Estimation and detection from coded signals

Joint research:
- UGent (N. Noels, H. Wymeersch, H. Steendam, M. Moeneclaey)
- UCL (C. Herzet, L. Vandendorpe)

Presented by Marc Moeneclaey, UGent - TELIN dept.
Marc.Moeneclaey@telin.UGent.be
Outline

• Coding + linear modulation, AWGN channel, unknown parameter $\theta$ (carrier phase, ...)
• MAP detection of information bits
  * $\theta$ known
  * $\theta$ unknown
• Low-complexity alternatives to MAP detection for unknown $\theta$
  * iterative ML estimation of $\theta$ (EM algorithm)
  * simplified sum-product algorithm
• Numerical results
• Conclusions
Coding + linear modulation, AWGN channel

infobits \( b \) → encoding, interleaving, mapping → coded symbols \( a \) → pulse \( h(t) \) → linear modulation → AWGN channel

\[ r(t) = e^{j\theta} \sum_k a_k h(t - kT) + w(t) \]

\( b \rightarrow a = \chi(b) \): turbo code, LDPC code, BICM, ....

\( h(t) \): square-root Nyquist pulse

\( \theta \): unknown parameter (carrier phase, time delay, freq. offset, gain)
MAP detection, $\theta$ known

MAP detection (on individual infobits) achieves minimum BER

$$\hat{b}_k = \arg \max_{b_k \in \{0,1\}} p(b_k \mid r) \quad \text{r : vector representation of } r(t)$$

Assuming $\theta = \theta_0$ is known, bit APP $p(b_k \mid r)$ is given by

$$p(b_k \mid r) \propto \sum_{b: b_k} p(r \mid a = \chi(b), \theta_0) \quad \text{(many terms !)}$$

Efficient computation of APPs: *sum-product (SP) algorithm + message-passing in factor graph (FG)*
Factor graph for known $\theta$

\[
g(z_1, a_1, \theta_0) \quad g(z_k, a_k, \theta_0) \quad g(z_K, a_K, \theta_0)
\]

Encoding, interleaving, mapping

\[
\begin{align*}
b_1 & \downarrow & a_1 \\
b_\ell & \downarrow & a_k \\
b_L & \downarrow & a_K
\end{align*}
\]

\[
g(z_k, a_k, \theta_0) \quad \text{channel observation} \quad g(z_k, a_k, \theta_0) = \exp\left(\frac{2}{N_0} \Re[a_k^* z_k e^{-j\theta}]\right)
\]
Factor graph for known $\theta$

\[
g(z_1, a_1, \theta_0) \quad g(z_k, a_k, \theta_0) \quad g(z_K, a_K, \theta_0)
\]

Encoding, interleaving, mapping

\[
\begin{align*}
l \quad b_1 \quad \downarrow \quad & g(z_k, a_k, \theta_0) \quad \text{channel observation} \\
& \quad p_e(a_k) \quad \text{extrinsic information on } a_k \\
& \quad p(b_\ell | r) \quad \text{information bit APP}
\end{align*}
\]
Factor graph for known $\theta$

Because of cycles in FG, MAP detection is iterative
Factor graph, $\theta$ known

- $g(z_1,a_1,\theta_0)$
- $g(z_k,a_k,\theta_0)$
- $g(z_K,a_K,\theta_0)$

Encoding, interleaving, mapping

- $a_1$
- $a_k$
- $a_K$

- $b_1$
- $b_\ell$
- $b_L$

- $g(z_k,a_k,\theta_0)$ channel observation
- $g(z_k,a_k,\theta_0) = \exp\left(\frac{2}{N_0} \text{Re}[a_k^* z_k e^{-j\theta}]\right)$

- $p_e^{(n)}(a_k)$ extrinsic information on $a_k$

- $p^{(n)}(b_\ell|r)$ information bit APP

$n$: iteration index of MAP detector

iterate until convergence
Receiver structure, $\theta$ known

\[ r(t) \xrightarrow{h^*(-t)} Z_k \xrightarrow{	ext{exp}(-j\theta_0)} \text{"conventional" MAP detector} \]

(SP algorithm, iterate until convergence)
MAP detection, $\theta$ unknown

When $\theta$ is unknown (uniform distrib.), bit APP is given by

$$p(b_k \mid r) \propto \sum_{b: b_k} \left( \int p(r \mid a = \chi(b), \theta) d\theta \right)$$

