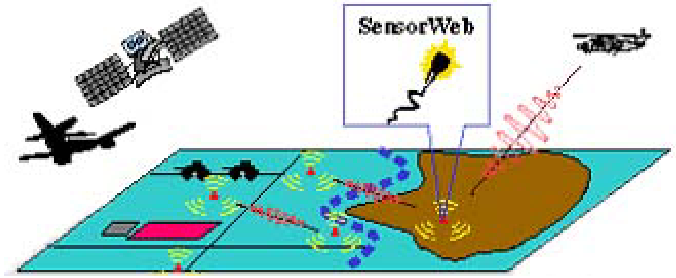
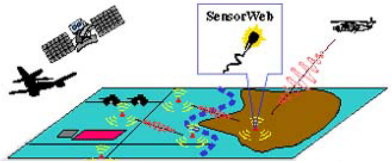


Data Association/Fusion for Heterogenous Sensors in Nonlinear and Dispersive Media

John Fisher

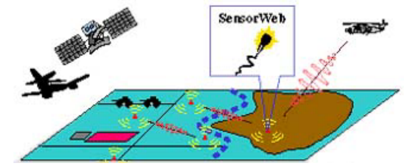
SensorWeb MURI Review Meeting

June 14, 2002

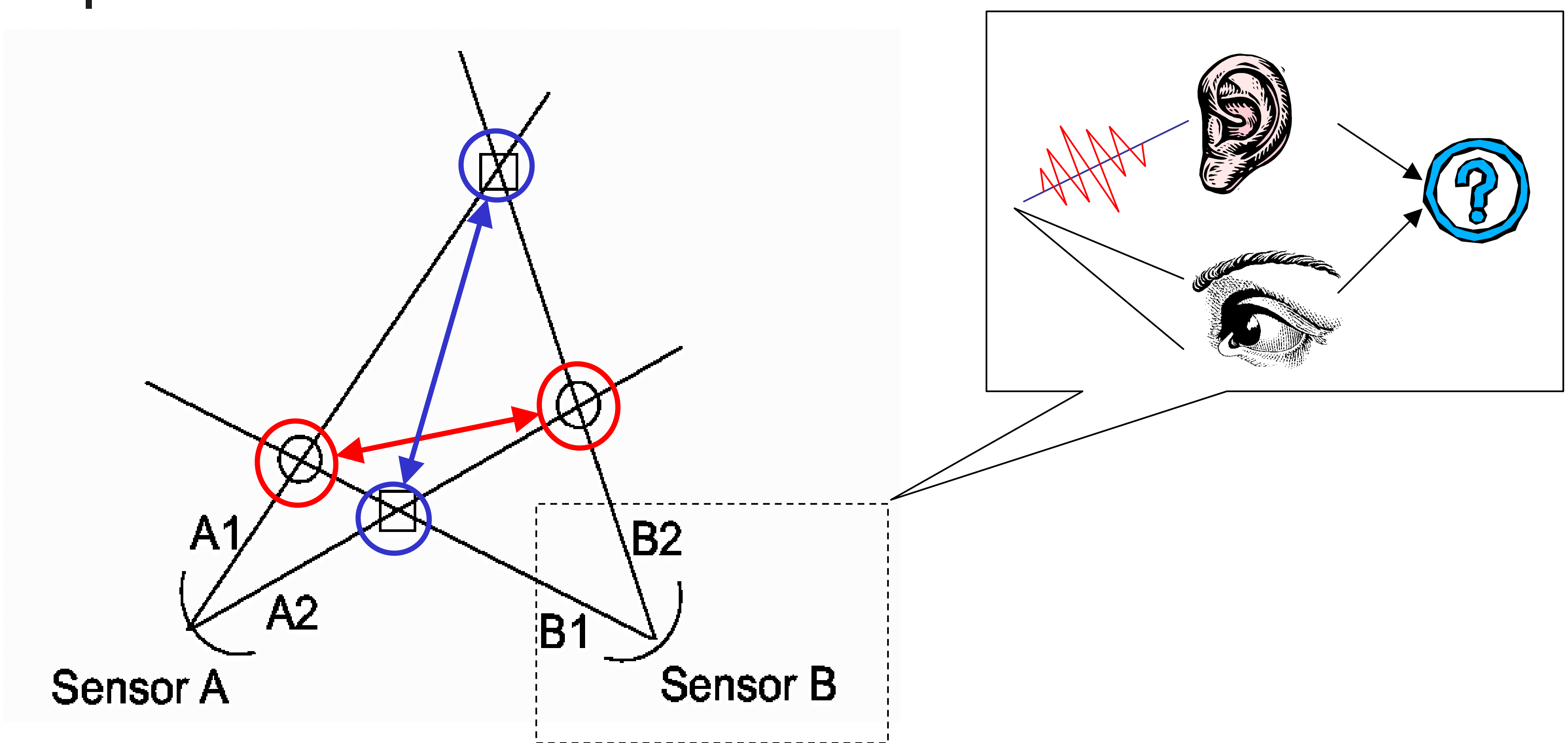


Goals

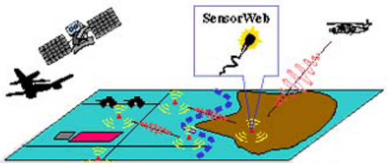
- Data association across nonlinear and/or dispersive channels
- Statistical models of heterogeneous sensors

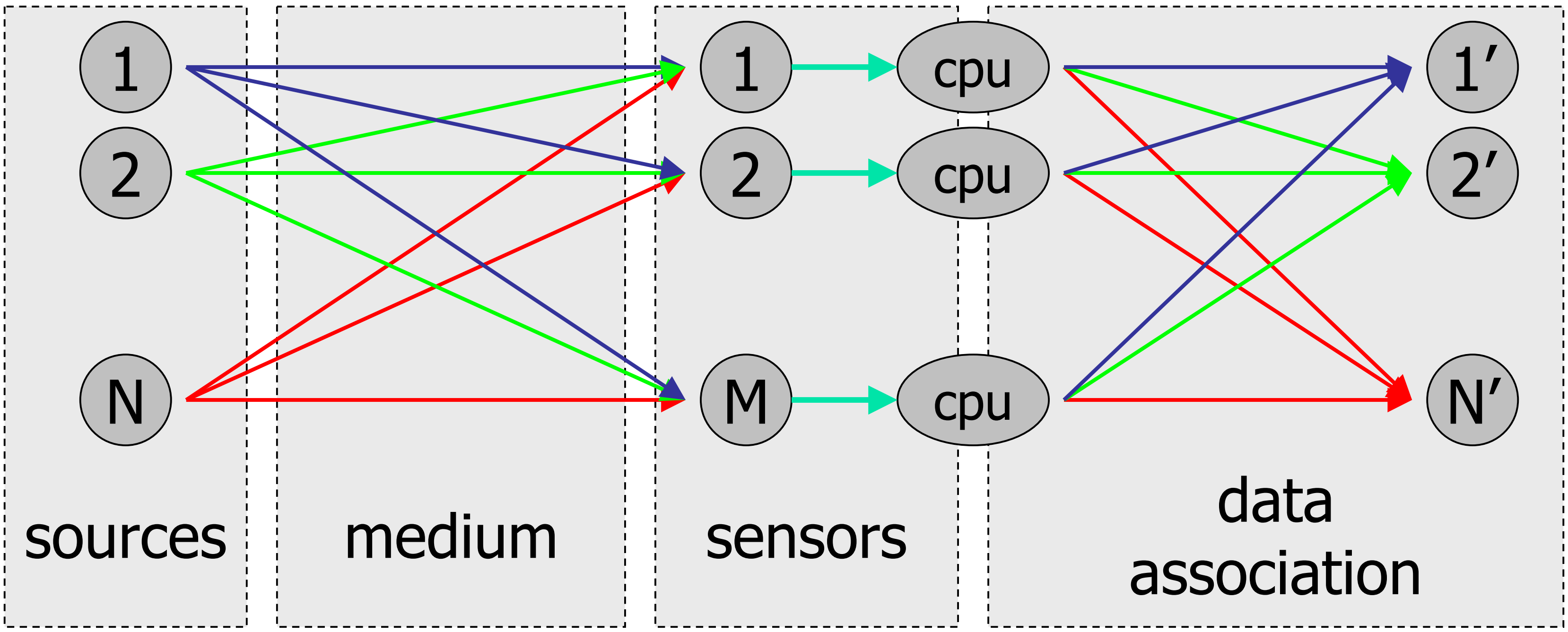


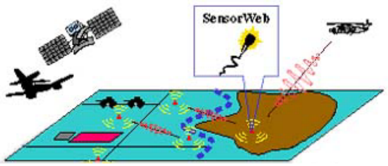
Data Association and Fusion



Fusion and Data Association



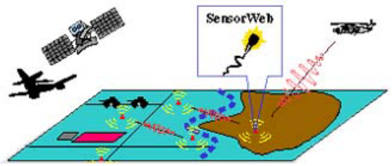




Issues and Challenges

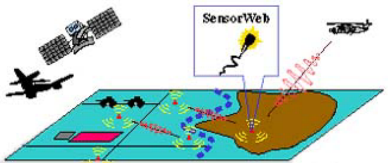
In the absence of a prior statistical model and given nonlinear and/or dispersive media

- How do we
  - solve data association problem in a principled manner?
  - fuse/utilize multi-modal data?
- Can we
  - propagate the results to higher level algorithms in the form of likelihoods or scores?

