Blind Single Channel Deconvolution using Nonstationary Signal Processing

Reverberation Cancellation in Acoustic Environments

James R. Hopgood
`http://www-sigproc.eng.cam.ac.uk/~jrh1008/`
`jrh1008@eng.cam.ac.uk`

Signal Processing Laboratory, Department of Engineering, University of Cambridge

UNIVERSITY OF CAMBRIDGE

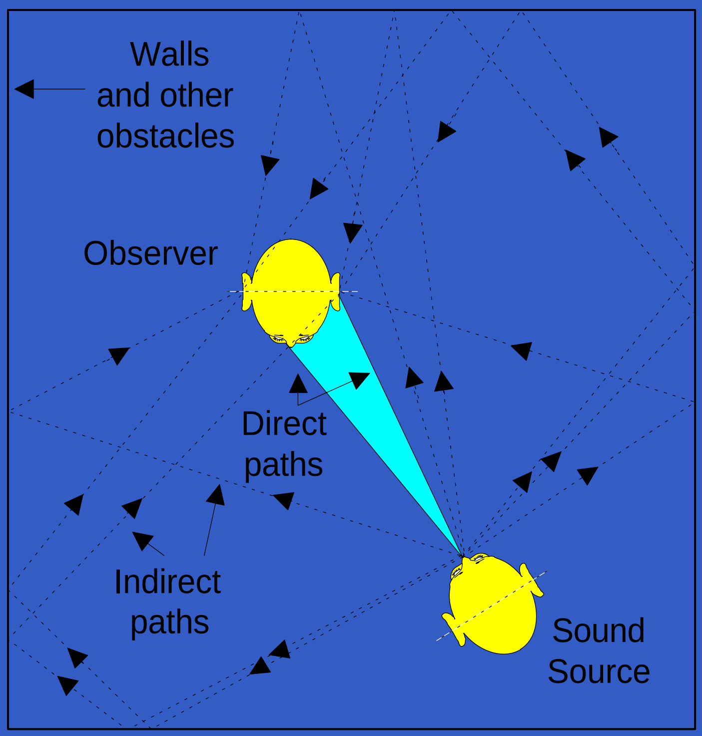
Go to full-screen mode now by hitting CTRL-L

Introduction to Blind Deconvolution

- Blind deconvolution fundamental in signal processing
- Observation, $\mathbf{x} = \{x(t),\, t \in \mathcal{T}\}$, modelled as the convolution of unknown source, $\{s(t),\, t \in \mathcal{T}\}$, with unknown distortion operator, $\mathcal{A}$; *i.e.* $x(t) \triangleq s(t) \star \mathcal{A}$

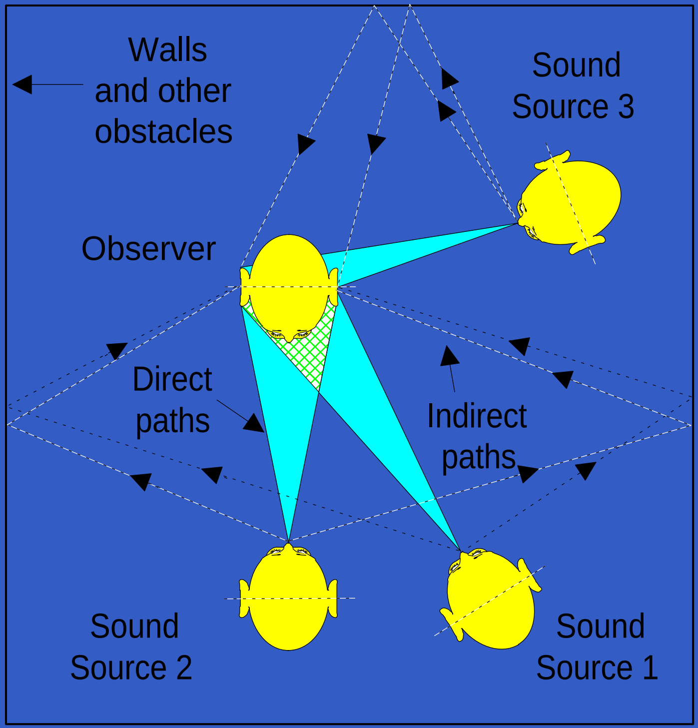
- Estimate $\mathcal{A}$, *or* $\hat{s}(t) = a\, s(t - \tau)$, a scaled shifted version of $s(t)$, where $a,\, \tau \in \mathbb{R}$, given only the observations, $\mathbf{x}$