$\neq$ conventional MAP detector

hard to compute, because of integration over $\theta$
Factor graph, $\theta$ unknown

messages are functions of $\theta$

downward messages involve integration over $\theta$

\[
g(z_1, a_1, \theta) \quad \cdots \quad g(z_K, a_K, \theta)\]

Encoding, interleaving, mapping

\[
b_1 \quad \cdots \quad b_L\]
Alternatives to exact MAP detection when $\theta$ is unknown

**Alternative 1 :** ML parameter estimation
Provide estimate of $\theta$ to conventional MAP detector

$$p(b_k | r) = \int p(b_k | r, \theta)p(\theta | r)d\theta$$

$$\approx p(b_k | r, \hat{\theta}(r))$$

$$p(\theta | r) \approx \delta(\theta - \hat{\theta}(r))$$

**Alternative 2 :** Simplified sum-product algorithm
Approximation of $\theta$-dependent messages
Alternative 1:  
ML parameter estimation
Receiver structure when estimating $\theta$

**Strategy**: use conventional MAP detector, but provide *estimate* (instead of correct value) of $\theta$
ML parameter estimation

ML estimation of $\theta$ : $\hat{\theta}_{ML} = \arg \max_\theta p(r | \theta)$

For long observation intervals, mean-square error (MSE) of ML estimate converges to Cramer-Rao lower bound (CRB) on MSE.

Computation of CRB is hard when data symbols are not a priori known to receiver.

Simpler but less tight bound : modified CRB (MCRB) (assumes data symbols are known to receiver)
Cramer-Rao bound

\[ CRB^{-1} = E_r \left[ \left( \frac{\partial}{\partial \theta} \ln p(r; \theta) \right)^2 \right] = -E_r \left[ \frac{\partial^2}{\partial \theta^2} \ln p(r; \theta) \right] \]

\[ p(r; \theta) = \sum_a p(r | a; \theta)p(a) = E_a [p(r | a; \theta)] \]

average over (coded) symbol sequence \( a \)

(\text{many terms}!!)

analytical difficulty: \( E_a[.] \) “inside” \( \ln(.) \), so that \( \ln(p(r;\theta)) \) is not a simple function of \( r \) and \( \theta \)
Modified Cramer-Rao bound

$$\text{MCRB}^{-1} = E_{r,a} \left[ \left( \frac{\partial}{\partial \theta} \ln p(r | a; \theta) \right)^2 \right] = -E_{r,a} \left[ \frac{\partial^2}{\partial \theta^2} \ln p(r | a; \theta) \right]$$

$E_{r,a}[]. \text{“outside” } \ln(.) \Rightarrow \text{MCRB easier to evaluate than CRB,}$

$\text{when } \ln(p(r|a;\theta)) \text{ is simple function of } a \text{ and } r$

(e.g., linear modulation on AWGN channel)
MCRB for phase estimation

\[ r = a \exp(j\theta) + w \quad \text{E}[ww^T] = \frac{N_0}{2} I \]

\[
\ln p(r | a; \theta) = \frac{2}{N_0} \sum_k \Re[a_k^* r_k \exp(-j\theta)]
\]

\[
\frac{\partial^2}{\partial \theta^2} \ln p(r | a; \theta) = \frac{-2}{N_0} \sum_k \Re[a_k^* r_k \exp(-j\theta)]
\]

\[
\text{MCRB}^{-1} = -\text{E}_{r,a} \left[ \frac{\partial^2}{\partial \theta^2} \ln p(r | a; \theta) \right] = \frac{2}{N_0} \sum_k \text{E}_{a_k} \left[ |a_k|^2 \right] = \frac{2E_s K}{N_0}
\]
CRB for phase estimation: uncoded transmission

\[ p(a) = \prod_k p(a_k) \quad \text{(i.i.d. symbols)} \]

\[ \Rightarrow \quad p(r; \theta) = \prod_k p(r_k; \theta) = \prod_k E_{a_k}[p(r_k | a_k; \theta)] \]

\[ \text{CRB}^{-1} = K \cdot E_{r_k} \left[ \left( \frac{\partial}{\partial \theta} \ln(E_{a_k}[p(r_k | a_k; \theta)]) \right)^2 \right] \]