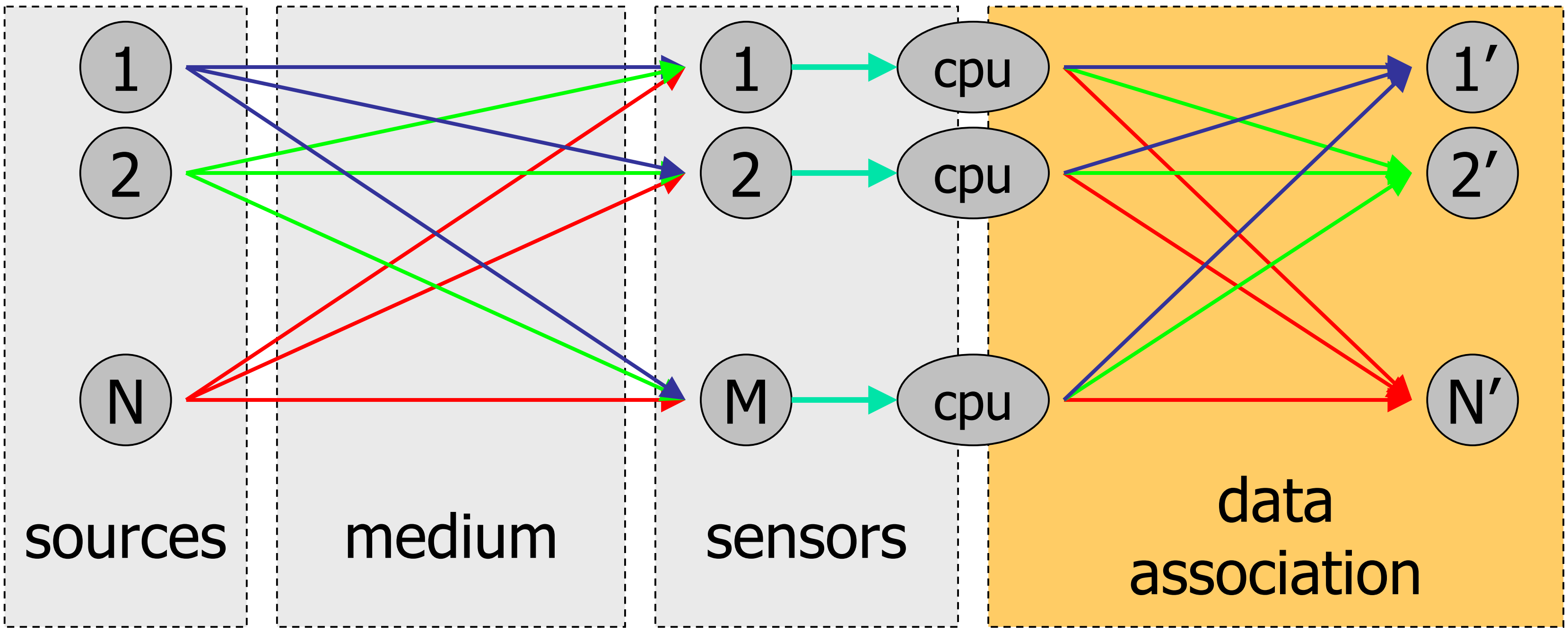


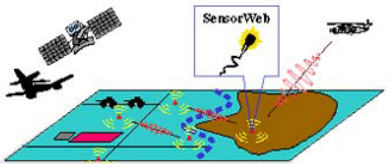
Vital Statistics

- IT-2 (fusion of heterogeneous sensors in unstructured and uncertain environments)
- RCA-1 (self-calibration)
- RCA-5 (fusion algorithms)
- RCA-7 (data for experiments and demos)
- Ties to RCA-2&3 (Tradeoffs in local vs. global processing)
- Contributors
  - Fisher, Ihler, Çetin, Willsky
- Preliminary outputs
  - Several publications and talks
  - A number of academic, industrial, and DoD interactions

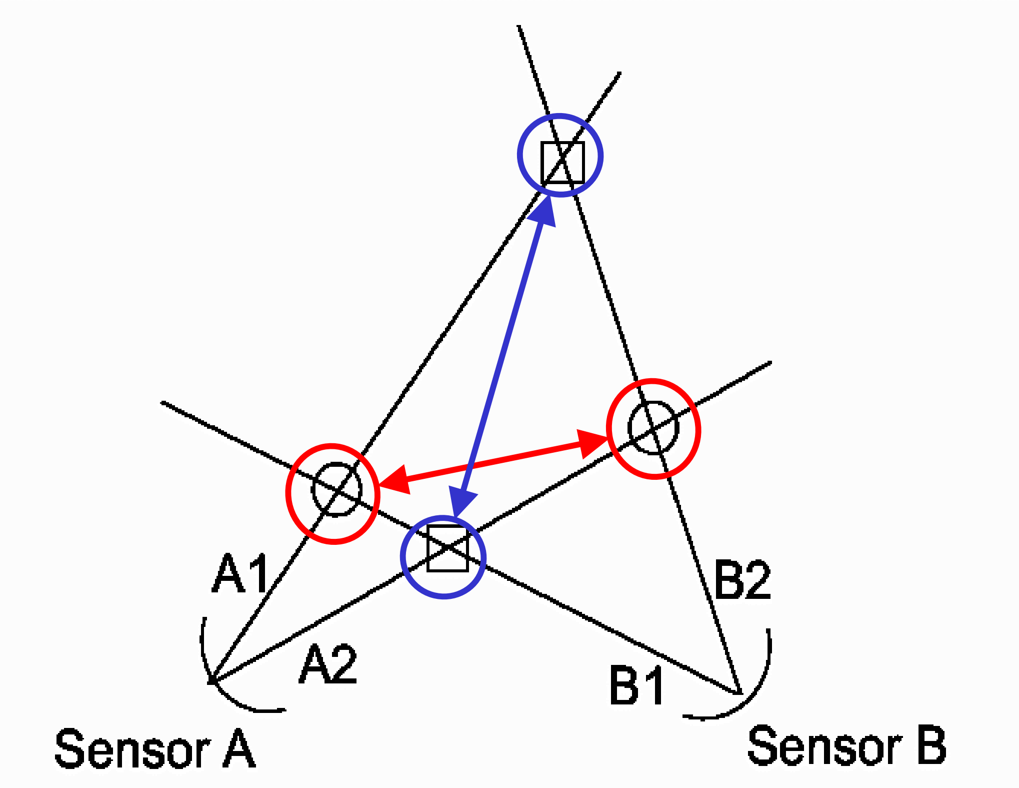


Data Association

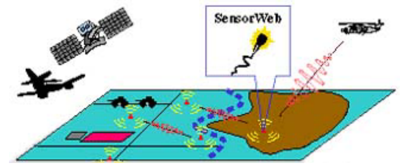




Data Association Problem



- Sensors receive data and associated bearing estimates
- Bearing alone results in an inherent ambiguity
- Correspondence problem between A1,2 and B1,2



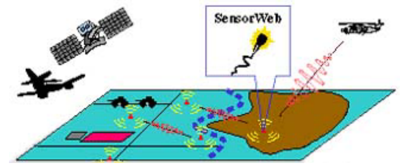
Data Association as a Hypothesis Testing Problem

