

Efficient Equalization for Wireless Communications in Hostile Environments

Thomas Strohmer

Department of Mathematics
University of California, Davis, USA
strohmer@math.ucdavis.edu
<http://math.ucdavis.edu/~strohmer>

Talk includes joint work with Arogyaswami Paulraj, Majid Emami (Stanford), Jan Hansen (Stanford/ETH), Raymond Guan (UC Davis)

Research supported by NSF DMS

Agenda

- What makes equalization difficult?
- Drawbacks of existing equalization methods
- Next generation equalizers
 - Banach algebras of matrices
 - Krylov subspaces, regularization, MMSE
 - Extension to oversampling, MIMO
 - Pre/post/joint equalization

What is a hostile environment?

- large AWGN (due to large bandwidth, interfering devices,...)
- low signal power (due to FCC regulations, equipment limitations,...)
- large delay spread
- deep fading channels
- no or incomplete channel knowledge
- Doppler spread
- interfering users
- low computing power/storage
- request for high data rates
- ...

Doppler spread

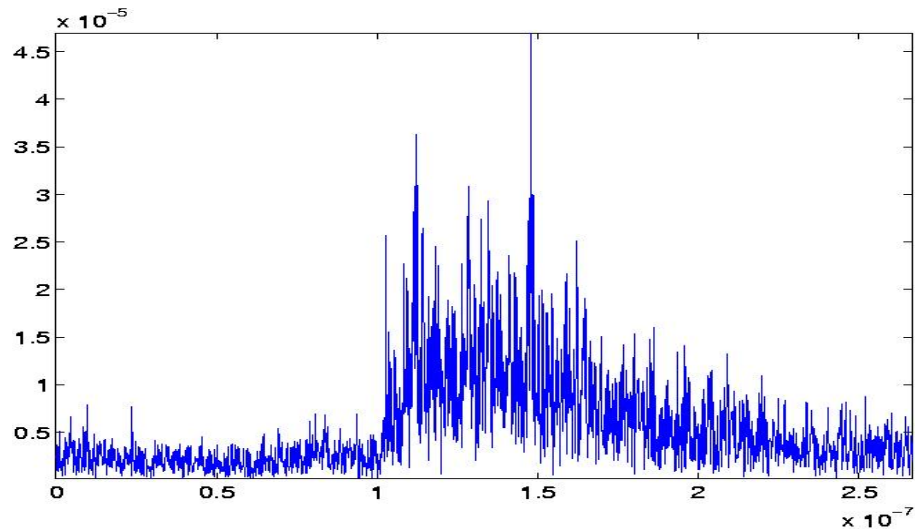
... will not be considered in this talk, but:
We have a fairly complete theory + algorithms
+ patents + modem

Wireless channel is modeled as pseudodifferential operator. Sparse representation of pseudodifferential operators gives if-and-only-if characterization of optimal class of transmission pulses.

Leads to efficient equalization (by embedding of spreading functions on Heisenberg group, the twisted convolution turns into regular convolution; use of *nonabelian FFTs*,...).

Modem designed for short radio wave communications, partially based on this framework, significantly outperforms other modems on the market.

Large delay spread, deep fading channels, low signal power



Current equalization methods (ML, even MMSE) can become prohibitively expensive for large delay spread channels, despite great progress such as Hassibi's fast spherical decoder.

Brute-force approaches to "reduce" delay spread: truncate channel impulse response h or use only N largest taps of h .

Alternatives to deal with large delay spread

- OFDM:

Advantage of OFDM: simple equalization, ...

Disadvantages: large PAR, sensible to carrier offset, ...

What if disadvantages dominate and we cannot (or do not want to) use OFDM?

Can we speed up and/or improve MMSE equalizers?

