Lecture 8:

Parameter Estimation in Subspace Problems

Singular Prediction Problems

Overview

- scalar case: sinusoids in noise
- multivariate case: blind MIMO channel estimation
- combined case: specular wireless channel models

SISO/Scalar Linear Prediction

- ∞^{th} order prediction error:

$$\widetilde{x}_{\infty,k} = x_k - \widehat{x}_{\infty,k} = x_k - \sum_{n=1}^{\infty} p_{\infty,n}\, x_{k-n} = P_\infty(q)\, x_k$$

$$q^{-1}x_k = x_{k-1},\;\; P_\infty(z) = \sum_{n=0}^{\infty} p_{\infty,n}\, z^{-1},\;\; p_{N,0} = 1 \text{ (monic)}$$

- ∞^{th} order prediction error variance: $\sigma^2_{\widetilde{x},\infty} = e^{\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln S_{xx}(f)\, df}$

- *singular* case: $\sigma^2_{\widetilde{x},\infty} = 0$ whenever spectrum *bandlimited* (parts are missing). Why perfect prediction? bandlimited $\Rightarrow$ can downsample $\Rightarrow$ can obtain some samples by linear interpolation from others, interpolation can be causal $\Rightarrow$ perfect prediction

Prototype Problem: Sinusoids

- $x_k = \sum_{n=1}^{N} A_n \, e^{j2\pi f_n k}$

$$\Rightarrow \underbrace{\prod_{n=1}^{N}(1 - e^{j2\pi f_n}\, q^{-1})}_{P_N(q)} \, x_k = 0 = \widetilde{x}_{N,k} = x_k + p_{N,1}\, x_{k-1} + \cdots + p_{N,N}\, x_{k-N}$$

- $\{A_n\}$ i.i.d. uniform phases $\Rightarrow$ x_k stationary, zero mean

- correlation sequence: $r_m = r_{xx}(m) = \mathrm{E}\, x_{k+m} x_k^H = \sum_{n=1}^{N} |A_n|^2 \, e^{j2\pi\, f_n\, m}$

  spectrum:

  $$S_{xx}(f) = \sum_{m=-\infty}^{\infty} r_{xx}(m)\, e^{-j2\pi\, f\, m} = \sum_{n=1}^{N} |A_n|^2 \, \delta(f - f_n)$$

Sinusoids: Signal Subspace Structure

- $$X_M(k) = \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-M+1} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \cdots & 1 \\ e^{-j2\pi f_1} & \cdots & e^{-j2\pi f_N} \\ \vdots & \cdots & \vdots \\ e^{-j2\pi f_1\,(M-1)} & \cdots & e^{-j2\pi f_N\,(M-1)} \end{bmatrix}}_{=\mathcal{V}} \underbrace{\begin{bmatrix} A_1\, e^{j2\pi f_1\, k} \\ \vdots \\ A_N\, e^{j2\pi f_N\, k} \end{bmatrix}}_{=S_k}$$

  $$= \mathcal{V}\, S_k$$

- one calls

$$\begin{aligned} Range\,\{\mathcal{V}\} &= \text{signal subspace} \\ (Range\,\{\mathcal{V}\})^{\perp} &= \text{noise subspace} \end{aligned}$$

Sinusoids: Covariance Subspace Structure

- covariance matrix

$$R_M = \mathrm{E} \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-M+1} \end{bmatrix} \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-M+1} \end{bmatrix}^H = \begin{bmatrix} r_0 & r_1 & \cdots & r_{M-1} \\ r_1^* & r_0 & \ddots & r_{M-2} \\ \vdots & \ddots & \ddots & \vdots \\ r_{M-1}^* & r_{M-2}^* & \cdots & r_0 \end{bmatrix}$$

$$= \mathrm{E} X_M(k)\, X_M^H(k) = \mathcal{V}\, R_{SS}\, \mathcal{V}^H\,, \quad R_{SS} = \mathrm{diag}\left\{|A_1|^2, \ldots, |A_N|^2\right\}$$

- $[x_k\; x_{k-1} \cdots x_{k-M+1}]$ = linear combination of $A_1, \cdots, A_N$
  $\Rightarrow$ rank$(R_M) = N$ for $M \geq N$

- Eigenvalues of R_M behave like $S_{xx}(f)$ for large M. However, singular sinusoid spectrum (support of measure zero) leads to singular R_M for finite M.

Sinusoids: Singular Prediction

- $\widetilde{x}_{M,k} = P_M(q)\, x_k = 0\,,\quad M \geq N$
- Hence, x_k is perfectly predictible from the previous N samples. $P_N(z) = \prod_{n=1}^{N}(1-e^{j2\pi f_n}\, z^{-1})$ and hence the f_n can be found by linear prediction: *Prony* method.
  Normal equations:
  $$P_N\, R_{XX} = [\sigma^2_{\widetilde{x},N}\; 0 \cdots 0]\,,\quad \sigma^2_{\widetilde{x},N} = 0$$
  where
  $$\begin{aligned} R_{XX} &= \mathrm{E}\, X_{N+1}(k)\, X^H_{N+1}(k) \\ P_N &= [p_{N,0}\; p_{N,1} \cdots p_{N,N}]\,,\quad p_{N,0} = 1 \end{aligned}$$
- $P_N\, R_{XX} = 0 \;\;\Rightarrow\;\; P_N\, \mathcal{V} = 0$ $\qquad R_{SS} > 0$

Sinusoids in Noise: Signal and Noise Subspaces

- additive white Gaussian noise v_k $\Rightarrow$ measure: $y_k = x_k + v_k$
- covariance structure

$$Y_M(k) = X_M(k) + V_M(k) = \mathcal{V}\, S_k + V_M(k) \quad \Rightarrow \quad R_{YY} = \mathcal{V}\, R_{SS}\, \mathcal{V}^H + \sigma_v^2 I$$

- Consider the eigendecomposition of R_{YY} $(\lambda_1 \geq \lambda_2 \geq \cdots)$:

$$R_{YY} = \sum_{i=1}^{N} \lambda_i V_i\, V_i^H + \sum_{i=N+1}^{M} \lambda_i V_i\, V_i^H = V_{\mathcal{S}} \Lambda_{\mathcal{S}} V_{\mathcal{S}}^H + V_{\mathcal{N}} \Lambda_{\mathcal{N}} V_{\mathcal{N}}^H$$

where $\Lambda_{\mathcal{N}} = \sigma_v^2 I_{M-N}$.

