Mob-Com Dept. Internal Seminar Series

Efficient Turbo-synchronization algorithms for phase noise and frequency offsets.

Giuseppe Caire, Eurecom Institute, France

Eurecom, September 23, 2004

Problem

- AWGN channel with phase noise and frequency offset

$$y_k = x_k e^{j(\theta_k + 2\pi\nu k)} + w_k, \quad k = 0, \dots, N-1 .$$

- The sequence $\mathbf{x} = (x_0, \dots, x_{N-1})$ is a codeword of the channel code $\mathcal{C}$ constructed over an M-ary modulation constellation $\mathcal{X} \subset \mathbb{C}$.

- We include pilot symbols, interleaving, preambles etc etc ... as part of the code.

- The channel SNR is given by $\mathsf{SNR} = \mathbb{E}[|x|^2]/N_0$, where x is uniform over $\mathcal{X}$.

Phase noise

- $\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_{N-1})$ is random, unknown to both transmitter and receiver, and statistically independent of $\mathbf{x}$ and $\mathbf{w}$.

- WORKING ASSUMPTION:

$$\theta_k = \theta_{k-1} + \Delta_k$$

where $\{\Delta_k\}$ is a white real Gaussian process with $\Delta_k \sim \mathcal{N}(0, \sigma_\Delta^2)$.

- Under this assumption, assuming $\theta_0 \sim$ Uniform$[0, 2\pi)$,

$$p(\theta_k | \theta_{k-1}, \theta_{k-2}, \ldots, \theta_0) = p(\theta_k | \theta_{k-1}) = p_\Delta(\theta_k - \theta_{k-1})$$

where $p_\Delta(\phi)$ is the pdf of $\Delta_k \bmod [0, 2\pi)$.

Frequency offset

- ν is assumed to be uniformly distributed in $[-\nu_{\max}, \nu_{\max}]$.
- $\nu_{\max} = f_{\max}T$, where $f_{\max}$ is the frequency specified tolerance and T is the symbol period.
- The one-sample per symbol model makes sense only if $\nu_{\max} \ll 1/2$ (for large frequency offset ISI appears).

Optimal bit-wise decoding

- Focus first on the case of phase noise only ($\nu_{\max} = 0$).
- The code $\mathcal{C}$ admits an encoding function $\mu_{\mathcal{C}} : \mathbb{F}_2^K \to \mathcal{X}^N$, mapping binary information messages $\mathbf{b} \in \mathbb{F}_2^K$ into the codewords.
- Optimal decision rule that minimizes the average bit-error probability:

$$\widehat{b}_i = \arg \max_{b \in \mathbb{F}_2} P_{\text{bit},i}(b|\mathbf{y})$$

Posterior probability marginalization $\Rightarrow$ BP

- Let $P(\mathbf{b}, \boldsymbol{\theta}|\mathbf{y})$ denote the joint posterior probability distribution function of the information bits and of the phase noise vector $\boldsymbol{\theta}$ given $\mathbf{y}$.

- The desired $P_{\text{bit},i}(b|\mathbf{y})$ can be obtained by marginalizing $P(\mathbf{b}, \boldsymbol{\theta}|\mathbf{y})$ with respect to $\boldsymbol{\theta}$ and to all b_j for $j \neq i$.

- This can be accomplished in an approximated but low-complexity way by BP applied on the FG of $P(\mathbf{b}, \boldsymbol{\theta}|\mathbf{y})$.

Factorization (1)

- $$\begin{aligned} P(\mathbf{b},\boldsymbol{\theta}|\mathbf{y}) &\propto P(\mathbf{b})p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta},\mathbf{b}) \\ &\propto \chi[\mathbf{x}=\mu_{\mathcal{C}}(\mathbf{b})]p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta},\mathbf{x}=\mu_{\mathcal{C}}(\mathbf{b})) \\ &\propto \chi[\mathbf{x}=\mu_{\mathcal{C}}(\mathbf{b})]p(\boldsymbol{\theta})\prod_{k=0}^{N-1} p(y_k|x_k,\theta_k) \\ &\propto \chi[\mathbf{x}=\mu_{\mathcal{C}}(\mathbf{b})]p(\boldsymbol{\theta})\prod_{k=0}^{N-1} f_k(x_k,\theta_k) \end{aligned}$$

- We have defined the functions:

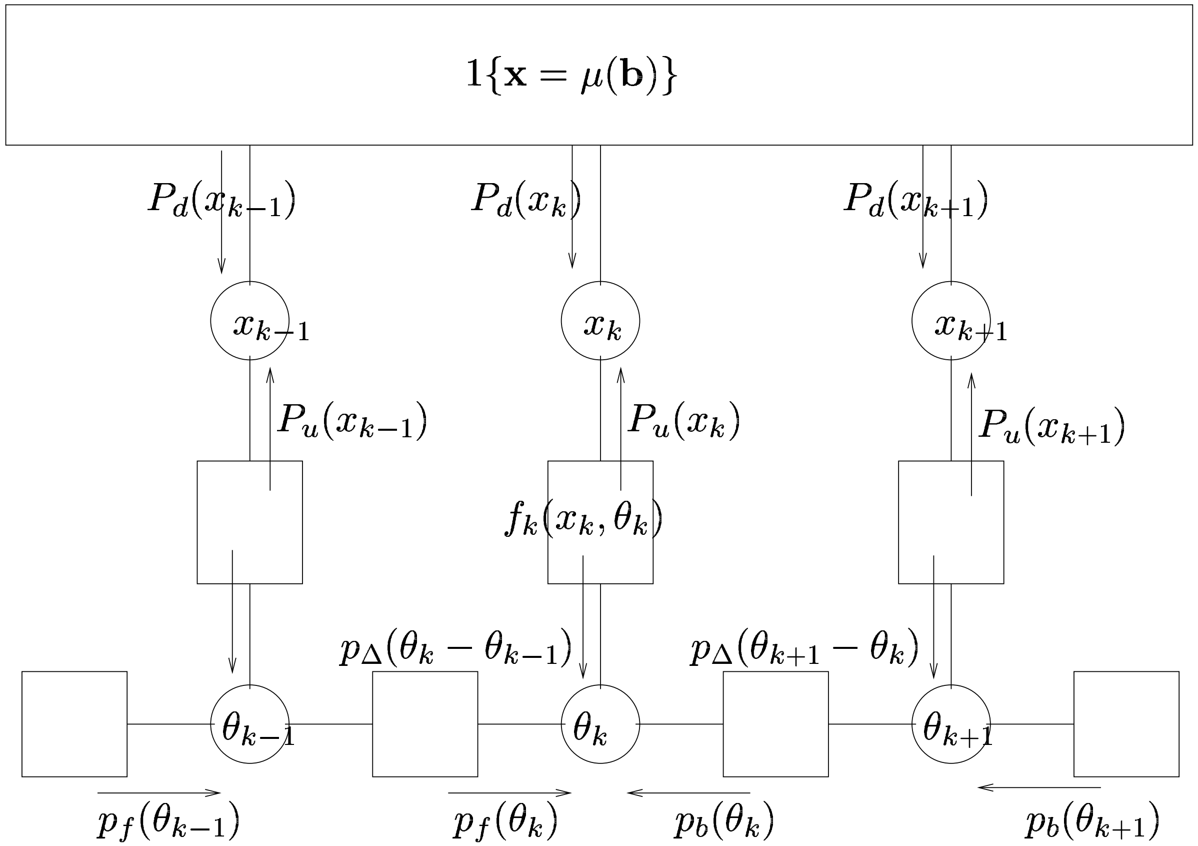
$$f_k(x_k,\theta_k) \overset{\text{def}}{=} \exp\left\{\frac{2}{N_0}\text{Re}[y_k x_k^* e^{-j\theta_k}] - \frac{|x_k|^2}{N_0}\right\} \propto \exp\left\{-\frac{1}{N_0}|y_k - x_k e^{j\theta_k}|^2\right\}.$$

Factorization (2)

Under the assumption of 1st order Markov model for the phase noise:

$$P(\mathbf{b}, \boldsymbol{\theta}|\mathbf{y}) \propto \chi[\mathbf{x} = \mu_{\mathcal{C}}(\mathbf{b})]p(\theta_0) \prod_{k=1}^{N-1} p_{\Delta}(\theta_k - \theta_{k-1}) \prod_{k=0}^{N-1} f_k(x_k, \theta_k)\,.$$

Factor Graph



Belief Propagation (1)

- $P_d(x_k)$: message from variable node x_k to factor node f_k (decoder soft-output);
- $P_u(x_k)$: message from factor node f_k to variable node x_k (decoder soft-input);
- $p_d(\theta_k)$: message from factor node f_k to variable node θ_k:

$$p_d(\theta_k) \propto \sum_{x \in \mathcal{X}} P_d(x_k = x) f_k(x_k = x, \theta_k) \, .$$

- We assume that in the lower part of the FG a forward-backward node activation schedule is adopted.

Belief Propagation (2)

- Messages $p_f(\theta_k)$ from factor node $p_\Delta(\theta_k - \theta_{k-1})$ to variable node θ_k, and $p_b(\theta_k)$ from factor node $p_\Delta(\theta_{k+1} - \theta_k)$ to variable node θ_k, can be recursively computed as follows:

$$p_f(\theta_k) \propto \int_0^{2\pi} p_d(\theta_{k-1}) p_f(\theta_{k-1}) p_\Delta(\theta_k - \theta_{k-1})\, d\theta_{k-1}$$

$$p_b(\theta_k) \propto \int_0^{2\pi} p_d(\theta_{k+1}) p_b(\theta_{k+1}) p_\Delta(\theta_{k+1} - \theta_k)\, d\theta_{k+1}$$

- The message $P_u(x_k)$ from f_k to x_k is given by

$$P_u(x_k) \propto \int_0^{2\pi} p_f(\theta_k) p_b(\theta_k) f_k(x_k, \theta_k)\, d\theta_k\,.$$

Belief Propagation (3)

- Decoder output: $\{P_d(x_k) : k = 0, \ldots, N-1\}$.
- Decoder input: $\{P_u(x_k) : k = 0, \ldots, N-1\}$.
- Pilot symbols have probability $P_d(x_k = a_k) = 1$, where a_k is the a-priori known k-th pilot value.

