Capon Redux

Louis L. Scharf\textsuperscript{1}  
Departments of Electrical  
Comp. Eng. and Statistics  
Colorado State University  
Fort Collins, CO 80523  
Phone: 970-491-2979  
Fax: 970-491-2249  
scharf@engr.colostate.edu

Todd McWhorter\textsuperscript{2}  
Mission Research Corporation  
3665 JFK Parkway  
Building 1, Suite 206  
Fort Collins, CO 80525  
Phone: 970-282-4400 ext. 25  
Fax: 970-282-9444  
Email: mcwhorter@aster.com

Shawn Kraut  
Departments of Physics  
and Elec. Eng.  
Duke University  
Durham, NC 27708-0291  
Phone: 919-660-5419  
Fax: 919-660-5293  
Email: kraut@ee.duke.edu

\textsuperscript{1} This work supported by ONR under contract N00014-01-1-1019-P0001.  
\textsuperscript{2} This work supported by ONR under contract N00014-00-C-0145.
Introduction

• Basic Experiment

• Capon Representations
  – Geometric Interpretations
  – Continued Fraction Expansions

• Variations on Experiments
  – Adaptive: Full Rank and Rank-Deficient
  – Mixed Coherent/Incoherent
  – Vector Extensions

• Iterative Approximations

• Properties
Basic Experiment

Begin with a 2-channel decomposition

\[ \begin{array}{c}
  b_0 \\
  B_0 \\
\end{array} \rightarrow d_0 \]

\[ x_0 : R_0 \rightarrow B_0 \rightarrow x_1 \]

\[ A_0 = [b_0 \ B_0] : \text{unitary} \]

\[ B_0^*b_0 = 0 \]

Problem: Minimize, with respect to \( v \),

\[ Q = E[1 - v^*]A_0^*x_0x_0^*A_0[1 - v^*]^*. \]

The composite covariance matrix for \( d_0 \) and \( x_1 \) is

\[ \begin{bmatrix}
  d_0 \\
  x_1
\end{bmatrix} : A_0^*R_0A_0 = \begin{bmatrix}
  b_0^*R_0b_0 & b_1^* \\
  b_1 & R_1
\end{bmatrix} ; \]

\[ \begin{array}{c}
  b_1 = B_0^*R_0b_0 \equiv H_0^*h_0 \\
  R_1 = B_0^*R_1B_0 \equiv H_0^*H_0 \\
  h_0 \triangleq R_0^{1/2}b_0 \\
  H_0 \triangleq R_0^{1/2}B_0
\end{array} \]
Two Expressions for the Capon Formula

The composite covariance matrix for $d_0$ and $x_1$ has two expressions for its inverse:

$$(A_0^* R_0 A_0)^{-1} = A_0^* R_0^{-1} A_0 = \begin{bmatrix} b_0^* R_0^{-1} b_0 & * \\ * & * \end{bmatrix} \quad \text{Form 1}$$

$$= \begin{bmatrix} b_0^* R_0 b_0 & b_1^* \\ b_1 & R_1 \end{bmatrix}^{-1} = \begin{bmatrix} Q_0^{-1} & * \\ * & * \end{bmatrix} \quad \text{Form 2}$$

The minimum MSE for estimating $d_0$ from $x_1$ is

$$Q_0 = \frac{1}{b_0^* R_0^{-1} b_0} \quad \text{Capon Form 1}$$

$$= b_0^* R_0 b_0 - b_1^* R_1^{-1} b_1 = h_0^* (I - P_{H_0}) h_0 \quad \text{Capon Form 2}$$
Subspace Geometry

\[ \langle b_0 \rangle : \text{beamformer subspace} \]
\[ \langle h_0 \rangle : \text{colored beamformer subspace} \]
\[ \langle H_0 \rangle : \text{colored GSC subspace} \]

\[ \langle B_0 \rangle : \text{GSC subspace} \]

Pythagoras:

\[ (I - P_{H_0})h_0 \]

\[ Q_0 = \text{CAPON} = \frac{1}{b_0^* R_0^{-1} b_0} = \| (I - P_{H_0})h_0 \|^2 \]
\[ = \sin^2 \theta \| h_0 \|^2 = \sin^2 \theta \ b_0^* R_0 b_0 \]
\[ = \sin^2 \theta \ \text{BARTLETT} \]

Kantorovich:

\[ \| (I - P_{H_0})h_0 \|^2 = \frac{1}{b_0^* R_0^{-1} b_0} \leq b_0^* R_0 b_0 = \| h_0 \|^2 \]
Continued Fraction Expansion

• In general, \( \|b_i\|^2 \neq 1 \).