- Assuming $\mathcal{V}_{\mathcal{S}}$ and R_{SS} to have full rank, the sets of eigenvectors $V_{\mathcal{S}}$ and $V_{\mathcal{N}}$ are orthogonal: $V_{\mathcal{S}}^H V_{\mathcal{N}} = 0$, and $\lambda_i > \sigma_v^2\,,\ i = 1, \ldots, N$.

Sinusoids in Noise: Linear Prediction

- Equivalent descriptions of the signal and noise subspaces:

$$Range\,\{V_{\mathcal{S}}\} \;=\; Range\,\{\mathcal{V}\} \quad , \quad V_{\mathcal{N}}^{H}\mathcal{V} = 0$$

- Linear prediction in the noisy case: minimize variance subject to norm constraint: *Pisarenko* method
  with $M = N + 1$: noise subspace dimension $= 1$

$$\begin{aligned}
\min_{\|P\|=1} P\,R_{YY}P^{H} &= \min_{\|P\|=1} \left\{P\,R_{XX}P^{H} + P\,R_{VV}P^{H}\right\} \\
&= \min_{\|P\|=1} \left\{P\,R_{XX}P^{H} + \sigma_v^2\,\|P\|^2\right\} \\
&= \sigma_v^2 + \min_{\|P\|=1} P\,R_{XX}P^{H}
\end{aligned}$$

$$\Rightarrow\; P\,R_{XX} = [0\cdots 0]\,,\; P^{H} = V_{min}(R_{YY}) = V_{min}(R_{XX})$$

Sinusoids in Noise: Signal Subspace Fitting

- two equivalent signal subspace descriptions: $\mathcal{V}$ and $V_{\mathcal{S}}$
- with an estimated covariance matrix, $\widehat{V}_{\mathcal{S}}$ is approximate, so consider

$$\min_{\mathbf{f},T} \|\mathcal{V}(\mathbf{f}) - \widehat{V}_{\mathcal{S}} T\|_F$$

where $\mathbf{f} = [f_1 \cdots f_N]$, $\|A\|_F^2 = \mathrm{tr}\,\left\{A^H A\right\}$.

- separable problem $\Rightarrow T = \widehat{V}_{\mathcal{S}}^H \mathcal{V}$, $\mathcal{V} - \widehat{V}_{\mathcal{S}} T = \mathrm{P}^{\perp}_{\widehat{V}_{\mathcal{S}}} \mathcal{V}$ and hence

$$\begin{aligned}
\|\mathrm{P}^{\perp}_{\widehat{V}_{\mathcal{S}}} \mathcal{V}\|_F^2 &= \mathrm{tr}\,\mathcal{V}^H \mathrm{P}^{\perp}_{\widehat{V}_{\mathcal{S}}} \mathcal{V} = \mathrm{tr}\,\mathcal{V}^H \mathrm{P}_{\widehat{V}_{\mathcal{N}}} \mathcal{V} = \|\widehat{V}_{\mathcal{N}}^H \mathcal{V}\|_F^2 \\
&= \sum_{i=N+1}^{M} \|\widehat{V}_i^H \mathcal{V}\|^2 = \sum_{j=1}^{N} \sum_{i=N+1}^{M} |\widehat{V}_i(f_j)|^2
\end{aligned}$$

$$\mathrm{P}_V = V(\underbrace{V^H V}_{=I})^{-1} V^H = V\,V^H \,,\ \mathrm{P}^{\perp}_V = I - \mathrm{P}_V \,,\ \mathrm{P}^H = \mathrm{P},\, \mathrm{PP} = \mathrm{P}$$

Sinusoids in Noise: Signal Subspace Fitting (2)

- $\widehat{V}_i(f)$ = Fourier Transform of $\widehat{V}_i$ components
- exact solution: joint optimization

$$\min_{f_1,\ldots,f_N} \sum_{j=1}^{N} \sum_{i=N+1}^{M} |\widehat{V}_i(f_j)|^2$$

- approximate solution: plot as a function of f and find N largest peaks of

$$\frac{1}{\displaystyle\sum_{i=N+1}^{M} |\widehat{V}_i(f)|^2}$$

MUSIC! (MUltiple SIgnal Classification algorithm)

Sinusoids in Noise: Noise Subspace Parameterization

- $P(q)\, e^{j2\pi f_i k} = 0, \;\Rightarrow\;\; \mathcal{G}^H(P)\, \mathcal{V} = 0$ where $\mathcal{G}(P) :\; M \times (M-N)$

$$\mathcal{G}^H(P) = \begin{bmatrix} p_0 & p_1 & \cdots & p_N & 0 & \cdots & 0 \\ 0 & p_0 & p_1 & \cdots & p_N & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \vdots \\ 0 & \cdots & 0 & p_0 & p_1 & \cdots & p_N \end{bmatrix} \qquad \text{Toeplitz}$$

hence $Range\ \{\mathcal{G}(P)\} = Range\ \{\mathcal{V}_{\mathcal{N}}\} =$ noise subspace

- noise subspace fitting:

$$\min_{P,T} \|\mathcal{G}(P) - \widehat{V}_{\mathcal{N}} T\|_F$$

Sinusoids in Noise: Noise Subspace Fitting

- separable problem $\Rightarrow T = \widehat{V}_{\mathcal{N}}^H \mathcal{G}\,,\ \ \mathcal{G} - \widehat{V}_{\mathcal{N}} T = P^{\perp}_{\widehat{V}_{\mathcal{N}}} \mathcal{G}$ and hence

$$\begin{aligned} \|P^{\perp}_{\widehat{V}_{\mathcal{N}}} \mathcal{G}\|_F^2 &= \operatorname{tr} \mathcal{G}^H P^{\perp}_{\widehat{V}_{\mathcal{N}}} \mathcal{G} = \operatorname{tr} \mathcal{G}^H P_{\widehat{V}_{\mathcal{S}}} \mathcal{G} = \|\widehat{V}_{\mathcal{S}}^H \mathcal{G}\|_F^2 \\ &= \sum_{i=1}^{2M} \|\mathcal{G}^H \widehat{V}_i\|^2 \end{aligned}$$

Let $\mathcal{G}^H \widehat{V}_i = \widehat{\mathcal{W}_i}\, P^H$ where $\widehat{\mathcal{W}_i} = \mathcal{W}(\widehat{V}_i)$ is Hankel, then we get

$$\min_P P \left(\sum_{i=1}^{N} \widehat{\mathcal{W}_i}^H \widehat{\mathcal{W}_i}) \right) P^H$$

subject to $P_0 = 1$ or $\|P\| = 1$.