Approaches to practical low-complexity algorithms

- Discretization: propagate pmfs instead of pdfs.
- Canonical distributions: propagate the parameters of the distributions.
- Remark: There are other standard approaches in the literature, e.g., particle filters (generate Monte Carlo samples from a distribution).

Discretization

- We assume that the channel phase θ_k may take $\Theta = \{0, 2\pi/L, \ldots, 2\pi(L-1)/L\}$.

- This approach becomes "optimal" (in the sense that it approaches the performance of the exact BP) for a sufficiently large number of discretization levels.

- Rule of thumb: for M-PSK signals, $L = 8M$ values are sufficient to have no practical performance loss w.r.t. exact BP.

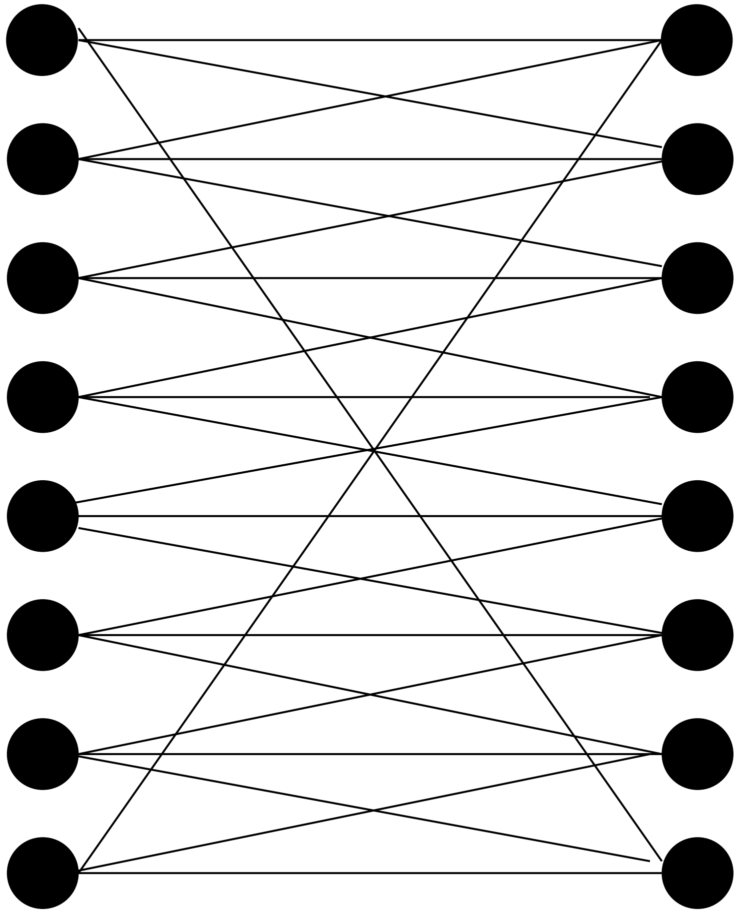
- Main limitation: complexity (it is still quite computationally intensive).

Phase trellis (1)

- State space Θ of size L, isomorphic to $\mathbb{Z}_L$.
- A sensible choice for the phase dynamics is to assume that state ℓ_k has transitions with non-zero probability only to states $\ell_{k+1} = \ell_k$, $\ell_{k+1} = \ell_k + 1$, $\ell_{k+1} = \ell_k - 1$, modulo L.
- By symmetry of the phase noise distribution, we have that

$$P(\ell_k \to \ell_{k+1}) = \begin{cases} P_\Delta/2 & \ell_{k+1} = \ell_k - 1 \\ 1 - P_\Delta & \ell_{k+1} = \ell_k \\ P_\Delta/2 & \ell_{k+1} = \ell_k + 1 \end{cases}$$

Phase trellis (2)



Matching the discrete process variance

- The value of P_Δ is chosen such that the variance of the phase increment is equal to the variance of the phase difference of the continuous (non-discretized) phase noise process.

- For example, for the Wiener model defined before we have

$$\sigma^2_{\mathrm{discr}} = P_\Delta \left(\frac{2\pi}{L}\right)^2$$

and by letting $\sigma^2_{\mathrm{discr}} = \sigma^2_\Delta$ we obtain

$$P_\Delta = \left(\frac{\sigma_\Delta L}{2\pi}\right)^2$$

Phase BCJR: definitions

- Denote the message from $p(\theta_k|\theta_{k-1})$ to θ_k, in the log-probability domain, as $\{\alpha_k(\ell) : \ell \in \mathbb{Z}_L\}$.
- Denote the message from $p(\theta_{k+1}|\theta_k)$ to θ_k, in the log-probability domain, as $\{\beta_k(\ell) : \ell \in \mathbb{Z}_L\}$.
- The branch weight for the trellis section $(k, k+1)$ is given by

$$\gamma_k(\ell, m) = \text{LogSum}_{x \in \mathcal{X}} \left\{ \log P_d(x_k = x) - \frac{1}{N_0} \left| y_k - x e^{-j2\pi\ell/L} \right|^2 \right\} + \log P(\ell \to m)$$

where $P(\ell \to m)$ is the trellis transition probability $\ell_k = \ell$ and $\ell_{k+1} = m$.

Phase BCJR: forward recursion

- Initialize $\alpha_0(\ell) = 0$ for all $\ell \in \mathbb{Z}_L$.
- For $k = 0, \ldots, N-1$ let

$$\alpha_{k+1}(m) = \mathrm{LogSum}_{\ell \in \mathbb{Z}_L} \; \{\gamma_k(\ell, m) + \alpha_k(\ell)\}$$

Phase BCJR: backward recursion

- Initialize $\beta_{N-1}(\ell) = 0$ for all $\ell \in \mathbb{Z}_L$.
- For $k = N-1, \ldots, 1$ let

$$\beta_{k-1}(m) = \mathrm{LogSum}_{\ell \in \mathbb{Z}_L} \ \{\gamma_k(\ell, m) + \beta_k(\ell)\}$$

Phase BCJR: output

Finally, the BCJR output is given by the log-probabilities $\{\eta_k(x) : x \in \mathcal{X}\}$, with

$$\eta_k(x) = \text{LogSum}_{\ell \in \mathbb{Z}_L} \left\{ \alpha_k(\ell) + \beta_k(\ell) - \frac{1}{N_0} \left| y_k - x e^{-j2\pi\ell/L} \right|^2 \right\}$$

for $k = 0, \ldots, N-1$.

Gaussian approximation of $p_d(\theta_k)$

- If the messages $P_d(x_k)$ were the exact a posteriori probabilities of the code symbols, it would be

$$p_d(\theta_k) \propto \sum_{x \in \mathcal{X}} P_d(x_k = x) f_k(x_k = x, \theta_k) \propto p(y_k|\theta_k)\,.$$

The pdf $p(y_k|\theta_k) = \sum_{x \in \mathcal{X}} P_d(x_k = x) g_{\mathbb{C}}(xe^{j\theta_k}, N_0, y_k)$ is a linear combination of Gaussian pdfs.

- We approximate $p(y_k|\theta_k)$ by the Gaussian pdf at minimum *divergence*, given by $g_{\mathbb{C}}(\alpha_k e^{j\theta_k}, N_0 + \beta_k - |\alpha_k|^2, y_k)$, where

$$\alpha_k \stackrel{\text{def}}{=} \sum_{x \in \mathcal{X}} x P_d(x_k = x)$$

and

$$\beta_k \stackrel{\text{def}}{=} \sum_{x \in \mathcal{X}} |x|^2 P_d(x_k = x)$$

Tikhonov parameterization

- With the above Gaussian approximation, we obtain

$$\begin{aligned} p_d(\theta_k) &\propto p(y_k|\theta_k) \\ &\simeq g_{\mathbb{C}}\left(\alpha_k e^{j\theta_k}, N_0 + \beta_k - |\alpha_k|^2, y_k\right) \\ &\propto \exp\left\{2\frac{\mathrm{Re}[y_k \alpha_k^* e^{-j\theta_k}]}{N_0 + \beta_k - |\alpha_k|^2}\right\} \end{aligned}$$

- Substituting in the forward and backward recursion, we obtain

$$p_f(\theta_k) \simeq \int_0^{2\pi} \exp\left\{2\frac{\mathrm{Re}[y_{k-1}\alpha_{k-1}^* e^{-j\theta_{k-1}}]}{N_0 + \beta_{k-1} - |\alpha_{k-1}|^2}\right\} p_f(\theta_{k-1}) p_\Delta(\theta_k - \theta_{k-1})\, d\theta_{k-1}\,.$$

$$p_b(\theta_k) \simeq \int_0^{2\pi} \exp\left\{2\frac{\mathrm{Re}[y_{k+1}\alpha_{k+1}^* e^{-j\theta_{k+1}}]}{N_0 + \beta_{k+1} - |\alpha_{k+1}|^2}\right\} p_f(\theta_{k+1}) p_\Delta(\theta_k - \theta_{k+1})\, d\theta_{k+1}\,.$$

Facts about Tikhonov pdfs

- Let
$$t(\zeta;x) = \frac{1}{2\pi \mathrm{I}_0(|\zeta|)} e^{\mathrm{Re}[\zeta e^{-jx}]}$$
denote a Tikhonov pdf with parameter ζ.

