

IDENTIFICAZIONE SET MEMBERSHIP CON DINAMICHE NON MODELLATE

Michele TARAGNA

Dipartimento di Automatica e Informatica

Politecnico di Torino

`michele.taragna@polito.it`



Corso di III livello

“Experimental modeling: costruzione di modelli da dati sperimentali”

Sommario

- Motivazioni dell'identificazione di modelli d'incertezza dinamici.
- Teoria dell'identificazione di model set.
- Errori di identificazione e concetti di ottimalità.
- Assunzioni su sistema e rumore e loro validazione.
- Norme per misurare l'errore di identificazione: \mathcal{H}_∞ , ℓ_1 , \mathcal{H}_2 .
- Identificazione \mathcal{H}_∞ :
 - metodi di validazione;
 - algoritmi lineari, “two-stage”, interpolatori; algoritmi ottimi e “nearly-optimal”;
 - modelli di incertezza di ordine ridotto;
 - (interazione identificazione-controllo.)
- Appendice: norme, spazi di segnali e di sistemi.

Motivazioni

Problema: fare *inferenze* su un sistema dinamico reale S^o a partire da informazioni generali su esso e da N misure sperimentali corrotte da rumore

$$y^N = F_N(S^o) + e^N$$

↑ ↑ ↓

vettore delle misure (noto)	operatore informazione (noto)	rumore di misura (non noto)
-----------------------------------	-------------------------------------	-----------------------------------

Esempi di inferenza:

- predizione delle risposte future di S^o ;
- controllo delle risposte future di S^o ;
- fault detection;
- diagnostica;
- ...

- **Informazione sperimentale**

- *Misure nel tempo*: N campioni dell'uscita y del sistema S^o , inizialmente a riposo, cui si applica un ingresso noto u

$$y_\ell = \sum_{k=0}^{\ell} h_k^{S^o} u_{\ell-k} + e_\ell, \quad \ell = 0, \dots, N-1$$

⇓

operatore informazione : $F_N(S) = F_N h^S$

$$\text{con } F_N := \begin{bmatrix} U_N & 0_{N \times \infty} \end{bmatrix} \in \mathbb{R}^{N \times \infty}, U_N := \begin{bmatrix} u_0 & 0 & \cdots & 0 \\ u_1 & u_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ u_{N-1} & u_{N-2} & \cdots & u_0 \end{bmatrix} \in \mathbb{R}^{N \times N}$$

- *Misure in frequenza*: parte reale ed immaginaria di $N/2$ campioni della risposta in frequenza del sistema $S^o(\omega_k)$, $k = 1, \dots, N/2$:

$$\begin{cases} y_{2k-2} &= \Re(S^o(\omega_k)) + e_{2k-2}, \\ y_{2k-1} &= \Im(S^o(\omega_k)) + e_{2k-1}, \end{cases} \quad k = 1, \dots, N/2$$

$$\Downarrow$$

operatore informazione : $F_N(S) = F_N h^S$

$$F_N := \left[\Omega^T(\omega_1) \cdots \Omega^T(\omega_{N/2}) \right]^T \in \mathbb{R}^{N \times \infty}$$

$$\Omega(\omega) := \begin{bmatrix} \Omega_1(\omega) \\ \Omega_2(\omega) \end{bmatrix} = \begin{bmatrix} \Re(\Psi(\omega)) \\ \Im(\Psi(\omega)) \end{bmatrix} \in \mathbb{R}^{2 \times \infty}$$

$$\Psi(\omega) := \begin{bmatrix} 1 & e^{-j\omega} & e^{-j2\omega} & \dots \end{bmatrix} \in \mathbb{C}^{1 \times \infty}$$

- **Approccio:**

- stimare un modello \hat{M}
- fare l'inferenza sul modello \hat{M}
- calcolare l'errore di inferenza $E(\hat{M})$ commesso facendo l'inferenza su \hat{M} anziché su S^o

- **Problemi:**

- dato un modello \hat{M} , calcolare l'errore $E(\hat{M})$
 - cercare un modello \hat{M} che dia un “piccolo” errore $E(\hat{M})$
 - analizzare il comportamento dell'errore $E(\hat{M})$ per $N \rightarrow \infty$
 - progettare l'esperimento in modo che $E(\hat{M})$ sia “piccolo”
- Senza fare ipotesi su S^o e su e^N , l'errore $E(\hat{M})$ è illimitato



quali informazioni “a priori” su S^o ed e^N si possono assumere?

Teoria dell'identificazione statistica

- **Ipotesi:**

- $S^o \in K = \{M_n(p), p \in \mathbb{R}^n\}$: insieme di modelli dinamici parametrici
- e^N : rumore stocastico con d.d.p. nota, eventualmente filtrato



problema di stima parametrica statistica:

$$y^N = F_N(M_n(p)) + e^N$$

- Se $S^o \notin K$, come avviene in molti casi, allora

$$y^N = F_N(M_n(p)) + e^N + \text{dinamiche non modellate}$$



come tener conto delle dinamiche non modellate?

Teoria dell'identificazione di model set

- **Ipotesi:**

- $S^o \in K$ = sottoinsieme dello spazio dei sistemi dinamici \mathcal{S} avente una parametrizzazione non necessariamente finita (ad esempio, sistemi con risposta impulsiva esponenzialmente decrescente)
- $e^N \in \mathcal{B}_e$ = insieme limitato

- **Feasible Systems Set:** insieme di tutti i sistemi consistenti con le ipotesi e con i dati sperimentali (*unfalsified systems set*)

$$FSS(K, \mathcal{B}_e, F_N, y^N) := \{S \in K : y^N = F_N(S) + \tilde{e}^N, \tilde{e}^N \in \mathcal{B}_e\}$$

- FSS riassume tutte e sole le informazioni sul sistema da identificare (ipotesi su S^o ed e^N , dati y^N)
- $S^o \in FSS$ se le ipotesi sono vere \Rightarrow validare le ipotesi \Leftrightarrow controllare se FSS non è vuoto
- se le ipotesi sono troppo deboli $\Rightarrow FSS$ risulta non limitato

- Problema:

FSS non è sovente rappresentato nella forma più adatta all'uso \Rightarrow
si cercano insiemi di modelli che *includano* FSS e siano di forma *opportuna*



- **Model set**: un insieme di modelli $\mathcal{M} \subseteq \mathcal{S}$ è un model set per S^o se

$$\mathcal{M} \supseteq FSS \ni S^o$$

- Si considerano model set *additivi*:

$$\mathcal{M} = \left\{ \hat{M} + \Delta : \|\Delta\|_{\mathcal{S}} \leq \gamma \right\}$$

$\hat{M} \in \mathcal{S}$: modello nominale

$\Delta = S^o - \hat{M}$: errore di modello (dinamica non modellata)

$\|\cdot\|_{\mathcal{S}}$: norma nello spazio dei sistemi dinamici \mathcal{S}

- **Algoritmo di identificazione ϕ :**

$$\phi(K, \mathcal{B}_e, F_N, y^N) = \hat{M}$$

- Misura dell'errore commesso facendo inferenze su \hat{M} anziché su S^o :

$$\delta(\hat{M}) = \|S^o - \hat{M}\|_{\mathcal{S}} \leq \sup_{S \in FSS} \|S - \hat{M}\|_{\mathcal{S}}$$

tale limite superiore è il più stringente ed è limitato solo grazie alle ipotesi su S^o



dato un modello \hat{M} , il model set additivo più “piccolo” è

$$\hat{\mathcal{M}} = \left\{ \hat{M} + \Delta : \|\Delta\|_{\mathcal{S}} \leq \sup_{S \in FSS} \|S - \hat{M}\|_{\mathcal{S}} \right\}$$

- **Errore di identificazione di un algoritmo ϕ ,
errore di identificazione di un model set $\hat{\mathcal{M}}$:**

$$E(\phi) = E(\hat{\mathcal{M}}) := \sup_{S \in FSS(K, \mathcal{B}_e, F_N, y^N)} \|S - \phi(K, \mathcal{B}_e, F_N, y^N)\|_S$$

- è il minimo errore di inferenza garantito dal modello $\hat{\mathcal{M}}$, identificato mediante l'algoritmo ϕ che opera sulle informazioni disponibili $(K, \mathcal{B}_e, F_N, y^N)$
- non è conservativo nei confronti del rumore, ma tiene conto della sua effettiva realizzazione verificatasi nell'esperimento preso in considerazione

