EM-based channel estimation in multicarrier systems

Rodrigo Carvajal

Departamento de Ingeniería Electrónica Universidad Técnica Federico Santa María.

Trabajo realizado en conjunto con Juan C. Agüero, Boris I. Godoy, Graham C. Goodwin y Juan I. Yuz.

3 de enero, 2013

Altamente utilizados en sistemas de comunicaciones: desde ADSL hasta LTE-A.

R. Carvajal **EM-based channel estimation in multicarrier systems** 2 / 45

Altamente utilizados en sistemas de comunicaciones: desde ADSL hasta LTE-A.

Idea básica: transmitir los datos en forma paralela (subportadoras).

Se basan en la aplicación de una Transformación Ortogonal, como DFT, DCT, DST, wavelet, etc.

Altamente utilizados en sistemas de comunicaciones: desde ADSL hasta LTE-A.

Idea básica: transmitir los datos en forma paralela (subportadoras).

Se basan en la aplicación de una Transformación Ortogonal, como DFT, DCT, DST, wavelet, etc.

Ventajas

- ♦ Eficiente uso del ancho de banda.
- ◊ Baja complejidad numérica.
- Robustez contra el desvanecimiento generado por multitrayectorias.

Desventajas

- Dificultad para mantener la ortogonalidad en el receptor
- Sensibilidad a errores de sincronización en el tiempo, frecuencia y fase.
- ◊ Susceptibilidad a la interferencia interportadora.

Phase distortion

Errores de sincronización en frecuencia y fase: *Carrier frequency offset* y *Phase noise*.

Carrier frequency offset (CFO) \rightarrow imperfección del oscilador (constante) y corrimiento Doppler.

Phase noise (PHN) \rightarrow imperfección del oscilador (aleatoria). Su comportamiento depende del tipo de oscilador (PLL ó *free running oscillator*).

Phase distortion

Errores de sincronización en frecuencia y fase: *Carrier frequency offset* y *Phase noise*.

Carrier frequency offset (CFO) \rightarrow imperfección del oscilador (constante) y corrimiento Doppler.

Phase noise (PHN) \rightarrow imperfección del oscilador (aleatoria). Su comportamiento depende del tipo de oscilador (PLL ó *free running oscillator*).

Estos errores de sincronización comprometen la ortogonalidad entre las subportadoras y crean interferencia interportadora (ICI).

R. Carvajal EM-based channel estimation in multicarrier systems 3 / 45



R. Carvajal **EM-based channel estimation in multicarrier systems** 4 / 45



R. Carvajal **EM-based channel estimation in multicarrier systems** 4 / 45

Signal model

Los datos binarios se asumen independientes e idénticamente distribuídos (iid).

R. Carvajal EM-based channel estimation in multicarrier systems 5 / 45

Signal model

Los datos binarios se asumen independientes e idénticamente distribuídos (iid).

Modulación compleja (comúnmente). $s \rightarrow d$

Signal model

Los datos binarios se asumen independientes e idénticamente distribuídos (iid).

Modulación compleja (comúnmente). $\mathbf{s} \rightarrow \mathbf{d}$

Señales de banda angosta en paralelo.

Signal model

Los datos binarios se asumen independientes e idénticamente distribuídos (iid).

Modulación compleja (comúnmente). $\mathbf{s} \rightarrow \mathbf{d}$

Señales de banda angosta en paralelo.

Modulación digital vía Transformación Ortogonal (DFT, DCT, Wavelet, etc.).

Signal model

Los datos binarios se asumen independientes e idénticamente distribuídos (iid).

```
Modulación compleja (comúnmente). \mathbf{s} \rightarrow \mathbf{d}
```

Señales de banda angosta en paralelo.

Modulación digital vía Transformación Ortogonal (DFT, DCT, Wavelet, etc.).

$$\mathbf{x} = \mathbf{F}^H \mathbf{d}$$

R. Carvajal EM-based channel estimation in multicarrier systems 5 / 45

Prefijo Cíclico

Se inserta un prefijo cíclico (CP) de largo N_g para combatir el desvanecimiento.

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{M}_{CP} \\ \mathbf{I}_{N_C} \end{bmatrix} \mathbf{x},$$



R. Carvajal EM-based channel estimation in multicarrier systems 6 / 45

Modelo de canal

El canal es modelado como un filtro FIR complejo, $\mathbf{h} \in \mathbb{C}^{L}$.

Con la eliminación del prefijo cíclico, el canal es modelado como una matriz circular¹. $\mathbf{r} = \begin{bmatrix} \mathbf{0}_{Y_1, \dots, Y_n} & \mathbf{I}_{Y_n} \end{bmatrix} \bar{\mathbf{r}}$

$$egin{aligned} \mathbf{r} &= \begin{bmatrix} \mathbf{0}_{N_C imes N_g} & \mathbf{I}_{N_C} \end{bmatrix} \mathbf{r} \ &= ilde{\mathbf{H}} \mathbf{x} + m{\eta} \ &= ilde{\mathbf{H}} \mathbf{F}^H \mathbf{d} + m{\eta} \,, \end{aligned}$$

 $\mathbf{\tilde{H}} \in \mathbb{C}^{N_C \times N_C}$ es la matriz circular de canal dada por:

$$\tilde{\mathbf{H}} = \begin{bmatrix} h_0 & \cdots & 0 & \cdots & h_1 \\ \vdots & h_0 & \vdots & \cdots & h_2 \\ h_{L-1} & \vdots & h_0 & \cdots & \vdots \\ \vdots & \ddots & \cdots & \ddots & 0 \\ 0 & \cdots & h_{L-1} & \cdots & h_0 \end{bmatrix}$$

¹Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Process. Mag.*, vol. 17, pp. 29–48, May 2000.

R. Carvajal EM-based channel estimation in multicarrier systems 7/45

Caso especial: OFDM

· · .

En OFDM, la transformación ortofonal está dada por la DFT.

La señal recibida es demodulada en el receptor a través de la DFT (traspaso al dominion de la frecuencia).

$$\tilde{\mathbf{r}} = \mathbf{F}\mathbf{r} = \mathbf{F}\tilde{\mathbf{H}}\mathbf{x} + \mathbf{F}\boldsymbol{\eta} = \mathbf{F}\tilde{\mathbf{H}}\mathbf{F}^{H}\mathbf{d} + \tilde{\boldsymbol{\eta}} = \mathbf{D}\mathbf{d} + \tilde{\boldsymbol{\eta}},$$

donde

$$\mathbf{D}_{\mathrm{OFDM}} = \begin{bmatrix} D_0 & & \\ & \ddots & \\ & & D_{N_C-1} \end{bmatrix},$$

con

$$D_k = \sum_{l=0}^{L-1} h_l e^{-j2\pi \frac{kl}{N_C}}, \quad k = 0, 1, ..., N_C - 1.$$

R. Carvajal EM-based channel estimation in multicarrier systems 8 / 45

Phase Noise

 \rightarrow En *free-running* oscillators, PHN es modelado como un movimiento Browniano continuo:

$$\phi_{k+1} = \phi_k + v_k,$$

donde:

- v_k : Variabe aleatoria Gaussiana, iid, $v_k \sim \mathcal{N}(0, 2\pi\beta T/N_C)$.
- β : Ancho de banda del PHN.
- *T*: duración del símbolo.

Carrier Frequency Offset

- → CFO se puede modelar como una matriz diagonal $C_{\varepsilon} = e^{j \operatorname{diag}\left(\frac{2\pi \varepsilon k}{N_{C}}\right)}$, con $k = 0, 1, \dots, N_{C} 1$.
- $\rightarrow \varepsilon$ es el corrimiento Doppler *normalizado* ($|\varepsilon| \le 1/2$).

R. Carvajal **EM-based channel estimation in multicarrier systems** 9 / 45

Signal Model

Se asume que el prefijo cíclico ha sido removido exitosamente.

R. Carvajal **EM-based channel estimation in multicarrier systems** 10 / 45

Signal Model

Se asume que el prefijo cíclico ha sido removido exitosamente.

