

**Università degli studi di Roma
“La Sapienza”**

Facoltà di Ingegneria
Corso di Laurea in Ingegneria delle Telecomunicazioni
Dipartimento Info-Com

Tesi in Teoria dei Segnali

**Rivelazione e stima parametrica di segnali
PPS: applicazioni alla demodulazione di
segnali CPM ed al telerilevamento**

RELATORE

Chiar.mo Prof. Ing.
Sergio Barbarossa

CORRELATORE

Ing. Anna Scaglione

LAUREANDO

Roberto Mameli
matr. 09082267

Anno Accademico 1996/97

Indice

Introduzione	iv
---------------------	----

Parte I - Segnali a fase polinomiale (stima dei parametri e rivelazione)

1. Stima dei parametri di un segnale a fase polinomiale	2
1.1 Definizioni preliminari.	2
1.2 Product High-Order Ambiguity Function (PHAF).	6
1.3 Algoritmo di stima basato sulla PHAF.	11
2. Analisi statistica della HAF	14
2.1 Introduzione.	14
2.2 Caratteristiche della ml-HAF in presenza di solo rumore.	14
2.3 Studio della ml-HAF in presenza di segnale.	20

3. Rivelazione di segnali PPS immersi in AWGN	27
3.1 Introduzione.	27
3.2 Probabilità di falso allarme ottenibile con impiego di una soglia adattiva.	29
3.3 Analisi delle prestazioni ($L=1$).	39
3.4 Estensione al caso $L \geq 2$.	47

Parte II - Applicazione dei segnali a fase polinomiale (modulazione numerica PMSK)

4. Modulazione numerica PMSK	54
4.1 Introduzione	54
4.2 Richiami sulle modulazioni numeriche con memoria: CPM e PMSK.	54
4.3 Demodulazione di segnali PMSK.	60
4.4 Algoritmo di Viterbi per la combinazione delle stime su intervalli consecutivi.	65
4.5 Prestazioni.	73

Appendici

Appendice A1- Determinazione della varianza della ml-HAF in assenza di segnale.	83
---	----

Appendice A2- Indipendenza statistica fra ml-HAF ottenute con set di lags diversi.	85
--	----

Appendice A3- Determinazione della varianza della ml-HAF in presenza di segnale.	88
Appendice A4- Determinazione della P_{FA} ottenuta tramite autogate nel caso di varianza non costante.	97
Appendice A5- Dipendenza della P_{FA} dalla frequenza.	104
Bibliografia	107

Parte I

Segnali a fase polinomiale

(stima dei parametri e rivelazione)

1. Stima dei parametri di un segnale a fase polinomiale

1.1 Definizioni preliminari.

Con il termine di segnali a fase polinomiale (Polynomial Phase Signals - PPS) vengono comunemente indicati tutti i segnali complessi del tipo:

$$(1.1.1) \quad s(n) = A \cdot e^{j \cdot 2\pi \cdot \sum_{m=0}^M a_m \cdot n^m} \quad n = 0, 1, \dots, N-1$$

Tali segnali sono caratterizzati da un'ampiezza costante (A) e da una fase continua, costituita da un polinomio di grado opportuno (M); M viene detto ordine del segnale a fase polinomiale, mentre gli ($M+1$) parametri a_0, a_1, \dots, a_M vengono indicati con il nome di coefficienti del segnale. Un segnale del tipo (1.1.1) viene detto mono-componente, e può essere considerato come caso particolare di segnale PPS multi-componente, definito come segue:

$$(1.1.2) \quad s(n) = \sum_{k=1}^K A_k \cdot e^{j \cdot 2\pi \cdot \sum_{m=0}^M a_{k,m} \cdot n^m}$$

L'importanza del modello a fase polinomiale discende dall'osservazione che qualunque segnale avente inviluppo costante e fase continua in un intervallo di durata limitata può essere approssimativamente considerato come PPS: è infatti sufficiente espandere in serie la fase di tale segnale e troncare lo sviluppo ad un ordine M opportuno (scelto in base all'accuratezza desiderata). Segnali del tipo predetto si presentano normalmente in una notevole varietà di applicazioni, sia nel campo delle telecomunicazioni (ad esempio nelle modulazioni numeriche CPM) che in quello del telerilevamento (echi radar nei SAR). Inoltre in alcuni casi il comportamento polinomiale della fase è da ritenersi esatto (e non semplicemente approssimato come negli esempi sopra citati): ciò accade ad esempio quando si considerano modulazioni numeriche del tipo PMSK (Polynomial Minimum

Shift Keying), nelle quali il segnale ricevuto in ciascun intervallo di simbolo è esattamente del tipo (1.1.1) (si veda a tal proposito [7]).

Si supponga ora assegnata una sequenza $s(n)$, non necessariamente del tipo a fase polinomiale. Si definisce multi-lag High Order Instantaneous Moment (ml-HIM) di ordine M di $s(n)$ la nuova sequenza così ottenuta:

$$(1.1.3) \quad s_M(n; \underline{\tau}_{M-1}) = \begin{cases} s(n) & M = 1 \\ s_{M-1}(n + \tau_{M-1}; \underline{\tau}_{M-2}) \cdot s_{M-1}^*(n - \tau_{M-1}; \underline{\tau}_{M-2}) & M \geq 2 \end{cases}$$

Nella precedente espressione $\underline{\tau}_{M-1} = [\tau_1 \tau_2 \dots \tau_{M-1}]$ rappresenta un vettore costituito da $(M-1)$ lags τ_i interi positivi arbitrariamente scelti. La definizione (1.1.3) è ricorsiva, per cui consente il calcolo dell'ml-HIM di ordine M a partire da quello di ordine $M-1$; si noti che l'ml-HIM di ordine 1 coincide con la sequenza $s(n)$, mentre quello di ordine M è in generale costituito dal prodotto di 2^{M-1} repliche di $s(n)$, opportunamente traslate e coniugate.

Esempio:

L'ml-HIM di ordine 2 di una sequenza $s(n)$ è dato da:

$$s_2(n; \tau) = s(n + \tau) \cdot s^*(n - \tau)$$

mentre quello di ordine 3 è definito come segue:

$$s_3(n; \tau_1, \tau_2) = s(n + \tau_1 + \tau_2) \cdot s(n - \tau_1 - \tau_2) \cdot s^*(n + \tau_1 - \tau_2) \cdot s^*(n - \tau_1 + \tau_2)$$

La proprietà fondamentale che giustifica l'introduzione della definizione (1.1.3) e che è stata illustrata in diversi lavori precedenti (ad esempio in [2], [3], [4] e [5]) è la seguente: l'ml-HIM di ordine M di un segnale a fase polinomiale del tipo (1.1.1) è un esponenziale complesso di frequenza direttamente proporzionale al coefficiente di grado massimo a_M :

$$(1.1.4) \quad s_M(n; \underline{\tau}_{M-1}) = A^{2^{M-1}} \cdot e^{j \cdot 2\pi \cdot P_M \cdot [a_{M-1} + a_M \cdot n]}$$

dove:

$$(1.1.5) \quad P_M = 2^{M-1} \cdot M! \cdot \prod_{i=1}^{M-1} \tau_i$$

Sulla base di tale proprietà in [2] e [3] è stato proposto un metodo per la stima dei coefficienti di un PPS molto efficiente dal punto di vista computazionale. Si supponga infatti noto l'ordine M del segnale (1.1.1); scelto opportunamente un set di lags è quindi possibile determinarne il coefficiente di grado massimo a_M stimando la frequenza del corrispondente ml-HIM di ordine M . Ciò può essere ottenuto ad esempio individuando la posizione del massimo del modulo della trasformata di Fourier del suddetto ml-HIM, che viene comunemente indicata con la notazione $S_M(f; \underline{\tau}_{M-1})$ e prende il nome di multi-lag High Order Ambiguity Function (ml-HAF) di ordine M della sequenza $s(n)$:

$$(1.1.6) \quad S_M(f; \underline{\tau}_{M-1}) = \sum_{n=\sum \tau_i}^{N-\sum \tau_i-1} s_M(n; \underline{\tau}_{M-1}) \cdot e^{-j \cdot 2\pi \cdot f \cdot n} =$$

$$= A \cdot 2^{M-1} \cdot e^{j\theta} \cdot \frac{\sin \left[\pi \cdot (f - f_0) \cdot \left(N - 2 \cdot \sum_{i=1}^{M-1} \tau_i \right) \right]}{\sin \left[\pi \cdot (f - f_0) \right]}$$

dove:

$$(1.1.7) \quad \theta = 2\pi \cdot P_M \cdot a_{M-1} - \pi \cdot (f - f_0) \cdot (N-1)$$

$$(1.1.8) \quad f_0 = P_M \cdot a_M$$

(da notare che la sommatoria viene estesa esclusivamente agli elementi non nulli dell' ml-HIM $s_M(n; \underline{\tau}_{M-1})$). Dopo aver stimato il coefficiente di grado massimo \tilde{a}_M ($\tilde{a}_M = f_0 / P_M$) se ne può effettuare la compensazione:

$$(1.1.9) \quad s(n) \leftarrow s(n) \cdot e^{-j \cdot 2\pi \cdot \tilde{a}_M \cdot n^M}$$

Si riduce in tal modo l'ordine del segnale $s(n)$ di un'unità, rendendo possibile la stima dei restanti coefficienti del PPS attraverso l'iterazione del procedimento appena descritto. L'algoritmo appena illustrato risulta molto efficiente dal punto di vista computazionale, ma risente di diversi inconvenienti a causa della non linearità della trasformazione definita dalla (1.1.3). In [5] e [6] ne è stata ad esempio esaminata l'applicazione in presenza di rumore additivo sovrapposto al segnale; in tale sede si è visto che esso continua a funzionare egregiamente in condizioni di elevato rapporto segnale-rumore (SNR), garantendo tra l'altro delle prestazioni molto vicine all'ottimo teorico costituito dal limite inferiore di Rao-Cramer; tuttavia al crescere del livello di rumore si manifesta un tipico effetto "soglia", con un brusco decadimento delle prestazioni originato dalla comparsa di picchi spuri nel calcolo della ml-HAF (originati dai prodotti incrociati fra le varie repliche di segnale e di rumore). Un secondo inconveniente legato alla non linearità è invece dovuto ai prodotti incrociati fra le varie componenti di segnale, che rende inapplicabile il suddetto metodo nel caso multi-componente (anche in presenza di alto SNR); il problema sorge in particolare quando i coefficienti di grado più elevato delle varie componenti coincidono fra di loro: in [5] si è infatti mostrato che ciò comporta la comparsa di sinusoidi spurie nell'espressione dell'ml-HAF che potrebbero essere erroneamente interpretate come componenti di segnale utile. In quel contesto la soluzione ad entrambi gli inconvenienti illustrati è stata ottenuta introducendo una estensione della ml-HAF precedentemente definita, che prende il nome di ml-HAF moltiplicativa (mml-HAF-multiplicative ml-HAF, o più semplicemente PHAF- Product HAF). Essa sarà oggetto del prossimo paragrafo.