Typically, one incorporates factorization *and* and an assumed parameterized density under either hypothesis

$$H_1 : [A_1, A_2, B_1, B_2]_i \sim p_{H_1}(A_1, A_2, B_1, B_2) = p_{H_1}(A_1, B_1)\, p_{H_1}(A_2, B_2)$$
$$H_2 : [A_1, A_2, B_1, B_2]_i \sim p_{H_2}(A_1, A_2, B_1, B_2) = p_{H_2}(A_1, B_2)\, p_{H_2}(A_2, B_1)$$

resulting in a log-likelihood of

$$\log L = \sum_i \log\left(\frac{p_{H_1}([A_1, B_1]_i)\, p_{H_1}([A_2, B_2]_i)}{p_{H_2}([A_1, B_2]_i)\, p_{H_2}([A_2, B_1]_i)} \right)$$
$$= \sum_i \log\left(\frac{p_{H_1}([A_1, B_1]_i)\, p_{H_1}([A_2, B_2]_i)}{p(A_{1i})\, p(A_{2i})\, p(B_{1i})\, p(B_{2i})} \cdot \frac{p(A_{1i})\, p(A_{2i})\, p(B_{1i})\, p(B_{2i})}{p_{H_2}([A_1, B_2]_i)\, p_{H_2}([A_2, B_1]_i)} \right)$$



Data Association as a Hypothesis Testing Problem

Expectation (or limit) under H_1 yields

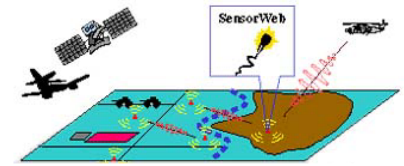
$$
\begin{aligned}
E_{H_1}\{\log L\} \propto\ & I(A_1;B_1)+I(A_2;B_2)+ \\
& D\left(p_{H_1}(A_1)\,p_{H_1}(B_2)\,\|\,p_{H_2}([A_1,B_2])\right)+ \\
& D\left(p_{H_1}(A_2)\,p_{H_1}(B_1)\,\|\,p_{H_2}([A_2,B_1])\right)
\end{aligned}
$$

under H_2 yields

$$
\begin{aligned}
E_{H_2}\{\log L\} \propto\ & -I(A_1;B_2)-I(A_2;B_1)- \\
& D\left(p_{H_2}(A_1)\,p_{H_2}(B_1)\,\|\,p_{H_1}([A_1,B_1])\right)- \\
& D\left(p_{H_2}(A_2)\,p_{H_2}(B_2)\,\|\,p_{H_1}([A_2,B_2])\right)
\end{aligned}
$$

Terms related to association (i.e. statistical dependency)

Terms related to density modeling assumptions



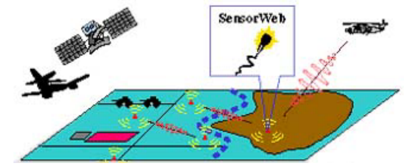
What if prior models are not available?

- Suppose we have "perfect" density estimates from the measurements.
- We can estimate the factorization under H_1 and H_2 from data, but the terms due to the assumed density go away.
- Data association becomes a hypothesis test over factorizations

$$E_{H_1}\{\log L\} \propto I(A_1, B_1) + I(A_2, B_2)$$

$$E_{H_2}\{\log L\} \propto -\left(I(A_1, B_2) + I(A_2, B_1)\right)$$

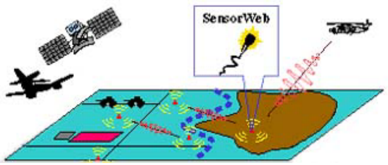
  - So we lose the benefit of parametric density terms, but …
  - We also don't suffer when the density model terms are wrong.



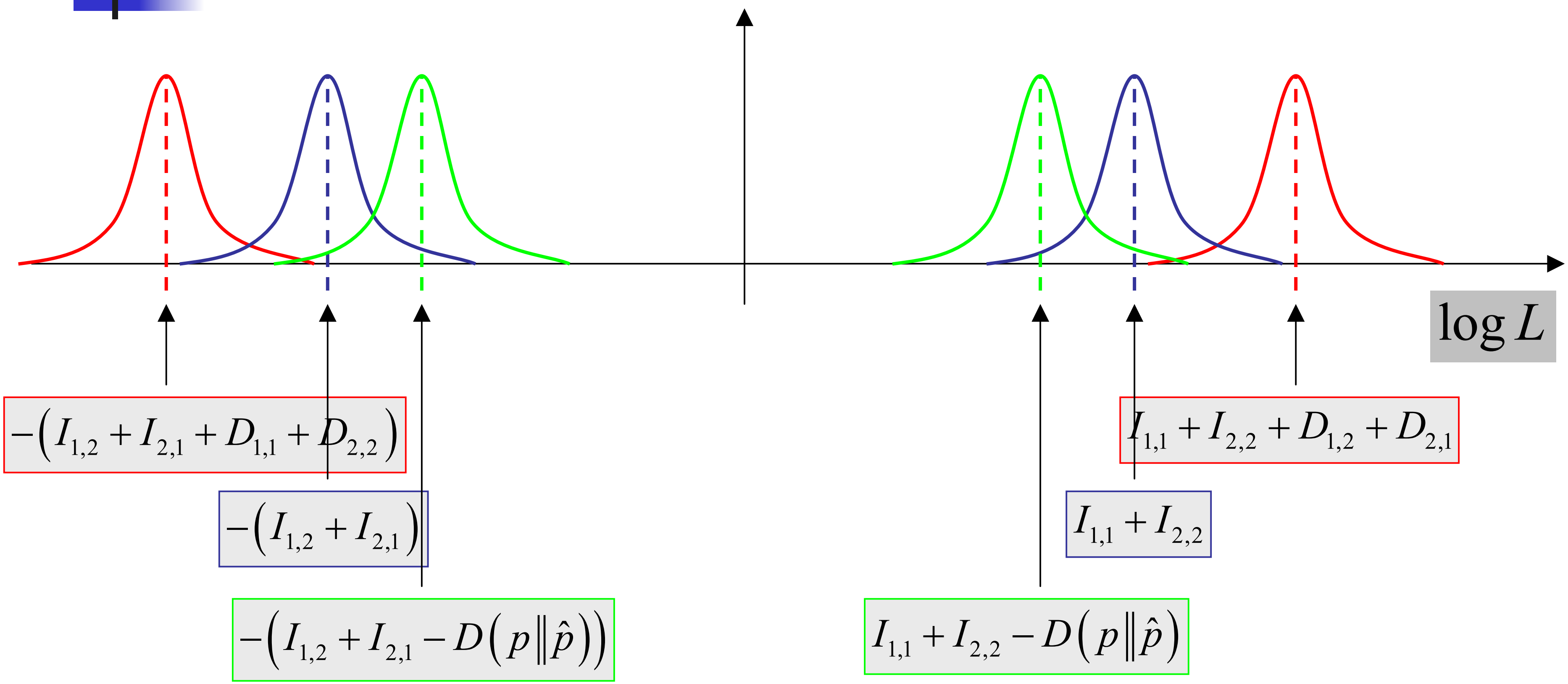
Density Estimation Effects

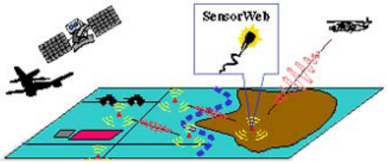
- Of course, we have imperfect density estimates introducing additional biases in our log-likelihood computation.

$$
\begin{aligned}
E_{H_1}\{\log L\} \propto\ & I(A_1,B_1)+I(A_2,B_2)- \\
& D\big(p(A_1,B_1)\,p(A_2,B_2)\,\big\|\,\hat{p}(A_1,B_1)\,\hat{p}(A_2,B_2)\big)+ \\
& D\big(p(A_1)\,p(A_2)\,p(B_1)\,p(B_2)\,\big\|\,\hat{p}(A_1,B_2)\,\hat{p}(A_2,B_1)\big) \\
E_{H_2}\{\log L\} \propto\ & -\big(I(A_1,B_2)+I(A_2,B_1)- \\
& D\big(p(A_1,B_2)\,p(A_2,B_1)\,\big\|\,\hat{p}(A_1,B_2)\,\hat{p}(A_2,B_1)\big)+ \\
& D\big(p(A_1)\,p(A_2)\,p(B_1)\,p(B_2)\,\big\|\,\hat{p}(A_1,B_1)\,\hat{p}(A_2,B_2)\big)\big)
\end{aligned}
$$



Distribution over log-likelihoods



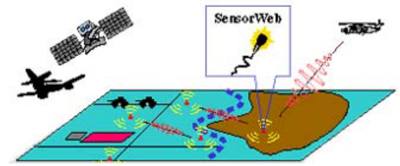


Dimensionality – Feature Extraction

- High dimensionality precludes direct density estimation.
- Dependence may be captured in a low-dimensional subspace (or manifold).
- Compromise by projecting to a lower dimensional space.

$$H_1 : [A_1, A_2, B_1, B_2] \rightarrow \left[f\left(A_1;\alpha_{11}\right), f\left(A_2;\alpha_{21}\right), f\left(B_1;\beta_{11}\right), f\left(B_2;\beta_{21}\right) \right]$$
$$H_2 : [A_1, A_2, B_1, B_2] \rightarrow \left[g\left(A_1;\alpha_{12}\right), g\left(A_2;\alpha_{22}\right), g\left(B_1;\beta_{12}\right), g\left(B_2;\beta_{22}\right) \right]$$

- Use kernel density estimator to compensate for complexity in the feature space.
- Objective is to do this without training.