• Define the “local” Capon and “local” Bartlett estimates

\[
Q_i = \frac{b_i^* b_i}{b_i^* R_i^{-1} b_i}; \quad P_i = \frac{b_i^* R_i b_i}{b_i^* b_i}.
\]

• Capon form 2 generates the recursion

\[
Q_i = P_i - \frac{b_i^* b_i}{Q_{i+1}}.
\]

• Then

\[
Q_0 = P_0 - \frac{b_1^* b_1}{P_1 - \frac{b_2^* b_2}{P_2 - \frac{b_3^* b_3}{P_3 - \cdots}}}
\]
Continued Fraction Recursions

• A graphic for remembering the recursions is

\[
\begin{align*}
\mathbf{x}_0 & \rightarrow \mathbf{R}_0 \mathbf{b}_0^* \rightarrow \mathbf{B}_0^* \\
& \quad \rightarrow \mathbf{R}_1 \mathbf{b}_1^* \rightarrow \mathbf{B}_1^* \\
& \quad \rightarrow \mathbf{R}_2 \mathbf{b}_2^* \rightarrow \mathbf{B}_2^* \\
& \quad \rightarrow \mathbf{R}_3 \mathbf{b}_3^* \rightarrow \mathbf{B}_3^* \\
& \quad \rightarrow \cdots
\end{align*}
\]

\[d_0 : \|\mathbf{b}_0\|^2 P_0\]
\[d_1 : \|\mathbf{b}_1\|^2 P_1\]
\[d_2 : \|\mathbf{b}_2\|^2 P_2\]
\[d_3 : \|\mathbf{b}_3\|^2 P_3\]

• This brings to mind the MSWF.

• This structure is only defined by the recursion

\[Q_i = P_i - \frac{\mathbf{b}_i^* \mathbf{b}_i}{Q_{i+1}}.\]
Intermediate Remarks

• The covariance matrix for the scalar outputs \( \{d_i\} \) is

\[
\begin{bmatrix}
  d_0 \\
  d_1 \\
  \vdots \\
  d_k
\end{bmatrix}
\begin{bmatrix}
  \|b_0\|^2 P_0 & \|b_1\|^2 & 0 \\
  \|b_1\|^2 & \|b_1\|^2 P_1 & \ddots \\
  \vdots & \ddots & \ddots & \|b_k\|^2 \\
  0 & \|b_k\|^2 & \|b_k\|^2 P_k
\end{bmatrix}.
\]

• The Schur complement recursions may be coded with a tree-structured diagram - which is the analysis stage of the MSWF. The same recursion decomposes the Capon beamformer into a continued fraction of Bartlett beamformers.
Continued Remarks

• To keep the Schur recursions going, or equivalently to tri-diagonalize, it is required that $b_i = B_{i-1}^* R_{i-1} b_{i-1}$ and $\langle B_i \rangle = \langle b_i \rangle^\perp$.

• The forced recursion makes $\|b_i\|^2$ and $P_i$ invariant to the basis for $\langle B_i \rangle$.

• In the Schur recursion and its equivalent MSWF realization, all parameters are invariant to the basis for $\langle B_i \rangle$ and there is no flexibility in the choice of the vectors $\{b_i\}$. The basis does influence $b_i$ but not its norm or $P_i$. 
Variations in the Final Stage

- If we terminate the MSWF at step $r$ where $r < (N - 1)$, we no longer need restrict the next “local” steering vector to maintain the tri-diagonalization recursion.

- We can exploit this property to better reduce MSE.
Design of the Final Basis Vector

- The MSE reduction in stage $r$ is

$$\Delta_r = \frac{g_r^* b_r b_r^* g_r}{g_r^* R_r g_r}$$

- The maximum reduction occurs with $g_r = R_r^{-1} b_r$.