- IIR equalizers: only efficient if inverse of channel matrix H has fast off-diagonal decay
- Block transmission
- FIR equalizers

Mathematical backbone for equalizer design:
Banach algebras of matrices with off-diagonal
decay

Sketch of Theorem:[Baskakov '90, '97] Let
 $A = [A_{k,l}]_{k,l \in \mathcal{I}}$ be a matrix from $\ell^2(\mathcal{I})$ to $\ell^2(\mathcal{I})$.
Assume that

$$|A_{k,l}| \leq d_\alpha(|k - l|)$$

where $d_\alpha(x)$ is a “decaying” function, such as
 $d_\alpha(x) = e^{-\alpha|x|}$ or $d_\alpha(x) = (1 + |x|)^\alpha$, $\alpha > 0$.
Such matrices form an inversely closed Banach
algebra, i.e.,

$$|(AB)_{k,l}| \leq d_\beta(|k - l|)$$

and

$$|(A^{-1})_{k,l}| \leq d_\gamma(|k - l|),$$

where β and γ depend on α and on $\text{cond}(A)$
(also true for pseudo-inverse).

Large delay spread + deep fades destroy decay
of H^{-1} .

Basic setup

After A/D conversion at receiver, received sample at time index k is

$$y(k) = \sum_{n=1}^L h(L-n)x(k-L+n)$$

In matrix notation $y^{(k)} = Hx^{(k)} + \varepsilon$, where

$$H = \begin{bmatrix} h(L-1) & h(L-2) & \dots & h(0) & \dots & 0 \\ 0 & h(L-1) & h(L-2) & \dots & h(0) & 0 \\ \vdots & \ddots & & \ddots & & 0 \\ 0 & \dots & 0 & h(L-1) & \dots & h(0) \end{bmatrix}$$

$$x^{(k)} = [x(k-L+1), \dots, x(k+N-1)]^T,$$
$$y^{(k)} = [y(k), \dots, y(k+N-1)]^T, \varepsilon \text{ is AWGN.}$$

H is $(N+L-1) \times N$ matrix, thus the system is underdetermined and x cannot be recovered from y .

Two possibilities to overcome this problem:

(i) precoding: choose $(N' + L - 1) \times N'$ matrix A , and let $N = N' + L - 1$, such that $\text{ran}(A) = \text{null}(H^+)^\perp$, and transmit Ax instead of x .

Simple choice $A = \begin{bmatrix} I \\ 0 \end{bmatrix}$ which is just zero-padding and renders H into tall matrix or into circulant matrix C_H .

Loss of spectral efficiency as in OFDM

(ii) Oversampling or antenna area at receiver: allows for FIR equalizers (no longer optimum w.r.t. AWGN).

Block transmission with zero-padding

Approach via circulant matrix [Falconer, Giannakis,...]: zero-padding allows to replace H by circulant matrix C . MMSE equalizer (with respect to matrix C !) computes

$$x_\sigma = P_N(C_\sigma^\dagger y_\sigma),$$

where $C_\sigma^\dagger = (C^*C + \sigma^2 I)^{-1}C^*$

Computational costs: $\mathcal{O}((N + L) \log(N + L))$ where L is number of taps, N is symbol block length

This is standard but is this a good idea?

No, modern numerical analysis provides better approach

Better approach

Note that zero-padding transforms H into lower triangular tall Toeplitz matrix. Solve $Hx = y$ iteratively via conjugate gradient (CG) method.

Matrix-vector multiplications Hx_n and H^*z_n can be done by FFTs

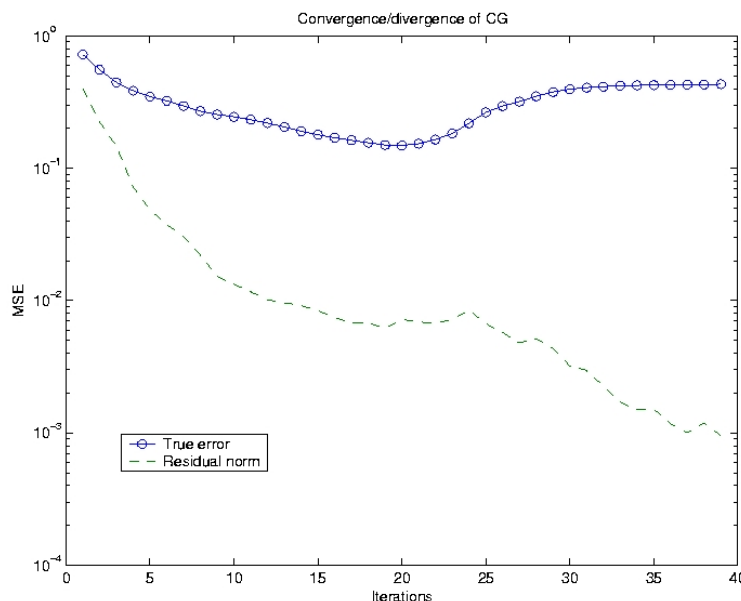
Why should this be better? It is definitely not faster.