- **Algoritmo e model set ottimi**

i) un model set \mathcal{M}^* è ottimo se minimizza l'errore di inferenza per l'*attuale* insieme di informazioni disponibili $(K, \mathcal{B}_e, F_N, y^N)$:

$$E(\mathcal{M}^*) = \inf_{\hat{\mathcal{M}} \subseteq \mathcal{S}} E(\hat{\mathcal{M}}) = r(K, \mathcal{B}_e, F_N, y^N)$$

$$\begin{aligned} r(K, \mathcal{B}_e, F_N, y^N) &:= \inf_{\phi} \sup_{S \in FSS} \|S - \phi(K, \mathcal{B}_e, F_N, y^N)\|_{\mathcal{S}} \\ &= \text{raggio di informazione} \end{aligned}$$

ii) un algoritmo ϕ^* è ottimo se minimizza l'errore di inferenza per *qualsiasi* possibile insieme di informazioni $(K, \mathcal{B}_e, F_N, y^N)$:

$$E(\phi^*) = r(K, \mathcal{B}_e, F_N, y^N), \quad \forall (K, \mathcal{B}_e, F_N, y^N)$$

- è ottimo l'*algoritmo centrale*

$$\phi^c (y^N) = M^c = \underbrace{\arg \inf_{M \in \mathcal{S}} \sup_{S \in FSS} \|S - M\|_{\mathcal{S}}}_{\text{centro di Chebicheff di } FSS}$$

- ⇒ è ottimo il corrispondente *model set centrale*

$$\mathcal{M}^c = \{M^c + \Delta : \|\Delta\|_{\mathcal{S}} \leq r\}$$

- in molti casi, non esistono metodi di calcolo del modello centrale M^c



si considerano **algoritmi subottimi**

- **Algoritmi subottimi**

- * algoritmi lineari

- * algoritmi non lineari a due passi, non parametrici

- Helmicki-Jacobson-Nett: *TAC* 1991; Partington: *IJC* 1991;

- Gu-Khargonekar: *Automatica* 1992, *TAC* 1992; . . .

- * algoritmi non lineari interpolatori, non parametrici

- Chen-Nett-Fan: *ACC* 1992, *TAC* 1995; Bai-Raman: *ACC* 1992;

- Chen-Nett: *CDC* 1993, *TAC* 1995; Gu-Xiong-Zhou: *S&CL* 1993; . . .

- * algoritmi misti parametrici - non parametrici

- Wahlberg-Ljung: *TAC* 1992; Kosut-Boyd: *TAC* 1992; Younce-Rohrs: *TAC* 1992;

- Zhou-Kimura: *TAC* 1995; Giarré-Milanese-Taragna: *TAC* 1997; . . .

- **Algoritmi subottimi**

- sono computazionalmente più semplici dell'algoritmo centrale, ma richiedono di rinunciare almeno in parte all'accuratezza nell'identificazione
- il modello nominale può non essere di forma adatta al successivo uso (ad esempio, ordine troppo elevato, ...) \Rightarrow
è opportuno trovare algoritmi che forniscano model set con modelli nominali di *ordine basso*, a costo di un possibile ulteriore *deterioramento* della qualità dell'identificazione



come si misura tale deterioramento?

- **Livello di ottimalità di un algoritmo ϕ :**

un algoritmo ϕ è α -ottimo se $\exists \alpha \in \mathbb{R}$ tale che

$$E(\phi(K, \mathcal{B}_e, F_N, y^N)) \leq \alpha \cdot r(K, \mathcal{B}_e, F_N, y^N), \quad \forall (K, \mathcal{B}_e, F_N, y^N)$$

- misura la qualità dell'algoritmo rispetto all'algoritmo ottimo nel *caso pessimo*, considerando ogni possibile informazione (ipotesi e dati)
- gli algoritmi non lineari interpolatori sono 2-ottimi per qualsiasi norma e per tale motivo sono detti “almost optimal”

- **Livello di ottimalità di un model set $\hat{\mathcal{M}}$:**

$$\alpha(\hat{\mathcal{M}}) = E(\hat{\mathcal{M}}) / r \geq 1$$

- misura la qualità *reale* del model set identificato rispetto a quello ottimo, considerando le sole informazioni disponibili (ipotesi e dati)
- la valutazione esatta di $\alpha(\hat{\mathcal{M}})$ è difficoltosa, poiché il calcolo di $E(\hat{\mathcal{M}})$ ed r è in generale un problema di complessa soluzione

- **Convergenza:**

un algoritmo ϕ è detto convergente se

$$\lim_{\substack{N \rightarrow \infty \\ \mathcal{B}_e \rightarrow \emptyset}} E(\phi) = 0$$

- **Convergenza robusta:**

un algoritmo ϕ è detto robustamente convergente se è convergente indipendentemente dalle informazioni a priori sul sistema e sul rumore.

Classificazione

- Norma $\|\cdot\|_S$ che misura l'errore di identificazione

$$\|S(z)\|_\infty^W := \sup_{0 \leq \omega \leq 2\pi} |W^{-1}(\omega)S(\omega)| \Rightarrow \text{identificazione } \mathcal{H}_\infty$$

$$\|S(z)\|_1 := \sum_{k=-\infty}^{\infty} |h_k^S| \Rightarrow \text{identificazione } \ell_1$$

$$\|S(z)\|_2 := \sqrt{\sum_{k=-\infty}^{\infty} |h_k^S|^2} \Rightarrow \text{identificazione } \mathcal{H}_2$$

- Informazione “a priori” sul rumore di misura e^N

$$e^N \in \mathcal{B}_{e,\infty} := \{\tilde{e}^N \in \mathbb{R}^N : \|A\tilde{e}^N\|_\infty^{W_e} \leq \varepsilon\}$$

$$e^N \in \mathcal{B}_{e,2} := \{\tilde{e}^N \in \mathbb{R}^N : \|\tilde{e}^N\|_2 \leq \varepsilon\}$$

$\mathcal{B}_{e,\infty}$ è più adatto a caratterizzare deterministicamente il rumore di misura, in modo simile a quanto avviene con l'identificazione statistica

- **Informazione “a priori” sul sistema S^o**

\mathcal{S} : spazio di Banach dei sistemi dinamici SISO, a tempo discreto, causali, LTI, BIBO-stabili, eventualmente a parametri distribuiti, con

$h^S := \{h_k^S\}_{k=0}^{\infty}$: risposta impulsiva di $S \in \mathcal{S}$

$S(z) := \sum_{k=0}^{\infty} h_k^S z^k$: funzione di trasferimento di S (o Λ -trasformata di h^S)

$S^o \in \mathcal{H}_{\infty}(\mathbb{D})$: sistema reale da identificare, parzialmente noto, con

$\mathcal{H}_{\infty}(\mathbb{D}) := \left\{ f : \mathbb{D} \rightarrow \mathbb{C} \mid f \text{ analitica in } \mathbb{D}, \|f\|_{\infty} := \operatorname{ess\,sup}_{z \in \mathbb{D}} |f(z)| < \infty \right\}$

$\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$: cerchio unitario aperto

1. S^o è un sistema esponenzialmente stabile, avente un guadagno stazionario massimo noto $M > 0$ a fronte di un ingresso $\rho^{-k} e^{i\omega k}$, con $\rho > 1$:

$$S^o \in K_{\rho, M}^{(1)} := \{S \in \mathcal{H}_\infty(\mathbb{D}) : S(z) \in \mathcal{H}_{\rho, M}(\mathbb{D})\}$$

$$\mathcal{H}_{\rho, M}(\mathbb{D}) := \left\{ f : \mathbb{D} \rightarrow \mathbb{C} \mid f \text{ analitica in } \mathbb{D}_\rho, \|f\|_{\infty, \rho} := \sup_{z \in \mathbb{D}_\rho} |f(z)| \leq M \right\}$$

$$\mathbb{D}_\rho := \{z \in \mathbb{C} : |z| \leq \rho\} : \text{cerchio chiuso di raggio } \rho$$

[Helmicki-Jacobson-Nett: *TAC*'91; Partington: *IJC*'91; Gu-Khargonekar: *TAC*'92;

Gu-Xiong-Zhou: *SCL*'93; Chen-Nett-Fan: *TAC*'95; . . .]

2. S^o è un sistema esponenzialmente stabile, con un inviluppo $M\rho^{-k}$ della risposta impulsiva, con $M > 0$ e $\rho > 1$ costanti note:

$$S^o \in K_{\rho, M}^{(2)} := \{S \in \mathcal{H}_\infty(\mathbb{D}) : |h_k^S| \leq M\rho^{-k}, \forall k \geq 0\}$$

[Giarré-Milanese-Taragna: *TAC*'97; Milanese-Taragna: *TAC*'02; . . .]

