1}{2}\sum_i r_{ii} x_i^2\end{aligned}$$

Optimal Detection of Vector Channel

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{n}$$

Best estimate for transmitted data:

$$\begin{aligned}\hat{\boldsymbol{x}} &= \operatorname*{argmin}_{\boldsymbol{x}\in\{\pm 1\}^K} ||\boldsymbol{y} - \boldsymbol{S}\boldsymbol{x}|| \\ &= \operatorname*{argmin}_{\boldsymbol{x}\in\{\pm 1\}^K} -\frac{1}{2}\boldsymbol{x}^\dagger \boldsymbol{R}\boldsymbol{x} - \boldsymbol{h}^\dagger \boldsymbol{x} \\ &= \operatorname*{argmin}_{\boldsymbol{x}\in\{\pm 1\}^K} -\sum_i \sum_{j<i} r_{ij} x_i x_j - \sum_i h_i x_i\end{aligned}$$

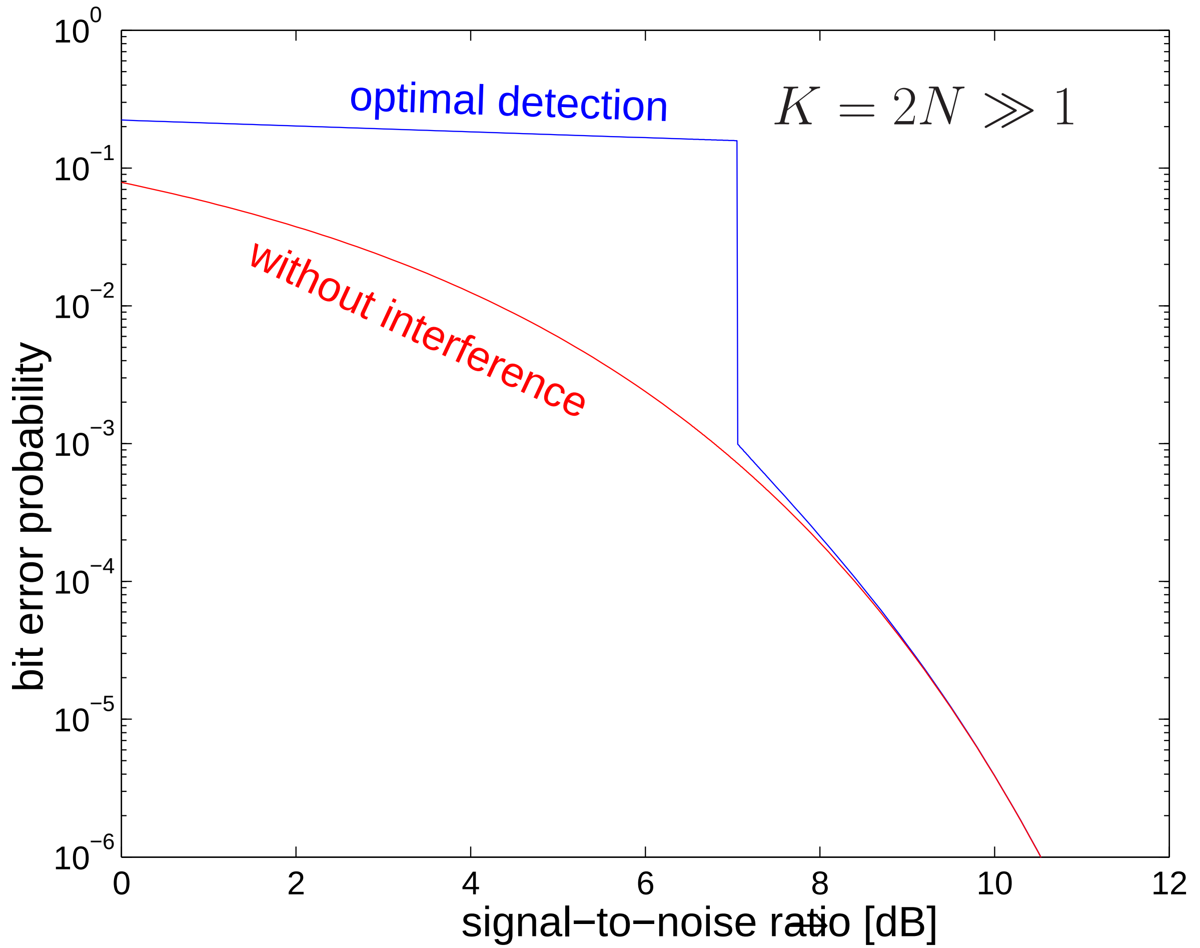
with

$$\begin{aligned}\boldsymbol{R} &= -2\boldsymbol{S}^\dagger \boldsymbol{S} \\ \boldsymbol{h} &= 2\boldsymbol{S}^\dagger \boldsymbol{y}\end{aligned}$$

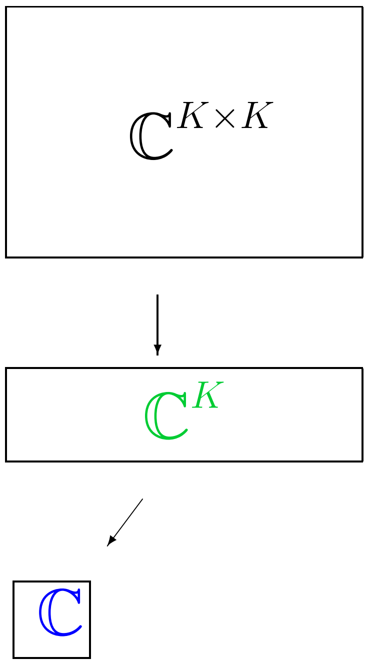
Minimization of the energy function of a spin glass!

 © Ralf R. Müller 2005-2006

A Phase Transition in Random CDMA



Large Systems



$\mathcal{O}\left(K^2\right)$ interactions

$\uparrow$

$\mathcal{O}(K)$ microscopic objects

$\downarrow$

$\mathcal{O}(1)$ macroscopic variables

Macroscopic variables are self-averaging.

© *Ralf R. Müller 2005-2006*

Boltzmann Distribution

The Thermodynamic Equilibrium maximizes the entropy

$$H(X) = -\sum_i \Pr(x_i) \log \Pr(x_i)$$

for given constant energy

$$E(X) = \sum_i ||x_i|| \Pr(x_i)$$

yielding the *Boltzmann distribution*

$$\Pr(x_i) = \frac{e^{-\frac{1}{T}||x_i||}}{\sum_i e^{-\frac{1}{T}||x_i||}}.$$

© Ralf R. Müller 2005-2006

Free Energy

Since the energy is constant, we can minimize the free energy

$$F(X) \stackrel{\triangle}{=} E(X) - TH(X)$$

instead of maximizing entropy. This is often less complicated.

With the Boltzmann distribution, the free energy is given by

$$F(X) = -T \log \left[\sum_i e^{-\frac{1}{T}||x_i||} \right].$$

It depends only on the partition function.

The free energy is self-averaging.

 © Ralf R. Müller 2005-2006

Energy vs. Entropy

The following two tasks are dual:

- Minimize the energy for fixed entropy
- Maximize the entropy for fixed energy

Consider free energy

$$F(X) = E(X) - TH(X)$$

and read the temperature (or its inverse) as Lagrange multiplier.

For the dual problem have

$$-\frac{1}{T}F(X) = H(X) - \frac{1}{T}E(X)$$

© Ralf R. Müller 2005-2006

The Meaning of the Energy Function

In physics, the energy function varies with the force causing the potential.

Theoretically speaking, the choice of the energy function is arbitrary as long as it is uniformly bounded from below.

Nature maximizes entropy for a given energy.

In communications engineering, the energy function is the metric used by the decoder.

The decoder does the dual job of nature, to minimize the metric for a given output entropy.

Since the decoder dictates the thermodynamics of our toy universe, the same holds true if the decoder uses a suboptimal (wrong) or insufficient metric, perhaps due to lack of knowledge about the channel state.

The free choice of the energy function allows to analyze mismatched receivers.

 © *Ralf R. Müller 2005-2006*

LMMSE Detector with Mismatched Powers

Theorem: *Let $(U_1, \ldots, U_K)$ be an arbitrary sequence of non-negative numbers such that, as $K \to \infty$, the empirical joint cdf of the pairs $\{(U_k, P_k) : k = 1, \ldots, K\}$ converges weakly to a given non-random cdf $F(u, p)$. Moreover, let the P_ks are uniformly bounded above and the U_ks are uniformly bounded below by a positive number for all K. Then, the output SINR of the mismatched LMMSE detector assuming powers $\{U_k\}$ instead of the true powers $\{P_k\}$ in the standard random spreading model converges as $K = \beta N \to \infty$ almost surely to*

$$\eta P_k \frac{1 + \beta \int \frac{u}{(1 + u\eta)^2} \, dF(u, p)}{1 + \beta \int \frac{p}{(1 + u\eta)^2} \, dF(u, p)} \quad \text{where} \quad \eta = \left(1 + \beta \int \frac{u}{1 + u\eta} \, dF(u, p) \right)^{-1}$$

is the multiuser efficiency of an LMMSE detector of a "virtual channel" having powers given by $\{U_k\}$ instead of $\{P_k\}$.

© Ralf R. Müller 2005-2006

Average Free Energy

When analyzing a random system, we evaluate the average free energy

© Ralf R. Müller 2005-2006

Average Free Energy

When analyzing a random system, we evaluate the average free energy

as Shannon analyzed the average performance of all codes.

© Ralf R. Müller 2005-2006

Free Energy for a Random Parameter

Consider a self-averaging random parameter, e.g. a spreading matrix.

$$F(X|y_j) = \mathop{\mathrm{E}}_{Y} F(X|Y)$$

© *Ralf R. Müller 2005-2006*

Free Energy for a Random Parameter

Consider a self-averaging random parameter, e.g. a spreading matrix.

$$
\begin{aligned}
F(X|y_j) &= \operatorname*{E}_{Y} F(X|Y) \\
&= -T \operatorname*{E}_{Y} \log \left[\sum_i \mathrm{e}^{-\frac{1}{T}||x_i||} \right]
\end{aligned}
$$

The energy function depends on the random parameter y_j.

The expectation of a logarithm is a hard problem.

© *Ralf R. Müller 2005-2006*

Replica Continuity

$$\mathop{\mathrm{E}}_{Y} \log(Y) = \lim_{n \to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} Y^n$$

Evalute n^{th} moments for integer n and assume analytic continuity for the limit.

 © *Ralf R. Müller 2005-2006*

Replica Continuity

$$\operatorname*{E}_{Y} \log(Y) = \lim_{n\to 0} \frac{\partial}{\partial n} \operatorname*{E}_{Y} Y^n$$

Evalute n^{th} moments for integer n and assume analytic continuity for the limit.