Dimensionality – Feature Extraction

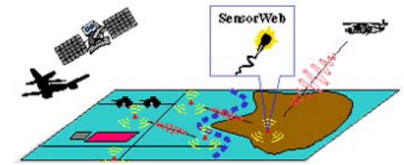
- By choosing the projection coefficients to maximize MI under the hypothesis we minimize the deviation of the LRT in the feature space from the LRT in the measurement space

$$H_1: \arg\max_{\alpha_{11},\alpha_{21},\beta_{11},\beta_{21}} \left[I\left(f\left(A_1;\alpha_{11}\right); f\left(B_1;\beta_{11}\right)\right) + I\left(f\left(A_2;\alpha_{21}\right); f\left(B_2;\beta_{21}\right)\right) \right]$$

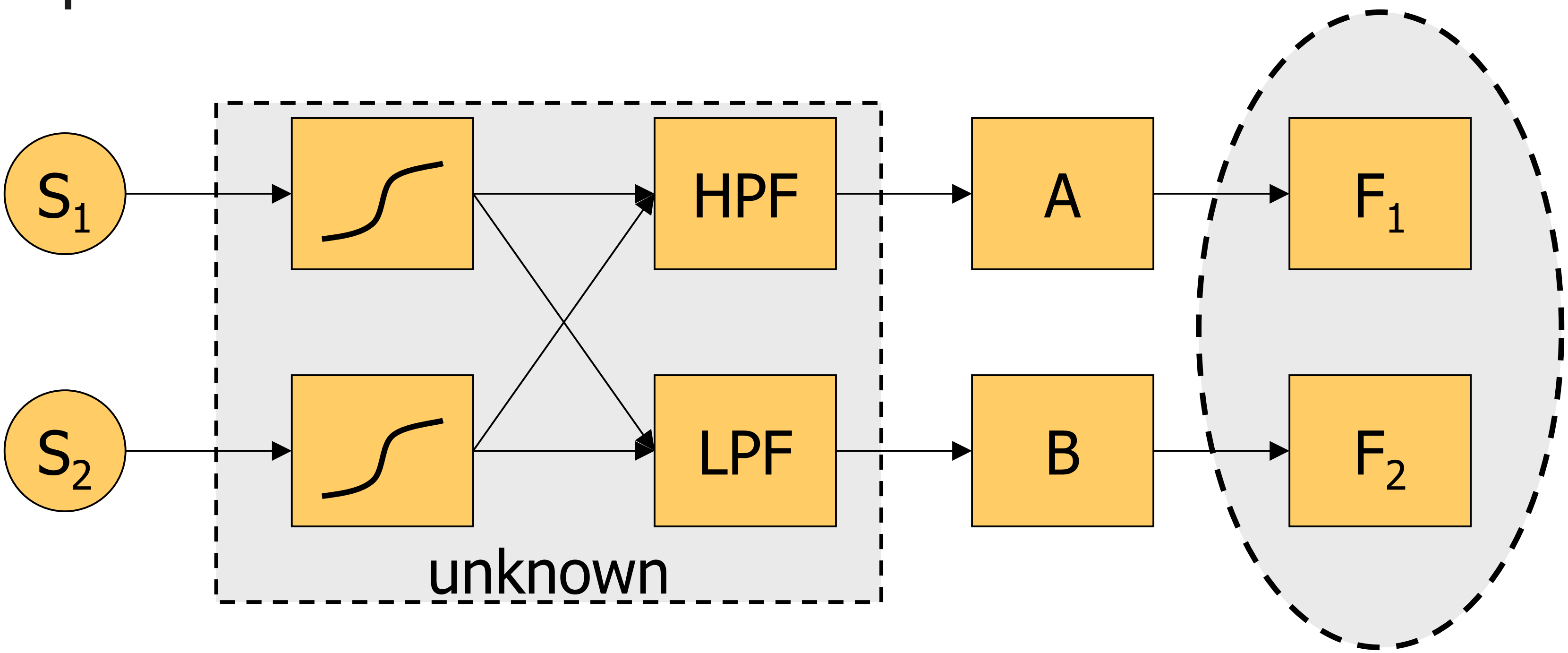
$$H_2: \arg\max_{\alpha_{12},\alpha_{22},\beta_{12},\beta_{22}} \left[I\left(g\left(A_1;\alpha_{12}\right); g\left(B_2;\beta_{22}\right)\right) + I\left(g\left(A_2;\alpha_{22}\right); g\left(B_1;\beta_{12}\right)\right) \right]$$

$$\sum_i \log\left(\frac{p_{H_1}\left(\left[A_1, B_1\right]_i\right) p_{H_1}\left(\left[A_2, B_2\right]_i\right)}{p_{H_2}\left(\left[A_1, B_2\right]_i\right) p_{H_2}\left(\left[A_2, B_1\right]_i\right)} \right)$$

$$\sum_i \log\left(\frac{\tilde{p}_{H_1}\left(\left[f\left(A_1;\alpha_{11}\right), f\left(B_1;\beta_{11}\right)\right]_i\right) \tilde{p}_{H_1}\left(\left[f\left(A_2;\alpha_{21}\right), f\left(B_2;\beta_{21}\right)\right]_i\right)}{\tilde{p}_{H_2}\left(\left[g\left(A_1;\alpha_{12}\right), g\left(B_2;\beta_{22}\right)\right]_i\right) \tilde{p}_{H_2}\left(\left[g\left(A_2;\alpha_{22}\right), g\left(B_1;\beta_{12}\right)\right]_i\right)} \right)$$