3. S^o ha un limite superiore noto $\gamma > 0$ del modulo della derivata della funzione di trasferimento:

$$S^o \in K_\gamma^{(3)} \doteq \left\{ S \in \mathcal{H}_\infty(\mathbb{D}) : \sup_{z \in \mathbb{D}} \left| \frac{dS(z)}{dz} \right| \leq \gamma \right\}$$

[Glaum-Lin-Zames: *CDC*'96]

4. S^o ha un limite superiore noto $\gamma > 0$ del modulo della parte reale e della parte immaginaria della derivata della funzione di trasferimento:

$$S^o \in K_\gamma^{(4)} \doteq \left\{ S \in \mathcal{H}_\infty(\mathbb{D}) : \sup_{\omega \in [0, 2\pi]} \left| \frac{dS_{R/I}(\omega)}{d\omega} \right| \leq \gamma \right\}$$

[Milanese-Novara-Taragna: *ECC*'01]

$$K_{\rho, M}^{(1)} \subset K_{\rho, M}^{(2)} \subset K_\gamma^{(3)} \subset K_\gamma^{(4)}, \text{ con } \gamma = M\rho / (\rho - 1)^2$$

Identificazione \mathcal{H}_∞

- Norma $\|\cdot\|_{\mathcal{S}}$ che misura l'errore di identificazione

$$\|S\|_{\mathcal{S}} = \|S(z)\|_{\infty}^W = \sup_{0 \leq \omega \leq 2\pi} |W^{-1}(\omega)S(\omega)|$$

$W^{-1}(\omega)$: funzione peso *opportunamente* scelta in base all'uso del modello
(ad esempio, per il controllo robusto H_∞)

Algoritmi lineari

- Gli algoritmi lineari operano linearmente sui dati sperimentali.
- Gli algoritmi lineari hanno una struttura estremamente semplice, essendo algoritmi dei minimi quadrati eventualmente pesati.
- Gli algoritmi lineari sono classificabili in due categorie:
 - gli algoritmi lineari “untuned”
 - gli algoritmi lineari “tuned”a seconda che siano indipendenti o no dalle informazioni a priori disponibili sul sistema e sul rumore.
- Le limitazioni degli algoritmi lineari giustificano lo studio dei più sofisticati algoritmi non lineari.

Limitazioni fondamentali degli algoritmi lineari

- Nel caso dell'identificazione \mathcal{H}_∞ non esistono algoritmi lineari α -ottimi per nessun valore di α comunque grande.
- Non esistono algoritmi lineari che siano robustamente convergenti, a causa sostanzialmente della non limitatezza dell'operatore di proiezione utilizzato.
- Per garantire almeno la convergenza, un algoritmo lineare deve essere “tuned”, cioè deve tener conto in modo esplicito dell'informazione a priori disponibile

Algoritmo lineare “tuned” nel caso di misure in frequenza, assumendo che

- il sistema $S^o \in K_{\rho, M}^{(1)}$, con $\rho > 1$ ed $M \geq 0$
- l'errore $e^N \in \mathcal{B}_{e, \infty}$, con $\varepsilon \geq 0$ ed $A = W_e = I_{N \times N}$

$$\hat{M}(z) = \sum_{k=0}^{n-1} q_k^* z^k$$

dove

$$q^* = [q_k^*] = \left[\frac{c_k(y^N)}{1 + \left(\frac{\varepsilon + M\rho^{-n}}{M} \right)^2 \rho^{2k}} \right]$$

essendo $c_k(y^N)$ i coefficienti della DFT inversa di y^N

$$c_k(y^N) = \frac{1}{N} \sum_{i=0}^{N-1} y_i \left(e^{j2\pi/N} \right)^{-ik}$$

Algoritmi non lineari “two-stage”

- Per ovviare alle limitazioni di convergenza degli algoritmi lineari, si cercano algoritmi non lineari “untuned” che realizzano la seguente procedura a due passi:
 - passo 1: si costruisce un modello preliminare non causale $\hat{M}^{(0)} \in \mathcal{L}_\infty$ con un algoritmo lineare “untuned” che opera una interpolazione bilaterale

$$\hat{M}^{(0)}(z) = \sum_{k=-n+1}^{n-1} w_{k,n} c_k(y) z^k$$

dove $\{w_{k,n}\}$ è una sequenza di pesi detta *window*

- passo 2: si sceglie il modello identificato come la miglior approssimazione in $\mathcal{H}_\infty(\mathbb{D})$ di $\hat{M}^{(0)}$ (problema di approssimazione non lineare di Nehari)

$$\hat{M}(z) = \arg \min_{M \in \mathcal{H}_\infty(\mathbb{D})} \|\hat{M}^{(0)} - M\|_\infty$$

Gli algoritmi two-stage proposti differiscono nel passo 1:

Algoritmo non lineare “two-stage” #1:

- passo 1: approssimazione mediante spline lineari e troncamento.

Algoritmo non lineare “two-stage” #2:

- passo 1: approssimazione con sequenze di pesi $\{w_{k,n}\}$ simmetriche pari rispetto a k (es. finestre triangolari, sinc, cos, trapezoidali), troncate per $k \geq n$.

Proprietà e limitazioni degli algoritmi non lineari “two-stage”

- Sono robustamente convergenti se la sequenza di pesi $\{w_{k,n}\}$ è indipendente dalle informazioni a priori.
- L'errore di identificazione e l'ordine del modello identificato dipendono dal tipo di window utilizzata.
- Mantengono una struttura ancora relativamente semplice.
- Il loro livello di ottimalità non è noto.
- Il modello identificato può non appartenere all'insieme FSS dei sistemi consistenti con tutte le informazioni sul sistema da identificare.

Algoritmi interpolatori

- Forniscono modelli identificati che appartengono all' FSS :

$$\phi^I (K, \mathcal{B}_e, F_N, y^N) = \hat{M}^I \in FSS$$

- Sono algoritmi non lineari e “tuned”.
- Sono in grado di interpolare in modo approssimato i dati sperimentali, tenendo esplicitamente conto delle informazioni a priori disponibili.
- Realizzano in generale una procedura a due passi:
 - passo 1: validazione delle informazioni a priori
 - passo 2: costruzione di un modello $\hat{M}^I \in FSS$ mediante tecniche di interpolazione non lineare

Problema della validazione:

date le informazioni a priori ed i dati sperimentali, FSS è vuoto oppure no?

Risultato #1 (interpolazione di Nevanlinna-Pick, per misure in frequenza).

Le informazione a priori

- sul sistema: $S^o \in K_{\rho, M}^{(1)}$, con $\rho > 1$ ed $M \geq 0$
- sull'errore: $e^N \in \mathcal{B}_{e, \infty}$, con $\varepsilon \geq 0$ ed $A = W_e = I_{N \times N}$

sono consistenti con il vettore dei dati sperimentali $y^N \in \mathbb{C}^N$ (cioè $FSS \neq \emptyset$)

se e solo se esiste un vettore $\eta^N \in \mathcal{B}_{e, \infty}$ tale che la matrice di Pick:

$$P_N := \left[\frac{1 - (1/M^2) (y_i - \eta_i) (\bar{y}_j - \bar{\eta}_j)}{1 - (1 - \rho^2) z_i \bar{z}_j} \right]$$

sia definita nonnegativa, ossia $P_N \geq 0$.

Tale problema di consistenza è riscrivibile in forma di LMI ed è risolvibile se e solo se esiste un vettore $\eta^N \in \mathcal{B}_{e,\infty}$ tale che:

$$A(\eta^N) := \begin{bmatrix} -Q & -\frac{1}{M}(D_y - D_\eta) \\ -\frac{1}{M}(D_y - D_\eta)^H & -Q^{-1} \end{bmatrix} \leq 0$$

con $D_y := \text{diag}(y_0, y_1, \dots, y_{N-1})$, $D_\eta := \text{diag}(\eta_0, \eta_1, \dots, \eta_{N-1})$ e

$$Q := \left[\frac{1}{1 - (1 - \rho^2) z_i \bar{z}_j} \right].$$

[Chen-Nett-Fan, *TAC* 1995]

Algoritmo interpolatorio per dati in frequenza:

- Passo 1: si trova la soluzione $\eta^N = \eta^*$ del problema di consistenza costituito dal precedente LMI, tale che

$$P_N^* = \left[\frac{1 - (1/M^2) (y_i - \eta_i^*) (\bar{y}_j - \bar{\eta}_j^*)}{1 - (1 - \rho^2) z_i \bar{z}_j} \right] \geq 0$$

- Passo 2: applicando l'algoritmo di Pick, si costruisce una funzione che interpoli in modo approssimato $\tilde{y}^N := y^N + \eta^N$ e la si usa come modello identificato.

L'algoritmo interpolatorio è molto oneroso da un punto di vista computazionale qualora il numero di dati sia elevato.

La matrice di Pick può facilmente risultare mal condizionata.

Risultato #2 (interpolazione di Carathéodory-Fejér, per misure nel tempo).