- Fact 1.
$$t(\zeta; x+\eta) = t(\zeta e^{-j\eta}; x)$$

- Fact 2.
$$t(\zeta_1;x)t(\zeta_2;x) \propto \frac{\mathrm{I}_0(|\zeta_1+\zeta_2|)}{2\pi \mathrm{I}_0(|\zeta_1|)\mathrm{I}_0(|\zeta_2|)} t(\zeta_1+\zeta_2;x)$$

- Fact 3. Let $g(y,\sigma^2;x)$ denote a real Gaussian distribution in x with mean value y and variance σ^2, then
$$\int t(\zeta;y)g(y,\sigma^2;x)dy \simeq \frac{\mathrm{I}_0\left(\frac{|\zeta|}{1+\sigma^2|\zeta|}\right)}{\mathrm{I}_0(|\zeta|)} \frac{e^{\frac{\sigma^2|\zeta|^2}{1+\sigma^2|\zeta|}}}{\sqrt{1+\sigma^2|\zeta|}} t(\frac{\zeta}{1+\sigma^2|\zeta|};x)$$

- Fact 3 can be easily shown by direct calculation, considering the approximation

$$t(\zeta; y) \simeq \frac{e^{|\zeta|}}{\sqrt{2\pi|\zeta|}\mathrm{I}_0(|\zeta|)} g\left(\arg(\zeta), \frac{1}{|\zeta|}; y\right)$$

which holds for values of $|\zeta|$ larger then few units. This is obtained by approximating $\cos(y - \arg(\zeta)) \approx 1 - \frac{1}{2}(y - \arg(\zeta))^2$, i.e., by using its Taylor expansion truncated to the first term.

Tikhonov algorithm: forward and backward recursions

- By using the above facts, we are able to obtain forward and backward recursions in the form:

$$p_f(\theta_k) \propto \exp\left\{\mathrm{Re}[a_{f,k} e^{-j\theta_k}]\right\}, \quad p_b(\theta_k) \propto \exp\left\{\mathrm{Re}[a_{b,k} e^{-j\theta_k}]\right\}$$

- The forward and backward parameters $a_{f,k}$ and $a_{b,k}$ can be recursively computed by

$$a_{f,k} = \frac{a_{f,k-1} + u_{k-1}}{1 + \sigma_\Delta^2 |a_{f,k-1} + u_{k-1}|}$$

$$a_{b,k} = \frac{a_{b,k+1} + u_{k+1}}{1 + \sigma_\Delta^2 |a_{b,k+1} + u_{k+1}|}$$

where

$$u_k = 2\frac{y_k \alpha_k^*}{N_0 + \beta_k - |\alpha_k|^2}$$

Tikhonov algorithm: output

The ouptut of the Tikhonov algorithm is given by

$$P_u(x_k) \propto \exp\left\{-\frac{|x_k|^2}{N_0}\right\} \mathrm{I}_0\left(\left|a_{f,k} + a_{b,k} + \frac{y_k x_k^*}{N_0/2}\right|\right)$$

for all $x_k \in \mathcal{X}$.

Estimating the phasor rather than the phase

- We model the phasor process $h_k \stackrel{\text{def}}{=} e^{j\theta_k}$ as a complex circularly symmetric Gauss-Markov process.
- We treat $\mathbf{h} = (h_0, \ldots, h_{K-1})$ and $\mathbf{y}$ as jointly Gaussian, by letting

$$p(y_k|h_k) \simeq g_{\mathbb{C}}\left(\alpha_k h_k, N_0 + \beta_k - |\alpha_k|^2, y_k\right)$$

- The underlying dynamical system is given by

$$\begin{aligned} h_{k+1} &= \rho h_k + v_k \\ y_k &= \alpha_k h_k + w_k \end{aligned}$$

where $v_k \sim \mathcal{N}_{\mathbb{C}}(0, 1-\rho^2)$ and $w_k \sim \mathcal{N}_{\mathbb{C}}(0, N_0 + \beta_k - |\alpha_k|^2)$.

Some facts about Gaussian pdfs (1)

- Fact 1.

$$g_{\mathbb{C}}(A_1, \Sigma_1, x) g_{\mathbb{C}}(A_2, \Sigma_2; x) \propto g_{\mathbb{C}}\left(\frac{\Sigma_2}{\Sigma_1+\Sigma_2}A_1 + \frac{\Sigma_1}{\Sigma_1+\Sigma_2}A_2, \frac{\Sigma_1\Sigma_2}{\Sigma_1+\Sigma_2}; x\right) .$$

- Fact 2. Let (x, y) be jointly Gaussian, such that $x \sim g_{\mathbb{C}}(A_1, \Sigma_1; x)$ and $y = A_2 x + z_2$ with $z_2 \sim g_{\mathbb{C}}(0, \Sigma_2)$, independent of x. Then

$$g_{\mathbb{C}}\left(\begin{bmatrix} A_1 \\ A_1 A_2 \end{bmatrix}, \begin{bmatrix} \Sigma_1 & A_1^* \Sigma_1 \\ A_1 \Sigma_1 & \Sigma_2 + |A_1|^2 \Sigma_1 \end{bmatrix}; \begin{bmatrix} x \\ y \end{bmatrix}\right) = g_{\mathbb{C}}(A_1, \Sigma_1; x) g_{\mathbb{C}}(A_2 x, \Sigma_2; y)$$

$$= g_{\mathbb{C}}(A_1 A_2, \Sigma_2 + |A_2|^2 \Sigma_1; y) g_{\mathbb{C}}\left(A_1 + \frac{A_2^* \Sigma_1}{\Sigma_2 + |A_2|^2 \Sigma_1}(y - A_2 A_1), \frac{\Sigma_2 \Sigma_1}{\Sigma_2 + |A_2|^2 \Sigma_1}; x\right)$$

Some facts about Gaussian pdfs (2)

It follows that the marginal pdf of y is given by

$$\int g_{\mathbb{C}}(A_1, \Sigma_1; x) g_{\mathbb{C}}(A_2 x, \Sigma_2; y) dx = g_{\mathbb{C}}\left(A_1 A_2, \Sigma_2 + |A_2|^2 \Sigma_1; y\right) .$$

and that $\mathbb{E}[x|y]$ (MMSE estimation of x given y) is given by

$$A_1 + \frac{A_2^* \Sigma_1}{\Sigma_2 + |A_2|^2 \Sigma_1}(y - A_2 A_1)$$

Moreover, $\mathbb{E}[x|y]$ is Gaussian with variance

$$\frac{\Sigma_2 \Sigma_1}{\Sigma_2 + |A_2|^2 \Sigma_1}$$

Some facts about Gaussian pdfs (3)

- Fact 3. Let $X \sim \mathcal{N}_{\mathbb{C}}(A, \Sigma)$ be expressed in magnitude and phase as $X = Re^{j\theta}$. Hence, the joint pdf of R and θ is given by

$$f(R, \theta) = \frac{1}{\pi\Sigma} R \exp\left(-\frac{R^2 + |A|^2}{\Sigma}\right) \exp\left(\frac{2R|A|\cos(\theta - \phi)}{\Sigma}\right)$$

where $A = |A|e^{j\phi}$.

- Fact 4. Let $a, b \in \mathbb{R}_+$. We have,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left(a\cos(\theta - \alpha) + b\cos(\theta - \beta)\right) d\theta = I_0\left(\sqrt{a^2 + b^2 + 2ab\cos(\alpha - \beta)}\right)$$

where $I_0(z)$ is the modified Bessel function of the first kind and order zero.

Kalman algorithm: forward recursion (1)

- Let $m_{k|k-1} = \mathbb{E}[h_k|\{y_j : j = 0, \ldots, k-1\}], \Sigma_{k|k-1} = \text{var}(h_k|\{y_j : j = 0, \ldots, k-1\}$ (prediction).

- Let $m_{k|k} = \mathbb{E}[h_k|\{y_j : j = 0, \ldots, k\}]$ and $\Sigma_{k|k} = \text{var}(h_k|\{y_j : j = 0, \ldots, k\})$ (filtering).

- We let

$$p_{f,k}(h_k) \propto g_{\mathbb{C}}(m_{k|k-1}, \Sigma_{k|k-1}; h_k)$$

The forward recursion takes on the form ...

Kalman algorithm: forward recursion (2)

$$
\begin{aligned}
p_{f,k+1}(h_{k+1}) &= g_{\mathbb{C}}(m_{k+1|k}, \Sigma_{k+1|k}; h_{k+1}) \\
&\propto \int g_{\mathbb{C}}(m_{k|k-1}, \Sigma_{k|k-1}; h_k) \cdot \\
&\quad \cdot g_{\mathbb{C}}(\alpha_k h_k, N_0 + \beta_k - |\alpha_k|^2; y_k) g_{\mathbb{C}}(\rho h_k, 1 - \rho^2; h_{k+1}) \, dh_k \\
&\overset{(a)}{=} g_{\mathbb{C}}(\alpha_k m_{k|k-1}, |\alpha_k|^2 (\Sigma_{k|k-1} - 1) + N_0 + \beta_k; y_k) \cdot \\
&\quad \cdot \int g_{\mathbb{C}}(m_{k|k}, \Sigma_{k|k}; h_k) g_{\mathbb{C}}(\rho h_k, 1 - \rho^2; h_{k+1}) \, dh_k \\
&\propto g_{\mathbb{C}}(\rho m_{k|k}, \rho^2 (\Sigma_{k|k} - 1) + 1; h_{k+1})
\end{aligned}
$$

Kalman algorithm: forward recursion (3)

- (a) follows from Fact 2, the last line follows from Fact 1 and where we have identified

$$m_{k|k} = m_{k|k-1} + \frac{\alpha_k^* \Sigma_{k|k-1}}{|\alpha_k|^2(\Sigma_{k|k-1} - 1) + N_0 + \beta_k}(y_k - \alpha_k m_{k|k-1})$$

and

$$\Sigma_{k|k} = \frac{(N_0 + \beta_k - |\alpha_k|^2)\Sigma_{k|k-1}}{\Sigma_{k|k-1}(1 - |\alpha_k|^2) + N_0 + \beta_k}$$

by the MMSE estimation property stated in Fact 2.

Kalman algorithm: forward recursion summary

- $$\begin{aligned} m_{k|k} &= m_{k|k-1} + \frac{\Sigma_{k|k-1}\alpha_k^*}{|\alpha_k|^2(\Sigma_{k|k-1}-1)+N_0+\beta_k}\left(y_k - \alpha_k m_{k|k-1}\right) \\ \Sigma_{k|k} &= \frac{N_0+\beta_k-|\alpha_k|^2}{|\alpha_k|^2(\Sigma_{k|k-1}-1)+N_0+\beta_k}\Sigma_{k|k-1} \\ m_{k+1|k} &= \rho m_{k|k} \\ \Sigma_{k+1|k} &= \rho^2(\Sigma_{k|k}-1)+1 \end{aligned}$$

  for $k = 0, \ldots, N-1$, with initial conditions $\Sigma_{0|-1} = 1$ and $m_{0|-1} = 0$.