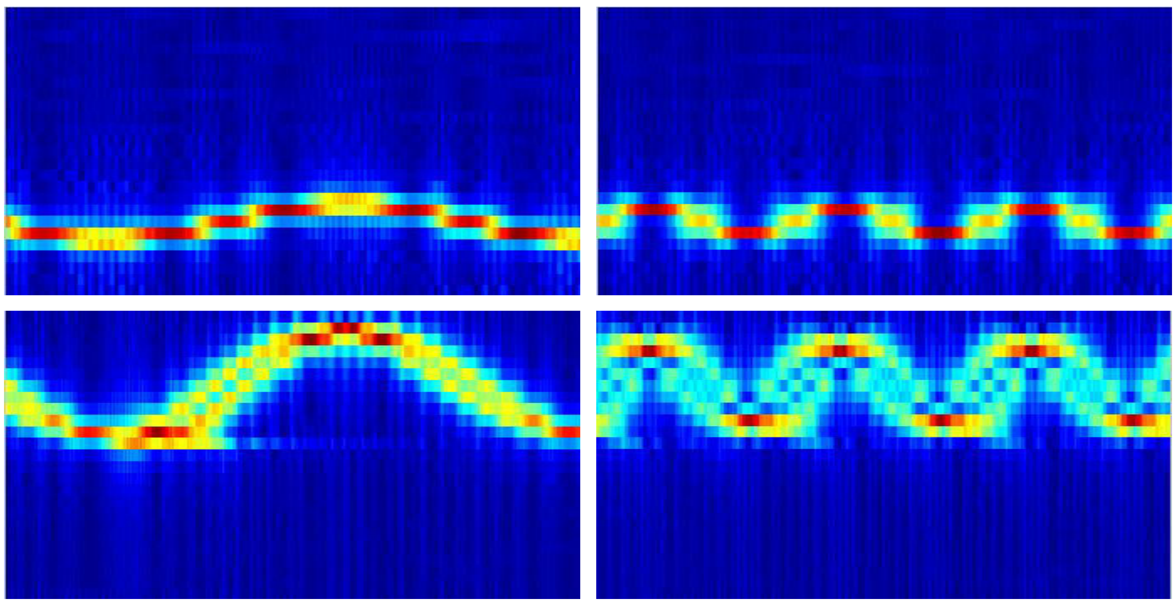


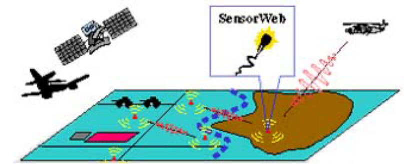
Nonlinear Channel



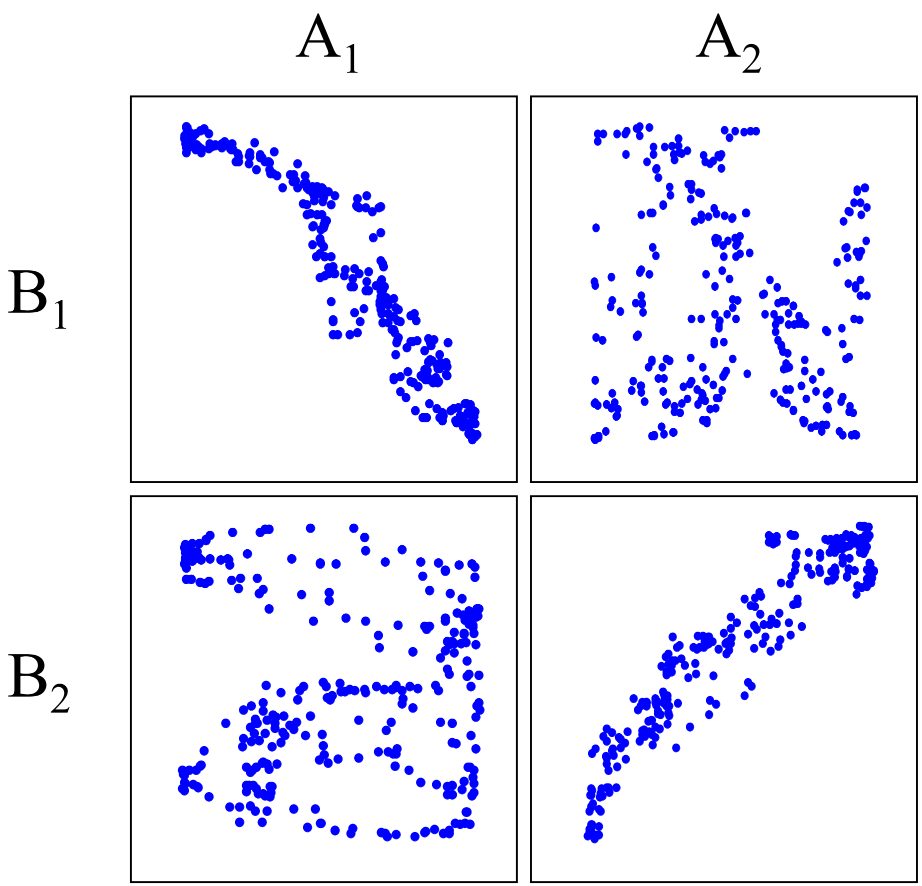
Received signals are uncorrelated but *not* independent

Narrowband, Uncorrelated Signals

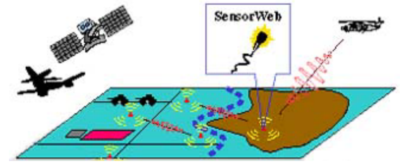




Feature Space



- Spectra projected down to single scalar value.
- Density estimated in MI optimized feature space.
- MI (data association log-likelihood) computed over feature space.
- Dependence is clear.