Le informazione a priori

- sul sistema: $S^o \in K_{\rho, M}^{(1)}$, con $\rho > 1$ ed $M \geq 0$
- sull'errore: $e^N \in \mathcal{B}_{e, \infty}$, con $\varepsilon \geq 0$ ed $A = W_e = I_{N \times N}$

sono consistenti con il vettore dei dati sperimentali $y^N \in \mathbb{R}^N$ (cioè $FSS \neq \emptyset$)
se e solo se esiste un vettore $\eta^N \in \mathcal{B}_{e, \infty}$ tale che:

$$(T_y - T_\eta)^T D_\rho^2 (T_y - T_\eta) \leq M^2 T_u^T D_\rho^2 T_u$$

con $D_\rho := \text{diag}(1, \rho, \dots, \rho^{N-1})$ e T_u, T_y e T_η pari alle matrici di Toeplitz triangolari inferiori associate rispettivamente ad u^N, y^N ed η^N .

Tale problema di consistenza è riscrivibile in forma di LMI ed è risolvibile se e solo se esiste un vettore $\eta^N \in \mathcal{B}_{e,\infty}$ tale che:

$$A(\eta^N) := \begin{bmatrix} -T_u^T D_\rho^2 T_u & -(T_y - T_\eta)^T \\ -(T_y - T_\eta) & -M^2 D_\rho^{-2} \end{bmatrix} \leq 0$$

Avendo a disposizione la soluzione η^* di tale LMI, si può ricavare un modello interpolatorio utilizzando la procedura di Carathéodory-Fejér.

[Chen-Nett, *TAC* 1995]

Algoritmo interpolatorio per dati nel tempo

- Passo 1: si trova la soluzione $\eta^N = \eta^* \in \mathcal{B}_{e,\infty}$ del problema di consistenza costituito dal precedente LMI.

Se $A(\eta^N) > 0$, stop; altrimenti, si calcola la matrice

$$H := \frac{1}{M} D_\rho (T_y - T_{\eta^*}) T_u^{-1} D_\rho^{-1} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{N-1} & h_{N-2} & \cdots & h_0 \end{bmatrix}$$

- Passo 2: si sceglie un intero $n \leq N$ e si prende come modello identificato:

$$\hat{M}(z) := M \sum_{k=0}^{n-1} \frac{1 - \rho^{2(k-n)}}{\rho^k} h_k z^k$$

Proprietà degli algoritmi interpolatori

- Sono convergenti, ma non robustamente.
- Sono 2-ottimi e per tale motivo sono detti “almost optimal”.

Limitazioni fondamentali degli algoritmi interpolatori

- I modelli ottenuti sono di ordine elevato.
- Sono noti solo limiti superiori sull'errore di identificazione: $\|E(\phi^I)\| < B$.
- Malgrado i precedenti algoritmi interpolatori siano 2-ottimali, tali limiti superiori sull'errore di identificazione possono rivelarsi poco stringenti, e in particolare può risultare che $B/r \gg 2$.
- Si possono ricavare modelli ridotti, dei quali però non è noto il livello di ottimalità.

Algoritmo interpolatorio “nearly-optimal” per dati nel tempo e/o in frequenza

Ad ogni sistema appartenente all' FSS corrisponde una funzione di trasferimento ammissibile, rappresentabile sul piano polare (o \mathbb{R}^2) per ogni $\omega \in [0, 2\pi] \Rightarrow$

- **Value set** $V(\omega)$: fissata $\omega \in [0, 2\pi]$,

$$V(\omega) := \left\{ v \in \mathbb{R}^2 : v = \begin{bmatrix} \Re(S(\omega)) \\ \Im(S(\omega)) \end{bmatrix}, S \in FSS \right\} \subset \mathbb{R}^2$$

- Proprietà:

- $V(\omega)$ è un politopo, convesso e limitato $\forall \omega$
- *centro di Chebicheff* $c_2[V(\omega)]$ in norma euclidea:

$$c_2[V(\omega)] = \arg \inf_{s \in \mathbb{R}^2} \sup_{v \in V(\omega)} \|s - v\|_2$$

- *raggio di Chebicheff* $r_2[V(\omega)]$ in norma euclidea:

$$r_2[V(\omega)] = \inf_{s \in \mathbb{R}^2} \sup_{v \in V(\omega)} \|s - v\|_2$$

Problemi:

- è impossibile descrivere esattamente $V(\omega)$
- è irrealizzabile il calcolo esatto di $c_2[V(\omega)]$ e $r_2[V(\omega)]$



si calcolano **approssimazioni convergenti di $V(\omega)$** , basate sulla costruzione di due politopi in \mathbb{R}^2 con m vertici:

- il politopo $\overline{VO}_m^\nu(\omega)$, con vertici $\overline{v}_k(\omega)$, $k = 1, \dots, m$, “contiene” $V(\omega)$
- il politopo $\underline{VI}_m^\nu(\omega)$, con vertici $\underline{v}_k(\omega)$, $k = 1, \dots, m$, è contenuto in $V(\omega)$

tali politopi si ricavano mediante tecniche di programmazione lineare.

Risultato #1: convergenza delle approssimazioni di $V(\omega)$

[Milanese-Taragna, TAC 2002]

Sia $\overline{VO}_m^\nu(\omega) := \{s \in \mathbb{R}^2 : s = \bar{s} + \tilde{s}, \bar{s} \in VO_m^\nu(\omega), \|\tilde{s}\|_2 \leq \delta_\nu\}$, $\delta_\nu := \frac{L\rho^\nu}{1-\rho}$.

Fissati $\nu \in \mathbb{Z}$ ed $m \in \mathbb{Z}$,

$$\underline{VI}_m^\nu(\omega) \subseteq V(\omega) \subseteq \overline{VO}_m^\nu(\omega)$$

$$r_2[\underline{VI}_m^\nu(\omega)] \leq r_2[V(\omega)] \leq r_2[\overline{VO}_m^\nu(\omega)] \leq r_2[\overline{VO}_m^\nu(\omega)] + \delta_\nu$$

$$\lim_{\nu, m \rightarrow \infty} r_2[\underline{VI}_m^\nu(\omega)] = \lim_{\nu, m \rightarrow \infty} \left\{ r_2[\overline{VO}_m^\nu(\omega)] + \delta_\nu \right\} = r_2[V(\omega)]$$

$$\lim_{\nu, m \rightarrow \infty} c_2[\underline{VI}_m^\nu(\omega)] = \lim_{\nu, m \rightarrow \infty} c_2[\overline{VO}_m^\nu(\omega)] = c_2[V(\omega)]$$

Risultato 2: calcolo dell'errore di identificazione $E(\hat{\mathcal{M}})$

[Milanese-Taragna, *TAC* 2002]

Dato un modello $\hat{\mathcal{M}}$, fissati $\nu \in \mathbb{Z}$ ed $m \in \mathbb{Z}$:

$$\underline{E}_m^\nu(\hat{\mathcal{M}}) \leq E(\hat{\mathcal{M}}) \leq \overline{E}_m^\nu(\hat{\mathcal{M}})$$

$$\lim_{\nu, m \rightarrow \infty} \overline{E}_m^\nu(\hat{\mathcal{M}}) = \lim_{\nu, m \rightarrow \infty} \underline{E}_m^\nu(\hat{\mathcal{M}}) = E(\hat{\mathcal{M}})$$

$$\underline{E}_m^\nu(\hat{\mathcal{M}}) = \sup_{0 \leq \omega \leq 2\pi} \max_{k=1, \dots, m} \left\| v_k(\underline{VI}_m^\nu(\omega)) - \hat{\mathcal{M}}(\omega) \right\|_2$$

$$\overline{E}_m^\nu(\hat{\mathcal{M}}) = \sup_{0 \leq \omega \leq 2\pi} \max_{k=1, \dots, m} \left\| v_k(\overline{VO}_m^\nu(\omega)) - \hat{\mathcal{M}}(\omega) \right\|_2 + \delta_\nu$$

Risultato 3: calcolo del livello di ottimalità $\alpha(\hat{\mathcal{M}})$

[Milanese-Taragna, *TAC* 2002]

Dato un model set $\hat{\mathcal{M}}$, fissati $\nu \in \mathbb{Z}$ ed $m \in \mathbb{Z}$:

$$\underline{\alpha}_m^\nu(\hat{\mathcal{M}}) \leq \alpha(\hat{\mathcal{M}}) \leq \bar{\alpha}_m^\nu(\hat{\mathcal{M}})$$

$$\lim_{\nu, m \rightarrow \infty} \bar{\alpha}_m^\nu(\hat{\mathcal{M}}) = \lim_{\nu, m \rightarrow \infty} \underline{\alpha}_m^\nu(\hat{\mathcal{M}}) = \alpha(\hat{\mathcal{M}}).$$

$$\underline{\alpha}_m^\nu(\hat{\mathcal{M}}) = \max \left\{ 1, \underline{E}_m^\nu(\hat{\mathcal{M}}) / \left[\sup_{0 \leq \omega \leq 2\pi} r_2[\overline{VO}_m^\nu(\omega)] + \delta_\nu \right] \right\}$$

$$\bar{\alpha}_m^\nu(\hat{\mathcal{M}}) = \overline{E}_m^\nu(\hat{\mathcal{M}}) / \sup_{0 \leq \omega \leq 2\pi} r_2[\underline{VI}_m^\nu(\omega)].$$