Kalman algorithm: backward recursion summary

- $$\begin{aligned}
\mu_{k|k} &= \mu_{k|k+1} + \frac{\Xi_{k|k+1}\alpha_k^*}{|\alpha_k|^2(\Xi_{k|k+1} - 1) + N_0 + \beta_k}\left(y_k - \alpha_k \mu_{k|k+1}\right) \\
\Xi_{k|k} &= \frac{N_0 + \beta_k - |\alpha_k|^2}{|\alpha_k|^2(\Xi_{k|k+1} - 1) + N_0 + \beta_k}\Xi_{k|k+1} \\
\mu_{k-1|k} &= \rho\mu_{k|k} \\
\Xi_{k-1|k} &= \rho^2(\Xi_{k|k} - 1) + 1
\end{aligned}$$

  for $k = N-1, \ldots, 0$, with initial conditions $\Xi_{N-1|N} = 1$ and $\mu_{N-1|N} = 0$.

Kalman algorithm: output

- The smoothed estimate is given by

$$m_k = \frac{\Xi_{k|k+1}}{\Sigma_{k|k-1} + \Xi_{k|k+1}} m_{k|k-1} + \frac{\Sigma_{k|k-1}}{\Sigma_{k|k-1} + \Xi_{k|k+1}} \mu_{k|k+1}$$

$$\Sigma_k = \frac{\Sigma_{k|k-1} \Xi_{k|k+1}}{\Sigma_{k|k-1} + \Xi_{k|k+1}}.$$

- We let $h_k = R_k e^{j\theta_k}$ and we obtain

$$P_u(x_k) \propto \int \exp\left(-\frac{1}{N_0}|y_k - x_k e^{j\theta_k}|^2\right) g_{\mathbb{C}}\left(m_k, \Sigma_k, R_k e^{j\theta_k}\right) R_k \, dR_k \, d\theta_k$$

$$\propto e^{-\frac{|x_k|^2}{N_0}} \int_0^\infty e^{-z} I_0\left(2\sqrt{\frac{|y_k x_k^*|^2}{N_0^2} + \frac{z|m_k|^2}{\Sigma_k} + 2\frac{\sqrt{z}}{\sqrt{\Sigma_k} N_0} \mathrm{Re}\{y_k x_k^* m_k^*\}}\right) dz$$

Handling constant carrier frequency offsets

- We propose to handle phase and frequency in two different ways: using canonical parameterization for the phase and discretization for the frequency.
- Frequency is constrained to take values on the grid of points (frequency "states"):

$$\mathcal{V} = \{\nu_\ell = \ell\Delta_\nu - \nu_{\max} \ : \ \ell = 0, \ldots, 2M\}$$

with $\Delta_\nu = \nu_{\max}/M$.

Brute-force approach: parallel decoders

- Estimate $P_{\text{bit},i}(b|\mathbf{y}, \nu_\ell)$ for all $\ell = 0, \ldots, 2M$ by applying any of the BP approximations seen before conditionally to the hypothesis $\nu = \nu_\ell$.
- The symbol-by-symbol decisions on the information bits are made using the probabilities

$$P_{\text{bit},i}(b|\mathbf{y}) \approx \frac{1}{2M} \sum_{\ell=0}^{2M} P_{\text{bit},i}(b|\mathbf{y}, \nu_\ell)$$

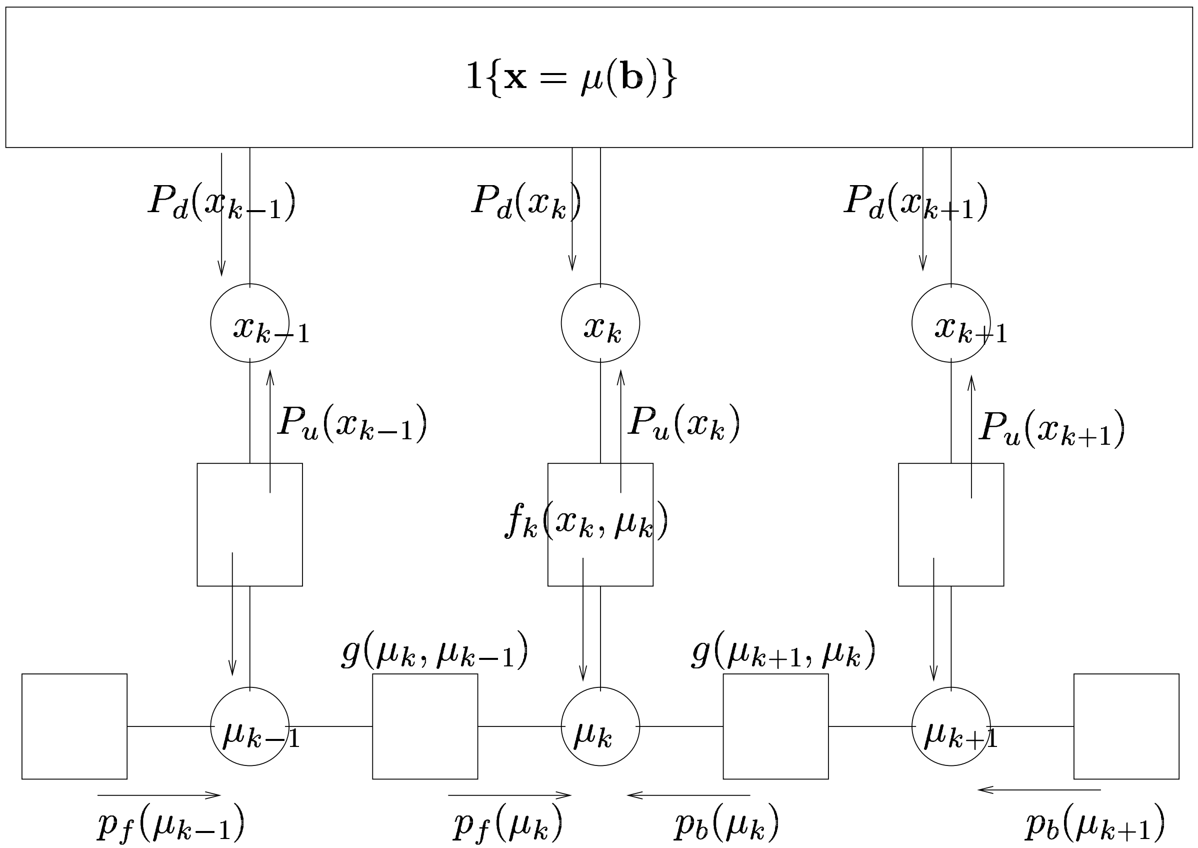
- This amounts to running $2M + 1$ decoders in parallel, one for each frequency value, and combining their soft-output symbol-by-symbol decision metrics.
- It might be an attractive approach if some pre-estimation technique yields a small set of frequency values containing the true value with high probability!

Low-complexity approaches: parallel detectors, single decoder

- We define a new channel state variable $\mu_k = (\theta_k, \nu_k)$.
- We define the function nodes

$$g_k(\theta_k, \theta_{k-1}, \nu_k, \nu_{k-1}) = p_\Delta(\theta_k - \theta_{k-1}) 1\{\nu_k = \nu_{k-1}\}$$

Factor graph for phase and frequency



The parallel Tikhonov algorithm (1)

- Applying BP we have

$$
\begin{aligned}
p_f(\mu_k) &= p(\theta_k, \nu|\mathbf{y}_0^{k-1}) = p(\nu|\mathbf{y}_0^{k-1})p(\theta_k|\nu, \mathbf{y}_0^{k-1}) \\
p_b(\mu_k) &= p(\theta_k, \nu|\mathbf{y}_{k+1}^{N-1}) = p(\nu|\mathbf{y}_{k+1}^{N-1})p(\theta_k|\nu, \mathbf{y}_{k+1}^{N-1})
\end{aligned}
$$

(notation: $\mathbf{y}_i^j \stackrel{\text{def}}{=} \{y_k\}_{k=i}^j$, $i < j$).

- With discretization,

$$
\begin{aligned}
P(\nu = \nu_\ell|\mathbf{y}_0^{k-1}) &= \gamma_{f,k}^{(\ell)} \\
P(\nu = \nu_\ell|\mathbf{y}_{k+1}^{N-1}) &= \gamma_{b,k}^{(\ell)}
\end{aligned}
$$

The parallel Tikhonov algorithm (2)

- As before, we assume

$$\begin{aligned} p(\theta_k|\nu_\ell, \mathbf{y}_0^{k-1}) &\simeq t(a_{f,k}^{(\ell)}; \theta_k) \\ p(\theta_k|\nu_\ell, \mathbf{y}_{k+1}^{N-1}) &\simeq t(a_{b,k}^{(\ell)}; \theta_k) \end{aligned}$$

- Hence, in order to update the messages, we have simply to propagate the variables $\gamma_{f,k}^{(\ell)}$, $\gamma_{b,k}^{(\ell)}$, $a_{f,k}^{(\ell)}$, and $a_{b,k}^{(\ell)}$, for all $\nu_\ell \in \mathcal{V}$ along the phase trellis.