More general, we have

$$\operatorname*{E}_{Y} \log \int_{\mathbb{R}} f(x,Y)dx = \lim_{n\to 0} \frac{\partial}{\partial n} \operatorname*{E}_{Y} \left[\int_{\mathbb{R}} f(x,Y)dx \right]^n$$

© *Ralf R. Müller 2005-2006*

Replica Continuity

$$\mathop{\mathrm{E}}_{Y} \log \int_{\mathbb{R}} f(x,Y)dx = \lim_{n\to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} \left[\int_{\mathbb{R}} f(x,Y)dx \right]^n$$

Replica Continuity

$$\mathop{\mathrm{E}}_{Y} \log \int_{\mathbb{R}} f(x,Y)dx = \lim_{n\to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} \left[\int_{\mathbb{R}} f(x,Y)dx \right]^n$$

With

$$\left(\int_{\mathbb{R}} g(x)dx \right)^n = \prod_{a=1}^{n} \int_{\mathbb{R}} g(x_a)dx_a$$

© Ralf R. Müller 2005-2006

Replica Continuity

$$\mathop{\mathrm{E}}_{Y} \log \int_{\mathbb{R}} f(x,Y)dx = \lim_{n\to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} \left[\int_{\mathbb{R}} f(x,Y)dx \right]^n$$

With

$$\left(\int_{\mathbb{R}} g(x)dx \right)^n = \prod_{a=1}^{n} \int_{\mathbb{R}} g(x_a)dx_a$$

we finally get

$$\mathop{\mathrm{E}}_{Y} \log \int_{\mathbb{R}} f(x,Y)dx = \lim_{n\to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} \prod_{a=1}^{n} \int_{\mathbb{R}} f(x_a,Y)dx_a$$

© Ralf R. Müller 2005-2006

Replica Symmetry

Throughout the calaculations, we solve integrals of the form

$$I = \frac{1}{K} \log \int_{\mathbb{R}^2} \mathrm{e}^{Kf(x_1,x_2)} dx_1 dx_2$$

for $K \to \infty$ by *saddle point integration*.

© *Ralf R. Müller 2005-2006*

Replica Symmetry

Throughout the calaculations, we solve integrals of the form

$$I = \frac{1}{K} \log \int_{\mathbb{R}^2} \mathrm{e}^{K f(x_1, x_2)} dx_1 dx_2 \rightarrow \max_{x_1, x_2} f(x_1, x_2)$$

for $K \rightarrow \infty$ by *saddle point integration*.

© *Ralf R. Müller 2005-2006*

Replica Symmetry

Throughout the calculations, we solve integrals of the form

$$I = \frac{1}{K} \log \int_{\mathbb{R}^2} \mathrm{e}^{Kf(x_1,x_2)} dx_1 dx_2 \to \max_{x_1,x_2} f(x_1,x_2)$$

for $K \to \infty$ by *saddle point integration*.

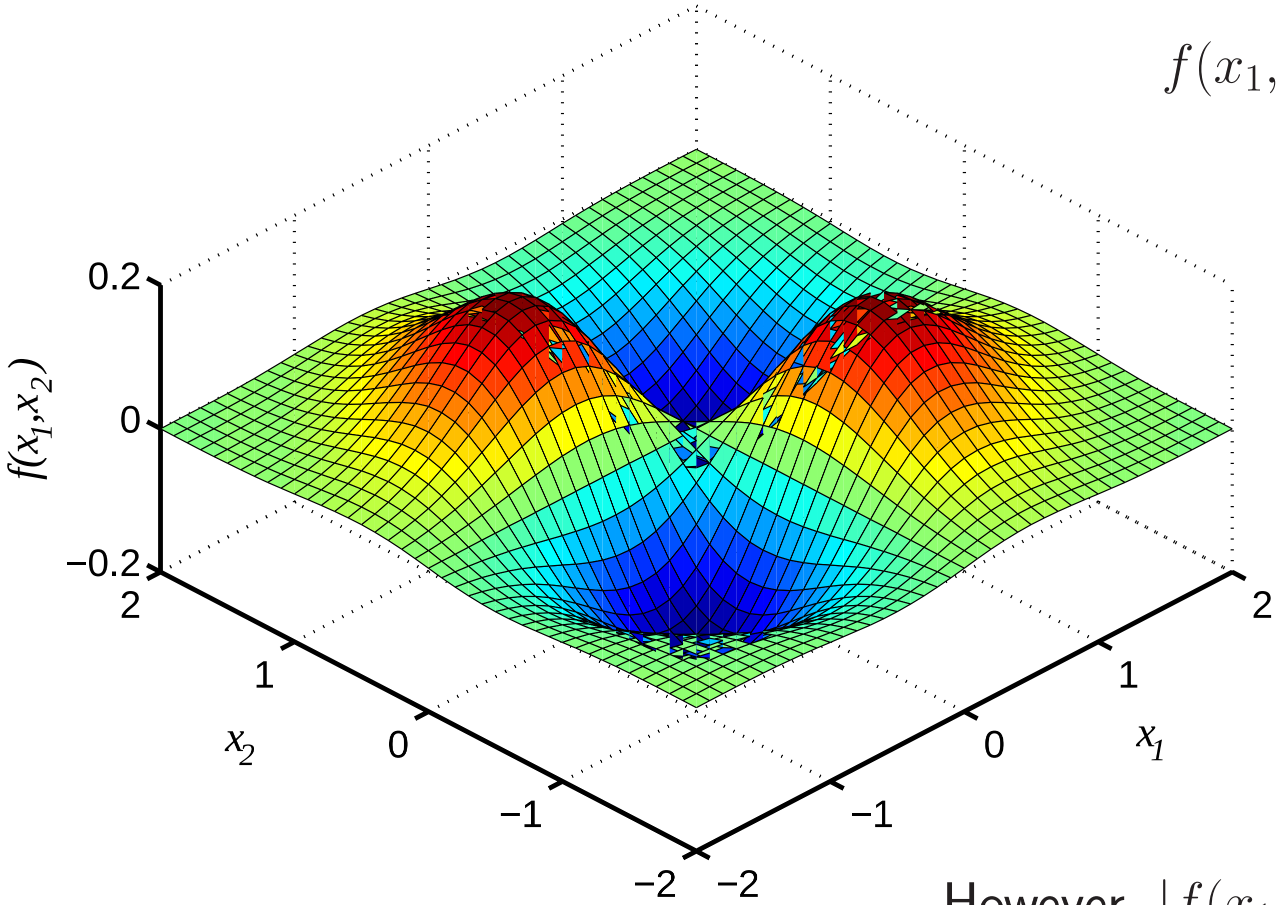
If the maximization is too tedious, we assume *replica symmetry*:

$$\max_{x_1,x_2} f(x_1,x_2) = \max_{x} f(x,x)$$

Replica symmetry is a strong assumption and not always valid.

© Ralf R. Müller 2005-2006

A Counterexample to Replica Symmetry



$$f(x_1, x_2) = -\sin(x_1 x_2) \mathrm{e}^{-x_1^2 - x_2^2}$$

However, $|f(x_1, x_2)|$ is replica symmetric.

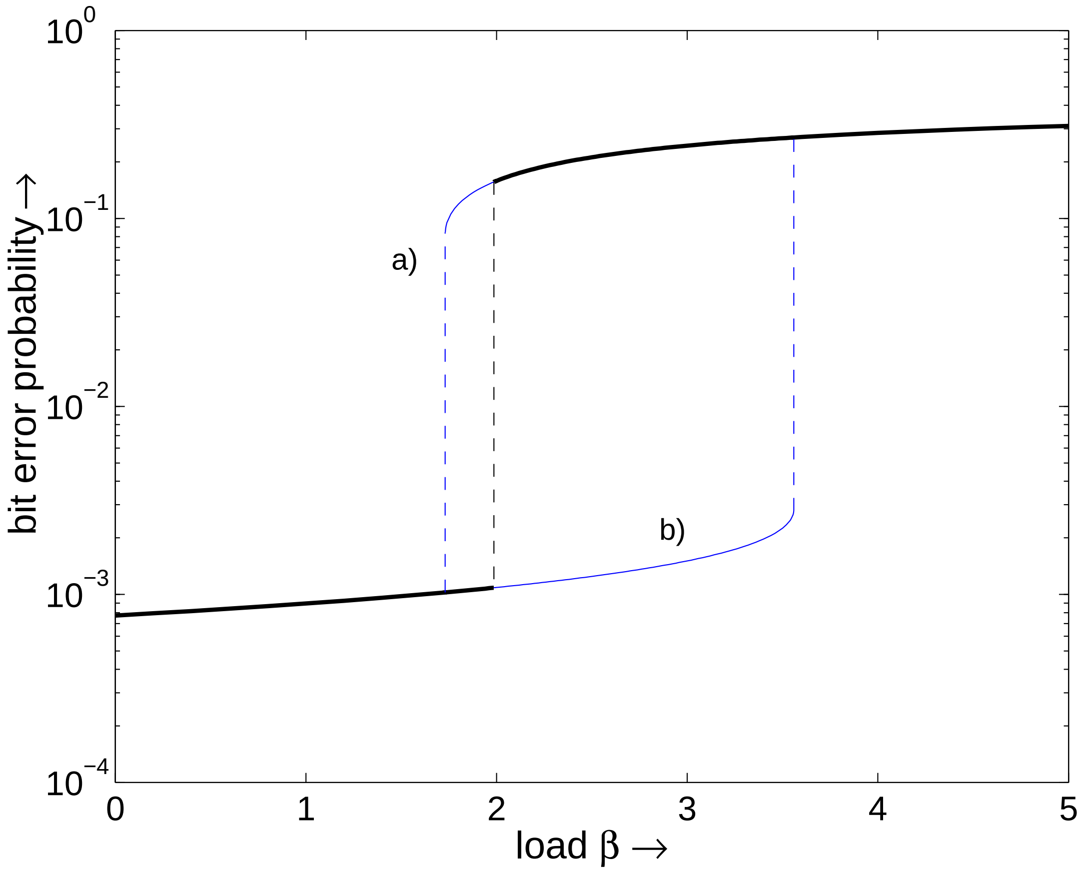
© *Ralf R. Müller 2005-2006*

Phase Transitions

If the final equations allow for multiple solutions, the correct solution is identified by minimizing the free energy.

 © *Ralf R. Müller 2005-2006*

Phase Transitions and Neural Networks



Individually Optimum Maximum Likelihood Detector

Let $\mathcal{A} = \{+1; -1\}$, the chips of any user be i.i.d. random variables with finite variance and vanishing odd moments, the powers of all users identical, and $N, K \to \infty$, but $\beta = K/N$ fixed. Then, the multiuser efficiency is a solution to the fixed point equation

$$\frac{1}{\eta_{\mathrm{IO}}} = 1 + \frac{\beta}{\sigma_n^2}\left[1 - \sqrt{\frac{\eta_{\mathrm{IO}}}{2\pi\sigma_n^2}} \int_{\mathbb{R}} \tanh\left(\frac{\eta_{\mathrm{IO}}}{\sigma_n^2}x\right)\exp\left(-\frac{\eta_{\mathrm{IO}}(x-1)^2}{2\sigma_n^2}\right)\mathrm{d}x\right].$$

In case the fixed point equation has multiple solutions, the correct one is that solution for which the term

$$\frac{\eta_{\mathrm{IO}}}{\sigma_n^2} + \frac{\eta_{\mathrm{IO}} - \log\eta_{\mathrm{IO}}}{2\beta} - \sqrt{\frac{\eta_{\mathrm{IO}}}{2\pi\sigma_n^2}}\int_{\mathbb{R}} \log\left[\cosh\left(\frac{\eta_{\mathrm{IO}}}{\sigma_n^2}x\right)\right]\exp\left(-\frac{\eta_{\mathrm{IO}}(x-1)^2}{2\sigma_n^2}\right)\mathrm{d}x$$

is smallest.

Example for Replica Calculations

Consider the analysis of an asymptotically large CDMA systems with arbitrary joint distribution of the variances of the random chips. It includes the practically imporant case of multi-carrier CDMA transmission with users of arbitrary powers in frequency-selective fading as a special case.

The vector-valued, real additive white Gaussian noise channel is characterized by the conditional pdf

$$\mathrm{p}_{\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H}}(\boldsymbol{y},\boldsymbol{x},\boldsymbol{H}) = \frac{\mathrm{e}^{-\frac{1}{2\sigma_0^2}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})}}{(2\pi\sigma_0^2)^{\frac{N}{2}}}. \tag{1}$$

Moreover, let the detector be characterized by the assumed conditional probability distribution

$$\breve{\mathrm{p}}_{\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H}}(\boldsymbol{y},\boldsymbol{x},\boldsymbol{H}) = \frac{\mathrm{e}^{-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})}}{(2\pi\sigma^2)^{\frac{N}{2}}}$$

and the assumed prior distribution $\breve{\mathrm{p}}_{\boldsymbol{x}}(\boldsymbol{x})$.