1.2 Product High-Order Ambiguity Function (PHAF).

Nel paragrafo precedente è stato illustrato l'algoritmo di stima dei coefficienti di un PPS mono-componente basato sulla ml-HAF definita in (1.1.6). Sono stati successivamente evidenziati i limiti di tale algoritmo, legati essenzialmente alla non linearità della suddetta trasformazione. Nel presente paragrafo viene presentata una generalizzazione della definizione suindicata che rende possibile la risoluzione degli inconvenienti citati.

Si consideri nuovamente assegnata una generica sequenza $s(n)$ di N elementi; supponiamo inoltre dati L diversi set di lags, del tipo:

$$(1.2.1) \quad \underline{\tau}_{M-1}^{(h)} = [\tau_1^{(h)} \tau_2^{(h)} \dots \tau_{M-1}^{(h)}] \quad h = 1, 2, \dots, L$$

La ml-HAF moltiplicativa (mml-HAF o PHAF) associata agli L set di lags (1.2.1) è definita nel modo seguente:

$$(1.2.2) \quad S_{M-1}^L (f; \underline{\tau}_{M-1}^L) = \prod_{h=1}^L S_M (\alpha_h \cdot f; \underline{\tau}_{M-1}^{(h)})$$

dove:

$$(1.2.3) \quad \underline{\tau}_{M-1}^L = [\underline{\tau}_{M-1}^{(1)} \underline{\tau}_{M-1}^{(2)} \dots \underline{\tau}_{M-1}^{(L)}]$$

$$(1.2.4) \quad \alpha_h = \frac{\prod_{k=1}^{M-1} \tau_k^{(h)}}{\prod_{k=1}^{M-1} \tau_k^{(1)}}$$

Pertanto essa risulta semplicemente data dal prodotto delle ml-HAF ottenute utilizzando gli L differenti set di lags (1.2.1), opportunamente scalate in frequenza attraverso l'introduzione dei fattori α_h . Lo scopo di quest'ultima operazione può essere compreso solo dopo aver richiamato il seguente teorema, la cui dimostrazione può essere trovata in [5]:

Teorema:

Dato un segnale PPS multi-componente di ordine M del tipo (1.1.2), il corrispondente ml-HIM di ordine M è ancora un segnale PPS multi-componente. Più esattamente risulta:

$$(1.1.2) \quad s(n) = \sum_{k=1}^K A_k \cdot e^{j2\pi \cdot \sum_{m=0}^M a_{k,m} \cdot n^m}$$

$$(1.2.5) \quad s_M(n; \underline{\tau}_{M-1}) = \sum_{k_1=1}^K \sum_{k_2=1}^K \dots \sum_{k_p=1}^K A_{k_1} \cdot A_{k_2} \cdot \dots \cdot A_{k_p} \cdot e^{j2\pi \cdot \sum_{m=0}^M b_{k_1 \dots k_p; m}^{(M)} \cdot n^m}$$

dove $p = 2^{M-1}$; i coefficienti $b_{k_1 \dots k_p; m}^{(M)}$ possono essere calcolati attraverso le seguenti relazioni ricorsive:

$$(1.2.6) \quad b_{k_1; m}^{(1)} = a_{k_1, m} \quad m = 0, 1 \dots M$$

$$(1.2.7) \quad b_{k_1, k_2; m}^{(2)} = \sum_{h=0}^{M-m} \binom{m+h}{m} \cdot \left[b_{k_1; m+h}^{(1)} - (-1)^h \cdot b_{k_2; m+h}^{(1)} \right] \cdot \tau_1^h$$

.....

$$(1.2.8) \quad b_{k_1, \dots, k_p; m}^{(M)} = \sum_{h=0}^{M-m} \binom{m+h}{m} \cdot \left[b_{k_1, \dots, k_{2^{M-2}}; m+h}^{(M-1)} - (-1)^h \cdot b_{k_{2^{M-2}+1}, \dots, k_{2^{M-1}}; m+h}^{(M-1)} \right] \cdot \tau_{M-1}^h$$

Nella espressione (1.2.5) compaiono vari termini: quelli per i quali risulta $k_1 = k_2 = \dots = k_p$ prendono il nome di autotermini, mentre gli altri vengono definiti crosstermini. Dalle (1.2.6) - (1.2.8) segue che gli autotermini dell'ml-HIM di ordine M di un segnale multicomponent PPS di ordine M sono delle sinusoidi complesse di frequenze proporzionali ai coefficienti di grado massimo del segnale (1.1.2) e date rispettivamente da¹:

$$(1.2.9) \quad f_k = P_M \cdot a_{k; M} \quad k = 1, 2, \dots, K$$

¹ Questo teorema costituisce evidentemente una generalizzazione della proprietà vista nel paragrafo precedente.

dove P_M è dato dalla (1.1.5), per cui f_k risulta proporzionale anche al prodotto dei lags utilizzati. I crosstermini di (1.2.5) sono in generale di ordine superiore al primo; è tuttavia possibile dimostrare² che nel caso in cui le varie componenti di (1.1.2) abbiano i coefficienti di grado più elevato coincidenti (tutti o in parte) allora nell'ml-HIM (1.2.5) compariranno delle sinusoidi spurie (associate cioè a crosstermini sinusoidali); queste ultime avranno tuttavia delle frequenze legate ai coefficienti del PPS ed al prodotto dei lags da un legame che non risulterà più di semplice proporzionalità diretta, a differenza di quanto accadeva per gli autotermini sinusoidali. A causa di questa proprietà l'effetto della scalatura dell'asse delle frequenze è in ultima analisi quello di riallineare i soli picchi associati alle componenti utili di segnale (solo questi ultimi, infatti, cadono in corrispondenza di valori di frequenza direttamente proporzionali al prodotto dei lags). Viceversa i picchi spuri, dovuti sia ad eventuali crosstermini sinusoidali che alla presenza di rumore additivo, non subiscono lo stesso effetto. Il prodotto delle varie ml-HAF opportunamente scalate consente quindi di esaltare i picchi associati alle componenti di segnale utile, mentre quelli spuri vengono pesantemente attenuati rispetto ai primi.

Esempio:

Si prenda in considerazione il seguente segnale multi-component PPS, a cui viene sommato rumore additivo Gaussiano bianco $w(n)$:

$$x(n) = e^{+j \cdot 2\pi \cdot [0.1 + 0.3 \cdot n + 0.001 \cdot n^2]} + e^{+j \cdot 2\pi \cdot [0.25 - 0.12 \cdot n + 0.001 \cdot n^2]} + w(n)$$

con $\sigma_w^2 = 0.32$. Sia poi $N = 256$, $M = 2$, $\tau^{(1)} = 64$ e $\tau^{(2)} = 48$.

Nella figura 1 è riportato il modulo della ml-HAF ($L=1$, $\tau^{(1)} = 64$) di $x(n)$. Si noti la comparsa di 3 picchi principali: di questi il primo a destra è quello utile ($f_0 = 2 \cdot 2! \cdot \tau^{(1)} \cdot 0.001 = 0.256$), mentre gli altri due sono associati alla presenza di crosstermini sinusoidali (dovuti al fatto che le due componenti di $x(n)$ possiedono lo stesso coefficiente di grado massimo).

² Anche questa dimostrazione si trova in [5].

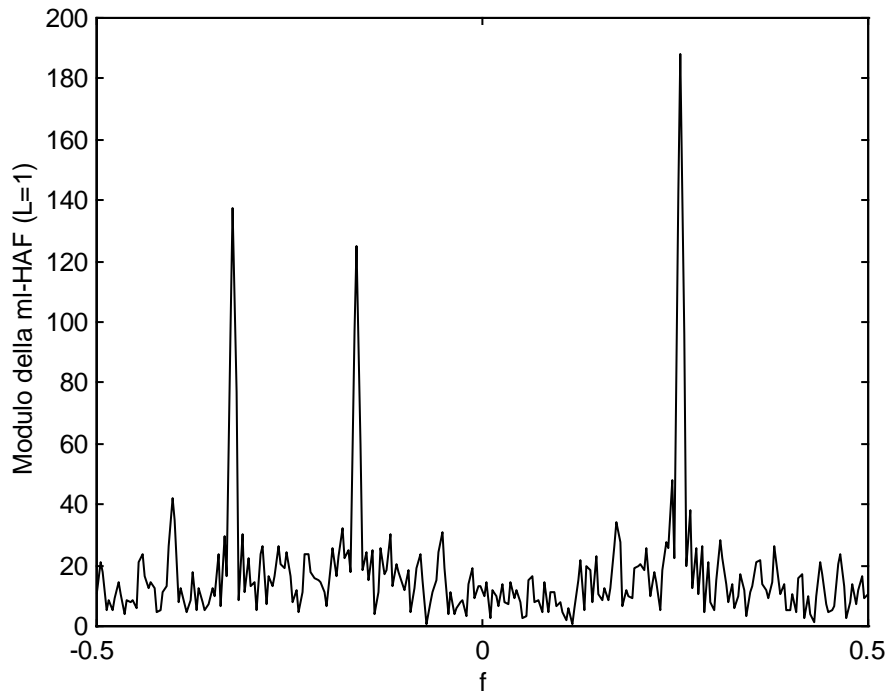


fig. 1

In base alle considerazioni precedentemente effettuate è lecito sperare in un decisivo miglioramento attraverso l'introduzione di un secondo lag ($L = 2$; $\tau^{(2)} = 48$). Il risultato ottenuto è riportato nella figura 2 della pagina seguente. Come si può notare i picchi spuri vengono quasi eliminati, ed anche il rumore di sottofondo subisce una notevole attenuazione.