The parallel Tikhonov algorithm (3)

- As for the standard Tikhonov algorithm, we use the min-divergence Gaussian approximation and let

$$p_d(\theta_k) \simeq t(\frac{2y_k\alpha_k^*}{N_0 + \beta_k - |\alpha_k|^2}; \theta_k)$$

- Also, for notational convenience, we introduce the term

$$u_k = \frac{2y_k\alpha_k^*}{N_0 + \beta_k - |\alpha_k|^2}$$

The parallel Tikhonov algorithm (4)

- By using again the properties of the Tikhonov pdf, after some algebra, we obtain the forward recursion in the form

$$\gamma_{f,k+1}^{(\ell)} t(a_{f,k+1}^{(\ell)};\theta_{k+1}) \quad \propto \quad \gamma_{f,k}^{(\ell)} \frac{\mathrm{I}_0\left(\frac{\left|a_{f,k}^{(\ell)}+u_k\right|}{1+\sigma_\Delta^2\left|a_{f,k}^{(\ell)}+u_k\right|}\right)}{\mathrm{I}_0\left(\left|a_{f,k}^{(\ell)}\right|\right)}$$

$$\cdot \frac{e^{\frac{\sigma_\Delta^2\left|a_{f,k}^{(\ell)}+u_k\right|^2}{1+\sigma_\Delta^2\left|a_{f,k}^{(\ell)}+u_k\right|}}}{\sqrt{1+\sigma_\Delta^2\left|a_{f,k}^{(\ell)}+u_k\right|}} t\left(\frac{\left(a_{f,k}^{(\ell)}+u_k\right)e^{j2\pi\nu_\ell}}{1+\sigma_\Delta^2|a_{f,k}^{(\ell)}+u_k|};\theta_{k+1}\right)$$

The parallel Tikhonov algorithm: forward recursion

- Since $|a_{f,k}^{(\ell)} + u_k|$ is usually larger then a few units and $\sigma_\Delta^2 \ll 1$, we use the large argument approximation $\mathrm{I}_0(x) \simeq e^x$ and neglect $\sigma_\Delta^2 |a_{f,k}^{(\ell)} + u_k|$ with respect to 1 in the square root.

- We obtain

$$
\begin{aligned}
a_{f,k+1}^{(\ell)} &= \frac{a_{f,k}^{(\ell)} + u_k}{1 + \sigma_\Delta^2 |a_{f,k}^{(\ell)} + u_k|} e^{j2\pi\nu_\ell} \\
\gamma_{f,k+1}^{(\ell)} &= \gamma_{f,k}^{(\ell)} \exp\left\{ \left| a_{f,k}^{(\ell)} + u_k \right| - \left| a_{f,k}^{(\ell)} \right| \right\}
\end{aligned}
$$

for each $\ell = 0, \ldots, 2M$.

The parallel Tikhonov algorithm: backward recursion

- Similarly, the backward recursion is given by

$$
\begin{aligned}
a_{b,k-1}^{(\ell)} &= \frac{a_{b,k}^{(\ell)} + u_k}{1 + \sigma_\Delta^2 |a_{b,k}^{(\ell)} + u_k|} e^{-j2\pi\nu_\ell} \\
\gamma_{b,k-1}^{(\ell)} &= \gamma_{b,k}^{(\ell)} \exp\left\{ \left| a_{b,k}^{(\ell)} + u_k \right| - \left| a_{b,k}^{(\ell)} \right| \right\}
\end{aligned}
$$

The parallel Tikhonov algorithm: output

- By applying sum-product algorithm we find

$$P_u(x_k) = \sum_{\ell} \int p_f(\theta_k, \nu_\ell) p_b(\theta_k, \nu_\ell) p(y_k | x_k, \theta_k) d\theta_k$$

$$\propto e^{-\frac{|x_k|^2}{N_0}} \sum_{\ell} \gamma_{f,k}^{(\ell)} \gamma_{b,k}^{(\ell)} \int t(a_{f,k}^{(\ell)}; \theta_k) t(a_{b,k}^{(\ell)}; \theta_k) e^{\mathrm{Re}\left[\frac{y_k c_k^*}{N_0/2} e^{-j\theta_k}\right]} d\theta_k$$

- This yields

$$P_u(x_k) \propto e^{-\frac{|x_k|^2}{N_0}} \sum_{\ell} \gamma_{f,k}^{(\ell)} \gamma_{b,k}^{(\ell)} \frac{\mathrm{I}_0\left(\left|a_{f,k}^{(\ell)} + a_{b,k}^{(\ell)} + \frac{y_k c_k^*}{N_0/2}\right|\right)}{\mathrm{I}_0\left(\left|a_{f,k}^{(\ell)}\right|\right) \mathrm{I}_0\left(\left|a_{b,k}^{(\ell)}\right|\right)} .$$

The parallel Kalman algorithm (1)

- Using the phasor $h_k = e^{j2\pi\theta_k}$ rather than θ_k itself, the resulting re-defined state variable is given by $\mu_k = (h_k, \nu_k)$ with $h_k \in \mathbb{C}$ and $\nu_k \in \mathcal{V}$.

- If follows that, for each value $\nu = \nu_\ell$, the corresponding conditional estimator for $\{h_k\}$ is given by a Kalman smoother identical to what seen before, when replacing α_k by the rotated version $\alpha_k e^{j2\pi\nu_\ell k}$.

- The $2M+1$ Kalman smoothers are run in parallel, and produce estimates that are assumed to be Gaussian $\sim \mathcal{N}_{\mathbb{C}}(m_k(\ell), \Sigma_k)$.

- Notice that the covariance sequence of the Kalman smoothers does not depend on ℓ, and hence only one covariance recursion instead of $2M+1$ can be implemented.

The parallel Kalman algorithm (2)

- It remains to see how the terms $\gamma_{f,k}(\ell)$ and $\gamma_{b,k}(\ell)$ are recursively calculated.
- By using again the properties fo the Gaussian pdfs:

$$\begin{aligned} p_{f,k+1}(\mu_{k+1}) &= \gamma_{f,k+1}(\ell) g_{\mathbb{C}}(m_{k+1|k}(\ell), \Sigma_{k+1|k}; h_{k+1}) \\ &= \gamma_{f,k}(\ell) g_{\mathbb{C}}(\alpha_k e^{j2\pi\nu_\ell k} m_{k|k-1}(\ell), |\alpha_k|^2(\Sigma_{k|k-1} - 1) + N_0 + \beta_k; y_k) \\ &\quad \cdot g_{\mathbb{C}}(\rho m_{k|k}(\ell), \rho^2(\Sigma_{k|k} - 1) + 1; h_{k+1}) \end{aligned}$$

The parallel Kalman algorithm (3)

- It follows that the forward update is

$$\gamma_{f,k+1}(\ell) \propto \gamma_{f,k}(\ell) \exp\left(-\frac{1}{|\alpha_k|^2(\Sigma_{k|k-1}-1)+N_0+\beta_k}\left|y_k-\alpha_k e^{j2\pi\nu_\ell k} m_{k|k-1}(\ell)\right|^2\right)$$

- Similarly, the backward update is

$$\gamma_{b,k-1}(\ell) \propto \gamma_{b,k}(\ell) \exp\left(-\frac{1}{|\alpha_k|^2(\Sigma_{k|k+1}-1)+N_0+\beta_k}\left|y_k-\alpha_k e^{j2\pi\nu_\ell k} m_{k|k+1}(\ell)\right|^2\right)$$

The parallel Kalman algorithm: output

- The messages to be sent to the decoder are immediately obtained as

$$
\begin{aligned}
P_u(x_k) \quad &\propto \quad e^{-\frac{|x_k|^2}{N_0}} \sum_{\ell=0}^{2M} \gamma_{f,k}(\ell)\gamma_{b,k}(\ell) \\
&\cdot \int_0^{\infty} e^{-z} I_0 \left(2\sqrt{\frac{|y_k x_k^*|^2}{N_0^2} + \frac{z|m_k(\ell)|^2}{\Sigma_k} + 2\frac{\sqrt{z}}{\sqrt{\Sigma_k}N_0}\mathrm{Re}\{y_k x_k^* m_k^*(\ell) e^{-j2\pi\nu_\ell k}\}} \right)
\end{aligned}
$$

with

$$
\begin{aligned}
m_k(\ell) &= \frac{\Xi_{k|k+1}}{\Sigma_{k|k-1} + \Xi_{k|k+1}} m_{k|k-1}(\ell) + \frac{\Sigma_{k|k-1}}{\Sigma_{k|k-1} + \Xi_{k|k+1}} \mu_{k|k+1}(\ell) \\
\Sigma_k &= \frac{\Sigma_{k|k-1}\Xi_{k|k+1}}{\Sigma_{k|k-1} + \Xi_{k|k+1}} .
\end{aligned}
$$