In theory CG converges to H^+y_σ , which is the zero-forcing solution. But we want MMSE type solution.

Important fact: CG is a regularization method, i.e., it converges first to direction of singular vectors associated to large singular values, and later to singular vectors of small singular values.

Thus: use number of iterations as stopping criterion.

But: for noisy right-hand-sides monotone convergence is lost



Right choice of stopping criterion is delicate and crucial for success of CG!

We have found such stopping criteria

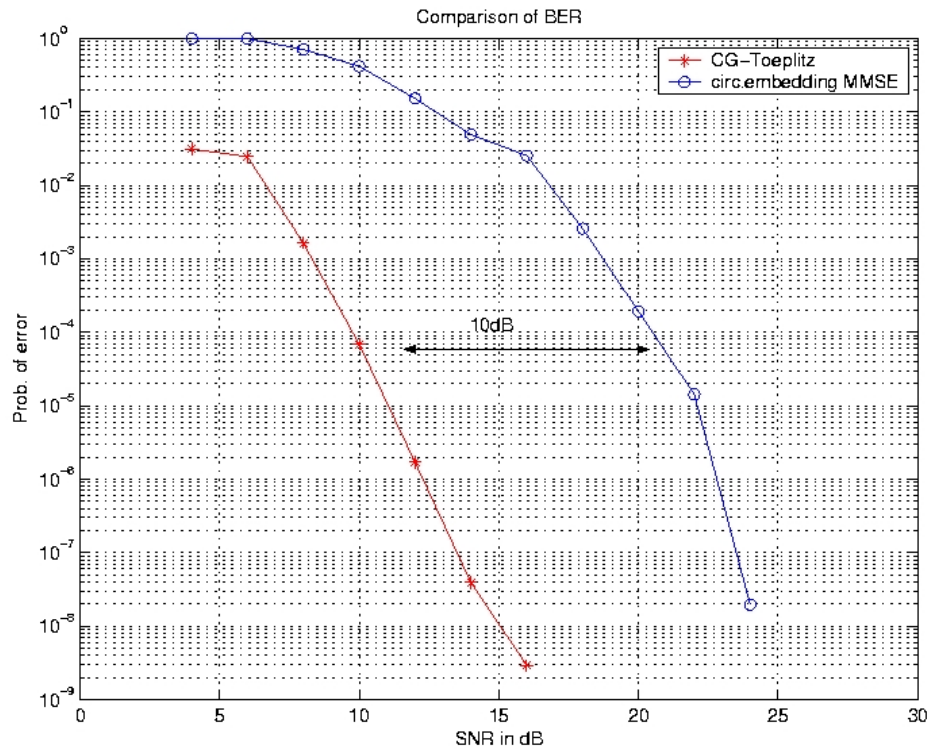
Other important observations:

(i) Circulant matrix C is a square matrix, thus its left-inverse has no null-space. Left-inverse of H has null-space, thus fraction of noise residing in null space gets completely eliminated.

(ii) C has larger condition number than H . Proof via Cauchy's interlace theorem.

Thus for noisy, deep fading channels CG-approach should outperform standard circulant-matrix based equalization.