La señal transmitida \mathbf{x} se asume compuesta de una señal determinística (conocida) y una señal estocástica (desconocida).

Nos enfocamos en las señales en el dominio del tiempo, i.e.,

$$\mathbf{r} = (\mathbf{C}_{\varepsilon} \Phi) \, \mathbf{\tilde{H}} \mathbf{x} + \boldsymbol{\eta},$$

$$\Phi = \begin{bmatrix} e^{j\phi_0} & & \\ & \ddots & \\ & & e^{j\phi_{N_C-1}} \end{bmatrix}, \quad \mathbf{C}_{\varepsilon} = e^{j\operatorname{diag}\left(\frac{2\pi\varepsilon k}{N_C}\right)}.$$

R. Carvajal **EM-based channel estimation in multicarrier systems** 10 / 45



Diagrama en bloques de un sistema multiportadora.

R. Carvajal **EM-based channel estimation in multicarrier systems** 11/45

Señales propias e impropias

En general, se intenta explotar la representación compleja de las señales. Las señales complejas se pueden clasificar en *propias* e *impropias*². Señales propias: La matrix de autocorrelación es la misma para la parte real y la parte imaginaria. La matrix de *pseudocovarianza* es igual a cero.

²K. S. Miller, *Complex stochastic processes: An introduction to theory and applications,* London, UK: Addison-Wesley Publishing Company, Inc., 1974.

Señales propias e impropias

En general, se intenta explotar la representación compleja de las señales. Las señales complejas se pueden clasificar en *propias* e *impropias*². Señales propias: La matrix de autocorrelación es la misma para la parte real y la parte imaginaria. La matrix de *pseudocovarianza* es igual a cero. La mayoría de los esquemas de modulación son *propios*. Algunos no lo son, como BPSK y GMSK.

Ruido aditivo impropio puede estar presente en sistemas multiportadoras.

²K. S. Miller, *Complex stochastic processes: An introduction to theory and applications,* London, UK: Addison-Wesley Publishing Company, Inc., 1974.

Señales propias e impropias

En general, se intenta explotar la representación compleja de las señales. Las señales complejas se pueden clasificar en *propias* e *impropias*². Señales propias: La matrix de autocorrelación es la misma para la parte real y la parte imaginaria. La matrix de *pseudocovarianza* es igual a cero. La mayoría de los esquemas de modulación son *propios*. Algunos no lo son, como BPSK y GMSK. Ruido aditivo *impropio* puede estar presente en sistemas multiportadoras.

- : Utilizamos una representación real de las señales complejas.
- ⇒ Extendemos los resultados a cualquier tipo de modulación y a cualquier naturaleza del ruido aditivo.

²K. S. Miller, *Complex stochastic processes: An introduction to theory and applications*, London, UK: Addison-Wesley Publishing Company, Inc., 1974.

Señal de entrenamiento y señal de datos

En general, se utiliza una señal (secuencia) de entrenamiento para estimación y ecualización.

La transmisión de secuencias de entrenamiento reduce la cantidad de datos *útiles* a transmitir.

Señal de entrenamiento y señal de datos

En general, se utiliza una señal (secuencia) de entrenamiento para estimación y ecualización.

La transmisión de secuencias de entrenamiento reduce la cantidad de datos *útiles* a transmitir.

:. Expresamos la señal transmitida en función de la parte real e imaginaria y de la parte conocida y desconocida.

$$\Rightarrow \quad \bar{\mathbf{x}} = \left[\mathbf{x}_{\mathsf{R}}^{(\mathsf{T})^{T}} \mathbf{x}_{\mathsf{R}}^{(\mathsf{U})^{T}} \mathbf{x}_{\mathsf{I}}^{(\mathsf{T})^{T}} \mathbf{x}_{\mathsf{I}}^{(\mathsf{U})^{T}} \right]^{T} \in \mathbb{R}^{2N_{C}}$$

R. Carvajal EM-based channel estimation in multicarrier systems 13 / 45

Modelo en variables de estado

De la ecuación del sistema, $\mathbf{r} = (\mathbf{C}_{\varepsilon} \Phi) \tilde{\mathbf{H}} \mathbf{x} + \eta$, obtenemos:

$$\begin{split} \phi_{k+1} &= \phi_k + v_k, \\ \bar{\mathbf{x}}_{k+1}^{(\mathsf{U})} &= \bar{\mathbf{x}}_k^{(\mathsf{U})} = \bar{\mathbf{x}}_k^{(\mathsf{U})} \\ \mathbf{y}_k &= \mathbf{M}_k \bar{\mathbf{x}}_k + \bar{\eta}_k \end{split}$$

,

donde $\mathbf{y}_k = [\mathfrak{Re}\{r_k\} \ \mathfrak{Im}\{r_k\}]^T, \ \bar{\boldsymbol{\eta}}_k = [\mathfrak{Re}\{\boldsymbol{\eta}_k\} \ \mathfrak{Im}\{\boldsymbol{\eta}_k\}]^T, \ k = 0, 1, ..., N_C - 1,$

$$\begin{split} \mathbf{M}_{k} &= \begin{bmatrix} \mathbf{a}_{k} & -\mathbf{b}_{k} \\ \mathbf{b}_{k} & \mathbf{a}_{k} \end{bmatrix}, \\ \mathbf{a}_{k} &= (\cos \psi_{k}) \mathbf{q}_{k+1}^{T} \mathfrak{Re} \left\{ \tilde{\mathbf{H}} \right\} - (\sin \psi_{k}) \mathbf{q}_{k+1}^{T} \mathfrak{Im} \left\{ \tilde{\mathbf{H}} \right\}, \\ \mathbf{b}_{k} &= (\sin \psi_{k}) \mathbf{q}_{k+1}^{T} \mathfrak{Re} \left\{ \tilde{\mathbf{H}} \right\} + (\cos \psi_{k}) \mathbf{q}_{k+1}^{T} \mathfrak{Im} \left\{ \tilde{\mathbf{H}} \right\}, \\ \psi_{k} &= \phi_{k} + \frac{2\pi k\varepsilon}{N_{C}}. \end{split}$$

R. Carvajal EM-based channel estimation in multicarrier systems 14 / 45

Maximum Likelihood (ML) es una técnica usada para estimar los parámetros (θ) de una distribución.

R. Carvajal **EM-based channel estimation in multicarrier systems** 15 / 45

Maximum Likelihood (ML) es una técnica usada para estimar los parámetros (θ) de una distribución.

Propiedades

- **P1.** <u>Consistencia</u>: $\lim_{N\to\infty} \hat{\theta}_{ML} = \theta_0$.
- **P2.** <u>Eficiencia</u>: se obtiene la menor varianza posible (Cramér-Rao lower bound).
- **P3.** <u>Invarianza</u>: $\alpha = f(\theta) \Rightarrow \hat{\alpha}_{ML} = f(\hat{\theta}_{ML}).$
- P4. Sesgo: Existe sesgo. Asintóticamente centrado.

Maximum Likelihood (ML) es una técnica usada para estimar los parámetros (θ) de una distribución.

Propiedades

- **P1.** <u>Consistencia</u>: $\lim_{N\to\infty} \hat{\theta}_{ML} = \theta_0$.
- **P2.** <u>Eficiencia</u>: se obtiene la menor varianza posible (Cramér-Rao lower bound).
- **P3.** <u>Invarianza</u>: $\alpha = f(\theta) \Rightarrow \hat{\alpha}_{ML} = f(\hat{\theta}_{ML}).$
- P4. Sesgo: Existe sesgo. Asintóticamente centrado.

Procedimiento:	$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \arg \max_{\boldsymbol{\theta}} p(\mathbf{y} \boldsymbol{\theta}).$	
Alternativamente,	$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}),$	

R. Carvajal EM-based channel estimation in multicarrier systems 15 / 45

Maximum Likelihood (ML) es una técnica usada para estimar los parámetros (θ) de una distribución.