\[ E_{a_k}[\cdot] : 1-\text{D summation} \]

\[ E_{r_k}[\cdot] : 1-\text{D numerical integration} \]
CRB for phase estimation: coded transmission

\[
\frac{\partial}{\partial \theta} \ln p(r; \theta) = \sum_a \left( \frac{\partial}{\partial \theta} \ln p(r | a; \theta) \right) p(a | r; \theta)
\]

\[
r = a \exp(j\theta) + w \quad \Rightarrow \quad \frac{\partial}{\partial \theta} \ln p(r | a; \theta) = \frac{2}{N_0} \sum_k \text{Im}[a_k^* r_k \exp(-j\theta)]
\]

\[
\frac{\partial}{\partial \theta} \ln p(r; \theta) = \frac{2}{N_0} \sum_k \text{Im}[\tilde{a}_k^* (r; \theta) r_k \exp(-j\theta)]
\]

\[
\tilde{a}_k (r; \theta) = \sum_a a_k p(a | r; \theta) = \sum_{a_k} a_k \sum_{a} p(a_k | r; \theta) \quad \text{marginal symbol APPs (from MAP detector operating on } r)\]

\[
\text{CRB}^{-1} = E_r \left[ \left( \frac{\partial}{\partial \theta} \ln p(r; \theta) \right)^2 \right] \quad \text{E}_{a_k | r}[\cdot] : 1-D summation per symbol}

\[
E_r[\cdot] : \text{Monte Carlo simulation}
\]
CRB, MCRB : numerical results

Large SNR : CRB → MCRB
Small SNR : CRB uncoded > CRB coded > MCRB

\[ MCRB = \frac{N_0}{2LE_s} \]

(phase estimation)

⇒ code properties should be exploited during estimation
Computation of ML estimate

Direct application of ML estimation is complicated:

$$p(r | \theta) \propto \sum_b p(r | a = \chi(b), \theta) \quad \text{many terms!}$$

$$\Rightarrow \text{compute ML estimate iteratively (EM algorithm)}: \quad \hat{\theta}_{\text{ML}} = \lim_{i \to \infty} \hat{\theta}^{(i)}$$

$$\hat{\theta}^{(i+1)} = \arg \max_{\theta} E[p(r | a, \theta) | r, \hat{\theta}^{(i)}] = \arg \left( \sum_k \tilde{a}_k^{(i)}* z_k \right)$$

$$\tilde{a}_k^{(i)} = E[a_k | r, \hat{\theta}^{(i)}] \quad \text{soft decision (SD) on symbols}$$

symbol APPs computed by MAP detector
EM algorithm : symbol APP computation

outer iteration index $i$ (EM algorithm),
inner iteration index $n$ (MAP detector)

for each $i$:

MAP detector is reset and iterated until convergence ($n \to \infty$)

\[ g(z_k, a_k, \hat{\theta}^{(i)}) \]

Encoding, interleaving, mapping

\[ p_e^{(i,n)}(a_k) \]

\[ p(a_k | r, \hat{\theta}^{(i)}) = g(z_k, a_k, \hat{\theta}^{(i)})p_e^{(i,\infty)}(a_k) \Rightarrow SD \ \tilde{a}_k^{(i)} \]
Receiver structure for EM algorithm

symbol APP: \( p(a_k | r, \hat{\theta}^{(i)}) = g(z_k, a_k, \hat{\theta}^{(i)})p_e^{(i,\infty)}(a_k) \Rightarrow SD \tilde{a}_k^{(i)} \)
Reduced-complexity EM algorithm

For each EM iteration, MAP detector is iterated till convergence
⇒ computational complexity too high

Solution: “merging” of EM iterations and MAP detector iterations. For each EM iteration, only one MAP detector iteration is performed, without resetting MAP detector
⇒ reduced complexity (at expense of reduced convergence speed)

symbol APP: \[ p(a_k \mid r, \hat{\theta}^{(i)}) \propto g(z_k, a_k, \hat{\theta}^{(i)}) p_{e^{(i,l)}}(a_k) \quad \Rightarrow \quad \text{SD} \; \hat{a}_k^{(i)} \]
EM algorithm: exploiting code properties

- Partial (or no) exploitation of code properties \(\Rightarrow\) degradation of MSE
- Does not exploit code properties
- Uses infobit APPs only, assumes parity bits are uncoded
- Uses infobit APPs and parity bit APPs