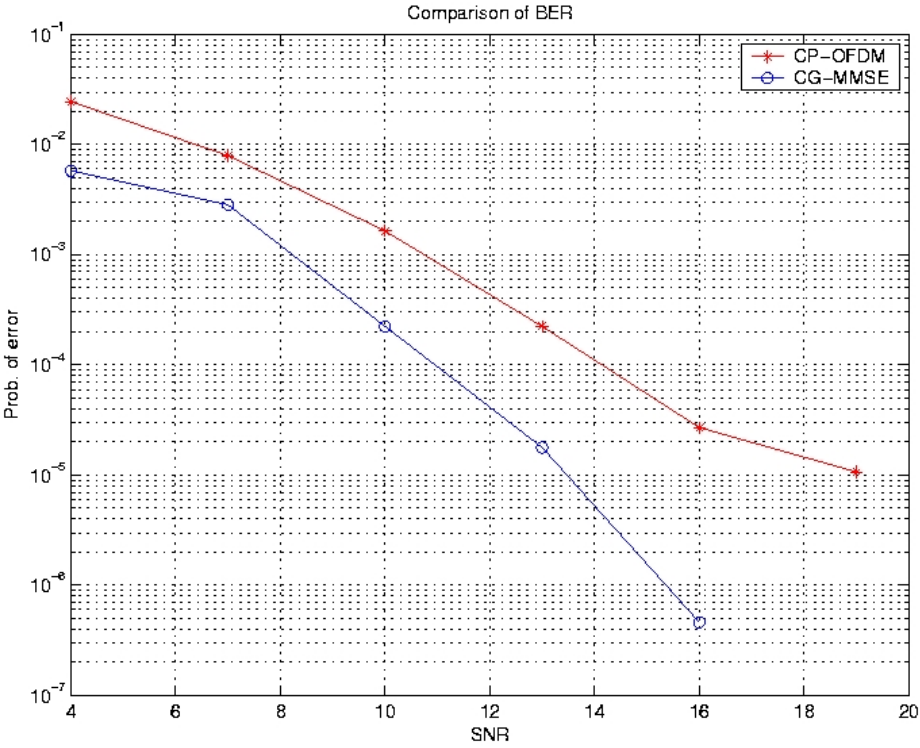
Comparison of two MMSE equalization schemes



Nice fact: CG needs less iterations with increasing noise, often only 2-3 iterations are enough. Therefore computational costs are $\mathcal{O}(K(N+L)\log(N+L))$, where K is often very small.

Additional right preconditioning can be used.

Comparison of CP-OFDM and CG-MMSE



Further speed-up of CG-Toeplitz approach

Gohberg-Semencul formula for inverse Toeplitz matrix:

Let T be invertible, hermitian Toeplitz matrix. T^{-1} is not Toeplitz, but

$$T^{-1} = L_1^* L_1 - L_2^* L_2$$

where L_1 and L_2 are lower triangular Toeplitz matrices with entries taken from the vector a where $a = T^{-1}e_1$. Solve $Ta = e_1$ via CG.

Hence storage requirement for T^{-1} is n and $T^{-1}x$ can be computed via FFTs.

CG method combined with Gohberg-Semencul formula this gives a very fast, storage-efficient, robust equalization method.

Patent pending ...

FIR equalization, oversampling, SIMO

Well-known: oversampling or multiple receive antennas allows for FIR equalizers. Standard FIR equalization leads to noise enhancement:

(i) nullspace of H^* has to be utilized to construct FIR filter. Also condition number of receive FIR filter bank may become large. For ill-conditioned H and large delay spread, this can be serious drawback.

(ii) MMSE FIR equalizer is optimum among all FIR filters (of same length), but filter structure does not lead to optimum MMSE equalization.

Two possibilities to improve FIR equalizer performance and keep linear structure:

(i) Use similar approach as for block transmission (CG+Gohberg-Semencul)

(ii) use joint equalization (pre-equalization + post-equalization)

Even better: combine (i) and (ii)

CG approach to FIR equalization

CG approach can be easily extended to case of oversampling at receiver as well as to multiple receive antennas.

Scalar entries of Toeplitz matrices are replaced by (non-commuting) matrices, thus we get block-Toeplitz-type matrices.