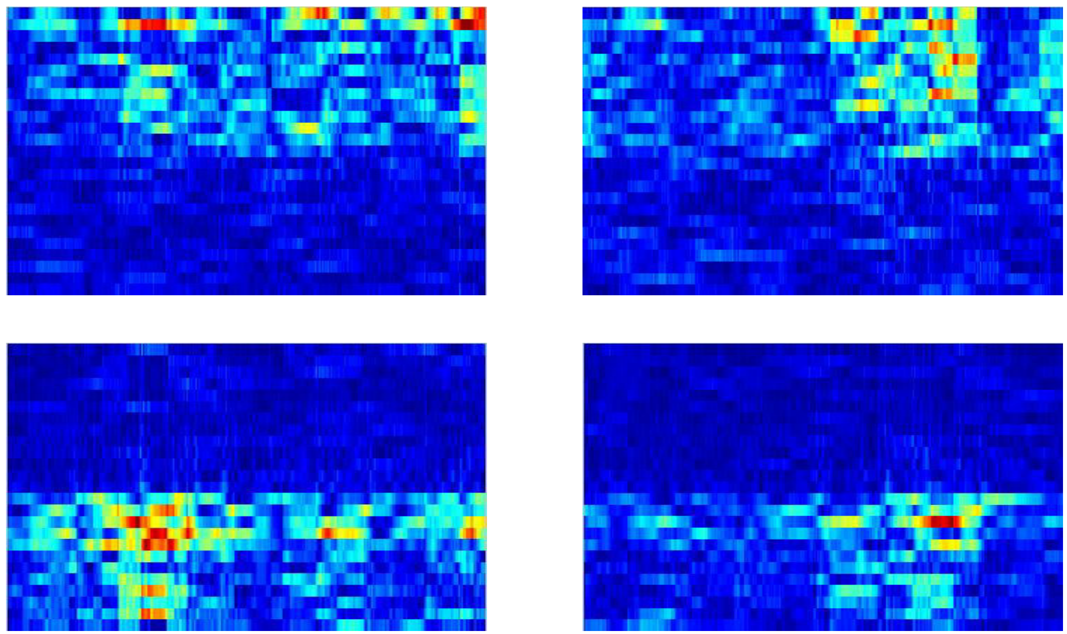


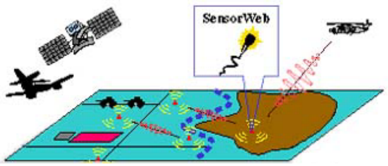
Wideband, Uncorrelated Signals

Signal 1 | Signal 2

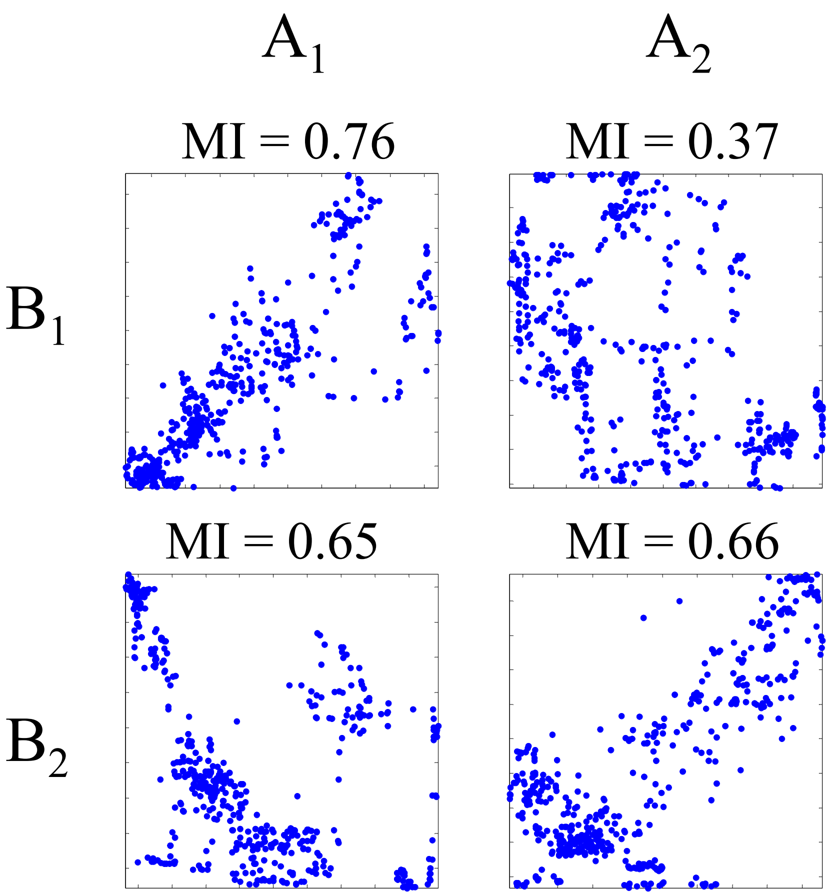
Sensor A

Sensor B



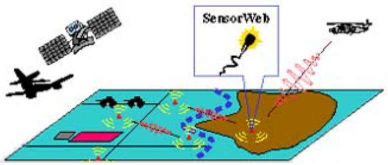


Feature Space

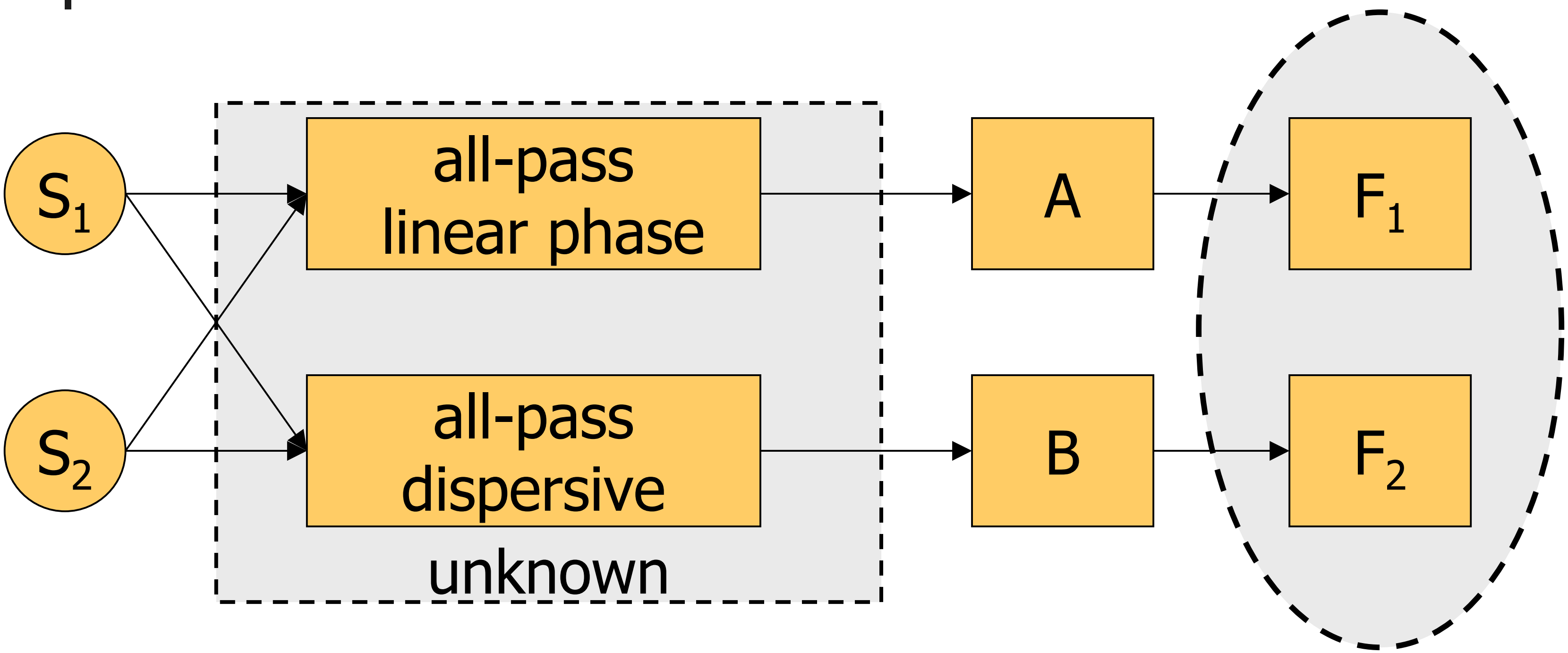


- Association is less obvious for wideband case.
- This is also reflected in MI values.
- Combined score still chooses correct association

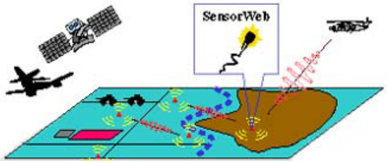
0.143 (1-1,2-2) > 0.102 (1-2,2-1)