Propiedades

- **P1.** <u>Consistencia</u>: $\lim_{N\to\infty} \hat{\theta}_{ML} = \theta_0$.
- **P2.** <u>Eficiencia</u>: se obtiene la menor varianza posible (Cramér-Rao lower bound).
- **P3.** <u>Invarianza</u>: $\alpha = f(\theta) \Rightarrow \hat{\alpha}_{ML} = f(\hat{\theta}_{ML}).$
- P4. Sesgo: Existe sesgo. Asintóticamente centrado.

Procedimiento:	$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \arg \max_{\boldsymbol{\theta}} p(\mathbf{y} \boldsymbol{\theta}).$	
Alternativamente,	$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \arg \max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}),$	$ con \ell(\boldsymbol{\theta}) = \log p(\mathbf{y} \boldsymbol{\theta}). $

R. Carvajal EM-based channel estimation in multicarrier systems 15 / 45

Dificultad: presencia de variables escondidas

³A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. B*, vol. 39, no. 1, pp. 1–38, 1977.

R. Carvajal EM-based channel estimation in multicarrier systems 16 / 45

Dificultad: presencia de variables escondidas

Podemos utilizar Expectation-Maximization (EM) algorithm³: EM es un algoritmo iterativo que, en general, se utiliza para obtener el estimador de ML.

Las señales desconocidas se tratan como variables escondidas.



³A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. B*, vol. 39, no. 1, pp. 1–38, 1977.

R. Carvajal **EM-based channel estimation in multicarrier systems** 16 / 45

El algoritmo EM en sistemas multiportadora

Definición de los parámetros.

Definición de las variables escondidas.

El algoritmo EM en sistemas multiportadora

Definición de los parámetros.

Definición de las variables escondidas.

- Posibilidades
- ◊ h, x: variables aleatorias
- ◊ h, x: parámetros constantes
- h: variable aleatoria ,
 x: parámetro constante
- h: parámetro constante ,
 x: variable aleatoria

♦ <u>Dificultades</u>

- ◊ Nolinealidades
- Función de densidad de probabilidad singular
- Número de mediciones (N_C) menor que el número de parámetros

El algoritmo EM en sistemas multiportadora

Definición de los parámetros.

Definición de las variables escondidas.

Posibilidades

- Dificultades
- ◊ Nolinealidades

El algoritmo EM en sistemas multiportadora

Definición de los parámetros.

Definición de las variables escondidas.

- Posibilidades
- ◊ h, x: variables aleatorias

- ♦ <u>Dificultades</u>
- ♦ Nolinealidades
- Función de densidad de probabilidad singular

R. Carvajal **EM-based channel estimation in multicarrier systems** 17 / 45
El algoritmo EM en sistemas multiportadora

Definición de los parámetros.

Definición de las variables escondidas.

- Posibilidades
- ♦ h, x: parámetros constantes

- ♦ <u>Dificultades</u>
- ♦ Nolinealidades

 Número de mediciones (N_C) menor que el número de parámetros

R. Carvajal EM-based channel estimation in multicarrier systems 17 / 45

El algoritmo EM en sistemas multiportadora

Definición de los parámetros.

Definición de las variables escondidas.

Posibilidades

h: variable aleatoria ,
 x: parámetro constante

♦ <u>Dificultades</u>

- ♦ Nolinealidades
- Función de densidad de probabilidad singular
- Número de mediciones (N_C) menor que el número de parámetros

R. Carvajal EM-based channel estimation in multicarrier systems 17 / 45

El algoritmo EM en sistemas multiportadora

Definición de los parámetros.

Definición de las variables escondidas.

Posibilidades

Dificultades

- Nolinealidades
- Función de densidad de probabilidad singular

h: parámetro constante ,
 x: variable aleatoria

R. Carvajal EM-based channel estimation in multicarrier systems 17 / 45

El algoritmo EM en sistemas multiportadora

Definición de los parámetros.

Definición de las variables escondidas.

- Posibilidades
- ◊ h, x: variables aleatorias
- ◊ h, x: parámetros constantes
- h: variable aleatoria ,
 x: parámetro constante
- h: parámetro constante ,
 x: variable aleatoria

Dificultades

- ◊ Nolinealidades
- Función de densidad de probabilidad singular
- Número de mediciones (N_C) menor que el número de parámetros

$$\boldsymbol{\theta} = [\mathbf{h}, \boldsymbol{\varepsilon}, (\boldsymbol{\beta}T)^{-1}, \boldsymbol{\sigma}_{\boldsymbol{\eta}}^2], \\ p(\bar{\mathbf{x}}^{(\mathsf{U})}, \boldsymbol{\phi}, \mathbf{y} | \boldsymbol{\theta}) = p(\mathbf{y} | \bar{\mathbf{x}}^{(\mathsf{U})}, \boldsymbol{\phi}, \boldsymbol{\theta}) p(\boldsymbol{\phi}) p(\bar{\mathbf{x}}^{(\mathsf{U})}).$$

 $\Rightarrow \quad \mathcal{Q}(\theta, \hat{\theta}^{(i)}) = \mathbb{E}[\log p(\mathbf{y}|\bar{\mathbf{x}}^{(\mathsf{U})}, \phi, \theta)|\mathbf{y}, \hat{\theta}^{(i)}] + \mathbb{E}[\log p(\phi|\theta)|\mathbf{y}, \hat{\theta}^{(i)}] + \mathbb{E}[\log p(\bar{\mathbf{x}}^{(\mathsf{U})})|\mathbf{y}, \hat{\theta}^{(i)}].$ R. Carvajal **EM-based channel estimation in multicarrier systems** 17 / 45

Propiedades estadísticas de las señales

Phase noise:
$$p(\phi | \theta) = \frac{1}{\sqrt{|2\pi\Sigma_{\phi}|}} e^{-\frac{1}{2}\phi^{T}\Sigma_{\phi}^{-1}\phi}, \Sigma_{\phi} = \frac{2\pi\beta T}{N_{C}}\mathbf{R},$$

 $R_{p,q} = \min(p,q) + N_{g}.$

Propiedades estadísticas de las señales

Phase noise:
$$p(\phi | \theta) = \frac{1}{\sqrt{|2\pi\Sigma_{\phi}|}} e^{-\frac{1}{2}\phi^{T}\Sigma_{\phi}^{-1}\phi}, \Sigma_{\phi} = \frac{2\pi\beta T}{N_{C}}\mathbf{R},$$

 $R_{p,q} = \min(p,q) + N_{g}.$

<u>Señal transmitida</u>: Mapeo complejo - Modulación a través de una transformación ortogonal

R. Carvajal EM-based channel estimation in multicarrier systems 18 / 45

Propiedades estadísticas de las señales

Phase noise:
$$p(\phi | \theta) = \frac{1}{\sqrt{|2\pi\Sigma_{\phi}|}} e^{-\frac{1}{2}\phi^{T}\Sigma_{\phi}^{-1}\phi}, \Sigma_{\phi} = \frac{2\pi\beta T}{N_{C}}\mathbf{R},$$

 $R_{p,q} = \min(p,q) + N_{g}.$

<u>Señal transmitida</u>: Mapeo complejo - Modulación a través de una transformación ortogonal

 \Rightarrow CLT: $\bar{\mathbf{x}}$ se distribuye Normal.

Propiedades estadísticas de las señales

Phase noise:
$$p(\phi | \theta) = \frac{1}{\sqrt{|2\pi\Sigma_{\phi}|}} e^{-\frac{1}{2}\phi^{T}\Sigma_{\phi}^{-1}\phi}, \Sigma_{\phi} = \frac{2\pi\beta T}{N_{C}}\mathbf{R},$$

 $R_{p,q} = \min(p,q) + N_{g}.$

<u>Señal transmitida</u>: Mapeo complejo - Modulación a través de una transformación ortogonal

 \Rightarrow CLT: $\bar{\mathbf{x}}$ se distribuye Normal.