Gohberg-Semencul formula can be extended to block-Toeplitz matrices.

While in block transmission H^*H was Toeplitz, here H^*H is *not* block Toeplitz.

Use other form of left inverse:

$$H_{\sigma}^{\dagger} = (H^*H + \sigma I)^{-1}H^* = H^*(HH^* + \sigma I)^{\dagger}.$$

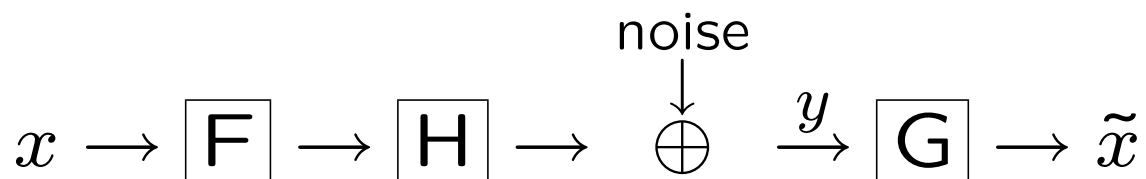
One can show that HH^* is block-Toeplitz.

Rest: straightforward

Precoding / Joint equalization

Well-known idea: can move equalizer from receiver to transmitter: precoding/pre-equalizer.

Or: can have equalizer at transmitter and at receiver: joint equalization (joint precoding: [Scaglione et al., '02])



F: pre-equalizer G: post-equalizer

Assume power constraint at transmitter $P \leq 1$.
Let λ_{\min} be minimum singular value of H .

Ignoring G the MSE-optimum precoder F is

$$F = H^*(HH^* + \nu^2 I)^{-1}, \quad \nu = \frac{\sigma^2}{P} = \sigma^2$$

We have $\tilde{x} = HFx + \varepsilon$ and the MSE is bounded by

$$\|x - \tilde{x}\|_2 \leq \frac{\lambda_{\min}(H)^2}{\lambda_{\min}(H)^2 + \sigma^2} + \sigma$$

Compare this to MMSE equalizer (no F , only G): $\tilde{x} = H_\sigma^\dagger Hx + H_\sigma^\dagger \varepsilon$ and the MSE is bounded by

$$\|x - \tilde{x}\|_2 \leq \frac{\lambda_{\min}(H)^2}{\lambda_{\min}(H)^2 + \sigma^2} + \min\left\{\frac{1}{2}, \frac{\lambda_{\min}(H)\sigma}{\lambda_{\min}^2(H) + \sigma^2}\right\}$$

Thus for ill-conditioned H and large noise, post-equalization results in worse performance.

Joint equalization: offers two potential benefits: (i) better MSE performance (ii) better FIR filters

For given F the MSE-optimum G is

$$G = [(HF)^*(HF) + \lambda^2 I]^{-1} (HF)^*.$$

$\tilde{x} = G(HFx + \varepsilon)$ and the MSE is bounded by

$$\|x - \tilde{x}\|_2 \leq \frac{\lambda_{\min}^2(HF)}{\lambda_{\min}^2(HF) + \sigma^2} + \min\left\{\frac{1}{2}, \frac{\lambda_{\min}(HF)\sigma}{\lambda_{\min}^2(HF) + \sigma^2}\right\}$$

When is this better?

Good rule of thumb (backed up by theory): if

$$\lambda_{\min}(H) + \sigma^2 < 1$$

joint equalization is better.

Joint equalization helps FIR equalizer design

Want receive equalizer to have limited complexity. Measure complexity of equalizer by number of taps.

Thus: want receive equalizer to have only R taps. Hence: have to find F, G such that

$$\|G(HFx + \varepsilon) - x\|_2 = \min!$$

subject to $\text{trace}(F^*F) \leq P$ and G is R -banded matrix.

This leads to complicated non-convex optimization problem. Even if it can be solved, its computational costs are too high for practice.

Different approach:

Matrix algebra theory gives insight in difficulties and how to proceed: need F and HF to have as small condition number as possible.