Acoustic Reverberation Cancellation



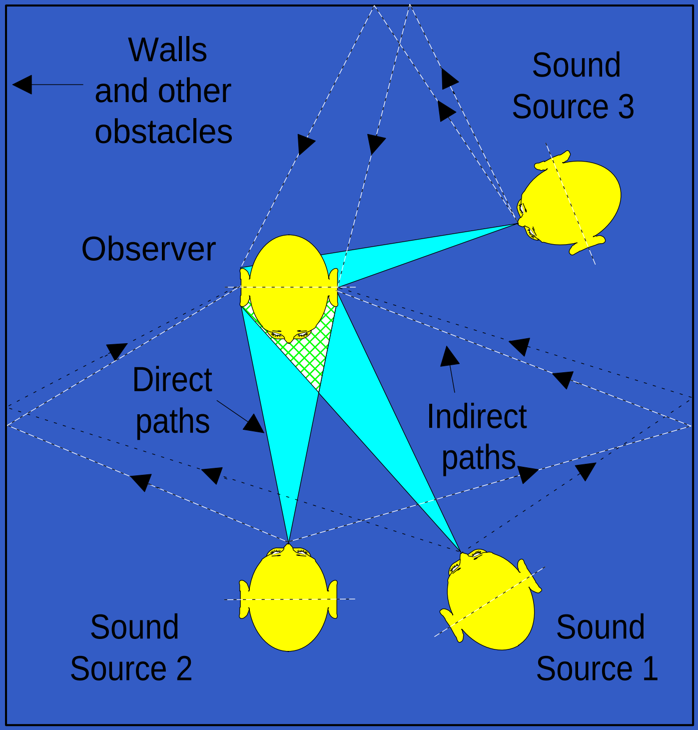
- Normal hearing: can concentrate on original sound despite:
  - reverberation

Acoustic Reverberation Cancellation



- Normal hearing: can concentrate on original sound despite:
  - reverberation
  - environmental noise

Acoustic Reverberation Cancellation



- Normal hearing: can concentrate on original sound despite:
  - reverberation
  - environmental noise
- Hearing aid users unable to distinguish one voice from another
- Sensori-neuro loss cannot be compensated for by simple amplifying hearing aids or surgery

UNIVERSITY OF CAMBRIDGE

Bayesian Blind Deconvolution

Bayes's Theorem

- The *posterior probability*, $p\left(\boldsymbol{\theta} \mid \mathbf{x}, \mathcal{I}\right)$, of the system parameters, $\boldsymbol{\theta}$, given the state of the system, $\mathbf{x}$, and an underlying model, $\mathcal{I}$, is given by Bayes' theorem:

$$p\left(\boldsymbol{\theta} \mid \mathbf{x}, \mathcal{I}\right) = \frac{p\left(\mathbf{x} \mid \boldsymbol{\theta}, \mathcal{I}\right) p\left(\boldsymbol{\theta} \mid \mathcal{I}\right)}{p\left(\mathbf{x} \mid \mathcal{I}\right)}$$

Bayes's Theorem

- The *posterior probability*, $p\,(\boldsymbol{\theta} \mid \mathbf{x}, \mathcal{I})$, of the system parameters, $\boldsymbol{\theta}$, given the state of the system, $\mathbf{x}$, and an underlying model, $\mathcal{I}$, is given by Bayes' theorem:

$$p\,(\boldsymbol{\theta} \mid \mathbf{x}, \mathcal{I}) = \frac{p\,(\mathbf{x} \mid \boldsymbol{\theta}, \mathcal{I})\, p\,(\boldsymbol{\theta} \mid \mathcal{I})}{p\,(\mathbf{x} \mid \mathcal{I})}$$

- $p\,(\mathbf{x} \mid \boldsymbol{\theta}, \mathcal{I})$ is the *likelihood*

Bayes's Theorem

- The *posterior probability*, $p(\boldsymbol{\theta} \mid \mathbf{x}, \mathcal{I})$, of the system parameters, $\boldsymbol{\theta}$, given the state of the system, $\mathbf{x}$, and an underlying model, $\mathcal{I}$, is given by Bayes' theorem:

$$p(\boldsymbol{\theta} \mid \mathbf{x}, \mathcal{I}) = \frac{p(\mathbf{x} \mid \boldsymbol{\theta}, \mathcal{I})\, p(\boldsymbol{\theta} \mid \mathcal{I})}{p(\mathbf{x} \mid \mathcal{I})}$$

- $p(\mathbf{x} \mid \boldsymbol{\theta}, \mathcal{I})$ is the *likelihood*
- $p(\boldsymbol{\theta} \mid \mathcal{I})$ represents prior knowledge

Bayes's Theorem

- The *posterior probability*, $p\left(\boldsymbol{\theta} \mid \mathbf{x}, \mathcal{I}\right)$, of the system parameters, $\boldsymbol{\theta}$, given the state of the system, $\mathbf{x}$, and an underlying model, $\mathcal{I}$, is given by Bayes' theorem:

$$p\left(\boldsymbol{\theta} \mid \mathbf{x}, \mathcal{I}\right) = \frac{p\left(\mathbf{x} \mid \boldsymbol{\theta}, \mathcal{I}\right) p\left(\boldsymbol{\theta} \mid \mathcal{I}\right)}{p\left(\mathbf{x} \mid \mathcal{I}\right)}$$