Señal recibida: Asumiendo ruido aditivo Gaussiano (media cero),

$$p(\mathbf{y} | \bar{\mathbf{x}}, \boldsymbol{\theta}) = \frac{1}{(2\pi)^{N_c} |\Sigma_y|^{1/2}} e^{-0.5(\mathbf{y} - \mathbf{M}\bar{\mathbf{x}})^T \Sigma_y^{-1}(\mathbf{y} - \mathbf{M}\bar{\mathbf{x}})},$$
$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_0^T & \cdots & \mathbf{M}_{N_c-1}^T \end{bmatrix}^T, \quad \Sigma_y = \mathbb{E}[\bar{\boldsymbol{\eta}}\bar{\boldsymbol{\eta}}^T]$$

R. Carvajal EM-based channel estimation in multicarrier systems 18 / 45

Establece una cota inferior para la matriz de covarianza del error para cualquier estimador sin sesgo. Se puede obtener como el inverso de la matriz de información de Fisher $(FIM)^4$.

⁴A. Stuart, K. Ord, and S. Arnold, *Kendall's advanced theory of statistics*, 6th ed., vol. 2A, London: Arnold, 1999.

Establece una cota inferior para la matriz de covarianza del error para cualquier estimador sin sesgo. Se puede obtener como el inverso de la matriz de información de Fisher $(FIM)^4$.

Fisher information matrix

$$\mathcal{I}_{\theta} = \mathbb{E}\left[\left.\frac{\partial \ell(\theta)}{\partial \theta}\right|_{\theta_{0}} \left(\left.\frac{\partial \ell(\theta)}{\partial \theta}\right|_{\theta_{0}}\right)^{T}\right]$$

⁴A. Stuart, K. Ord, and S. Arnold, *Kendall's advanced theory of statistics*, 6th ed., vol. 2A, London: Arnold, 1999.

Establece una cota inferior para la matriz de covarianza del error para cualquier estimador sin sesgo. Se puede obtener como el inverso de la matriz de información de Fisher $(FIM)^4$.

Fisher information matrix

$$\mathcal{I}_{\theta} = \mathbb{E}\left[\left.\frac{\partial \ell(\theta)}{\partial \theta}\right|_{\theta_{0}} \left(\left.\frac{\partial \ell(\theta)}{\partial \theta}\right|_{\theta_{0}}\right)^{T}\right] = \mathbb{E}\left[\left.\frac{\partial \mathcal{Q}(\theta, \hat{\theta}^{(i)})}{\partial \theta}\right|_{\theta=\theta_{0}} \left(\left.\frac{\partial \mathcal{Q}(\theta, \hat{\theta}^{(i)})}{\partial \theta}\right|_{\theta=\theta_{0}}\right)^{T}\right]$$

R. Carvajal EM-based channel estimation in multicarrier systems 19 / 45

⁴A. Stuart, K. Ord, and S. Arnold, *Kendall's advanced theory of statistics*, 6th ed., vol. 2A, London: Arnold, 1999.

Establece una cota inferior para la matriz de covarianza del error para cualquier estimador sin sesgo. Se puede obtener como el inverso de la matriz de información de Fisher $(FIM)^4$.

Fisher information matrix

$$\mathcal{I}_{\theta} = \mathbb{E}\left[\left.\frac{\partial \ell(\theta)}{\partial \theta}\right|_{\theta_{0}} \left(\left.\frac{\partial \ell(\theta)}{\partial \theta}\right|_{\theta_{0}}\right)^{T}\right] = \mathbb{E}\left[\left.\frac{\partial \mathcal{Q}(\theta, \hat{\theta}^{(i)})}{\partial \theta}\right|_{\theta=\theta_{0}} \left(\left.\frac{\partial \mathcal{Q}(\theta, \hat{\theta}^{(i)})}{\partial \theta}\right|_{\theta=\theta_{0}}\right)^{T}\right]$$

CRLB para $(\beta T)^{-1}$

Teorema:
$$\mathcal{I}_{(\beta T)^{-1}} < \frac{N_C(\beta T)^2}{2} \Rightarrow \text{CRLB}_{(\beta T)^{-1}} > \frac{2}{N_C(\beta T)^2}.$$

⁴A. Stuart, K. Ord, and S. Arnold, *Kendall's advanced theory of statistics*, 6th ed., vol. 2A, London: Arnold, 1999.

R. Carvajal EM-based channel estimation in multicarrier systems 19 / 45

Ejemplos numéricos: OFDM

 $\frac{\text{CRLB para } (\beta T)^{-1}}{\text{Simulaciones de MC: } \mathcal{I}_{(\beta T)^{-1}} = \int \left(\frac{\partial \mathcal{Q}(\theta, \hat{\theta}^{(i)})}{\partial (\beta T)^{-1}}\right)^2 p(\mathbf{y}) d\mathbf{y}.$

La señal transmitida se considera conocida.

- ► $\mathbf{h} = [0.5654 + 0.6846i, -2.1151 + 0.7739i, -0.9695 0.4545i, 0.0933 0.3718i]^T$,
- $\blacktriangleright \quad \varepsilon = 0.2537,$
- $\bullet \ \sigma_{\eta}^2 = 0.1,$
- ► SNR = 10[dB].

R. Carvajal **EM-based channel estimation in multicarrier systems** 20 / 45

Fisher information and Cramér-Rao lower bound for different number of subcarriers and different values of $(\beta T)^{-1}$.

N _C	$(\beta T)^{-1}$	$\mathcal{I}_{(\beta T)^{-1}}$	CRLB	$\sigma_{(\hat{eta T})^{-1}}$	% error
64	10	$2.86 imes10^{-1}$	3.50	1.87	18.70
128	10	$5.8 imes10^{-1}$	1.73	1.31	13.13
64	50	$1.02 imes 10^{-2}$	97.42	9.87	19.74
128	50	$2.07 imes 10^{-2}$	48.11	6.94	13.88
256	50	$4.62 imes 10^{-2}$	21.65	4.65	9.3
64	100	$1.6 imes10^{-3}$	633.39	25.17	25.17
128	100	$5.06 imes 10^{-3}$	197.45	14.05	14.05
256	100	$1.15 imes 10^{-2}$	86.85	9.31	9.31
64	500	$9.96 imes 10^{-6}$	1×10^5	317	63.37
128	500	$5.07 imes 10^{-5}$	$1.97 imes 10^4$	140	28.09
256	500	2.99×10^{-4}	3.34×10^3	57.8	11.56

R. Carvajal **EM-based channel estimation in multicarrier systems** 21 / 45

MLE performance under different training levels

Complete data: $(\boldsymbol{\phi}, \bar{\mathbf{x}}^{(U)}) \Rightarrow \mathcal{Q}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) = \mathbb{E}[\log p(\boldsymbol{\phi}, \bar{\mathbf{x}}^{(U)}, \mathbf{y} | \boldsymbol{\theta}) | \mathbf{y}, \hat{\boldsymbol{\theta}}^{(i)}]$

MLE performance under different training levels

Complete data: $(\boldsymbol{\phi}, \bar{\mathbf{x}}^{(U)}) \Rightarrow \mathcal{Q}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) = \mathbb{E}[\log p(\boldsymbol{\phi}, \bar{\mathbf{x}}^{(U)}, \mathbf{y}|\boldsymbol{\theta})|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(i)}]$

 $\Rightarrow p(\boldsymbol{\phi}, \bar{\mathbf{x}}^{(\mathsf{U})} | \mathbf{y}) ?$

MLE performance under different training levels

Complete data: $(\boldsymbol{\phi}, \bar{\mathbf{x}}^{(U)}) \Rightarrow \mathcal{Q}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) = \mathbb{E}[\log p(\boldsymbol{\phi}, \bar{\mathbf{x}}^{(U)}, \mathbf{y}|\boldsymbol{\theta})|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(i)}]$

 $\Rightarrow p(\phi, \bar{\mathbf{x}}^{(\mathsf{U})} | \mathbf{y}) ?$

◊ La parte desconocida (estocástica) de la señal transmitida es asumida Gaussiana (constante).