Sinusoids in Noise: Maximum Likelihood Estimation

- additive noise v_k white and Gaussian $\rightarrow$ likelihood criterion

$$\min_{\mathbf{f},S} \|Y - \mathcal{V}(\mathbf{f})\, S\|^2$$

- separable $\Rightarrow \quad S = (\mathcal{V}^H \mathcal{V})^{-1} \mathcal{V}^H Y$

  $\Rightarrow \; \|Y - \mathcal{V}\, S\|^2 = Y^H \mathrm{P}^{\perp}_{\mathcal{V}} Y = Y^H \mathrm{P}_{\mathcal{G}(P)} Y = P \mathcal{Y}^H (\mathcal{G}^H(P)\, \mathcal{G}(P))^{-1} \mathcal{Y} P^H$

  where $\mathcal{G}^H(P)\, Y = \mathcal{Y}(Y)\, P^H$ (commutativity of convolution, $\mathcal{Y}$ Hankel)

- IQML (Iterative Quadratic Maximum Likelihood), iteration n:

$$\min_{P^{(n)}} P^{(n)} \left(\mathcal{Y}^H (\mathcal{G}^H(P^{(n-1)})\, \mathcal{G}(P^{(n-1)}))^{-1} \mathcal{Y}\right) P^{(n)\,H}$$

  subject to $P_0 = 1$ or $\|P\| = 1$

- initialization: Pisarenko: $\min_{P^{(0)}} P^{(0)} \left(\mathcal{Y}^H \mathcal{Y}\right) P^{(0)\,H}$

Sinusoids in Noise: Maximum Likelihood Estimation (2)

- IQML leads to biased estimates
- Denoised IQML (DIQML):
  since $\text{tr}\{P_{\mathcal{G}}\}$ = dimension noise subspace = constant

$$\begin{aligned}
\arg\min_P Y^H \mathrm{P}_{\mathcal{G}(P)} Y &= \arg\min_P \text{tr}\left\{\mathrm{P}_{\mathcal{G}(P)}\, YY^H\right\} \\
&= \arg\min_P \text{tr}\left\{\mathrm{P}_{\mathcal{G}(P)}\, \left(YY^H - \sigma_v^2\, I\right)\right\}
\end{aligned}$$

hence we get $\min_{P^{(n)}} P^{(n)} \left(B^{(n-1)} - \sigma_v^2\, C^{(n-1)}\right) P^{(n)\,H}$ where

$$\begin{aligned}
B^{(n-1)} &= \mathcal{Y}^H (\mathcal{G}^H(P^{(n-1)})\, \mathcal{G}(P^{(n-1)}))^{-1} \mathcal{Y} \\
C^{(n-1)} &= \sum_{k=1}^{M-N} \left[(\mathcal{G}^H(P^{(n-1)})\, \mathcal{G}(P^{(n-1)}))^{-1}\right]_{k:k+N,k:k+N}
\end{aligned}$$

Sinusoids in Noise: Maximum Likelihood Estimation (3)

- Pseudo Quadratic ML (PQML): actual ML gradient

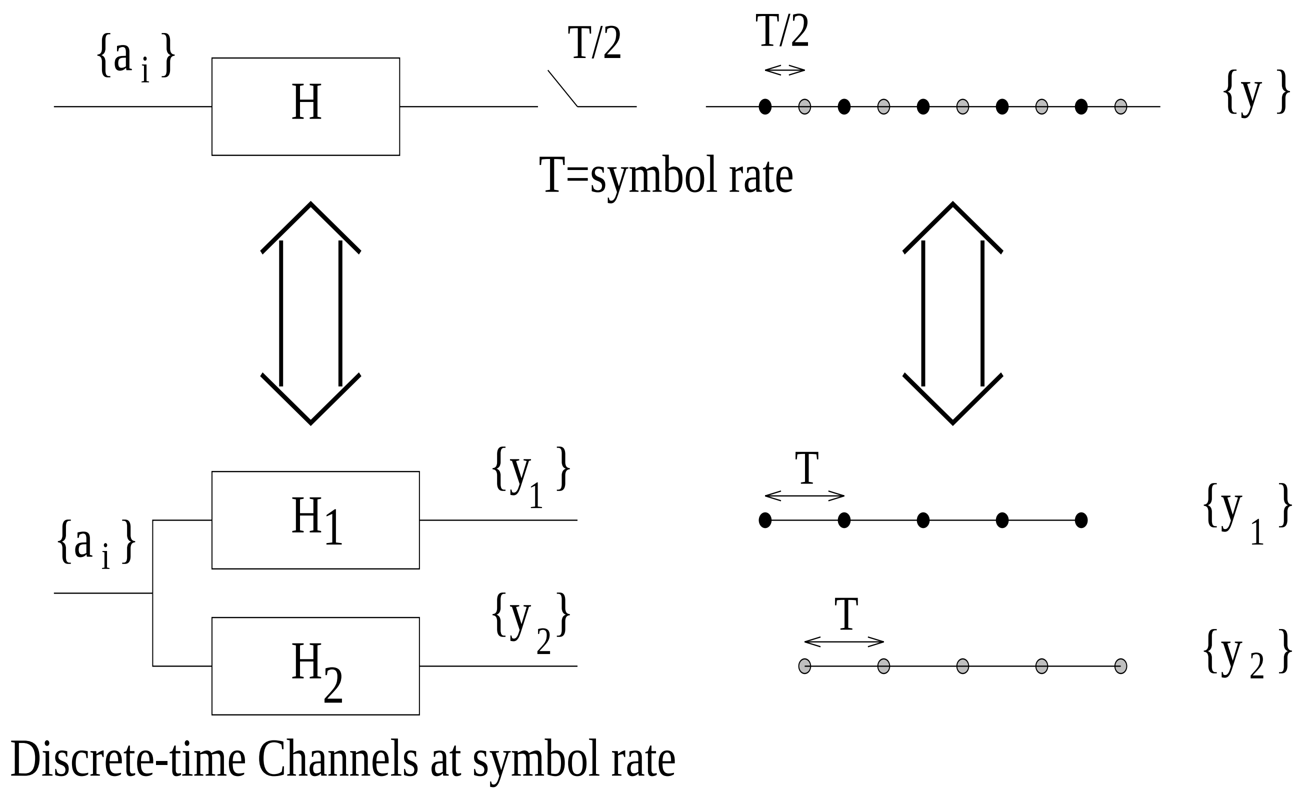
$$\nabla_P \, Y^H \mathrm{P}_{\mathcal{G}(P)} Y = Q\, P \;, \quad Q = B - D \geq 0$$