- $p\left(\mathbf{x} \mid \boldsymbol{\theta}, \mathcal{I}\right)$ is the *likelihood*
- $p\left(\boldsymbol{\theta} \mid \mathcal{I}\right)$ represents prior knowledge
- $p\left(\mathbf{x} \mid \mathcal{I}\right)$ is the *evidence* and, although usually regarded as a normalising constant, is of interest for model selection

Likelihood Function for System

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \,|\, \boldsymbol{\theta}, \boldsymbol{\phi}, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{(\sqrt{2\pi}\sigma_i)^{T_i}} \exp\left\{ -\frac{(\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i)^T (\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i)}{2{\sigma_i}^2} \right\}$$

- Where $s(t) \equiv s(t, \mathbf{a}, \mathbf{x})$ given by:

$$s(t) = x(t) + \sum_{p \in \mathcal{P}} a(p)\, x(t-p)$$

Likelihood Function for System

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\phi}, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{(\sqrt{2\pi}\sigma_i)^{T_i}} \exp\left\{ -\frac{\left(\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i\right)^T \left(\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i\right)}{2{\sigma_i}^2} \right\}$$

- Data vector, $\mathbf{s}_i$: $[\mathbf{s}_i]_{t-t_i+1} = s(t)$, $t \in \mathcal{T}_i$, $i \in \mathcal{M}$

Likelihood Function for System

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \phi, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{(\sqrt{2\pi}\sigma_i)^{T_i}} \exp\left\{ -\frac{(\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i)^T (\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i)}{2{\sigma_i}^2} \right\}$$

- Data vector, $\mathbf{s}_i$: $[\mathbf{s}_i]_{t-t_i+1} = s(t),\ t \in \mathcal{T}_i,\ i \in \mathcal{M}$
- Data matrix, S_i: $[\mathrm{S}_i]_{t-t_i+1,q} = \mathbf{s}(t-q),\ t \in \mathcal{T}_i, q \in \mathcal{Q}_i,\ i \in \mathcal{M}$

UNIVERSITY OF CAMBRIDGE

Likelihood Function for System

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\phi}, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{(\sqrt{2\pi}\sigma_i)^{T_i}} \exp\left\{ -\frac{\left(\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i\right)^T \left(\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i\right)}{2{\sigma_i}^2} \right\}$$

- Data vector, $\mathbf{s}_i$: $[\mathbf{s}_i]_{t-t_i+1} = s(t)$, $t \in \mathcal{T}_i$, $i \in \mathcal{M}$
- Data matrix, S_i: $[\mathrm{S}_i]_{t-t_i+1,q} = s(t-q)$, $t \in \mathcal{T}_i, q \in \mathcal{Q}_i$, $i \in \mathcal{M}$
- Source parameters, $\mathbf{b} = \{\mathbf{b}_i,\ i \in \mathcal{M}\}$: $[\mathbf{b}_i]_q = b_i(q)$, $q \in \mathcal{Q}_i$

Likelihood Function for System

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\phi}, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{\left(\sqrt{2\pi}\sigma_i\right)^{T_i}} \exp\left\{ -\frac{\left(\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i\right)^T \left(\mathbf{s}_i + \mathrm{S}_i \mathbf{b}_i\right)}{2{\sigma_i}^2} \right\}$$

- Data vector, $\mathbf{s}_i$: $[\mathbf{s}_i]_{t-t_i+1} = s(t),\ t \in \mathcal{T}_i,\ i \in \mathcal{M}$
- Data matrix, S_i: $[\mathrm{S}_i]_{t-t_i+1,q} = \mathbf{s}(t-q),\ t \in \mathcal{T}_i, q \in \mathcal{Q}_i,\ i \in \mathcal{M}$
- Source parameters, $\mathbf{b} = \{\mathbf{b}_i,\ i \in \mathcal{M}\}$: $[\mathbf{b}_i]_q = b_i(q),\ q \in \mathcal{Q}_i$
- Excitation Variances, $\boldsymbol{\sigma} = \{\sigma_i^2,\ i \in \mathcal{M}\}$

Likelihood Function for System

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \phi, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{\left(\sqrt{2\pi}\sigma_i\right)^{T_i}} \exp\left\{ -\frac{\left(\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i\right)^T\left(\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i\right)}{2{\sigma_i}^2} \right\}$$