MLE performance under different training levels

Complete data: $(\boldsymbol{\phi}, \bar{\mathbf{x}}^{(U)}) \Rightarrow \mathcal{Q}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) = \mathbb{E}[\log p(\boldsymbol{\phi}, \bar{\mathbf{x}}^{(U)}, \mathbf{y}|\boldsymbol{\theta})|\mathbf{y}, \hat{\boldsymbol{\theta}}^{(i)}]$

 $\Rightarrow p(\phi, \bar{\mathbf{x}}^{(\mathsf{U})}|\mathbf{y}) ?$

◊ La parte desconocida (estocástica) de la señal transmitida es asumida Gaussiana (constante).

Solución: Rao-Blackwellization.

Rao-Blackwellization

=

Rao-Blackwellization corresponde a marginalizar el estado que aparece lineal en la ecuación de salida 5,6

$$\Rightarrow \qquad p(\bar{\mathbf{x}}^{(\mathsf{U})}|\phi,\mathbf{y}), p(\phi|\mathbf{y})? \quad \Leftrightarrow \quad p(\phi,\bar{\mathbf{x}}^{(\mathsf{U})}|\mathbf{y}) = p(\bar{\mathbf{x}}^{(\mathsf{U})}|\phi,\mathbf{y})p(\phi|\mathbf{y})$$

Rao-Blackwellized Particle Filter:

$$p(\boldsymbol{\phi}_{0:k}, \bar{\mathbf{x}}^{(\mathsf{U})} | \mathbf{y}_{0:k}) = \underbrace{p(\bar{\mathbf{x}}^{(\mathsf{U})} | \boldsymbol{\phi}_{0:k}, \mathbf{y}_{0:k})}_{\text{Optimal KF}} \underbrace{p(\boldsymbol{\phi}_{0:k} | \mathbf{y}_{0:k})}_{\text{PF}}.$$

<u>Particle Filter</u>: $p(\phi_{0:k}|\mathbf{y}_{0:k}) = \frac{p(y_k|\phi_{0:k},\mathbf{y}_{0:k-1})p(\phi_k|\phi_{0:k-1},\mathbf{y}_{0:k-1})}{p(y_k|\mathbf{y}_{0:k-1})} \times p(\phi_{0:k-1}|\mathbf{y}_{0:k-1}|).$

<u>Particle Smoother</u>: Se asume $\bar{\mathbf{x}}^{(U)}$ conocido.

⁵A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo methods in practice*, Springer-Verlag, 2001.

⁶J. S. Liu, *Monte Carlo strategies in scientific computing*, Springer, 2004.

R. Carvajal **EM-based channel estimation in multicarrier systems** 23 / 45



Frequency response difference between the true channel and the estimated channel for different PHN bandwidths and different training levels. $N_C = 64$.

R. Carvajal **EM-based channel estimation in multicarrier systems** 24 / 45



Frequency response difference between the true channel and the estimated channel for different PHN bandwidths and different training levels. $N_C = 256$.

R. Carvajal EM-based channel estimation in multicarrier systems 25 / 45

Channel estimation mean squared error (MSE) in OFDM systems for different number of subcarriers, training factors and PHN bandwidth.

N _C	Training	$(\beta T)^{-1}$	MSE
64	100%	100	0.0097
64	100%	10	0.0468
64	87.5%	100	0.1835
64	87.5%	10	0.2284
256	100%	100	$0.0057 \\ 0.0300$
256	100%	10	
256	87.5%	100	0.0103
256	87.5%	10	0.0922

R. Carvajal **EM-based channel estimation in multicarrier systems** 26 / 45

Parameter estimation and parameter estimation variance in OFDM systems with 64 subcarriers for different training factors and PHN bandwidth.

Training	$(\beta T)^{-1}$	$(\hat{\boldsymbol{\beta T}})^{-1}$	$\sigma_{(\hat{eta T})^{-1}}$	ê	$\sigma_{\hat{arepsilon}}$	$\hat{\sigma}_{\eta}^2$	$\sigma_{\hat{\sigma}_{\eta^2}}$
100%	100	8943.4	15163	0.3120	0.0068	0.1015	0.0070
100%	10	12.933	4.7770	0.4183	0.0177	0.0938	0.0104
87.5%	100	672.69	1841.7	0.2863	0.0508	0.0955	0.0823
87.5%	10	18.586	7.1489	0.4144	0.0471	0.1542	0.0496

Parameter estimation and parameter estimation variance in OFDM systems with 256 subcarriers for different training factors and PHN bandwidth.

Training	$(\beta T)^{-1}$	$(\hat{\boldsymbol{\beta T}})^{-1}$	$\sigma_{(\hat{\beta T})^{-1}}$	Ê	$\sigma_{\hat{arepsilon}}$	$\hat{\sigma}_{\eta}^2$	$\sigma_{\hat{\sigma}_{\eta^2}}$
100%	100	138.96	28.281	0.2891	0.0148	0.0986	0.0017
100%	10	12.619	2.0362	0.2318	0.0381	0.1006	0.0037
87.5%	100	728.09	268.98	0.2450	0.0084	0.2519	0.0049
87.5%	10	48.385	6.9522	0.1966	0.0182	0.2789	0.0065

R. Carvajal **EM-based channel estimation in multicarrier systems** 27 / 45

Maximum a Posteriori Estimation

Maximum a Posteriori es una técnica de optimización que trata los parámetros como variables aleatorias⁷.

Propiedades

- **P1.** <u>Consistencia</u>: $\lim_{N\to\infty} \hat{\theta}_{ML} = \theta_0$.
- **P2.** <u>Eficiencia</u>: se obtiene la menor varianza posible (Cramér-Rao lower bound) (asintóticamente eficiente y equivalente al estimador de ML bajo condiciones de regularidad).
- P3. Invarianza: en general, los estimadores de MAP no son invariantes.

Procedimiento:	$\hat{\boldsymbol{\theta}}_{\text{MAP}} = \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta} \mathbf{y}).$
Alternativamente,	$\hat{\boldsymbol{\theta}}_{\text{MAP}} = \arg \max_{\boldsymbol{\theta}} p(\mathbf{y} \boldsymbol{\theta}) p(\boldsymbol{\theta}),$

⁷C. Gourieroux, A. Monfort, *Statistics and econometric models*, vol. 1, Cambridge University Press, 1995.

R. Carvajal EM-based channel estimation in multicarrier systems 28 / 45

Maximum a Posteriori Estimation

Maximum a Posteriori es una técnica de optimización que trata los parámetros como variables aleatorias⁷.

Propiedades

- **P1.** <u>Consistencia</u>: $\lim_{N\to\infty} \hat{\theta}_{ML} = \theta_0$.
- **P2.** <u>Eficiencia</u>: se obtiene la menor varianza posible (Cramér-Rao lower bound) (asintóticamente eficiente y equivalente al estimador de ML bajo condiciones de regularidad).
- P3. Invarianza: en general, los estimadores de MAP no son invariantes.

Procedimiento:	$\hat{\boldsymbol{\theta}}_{\text{MAP}} = \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta} \mathbf{y}).$
Alternativamente,	$ \hat{\boldsymbol{\theta}}_{\text{MAP}} = \arg \max_{\boldsymbol{\theta}} p(\mathbf{y} \boldsymbol{\theta}) p(\boldsymbol{\theta}), \\ \hat{\boldsymbol{\theta}}_{\text{MAP}} = \arg \max_{\boldsymbol{\theta}} \log p(\mathbf{y} \boldsymbol{\theta}) + g(\boldsymbol{\theta}). $

⁷C. Gourieroux, A. Monfort, *Statistics and econometric models*, vol. 1, Cambridge University Press, 1995.

R. Carvajal EM-based channel estimation in multicarrier systems 28 / 45

Motivación

- Históricamente se han utilizado distribuciones para caracterizar los canales de comunicaciones (i.e. Rayleigh, Rice).
- Sparse channel estimation es un área de rápido crecimiento (UWC, hilly terrain, HDTV).
- Sparse parameter estimation ha sido abordado desde el punto de vista de la regularización (penalización).
- ♦ Es posible asignar una pdf que describa el nivel de *sparsity*.



R. Carvajal **EM-based channel estimation in multicarrier systems** 29 / 45

Sparse Prior Distribution

Sparsity se representa usualmente en términos de las pseudo-norma ℓ_0 , también denominada "norma" ℓ_0 (ℓ_0 -norm).