 © Ralf R. Müller 2005-2006

Let the entries of $\boldsymbol{H}$ be independent zero-mean with vanishing odd order moments and variances w_{ck}^2/N for row c and column k.

Applying Bayes' law, we find

$$\breve{\mathrm{p}}_{\boldsymbol{x}|\boldsymbol{y},\boldsymbol{H}}\left(\boldsymbol{x},\boldsymbol{y},\boldsymbol{H}\right) = \frac{\mathrm{e}^{-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})+\log \breve{\mathrm{p}}_{\boldsymbol{x}}(\boldsymbol{x})}}{\int \mathrm{e}^{-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})}\, \mathrm{d}\breve{\mathrm{P}}_{\boldsymbol{x}}\left(\boldsymbol{x}\right)}.$$

Since the Boltzmann distribution holds for any temperature T, we set w.l.o.g. $T = 1$ and find the appropriate energy function to be

$$||\boldsymbol{x}|| = \frac{1}{2\sigma^2}\left(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\right)^{\mathrm{T}}\left(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\right) - \log \breve{\mathrm{p}}_{\boldsymbol{x}}\left(\boldsymbol{x}\right). \tag{2}$$

This choice of the energy function ensures that the thermodynamic equilibrium models the detector defined by the assumed conditional and prior distributions.

© Ralf R. Müller 2005-2006

With (1) and (2), the definition of free energy, and replica continuity, we find for the free energy per user

$$\begin{aligned}
\frac{\mathrm{F}(\boldsymbol{x})}{K} &= -\frac{1}{K}\mathop{\mathrm{E}}_{\boldsymbol{H}}\int\int_{\mathbb{R}^N}\log\left(\int \mathrm{e}^{-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})}\mathrm{d}\breve{\mathrm{P}}_{\boldsymbol{x}}(\boldsymbol{x})\right)\frac{\mathrm{e}^{-\frac{1}{2\sigma_0^2}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})}}{(2\pi\sigma_0^2)^{\frac{N}{2}}}\mathrm{d}\boldsymbol{y}\mathrm{d}\mathrm{P}_{\boldsymbol{x}}(\boldsymbol{x}) \\
&= -\frac{1}{K}\lim_{n\to 0}\frac{\partial}{\partial n}\log\mathop{\mathrm{E}}_{\boldsymbol{H}}\int\int_{\mathbb{R}^N}\left(\int \mathrm{e}^{-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})}\mathrm{d}\breve{\mathrm{P}}_{\boldsymbol{x}}(\boldsymbol{x})\right)^n \\
&\qquad \frac{\mathrm{e}^{-\frac{1}{2\sigma_0^2}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})}}{(2\pi\sigma_0^2)^{\frac{N}{2}}}\mathrm{d}\boldsymbol{y}\mathrm{d}\mathrm{P}_{\boldsymbol{x}}(\boldsymbol{x}) \\
&= -\frac{1}{K}\lim_{n\to 0}\frac{\partial}{\partial n}\log\underbrace{\int\frac{\int_{\mathbb{R}^N}\mathop{\mathrm{E}}_{\boldsymbol{H}}\prod_{a=0}^{n}\mathrm{e}^{-\frac{1}{2\sigma_a^2}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}_{\boldsymbol{a}})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}_{\boldsymbol{a}})}\mathrm{d}\boldsymbol{y}}{(2\pi\sigma_0^2)^{\frac{N}{2}}}\prod_{a=0}^{n}\mathrm{d}\mathrm{P}_a(\boldsymbol{x}_a)}_{\triangleq\Xi_n} \qquad (3)
\end{aligned}$$

with $\sigma_a = \sigma, \forall a \geq 1$, $\mathrm{P}_0(\boldsymbol{x}) = \mathrm{P}_{\boldsymbol{x}}(\boldsymbol{x})$, and $\mathrm{P}_a(\boldsymbol{x}) = \breve{\mathrm{P}}_{\boldsymbol{x}}(\boldsymbol{x}), \forall a \geq 1$.

© Ralf R. Müller 2005-2006

The integral in (3) is given by

$$\Xi_n = \int \prod_{c=1}^{N} \frac{\int\limits_{\mathbb{R}} \mathop{\mathrm{E}}_{\boldsymbol{H}} \prod_{a=0}^{n} \exp\left[-\frac{1}{2\sigma_a^2} \left(y_c - \sum_{k=1}^{K} h_{ck} x_{ak} \right)^2 \right] \mathrm{d}y_c}{\sqrt{2\pi}\sigma_0} \prod_{a=0}^{n} \mathrm{dP}_a\left(\boldsymbol{x}_a\right), \qquad (4)$$

with y_c, x_{ak}, and h_{ck} denoting the c^{th} component of $\boldsymbol{y}$, the k^{th} component of $\boldsymbol{x}_a$, and the $(c,k)^{\mathrm{th}}$ entry of $\boldsymbol{H}$, respectively. The integrand depends on $\boldsymbol{x}_a$ only through

$$v_{ac} \overset{\triangle}{=} \frac{1}{\sqrt{\beta}} \sum_{k=1}^{K} h_{ck} x_{ak}, \qquad a = 0, \ldots, n.$$

These quantities v_{ac} can be regarded, in the limit $K \to \infty$ as jointly Gaussian random variables with zero mean and covariances

$$Q_{ab}[c] = \mathop{\mathrm{E}}_{\boldsymbol{H}} v_{ac} v_{bc} = \frac{1}{K} \boldsymbol{x}_a \overset{(c)}{\bullet} \boldsymbol{x}_b \qquad (5)$$

where the parametric inner products are defined by $\boldsymbol{x}_a \overset{(c)}{\bullet} \boldsymbol{x}_b \overset{\triangle}{=} \sum_{k=1}^{K} x_{ak} x_{bk} w_{ck}^2$.

In order to perform the integration in (4), the $K(n+1)$-dimensional space spanned by the replicas and the vector $\boldsymbol{x}_0$ is split into subshells

$$S\{Q[\cdot]\} \triangleq \left\{ \boldsymbol{x}_0, \ldots, \boldsymbol{x}_n \middle| \boldsymbol{x}_a \overset{(c)}{\bullet} \boldsymbol{x}_b = KQ_{ab}[c] \right\}$$

where the inner product of two different vectors $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$ is constant.[1]

The splitting of the $K(n+1)$-dimensional space is depending on the chip time c. With this splitting of the space, we find[2]

$$\Xi_n = \int\limits_{\mathbb{R}^{N(n+1)(n+2)/2}} \mathrm{e}^{K\mathcal{I}\{Q[\cdot]\}} \prod_{c=1}^{N} \mathrm{e}^{\mathcal{G}\{Q[c]\}} \prod_{a \le b} \mathrm{d}Q_{ab}[c], \qquad (6)$$

with appropriate choices of the function $\mathcal{I}\{Q[\cdot]\}$ and $\mathcal{G}\{Q[c]\}$.

[1] The notation $f\{Q[\cdot]\}$ expresses the dependency of the function $f(\cdot)$ on all $Q_{ab}[c], 0 \le a \le b \le n, 1 \le c \le N$.
[2] The notation $\prod_{a \le b}$ is used as shortcut for $\prod_{a=0}^{n} \prod_{b=a}^{n}$.

In (6),

$$e^{K\mathcal{I}\{Q[\cdot]\}} = \int \left[\prod_{a\le b}\prod_{c=1}^{N} \delta\left(\frac{\boldsymbol{x}_a \overset{(c)}{\bullet} \boldsymbol{x}_b}{N} - \beta Q_{ab}[c]\right)\right] \prod_{a=0}^{n} \mathrm{dP}_a(\boldsymbol{x}_a)$$

denotes the probability weight of the subshell and

$$e^{\mathcal{G}\{Q[c]\}} = \frac{1}{\sqrt{2\pi}\sigma_0} \int_{\mathbb{R}} \mathop{\mathrm{E}}_{\boldsymbol{H}} \prod_{a=0}^{n} \exp\left[-\frac{\beta}{2\sigma_a^2}\left(\frac{y_c}{\sqrt{\beta}} - v_{ac}\{Q[c]\}\right)^2\right] \mathrm{d}y_c + \mathcal{O}(K^{-1}).$$

This procedure is a change of integration variables in multiple dimensions where the integration of an exponential function over the replicas has been replaced by integration over the variables $Q_{ab}[\cdot]$. In the following the blue and green terms in (6) are evaluated separately.

© *Ralf R. Müller 2005-2006*

First, we calculate the measure $\mathrm{e}^{K\mathcal{I}\{Q[\cdot]\}}$. As for some $t \in \mathbb{R}$, we have the Fourier expansion of the Dirac measure

$$\delta\left(\frac{\boldsymbol{x}_a \overset{(c)}{\bullet} \boldsymbol{x}_b}{N} - \beta Q_{ab}[c]\right) = \frac{1}{2\pi \mathrm{j}} \int_{\mathcal{J}} \exp\left[\tilde{Q}_{ab}[c]\left(\frac{\boldsymbol{x}_a \overset{(c)}{\bullet} \boldsymbol{x}_b}{N} - \beta Q_{ab}[c]\right)\right] \mathrm{d}\tilde{Q}_{ab}[c]$$

with $\mathcal{J} = (t - \mathrm{j}\infty; t + \mathrm{j}\infty)$, the measure $\mathrm{e}^{K\mathcal{I}\{Q[\cdot]\}}$ can be expressed as

$$\mathrm{e}^{K\mathcal{I}\{Q[\cdot]\}} = \int \left[\prod_{c=1}^{N} \prod_{a \le b} \int_{\mathcal{J}} \mathrm{e}^{\tilde{Q}_{ab}[c]\left(\frac{\boldsymbol{x}_a \overset{(c)}{\bullet} \boldsymbol{x}_b}{N} - \beta Q_{ab}[c]\right)} \frac{\mathrm{d}\tilde{Q}_{ab}[c]}{2\pi \mathrm{j}}\right] \prod_{a=0}^{n} \mathrm{dP}_a(\boldsymbol{x}_a)$$

$$= \int_{\mathcal{J}^{N(n+2)(n+1)/2}} \mathrm{e}^{-\beta \sum\limits_{c=1}^{N} \sum\limits_{a \le b} \tilde{Q}_{ab}[c] Q_{ab}[c]} \left(\prod_{k=1}^{K} M_k\left\{\tilde{Q}[\cdot]\right\}\right) \prod_{c=1}^{N} \prod_{a \le b} \frac{\mathrm{d}\tilde{Q}_{ab}[c]}{2\pi \mathrm{j}} \qquad (7)$$

with

$$M_k\left\{\tilde{Q}[\cdot]\right\} = \int \exp\left(\frac{1}{N} \sum_{a \le b} \sum_{c=1}^{N} \tilde{Q}_{ab}[c] x_{ak} x_{bk} w_{ck}^2\right) \prod_{a=0}^{n} \mathrm{dP}_a(x_{ak}).$$

 © Ralf R. Müller 2005-2006

In the limit of $K \to \infty$ one of the exponential terms in (6) will dominate over all others. Only the maximum value of the correlation $Q_{ab}[c]$ is relevant for calculation of the integral.

We assume that the replicas within the dominant subshell are symmetric (replica symmetry). Thus, the maximum values of the correlations $Q_{ab}[c]$ are identical for all positive $a \neq b$. The same applies to the the correlations $Q_{a0}[c]$.

Hereby, we reduce the number of different correlation variables from $(n+1)(n+2)/2$ to four per chip time and set $Q_{00}[c] = p_{0c}$, $Q_{0a}[c] = m_c, \forall a \neq 0$, $Q_{aa}[c] = p_c, \forall a \neq 0$, $Q_{ab}[c] = q_c, \forall 0 \neq a \neq b \neq 0$.