- Data vector, $\mathbf{s}_i$: $[\mathbf{s}_i]_{t-t_i+1} = s(t)$, $t \in \mathcal{T}_i$, $i \in \mathcal{M}$
- Data matrix, S_i: $[\mathrm{S}_i]_{t-t_i+1,q} = \mathbf{s}(t-q)$, $t \in \mathcal{T}_i, q \in \mathcal{Q}_i$, $i \in \mathcal{M}$
- Source parameters, $\mathbf{b} = \{\mathbf{b}_i,\ i \in \mathcal{M}\}$: $[\mathbf{b}_i]_q = b_i(q)$, $q \in \mathcal{Q}_i$
- Excitation Variances, $\boldsymbol{\sigma} = \{\sigma_i^2,\ i \in \mathcal{M}\}$
- All source and channel parameters, $\boldsymbol{\theta} = \{\mathbf{a},\ \boldsymbol{\sigma}, \mathbf{b}\}$

Likelihood Function

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \phi, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{(\sqrt{2\pi}\sigma_i)^{T_i}} \exp\left\{ -\frac{(\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i)^T (\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i)}{2\sigma_i^2} \right\}$$

- Other parameters: $\phi = \{\boldsymbol{\tau}, \boldsymbol{\Xi}, \boldsymbol{\delta}, \boldsymbol{\nu}, \gamma\}$

Likelihood Function

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\phi}, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{(\sqrt{2\pi}\sigma_i)^{T_i}} \exp\left\{ -\frac{(\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i)^T (\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i)}{2\sigma_i^2} \right\}$$

- Other parameters: $\boldsymbol{\phi} = \{\boldsymbol{\tau}, \boldsymbol{\Xi}, \boldsymbol{\delta}, \boldsymbol{\nu}, \boldsymbol{\gamma}\}$
  - vector of changepoints: $\boldsymbol{\tau} = \{t_i, i \in \mathcal{M}\}$

Likelihood Function

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\phi}, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{(\sqrt{2\pi}\sigma_i)^{T_i}} \exp\left\{-\frac{\left(\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i\right)^T\left(\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i\right)}{2\sigma_i^2}\right\}$$

- Other parameters: $\boldsymbol{\phi} = \{\boldsymbol{\tau}, \boldsymbol{\Xi}, \boldsymbol{\delta}, \boldsymbol{\nu}, \gamma\}$
  - vector of changepoints: $\boldsymbol{\tau} = \{t_i, i \in \mathcal{M}\}$
  - vector of model orders: $\boldsymbol{\Xi} = \{Q_i, i \in \mathcal{M}\}$

Likelihood Function

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\phi}, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{(\sqrt{2\pi}\sigma_i)^{T_i}} \exp\left\{ -\frac{(\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i)^T (\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i)}{2\sigma_i^2} \right\}$$

- Other parameters: $\boldsymbol{\phi} = \{\boldsymbol{\tau}, \boldsymbol{\Xi}, \boldsymbol{\delta}, \boldsymbol{\nu}, \gamma\}$
  - vector of changepoints: $\boldsymbol{\tau} = \{t_i, i \in \mathcal{M}\}$
  - vector of model orders: $\boldsymbol{\Xi} = \{Q_i, i \in \mathcal{M}\}$
  - vectors of hyperparameters: $\boldsymbol{\delta} = \{\delta_i, i \in \mathcal{M}\}$

Likelihood Function

The likelihood function for the observed signal, $\mathbf{x}$, is:

$$p\left(\mathbf{x} \mid \boldsymbol{\theta}, \phi, \mathcal{I}\right) = \prod_{i=1}^{M} \frac{1}{(\sqrt{2\pi}\sigma_i)^{T_i}} \exp\left\{-\frac{(\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i)^T(\mathbf{s}_i + \mathrm{S}_i\mathbf{b}_i)}{2\sigma_i^2}\right\}$$

- Other parameters: $\phi = \{\boldsymbol{\tau}, \boldsymbol{\Xi}, \boldsymbol{\delta}, \boldsymbol{\nu}, \boldsymbol{\gamma}\}$
  - vector of changepoints: $\boldsymbol{\tau} = \{t_i, i \in \mathcal{M}\}$
  - vector of model orders: $\boldsymbol{\Xi} = \{Q_i, i \in \mathcal{M}\}$
  - vectors of hyperparameters: $\boldsymbol{\delta} = \{\delta_i, i \in \mathcal{M}\}$
  - vectors of hyper-hyperparameters: $\boldsymbol{\nu} = \{\nu_i, i \in \mathcal{M}\}$, and $\boldsymbol{\gamma} = \{\gamma_i, i \in \mathcal{M}\}$

Posterior Distribution for System

- Apply Bayes's rule to obtain the posterior pdf for the unknown parameters $\boldsymbol{\theta}$ (assuming ϕ is known):

$$p(\boldsymbol{\theta} \mid \mathbf{x}, \phi, \mathcal{I}) \propto p(\mathbf{x} \mid \boldsymbol{\theta}, \phi, \mathcal{I})\, p(\boldsymbol{\theta} \mid \phi, \mathcal{I})$$