Intuition: why it works

- We plot

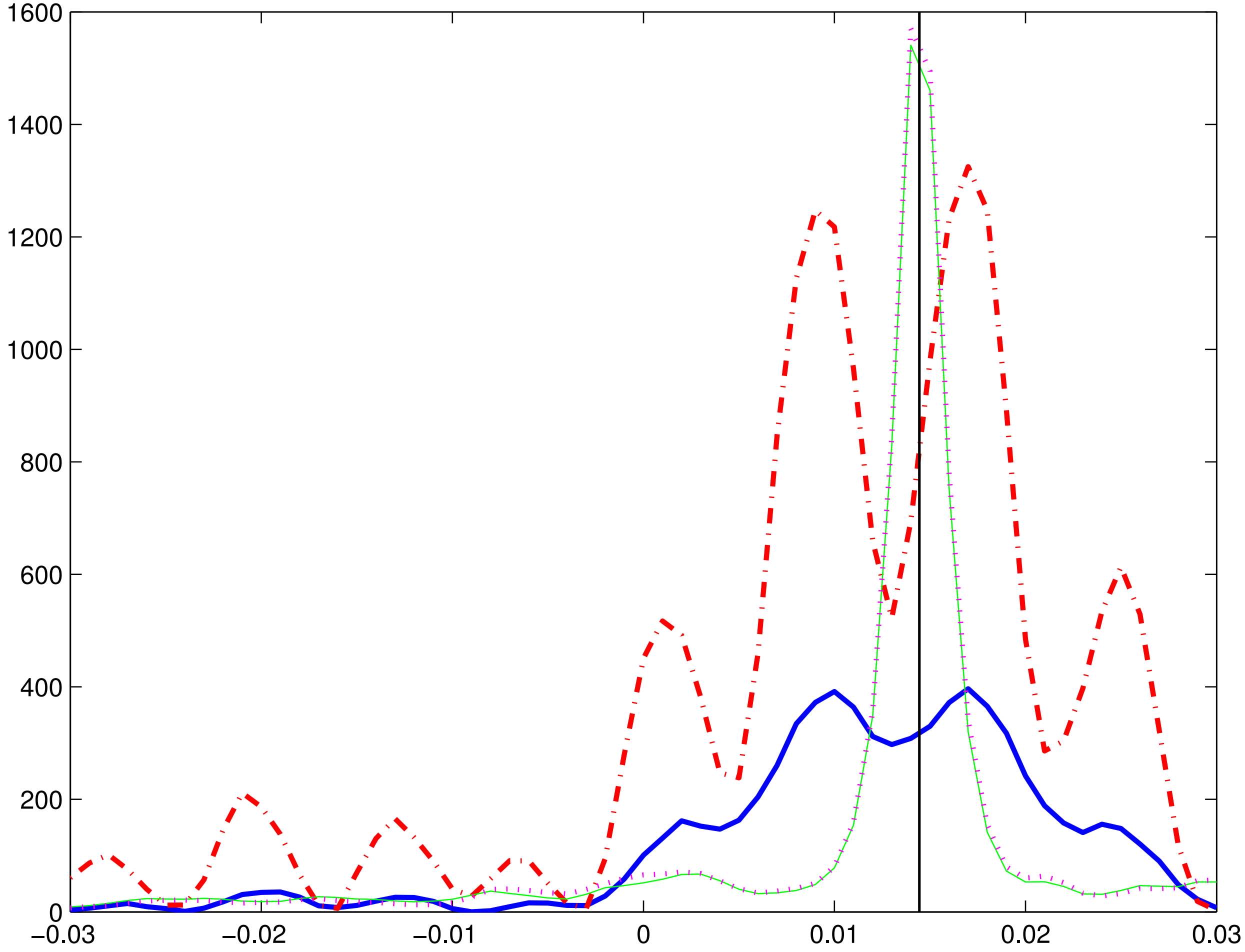
$$\mathcal{M}(\ell) = \sum_{k=0}^{N-1} |m_k(\ell)|^2$$

  vs. the frequency values $\nu_\ell \in \mathcal{V}$, for one snapshot.

- For ν_ℓ far from the true value $\nu^\star$ of the frequency offset, $m_k(\ell) \approx 0$.

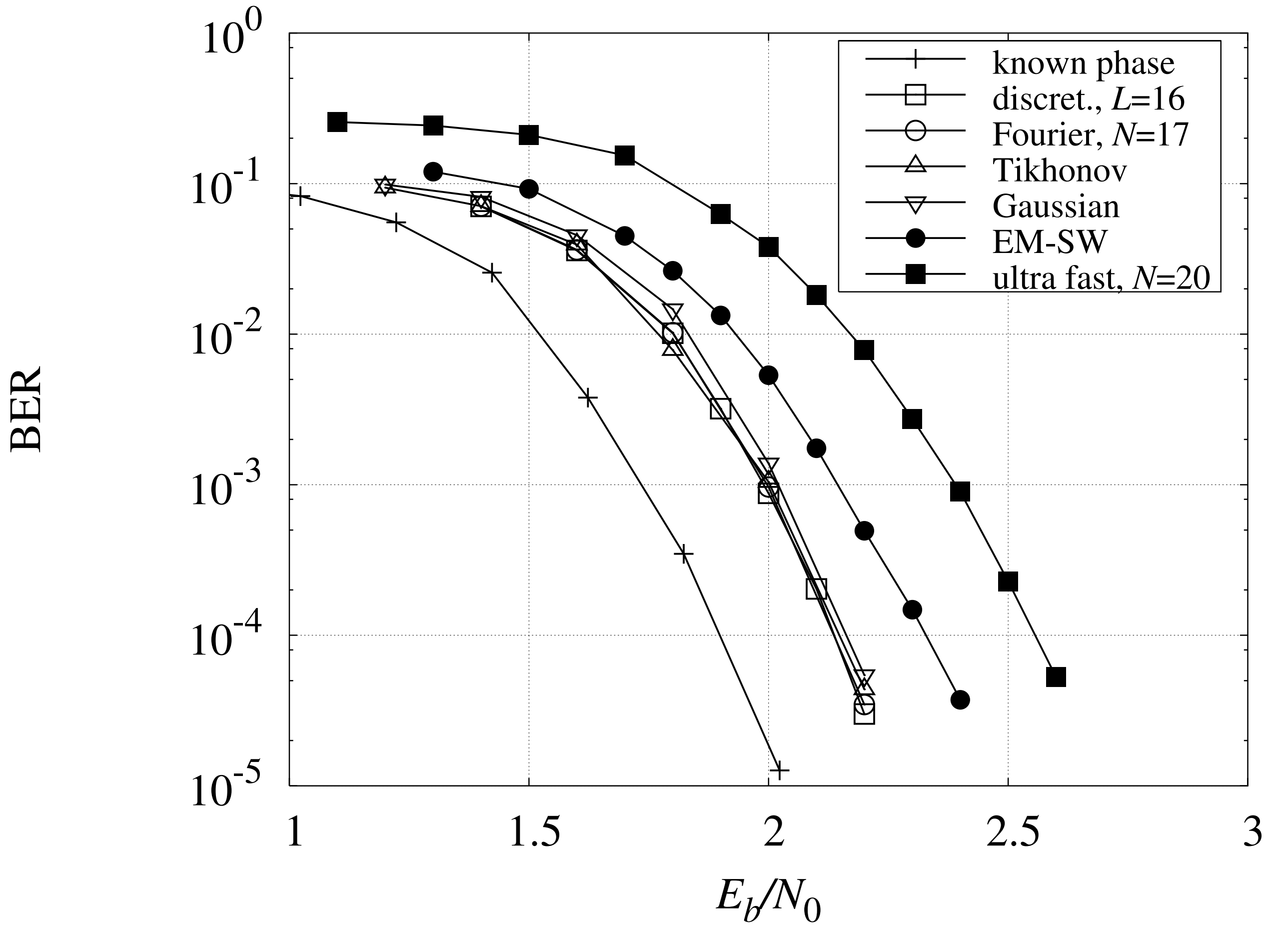
- For ν_ℓ close to $\nu^\star$, $|m_k(\ell)| \approx 1$.

Snapshot: DVB-RCS, TC (frame 1712 bits), rate $1/2$, $E_b/N_0 = 2$ dB

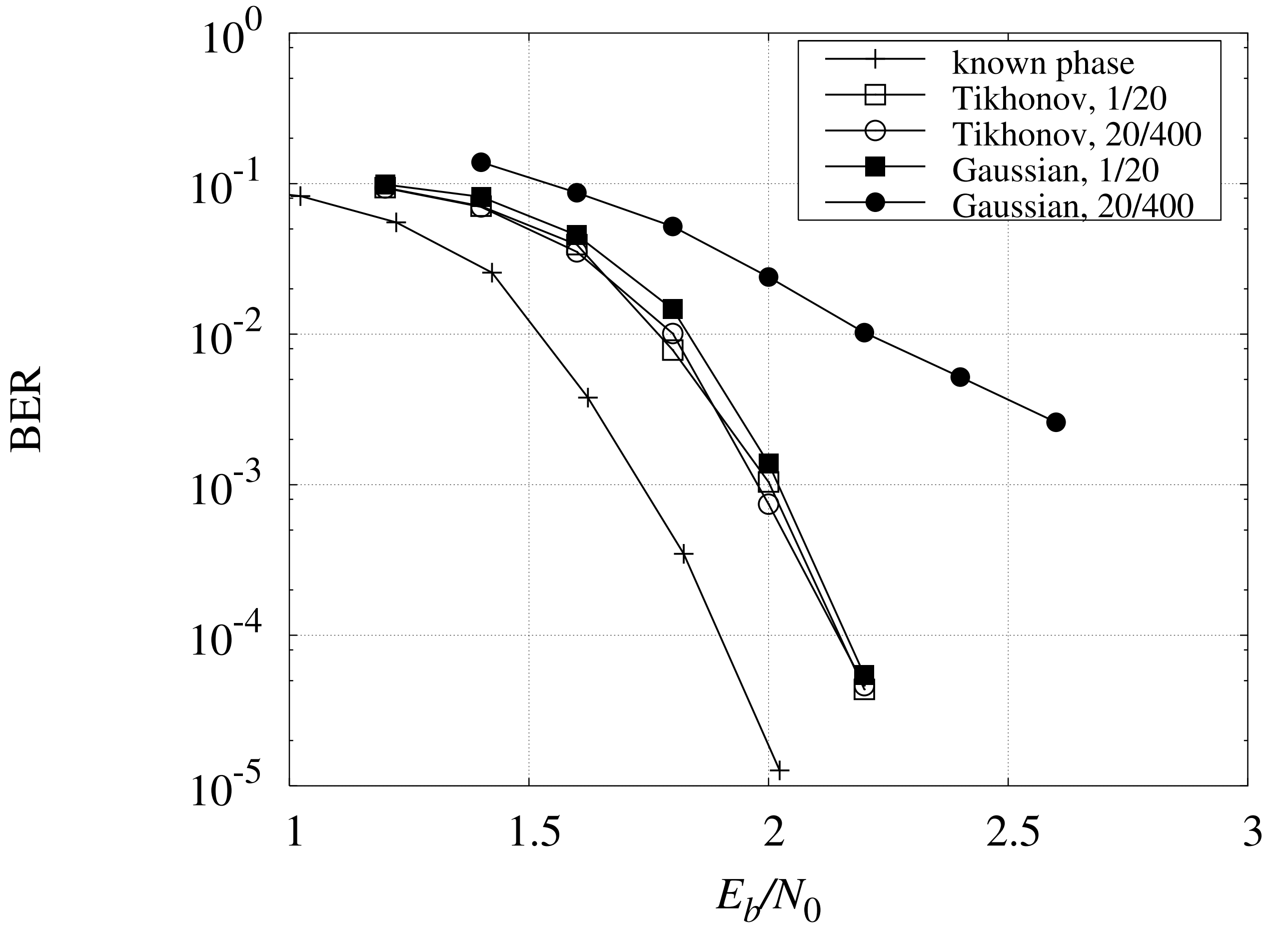


Some numerical example

Comparisons (binary LDPC $(3,6)$, $N = 4000$)