Algoritmo interpolatorio “nearly optimal”

Conoscendo i centri ed i raggi dei value set $V(\omega)$ su di un insieme di frequenze, si calcola un model set “nearly optimal” \hat{M}^{no} che approssima il model set centrale \hat{M}^c

- il modello nominale \hat{M}^{no} è soluzione di un problema di programmazione lineare
- l'algoritmo $\phi^{no}(y^N) = \hat{M}^{no}$ è $\sqrt{2}$ -ottimo
- il modello nominale ottenuto è però tipicamente di ordine elevato



model set aventi modelli nominali di ordine ridotto si possono ricavare mediante riduzione d'ordine del model set “nearly optimal”

Model set di ordine ridotto

Model set aventi modelli nominali di ordine ridotto si possono ricavare mediante riduzione d'ordine del model set “nearly optimal”



riduzione d'ordine di model set con inclusione minima e calcolo dell' α -ottimalità



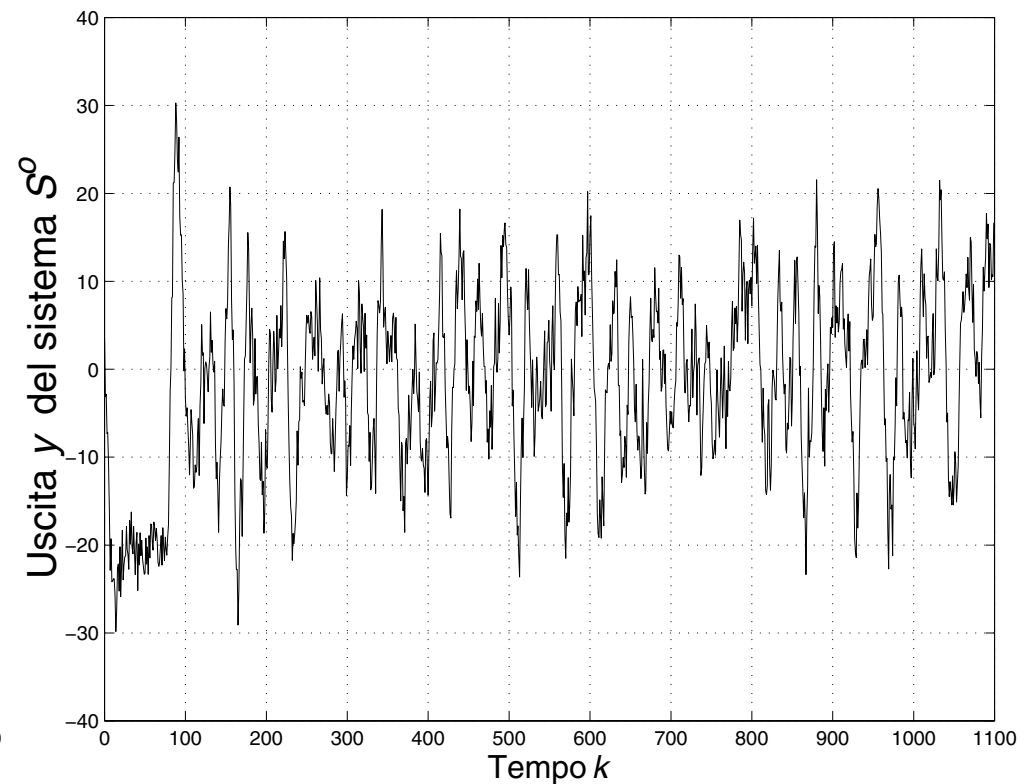
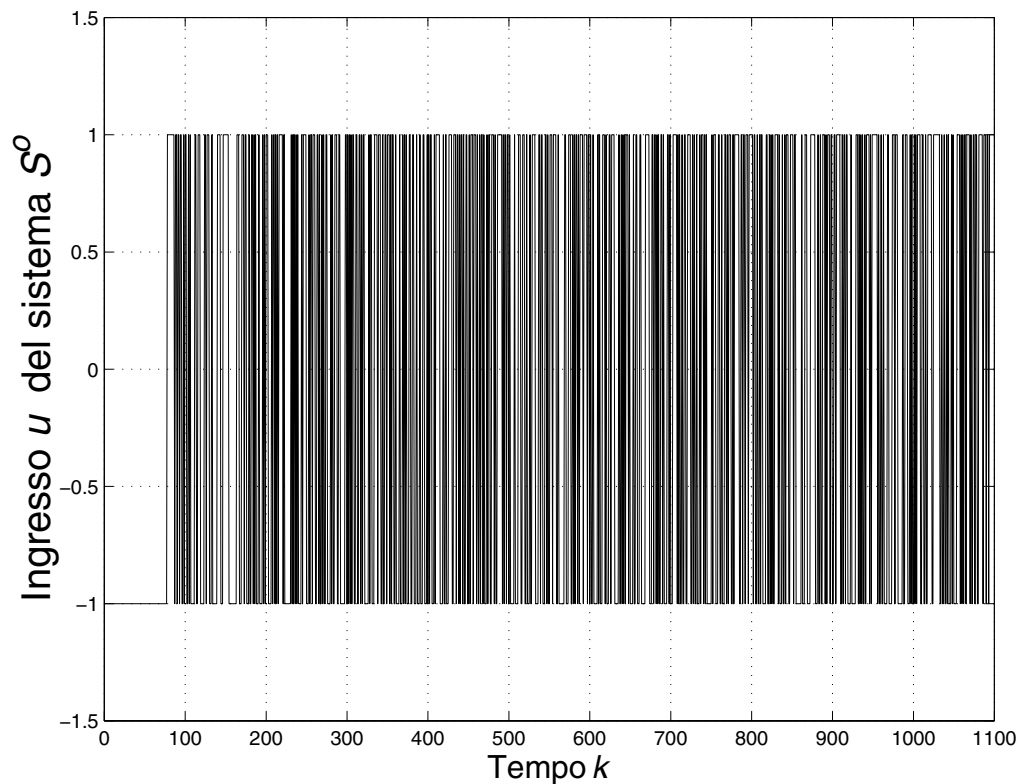
scelta dell'ordine del model set come bilanciamento fra il livello di ottimalità α e la complessità del model set

Esempio numerico

- $N = 1100$ campioni della risposta del sistema

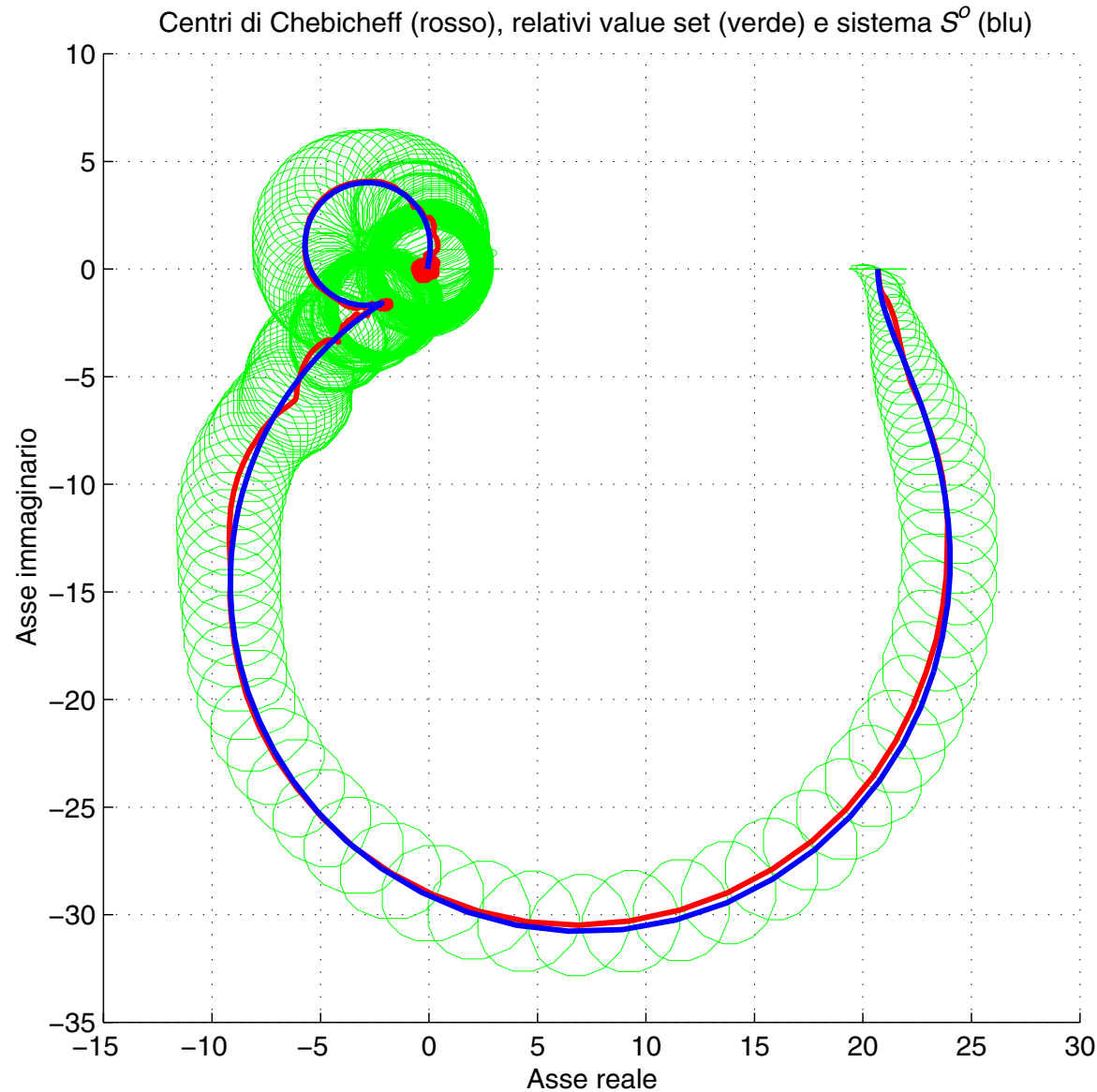
$$S^o(z) = \frac{z + 0.5z^2}{1 - 2.2z + 2.42z^2 - 1.87z^3 + 0.7225z^4}$$