- Turbo code, rate 1/2, 4-PSK
- Phase estimation

\[
\begin{align*}
\text{MSE/MCRB, CRB/MCRB} & \quad \text{(dB)} \\
0 & \quad 0.5 \quad 1 \quad 1.5 \quad 2 \quad 2.5 \quad 3 \\
\end{align*}
\]
EM algorithm : phase estimation (1/2)

Turbo code (I=240, rate 1/3), BICM, QPSK
Phase estimation
EM algorithm: phase estimation (2/2)

Turbo code (I=240, rate 1/3), BICM, QPSK
Phase estimation

- perfect synchronization
- V&V
- EM (10 it)
EM algorithm : timing estimation (1/2)

Turbo code \( (I=240, \text{rate } 1/3), \text{BICM, QPSK} \)
No pilot symbols

Timing estimation
\textit{(assume perfect frame synchronization)}
EM algorithm : timing estimation (2/2)

Turbo code (I=240, rate 1/3), BICM, QPSK
Timing estimation

- perfect synchronization
- O&M
- EM (10 it)
Alternative 2 :
Simplified sum-product algorithm
Simplified sum-product algorithm

Factor graph corresponding to unknown $\theta$

- messages are functions of $\theta$
- downward messages involve integration over $\theta$

Encoding, interleaving, mapping

$$g(z_1, a_1, \theta) \quad g(z_k, a_k, \theta) \quad g(z_K, a_K, \theta)$$
Simplified sum-product algorithm

Messages depending on $\theta$ are approximated by $\delta(\theta - \hat{\theta}^{(n)})$

$n : \text{iteration index of MAP detector}$

$$\hat{\theta}^{(n)} = \arg\max_{\theta} p^{(n)}(r \mid \theta) = \arg\max_{\theta} \sum_{a} p(r \mid a; \theta)p_e^{(n)}(a)$$

$$p(r \mid a; \theta) = \prod_{k} g(z_k, a_k, \theta)$$

$$p_e^{(n)}(a) = \prod_{k} p_e^{(n)}(a_k)$$
Simplified sum-product algorithm

Messages depending on \( \theta \) are \textit{approximated} by \( \delta(\theta - \hat{\theta}^{(n)}) \) 

\( \hat{\theta}^{(n)} = \arg \max_{\theta} p^{(n)}(r \mid \theta) = \arg \max_{\theta} \sum_{a} p(r \mid a; \theta)p^{(n)}_e(a) \)

\[ p(r \mid a; \theta) = \prod_{k} g(z_k, a_k, \theta) \]

\[ p^{(n)}_e(a) = \prod_{k} p^{(n)}_e(a_k) \]
Simplified sum-product algorithm

Messages depending on $\theta$ are *approximated* by $\delta(\theta - \hat{\theta}^{(n)})$

($n$ : iteration index of MAP detector)

$$
\hat{\theta}^{(n)} = \arg \max_{\theta} p^{(n)}(r | \theta) = \arg \max_{\theta} \sum_{a} p(r | a; \theta)p_e^{(n)}(a)
$$

$$
p(r | a; \theta) = \prod_{k} g(z_k, a_k, \theta)
$$

$$
p_e^{(n)}(a) = \prod_{k} p_e^{(n)}(a_k)
$$

$\uparrow p_e^{(n)}(a_k)$

$\uparrow p^{(n)}(r | \theta)$
Simplified sum-product algorithm

Messages depending on $\theta$ are approximated by $\delta(\theta - \hat{\theta}^{(n)})$
($n$ : iteration index of MAP detector)

$$\hat{\theta}^{(n)} = \arg \max_{\theta} p^{(n)}(r \mid \theta) = \arg \max_{\theta} \sum_{a} p(r \mid a; \theta)p_{e}^{(n)}(a)$$

$$p(r \mid a; \theta) = \prod_{k} g(z_{k}, a_{k}, \theta)$$

$$p_{e}^{(n)}(a) = \prod_{k} p_{e}^{(n)}(a_{k})$$
Simplified sum-product algorithm