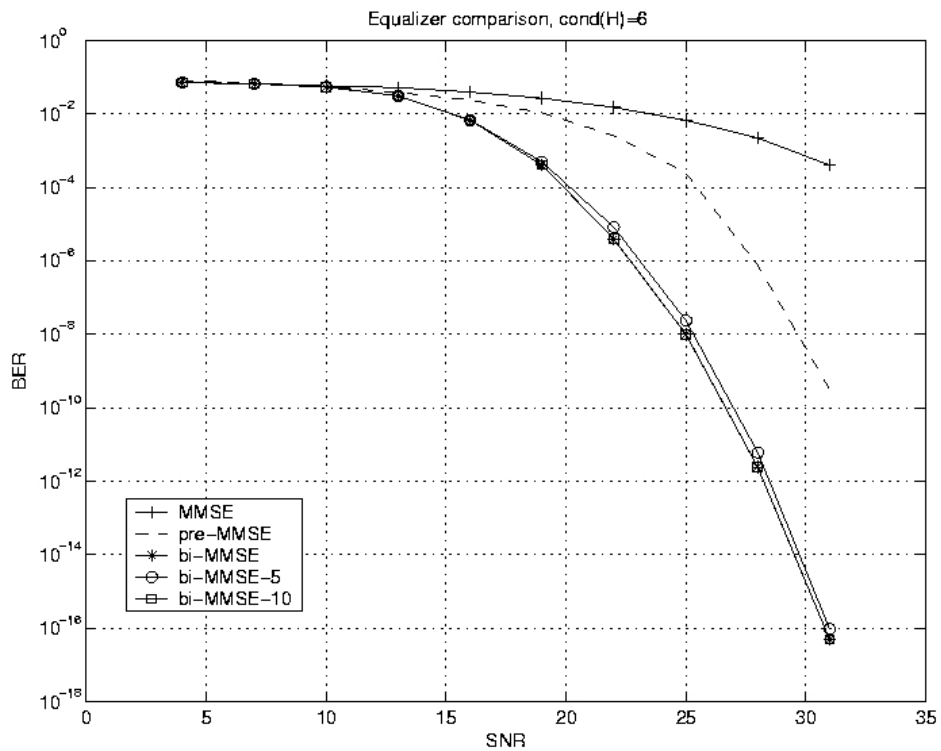
But if H is ill-conditioned and noise is large we cannot have $\text{cond}(F)$ and $\text{cond}(HF)$ small.

Further difficulties:

CSI may not be as accurate at transmitter.

Part of computational burden is shifted to transmitter, but we still need computationally efficient methods.

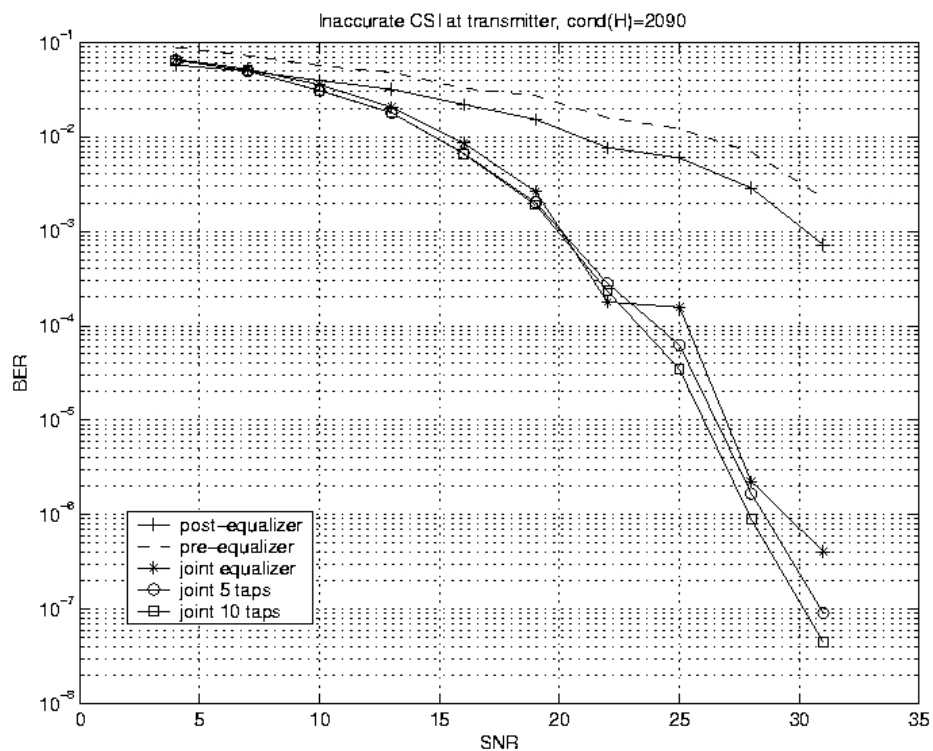
We can use fast CG-Toeplitz algorithms to implement joint equalizers