We apply the same idea to the correlation variables in the Fourier domain and set $\tilde{Q}_{00}[c] = G_{0c}/2$, $\tilde{Q}_{aa}[c] = G_c/2, \forall a \neq 0$, $\tilde{Q}_{0a}[c] = E_c, \forall a \neq 0$, and $\tilde{Q}_{ab}[c] = F_c, \forall 0 \neq a \neq b \neq 0$.

At this point the crucial benefit of the replica method becomes obvious. Assuming replica continuity, we have managed to reduce the evaluation of a continuous function to sampling it at integer points. Assuming replica symmetry we have reduced the task of evaluating infinitely many integer points to calculating 8 different correlations (4 of them in the original and 4 of them in the Fourier domain).

© Ralf R. Müller 2005-2006

The assumption of replica symmetry leads to

$$\sum_{a\leq b} \tilde{Q}_{ab}[c]Q_{ab}[c] = nE_c m_c + \frac{n(n-1)}{2} F_c q_c + \frac{G_{0c}p_{0c}}{2} + \frac{n}{2} G_c p_c \qquad (8)$$

and

$$M_k\{E,F,G,G_0\} = \int\limits_{\mathbb{R}^{n+1}} \mathrm{e}^{\frac{1}{N}\sum\limits_{c=1}^{N} w_{ck}^2 \left(\frac{G_{0c}}{2}x_{0k}^2 + \sum\limits_{a=1}^{n} E_c x_{0k}x_{ak} + \frac{G_c}{2}x_{ak}^2 + \sum\limits_{b=a+1}^{n} F_c x_{ak}x_{bk}\right)} \prod_{a=0}^{n} \mathrm{dP}_a(x_{ak})$$

$$= \int\limits_{\mathbb{R}^{n+1}} \mathrm{e}^{\frac{\tilde{G}_{0k}}{2}x_{0k}^2 + \sum\limits_{a=1}^{n} \tilde{E}_k x_{0k}x_{ak} + \frac{\tilde{G}_k}{2}x_{ak}^2 + \sum\limits_{b=a+1}^{n} \tilde{F}_k x_{ak}x_{bk}} \prod_{a=0}^{n} \mathrm{dP}_a(x_{ak}) \qquad (9)$$

where

$$\tilde{E}_k \triangleq \frac{1}{N}\sum_{c=1}^{N} E_c w_{ck}^2 \;, \qquad \tilde{F}_k \triangleq \frac{1}{N}\sum_{c=1}^{N} F_c w_{ck}^2 \qquad (10)$$

$$\tilde{G}_k \triangleq \frac{1}{N}\sum_{c=1}^{N} G_c w_{ck}^2 \;, \qquad \tilde{G}_{0k} \triangleq \frac{1}{N}\sum_{c=1}^{N} G_{0c} w_{ck}^2. \qquad (11)$$

(9) cannot be simplified further for a general prior distribution $\mathrm{dP}_a(x_{ak})$.

 © Ralf R. Müller 2005-2006

Second, we evaluate $\mathrm{e}^{\mathcal{G}\{Q[c]\}}$ in (6). We use the replica symmetry to construct the correlated Gaussian random variables v_{ac} out of independent zero-mean, unit-variance Gaussian random variables u_c, t_c, z_{ac} by

$$\begin{aligned} v_{0c} &= u_c \sqrt{p_{0c} - \frac{m_c^2}{q_c}} - t_c \frac{m_c}{\sqrt{q_c}} \\ v_{ac} &= z_{ac} \sqrt{p_c - q_c} - t_c \sqrt{q_c}, \quad a > 0. \end{aligned}$$

With that substitution, we get

$$\begin{aligned} \mathrm{e}^{\mathcal{G}(m_c, q_c, p_c, p_{0c})} &= \frac{1}{\sqrt{2\pi}\sigma_0} \int\limits_{\mathbb{R}^2} \int\limits_{\mathbb{R}} \exp\left[-\frac{\beta}{2\sigma_0^2} \left(u_c \sqrt{p_{0c} - \frac{m_c^2}{q_c}} - \frac{t_c m_c}{\sqrt{q_c}} - \frac{y_c}{\sqrt{\beta}} \right)^2 \right] \mathrm{D}u_c \\ &\quad \times \left[\int\limits_{\mathbb{R}} \exp\left[-\frac{\beta}{2\sigma^2} \left(z_c \sqrt{p_c - q_c} - t_c \sqrt{q_c} - \frac{y_c}{\sqrt{\beta}} \right)^2 \right] \mathrm{D}z_c \right]^n \mathrm{D}t_c \, \mathrm{d}y_c \\ &= \sqrt{\frac{\left(1 + \frac{\beta}{\sigma^2}(p_c - q_c)\right)^{1-n}}{1 + \frac{\beta}{\sigma^2}(p_c - q_c) + n\frac{\beta}{\sigma^2}\left(\frac{\sigma_0^2}{\beta} + p_{0c} - 2m_c + q_c\right)}} \end{aligned} \tag{12}$$

with the Gaussian measure $\mathrm{D}z = \exp(-z^2/2)\,\mathrm{d}z/\sqrt{2\pi}$.

© Ralf R. Müller 2005-2006

Since the integral in (6) is dominated by the maximum argument of the exponential function, the derivatives of

$$\frac{1}{N}\sum_{c=1}^{N}\left(\mathcal{G}\{Q[c]\}-\beta\sum_{a\leq b}\tilde{Q}_{ab}[c]Q_{ab}[c]\right) \tag{13}$$

with respect to m_c, q_c, p_c and p_{0c} must vanish as $N \to \infty$. Taking derivatives after plugging (8) and (12) into (13), solving for E_c, F_c, G_c, and G_{0c} and letting $n \to 0$ yields for all c

$$E_c = \frac{1}{\sigma^2 + \beta(p_c - q_c)} \tag{14}$$

$$F_c = \frac{\sigma_0^2 + \beta\left(p_{0c} - 2m_c + q_c\right)}{\left[\sigma^2 + \beta(p_c - q_c)\right]^2}$$

$$G_c = F_c - E_c \tag{15}$$

$$G_{0c} = 0. \tag{16}$$

In order to proceed with the calculations, a prior distribution needs to be specified.

Gaussian Prior Distribution

Assume the Gaussian prior

$$\mathrm{p}_a\left(x_{ak}\right) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x_{ak}^2/2} \qquad \forall a.$$

The integration in (9) can be performed explicitly and we find

$$M_k\left(\tilde{E}_k, \tilde{F}_k, \tilde{G}_k, \tilde{G}_{0k}\right) = \sqrt{\frac{\left(1+\tilde{F}_k-\tilde{G}_k\right)^{1-n}}{\left(1-\tilde{G}_{0k}\right)\left(1+\tilde{F}_k-\tilde{G}_k-n\tilde{F}_k\right)-n\tilde{E}_k^2}}. \qquad (17)$$

In the large system limit, the integral in (7) is dominated by that value of the integration variable which maximizes the argument of the exponential function. Thus, partial derivations of

$$\log \prod_{k=1}^{K} M_k\left(\tilde{E}_k, \tilde{F}_k, \tilde{G}_k, \tilde{G}_{0k}\right) - \beta \sum_{c=1}^{N} nE_c m_c + \frac{n(n-1)}{2} F_c q_c + \frac{G_{0c} p_{0c}}{2} + \frac{n}{2} G_c p_c \quad (18)$$

with respect to E_c, F_c, G_c, G_{0c} must vanish for all c as $N \to \infty$.

 © Ralf R. Müller 2005-2006

An explicit calculation of these derivatives yields

$$m_c = \frac{1}{K}\sum_{k=1}^{K} w_{ck}^2 \frac{\tilde{E}_k}{1+\tilde{E}_k} \tag{19}$$

$$q_c = \frac{1}{K}\sum_{k=1}^{K} w_{ck}^2 \frac{\tilde{E}_k^2+\tilde{F}_k}{\left(1+\tilde{E}_k\right)^2} \tag{20}$$

$$p_c = \frac{1}{K}\sum_{k=1}^{K} w_{ck}^2 \frac{\tilde{E}_k^2+\tilde{E}_k+\tilde{F}_k+1}{\left(1+\tilde{E}_k\right)^2} \tag{21}$$

$$p_{0c} = \frac{1}{K}\sum_{k=1}^{K} w_{ck}^2 \tag{22}$$

in the limit $n \to 0$ with (15) and (16).

Surprisingly, if we let the true prior to be binary and only the replicas to be Gaussian we also find (19) to (22). Note from Chapter 1 that this setting corresponds to linear MMSE detection.

Collecting our previous results to evaluate the free energy, we find

$$
\begin{aligned}
-\frac{1}{K}\frac{\partial}{\partial n}\log \Xi_n &= \frac{1}{K}\frac{\partial}{\partial n}\sum_{c=1}^{N}\left[-\mathcal{G}(m_c,q_c,p_c,p_{0c}) + \beta n E_c m_c + \frac{\beta n(n-1)}{2}F_c q_c + \frac{\beta n}{2}G_c p_c\right] \\
&\quad - \sum_{k=1}^{K}\log M_k\left(\tilde{E}_k,\tilde{F}_k,\tilde{G}_k,0\right) \\
&= \frac{1}{2K}\left[\sum_{c=1}^{N}\log\left(1+\frac{\beta}{\sigma^2}(p_c-q_c)\right) + 2\beta E_c m_c + \beta(2n-1)F_c q_c + \beta G_c p_c \right. \\
&\quad \left. + \frac{\sigma_0^2+\beta(p_{0c}-2m_c+q_c)}{\sigma^2+\beta(p_c-q_c)+n\sigma_0^2+n\beta(p_{0c}-2m_c+q_c)}\right] \\
&\quad + \frac{1}{2K}\sum_{k=1}^{K}\log\left(1+\tilde{E}_k\right) - \frac{\tilde{E}_k^2+\tilde{F}_k}{1+\tilde{E}_k-n\tilde{E}_k^2-n\tilde{F}_k} \\
&\xrightarrow{n\to 0} \frac{1}{2K}\left[\sum_{c=1}^{N}\log\left(1+\frac{\beta}{\sigma^2}(p_c-q_c)\right) + \frac{E_c}{F_c} + 2\beta E_c m_c - \beta F_c q_c + \beta G_c p_c\right] \\
&\quad + \frac{1}{2K}\sum_{k=1}^{K}\log\left(1+\tilde{E}_k\right) - \frac{\tilde{E}_k^2+\tilde{F}_k}{1+\tilde{E}_k} \\
&= \frac{\mathrm{F}(\boldsymbol{x})}{K}.
\end{aligned}
$$

 © Ralf R. Müller 2005-2006

This is the final result for the mismatched detector. The six macroscopic parameters $E_c, F_c, G_c, m_c, q_c, p_c$ are implicitly given by the simultaneous solution of the system of equations (14) to (15) and (19) to (21) with the definitions (10) to (11) for all chip times c. This system of equations can only be solved numerically.

Specializing our result to the matched detector by letting $\sigma \to \sigma_0$, we have $F_c \to E_c$, $G_c \to G_{0c}$, $q_c \to m_c$, $p_c \to p_{0c}$. This makes the free energy simplify to

$$\frac{\mathrm{F}(\boldsymbol{x})}{K} = \frac{1}{2K}\left[\sum_{c=1}^{N} \log\left(1 + \frac{\beta}{\sigma_0^2}(p_{0c} - m_c)\right) + 1 + \beta E_c m_c\right] + \frac{1}{2K}\sum_{k=1}^{K} \log\left(1 + \tilde{E}_k\right) - \tilde{E}_k$$
$$= \frac{1}{2K}\left[\sum_{c=1}^{N} \sigma_0^2 E_c - \log\left(\sigma_0^2 E_c\right)\right] + \frac{1}{2K}\sum_{k=1}^{K} \log\left(1 + \tilde{E}_k\right)$$

with

$$E_c = \frac{1}{\sigma_0^2 + \frac{\beta}{K}\sum_{k=1}^{K} \frac{w_{ck}^2}{1+\tilde{E}_k}}. \tag{23}$$

This result is more compact and it requires only to solve (23) numerically which is conveniently done by fixed-point iteration.