ad un ingresso di tipo PRBS di ampiezza unitaria

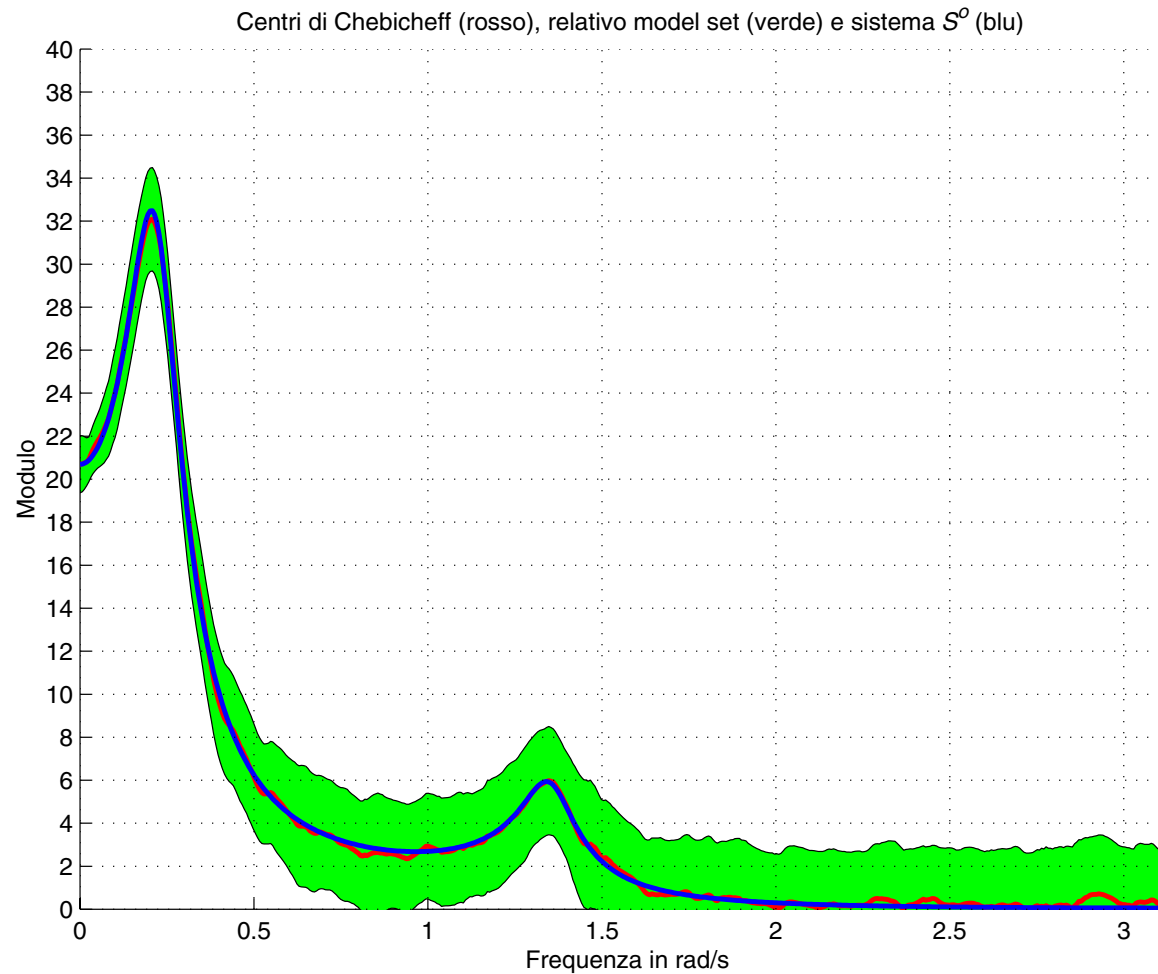


- Informazione “a priori” sul sistema: $S^o \in K_{\rho, M}^{(2)}$, con: $M = 6$, $\rho = 1/0.93$
- Informazione “a priori” sul rumore: $e^N \in \mathcal{B}_{e, \infty}$, con $\varepsilon = 4$, $A = W_e = I_{N \times N}$
(M, ρ, ε) validati secondo [Milanese-Taragna: *SYSID*'00]

- Value set $V(\omega)$: calcolati in 500 valori di ω equispaziati in $[0, \pi]$, con $\nu = 150$, $m = 16$

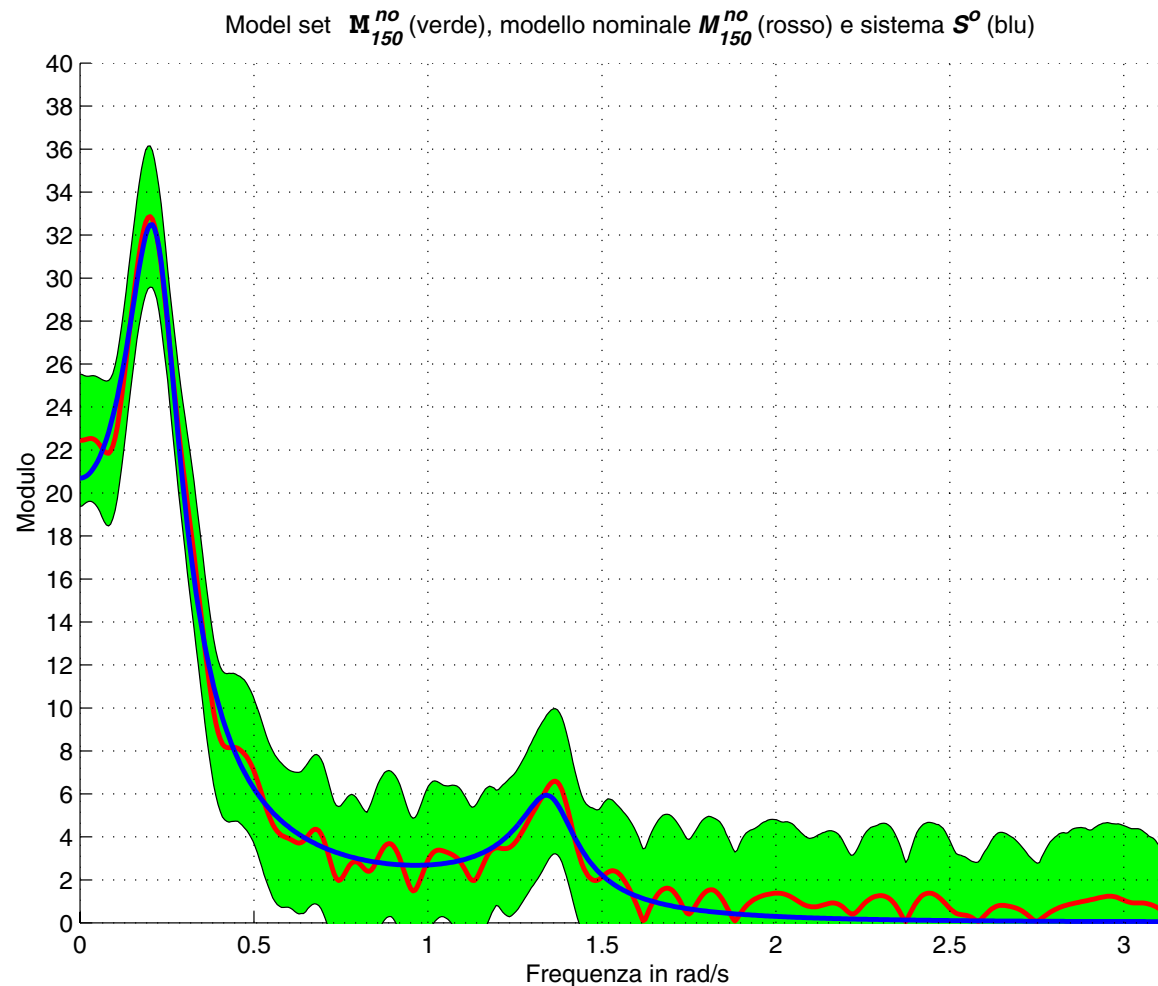


- Model set “centrale” ottenuto considerando:
 - come modello nominale, i centri di Chebicheff dei value set
 - come perturbazione, i raggi di Chebicheff dei value set