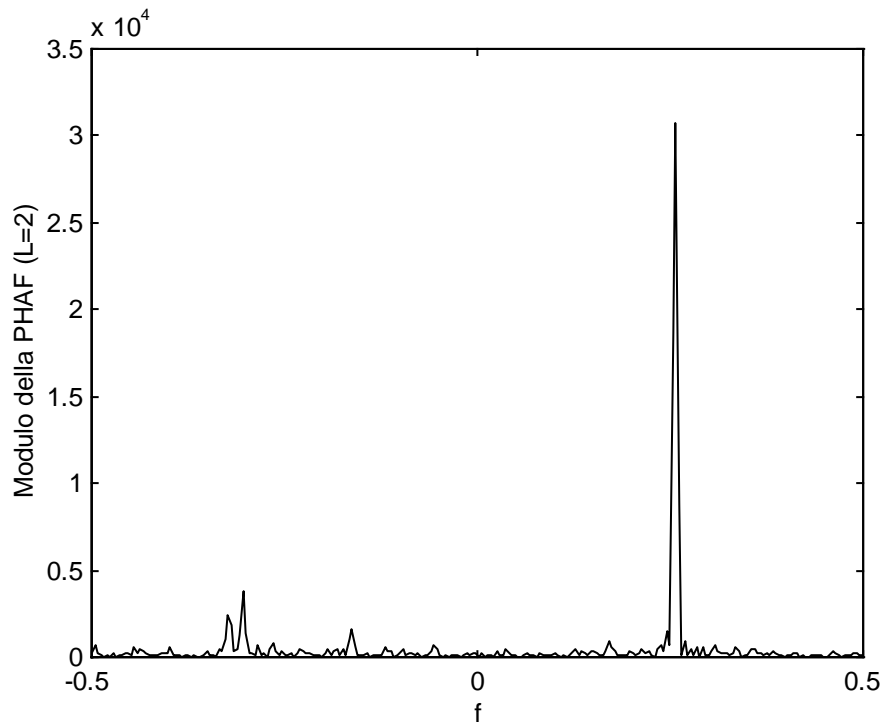


fig. 2

Occorre a questo punto mettere in evidenza un'osservazione: nonostante il motivo principale dell'introduzione della PHAF sia quello di rimuovere le spurie dovute a crosstermini sinusoidali, si ottengono comunque dei vantaggi anche nel caso mono-componente. Infatti l'esecuzione dei prodotti ha come effetto quello di limitare il rumore di sottofondo e ciò si traduce praticamente nell'abbassamento della soglia³ (in altri termini le prestazioni del metodo si avvicinano al limite inferiore di Rao-Cramer a partire da rapporti SNR tanto più bassi quanto maggiore è L). L'aumento di L comporta in realtà anche un lievissimo miglioramento in termini di errore di stima locale (ossia un ulteriore avvicinamento al già citato limite di Rao-Cramer), ma quest'ultimo effetto risulta certamente meno evidente del primo.

³ Si veda a tale proposito [5].

1.3 Algoritmo di stima basato sulla PHAF.

Come già accennato in precedenza l'introduzione della PHAF consente di generalizzare l'algoritmo di stima introdotto nel paragrafo 1.1 al caso di segnali multi-componente, ottenendo contemporaneamente un abbassamento della soglia, che si traduce nella possibilità di utilizzare il suddetto metodo a partire da rapporti segnale-rumore più bassi di quanto non fosse possibile prima. Sia allora assegnato un segnale multi-component PPS immerso in rumore additivo Gaussiano bianco:

$$(1.3.1) \quad x(n) = s(n) + w(n) = \sum_{k=1}^K A_k \cdot e^{j2\pi \cdot \sum_{m=0}^M a_{k,m} \cdot n^m} + w(n)$$

Il primo problema da affrontare consiste nel determinare l'ordine M del PPS, che si suppone incognito; ciò viene fatto attraverso i seguenti passi:

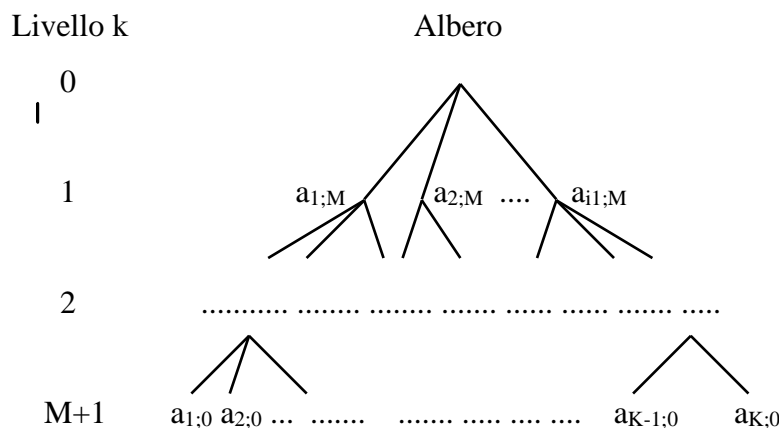
1. si pone $m = M^*$ (scegliendo M^* in modo che risulti sicuramente $M^* \geq M$);
2. si determina il modulo della PHAF ($L \geq 2$) di ordine m di $x(n)$ e si pone su di esso una soglia opportunamente scelta (il problema della scelta della soglia, che costituisce l'oggetto di interesse del presente lavoro, verrà trattato diffusamente nel seguito);
3. si valutano i superamenti della suddetta soglia; se essi avvengono tutti in corrispondenza di $f=0$ si decrementa m di un'unità ($m \leftarrow m-1$) e si torna al punto 2. In caso contrario si pone $M = m$ e si procede oltre.

Da notare che il superamento della soglia per $f = 0$ si presenta sicuramente quando $m > M$ (infatti un PPS di ordine M può sempre essere ritenuto di ordine $m > M$ purché i suoi $(m-M)$ coefficienti di grado massimo vengano considerati tutti nulli). Appare quindi evidente che il precedente algoritmo possa funzionare solo utilizzando la PHAF con $L \geq 2$: infatti i crosstermini sinusoidali a frequenze non nulle, che nel caso $L = 1$ si presenterebbero sicuramente a causa della coincidenza fra i coefficienti di grado massimo del PPS, potrebbero essere erroneamente interpretati come componenti di segnale. Una volta stimato l'ordine M del segnale l'algoritmo procede nel modo seguente:

1. si pone $m = M$ e $y(n) = x(n)$;

2. si determina il modulo della PHAF ($L \geq 2$) di ordine m di $y(n)$ e si pone su di esso una soglia opportunamente scelta;
3. si determinano i superamenti della suddetta soglia e si stimano i valori dei corrispondenti coefficienti di grado m ($a = a_{k,m}$);
4. per ciascuno dei coefficienti stimati al passo precedente si effettua dapprima la compensazione ($y(n) \leftarrow y(n) \cdot e^{-j \cdot 2\pi \cdot a \cdot n}$), quindi si decrementa m ; se $m \geq 1$ si torna (ricorsivamente) al passo 2, altrimenti si stima anche il termine noto e la ricorsione termina .

Il precedente algoritmo può essere compreso più semplicemente associando ad esso un albero opportunamente costruito: in esso la radice rappresenta un nodo fittizio, mentre ciascuno degli altri nodi rappresenta un opportuno coefficiente del PPS; in particolare i nodi di livello 1 (ossia i figli della radice) sono associati ai distinti coefficienti di grado M , quelli di livello 2 ai coefficienti di grado $M-1$, e così via (le foglie corrispondono ai termini noti).



Durante l'esecuzione dell'algoritmo di stima l'albero viene visitato ricorsivamente per livelli (padre \rightarrow 1° figlio da sx. \rightarrow 2° figlio da sx \rightarrow ... \rightarrow ultimo figlio a destra). Si parte inizialmente dalla radice ($m=M$ e $y(n)=x(n)$; vedi passo 1) e si stimano tutti i coefficienti distinti di grado massimo, ciascuno dei quali corrisponde ad un nodo di livello 1. Per ciascuno di essi si passa poi ad esplorare il corrispondente sotto-albero in modo ricorsivo (cfr. punto 4) fino ad arrivare al termine noto, ossia ad una delle foglie. Queste ultime sono in numero pari alle componenti di segnale e ciascuno dei cammini radice-foglia contiene i corrispondenti coefficienti (in ordine decrescente).

Si osservi che il metodo suindicato basa il suo funzionamento sulla fondamentale proprietà della PHAF ($L \geq 2$) di eliminare i picchi spuri; si noti altresì che in esso assume fondamentale importanza il problema della rivelazione, oggetto di studio della prima parte del lavoro. Esso si

traduce in pratica nella imposizione di una soglia opportunamente scelta sul modulo della PHAF ed era stato trascurato nel precedente algoritmo, essendo quest'ultimo applicabile al solo caso mono-componente, per cui una volta calcolata la ml-HAF era sufficiente valutarne il massimo e stimare da questo il corrispondente valore del coefficiente. Nella situazione attuale, invece, si hanno più componenti di segnale (al limite in numero non noto a priori), per cui non è più possibile limitarsi a determinare la posizione del massimo, ma occorre invece individuare tutti i picchi associati ai segnali utili. Inoltre l'imposizione di una soglia può rivelarsi utile anche nel caso mono-componente per bassi rapporti segnale-rumore; essa può infatti impedire che in tale situazione si ottenga una stima completamente errata. Ciò che accade in questo caso è che il segnale non viene neanche rivelato, in quanto completamente mascherato dal rumore, ma in alcune situazioni questa evenienza può essere preferibile a quella di una stima completamente errata.

2. Analisi statistica della HAF

2.1 Introduzione.

Nel capitolo precedente è stato effettuato un breve richiamo sui segnali a fase polinomiale e sul metodo di stima dei rispettivi parametri. Si è quindi fatto cenno al problema della rivelazione di segnali PPS immersi in rumore, mostrando il legame fra tale aspetto e quello della stima. L'analisi del problema della rivelazione di PPS immerso in rumore richiede preventivamente lo studio delle caratteristiche statistiche della ml-HAF. Pertanto in questo capitolo calcoleremo la densità di probabilità ed i principali momenti nei due casi di presenza del solo rumore e di presenza di segnale più rumore, rimandando al successivo lo studio dettagliato del problema della rivelazione. Gli argomenti affrontati in questa sede vanno considerati propedeutici per gli sviluppi successivi.