Girko's Law

Let the $N \times K$ random matrix $\boldsymbol{H}$ be composed of independent entries $(\boldsymbol{H})_{ij}$ with zero-mean and variances w_{ij}/N such that all w_{ij} are uniformly bounded from above. Assume that the empirical joint distribution of variances $w : [0,1] \times [0,\beta] \mapsto \mathbb{R}$ defined by $w(x,y) = w_{ij}$ for i,j satisfying

$$\frac{i}{N} \leq x \leq \frac{i+1}{N} \qquad \text{and} \qquad \frac{j}{N} \leq y \leq \frac{j+1}{N}$$

converges to a bounded joint limit distribution $w(x,y)$ as $K = \beta N \to \infty$. Then, for each $a,b \in [0,1], a < b$, and $\Im(s) > 0$

$$\frac{1}{N} \sum_{i=\lceil aN \rceil}^{\lceil bN \rceil} \left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} - s\mathbf{I}\right)_{ii}^{-1} \longrightarrow \int_a^b u(x,s)\mathrm{d}x$$

 © Ralf R. Müller 2005-2006

where convergence is in probability and $u(x,s)$ satisfies the fixed point equation

$$u(x,s) = \left[-s + \int_0^\beta \frac{w(x,y)\,\mathrm{d}y}{1 + \int_0^1 u(x',s) w(x',y)\,\mathrm{d}x'} \right]^{-1}$$

for every $x \in [0,1]$. The solution always exists and is unique in the class of functions $u(x,s) \geq 0$, analytic for $\Im(s) > 0$ and continuous on $x \in [0,1]$.

Moreover, almost surely, the empirical eigenvalue distribution of $\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}$ converges weakly to a limiting distribution whose Stieltjes transform is given by

$$G_{\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}}(s) = \int_0^1 u(x,s)\,\mathrm{d}x.$$

Comparing (23) to Girko's law, $\tilde{E}_k$ is recognized as the signal-to-interference and noise ratio of user k.

Using the similarity of free energy and the entropy of the channel output allows for the simple relationship

$$\frac{\mathrm{I}(\boldsymbol{x}, \boldsymbol{y})}{K} = \frac{\mathrm{F}(\boldsymbol{x})}{K} - \frac{1}{2\beta} \tag{24}$$

between the (normalized) free energy and the (normalized) mutual information between channel input signal $\boldsymbol{x}$ and channel output signal y given the channel matrix $\boldsymbol{H}$. Assuming that the channel is perfectly known to the receiver, but totally unknown to the transmitter, (24) gives the channel capacity per user.

Binary Prior Distribution

Consider a non-uniform binary prior

$$\mathrm{p}_a\left(x_{ak}\right)=\frac{1+t_k}{2}\,\delta(x_{ak}-1)+\frac{1-t_k}{2}\,\delta(x_{ak}+1). \tag{25}$$

Plugging the prior distribution into (9), we find

$$\begin{aligned}
M_k\left(\tilde{E}_k,\tilde{F}_k,\tilde{G}_k,\tilde{G}_{0k}\right) &= \int\limits_{\mathbb{R}^{n+1}} \mathrm{e}^{\frac{\tilde{G}_{0k}+n\tilde{G}_k}{2}+\sum\limits_{a=1}^{n}\tilde{E}_k x_{0k}x_{ak}+\sum\limits_{b=a+1}^{n}\tilde{F}_k x_{ak}x_{bk}} \prod_{a=0}^{n}\mathrm{dP}_a(x_{ak}) \\
&= \mathrm{e}^{\frac{1}{2}\left(\tilde{G}_{0k}+n\tilde{G}_k\right)} \sum_{\{x_{ak},a=1,\ldots,n\}} \left\{\frac{1+t_k}{2}\exp\left[\sum_{a=1}^{n}\tilde{E}_k x_{ak}+\sum_{b=a+1}^{n}\tilde{F}_k x_{ak}x_{bk}\right]\right. \\
&\quad \left. +\frac{1-t_k}{2}\exp\left[\sum_{a=1}^{n}-\tilde{E}_k x_{ak}+\sum_{b=a+1}^{n}\tilde{F}_k x_{ak}x_{bk}\right]\right\}\prod_{a=1}^{n}\Pr(x_{ak}) \\
&= \mathrm{e}^{\frac{1}{2}\left(\tilde{G}_{0k}+n\tilde{G}_k-n\tilde{F}_k\right)} \sum_{\{x_{ak},a=1,\ldots,n\}} \left\{\frac{1+t_k}{2}\exp\left[\frac{\tilde{F}_k}{2}\left(\sum_{a=1}^{n}x_{ak}\right)^2+\tilde{E}_k\sum_{a=1}^{n}x_{ak}\right]\right. \\
&\quad \left. +\frac{1-t_k}{2}\exp\left[\frac{\tilde{F}_k}{2}\left(\sum_{a=1}^{n}x_{ak}\right)^2-\tilde{E}_k\sum_{a=1}^{n}x_{ak}\right]\right\}\prod_{a=1}^{n}\Pr(x_{ak}).
\end{aligned} \tag{26}$$

We can use the following property of the Gaussian measure

$$\exp\left(\tilde{F}_k \frac{S^2}{2}\right) = \int \exp\left(\pm\sqrt{\tilde{F}_k} z S\right) \mathrm{D}z \qquad \forall S \in \mathbb{R}$$

to linearize the exponents

$$M_k\left(\tilde{E}_k, \tilde{F}_k, \tilde{G}_k, \tilde{G}_{0k}\right) = e^{\frac{1}{2}\left(\tilde{G}_{0k}+n\tilde{G}_k-n\tilde{F}_k\right)} \sum_{\{x_{ak}, a=1,\ldots,n\}} \int \frac{1+t_k}{2} \exp\left[\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k\right)\sum_{a=1}^n x_{ak}\right] + \frac{1-t_k}{2} \exp\left[-\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k\right)\sum_{a=1}^n x_{ak}\right] \mathrm{D}z \prod_{a=1}^n \Pr(x_{ak}).$$

Since

$$\begin{aligned} f_n &\triangleq \sum_{\{x_{ka}, a=1,\ldots,n\}} \exp\left[\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k\right)\sum_{a=1}^n x_{ka}\right]\prod_{a=1}^n \Pr(x_{ka}) \\ &= \sum_{x_{kn}} \Pr(x_{kn}) f_{n-1} \exp\left[\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k\right) x_{kn}\right] \\ &= f_{n-1} \frac{\cosh\left[\lambda_k/2 + z\sqrt{\tilde{F}_k}+\tilde{E}_k\right]}{\cosh\left(\lambda_k/2\right)} = \frac{\cosh^n\left[\lambda_k/2 + z\sqrt{\tilde{F}_k}+\tilde{E}_k\right]}{\cosh^n\left(\lambda_k/2\right)} \end{aligned}$$

 © *Ralf R. Müller 2005-2006*

with $t_k \triangleq \tanh(\lambda_k/2)$, we find

$$M_k\left(\tilde{E}_k, \tilde{F}_k, \tilde{G}_k, \tilde{G}_{0k}\right) = \frac{\int \frac{1+t_k}{2}\cosh^n\left(z\sqrt{\tilde{F}_k} + \tilde{E}_k + \frac{\lambda_k}{2}\right) + \frac{1-t_k}{2}\cosh^n\left(z\sqrt{\tilde{F}_k} + \tilde{E}_k - \frac{\lambda_k}{2}\right)\mathrm{D}z}{\cosh^n\left(\frac{\lambda_k}{2}\right)\exp\left(\frac{n\tilde{F}_k - \tilde{G}_{0k} - n\tilde{G}_k}{2}\right)}.$$

In the large system limit, the integral in (7) is dominated by that value of the integration variable which maximizes the argument of the exponential function. Thus, partial derivations of

$$\log\prod_{k=1}^{K} M_k\left(\tilde{E}_k, \tilde{F}_k, \tilde{G}_k, \tilde{G}_{0k}\right) - \beta\sum_{c=1}^{N} nE_c m_c + \frac{n(n-1)}{2}F_c q_c + \frac{G_{0c}p_{0c}}{2} + \frac{n}{2}G_c p_c$$

with respect to E_c, F_c, G_c, G_{0c} must vanish for all c as $N \to \infty$.

 © Ralf R. Müller 2005-2006

An explicit calculation of these derivatives gives

$$m_c = \frac{1}{K}\sum_{k=1}^{K} w_{ck}^2 \int \tfrac{1+t_k}{2}\tanh\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k+\tfrac{\lambda_k}{2}\right)+\tfrac{1-t_k}{2}\tanh\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k-\tfrac{\lambda_k}{2}\right)\mathrm{D}z \qquad (27)$$

$$q_c = \frac{1}{K}\sum_{k=1}^{K} w_{ck}^2 \int \tfrac{1+t_k}{2}\tanh^2\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k+\tfrac{\lambda_k}{2}\right)+\tfrac{1-t_k}{2}\tanh^2\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k-\tfrac{\lambda_k}{2}\right)\mathrm{D}z \qquad (28)$$

$$p_c = p_{0c} = \frac{1}{K}\sum_{k=1}^{K} w_{ck}^2 \qquad (29)$$

in the limit $n \to 0$.

In order to obtain (28), note from (26) that the first order derivative of $M_k \exp(n\tilde{F}_k/2)$ with respect to $\tilde{F}_k$ is identical to half of the second order derivative of $M_k \exp(n\tilde{F}_k/2)$ with respect to $\tilde{E}_k$.

 © Ralf R. Müller 2005-2006

Collecting our previous results to evaluate the free energy, we find

$$
\begin{aligned}
-\frac{1}{K}\frac{\partial}{\partial n}\log \Xi_n = & \frac{1}{K}\frac{\partial}{\partial n}\sum_{c=1}^{N}\left[-\mathcal{G}(m_c,q_c,p_c,p_{0c}) + \beta n E_c m_c + \frac{\beta n(n-1)}{2}F_c q_c + \frac{\beta n}{2}G_c p_c\right] \\
& -\sum_{k=1}^{K}\log M_k\left(\tilde{E}_k,\tilde{F}_k,\tilde{G}_k,0\right) \\
\xrightarrow{n\to 0} & \frac{1}{2K}\sum_{c=1}^{N}\left[\log\left(1+\frac{\beta}{\sigma^2}(p_c-q_c)\right)+\frac{E_c}{F_c}+2\beta E_c m_c - \beta F_c q_c + \beta G_c p_c\right] \\
& -\frac{1}{K}\sum_{k=1}^{K}\int \frac{1+t_k}{2}\log\cosh\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k+\frac{\lambda_k}{2}\right) \\
& +\frac{1-t_k}{2}\log\cosh\left(z\sqrt{\tilde{F}_k}+\tilde{E}_k-\frac{\lambda_k}{2}\right)\mathrm{D}z + \frac{1}{2}\log\left(1-t_k^2\right) - \frac{\tilde{F}_k+\tilde{G}_k}{2} \\
= & \frac{\mathrm{F}(\boldsymbol{x})}{K}.
\end{aligned}
$$

This is the final result for the free energy of the mismatched detector. The six macroscopic parameters $E_c, F_c, G_c, m_c, q_c, p_c$ are implicitly given by the simultaneous solution of the system of equations (14) to (15) and (27) to (29) with the definitions (10) to (11) for all chip times c. This system of equations can only be solved numerically.

In case of multiple solutions, the correct solution is that one which minimizes the free energy.