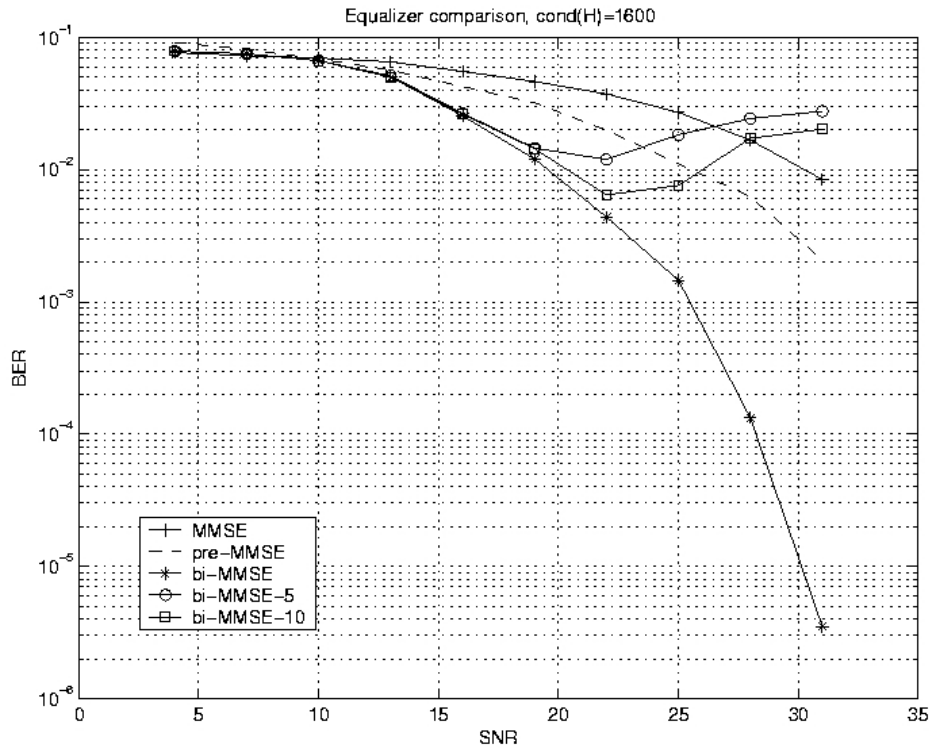
Joint equalization clearly outperforms the other two.

For well-conditioned channels essentially no loss for limited-tap receive equalizers.

Assume that CSI at transmitter is less precise than CSI at receiver.



In this case precoding alone is not very efficient, while joint equalization even with constraint number of taps works still fine.



Delicate behavior of limited-tap equalizers due to behavior of $\text{cond}(F)$ and $\text{cond}(HF)$.

Conclusion

Next generation equalizers for ill-conditioned channels with large delay spread and large AWGN:

- Gohberg-Semencul-type formulas give rise to filter-like structure for equalizer, but with much better robustness
- CG methods allow for fast numerical implementation with minimum storage requirement
- Joint precoding done right allows for simple receive equalizers

Thus: OFDM is no longer the only choice for channels with large delay spread