Sparse Prior Distribution

Sparsity se representa usualmente en términos de las pseudo-norma ℓ_0 , también denominada "norma" ℓ_0 (ℓ_0 -norm).

Problema de *sparsity*: $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{v}\|_0$ subject to $\mathbf{y} = \mathbf{A}\mathbf{v}$.

Sparse Prior Distribution

Sparsity se representa usualmente en términos de las pseudo-norma ℓ_0 , también denominada "norma" ℓ_0 (ℓ_0 -norm).

Problema de *sparsity*: $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{v}\|_0$ subject to $\mathbf{y} = \mathbf{A}\mathbf{v}$.

La solución es NP-hard⁸ \Rightarrow Relajación convexa: $\ell_0 \rightarrow \ell_1$.

Sparse Prior Distribution

Sparsity se representa usualmente en términos de las pseudo-norma ℓ_0 , también denominada "norma" ℓ_0 (ℓ_0 -norm).

Problema de sparsity: $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{v}\|_0$ subject to $\mathbf{y} = \mathbf{A}\mathbf{v}$.La solución es NP-hard⁸ \Rightarrow Relajación convexa: $\ell_0 \rightarrow \ell_1$.Nuevo problema: $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{v}\|_1$ subject to $\mathbf{y} = \mathbf{A}\mathbf{v}$.Fomulación más común:LASSO⁹ $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{y} - \mathbf{A}\mathbf{v}\|_2^2 + \lambda \|\mathbf{v}\|_1$, ó $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{y} - \mathbf{A}\mathbf{v}\|_2^2$ subject to $\|\mathbf{v}\|_1 \leq t$.

Solución: CVX, LARS.

⁸D. L. Donoho, "For most large underdetermined systems of linear equations the minimal ℓ_1 -norm solution is also the sparsest solution," *Commun. Pure Appl. Math.*, vol. 59, no. 7, pp.797–829, 2006.

⁹R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal. Statist. Soc B*, vol. 58, no. 1, pages 267–288, 1996.

R. Carvajal EM-based channel estimation in multicarrier systems 30 / 45

Soluciones comunes

Otros métodos:

- Métodos proximales (*Proximal methods*, PM)¹⁰,
- Aproximación lineal local (Local linear approximation, LLA)¹¹,
- Aproximación cuadrática local (Local quadratic approximation, LQA)¹².

R. Carvajal **EM-based channel estimation in multicarrier systems** 31 / 45

¹⁰R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM J. Control Opt.*, vol. 14, no. 5, pp. 877–898, 1976.

¹¹H. Zou and R. Li, "One-step sparse estimates in nonconcave penalized likelihood models," *The Annals of Statistics*, vol. 36, no. 4, pp. 1509–1533, 2008.

¹²D. Böhning and B. G. Lindsay, "Monotonicity of quadratic-approximation algorithms," *Ann. Inst. Statist. Math.*, vol. 40, no. 4, pp. 641–663, 1988.

Métodos proximales

- \rightarrow Se requiere obtener un "operador proximal" que entregue una solución cerrada.
- \rightarrow Mayorización / Minimización de la función costo por una función cuadrática.
- → <u>Limitación</u>: $\ell(\theta)$ debe tener primera derivada acotada (condición de Lipschitz de orden 1) y se debe poder evaluar.

$$|\dot{\ell}(\mathbf{v}_1) - \dot{\ell}(\mathbf{v}_2)| \le L(\mathbf{v}_1 - \mathbf{v}_2)$$

Métodos proximales

- \rightarrow Se requiere obtener un "operador proximal" que entregue una solución cerrada.
- \rightarrow Mayorización / Minimización de la función costo por una función cuadrática.
- → <u>Limitación</u>: $\ell(\theta)$ debe tener primera derivada acotada (condición de Lipschitz de orden 1) y se debe poder evaluar.

$$|\dot{\ell}(\mathbf{v}_1) - \dot{\ell}(\mathbf{v}_2)| \le L(\mathbf{v}_1 - \mathbf{v}_2)$$

Punto de vista Bayesiano

La penalización de norma ℓ_1 corresponde a una pdf Laplaciana.

$$p(\theta) = \frac{\lambda}{2} e^{-\lambda \|\theta\|_1}$$

R. Carvajal EM-based channel estimation in multicarrier systems 32 / 45

The EM algorithm for MAP estimation

Implementación tradicional: $Q_{\text{MAP}}(\theta, \hat{\theta}^{(i)}) = E[\log p(\mathbf{z}, \mathbf{y}|\theta)|\mathbf{y}, \hat{\theta}^{(i)}] + g(\theta).$

R. Carvajal **EM-based channel estimation in multicarrier systems** 33 / 45

¹³N. G. Polson and J. G. Scott, "Sparse Bayes estimation in non-Gaussian models via data augmentation," to appear *Biometrika*. Available online at http://arxiv.org/abs/1103.5407v2.

¹⁴O. Barndorff-Nielsen, J. Kent, and M. Sorensen, "Normal variance-mean mixtures and z distributions," *Int. Stat. Review*, vol. 50, no. 2, pp. 145–159, 1982.

¹⁵M. West, "On scale mixtures of normal distributions," *Biometrika*, vol. 74, no. 3, pp. 646–648, 1987.

The EM algorithm for MAP estimation

Implementación tradicional: $Q_{\text{MAP}}(\theta, \hat{\theta}^{(i)}) = E[\log p(\mathbf{z}, \mathbf{y}|\theta)|\mathbf{y}, \hat{\theta}^{(i)}] + g(\theta).$

Alternativa Bayesiana: Infinite Mixtures and the EM Algorithm

$$p(\theta) = \int p(\theta|\lambda) p(\lambda) d\lambda \quad \cos \theta |\lambda \sim \mathcal{N}(\mu_{\theta}, \Sigma_{\theta})$$

Diferentes nombres: variance-mean Gaussian mixture (VMGM)¹³, normal variance-mean mixture (NVMM)¹⁴, o normal scale mixture¹⁵.

¹³N. G. Polson and J. G. Scott, "Sparse Bayes estimation in non-Gaussian models via data augmentation," to appear *Biometrika*. Available online at http://arxiv.org/abs/1103.5407v2.

¹⁴O. Barndorff-Nielsen, J. Kent, and M. Sorensen, "Normal variance-mean mixtures and z distributions," *Int. Stat. Review*, vol. 50, no. 2, pp. 145–159, 1982.

¹⁵M. West, "On scale mixtures of normal distributions," *Biometrika*, vol. 74, no. 3, pp. 646–648, 1987.

R. Carvajal **EM-based channel estimation in multicarrier systems** 33 / 45

MAP-EM

$$\begin{split} \mathcal{Q}_{\text{MAP-EM}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) &= \mathcal{Q}_{\text{ML}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) + \mathcal{Q}_{\text{prior}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}), \\ \hat{\boldsymbol{\theta}}^{(i+1)} &= \arg\max_{\boldsymbol{\theta}} \mathcal{Q}_{\text{MAP-EM}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}). \end{split}$$
MAP-EM

$$\begin{split} \mathcal{Q}_{\text{MAP-EM}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) &= \mathcal{Q}_{\text{ML}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) + \mathcal{Q}_{\text{prior}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}), \\ \hat{\boldsymbol{\theta}}^{(i+1)} &= \arg\max_{\boldsymbol{\theta}} \mathcal{Q}_{\text{MAP-EM}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}). \end{split}$$

con

$$\mathcal{Q}_{\text{prior}}(\theta, \hat{\theta}^{(i)}) = \int \log p(\theta, \lambda) p(\lambda | \hat{\theta}^{(i)}) d\lambda$$
$$\frac{d}{d\theta} \mathcal{Q}_{\text{prior}}(\theta, \hat{\theta}^{(i)}) = \int \frac{d}{d\theta} \log p(\theta | \lambda) p(\lambda | \hat{\theta}^{(i)}) d\lambda$$