Specializing our result to the matched detector by letting $\sigma \to \sigma_0$, we have $F_c \to E_c$, $G_c \to G_{0c}$, $q_c \to m_c$. This makes the free energy simplify to

$$\begin{aligned}
\frac{\mathrm{F}(\boldsymbol{x})}{K} &= \frac{1}{2K}\sum_{c=1}^{N}\left[\log\left(1+\frac{\beta}{\sigma_0^2}(p_{0c}-m_c)\right)+1+\beta E_c m_c\right]-\frac{1}{K}\sum_{k=1}^{K}\log\sqrt{1-t_k^2}-\frac{\tilde{E}_k}{2}\\
&\quad+\int \tfrac{1+t_k}{2}\log\cosh\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k+\tfrac{\lambda_k}{2}\right)+\tfrac{1-t_k}{2}\log\cosh\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k-\tfrac{\lambda_k}{2}\right)\mathrm{D}z\\
&= \frac{1}{2K}\sum_{c=1}^{N}\left[\sigma_0^2 E_c-\log\left(\sigma_0^2 E_c\right)\right]-\frac{1}{K}\sum_{k=1}^{K}\log\sqrt{1-t_k^2}-\tilde{E}_k\\
&\quad+\int \tfrac{1+t_k}{2}\log\cosh\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k+\tfrac{\lambda_k}{2}\right)+\tfrac{1-t_k}{2}\log\cosh\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k-\tfrac{\lambda_k}{2}\right)\mathrm{D}z
\end{aligned}$$

where the macroscopic parameters E_c are given by

$$\begin{aligned}
\frac{1}{E_c} &= \sigma_0^2+\frac{\beta}{K}\sum_{k=1}^{K}w_{ck}^2\left[1-\int\frac{1+t_k}{2}\tanh\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k+\frac{\lambda_k}{2}\right)\right.\\
&\qquad\left.+\frac{1-t_k}{2}\tanh\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k-\frac{\lambda_k}{2}\right)\mathrm{D}z\right]\\
&= \sigma_0^2+\frac{\beta}{K}\sum_{k=1}^{K}w_{ck}^2\left(1-t_k^2\right)\int\frac{1-\tanh\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k\right)}{1-t_k^2\tanh^2\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k\right)}\mathrm{D}z.
\end{aligned}$$

 © Ralf R. Müller 2005-2006

Similar to the case of Gaussian priors, $\tilde{E}_k$ can be shown to be a kind of signal-to-interference and noise ratio, in the sense that the bit error probability of user k is given by

$$\Pr(\hat{x}_k \neq x_k) = \int_{\sqrt{\tilde{E}_k}}^{\infty} \mathrm{D}z.$$

An equivalent additive white Gaussian noise channel can be defined to model the multiuser interference for any prior.

For any input alphabet to the channel mutual information is given by (24) with the free energy corresponding to that input alphabet.

 © *Ralf R. Müller 2005-2006*

MC-CDMA in Multipath Fading

Equivalent baseband vector channel in frequency domain:

$$\underbrace{\boldsymbol{y}}_{\substack{N \times 1 \\ \text{frequency chips}}} = \Big(\underbrace{\boldsymbol{W}}_{\substack{N \times K \\ \text{channel matrix}}} \underbrace{\odot}_{\text{Hadamard product}} \underbrace{\boldsymbol{S}}_{\substack{N \times K \\ \text{spreading matrix}}} \Big) \underbrace{\boldsymbol{x}}_{\substack{K \times 1 \\ \text{users' data}}} + \underbrace{\boldsymbol{n}}_{\substack{N \times 1 \\ \text{noise vector}}}$$

- The noise $\boldsymbol{n}$ has i.i.d. Gaussian entries with *zero-mean* and *unit variance*.
- The columns of $\boldsymbol{S}$ are the random spreading sequences of the users.
- The columns of $\boldsymbol{W}$ are the random frequency responses of the users.

© *Ralf R. Müller 2005-2006*

Minimum Probability of Error for MAP Detector

Maximum a-posteriori detector:

$$\hat{x}_k = \arg\max_{x_k} \Pr(x_k | \boldsymbol{y}, \boldsymbol{W})$$

In the large system limit, there is an equivalent AWGN channel with SINR $\tilde{E}_k$ such that

$$\Pr(\hat{x}_k \neq x_k | \boldsymbol{W}) = \int_{\sqrt{\tilde{E}_k}}^{\infty} Dz = \mathrm{Q}\left(\sqrt{\tilde{E}_k}\right)$$

and

$$\Pr(\hat{x}_k \neq x_k) = \mathop{\mathrm{E}}_{\boldsymbol{W}} \Pr(\hat{x}_k \neq x_k | \boldsymbol{W}) = \mathop{\mathrm{E}}_{\boldsymbol{W}} \mathrm{Q}\left(\sqrt{\tilde{E}_k}\right)$$

© Ralf R. Müller 2005-2006

SINR of Equivalent AWGN Channel

For N, K large, solve the fixed-point system of equations

$$\tilde{E}_k = \frac{1}{N}\sum_{c=1}^{N} E_c w_{ck}^2$$

$$E_c = \frac{1}{\sigma_n^2 + \frac{\beta}{K}\sum_{k=1}^{K}(1-t_k)^2 w_{ck}^2 \int \frac{1-\tanh\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k\right)}{1-t_k^2\tanh^2\left(z\sqrt{\tilde{E}_k}+\tilde{E}_k\right)} Dz}$$

In practice, the fading statistics obey some structure:

- Asymptotic frequency-invariance on the uplink (reverse link)
- Rank-1 statistics on the downlink (forward link)

© Ralf R. Müller 2005-2006

Asymptotic Frequency Invariance (Uplink)

$$E_c = E \qquad \forall c$$

The fading is ergodic across the user population for each frequency c.

$$\tilde{E}_k = \frac{P_k}{\sigma_n^2 + \frac{\beta}{K}\sum_{n=1}^{K}(1-t_n)^2 P_n \int \frac{1-\tanh\left(z\sqrt{\tilde{E}_n}+\tilde{E}_n\right)}{1-t_n^2\tanh^2\left(z\sqrt{\tilde{E}_n}+\tilde{E}_n\right)} Dz}$$

with

$$P_k = \frac{1}{N}\sum_{c=1}^{N} w_{ck}^2$$

The spectrum of the received signal is white (frequency-invariant).

Full diversity is achieved.

© *Ralf R. Müller 2005-2006*

Rank 1 Statistics (Downlink)

$$\boldsymbol{W} = \boldsymbol{f}\boldsymbol{u}^{\mathrm{T}} \qquad \Longleftrightarrow \qquad w_{ck} = f_c u_k$$

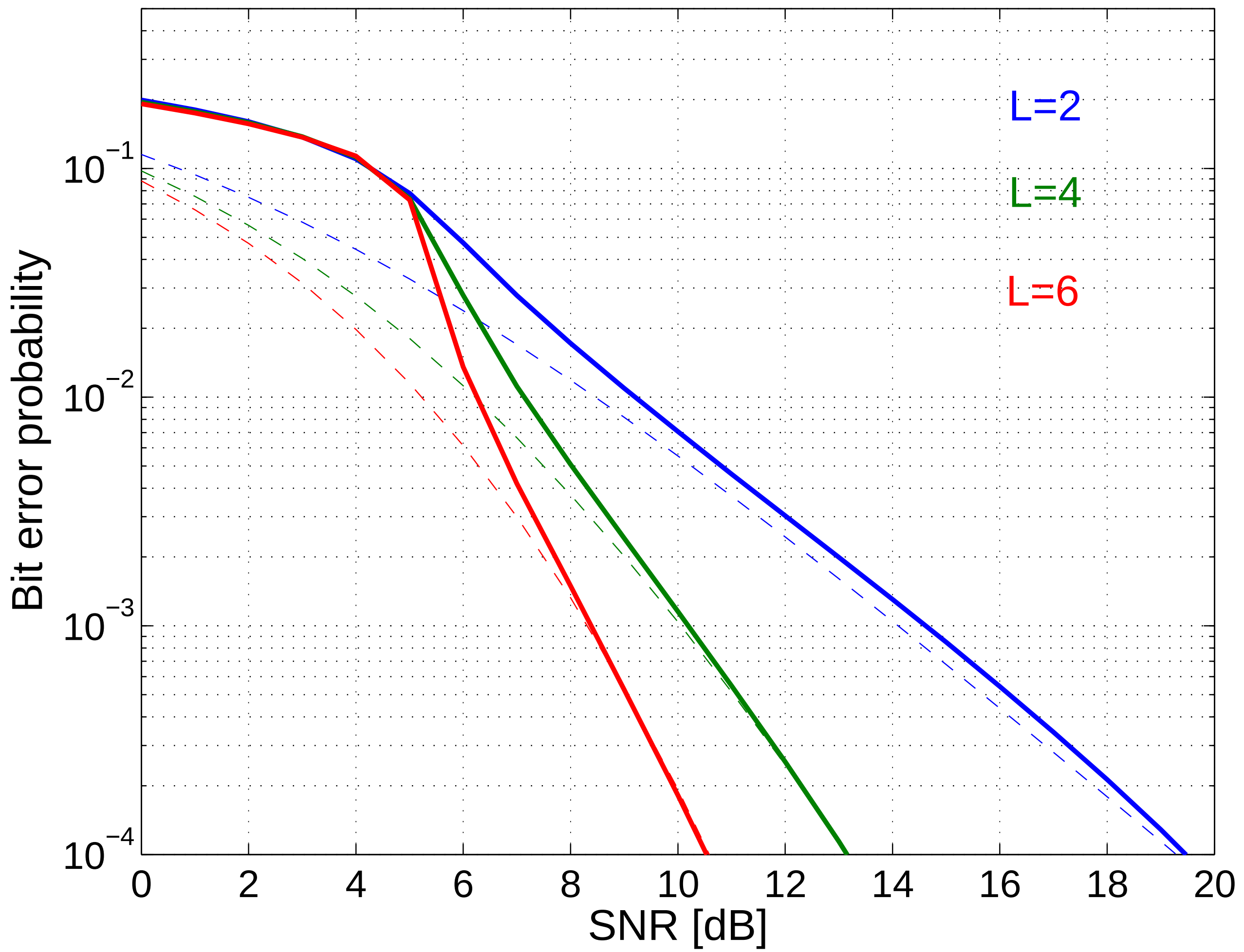
All users experience the same fading channel except for a scalar factor u_k.

$$\tilde{E}_k = \frac{u_k^2}{N} \sum_{c=1}^{N} \frac{1}{\frac{\sigma_n^2}{f_c^2} + \frac{\beta}{K} \sum_{n=1}^{K} (1-t_n)^2 u_n^2 \int \frac{1 - \tanh\left(z\sqrt{\tilde{E}_n} + \tilde{E}_n\right)}{1 - t_n^2 \tanh^2\left(z\sqrt{\tilde{E}_n} + \tilde{E}_n\right)} Dz}$$

Full diversity is achieved.

The spectrum of the received signal is colored $\Longrightarrow$ degradation.