2.2 Caratteristiche della ml-HAF in presenza di solo rumore.

Si prenda in considerazione una sequenza di campioni di rumore Gaussiano bianco di lunghezza N:

$$(2.2.1) \quad x(n) = w(n) \quad |n| \leq (N-1) / 2$$

con:

$$(2.2.2) \quad E [w(n)] = E [w] = 0$$

$$(2.2.3) \quad E [w^*(n) \cdot w(n+\tau)] = R_w (\tau) = \sigma_w^2 \cdot \delta(\tau)$$

Nel capitolo precedente è stata introdotta una definizione ricorsiva dell'ml-HIM; essa si presta bene ad essere implementata al calcolatore, ma nelle considerazioni di carattere teorico spesso risulta più utile fare riferimento alla definizione alternativa riportata di seguito:

$$(2.2.4) \quad s_M(n; \underline{\tau}_{M-1}) = \prod_{k=1}^{2^{M-1}} s^{*p_M(k)}[n + z_M(k)]$$

dove:

$$(2.2.5) \quad \underline{C}_M = \begin{bmatrix} 1 & 1 & \dots & -1 \\ 1 & 1 & \dots & -1 \\ \dots & \dots & \dots & -1 \\ 1 & -1 & \dots & -1 \end{bmatrix}$$

$$(2.2.6) \quad \underline{z}_M = [z_M(1) \quad z_M(2) \quad \dots \quad z_M(2^{M-1})] = \underline{\tau}_{M-1} \cdot \underline{C}_M$$

$$(2.2.7) \quad p_M(k) = \prod_{h=1}^{M-1} c_{h,k}$$

La matrice \underline{C}_M ha dimensioni $(M-1) \times 2^{M-1}$ e le sue colonne sono costituite da tutte e sole le possibili $(M-1)$ -uple ottenibili combinando i valori $+1$ e -1 ; il vettore riga \underline{z}_M contiene le somme dei lags (ottenute combinando i segni dei lags in tutti i modi possibili); $p_M(k)$ è invece definito come prodotto degli elementi della k -esima colonna di \underline{C}_M ($p_M(k) = \pm 1$). Infine la notazione $s^{*p_M(k)}[n]$ significa $s[n]$ se $p_M(k) = +1$, $s^*[n]$ se $p_M(k) = -1$. Seguendo tali notazioni possiamo esprimere l'ml-HIM e la ml-HAF della sequenza (2.2.1) nel modo seguente:

$$(2.2.8) \quad x_M(n; \underline{\tau}_{M-1}) = \prod_{k=1}^{2^{M-1}} w^{*p_M(k)}[n + z_M(k)]$$

$$(2.2.9) \quad X_M(f; \underline{\tau}_{M-1}) = \sum_n \prod_{k=1}^{2^{M-1}} w^{*p_M(k)}[n + z_M(k)] \cdot e^{-j \cdot 2\pi \cdot f \cdot n}$$

La sommatoria in (2.2.9) è da intendersi estesa agli elementi non nulli di $x_M(n; \underline{\tau}_{M-1})$ (ossia $|n| \leq (N-1)/2 - \sum \tau_i$ nel caso simmetrico; $\sum \tau_i \leq n \leq N-1 - \sum \tau_i$ in quello non simmetrico). Nell'ipotesi in cui N risulti molto grande ($N \gg 1$) la variabile aleatoria (2.2.9) può essere approssimativamente

ritenuta Gaussiana complessa; ciò in virtù del teorema del limite centrale ed indipendentemente dalla distribuzione dell'ml-HIM (2.2.8). E' abbastanza agevole provare che il valore atteso della ml-HAF è sempre nullo (indipendentemente da f e per qualunque valore di M). Infatti:

$$(2.2.10) \quad E [X_M(f; \underline{t}_{M-1})] = \sum_n E \left\{ \prod_{k=1}^{2^{M-1}} w^{*p_M(k)} [n + z_M(k)] \right\} \cdot e^{-j \cdot 2\pi \cdot f \cdot n}$$

e, in conseguenza delle ipotesi fatte su w(n), risulta⁴:

$$(2.2.11) \quad E \left\{ \prod_{k=1}^{2^{M-1}} w^{*p_M(k)} [n + z_M(k)] \right\} = 0$$

Sostituendo la precedente in (2.2.10) si ottiene:

$$(2.2.12) \quad E [X_M(f; \underline{t}_{M-1})] = 0 \quad \forall f$$

Purtroppo il calcolo della varianza di (2.2.9) non risulta altrettanto semplice. Si ha infatti:

$$(2.2.13) \quad \sigma_M^2(f) = E [|X_M(f; \underline{t}_{M-1})|^2] = E [X_M(f; \underline{t}_{M-1}) \cdot X_M^*(f; \underline{t}_{M-1})] = \\ = \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [m-n]} \cdot E \left\{ \prod_{k=1}^{2^{M-1}} w^{*p_M(k)} [n + z_M(k)] \cdot w^{[-p_M(k)]} [m + z_M(k)] \right\}$$

Nella (2.2.13) compare il valore atteso del prodotto di 2^M variabili aleatorie Gaussiane; in base alle note proprietà dei momenti di una Gaussiana il suo calcolo può essere opportunamente ricondotto a

⁴Infatti campioni di rumore distinti sono statisticamente indipendenti ed hanno valore atteso nullo. Pertanto una condizione necessaria (ma non sufficiente) affinché il valore atteso (2.2.11) sia diverso da 0 è che esista una coppia di indici distinti (k,h) per cui $z_M(k) = z_M(h)$ { $k, h = 1, 2, \dots, 2^{M-1}$; $k \neq h$ }; dalle definizioni (2.2.5) e (2.2.6) e dall'uguaglianza $z_M(k) = z_M(h)$ seguirebbe allora l'esistenza di un sottoinsieme di lags a somma nulla, ma ciò risulta impossibile essendo i suddetti lags interi positivi. Di conseguenza tutti i campioni di rumore che compaiono in (2.2.11) sono distinti, ed il valore atteso si fattorizza nel prodotto dei valori attesi (per l'indipendenza statistica); questi ultimi sono tutti nulli (cfr. (2.2.2)), e da ciò segue (2.2.11).

quello dei momenti del secondo ordine⁵, ma ciò nonostante risulta comunque arduo determinarne una espressione in forma chiusa. Per tale motivo risulta più semplice valutare per esteso la (2.2.13) nei due casi M=2 ed M=3 (si rimanda all'appendice A1 per maggiori dettagli); risulta:

$$(2.2.14) \quad \sigma_2^2(f) = \sigma_2^2 = \sigma_w^4 \cdot [N-2 \cdot \tau]$$

$$(2.2.15) \quad \sigma_3^2(f) = \sigma_3^2 = \sigma_w^8 \cdot [N-2 \cdot (\tau_1 + \tau_2)] \cdot [1 + \delta(\tau_1 - \tau_2)]$$

Da notare che in ambedue i casi la varianza risulta indipendente da f; con altra terminologia si può affermare che il processo (2.2.9) è “stazionario” in frequenza.

Un'altra proprietà molto importante che può essere facilmente provata è l'indipendenza statistica fra i campioni di due diverse ml-HAF, calcolate sulla stessa sequenza di rumore, ma con diversi lags. Si ha cioè:

$$(2.2.16) \quad E [X_M^*(f; \underline{\tau}_{M-1}) \cdot X_M(f; \underline{\tau}'_{M-1})] = 0$$

purché si abbia $\underline{\tau}_{M-1} \neq \underline{\tau}'_{M-1}$ (la dimostrazione della suddetta proprietà può essere trovata nell'appendice A2).

Le precedenti considerazioni consentono di valutare la densità di probabilità del modulo quadro della ml-HAF di x(n). $X_M(f; \underline{\tau}_{M-1})$ risulta infatti Gaussiana complessa, con valore atteso nullo e varianza $\sigma_M^2(f) = \sigma_M^2$ data dalla (2.2.13), per cui il suo modulo quadro avrà distribuzione esponenziale del tipo:

$$(2.2.17) \quad p_Q(q) = \lambda \cdot e^{-\lambda \cdot q} \cdot u_1(q)$$

dove:

$$(2.2.18) \quad q = | X_M(f; \underline{\tau}_{M-1}) |^2$$

$$(2.2.19) \quad \lambda = 1/\sigma_M^2$$

⁵ Il valore atteso del prodotto di 2^M variabili Gaussiani è formato dalla somma di $\prod_{h=0}^{2^{M-1}-1} (2^M - 1 - 2 \cdot h)$ termini (3 termini per M=2, 105 per M=3, e via dicendo); ciascuno di essi risulta dato dal prodotto di 2^{M-1} fattori, costituiti dai valori attesi di tutte le possibili coppie di variabili estratte dalle 2^M assegnate.

Anche la densità di probabilità del modulo quadro della PHAF ($L = 2$) può essere determinata in modo molto semplice. Posto infatti:

$$(2.2.18) \quad q = |X_M(f; \underline{x}_{M-1})|^2$$

$$(2.2.20) \quad q' = |X_M(\alpha \cdot f; \underline{x}'_{M-1})|^2$$

risulta:

$$(2.2.17) \quad p_Q(q) = \lambda \cdot e^{-\lambda \cdot q} \cdot u_1(q)$$

$$(2.2.21) \quad p_{Q'}(q') = \lambda' \cdot e^{-\lambda' \cdot q'} \cdot u_1(q')$$

dove:

$$(2.2.19) \quad \lambda = 1/\sigma_M^2$$

$$(2.2.22) \quad \lambda' = 1/\sigma_{M'}^2$$

Indicato con y il modulo quadro della PHAF, che risulta dato dal prodotto dei moduli quadri delle m -HAF componenti, e ricordando l'indipendenza statistica fra queste ultime (cfr. appendice A2) si può scrivere:

$$(2.2.23) \quad p_Y(y) = \int_{-\infty}^{+\infty} \frac{1}{|q|} \cdot p_{Q,Q'}\left(q, \frac{y}{q}\right) \cdot dq =$$

$$= \int_{-\infty}^{+\infty} \frac{1}{|q|} \cdot p_Q(q) \cdot p_{Q'}\left(\frac{y}{q}\right) \cdot dq =$$

$$= \begin{cases} \lambda \cdot \lambda' \cdot \int_0^{+\infty} \frac{1}{q} \cdot e^{-\left(\lambda \cdot q + \frac{\lambda' \cdot y}{q}\right)} dq & y \geq 0 \\ 0 & y < 0 \end{cases}$$

L'integrale precedente si risolve per sostituzione; ponendo infatti:

$$(2.2.24) \quad e^x = \sqrt{\frac{\lambda}{\lambda' \cdot y}} \cdot q$$

ed effettuando alcuni passaggi si ottiene alla fine:

$$\begin{aligned}
 (2.2.25) \quad \lambda \cdot \lambda' \cdot \int_0^{+\infty} \frac{1}{q} \cdot e^{-\left(\lambda \cdot q + \frac{\lambda' \cdot y}{q}\right)} dq &= \lambda \cdot \lambda' \cdot \int_{-\infty}^{+\infty} e^{-2\sqrt{\lambda \cdot \lambda' \cdot y} \cdot \cosh(x)} dx = \\
 &= 2 \cdot \lambda \cdot \lambda' \cdot \int_0^{+\infty} e^{-2\sqrt{\lambda \cdot \lambda' \cdot x} \cdot \cosh(x)} dx = 2 \cdot \lambda \cdot \lambda' \cdot K_0 [2\sqrt{\lambda \cdot \lambda' \cdot y}]
 \end{aligned}$$

dove:

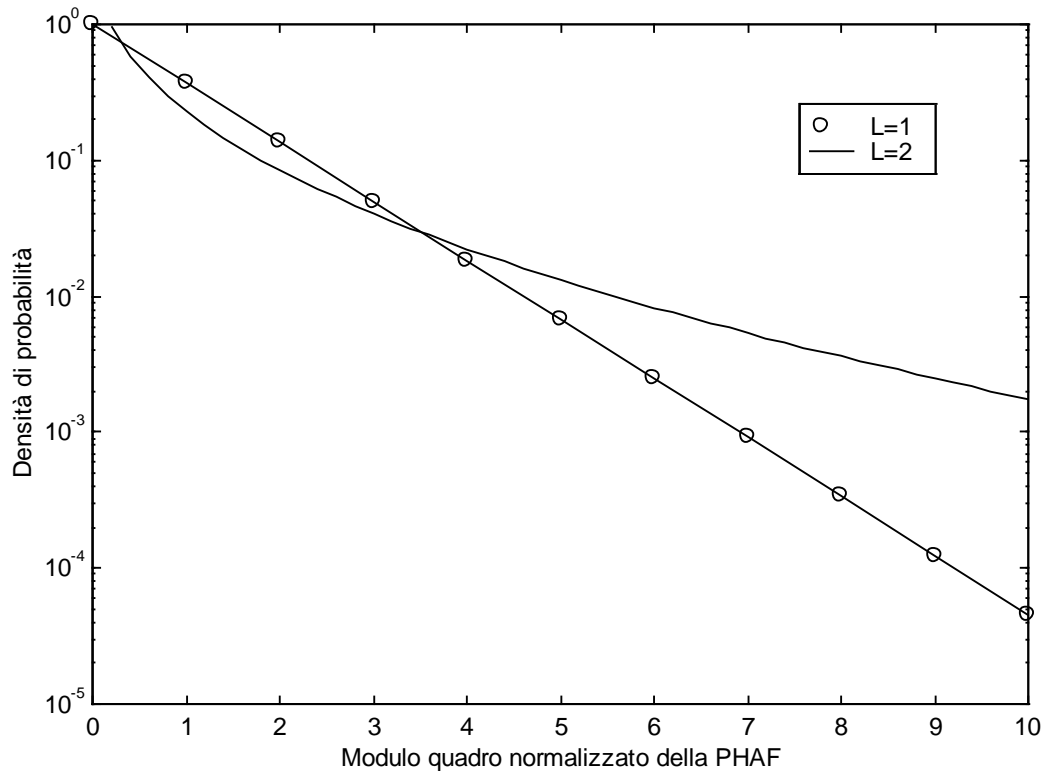
$$(2.2.26) \quad K_0 [\zeta] = \int_0^{+\infty} e^{-\zeta \cdot \cosh(x)} dx \quad (\zeta > 0)$$

rappresenta la funzione modificata di Bessel di ordine 0. Di conseguenza si ha:

$$(2.2.27) \quad p_Y(y) = 2 \cdot \lambda \cdot \lambda' \cdot K_0 [2\sqrt{\lambda \cdot \lambda' \cdot y}] \cdot u_{-1}[y]$$

Il caso $L = 3$ non risulta altrettanto semplice, in quanto la densità di una variabile aleatoria data dal prodotto di 3 variabili esponenziali non ammette una espressione in forma chiusa.

Nella figura seguente sono riportate le densità di probabilità (2.2.17) e (2.2.27), nelle ipotesi $\lambda = \lambda' = 1$ (densità normalizzate rispetto ai valori attesi):



Da notare che la densità (2.2.17) relativa al caso L=1 appare come una retta (si tratta infatti di un'esponenziale in scala semilogaritmica).

2.3 Studio della ml-HAF in presenza di segnale.

Si supponga ora di considerare un segnale a fase polinomiale immerso in rumore additivo Gaussiano bianco:

$$(2.3.1) \quad x(n) = s(n) + w(n) = A \cdot e^{j \cdot 2\pi \cdot \sum_{m=0}^M a_m \cdot n^m} + w(n)$$

Per semplicità si limita l'analisi alle ipotesi di segnale mono-componente e di elevato rapporto segnale-rumore ($SNR = A^2/\sigma_w^2 \gg 1$). Sotto questa condizione l'ml-HIM di ordine M di (2.3.1) può essere espresso come somma di due contributi, dati rispettivamente dall'ml-HIM del segnale utile ($s_M(n; \underline{\tau}_{M-1})$) e da un termine di "perturbazione" ($\delta_{s_M}(n; \underline{\tau}_{M-1})$) legato alla presenza di rumore. Più precisamente risulta:

$$\begin{aligned}
(2.3.2) \quad x_M(n; \underline{\tau}_{M-1}) &= \prod_{k=1}^{2^{M-1}} x^{*p_M(k)} [n + z_M(k)] = \\
&= \prod_{k=1}^{2^{M-1}} \left\{ s^{*p_M(k)} [n + z_M(k)] + w^{*p_M(k)} [n + z_M(k)] \right\} = \\
&= \prod_{k=1}^{2^{M-1}} s^{*p_M(k)} [n + z_M(k)] \cdot \left\{ 1 + \frac{w^{*p_M(k)} [n + z_M(k)]}{s^{*p_M(k)} [n + z_M(k)]} \right\} = \\
&\cong s_M(n; \underline{\tau}_{M-1}) \cdot \left\{ 1 + \sum_{k=1}^{2^{M-1}} \frac{w^{*p_M(k)} [n + z_M(k)]}{s^{*p_M(k)} [n + z_M(k)]} \right\} = \\
&= s_M(n; \underline{\tau}_{M-1}) + \delta s_M(n; \underline{\tau}_{M-1})
\end{aligned}$$

dove:

$$(2.3.3) \quad s_M(n; \underline{\tau}_{M-1}) = A^{2^{M-1}} \cdot e^{j \cdot 2\pi \cdot P_M \cdot [a_{M-1} + a_M \cdot n]}$$

$$(2.3.4) \quad \delta s_M(n; \underline{\tau}_{M-1}) = s_M(n; \underline{\tau}_{M-1}) \cdot \sum_{k=1}^{2^{M-1}} \frac{w^{*p_M(k)} [n + z_M(k)]}{s^{*p_M(k)} [n + z_M(k)]}$$

Nella (2.3.2) sono stati trascurati tutti i termini contenenti il prodotto di due o più componenti di rumore in virtù dell'ipotesi di alto SNR. La ml-HAF di ordine M della (2.3.1) è allora approssimativamente data da:

$$(2.3.5) \quad X_M(f; \underline{\tau}_{M-1}) \cong S_M(f; \underline{\tau}_{M-1}) + \delta S_M(f; \underline{\tau}_{M-1})$$

essendo $S_M(f; \underline{\tau}_{M-1})$ e $\delta S_M(f; \underline{\tau}_{M-1})$ fornite rispettivamente dalle trasformate di Fourier delle (2.3.3) e (2.3.4). Si ha quindi:

$$(2.3.6) \quad S_M(f; \underline{\tau}_{M-1}) = A^{2^{M-1}} \cdot e^{j\theta} \cdot \frac{\sin \left[\pi \cdot (f - f_0) \cdot \left(N - 2 \cdot \sum_{i=1}^{M-1} \tau_i \right) \right]}{\sin \left[\pi \cdot (f - f_0) \right]}$$

$$(2.3.7) \quad \delta S_M(f; \underline{\tau}_{M-1}) = \sum_{n=z_M(1)}^{N-z_M(1)-1} \delta s_M(n; \underline{\tau}_{M-1}) \cdot e^{-j \cdot 2\pi \cdot f \cdot n}$$

$$= \sum_n s_M(n; \underline{\tau}_{M-1}) \cdot e^{-j \cdot 2\pi \cdot f \cdot n} \cdot \sum_{k=1}^{2^{M-1}} \frac{w^{*p_M(k)}[n + z_M(k)]}{s^{*p_M(k)}[n + z_M(k)]}$$

con:

$$(2.3.8) \quad \theta = 2\pi \cdot P_M \cdot a_{M-1} - \pi \cdot (f - f_0) \cdot (N-1)$$

$$(2.3.9) \quad f_0 = P_M \cdot a_M$$

Anche nel caso in esame si suppone $N \gg 1$, per cui si può assumere valido il teorema del limite centrale e $\delta S_M(n; \underline{\tau}_{M-1})$ può essere ritenuto Gaussiano. Estrahendo poi il valore atteso di (2.3.7) si ricava facilmente:

$$(2.3.10) \quad E[\delta S_M(f; \underline{\tau}_{M-1})] = 0$$

La varianza di (2.3.7), per la cui determinazione si rimanda all'appendice A3, è invece data dalla seguente espressione:

$$(2.3.11) \quad E[|\delta S_M(f; \underline{\tau}_{M-1})|^2] = \sigma_M^2(f) = \frac{A^2}{\text{SNR}} \cdot \sum_{k=1}^{2^{M-1}} \sum_{h=1}^{2^{M-1}} \delta[p_M(k) - p_M(h)] \cdot \sum_n \sum_m \delta[n - m + z_M(k) - z_M(h)] \cdot e^{j2\pi(f - f_0)[m - n]}$$

in cui per $p_M(h)$, $z_M(h)$, SNR ed f_0 valgono le definizioni date in precedenza. Le sommatorie rispetto ad n e ad m sono da intendersi fra $-(N-1)/2 + \Sigma\tau_i$ e $(N-1)/2 - \Sigma\tau_i$ nel caso simmetrico, ovvero fra $\Sigma\tau_i$ e $N-1 - \Sigma\tau_i$ in quello non simmetrico. Per $M = 2$ ed $M = 3$ si ha in particolare (cfr. appendice A3):

$$(2.3.12) \quad \sigma_2^2(f) = \sigma_2^2 = 2 \cdot \frac{A^4}{\text{SNR}} \cdot [N - 2\tau]$$

$$(2.3.13) \quad \sigma_3^2(f) = E[|\delta S_3(f; \underline{\tau}_2)|^2] = 4 \cdot \frac{A^8}{\text{SNR}}$$

$$\left\{ \begin{aligned} & \left[N - 2 \cdot (\tau_1 + \tau_2) \right] \cdot u_{-1} \left[N - 2 \cdot (\tau_1 + \tau_2) \right] + \\ & + \frac{1}{2} \cdot \cos \left[4\pi \cdot (f - f_0) \cdot (\tau_1 - \tau_2) \right] \cdot \left[N - 4 \cdot \max(\tau_1, \tau_2) \right] \cdot u_{-1} \left[N - 4 \cdot \max(\tau_1, \tau_2) \right] + \\ & + \frac{1}{2} \cdot \cos \left[4\pi \cdot (f - f_0) \cdot (\tau_1 + \tau_2) \right] \cdot \left[N - 4 \cdot (\tau_1 + \tau_2) \right] \cdot u_{-1} \left[N - 4 \cdot (\tau_1 + \tau_2) \right] \end{aligned} \right\}$$