- If this is too expensive restrict $g_r$

$$g_r = \alpha b_r + \beta R_r b_r.$$

- Under this restriction, minimization of $\Delta$ reduces to solving a $2 \times 2$ matrix pencil.
Data Variations

• Replace the input vector by a sequence of vectors

\[ \mathbf{x}_0[1] \ldots \mathbf{x}_0[m] \rightarrow \mathbf{d}_0[1] \ldots \mathbf{d}_0[m] \]

\[ \mathbf{b}_0^* \rightarrow \mathbf{x}_1[1] \ldots \mathbf{x}_1[m] \rightarrow \mathbf{B}_0^* \]

• Build a filter to minimize the weighted mean-squared error

\[
Q = \sum_i \sum_j (d[i] - \mathbf{v}^* \mathbf{x}_1[i]) \mathbf{M}_{ij} (d[j] - \mathbf{v}^* \mathbf{x}_1[j])^*
\]

\[
= (1 - \mathbf{v}^*) \mathbf{A}^* \mathbf{X} \mathbf{X}^* \mathbf{A} \begin{pmatrix} 1 \\ -\mathbf{v} \end{pmatrix}
\]
Data Variations - Cont.

• The decompositions are identical with $R_0$ replaced by $XMX^*$. In particular

$$Q_0 = \frac{1}{b_0^*(XMX^*)^{-1}b_0} = h_0^*(I - P_{H_0})h_0.$$  

where

$$h_0 = (XMX^*)^{1/2}b_0; \quad H_0 = (XMX^*)^{1/2}B_0.$$  

• If $XMX^*$ is rank-deficient, limit the number of stages to the rank of the data matrix or less.

• It remains to choose the weighting matrix $M$. 
Trading Diversity Gain for Beamforming Gain

- The matrix $M$ determines the trade-off between coherent gain and incoherent gain.

- Choose $M = I$ for full diversity, i.e. no coherence may be exploited from snapshot to snapshot.

- Choose $M = 11^T$ for the fully-coherent matched filter case when coherence may be exploited.

- Choose $M$: rank-$r$ to trade diversity $r$ for matched filtering gain $m/r$ when partial coherence may be exploited.
MSWF Properties

• The Wiener filter is the solution to the minimization problem

\[
\text{Minimize } J = [1 - \mathbf{v}^*] \mathbf{A}_0^* \mathbf{R}_0 \mathbf{A}_0 [1 - \mathbf{v}^*].
\]

• Multi-stage filters where \( \mathbf{v}_k = \alpha \mathbf{v}_{k-1} + \beta \mathbf{p}_k \) result from iteratively minimizing \( J \).

• The multi-stage Wiener filter corresponds to the conjugate gradient minimization method. Factor the tri-diagonal covariance matrix for \( \{d_i\} \) as \( \mathbf{T} = \mathbf{LDL}^* \) where \( \mathbf{L} \) is lower bi-diagonal. The matrix

\[
[\mathbf{b}_0 \ \mathbf{p}_1 \ \cdots \ \mathbf{p}_k] = [\mathbf{b}_0 \ \mathbf{B}_0 \mathbf{b}_1 \ \cdots \ \mathbf{B}_0 \mathbf{B}_1 \cdots \mathbf{B}_{k-1} \mathbf{b}_k] \mathbf{L}^{-1}
\]

has \( \mathbf{R}_0 \)-conjugate columns.

• There exists an efficient 3-term recursion (Lanczos - Golub and Van Loan) for generating \( \{\mathbf{p}_k\} \) - don’t have to compute expensive \( \mathbf{B}_0 \mathbf{B}_1 \cdots \mathbf{B}_{k-1} \mathbf{b}_k \).
Multi-Stage Variations

- Different iterative minimization methods produce different multi-stage filters (Pre-Conditioning? Chebychev iterations?).
- We can do a minimization with the constraint $\mathbf{v}^* \mathbf{v} < g^2$ to get an iterative white-noise gain constraint multi-stage filter.
Summary

• Standard Capon formula is a bulk formula in $\mathbf{R}^{-1}$.

• Schur formula is bulk formula in $\mathbf{R}^{-1}$.

• However, Schur may be recursively resolved to get a continued fraction representation of Capon in terms of Bartlett - Multi-Stage.

• Bulk Capon may be iteratively resolved with a conjugate-gradient search.