Uplink

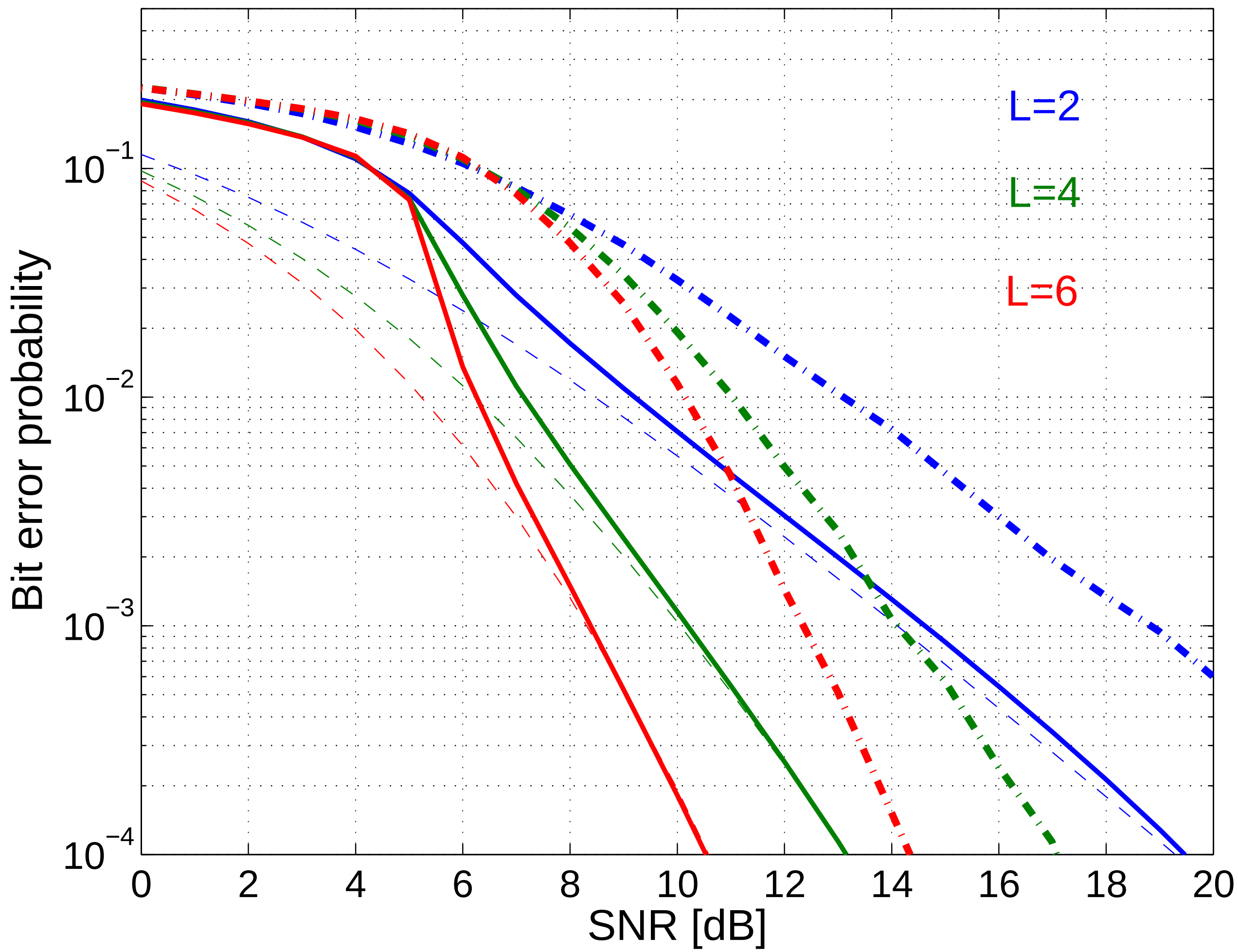


Uniform priors

L equal power paths

$$\frac{K}{N} = 1.5$$

Uplink vs. Downlink



Uniform priors

L equal power paths

$$\frac{K}{N} = 1.5$$

4 dB difference!

Part II:

Multistage Detection

© *Ralf R. Müller 2005-2006*

Matrix Inversion

Let $\boldsymbol{R}$ be non–singular.

Let λ_i denote the eigenvalues of $\boldsymbol{R}$.

Then,

$$\prod_{k=1}^{K}\left(\boldsymbol{R} - \lambda_k \mathbf{I}\right) = \mathbf{0} \qquad \Longrightarrow \qquad -\mathbf{I} + \sum_{k=1}^{K} \alpha_k \boldsymbol{R}^k = \mathbf{0}$$

Cayley–Hamilton Theorem — with appropriate α_k s.

Solution to matrix inversion problem given the eigenvalues:

$$\boldsymbol{R}^{-1} = -\sum_{k=1}^{K} \alpha_k \boldsymbol{R}^{k-1}$$

© Ralf R. Müller 2005-2006

Linear Multistage Detection

Linear MMSE filter: $\boldsymbol{L}_{\mathrm{MMSE}} = \left(\boldsymbol{R} + \sigma_n^2 \mathbf{I}\right)^{-1}$

Approximation by power series:

Cayley–Hamilton theorem yields:

$$\begin{aligned}\left(\boldsymbol{R} + \sigma_n^2 \mathbf{I}\right)^{-1} &= \sum_{i=0}^{K-1} \tilde{w}_i \boldsymbol{R}^i \\ &\approx \sum_{i=0}^{D-1} w_i \boldsymbol{R}^i \qquad \text{for } D < K.\end{aligned}$$

For random spreading the optimum weights converge almost surely, as $K, N \to \infty$ with $\beta = \frac{K}{N}$, and can be given in closed form.

 © *Ralf R. Müller 2005-2006*

Semi-Universal Weights

Filter shall be independent from the realization of the random matrix $\boldsymbol{S}$, but may use its statistics.

For most large random matrices, as $K = \beta N \to \infty$, many finite dimensional functions of the eigenvalues, e.g. the filter coefficients, freeze.

The asymptotic limits depend only on parts of the statistics of the random matrix.

The weights can be calculated off-line with the help of random matrix and free probability theory.

© *Ralf R. Müller 2005-2006*

Weight Design

is given by Yule–Walker equations:

$$\begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_{D+1} \end{bmatrix} = \begin{bmatrix} m_2+\sigma^2 m_1 & m_3+\sigma^2 m_2 & \dots & m_{D+2}+\sigma^2 m_{D+1} \\ m_3+\sigma^2 m_2 & m_4+\sigma^2 m_3 & \dots & m_{D+3}+\sigma^2 m_{D+2} \\ \vdots & \vdots & \ddots & \vdots \\ m_{D+2}+\sigma^2 m_{D+1} & m_{D+3}+\sigma^2 m_{D+2} & \dots & m_{2D+2}+\sigma^2 m_{2D+1} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_D \end{bmatrix}$$

with the total moments

$$m_n \triangleq \mathrm{E}\{\lambda^n\} = \mathrm{Tr}\left(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}\right)^n$$

Example for Asymptotic Weight Design

Random matrix with i.i.d. entries:

$$m_n = \frac{1}{n}\sum_{i=1}^{n}\binom{n}{i}\binom{n}{i-1}\beta^i.$$

$D=2$	$w_0 = -\sigma^2 w_1 + 2 + 2\beta$ $w_1 = -1$
$D=3$	$w_0 = -\sigma^2 w_1 + 3 + 4\beta + 3\beta^2$ $w_1 = -\sigma^2 w_2 - 3 - 3\beta$ $w_2 = 1$
$D=4$	$w_0 = -\sigma^2 w_1 + 4 + 6\beta + 6\beta^2 + 4\beta^3$ $w_1 = -\sigma^2 w_2 - 6 - 9\beta - 6\beta^2$ $w_2 = -\sigma^2 w_3 + 4 + 4\beta$ $w_3 = -1$
$D=5$	$w_0 = -\sigma^2 w_1 + 5 + 8\beta + 9\beta^2 + 8\beta^3 + 5\beta^4$ $w_1 = -\sigma^2 w_2 - 10 - 18\beta - 18\beta^2 - 10\beta^3$ $w_2 = -\sigma^2 w_3 + 10 + 16\beta + 10\beta^2$ $w_3 = -\sigma^2 w_4 - 5 - 5\beta$ $w_4 = 1$

© Ralf R. Müller 2005-2006

Rate of Convergence

Theorem 1 *Let $\boldsymbol{A} = \mathbf{I}$ and the chips of any user be i.i.d. zero–mean random variables with finite fourth moment and the sequences of all users jointly independent. Then, the multi–user efficiencies of all users converge almost surely, as $N, K \to \infty$ but $\beta = \frac{K}{N}$ fixed, to*

$$\eta_{\mathrm{WLPIC},D+1} = \frac{1}{1 + \dfrac{\beta}{\sigma_n^2 + \eta_{\mathrm{WLPIC},D}}}$$

with $\eta_0 = 0$ for optimally chosen weights.

The approximation converges to the exact MMSE performance as a *continued fraction*.
For optimal coefficients w_i, the approximation error ϵ decreases exponentially with the number of stages D:

$$\epsilon < \mathrm{const.}\, (1 + \mathrm{SNR})^{-D}$$

There are even tighter bounds.

Individual Weight Design

Allow for different weights for different users

$$\left(\boldsymbol{R} + \sigma_n^2 \mathbf{I}\right)^{-1} = \sum_{i=0}^{K-1} \tilde{w}_i \boldsymbol{R}^i$$

$$\approx \sum_{i=0}^{D-1} \boldsymbol{W}_i \boldsymbol{R}^i \qquad \text{for } D < K \qquad \text{and all } \boldsymbol{W}_i \text{ diagonal.}$$

Weight design by the same Yule-Walker equations, but with the k-partial moments

$$m_n^{(k)} = \left[\left(\boldsymbol{S}^{\mathrm{H}} \boldsymbol{S}\right)^n\right]_{kk}.$$

For users with different powers, individual weight design is better.

Do the k-partial moments convergence asymptotically?

© Ralf R. Müller 2005-2006

Convergence of Partial Moments

Let the random matrix $\boldsymbol{H}$ fufill the same conditions as needed for the deformed quarter circle law. Let $\boldsymbol{A}$ be an $K \times K$ diagonal matrix such that its singular value distribution converges almost surely, as $K \to \infty$ to a non-random limit distribution. Let

$$\boldsymbol{R} = \boldsymbol{A}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \boldsymbol{A}.$$

Then, $\left(\boldsymbol{R}^{\ell}\right)_{kk}$, the k^{th} diagonal element of $\boldsymbol{R}^{\ell}$ converges, conditioned on a_{kk}, the k^{th} diagonal element of $\boldsymbol{A}$, almost surely, as $K = \beta N \to \infty$ to

$$R_{kk}^{(\ell)} = |a_{kk}|^2 \beta \sum_{q=1}^{\ell} R_{kk}^{(q-1)} m_{\ell-q}^{(\boldsymbol{R})}, \qquad \ell > 1$$

with

$$m_q^{(\boldsymbol{R})} = \mathrm{Tr}(\boldsymbol{R}^q).$$

The total moments of $\boldsymbol{R}$ are conveniently given by the recursion

$$m_\ell^{(\boldsymbol{R})} = \beta \sum_{q=1}^{\ell} m_{\ell-q}^{(\boldsymbol{R})} \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} |a_{kk}|^2 R_{kk}^{(q-1)}.$$

Structure of Proof

Let

$$\boldsymbol{R} = \boldsymbol{A}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{A}, \qquad \boldsymbol{T} = \boldsymbol{H}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}$$

- Bound the difference between the quadratic form and the trace as Silverstein and Bai did.
- Use Lyapunov inequality to bound the forth moment of the entries of the random matrix $\boldsymbol{H}$.
- Recursively proof convergence of the diagonal elements of $(\boldsymbol{R})^k$ and the traces of $(\boldsymbol{T})^k$ flipping between the two.

 © Ralf R. Müller 2005-2006

Multipath Fading Channels

Let the path differences be only a few chip intervals. Approximate the linear time shift by a cyclic shift modulo N. For large N this becomes more and more accurate.