Posterior Distribution for System

- Apply Bayes's rule to obtain the posterior pdf for the unknown parameters $\boldsymbol{\theta}$ (assuming $\boldsymbol{\phi}$ is known):

$$p(\boldsymbol{\theta} \mid \mathbf{x}, \boldsymbol{\phi}, \mathcal{I}) \propto p(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\phi}, \mathcal{I})\, p(\boldsymbol{\theta} \mid \boldsymbol{\phi}, \mathcal{I})$$

- Assuming $\{b_j, \sigma_j\}$ are independent between blocks, the assigned priors are:

$$\mathbf{b}_j \mid \sigma_j^2 \sim \mathcal{N}\left(\mathbf{0}_{Q_j}, \sigma_j^2\, \delta_j^2\, \mathrm{I}_{Q_j}\right), \; \delta_j \in \mathbb{R}^+ \text{ and } \sigma_j^2 \sim \mathcal{IG}\left(\frac{\nu_j}{2}, \frac{\gamma_j}{2}\right)$$

- Hence:

$$p(\boldsymbol{\theta} \mid \boldsymbol{\phi}, \mathcal{I}) = p(\mathbf{a} \mid \boldsymbol{\phi}, \mathcal{I})\, p(\mathbf{b} \mid \boldsymbol{\sigma}, \boldsymbol{\phi}, \mathcal{I})\, p(\boldsymbol{\sigma} \mid \boldsymbol{\phi}, \mathcal{I})$$

Posterior Distribution for Channel

- Only interested in estimating the channel, $\mathbf{a}$, so marginalise the *nuisance* parameters $\mathbf{b}$ and $\boldsymbol{\sigma}$ by integrating over $\boldsymbol{\theta}_{-\mathbf{a}}$:

$$p\left(\boldsymbol{\theta}_{-\mathbf{b}} \,|\, \mathbf{x}\right) = \int_{\mathbb{R}^{Q_1}} \cdots \int_{\mathbb{R}^{Q_M}} p\left(\boldsymbol{\theta}_{-\mathbf{b}}, \mathbf{b} \,|\, \mathbf{x}\right) \, d\mathbf{b}_M \, \ldots \, d\mathbf{b}_1$$

$$p\left(\mathbf{a} \,|\, \mathbf{x}\right) = \int_0^\infty \cdots \int_0^\infty p\left(\mathbf{a}, \boldsymbol{\sigma} \,|\, \mathbf{x}\right) \, d\sigma_M^2 \ldots d\sigma_1^2$$

Posterior Distribution for Channel

- Yields the posterior density for the channel parameters $\mathbf{a}$:

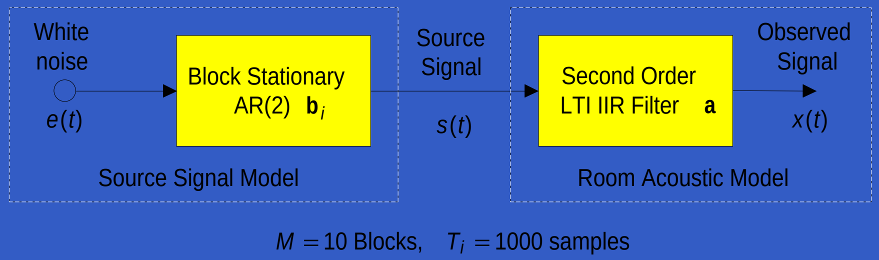
$$p\left(\mathbf{a} \,|\, \mathbf{x}, \boldsymbol{\phi}, \mathcal{I}\right) \propto p\left(\mathbf{a} \,|\, \boldsymbol{\phi}, \mathcal{I}\right) \times \prod_{i=1}^{M} \frac{\left\{\gamma_i + \mathbf{s}_i^T \mathbf{s}_i - \mathbf{s}_i^T \mathrm{S}_i \left(\mathrm{S}_i^T \mathrm{S}_i + \delta_i^{-2} \mathrm{I}_{Q_i}\right)^{-1} \mathrm{S}_i^T \mathbf{s}_i\right\}^{-R_i}}{\left|\mathrm{S}_i^T \mathrm{S}_i + \delta_i^{-2} \mathrm{I}_{Q_i}\right|^{\frac{1}{2}}}$$

  where $R_i = \frac{T_i + \nu_i + 1}{2}$, $i \in \mathcal{M}$

- Written in terms of $\mathbf{s}(t)$ to emphasise that it can be efficiently calculated by ‘inverse filtering’ the data, $\mathbf{x}(t)$

- MMAP estimate used *i.e.* $\arg_{\mathbf{a}} \max p\left(\mathbf{a} \,|\, \mathbf{x}, \boldsymbol{\phi}, P, \mathcal{I}\right)$