relation with DIQML: $\mathrm{E}_V \, D = \sigma_v^2 \, C$

- (D)IQML: asymptotically globally convergent
PQML: with a consistent initialization (e.g. DIQML), asymptotically only one iteration is required to get an estimate with ML performance

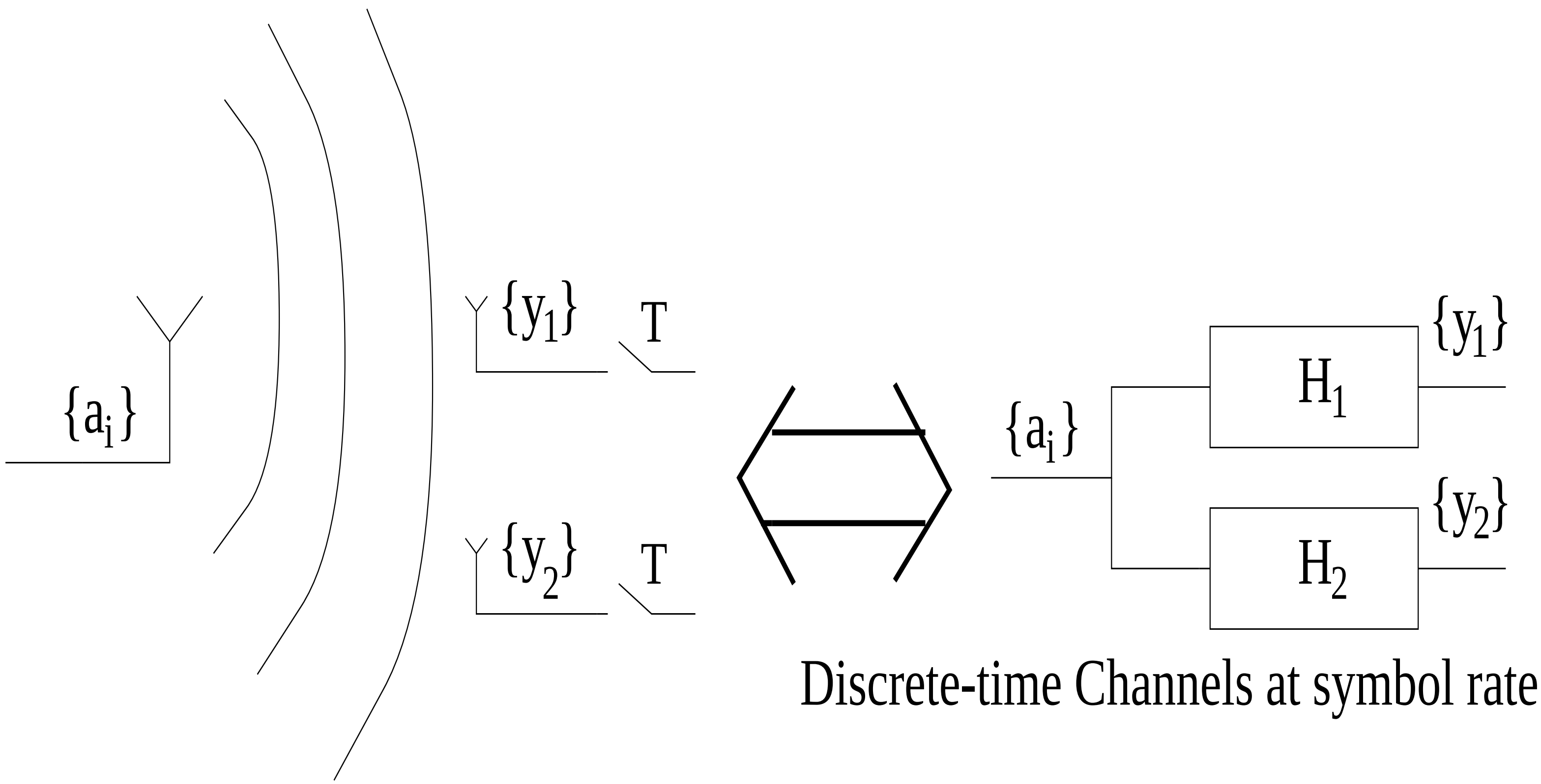
- So far: Deterministic ML (DML).
Gaussian ML (GML): postulate Gaussian distribution for S

SIMO/Multichannel Model (1)



- SIMO: Single Input Multiple Outputs
- oversampling received signal at twice the transmission rate

SIMO/Multichannel Model (2)



- single transmit (TX) antenna, multiple receive (RX) antennas

SIMO Singularity

- FIR channel input/output:

$$\underbrace{\mathbf{y}_k}_{p\times 1} = \underbrace{\mathbf{H}(q)}_{p\times 1}\,\underbrace{a_k}_{1\times 1}\,,\quad \mathbf{H}(z) = \sum_{n=0}^{N-1} \mathbf{h}_n\, z^{-n} \qquad\qquad \boldsymbol{H}(f) = \mathbf{H}(e^{j2\pi\, f})$$