R. Carvajal EM-based channel estimation in multicarrier systems 34 / 45

MAP-EM

$$\begin{split} \mathcal{Q}_{\text{MAP-EM}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) &= \mathcal{Q}_{\text{ML}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}) + \mathcal{Q}_{\text{prior}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}), \\ \hat{\boldsymbol{\theta}}^{(i+1)} &= \arg\max_{\boldsymbol{\theta}} \mathcal{Q}_{\text{MAP-EM}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)}). \end{split}$$

con

$$\mathcal{Q}_{\text{prior}}(\theta, \hat{\theta}^{(i)}) = \int \log p(\theta, \lambda) p(\lambda | \hat{\theta}^{(i)}) d\lambda$$
$$\frac{d}{d\theta} \mathcal{Q}_{\text{prior}}(\theta, \hat{\theta}^{(i)}) = \int \frac{d}{d\theta} \log p(\theta | \lambda) p(\lambda | \hat{\theta}^{(i)}) d\lambda$$

$$\begin{split} \frac{d}{d\theta} \mathcal{Q}_{\text{prior}}(\theta, \hat{\theta}^{(i)}) &= \int [-\Sigma_{\theta}^{-1}(\lambda)(\theta - \mu_{\theta}(\lambda))] p(\lambda|\hat{\theta}^{(i)}) d\lambda \\ &= E_{\lambda|\hat{\theta}^{(i)}}[-\Sigma_{\theta}^{-1}(\lambda)] \theta + E_{\lambda|\hat{\theta}^{(i)}}[\Sigma_{\theta}^{-1}(\lambda)\mu_{\theta}(\lambda)]. \end{split}$$

R. Carvajal EM-based channel estimation in multicarrier systems 34 / 45

MAP-EM

Caso especial:

$$\begin{split} \log p(\boldsymbol{\theta}) &= \sum_{j=1}^{p} g\left(\frac{\theta_{j}}{\tau s_{j}}\right) \\ \frac{\partial \mathcal{Q}_{\text{prior}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)})}{\partial \theta_{j}} &= \frac{1}{\theta_{j}^{(i)}} \left. \dot{g}(\theta_{j}) \right|_{\theta_{j} = \hat{\theta}_{j}^{(i)}} \theta_{j}. \end{split}$$

Selection of variance-mean mixture representations for penalty functions. $p(\theta_j) = \int_0^\infty \mathcal{N}_{\theta_j}(\mu_j + \lambda_j u_j, \tau^2 s_j^2 \lambda_j) p(\lambda_j) d\lambda_j^{13}.$

Penalty function	$g(heta_j)$	и _j	μ_j	$p(\lambda_j)$
Ridge	$(\theta_i/ au)^2$	0	0	$\lambda_i = 1$
Lasso	$ \theta_i/\tau $	0	0	Exponential
Bridge	$ \theta_i/\tau ^{\alpha}$	0	0	Stable
Generalized	$\begin{bmatrix} (1+\alpha) \end{bmatrix}_{1} (1+ \theta_{i})$	0	0	F C
Double-Pareto	$\left\lfloor \frac{\tau}{\tau} \right\rfloor \log \left(1 + \frac{\tau}{(\alpha \tau)} \right)$, 0	0	Exp-Gamma

¹³N. G. Polson and J. G. Scott, "Sparse Bayes estimation in non-Gaussian models via data augmentation," to appear Biometrika. Available online at http://arxiv.org/abs/1103.5407v2.

Ejemplos numéricos: OFDM con CFO.

$$\begin{split} \bar{\mathbf{x}}_{k+1}^{(\mathsf{U})} &= \bar{\mathbf{x}}_{k}^{(\mathsf{U})} = \bar{\mathbf{x}}^{(\mathsf{U})}, \\ \mathbf{y}_{k} &= \begin{bmatrix} \mathbf{a}_{k} & -\mathbf{b}_{k} \\ \mathbf{b}_{k} & \mathbf{a}_{k} \end{bmatrix} \bar{\mathbf{x}}_{k} + \begin{bmatrix} \mathfrak{Re}\left\{\boldsymbol{\eta}_{k}\right\} \\ \mathfrak{Im}\left\{\boldsymbol{\eta}_{k}\right\} \end{bmatrix}, \end{split}$$

con

$$\mathbf{a} = (\cos \psi_k) \mathbf{q}_{k+1}^T \mathfrak{Re} \left\{ \tilde{\mathbf{H}} \right\} - (\sin \psi_k) \mathbf{q}_{k+1}^T \mathfrak{Im} \left\{ \tilde{\mathbf{H}} \right\} \\ \mathbf{b} = (\sin \psi_k) \mathbf{q}_{k+1}^T \mathfrak{Re} \left\{ \tilde{\mathbf{H}} \right\} + (\cos \psi_k) \mathbf{q}_{k+1}^T \mathfrak{Im} \left\{ \tilde{\mathbf{H}} \right\} \\ \psi_k = \frac{2\pi k\varepsilon}{N_C},$$

R. Carvajal EM-based channel estimation in multicarrier systems 36 / 45

Entrenamiento parcial

$$\log p(\theta) = \sum_{j=1}^{p} g\left(\frac{\theta_{j}}{\tau s_{j}}\right),$$

N-MSE := $(\mathbf{h} - \hat{\mathbf{h}})^{H} (\mathbf{h} - \hat{\mathbf{h}}) / (\mathbf{h}^{H} \mathbf{h})$



N-MSE average value (30 Monte Carlo simulations) for the estimates obtained by using (no regularized) ML and MAP, 50% training.

R. Carvajal EM-based channel estimation in multicarrier systems 37 / 45

Dificultades

Varianza del ruido desconocida

Varianza desconocida \Rightarrow Escalamiento, problema no-convexo.

<u>Cambio de variables</u>¹⁶: $\mathbf{\bar{h}} = \mathbf{h}/\sigma_{\eta}, \, \rho = \sigma_{\eta}^{-1}$

R. Carvajal EM-based channel estimation in multicarrier systems 38 / 45

¹⁶N. Städler, P. Bühlmann, and S. van de Geer, " ℓ_1 -penalization for mixture regression models," *Test*, vol. 19, no. 2, pp. 209–285, 2010.

¹⁷B. P. Carlin and T. A. Louis, *Bayes and empirical Bayes methods for data analysis*, 2nd ed., Chapman & Hall/CRC, 2000.

Dificultades

Varianza del ruido desconocida

Varianza desconocida \Rightarrow Escalamiento, problema no-convexo.

<u>Cambio de variables</u>¹⁶: $\mathbf{\bar{h}} = \mathbf{h}/\sigma_{\eta}, \, \rho = \sigma_{\eta}^{-1}$

Estimación de au

Parámetro de la distribución *a priori* desconocido \Rightarrow Estimación.

Solución clásica: Empirical Bayes¹⁷

R. Carvajal EM-based channel estimation in multicarrier systems 38 / 45

¹⁶N. Städler, P. Bühlmann, and S. van de Geer, " ℓ_1 -penalization for mixture regression models," *Test*, vol. 19, no. 2, pp. 209–285, 2010.

¹⁷B. P. Carlin and T. A. Louis, *Bayes and empirical Bayes methods for data analysis*, 2nd ed., Chapman & Hall/CRC, 2000.

Dificultades

Varianza del ruido desconocida

Varianza desconocida \Rightarrow Escalamiento, problema no-convexo.

Cambio de variables¹⁶:
$$\mathbf{\bar{h}} = \mathbf{h}/\sigma_{\eta}, \ \rho = \sigma_{\eta}^{-1}$$

Estimación de au

Parámetro de la distribución *a priori* desconocido \Rightarrow Estimación.

¹⁶N. Städler, P. Bühlmann, and S. van de Geer, " ℓ_1 -penalization for mixture regression models," *Test*, vol. 19, no. 2, pp. 209–285, 2010.

¹⁷B. P. Carlin and T. A. Louis, *Bayes and empirical Bayes methods for data analysis*, 2nd ed., Chapman & Hall/CRC, 2000.