Principle Revisited

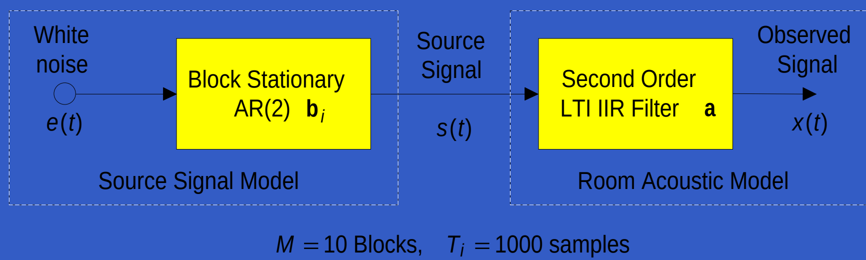


$P = 2,\ Q_i = 2,\ N = 10$ **and**

$$T_i = 1000,\ \forall i \in \mathcal{M}$$

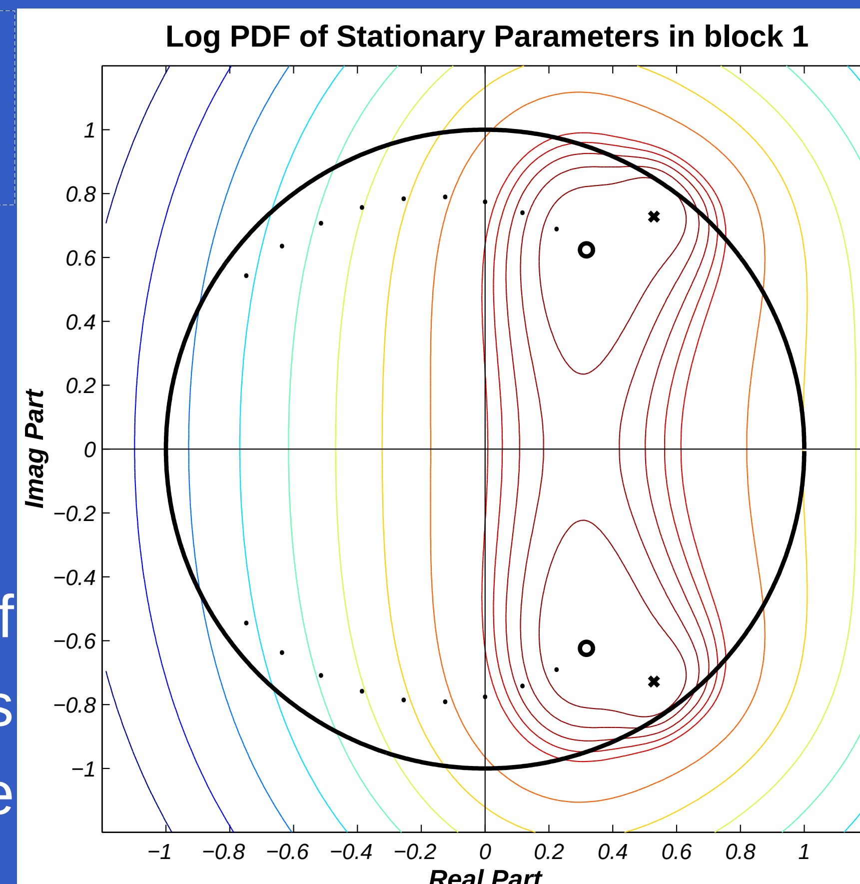
- Phase and magnitude of the pole locations for this BSAR(2) process change linearly with block number

Principle Revisited



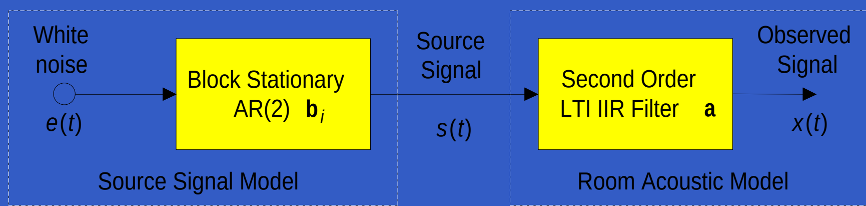
$$P = 2,\ Q_i = 2,\ N = 10 \textbf{ and}$$
$$T_i = 1000,\ \forall i \in \mathcal{M}$$

- Phase and magnitude of the pole locations for this BSAR(2) process change linearly with block number



$$\ln \hat{p}_1\left(\mathbf{r_a} \mid \mathbf{x}_1,\, \mathbf{x}_0,\, \phi,\, \mathcal{I}\right)$$