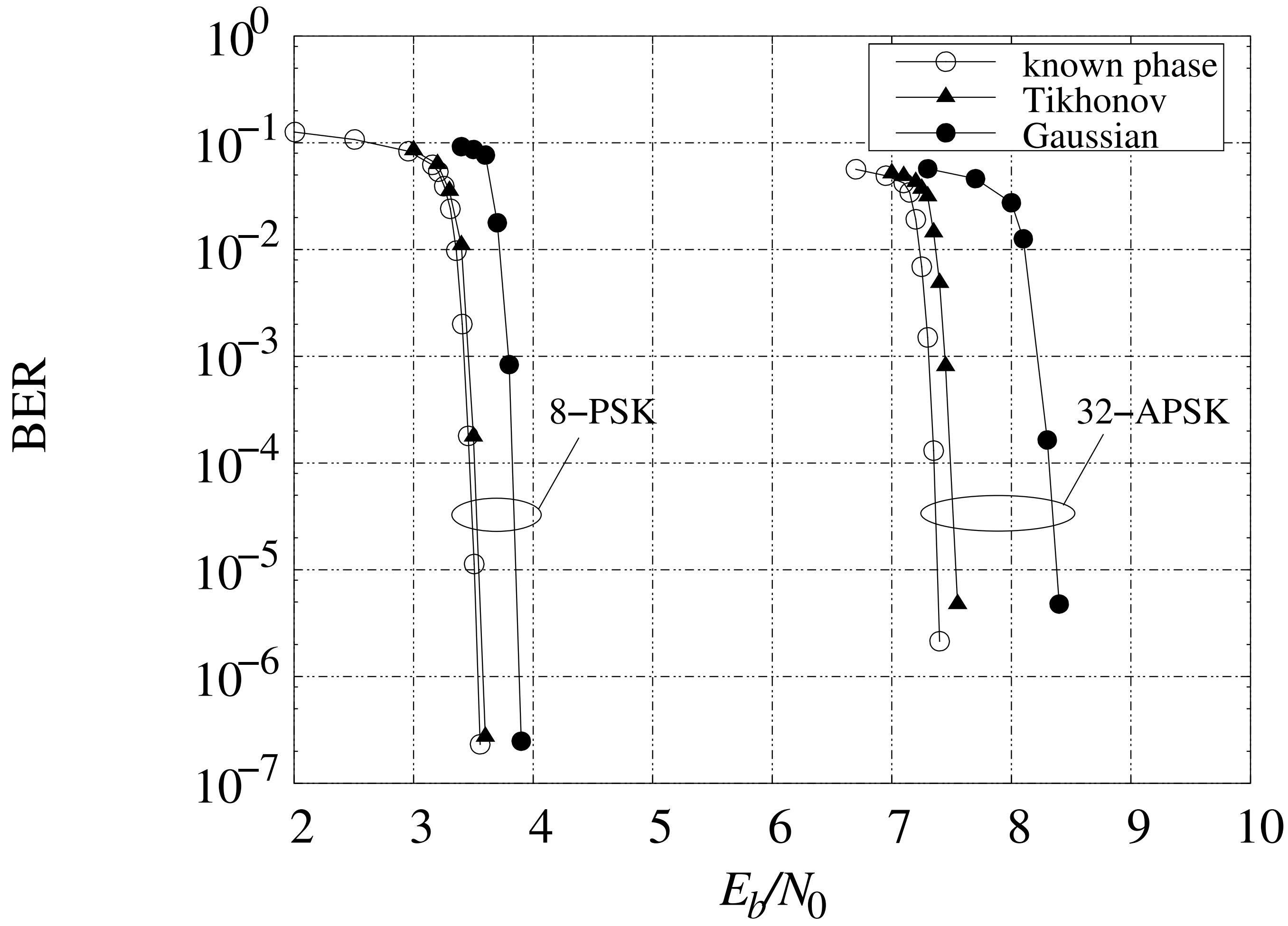


Comparisons (binary LDPC $(3,6)$, $N = 4000$)



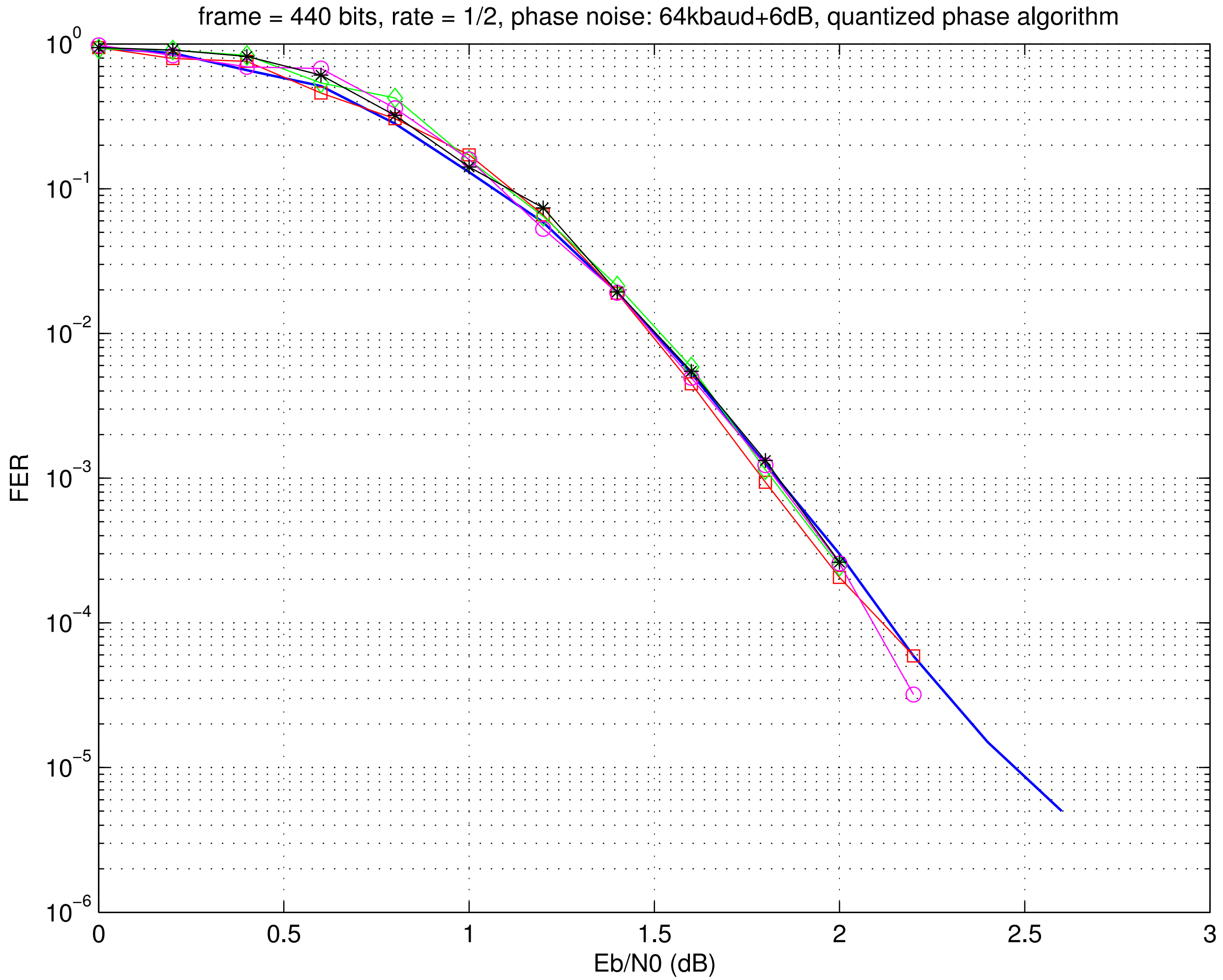
Effect of pilot symbols placement.