- matricial psdf:

$$\underbrace{S_{\mathbf{yy}}(f)}_{p\times p} = \boldsymbol{H}(f)\, S_{aa}(f)\, \boldsymbol{H}^H(f) = \sigma_a^2\, \boldsymbol{H}(f)\, \boldsymbol{H}^H(f) \qquad\qquad \text{rank 1!}$$

  singular spectrum for $p \geq 2$ subchannels even though channel input a_k is white noise, not sinusoids

SIMO Singularity (2)

- MIMO prediction: $\underbrace{\widetilde{\mathbf{y}}_k}_{p\times 1} = \underbrace{\mathbf{P}_\infty(q)}_{p\times p} \underbrace{\mathbf{y}_k}_{p\times 1}$
- prediction error covariance $\underbrace{\sigma^2_{\widetilde{\mathbf{y}},\infty}}_{p\times p} = \mathrm{E}\,\widetilde{\mathbf{y}}_{\infty,k}\widetilde{\mathbf{y}}^H_{\infty,k}$

$$\det(\sigma^2_{\widetilde{\mathbf{y}},\infty}) = e^{\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln\det(S_{\mathbf{yy}}(f))\,df} = 0 \;\Rightarrow\; \sigma^2_{\widetilde{\mathbf{y}},\infty} \text{ singular}$$

does not mean no randomness, but number of random sources $1 < p =$ vector dimension

Deterministic Blind Identifiability

- blind: determine H and A from Y only

$$\begin{bmatrix} \mathbf{y}_k \\ \mathbf{y}_{k-1} \\ \vdots \\ \mathbf{y}_{k-L+1} \end{bmatrix} = \begin{bmatrix} \mathbf{h}_0 & \cdots & \mathbf{h}_{N-1} & 0_{p\times 1} & \cdots & 0_{p\times 1} \\ 0_{p\times 1} & \mathbf{h}_0 & \cdots & \mathbf{h}_{N-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0_{p\times 1} \\ 0_{p\times 1} & \cdots & 0_{p\times 1} & \mathbf{h}_0 & \cdots & \mathbf{h}_{N-1} \end{bmatrix} \begin{bmatrix} a_k \\ a_{k-1} \\ \vdots \\ \vdots \\ a_{k-L-N+2} \end{bmatrix}$$

or $\underbrace{Y_L(k)}_{pL\times 1} = \underbrace{\mathcal{T}_L\left(\mathbf{H}_N\right)}_{pL\times(N+L-1)} \underbrace{A_{L+N-1}(k)}_{(L+N-1)\times 1}$

- $p = 1$ single channel: $L+2N-1$ unknowns from L equations

Deterministic Blind Identifiability (2)

- $p > 1$ multichannel: $L+(p+1)N-1$ unknowns from pL equations $\Rightarrow$ gets overdetermined for L big enough
- can easily eliminate A, e.g. for $p = 2$ subchannels:

$$
\begin{aligned}
[-H_2(q) \;\; H_1(q)] \begin{bmatrix} y_{1,k} \\ y_{2,k} \end{bmatrix} &= \underbrace{[-H_2(q) \;\; H_1(q)]}_{=\mathbf{H}^{\perp\,\dagger}(q)} \underbrace{\begin{bmatrix} H_1(q) \\ H_2(q) \end{bmatrix}}_{=\mathbf{H}(q)} a_k \\
&= (H_1(q)H_2(q) - H_2(q)H_1(q))\; a_k \equiv 0
\end{aligned}
$$

where $\mathbf{H}^\dagger(z) = \mathbf{H}^H(1/z^*)$ (matched filter)

Deterministic Blind Identification

- Subchannel Response Matching (SRM), Cross-Relation (CR) method

$$\underbrace{\mathbf{u}_k}_{(p-1)\times 1} = \underbrace{\widehat{\mathbf{H}}^{\perp\dagger}(q)}_{(p-1)\times p} \underbrace{\mathbf{y}_k}_{p\times 1}$$

$$\mathrm{tr}\,\{R_{\mathbf{uu}}\} = \sigma_a^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} \|\widehat{\boldsymbol{H}}^{\perp H}(f)\,\boldsymbol{H}(f)\|^2 df + \sigma_v^2 \|\widehat{\mathbf{H}}^{\perp}\|^2$$