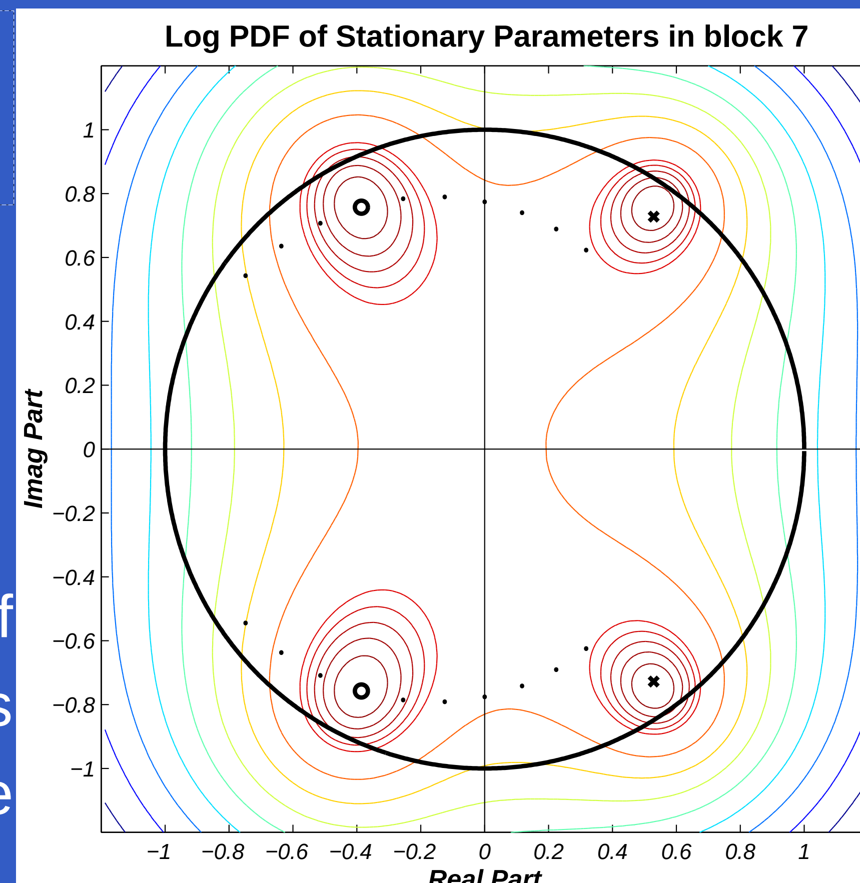
Principle Revisited



$M = 10$ Blocks, $T_i = 1000$ samples

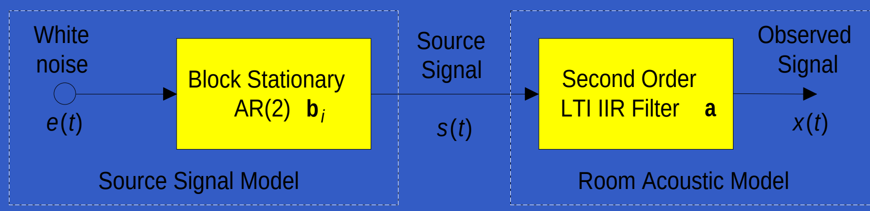
$$P = 2,\ Q_i = 2,\ N = 10 \text{ and}$$
$$T_i = 1000,\ \forall i \in \mathcal{M}$$

- Phase and magnitude of the pole locations for this BSAR(2) process change linearly with block number



$$\ln \hat{p}_7\left(\mathbf{r_a} \,\middle|\, \mathbf{x}_7, \mathbf{x}_6, \phi, \mathcal{I}\right)$$