R. Carvajal EM-based channel estimation in multicarrier systems 38 / 45



Comparison between Emp. Bayes (red) and ML-based (blue) estimates for τ , over 30 MonteCarlo simulations

R. Carvajal **EM-based channel estimation in multicarrier systems** 39 / 45

Más allá de la norma ℓ_1

→ <u>Elastic net</u>¹⁸: Combinación lineal entre ℓ_1 y ℓ_2 . Generalmente se desempeña mejor que LASSO.

$$g(\theta) = (1-\kappa) \|\theta\|_1 + \kappa \|\theta\|_2^2 \quad \Rightarrow \quad p_{EN}(\theta) = k_{EN} \left(\frac{\gamma}{2}\right)^p \frac{1}{\sqrt{|2\pi\sigma_{\theta}^2 \mathbf{I}|}} e^{-\gamma \|\theta\|_1} e^{-\frac{1}{2\sigma_{\theta}^2} \theta' \theta},$$

 \rightarrow Group LASSO¹⁹: *Sparsity* en grupos de elementos de θ .

$$g(\boldsymbol{\theta}) = \gamma \sum_{g=1}^{G} \sqrt{C_g} \|\boldsymbol{\theta}_g\|_2 \quad \Rightarrow \quad p_{\mathrm{GL}}(\boldsymbol{\theta}) = k_{GL} \prod_{g=1}^{G} e^{-\gamma \sqrt{C_g}} \|\boldsymbol{\theta}_g\|_2$$

 \rightarrow Opinion pool²⁰: opinión de expertos. $p_{OP}(\theta) = k \prod_m (p_m(\theta))^{\alpha_m}$.

¹⁸H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," J. R. Statist. Soc. B, vol. 67, no. 2, pp. 301–320, 2005.

¹⁹M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. R. Stat. Soc. B*, vol. 68, no. 1, pp. 49–67, 2007.

²⁰C. Genest, K. J. McConway, and M. J. Schervish, "Characterization of externally Bayesian pooling operators," *Annals of Statistics*, vol. 14, no. 2, pp. 487–501, 1986.

R. Carvajal **EM-based channel estimation in multicarrier systems** 40 / 45

Más allá de la norma ℓ_1

<u>Kernels</u>²¹: La distribución condicional $p(\theta|\lambda)$ puede ser interpretada como un kernel.

$$p(\boldsymbol{\theta}) = \int p(\boldsymbol{\theta}|\boldsymbol{\lambda}) p(\boldsymbol{\lambda}) d\boldsymbol{\lambda},$$

donde $p(\theta|\lambda)$ satisface

$$\frac{d\log p(\boldsymbol{\theta}|\boldsymbol{\lambda})}{d\boldsymbol{\theta}} = \mathbf{A}(\boldsymbol{\theta})b(\boldsymbol{\lambda}) + \mathbf{r}(\boldsymbol{\theta}).$$

$$\Rightarrow \frac{d\mathcal{Q}_{\text{prior}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(i)})}{d\boldsymbol{\theta}} = \mathbf{A}(\boldsymbol{\theta})\int \mathbf{b}(\boldsymbol{\lambda})p(\boldsymbol{\lambda}|\hat{\boldsymbol{\theta}}^{(i)})d\boldsymbol{\lambda} + \mathbf{r}(\boldsymbol{\theta}),$$

$$\frac{d\log p(\boldsymbol{\theta}|\boldsymbol{\lambda})}{d\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}^{(i)}} = \mathbf{A}(\hat{\boldsymbol{\theta}}^{(i)})\int \mathbf{b}(\boldsymbol{\lambda})p(\boldsymbol{\lambda}|\hat{\boldsymbol{\theta}}^{(i)})d\boldsymbol{\lambda} + \mathbf{r}(\hat{\boldsymbol{\theta}}^{(i)}).$$

R. Carvajal **EM-based channel estimation in multicarrier systems** 41 / 45

Más allá de la norma ℓ_1

<u>Kernels</u>²¹: La distribución condicional $p(\theta|\lambda)$ puede ser interpretada como un kernel.

$$p(\boldsymbol{\theta}) = \int p(\boldsymbol{\theta}|\boldsymbol{\lambda}) p(\boldsymbol{\lambda}) d\boldsymbol{\lambda},$$

donde $p(\theta|\lambda)$ satisface

$$\frac{d\log p(\boldsymbol{\theta}|\boldsymbol{\lambda})}{d\boldsymbol{\theta}} = \mathbf{A}(\boldsymbol{\theta})b(\boldsymbol{\lambda}) + \mathbf{r}(\boldsymbol{\theta}).$$

$$\Rightarrow \frac{d\mathcal{Q}_{\text{prior}}(\theta, \hat{\theta}^{(i)})}{d\theta} = \mathbf{A}(\theta) \int \mathbf{b}(\lambda) p(\lambda|\hat{\theta}^{(i)}) d\lambda + \mathbf{r}(\theta),$$

$$\frac{d\log p(\theta|\lambda)}{d\theta} \Big|_{\theta = \hat{\theta}^{(i)}} = \mathbf{A}(\hat{\theta}^{(i)}) \int \mathbf{b}(\lambda) p(\lambda|\hat{\theta}^{(i)}) d\lambda + \mathbf{r}(\hat{\theta}^{(i)}).$$

SISTEMA DE ECUACIONES !!!

²¹J. Keilson and F. W. Steutel, "Mixtures of distributions, moment inequalities and measures of exponentially and normality," *Ann. Probability*, vol. 2, no. 1, pp. 112–130, 1974.

R. Carvajal **EM-based channel estimation in multicarrier systems** 41 / 45

Más allá de la norma ℓ_1

Caso particular: Bridge regression ($\|\cdot\|_2^2$), prior Laplaciana, kernel Laplaciano.

$$p(\theta|\lambda) = \frac{1}{2\lambda} e^{-\lambda \|\theta\|_{1}}.$$

$$p(\lambda) = \frac{(n\gamma)^{1/q}}{2} S\left(\frac{(n\gamma)^{1/q}}{2} \frac{1}{\lambda}\right), \quad S(\cdot) \text{ es una distribución "estable".}$$

$$\Rightarrow \quad \mathcal{Q}_{\text{prior}}(\theta, \hat{\theta}^{(i)}) = \|\theta\|_{1} \dot{f}(\hat{\theta}^{(i)})$$

corresponde a la aproximación LLA de la penalización $f(\|\theta\|_1)$ en una regresión Bridge.

R. Carvajal EM-based channel estimation in multicarrier systems 42/45

Otra aplicación: Sparse FIR Filter Estimation With Quantized Output Data



R. Carvajal EM-based channel estimation in multicarrier systems 43/45

EM-based channel estimation in multicarrier systems

Publicaciones

- 1. R. Carvajal, J. C Agüero, B. I. Godoy, and G. C. Goodwin, "On the accuracy of phase noise bandwidth estimation in OFDM systems," in *Proc. of the 12th IEEE International Workshop on Signal Processing Advances in Wireless Communications, SPAWC* 2011.
- 2. R. Carvajal, J. C Agüero, B. I. Godoy, and G. C. Goodwin, "EM-based channel estimation in OFDM systems with phase noise," in *Proc. of IEEE Global Communications Conference, GLOBECOM* 2011.
- **3.** R. Carvajal, J. C Agüero, B. I. Godoy, G. C. Goodwin and J. I. Yuz, "EM-based identification of sparse FIR sysems having quantized data," in *Proc. of the 16th IFAC Symposium on System Identification, SySID* 2012.
- **4.** R. Carvajal, B. I. Godoy, J. C. Agüero, and G. C. Goodwin, "EM-based sparse channel estimation in OFDM systems," in *Proc. of the 13th IEEE International Workshop on Signal Processing Advances in Wireless Communications, SPAWC* 2012.
- **5.** R. Carvajal, J. C. Agüero, B. I. Godoy, and G. C. Goodwin, "EM-Based Maximum Likelihood channel estimation in multicarrier systems with phase distortion," to appear *IEEE Trans. Veh. Technol.*, 2013.

R. Carvajal EM-based channel estimation in multicarrier systems 44 / 45

Preguntas ?

R. Carvajal EM-based channel estimation in multicarrier systems 45 / 45