2 paths: All odd column of the $N \times 2K$ matrix $\boldsymbol{H}$ are i.i.d. Each even column of $\boldsymbol{H}$ is a cyclically shifted version of the adjacent column to the left.

$$\mathrm{E}\left\{\boldsymbol{b}\boldsymbol{b}^{\mathrm{H}}\right\} = \mathbf{I} \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

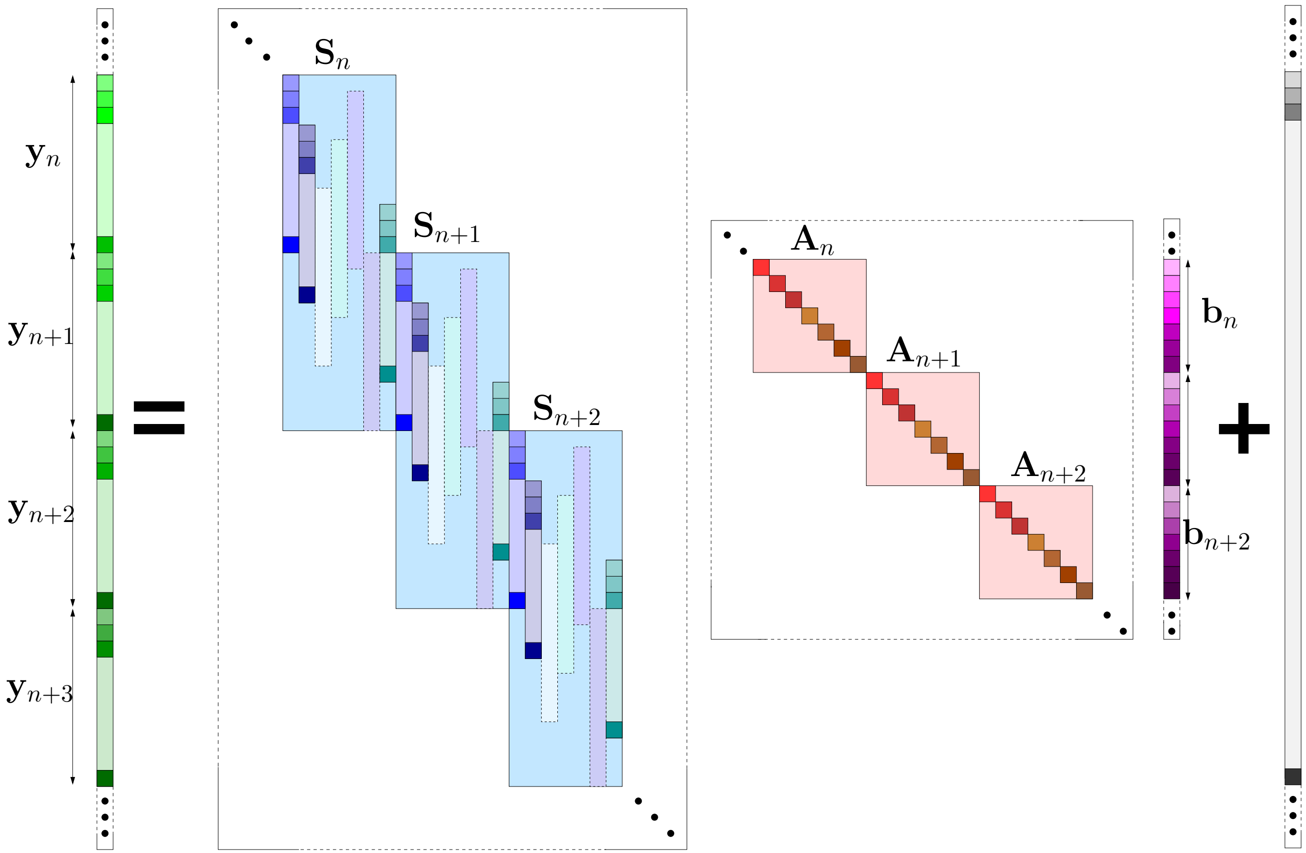
and $(\boldsymbol{A})_{kk}$ are independent zero-mean and complex Gaussian.

This setting is equivalent to the full i.i.d. setting in all asymptotic aspects if the users' powers follow the same distribution.

Equivalence holds for an arbitrary number of paths.

 © Ralf R. Müller 2005-2006

Asynchronous Users

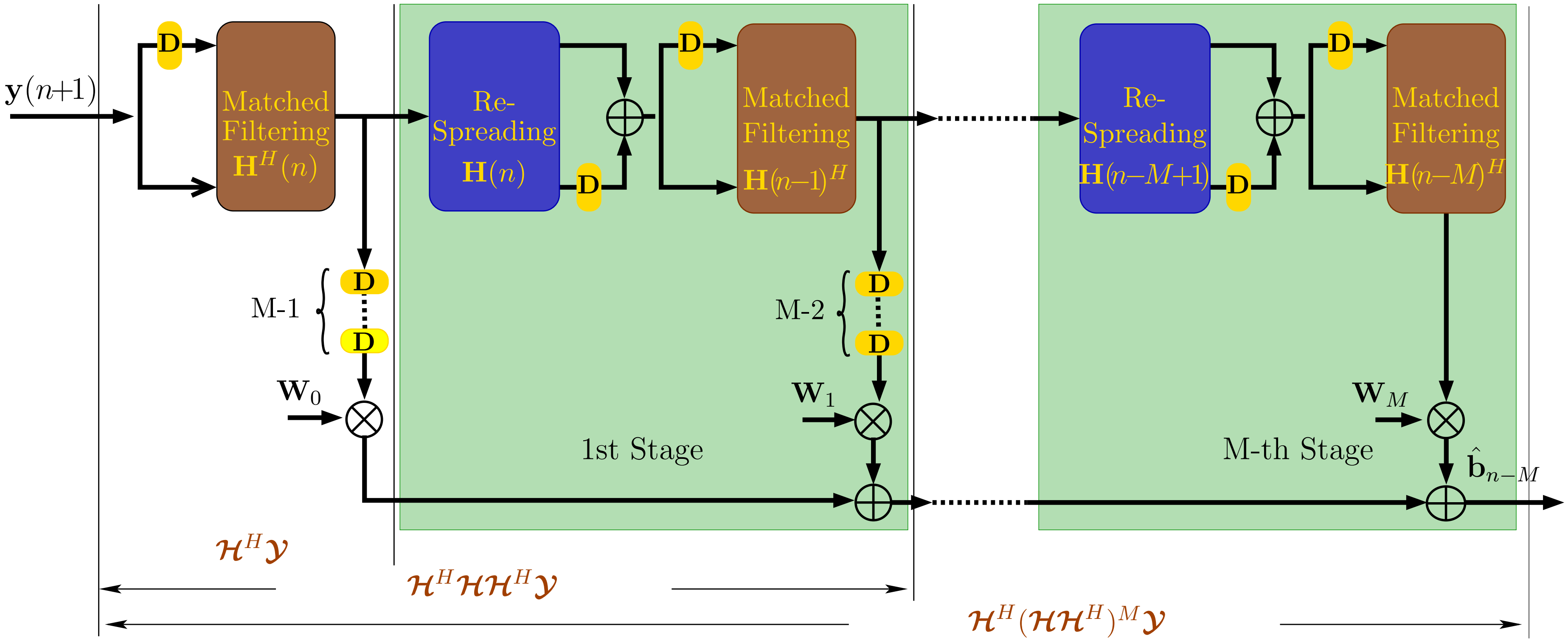


Convergence of k-partial moments proven. Recursive expressions to construct them known. Proof follows the same lines.

$$\mathcal{Y} = \underbrace{\mathcal{S}\mathcal{A}}_{\mathcal{H}}\mathcal{B} + \mathcal{N}$$

© Ralf R. Müller 2005-2006

Detector Structure for Asynchronous Users



No truncation effects.

Chip-Asynchronous Users

- Convergence of k-partial moments proven.
- Recursive expressions to construct them known.
- Proof follows the same lines, but makes use of properties of circulant matrices to cope with shifted bandlimited chip waveforms.

For non-zero roll-off factor, chip-asynchronicity gives significant improvements at high SNR.

 © *Ralf R. Müller 2005-2006*

CDMA with Dual Antenna Arrays

The system is described by the virtual $NR \times K$ spreading matrix

$$\tilde{\boldsymbol{S}} = \begin{bmatrix} h_{11}\boldsymbol{s}_1 & h_{12}\boldsymbol{s}_2 & \dots & h_{1K}\boldsymbol{s}_K \\ h_{21}\boldsymbol{s}_1 & h_{22}\boldsymbol{s}_2 & \dots & h_{2K}\boldsymbol{s}_K \\ \vdots & \vdots & \ddots & \vdots \\ h_{R1}\boldsymbol{s}_1 & h_{R2}\boldsymbol{s}_2 & \dots & h_{RK}\boldsymbol{s}_K \end{bmatrix}$$

Note that with the Kronecker product $\otimes$:

$$\tilde{\boldsymbol{s}}_k = \boldsymbol{h}_k \otimes \boldsymbol{s}_k$$

Note also that the entries of $\tilde{\boldsymbol{S}}$ are not jointly independent even if those ones of $\boldsymbol{S}$ and $\boldsymbol{H}$ are.

A Resource Pooling Result

Theorem 2 *Let the chips of any user be i.i.d. zero–mean random variables with finite 6th moment, the sequences of all users jointly independent, and the antenna array channel* h_{rk} *follow the i.i.d. complex Gaussian model. Then, the multi–user efficiency of the linear MMSE detector converges for all users almost surely, as* $N, K \to \infty$ *but* $\beta \triangleq \frac{K}{N}$ *and* R *fixed, to the deterministic unique positive solution of the fixed–point equation*

$$\frac{1}{\eta_{\text{MMSE}}} = 1 + \frac{\beta}{R} \int \frac{x}{\sigma_n^2 + \eta_{\text{MMSE}}\, x} \, \mathrm{dP}_{\tilde{A}^2}(x),$$

if the powers of the users converge weakly to the limit distribution $\mathrm{P}_{\tilde{A}^2}(x)$ *with*

$$|\tilde{A}_k|^2 = |A_k|^2 \sum_{r=1}^{R} |h_{rk}|^2.$$

Resource Pooling for Correlated MIMO Channels

Theorem 3 *Let the chips of any user be i.i.d. zero–mean random variables with finite 6th moment, the sequences of all users jointly independent, and the empirical distributions of the channel gains* h_{rk} *across the users converge, jointly for all receive antennas* r *to an* R*-dimensional joint limit distribution* $\mathrm{P}_{\boldsymbol{H}}(x)$*. Then, with linear MMSE detection, the SINR of user* k *converges, as* $N, K \to \infty$ *but* $\beta \triangleq \frac{K}{N}$ *and* R *fixed, conditioned on the channel gains of user* k *to*

$$\frac{\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{A} \boldsymbol{h}_k}{\sigma^2}$$

where $\boldsymbol{A}$ *is the deterministic unique positive definite solution of the matrix-valued fixed–point equation*

$$\boldsymbol{A}^{-1} = \mathbf{I} + \beta \int \frac{\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}}{\sigma_n^2 + \boldsymbol{x}^{\mathrm{H}} \boldsymbol{A} \boldsymbol{x}} \, \mathrm{dP}_{\boldsymbol{H}}(\boldsymbol{x}),$$

Asymptotic performance is characterized by an $R \times R$ matrix.

MS Detection for Correlated Resource Pooling

Let

- $\tilde{\boldsymbol{S}} = [\tilde{\boldsymbol{s}}_1, \tilde{\boldsymbol{s}}_2, \ldots, \tilde{\boldsymbol{s}}_K]$ with $\tilde{\boldsymbol{s}}_k = \boldsymbol{h}_k \otimes \boldsymbol{s}_k$ where $\boldsymbol{S}$ is i.i.d.
- the entries of $\boldsymbol{H}$ may be arbitrarily dependent as long as the rows have a joint limit distribution and are finite in number.

Then, as the dimensions of $\boldsymbol{S}$ grow

- the k-partial moments conditioned on $\boldsymbol{h}_k$ converge,
- recursive expressions for them are known,
- the proof follows along the same lines as before.

© *Ralf R. Müller 2005-2006*