Dispersive Medium

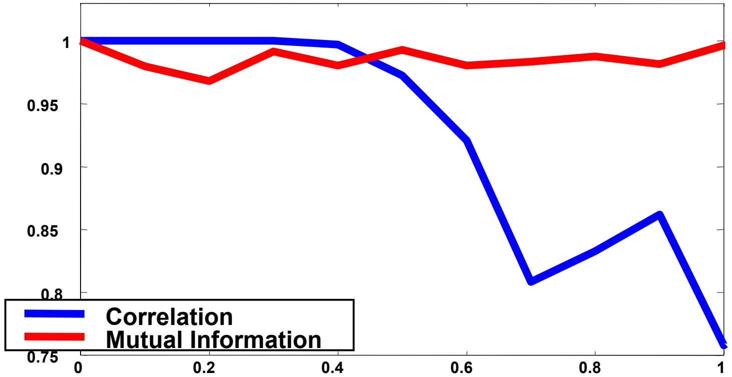


Both correlation and dependence degrade, but not at the same rate

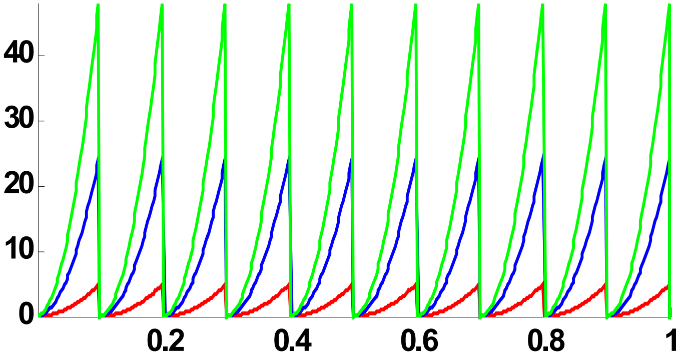


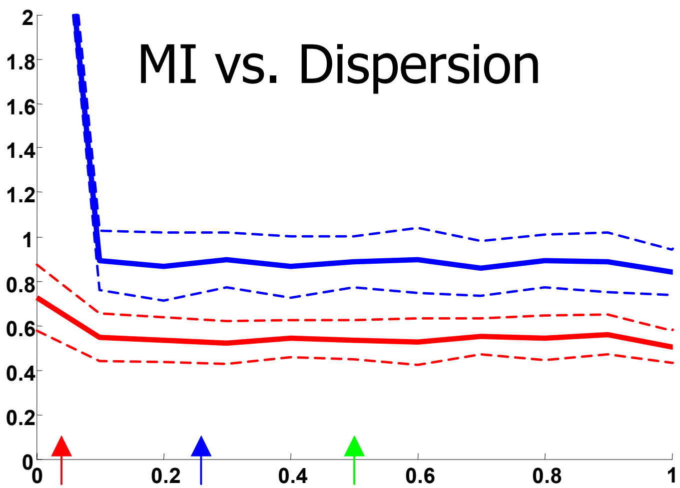
Dispersive Medium

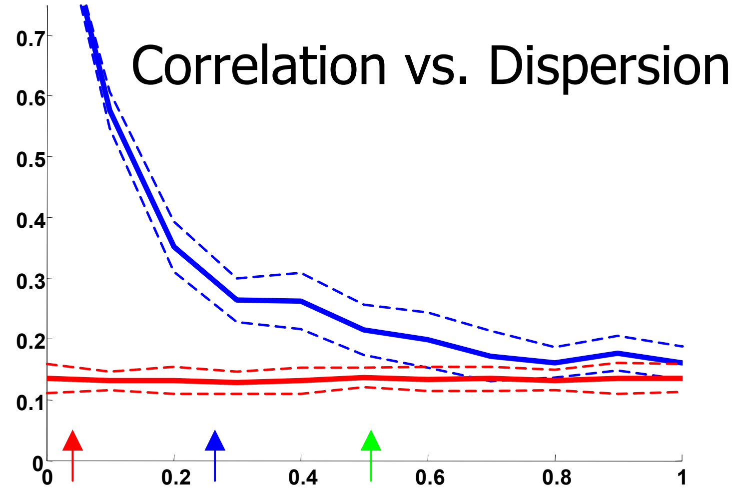
P{correct association}

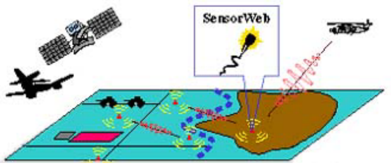


Phase response of media

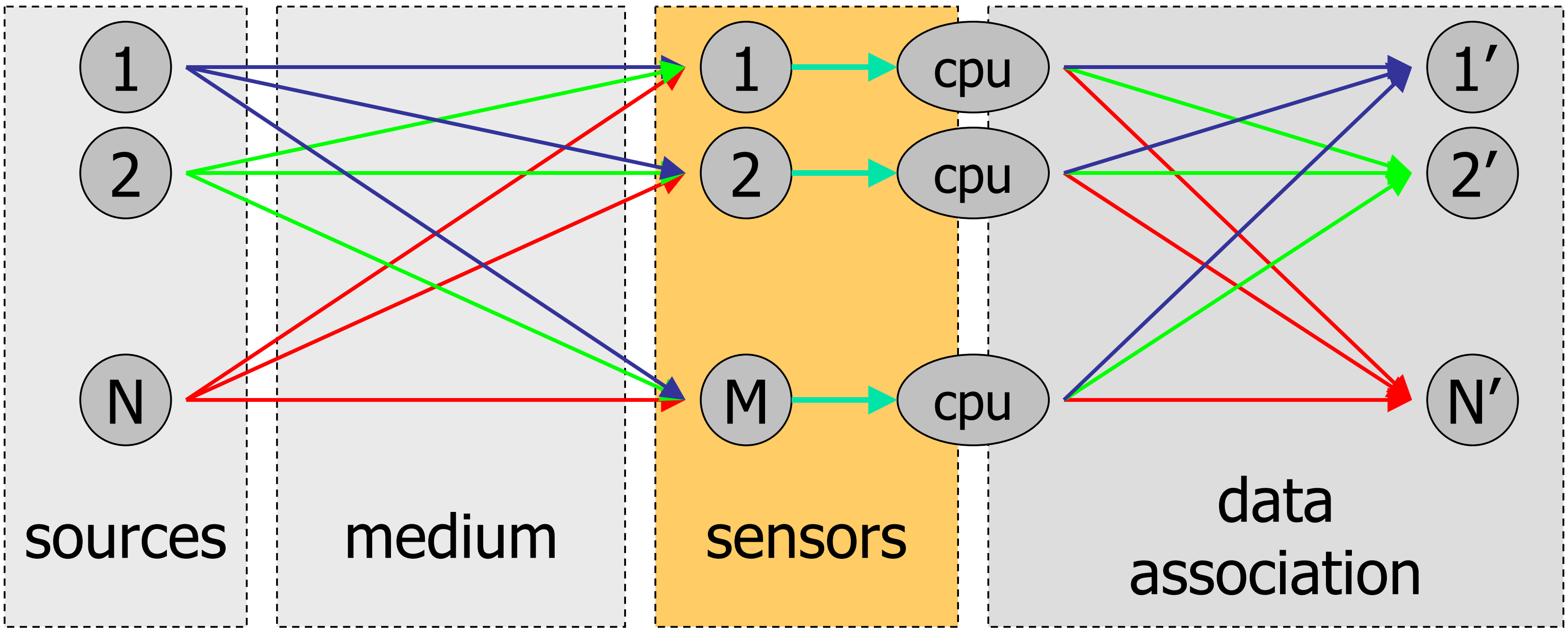


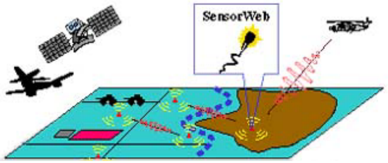




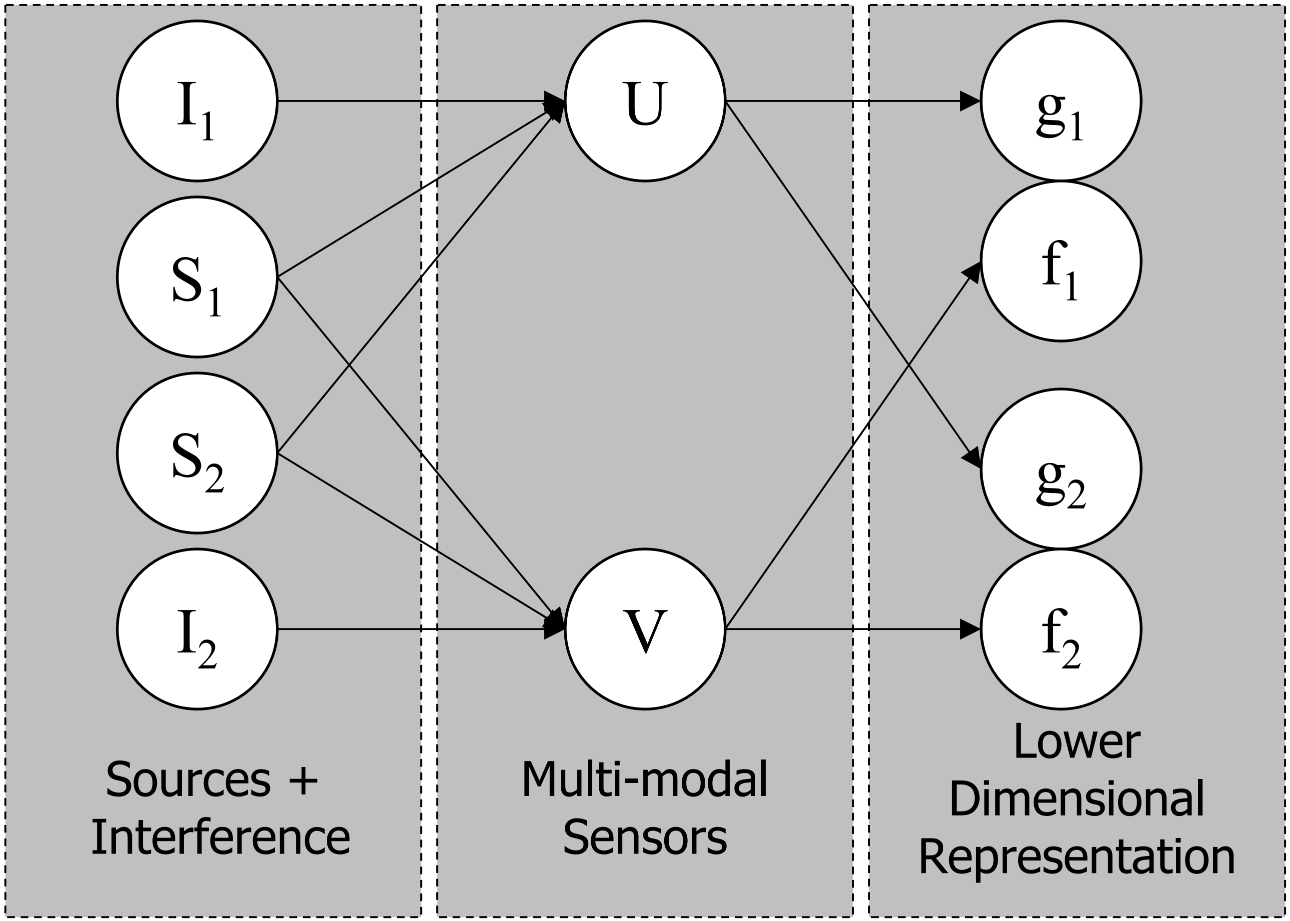


Information Theoretic Sensor Fusion

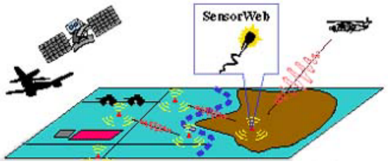




Last Year



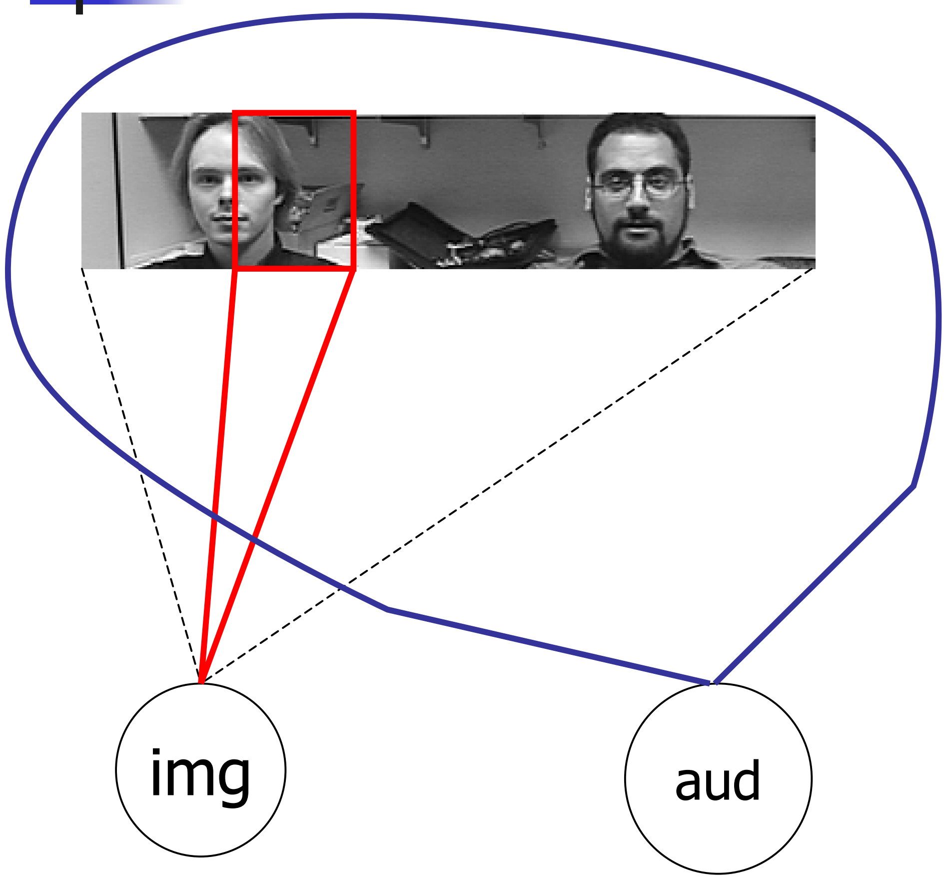
Maximizing $I(g_1;f_1)$, $I(g_2;f_2)$ and $H([g_1,f_1],[g_2,f_2])$ recovers a representation of the sources up to a permutation (re: data association)



Acoustically Steered Imaging Sensor

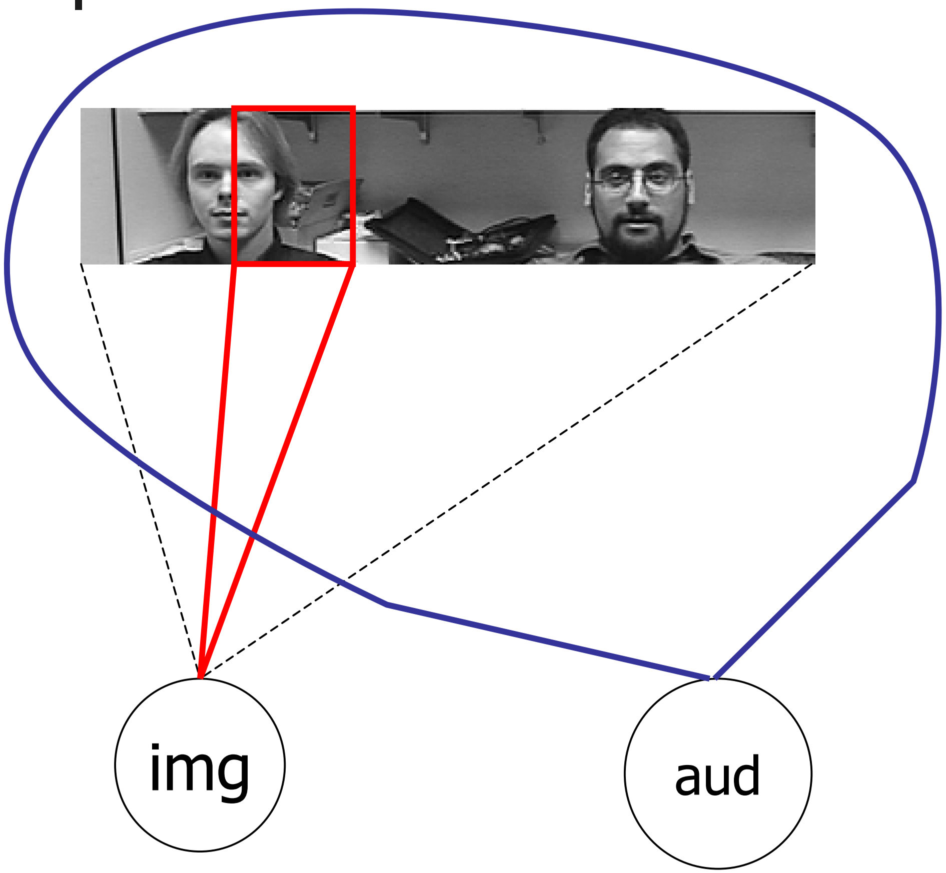