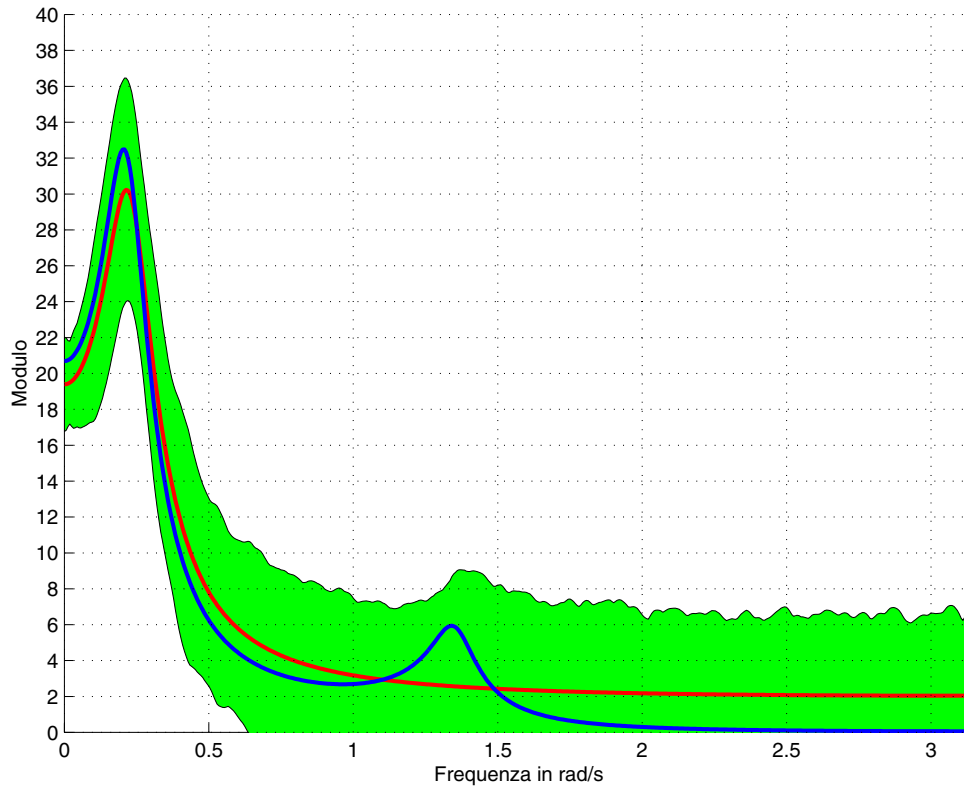
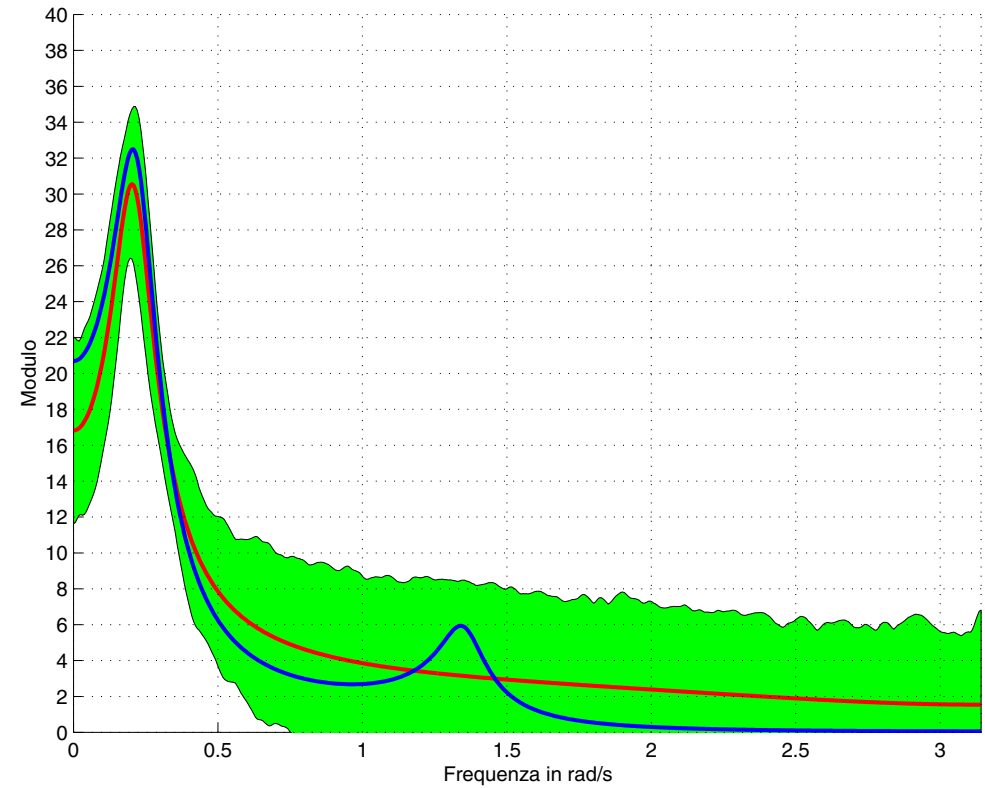


- Modelli nominali e corrispondenti model set utilizzati:

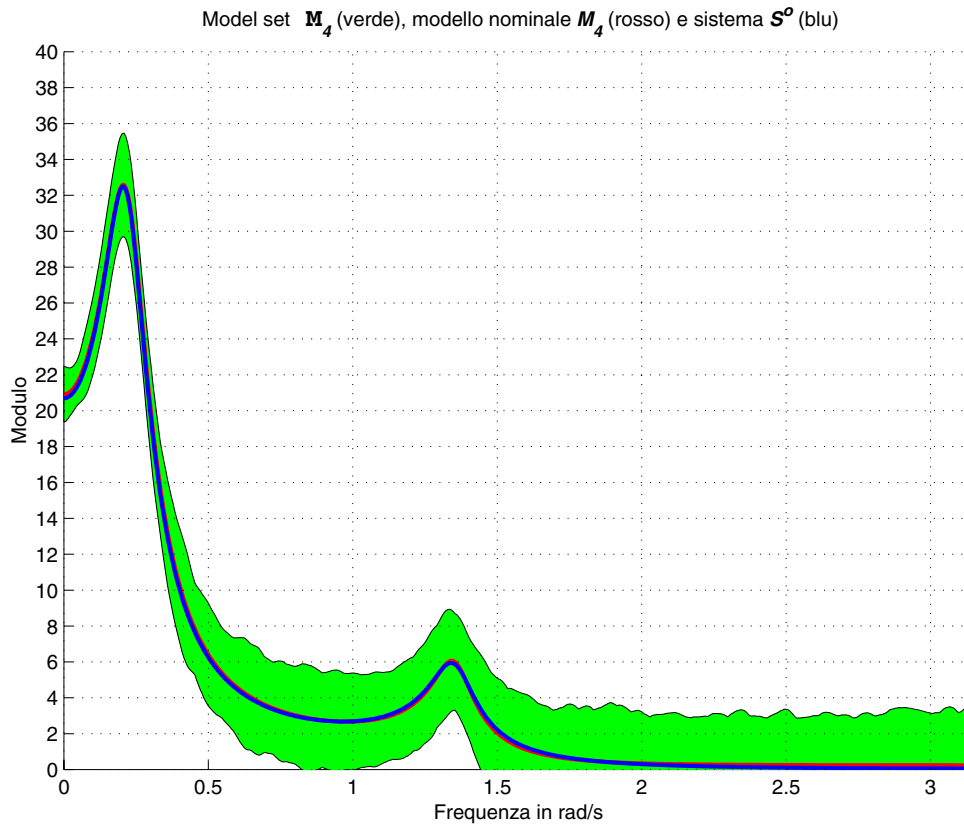
- 1) FIR \hat{M}_{150}^{no} di ordine 150 ottenuto mediante l'algoritmo "nearly optimal" a partire dai raggi e dai centri di Chebicheff dei value set $V(\omega)$ calcolati su 500 valori di ω equispaziati in $[0, \pi]$