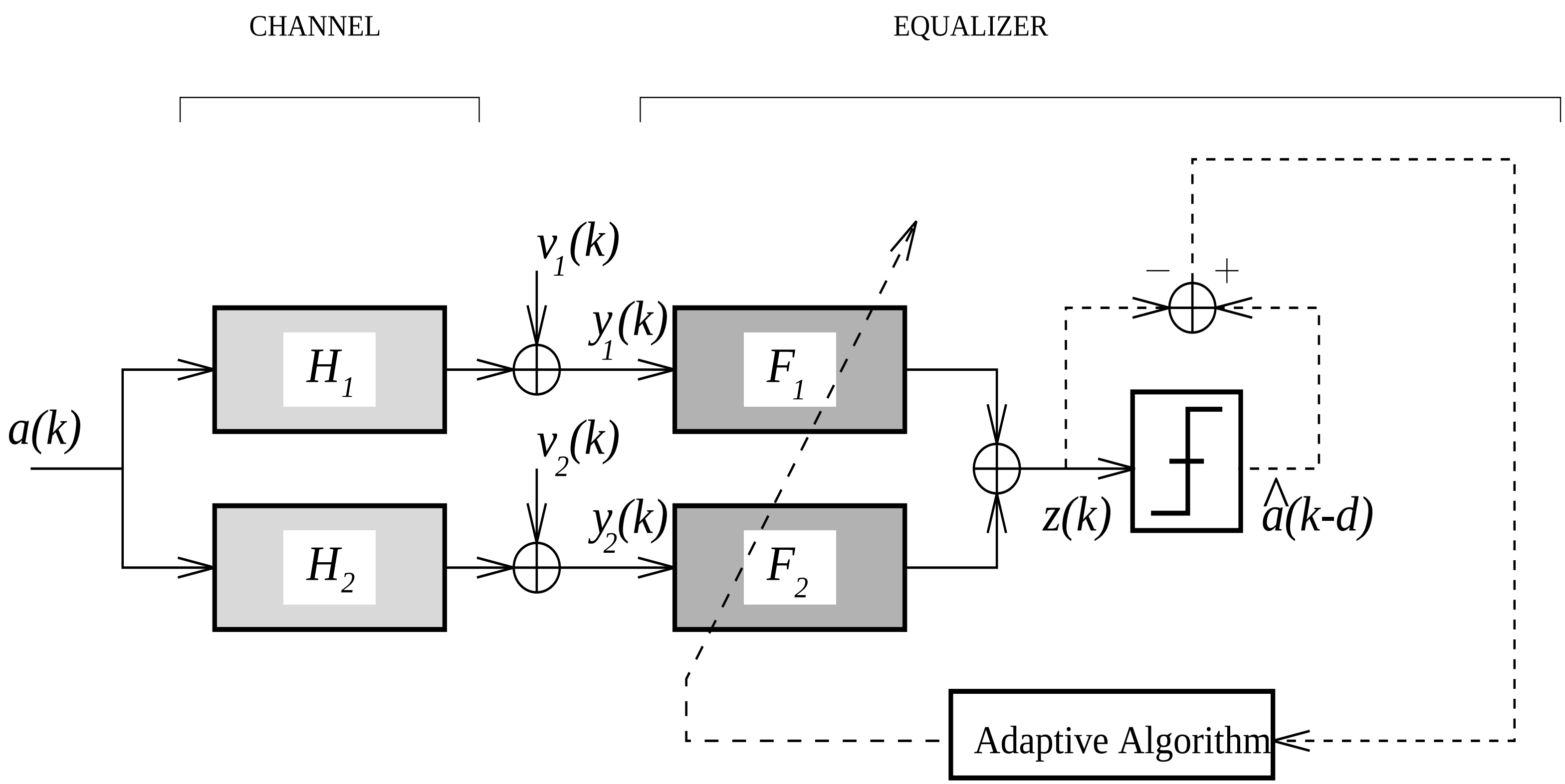
$$\Rightarrow \min_{\|\widehat{\mathbf{H}}^{\perp}\|=1} \sum_k \|\widehat{\mathbf{H}}^{\perp\dagger}(q)\,\mathbf{y}_k\|^2$$

  equivalent of Pisarenko

- for Signal or Noise Subspace fitting or DML:

$$\mathcal{V} \;\rightarrow\; \mathcal{T}(\mathbf{H}) \;,\quad \mathcal{G} \;\rightarrow\; \mathcal{T}(\mathbf{H}^{\perp})$$

SIMO Channel Equalization



- SISO equalization: $F_1(z) = \dfrac{1}{H_1(z)}$: ∞ causal and anti-causal portions

SIMO Channel Equalization (2)

- ZF condition in z domain: $n \in \{0, 1, \ldots, N{+}L{-}2\}$

$$\sum_{i=1}^{p} F_i(z)\, H_i(z) = [F_1(z) \cdots F_p(z)] \begin{bmatrix} H_1(z) \\ \vdots \\ H_p(z) \end{bmatrix} = \mathbf{F}(z)\, \mathbf{H}(z) = z^{-n}$$

  Bezout identity: FIR MISO equalizers exist for FIR SIMO channels

- condition: $\mathbf{H}(z)$ no zeros (otherwise SISO equalization problem for common factor between the $H_i(z)$)

SIMO Channel Equalization (3)

- in time domain:

$$
\left[\underbrace{\mathbf{f}_0}_{1\times p}\ \mathbf{f}_1 \cdots \mathbf{f}_{L-1}\right]
\begin{bmatrix}
\mathbf{h}_0 & \cdots & \mathbf{h}_{N-1} & 0_{p\times 1} & \cdots & 0_{p\times 1} \\
0_{p\times 1} & \mathbf{h}_0 & \cdots & \mathbf{h}_{N-1} & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & 0_{p\times 1} \\
0_{p\times 1} & \cdots & 0_{p\times 1} & \mathbf{h}_0 & \cdots & \mathbf{h}_{N-1}
\end{bmatrix}
$$

$$
= \ [0 \cdots 0\ 1\ 0 \cdots 0]
$$

$L+N-1$ eqs. to be satisfied with pL parameters $\Rightarrow L \geq \dfrac{N-1}{p-1}$

- $\mathcal{T}_L\left(\mathbf{H}_N\right)$: Sylvester matrix: full column rank if $\mathbf{H}(z)$ has no zeros

Linear Prediction in SIMO Channel

- linear prediction of order $L \geq \frac{N-1}{p-1} \Rightarrow \widehat{\mathbf{y}}_{L,k}$ can form any linear combination of $\{a_{k-1}, a_{k-2}, \ldots, a_{k-L-N+1}\}$, hence only $\mathbf{h}_0\, a_k$ left in $\mathbf{P}_L(q)\,\mathbf{y}_k = \widetilde{\mathbf{y}}_{L,k} = \mathbf{h}_0\, a_k$
- SIMO (FIR) channel $\Rightarrow$ singular prediction problem (at finite prediction order): $\sigma^2_{\widetilde{\mathbf{y}},L} = \sigma^2_a\, \mathbf{h}_0 \mathbf{h}_0^H$
- FIR channel $\Rightarrow$ FIR predictor:

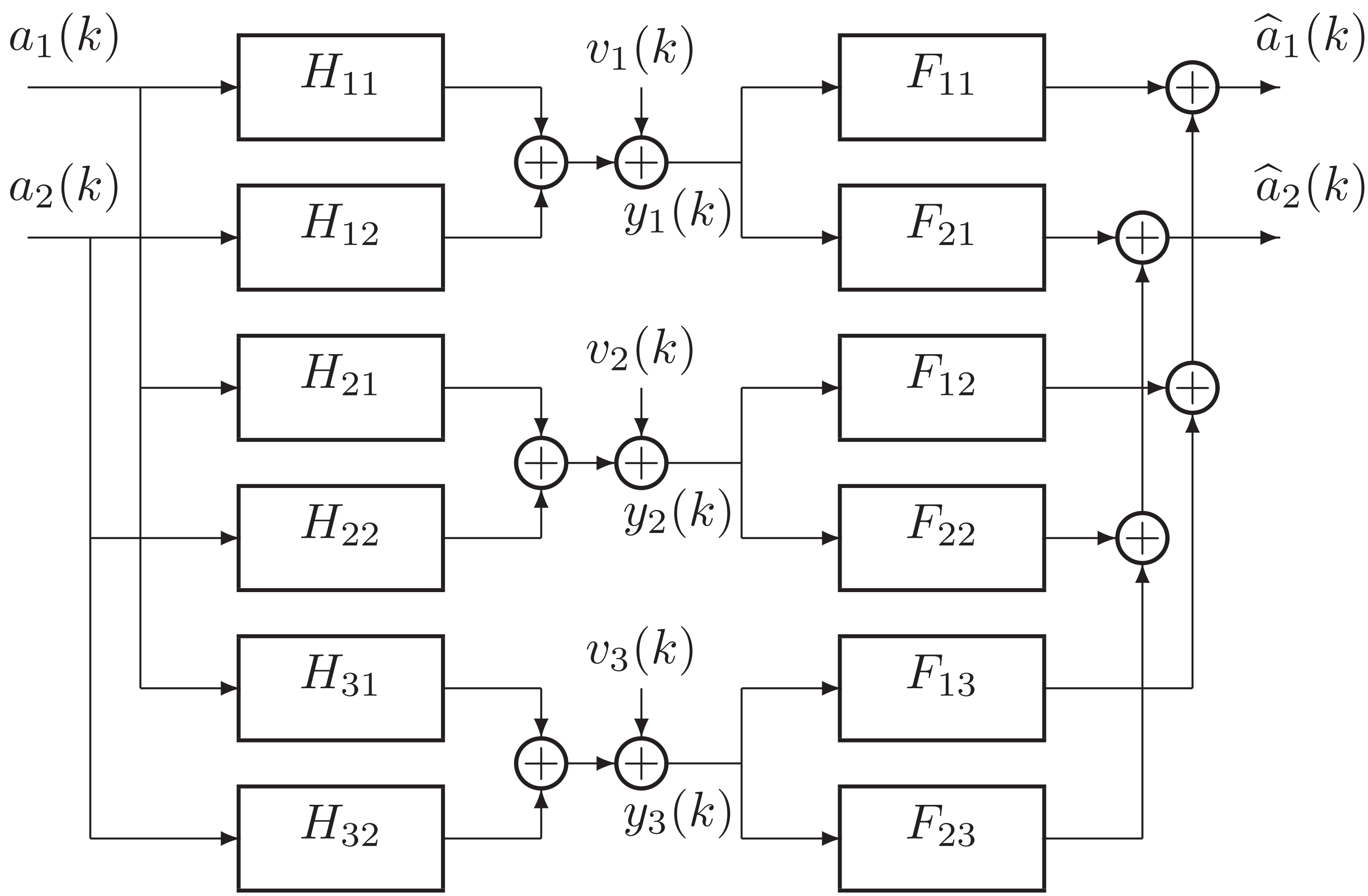
$$\mathbf{y}_k = \mathbf{H}_N(q)\, a_k \quad \Leftrightarrow \quad \mathbf{P}_L(q)\,\mathbf{y}_k = \mathbf{h}_0\, a_k$$

singular MA process $\Leftrightarrow$ singular AR process

Linear Prediction in SIMO Channel (2)

- let $\mathbf{h}_0^\perp$: $\left[\mathbf{h}_0\;\mathbf{h}_0^\perp\right]^H \left[\mathbf{h}_0\;\mathbf{h}_0^\perp\right] = \begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}$
- predictor has 2 components:
  - $\mathbf{h}_0^H \mathbf{P}(q)\,\mathbf{y}_k = \|\mathbf{h}_0\|^2\, a_k$: 0-delay equalizer
  - $\mathbf{h}_0^{\perp H}\,\mathbf{P}(q)\,\mathbf{y}_k = 0$: $\mathbf{h}_0^{\perp H}\,\mathbf{P}(z)$ characterizes the noise subspace whereas $\mathbf{H}(z)$ characterizes the signal subspace, can find $\mathbf{H}(z)$ from $\mathbf{h}_0^{\perp H}\,\mathbf{P}(z)\,\mathbf{H}(z) = 0$,
    we have $\mathbf{h}_0^{\perp H}\,\mathbf{P}(z) = \mathbf{H}^{\perp\dagger}(z)$