- Last time we presented our local fusion approach and justified it statistically.
- Here we focus on an application in the sensor domain considering a single narrow field of view imaging sensor (e.g. IR or video) guided by broad field of view acoustic sensors.

acoustically steered imaging sensor

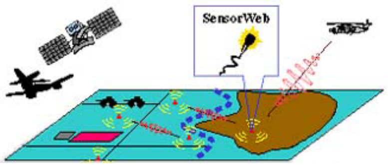


- Imaging sensor has a narrow field of view, but can be steered.
- Acoustic sensor has little directivity (broad field of view)

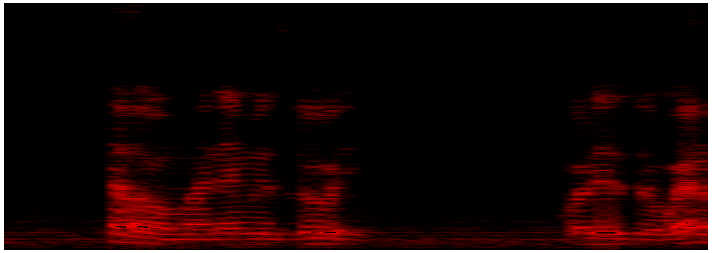
acoustically steered imaging sensor

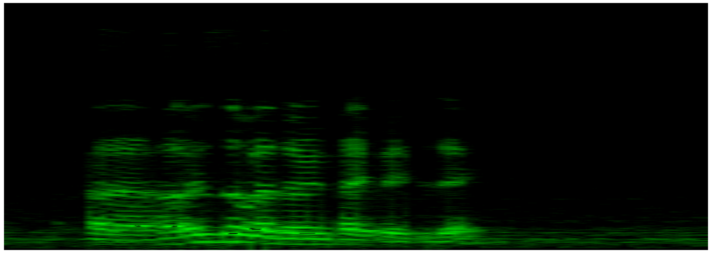


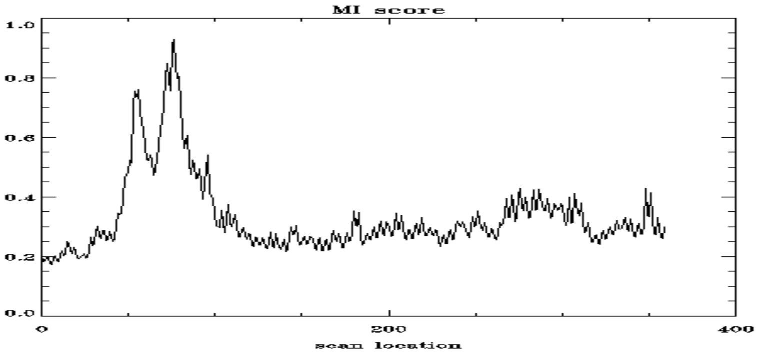
- Use local fusion to derive bearing to source when
  - One source is emitting an acoustic signal - opportunistic
  - Both sources emit acoustic signals simultaneously – local ambiguity

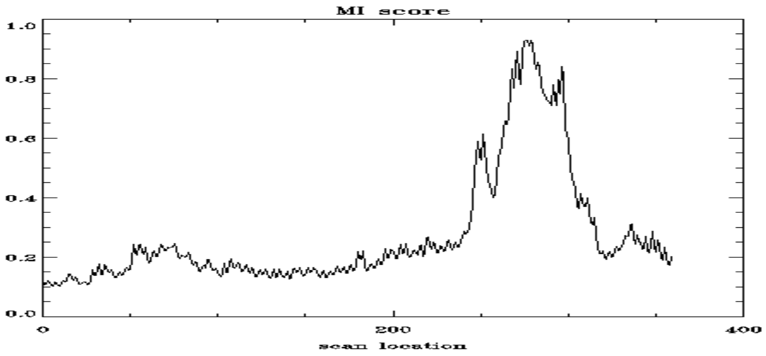


Single source detection









Multi Source Separation