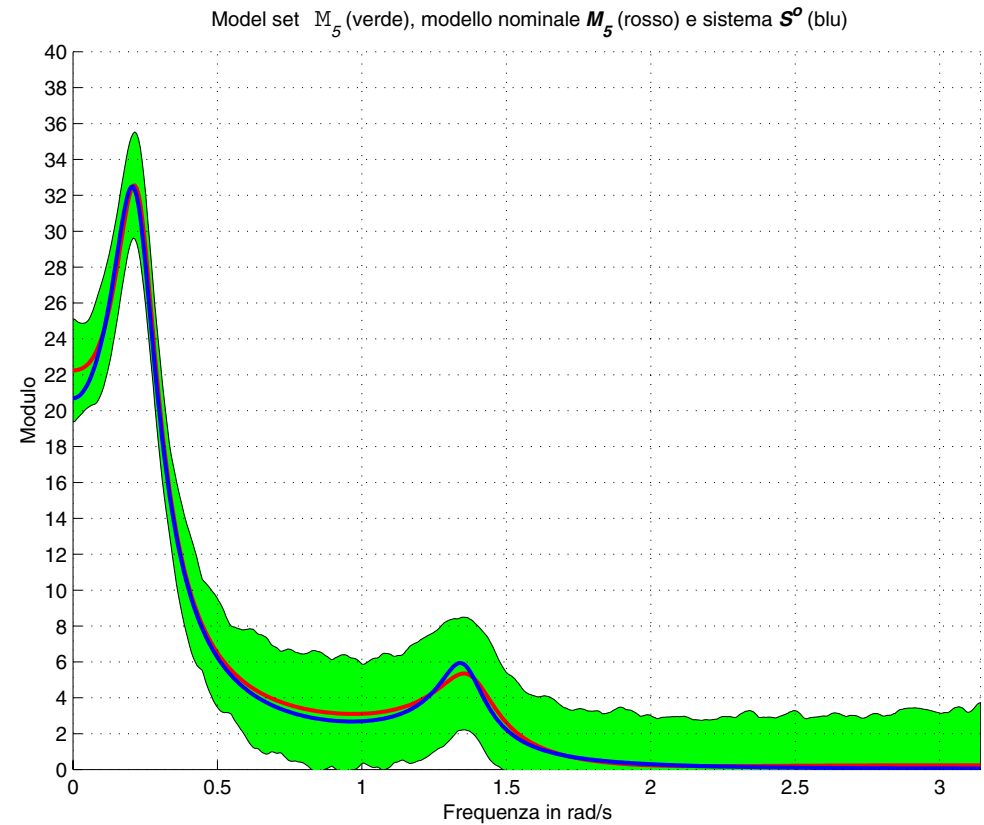
2) approssimazioni di \hat{M}_{150}^{no} di ordine $2 \div 5$ ($\hat{M}_2 \div \hat{M}_5$) ottenute con i metodi di riduzione del modello disponibili in MATLAB (approssimazioni ottime della norma di Hankel, troncamento bilanciato)

Model set \mathbf{M}_2 (verde), modello nominale M_2 (rosso) e sistema S^o (blu)model set \hat{M}_2 Model set \mathbf{M}_3 (verde), modello nominale M_3 (rosso) e sistema S^o (blu)model set \hat{M}_3

2) approssimazioni di \hat{M}_{150}^{no} di ordine $2 \div 5$ ($\hat{M}_2 \div \hat{M}_5$) ottenute con i metodi di riduzione del modello disponibili in MATLAB (approssimazioni ottime della norma di Hankel, troncamento bilanciato)



model set \hat{M}_4



model set \hat{M}_5

Errori di identificazione e
livelli di ottimalità dei model set identificati

$\hat{\mathcal{M}}$	$\underline{E}(\hat{\mathcal{M}})$	$\overline{E}(\hat{\mathcal{M}})$	$\underline{\alpha}_m^\nu(\hat{\mathcal{M}})$	$\overline{\alpha}_m^\nu(\hat{\mathcal{M}})$
$\hat{\mathcal{M}}_{150}^{no}$	3.57	3.61	1.00	1.04
$\hat{\mathcal{M}}_2$	6.49	6.52	1.77	1.87
$\hat{\mathcal{M}}_3$	5.31	5.34	1.45	1.53
$\hat{\mathcal{M}}_4$	3.48	3.49	1.00	1.01
$\hat{\mathcal{M}}_5$	3.48	3.49	1.00	1.01

Limiti sull'errore $E(\hat{\mathcal{M}})$ proposti in [Chen-Nett, *TAC* 95]:

$$\begin{cases} E(\hat{\mathcal{M}}) < 167.43 & \text{(Th. 4.2)} \\ E(\hat{\mathcal{M}}) < 16.99 & \text{(Th. 4.3)} \end{cases}$$

Riferimenti bibliografici

- J. Chen e G. Gu, *Control-Oriented System Identification: An H_∞ Approach*. New York: John Wiley & Sons, Inc., 2000.
- J. Chen, C. N. Nett e M. K. H. Fan, “Worst case system identification in H_∞ : validation of *a priori* information, essentially optimal algorithms, and error bounds,” *IEEE Transactions on Automatic Control*, vol. AC-40, no. 7, pp. 1260–1265, 1995.
- A. Garulli, A. Tesi e A. Vicino, eds., *Robustness in Identification and Control*, vol. 245 of *Lecture Notes in Control and Information Sciences*. Godalming, UK: Springer-Verlag, 1999.
- L. Giarré e M. Milanese, “SM identification of approximating models for H_∞ robust control,” *International Journal of Robust and Nonlinear Control*, vol. 9, pp. 319–332, 1999.
- L. Giarré, M. Milanese e M. Taragna, “ H_∞ identification and model quality evaluation,” *IEEE Transactions on Automatic Control*, vol. AC-42, no. 2, pp. 188–199, 1997.
- M. Glaum, L. Lin e G. Zames, “Optimal H_∞ approximation by systems of prescribed order using frequency response data,” in *Proc. of the 35th IEEE Conference on Decision and Control*, (Kobe, Japan), pp. 2318–2321, 1996.
- G. Gu e P. P. Khargonekar, “A class of algorithms for identification in H_∞ ,” *Automatica*, vol. 28, no. 2, pp. 299–312, 1992.
- G. Gu, D. Xiong e K. Zhou, “Identification in H_∞ using Pick’s interpolation,” *Systems & Control Letters*, vol. 20, pp. 263–272, 1993.

A. J. Helmicki, C. A. Jacobson e C. N. Nett, “Control oriented system identification: a worst-case/deterministic approach in H_∞ ,” *IEEE Transactions on Automatic Control*, vol. AC-36, no. 10, pp. 1163–1176, 1991.

M. Milanese, J. Norton, H. Piet-Lahanier e É. Walter, eds., *Bounding Approaches to System Identification*. New York: Plenum Press, 1996.

M. Milanese, C. Novara e M. Taragna, ““Fast” Set Membership H_∞ identification from frequency-domain data,” in *Proc. of European Control Conference ECC 2001*, (Porto, Portugal), pp. 1698–1703, 2001.

M. Milanese e M. Taragna, “Optimality, approximation, and complexity in Set Membership H_∞ identification,” *IEEE Transactions on Automatic Control*, vol. AC-47, no. 10, pp. 1682-1690, 2002.

M. Milanese e M. Taragna, “Nearly optimal model sets in H_∞ identification,” in *Proc. of the European Control Conference 2001*, (Porto, Portugal), pp. 1704–1709, 2001.

M. Milanese e A. Vicino, “Optimal estimation theory for dynamic systems with set membership uncertainty: an overview,” *Automatica*, vol. 27, no. 6, pp. 997–1009, 1991.

B. Ninness e G. C. Goodwin, “Estimation of model quality,” *Automatica*, vol. 31, no. 12, pp. 1771–1797, 1995.

J. R. Partington, *Interpolation, Identification, and Sampling*, vol. 17 of *London Mathematical Society Monographs New Series*. New York: Clarendon Press - Oxford, 1997.

T. Zhou e H. Kimura, “Structure of model uncertainty for a weakly corrupted plant,” *IEEE Transactions on Automatic Control*, vol. AC-40, no. 4, pp. 639–655, 1995.