Messages depending on $\theta$ are approximated by $\delta(\theta - \hat{\theta}^{(n)})$
$(n : \text{iteration index of MAP detector})$

$$\hat{\theta}^{(n)} = \arg\max_{\theta} p^{(n)}(r | \theta) = \arg\max_{\theta} \sum_{a} p(r | a; \theta)p_e^{(n)}(a)$$

$$p(r | a; \theta) = \prod_k g(z_k, a_k, \theta)$$  \hspace{2cm} $$p_e^{(n)}(a) = \prod_k p_e^{(n)}(a_k)$$

$\delta(\theta - \hat{\theta}^{(n)})$

$g(z_k, a_k, \hat{\theta}^{(n)})$
Simplified sum-product algorithm

(n+1)-th iteration of MAP detector

\[ g(z_k, a_k, \hat{\theta}^{(n)}) \]

Encoding, interleaving, mapping
Simplified sum-product algorithm

\[(n+1)\text{-th iteration of MAP detector}\]

\[g(z_k, a_k, \hat{\theta}^{(n)}) \quad p_e^{(n+1)}(a_k)\]

Encoding, interleaving, mapping
**Simplified sum-product algorithm:**

Maximization of $p^{(n)}(r | \theta)$

\[
\hat{\theta}^{(n)} = \operatorname{arg\,max}_\theta p^{(n)}(r | \theta) = \operatorname{arg\,max}_\theta \sum_a p(r | a; \theta) p_e^{(n)}(a)
\]

Similar to ML equation: $p(a)$ is replaced by $p_e^{(n)}(a)$

$\Rightarrow$ EM algorithm to maximize $p^{(n)}(r|\theta)$ iteratively

\[
\hat{\theta}^{(n)} = \lim_{i \to \infty} \hat{\theta}^{(n,i)}
\]

\[
\hat{\theta}^{(n,i+1)} = \operatorname{arg\,max}_\theta \sum_k E[g(z_k, a_k, \theta | r, \hat{\theta}^{(n,i)}] = \operatorname{arg}\left(\sum_k \tilde{a}_k^{(n,i)*} z_k\right)
\]

Symbol APP: $p^{(n)}(a_k | r; \hat{\theta}^{(n,i)}) \propto g(z_k, a_k, \hat{\theta}^{(n,i)}) p_e^{(n)}(a_k) \Rightarrow SD \tilde{a}_k^{(n,i)}$
Simplified sum-product algorithm: receiver structure

for each n: iterate EM algorithm until convergence ($i \to \infty$)

\[ r(t) h^*(-t) \]

Compute $\hat{\theta}^{(n,i)}$ ($i \to \infty$)

\[ \text{outer iterations (index n)} \]

\[ \{p_e^{(n)}(a_k)\} \]

extr. probs.

\[ \text{inner iterations (index i)} \]

\[ \text{conventional MAP detector} \]

\[ e^{-j\hat{\theta}^{(n-1,\infty)}} \]

symbol APP: $p^{(n)}(a_k \mid r; \hat{\theta}^{(n,i)}) \propto g(z_k, a_k, \hat{\theta}^{(n,i)})p_e^{(n)}(a_k) \Rightarrow SD \tilde{a}_k^{(n,i)}$
Simplified sum-product algorithm: estimator performance

BICM: rate 1/2 conv. code, 8-PSK (set part.)
1000 symbols, Eb/N0 = 5 dB

carrier phase unknown
Simplified sum-product algorithm:
BER performance

BICM: rate 1/2 conv. code, 8-PSK (set part.)
1000 symbols, $E_b/N_0 = 5$ dB

carrier phase unknown

# MAP detector iterations

V&V
EM
SP
perfect synchr.
Conclusions

Two alternatives for MAP detection when $\theta$ is unknown

- Iterative ML estimation of $\theta$ by means of EM algorithm
  symbol APP during $i$-th EM iteration
  
  $\propto g(z_k, a_k, \hat{\theta}^{(i)}) p_e^{(i,1)}(a_k)$

  one MAP detector iteration
  for each EM iteration

- Simplified SP algorithm (combined with EM algorithm)
  symbol APP during $i$-th EM iteration of $n$-th MAP decoder iteration

  $\propto g(z_k, a_k, \hat{\theta}^{(n,i)}) p_e^{(n)}(a_k)$

  EM algorithm iterated till convergence
  for each MAP detector iteration

Both algorithms yield similar MSE and BER after convergence, but simplified SP algorithm requires considerably less MAP decoder iterations to converge.