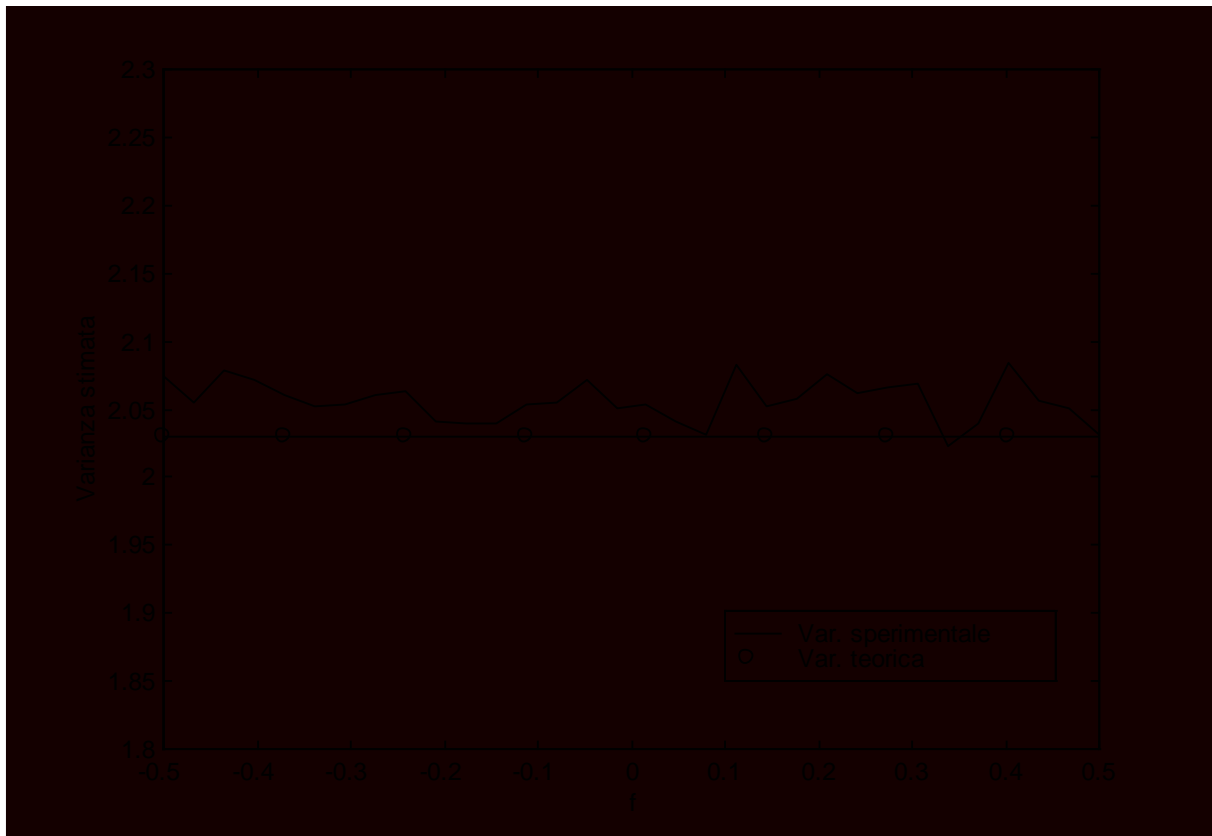
Si presti attenzione al fatto che nel caso in esame la varianza dipende in generale dalla frequenza f , a differenza di quanto accadeva in assenza di segnale. Negli esempi seguenti sono riportati i risultati di alcune simulazioni che provano la validità delle (2.3.12) e (2.3.13).

Esempio 1:

Con riferimento alla (2.3.1) si supponga:

$$\begin{array}{lll} A = 1 & N = 64 & M = 2 \\ L = 1 & \tau = 16 & SNR = 15 \text{ dB} \\ a_0 = 0.1 & a_1 = 0.3 & a_2 = 0.005 \end{array}$$

In queste condizioni la (2.3.12) fornisce $\sigma_2^2(f) = \sigma_2^2 \cong 2.024$. La varianza ottenuta per simulazione assume l'andamento riportato nella figura seguente:



Come si può notare l'accordo fra risultato teorico e sperimentale risulta molto buono; il leggero scarto esistente fra varianza teorica e varianza sperimentale è imputabile al fatto di aver trascurato i fattori contenenti il prodotto di due o più componenti di rumore nello sviluppo dell'*ml-HIM* (2.3.2).

Esempio 2:

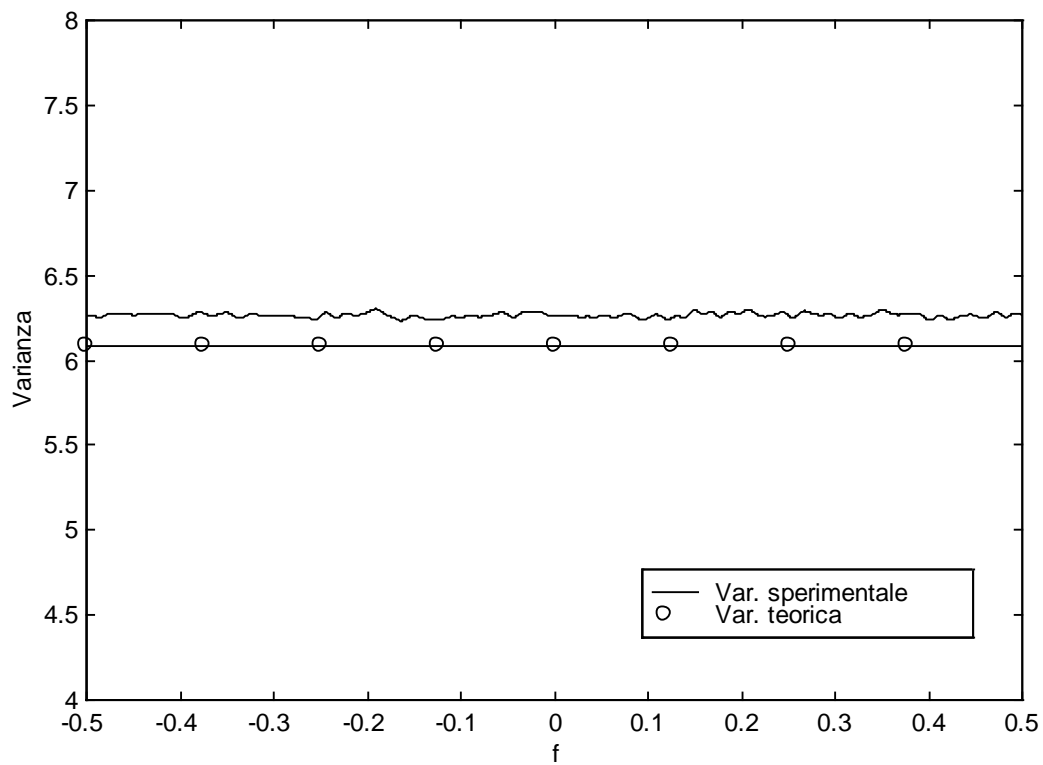
Questa volta si pone:

$$\begin{aligned}
 A &= 1 & N &= 192 & M &= 3 \\
 L &= 1 & \tau_1 &= \tau_2 = 32 \\
 a_0 &= -0.1 & a_1 &= 0.14 & a_2 &= 0.002 & a_3 &= 0.000006357 \\
 SNR &= 18 \text{ dB}
 \end{aligned}$$

La varianza teorica calcolata secondo la (2.3.13) vale:

$$\sigma_3^2(f) = 384/SNR \cdot A^8 \cong 6.086$$

mentre quella stimata assume il seguente andamento:



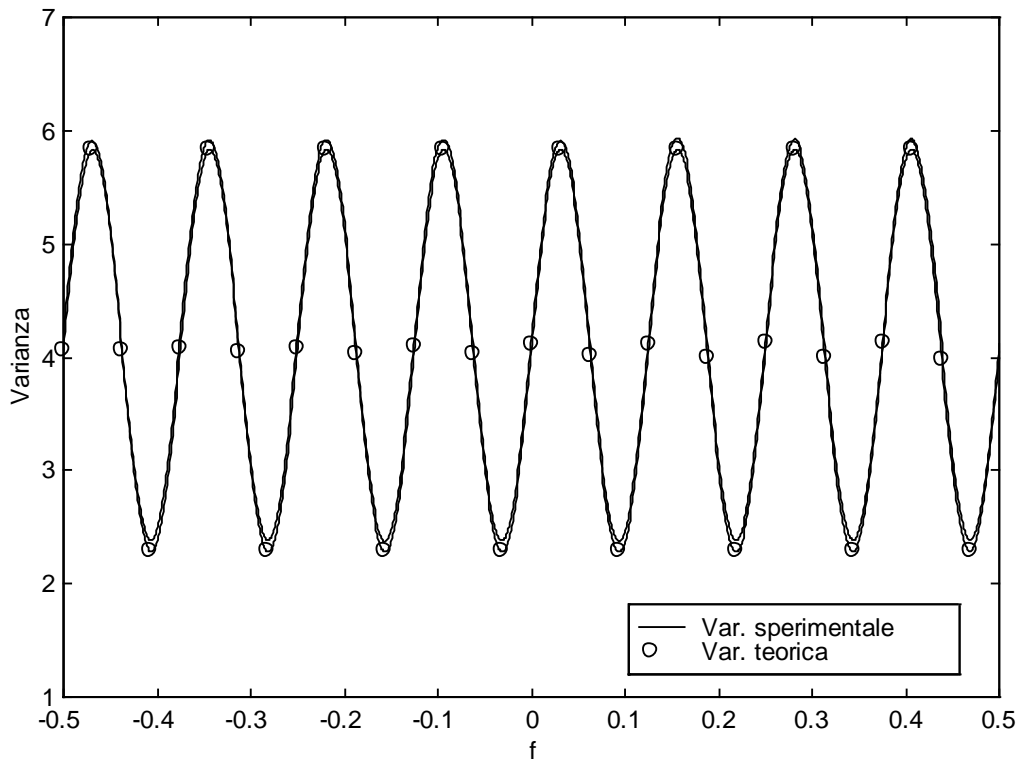
Anche in questo caso la relazione teorica si rivela ben confermata. Supponendo ora di modificare i lags nel modo seguente:

$$\tau_1 = 34\tau_2 = 30$$

si ottiene:

$$\sigma_s^2(f) = 4.057 + 1.775 \cdot \cos [16 \cdot \pi \cdot (f - 0.15625)]$$

La simulazione fornisce stavolta:



Quest'ultimo esempio conferma ulteriormente la validità della (2.3.13).