MIMO Case

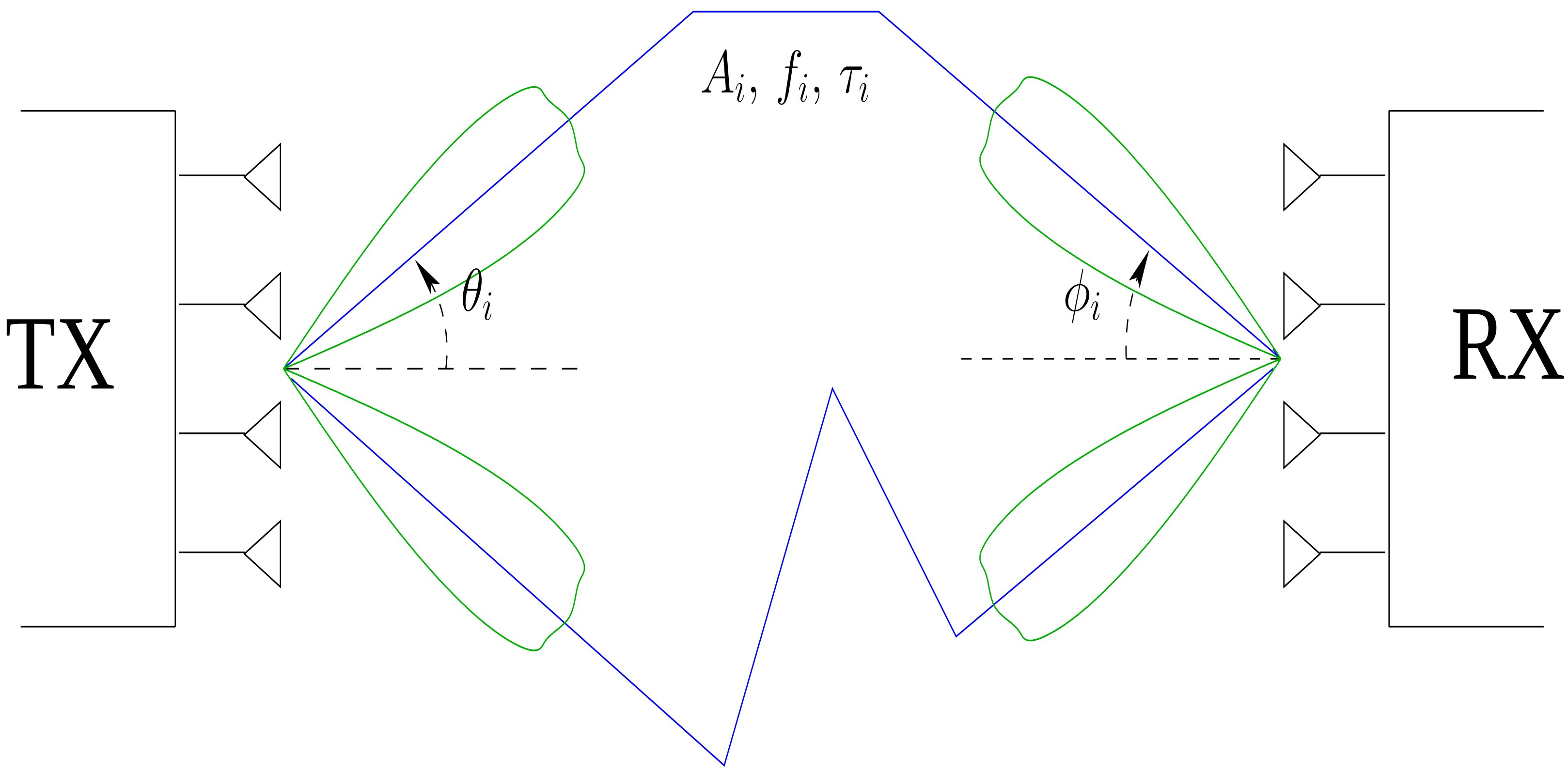

$a_1(k)$
$a_2(k)$
H_{11}
H_{12}
H_{21}
H_{22}
H_{31}
H_{32}
$v_1(k)$
$v_2(k)$
$v_3(k)$
$y_1(k)$
$y_2(k)$
$y_3(k)$
F_{11}
F_{21}
F_{12}
F_{22}
F_{13}
F_{23}
$\widehat{a}_1(k)$
$\widehat{a}_2(k)$

- MIMO input/output: $\underbrace{\mathbf{y}_k}_{p\times 1} = \underbrace{\mathbf{H}(q)}_{p\times q}\,\underbrace{\mathbf{a}_k}_{q\times 1}$
- MIMO channel MIMO prediction: $\mathbf{P}_L\,\mathbf{y}_k = \mathbf{h}_0\,\mathbf{a}_k$ for $L \geq \dfrac{\sum_{i=1}^{q} N_i - q}{p-q}$
- MIMO Bezout conditions:
  $\mathbf{H}(z)$ irreducible (full column rank $\forall\, z$) and column reduced

Multivariate Spectral Factorization

- $S_{\mathbf{yy}}(z) = \sigma_a^2\,\mathbf{H}(z)\,\mathbf{H}^\dagger(z)$ $\qquad\qquad \mathbf{H}^\dagger(z) = \mathbf{H}^H(1/z^*)$
- spectral factor $S_{\mathbf{yy}}^{1/2}(z)$ minimum phase and $p \times q$ since $S_{\mathbf{yy}}(z)$ of generic rank q
- then $\sigma_a\,\mathbf{H}(z) = S_{\mathbf{yy}}^{1/2}(z)\,Q$ for some unitary Q: $QQ^H = I_q$
  if $\mathbf{H}(z)$ not irreducible but only minimum-phase: |zeros| <1
- when $p > q$, it is easy to be minimum-phase (not for $p = q = 1$)

MIMO Transmission



- multiple transmit and receive antennas

MIMO Channel Model

- time-varying channel: $\mathbf{h}(t, \tau)$

$$\mathbf{h}(t, kT) = \sum_{i=1}^{I} A_i(t)\, e^{j2\pi\, f_i\, t}\, \mathbf{a}_R(\phi_i)\, \mathbf{a}_T^T(\theta_i)\, p(kT - \tau_i)$$

  $\mathbf{h}$ rank 1 in 3 dimensions; pathwise contributions:

- A_i: complex attenuation
- f_i: Doppler shift
- θ_i: angle of departure
- ϕ_i: angle of arrival
- τ_i: path delay
- $\mathbf{a}(.)$: antenna array response, $p(.)$ pulse shape (TX filter)

MIMO Channel Prediction

- $\underbrace{\underline{\mathbf{h}}(t)}_{M\times 1} = \mathrm{vec}\{\mathbf{h}(t,kT)\} = \sum_{i=1}^{I} \underline{\mathbf{h}}_i\, A_i(t)\, e^{j2\pi\, f_i\, t}$

  size M = # TX antenna x # RX antennas x delay spread
- $f_i \in (-f_D, f_D) \Rightarrow$ (fast fading) variation bandlimited $\Rightarrow$ perfectly predictible!?
- $S_{\underline{\mathbf{hh}}}(f)$ can be doubly singular:
  1. if $A_i(t) \equiv A_i$ and I finite: spectral support singularity: sinusoids!
  2. if $I < M$: matrix singularity, limited source of randomness (limited diversity)