Performance for the DVB-S2 standard

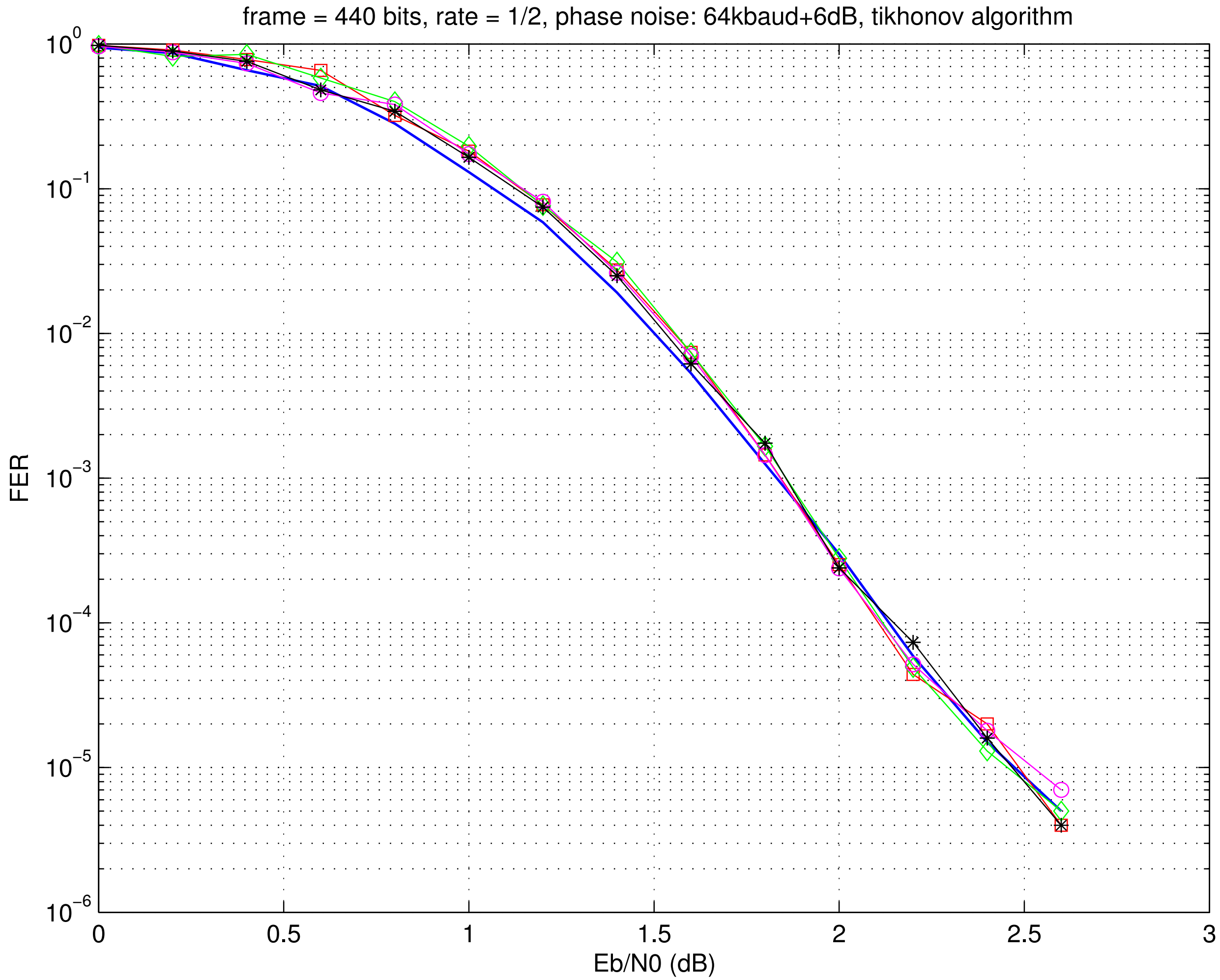


LDPC-coded modulation, binary codign rate 4/5, binary blocklength 64800, bursts of 36 pilots spaced by 1476 symbols.

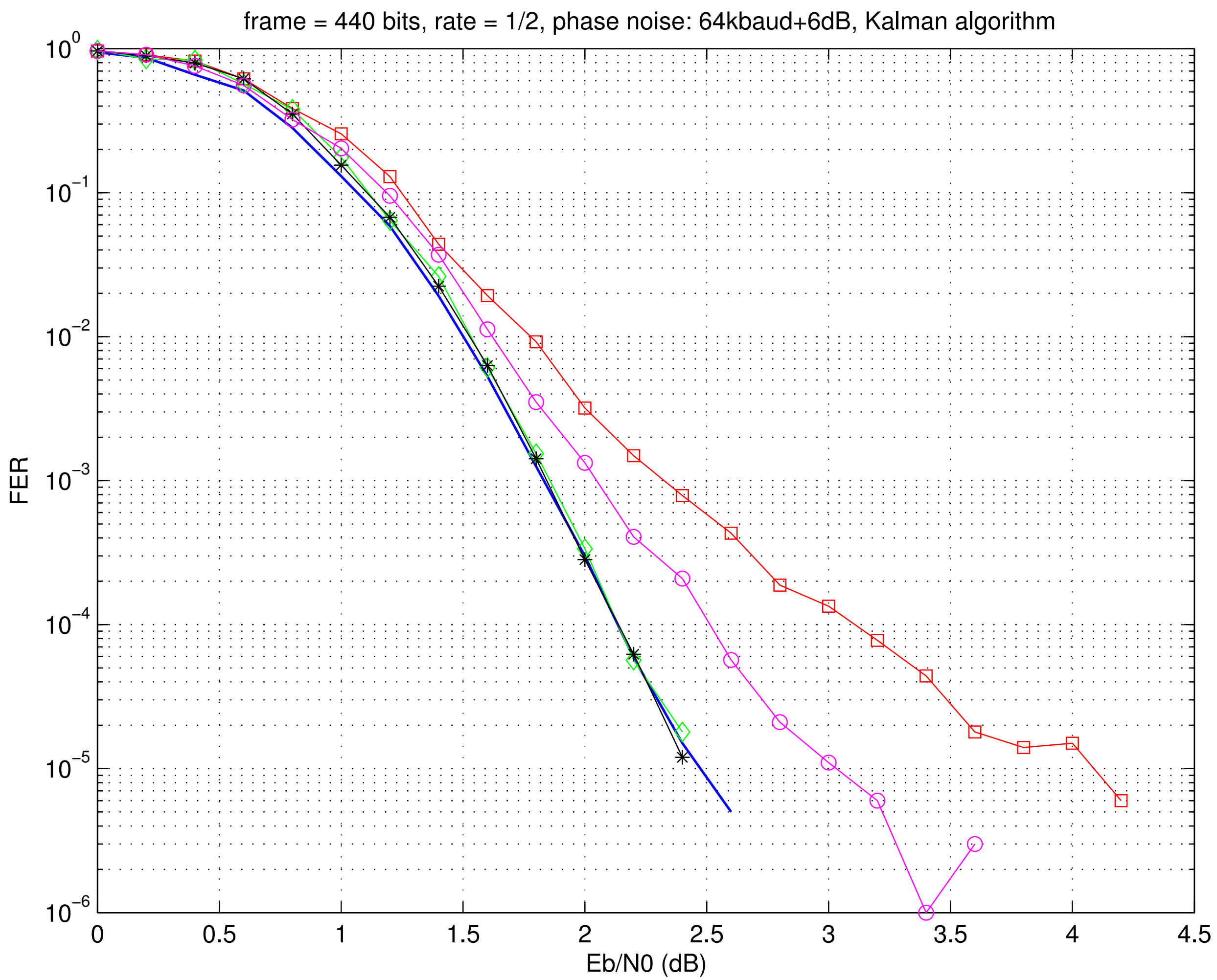
DVB-RCS standard: Phase BCJR



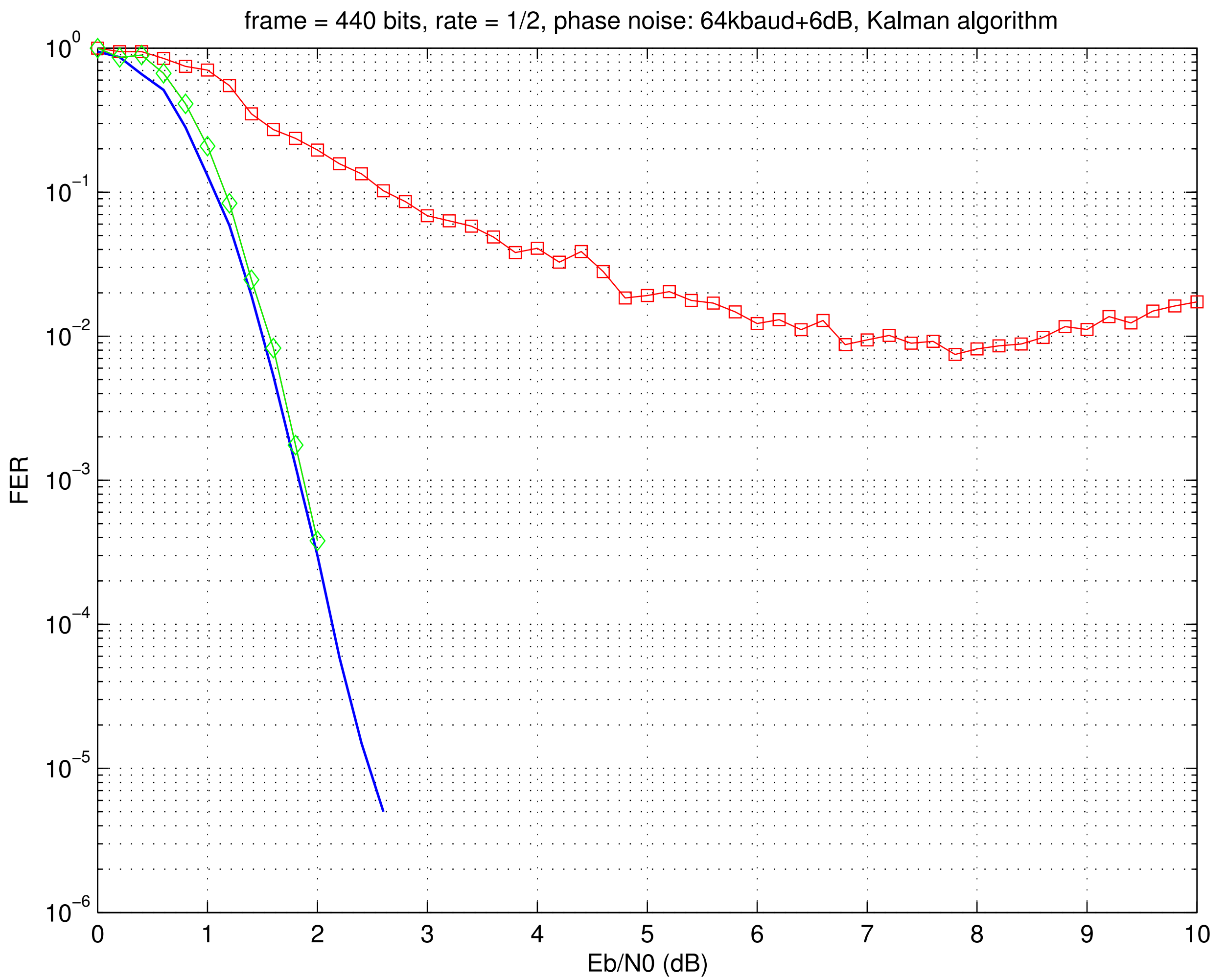
DVB-RCS standard: Tikhonov algorithm



DVB-RCS standard: Kalman algorithm



Parallel Kalman algorithm: $\nu_{\max} = 0.01$, $M = 10$



Parallel Kalman algorithm: $\nu_{\max} = 0.05$, $M = 50$