Principle Revisited

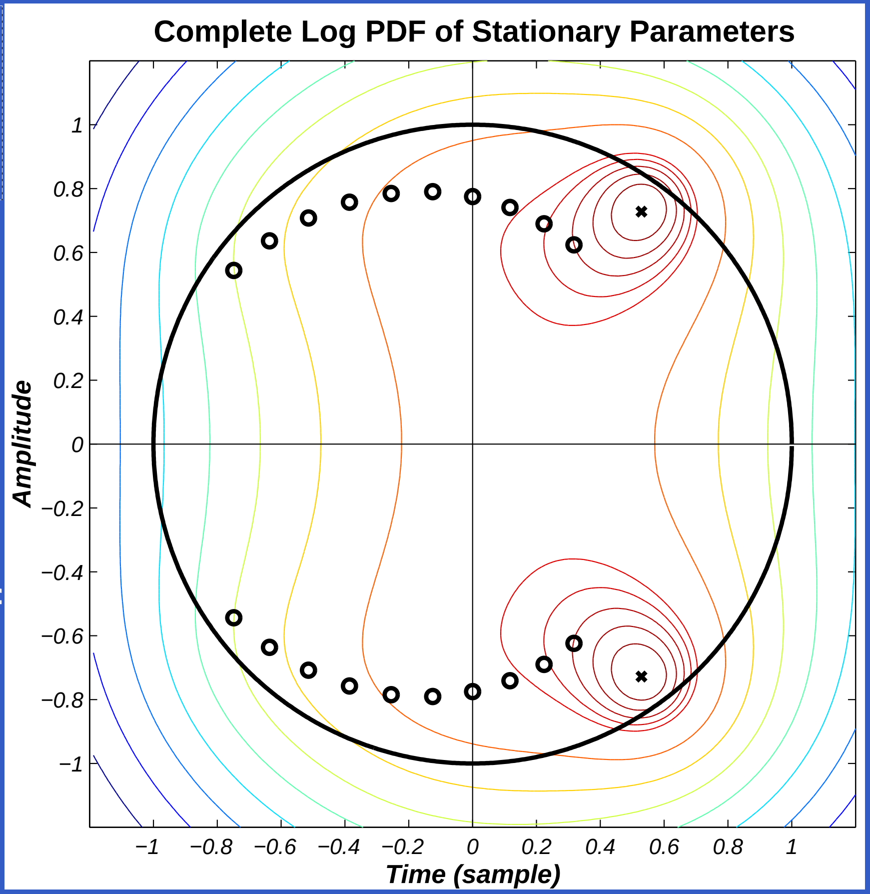


$M = 10$ Blocks, $T_i = 1000$ samples

$P = 2,\ Q_i = 2,\ N = 10$ **and**

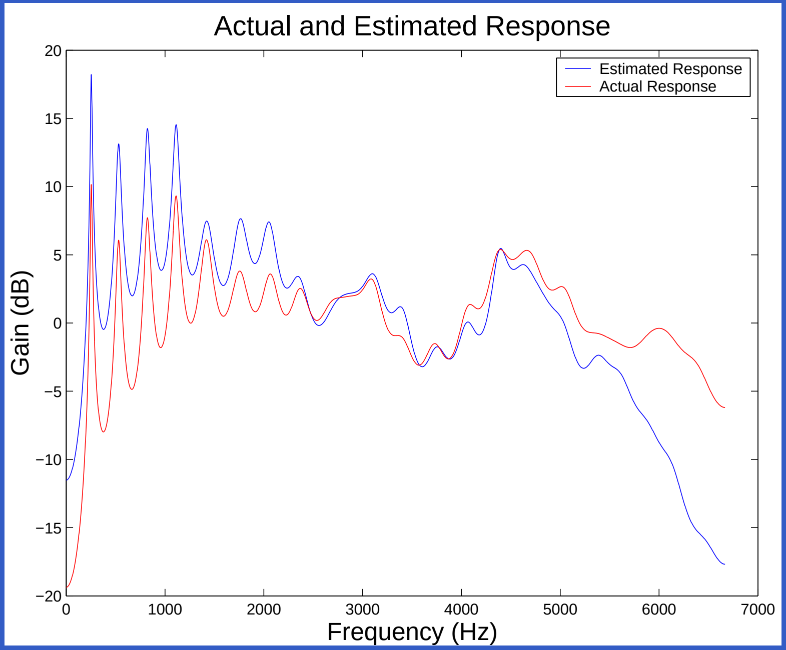
$T_i = 1000,\ \forall i \in \mathcal{M}$

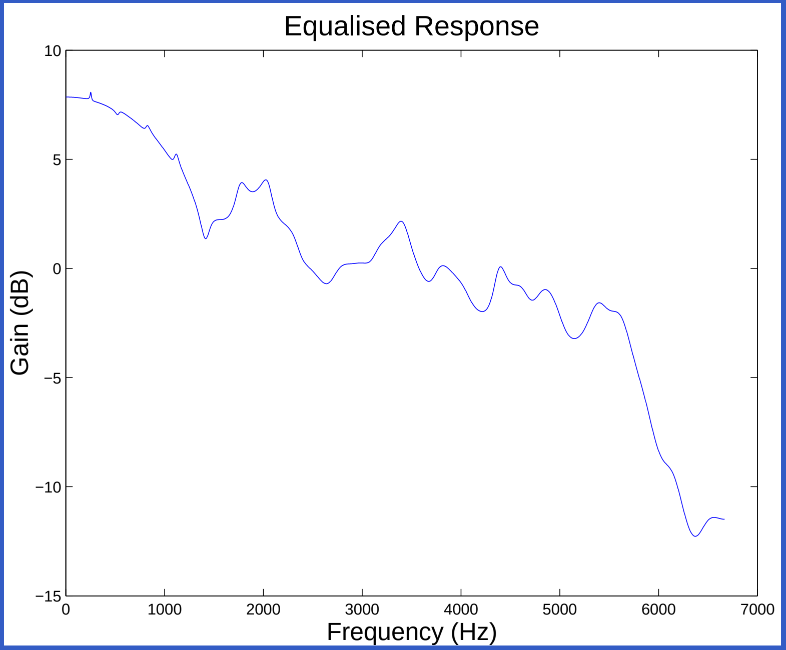
- Phase and magnitude of the pole locations for this BSAR(2) process change linearly with block number



$$\ln \hat{p}\left(\mathbf{r_a} \,\middle|\, \mathbf{x},\, \boldsymbol{\phi},\, \mathcal{I}\right)$$

A simple acoustic environment





- $\nu_i = \gamma_i = 0, \delta_i \sim 10^6, Q_i = 100, T_i = 1000$ and $M = 100$