Le considerazioni precedenti consentono di affermare che il processo $X_M(f; \underline{\tau}_{M-1})$ risulta Gaussiano, con valore atteso $S_M(f; \underline{\tau}_{M-1})$ dato da (2.3.6) e varianza (2.3.11). Da notare che sia l'uno che l'altra dipendono dalla frequenza f , per cui $X_M(f; \underline{\tau}_{M-1})$ non appare "stazionario" in frequenza. Inoltre la densità di probabilità del modulo quadro della ml-HAF non risulta facilmente determinabile, a causa della presenza del segnale che rende diverso da zero il valore medio.

3. Rivelazione di segnali PPS immersi in AWGN

3.1 Introduzione

Nei capitoli precedenti sono stati presentati dei richiami sui segnali a fase polinomiale ed è stato quindi affrontato lo studio delle caratteristiche statistiche della ml-HAF; risulta ora possibile esaminare con maggiore attenzione il già citato problema della rivelazione, che si presenta quando un segnale a fase polinomiale si trova immerso in rumore additivo. Ogni criterio di rivelazione si propone di discriminare la presenza di un determinato segnale, analizzando dati affetti da rumore additivo. L'esigenza di metodi robusti di rivelazione si riscontra in molteplici campi di applicazione, si pensi ad esempio alla rivelazione di bersagli radar. In questi casi la soluzione tipicamente adottata si basa sul criterio di Neyman-Pearson: si fissa cioè una soglia sul segnale (rumoroso), determinata sulla base di una prefissata probabilità di falso allarme (superamento della soglia dovuto al solo rumore) e si valuta poi la corrispondente probabilità di rivelazione corretta (detta probabilità di detezione), che risulta in genere dipendente dal rapporto segnale-rumore. Questo criterio presuppone implicitamente l'ipotesi di linearità, in quanto richiede che la probabilità di falso allarme determinata in assenza di segnale utile si conservi inalterata quando quest'ultimo è presente (principio di sovrapposizione degli effetti). Purtroppo il metodo di stima basato sulla ml-HAF non gode di questa proprietà, per cui non risulta possibile agire nel modo suindicato (cfr. esempio successivo).

Esempio:

Consideriamo un segnale PPS mono-componente immerso in AWGN:

$$x(n) = A \cdot e^{j \cdot 2\pi \cdot (a_0 + a_1 \cdot n + \dots + a_M \cdot n^M)} + w(n)$$

Per fissare le idee si supponga $M=2$. Nel capitolo 2 è stato dimostrato che in assenza di segnale il modulo quadro della ml-HAF possiede una distribuzione esponenziale del tipo:

$$p_Q(q) = \lambda \cdot e^{-\lambda q} \cdot u_1(q)$$

dove:

$$q = |X_2(f; \tau)|^2$$

$$\lambda = 1/\sigma_2^2 = 1 / [\sigma_w^4 \cdot (N-2 \cdot \tau)]$$

Se imponessimo una soglia T su q la corrispondente probabilità di falso allarme in assenza di segnale varrebbe:

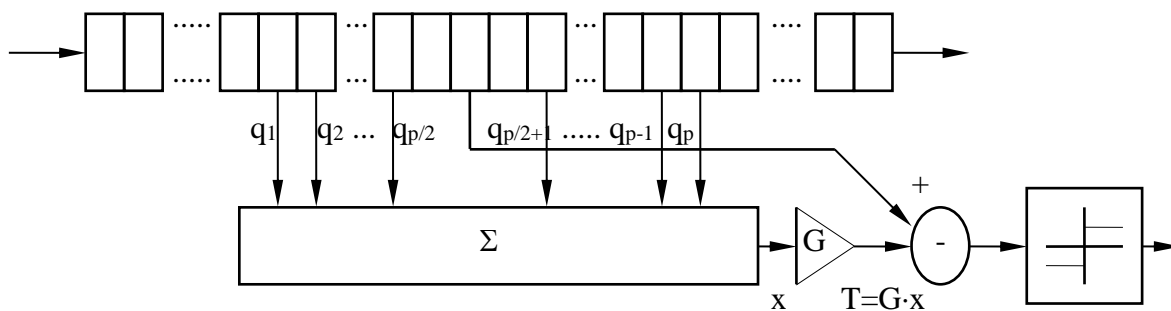
$$P_{FA} = e^{-\lambda T}$$

In presenza di segnale cambierebbero sia la densità di probabilità di q che i suoi momenti; ciò comporterebbe evidentemente anche una variazione della P_{FA} . Infatti, se anche ammettessimo valido in prima approssimazione l'andamento esponenziale, avremmo comunque un notevole incremento della varianza (a causa delle interazioni non lineari fra segnale e rumore quest'ultima passerà al nuovo valore $\sigma_2^2 = 2 \cdot [N-2\tau] \cdot A^2 \cdot \sigma_w^2$, con un incremento rispetto a quello precedente pari a $2 \cdot SNR$). Occorre poi considerare che il problema è reso ancora più complesso dalla non stazionarietà in frequenza del modulo quadro della ml-HAF ed in particolar modo dalla variabilità con f del suo valore atteso $S_M(f; \underline{\tau}_{M-1})$, che rende problematica l'imposizione di una soglia fissa.

I problemi della non linearità e della non stazionarietà in frequenza comportano quindi due inconvenienti: la non linearità implica l'impossibilità di fissare una soglia in assenza di segnale, perché la P_{FA} che si otterrebbe varierebbe in presenza di quest'ultimo (e dipenderebbe inoltre dal rapporto SNR); la non stazionarietà produce invece variabilità in frequenza della P_{FA} ottenuta imponendo una soglia fissa. Entrambi possono essere risolti ricorrendo ad una soglia adattiva; la restante parte del presente capitolo è quindi dedicata allo studio di questa soluzione, con particolare attenzione all'esame delle prestazioni del metodo, fornite attraverso le curve di rivelazione.

3.2 Probabilità di falso allarme ottenibile con impiego di una soglia adattiva (L=1).

Nel paragrafo precedente si è introdotto il problema della rivelazione e sono stati illustrati i principali inconvenienti legati all'imposizione di una soglia fissa; è stata quindi suggerita una possibile soluzione consistente nell'adozione di una soglia adattiva, che può essere ottenuta attraverso l'utilizzazione di un dispositivo del tipo seguente:



Il principio di funzionamento è molto semplice: sulla linea di ritardo posta superiormente scorrono i campioni in frequenza del modulo quadro della ml-HAF ($q_i = |X_M(f_i; \underline{T}_{M-1})|^2$; $i = 1, 2, \dots, p$); la soglia T posta sul campione centrale risulta quindi proporzionale alla somma dei p campioni ad esso contigui (quelli immediatamente adiacenti vengono in realtà scartati per evitare che il segnale utile possa contribuire al valore della soglia stessa). Un dispositivo del tipo riportato nella figura precedente viene normalmente utilizzato nei radar per ottenere una probabilità di falso allarme costante (condizione CFAR) ed è noto in quel contesto con il nome di autogate; nel seguito tale denominazione verrà conservata per semplicità.

Si supponga ora di impiegare effettivamente la struttura precedente per la rivelazione di un segnale a fase polinomiale immerso in rumore additivo Gaussiano bianco; il primo problema da affrontare riguarda la scelta dei parametri G e p (guadagno e numero di prese) necessari per ottenere una prefissata probabilità di falso allarme P_{FA} . Nel caso di segnali PPS mono-componente risulta abbastanza agevole determinare una espressione in forma chiusa del tipo $P_{FA} = P_{FA}(G, p)$. Si noti tuttavia che le approssimazioni fatte nel calcolare i parametri statistici della PHAF valgono nell'ipotesi di alto SNR; pertanto anche i seguenti sviluppi si baseranno su tale ipotesi. Come già visto in precedenza si ha:

$$(3.2.1) \quad X_M(f; \underline{T}_{M-1}) \cong S_M(f; \underline{T}_{M-1}) + \delta S_M(f; \underline{T}_{M-1})$$

dove:

$$(3.2.2) \quad S_M(f; \underline{\tau}_{M-1}) = A^{2^{M-1}} \cdot e^{j\theta} \cdot \frac{\sin \left[\pi \cdot (f - f_0) \cdot \left(N - 2 \cdot \sum_{i=1}^{M-1} \tau_i \right) \right]}{\sin \left[\pi \cdot (f - f_0) \right]}$$

$$(3.2.3) \quad \delta S_M(f; \underline{\tau}_{M-1}) = \sum_n \delta s_M(n; \underline{\tau}_{M-1}) \cdot e^{-j \cdot 2\pi \cdot f \cdot n} =$$

$X_M(f; \underline{\tau}_{M-1})$ risulta Gaussiano complesso con:

$$(3.2.4) \quad E[X_M(f; \underline{\tau}_{M-1})] = S_M(f; \underline{\tau}_{M-1})$$

$$(3.2.5) \quad E[|\delta S_M(f; \underline{\tau}_{M-1})|^2] = \sigma_M^2(f) = \frac{A^{2^M}}{\text{SNR}} \cdot \sum_{k=1}^{2^{M-1}} \sum_{h=1}^{2^{M-1}} \delta[p_M(k) - p_M(h)] \cdot \sum_n \sum_m \delta[n - m + z_M(k) - z_M(h)] \cdot e^{j2\pi(f - f_0)[m - n]}$$

Si faccia ora l'ulteriore ipotesi di trascurare i lobi laterali della (3.2.2) rispetto a quello principale. In queste condizioni $X_M(n; \underline{\tau}_{M-1})$ può essere ritenuto Gaussiano a valore atteso nullo (almeno per $f \neq f_0$), ed il suo modulo quadro assume allora distribuzione esponenziale:

$$(3.2.6) \quad p_{Q_i}(q_i) = \lambda_i \cdot e^{-\lambda_i \cdot q_i} \cdot u_{-1}(q_i)$$

con:

$$(3.2.7) \quad \lambda_i = \lambda(f_i) = 1/\sigma_M^2(f_i)$$

Nel caso particolare $M = 2$ la (3.2.5) fornisce una varianza indipendente dalla frequenza:

$$(3.2.8) \quad \sigma_2^2(f) = \sigma_2^2 = 2 \cdot [N - 2\tau] \cdot A^2 \cdot \sigma_w^2$$

per cui i campioni sulle p prese dell'autogate risultano identicamente distribuiti (sono infatti tutti esponenziali con identico parametro λ). Si assume poi per semplicità che essi siano anche statisticamente indipendenti⁶. Ricordando la definizione della soglia T:

$$(3.2.9) \quad T = G \cdot \sum_{i=1}^p q_i$$

ed osservando che la densità di probabilità della somma di p variabili aleatorie esponenziali identicamente distribuite ed indipendenti ha distribuzione gamma, si ricava:

$$(3.2.10) \quad p_T(T) = \frac{1}{(G \cdot \sigma_2^2)^p \cdot \Gamma(p)} \cdot T^{p-1} \cdot e^{-\frac{T}{G \cdot \sigma_2^2}} \cdot u_{-1}(T)$$

Detto quindi q il campione sulla presa centrale dell'autogate, che per $f \neq f_0$ ha ancora distribuzione del tipo (3.2.6), la probabilità di falso allarme può essere definita nel modo seguente:

$$(3.2.11) \quad P_{FA} = \text{Prob} \{ q > T \} = \text{Prob} \{ T - q < 0 \} = F_Z(0)$$

Nella precedente si è posto:

$$(3.2.12) \quad z = T - q$$

$$(3.2.13) \quad F_Z(z) = \text{Prob} \{ Z < z \} \equiv f_z. \text{ di distribuzione della v.a. } z$$

Supponendo ulteriormente che ci sia indipendenza statistica fra T e q si ricava:

$$(3.2.14) \quad F_Z(z) = \int_{-\infty}^{+\infty} p_T(T) \cdot [1 - F_Q(T - z)] \cdot dT$$

dove $F_Q(q)$ rappresenta la funzione di distribuzione della variabile q ed e' data dall'integrale della (3.2.6). Sostituendo in (3.2.14) si ottiene:

⁶ L'ipotesi di indipendenza è plausibile, visto che per simulazione si ricavano valori del coefficiente di correlazione non superiori a 10^{-2} .

$$(3.2.15) \quad P_{FA} = F_Z(0) = \int_0^{+\infty} \frac{1}{(G \cdot \sigma_2^2)^p \cdot \Gamma(p)} \cdot T^{p-1} \cdot e^{-\frac{T \cdot (G+1)}{G \cdot \sigma_2^2}} \cdot u_{-1}(T) \cdot dT$$

che risolto⁷ fornisce la seguente espressione per la probabilità di falso allarme:

$$(3.2.16) \quad P_{FA} = \frac{1}{(G+1)^p}$$

Il risultato ottenuto mostra che la P_{FA} dipende esclusivamente dai parametri dell'autogate (G e p); essa risulta pertanto indipendente da altri fattori, ed in modo particolare dal rapporto segnale-rumore (proprietà molto utile in virtù delle considerazioni fatte in precedenza). E' anche possibile osservare che la relazione (3.2.16), pur essendo stata ottenuta nel caso specifico $M = 2$, può essere approssimativamente ritenuta valida anche per valori di M superiori (in particolare per $M = 3$). In quest'ultimo caso, infatti, la varianza (3.2.5) non è più costante (almeno in generale), per cui i campioni sulle prese dell'autogate sono ancora esponenziali e statisticamente indipendenti, ma non identicamente distribuiti. Indicando allora con λ_i ($i = 1, \dots, p$) i reciproci delle varianze dei suddetti campioni (e con λ il reciproco della varianza del campione sulla presa centrale) è possibile provare che risulta⁸:

$$(3.2.17) \quad P_{FA} = \frac{1}{\prod_{i=1}^p \left[1 + G \cdot \frac{\lambda}{\lambda_i} \right]}$$

che coincide con la (3.2.16) nell'ipotesi $\lambda_i = \lambda$ ($i = 1, 2, \dots, p$). Da notare che il rapporto λ/λ_i non dipende né da A né da SNR ⁹, per cui anche la P_{FA} ottenuta gode della stessa proprietà; inoltre tale rapporto tende in genere ad assumere valori prossimi all'unità, per cui la (3.2.17) fornisce valori della probabilità di falso allarme all'incirca pari a quelli dati dalla (3.2.16). Quest'ultima

⁷ L'integrale (3.2.15) si risolve per sostituzione ponendo $t = T \cdot (G+1)/(G \cdot \sigma_2^2)$ e ricordando la definizione della funzione gamma:

$$\Gamma(p) = \int_0^{+\infty} t^{p-1} \cdot e^{-t} \cdot dt$$

⁸ La dimostrazione completa della (3.2.17) è riportata nell'appendice A4.

⁹ Questa affermazione risulta vera nel limite di validità dell'ipotesi di alto SNR.

osservazione consente anche di giustificare il fatto che la P_{FA} (3.2.17) risulti praticamente indipendente dalla frequenza f , come dovrebbe invece accadere essendo sia λ che i vari λ_i dipendenti da quest'ultima. Per maggiori dettagli si rimanda alla appendice A5.

La validità delle relazioni (3.2.16) - (3.2.17) è subordinata all'ipotesi di trascurare i lobi laterali della (3.2.2); occorre a questo punto verificare che la suddetta ipotesi sia lecita. A tale proposito si prenda in considerazione il seguente esempio:

Esempio 1:

Sia dato un segnale PPS mono-componente immerso in rumore additivo Gaussiano bianco:

$$x_1(n) = e^{j \cdot 2\pi \cdot (0.1 + 0.2 \cdot n + n^2 / 32768)} + w(n)$$

Sia poi:

$$\begin{aligned} N &= 256 & L &= 1 & \tau &= 64 \\ G &= 1 & p &= 6 \end{aligned}$$

In base alla (3.2.16) si dovrebbe avere $P_{FA} = (G+1)^{-p} = 1/64 = 1.5625 \cdot 10^{-2}$. Le simulazioni al calcolatore forniscono i seguenti valori (al variare di SNR):

SNR (dB)	0	5	10	15	20	25	30
P_{FA}	$1.438 \cdot 10^{-2}$	$1.452 \cdot 10^{-2}$	$1.483 \cdot 10^{-2}$	$1.456 \cdot 10^{-2}$	$1.457 \cdot 10^{-2}$	$1.483 \cdot 10^{-2}$	$1.466 \cdot 10^{-2}$

Come si può notare la probabilità di falso allarme si mantiene approssimativamente costante al variare del rapporto segnale-rumore in un intervallo che copre ben 3 ordini di grandezza (30 dB); inoltre il valore sperimentale della P_{FA} si avvicina notevolmente a quello teorico (la differenza è nell'ordine di grandezza del 5-6% ed è probabilmente dovuta all'approssimazione di trascurare quello che accade quando su una delle prese dell'autogate è presente il picco del segnale utile, che contribuisce ad innalzare notevolmente la soglia). Nel caso esaminato la (3.2.16) si dimostra valida e ciò conferma l'ipotesi fatta a proposito dei lobi laterali.

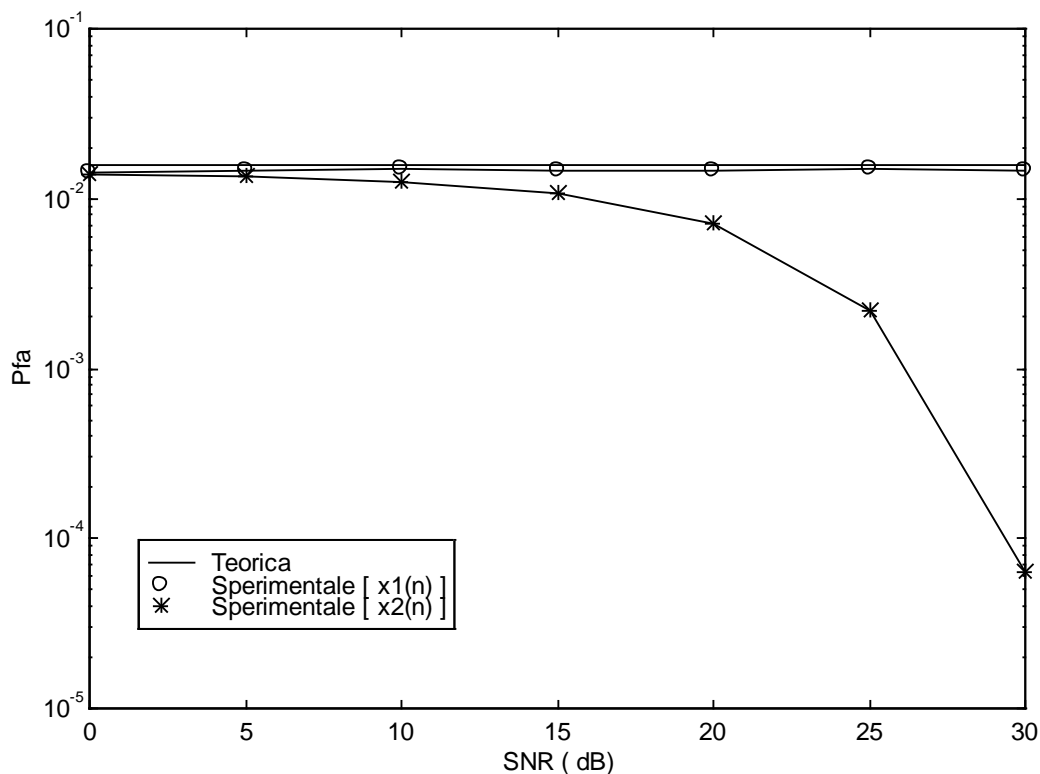
Si supponga ora di variare il solo coefficiente di grado più elevato, dimezzandolo; si avrà quindi:

$$x_2(n) = e^{j \cdot 2\pi \cdot (0.1 + 0.2 \cdot n + n^2 / 65536)} + w(n)$$

Secondo quanto visto in precedenza la probabilità di falso allarme dovrebbe conservarsi inalterata; i valori ottenuti per via sperimentale non sembrano però confermare questa osservazione (cfr. tabella seguente):

SNR (dB)	0	5	10	15	20	25	30
P_{FA}	$1.392 \cdot 10^{-2}$	$1.358 \cdot 10^{-2}$	$1.266 \cdot 10^{-2}$	$1.078 \cdot 10^{-2}$	$7.194 \cdot 10^{-3}$	$2.208 \cdot 10^{-3}$	$6.417 \cdot 10^{-5}$

I risultati precedenti sono riportati nella figura seguente:



Come si può notare nel secondo caso la relazione (3.2.16) non appare valida.

Per spiegare lo strano fenomeno visto nell'esempio precedente si può anzitutto osservare che il numero di campioni della ml-HAF coincide con la lunghezza dell'ml-HIM ed è quindi dato da $N_{HIM} = N \cdot 2 \cdot \tau$, che nel caso in esame vale $N_{HIM} = 256 \cdot 2 \cdot 64 = 128$. I campioni che scorrono nella linea di ritardo dell'autogate sono pertanto estratti a distanza $1/N_{HIM} = 1/128$ l'uno dall'altro. Tenendo presente il legame fra posizione del massimo e coefficiente di grado più elevato si ha che nel primo caso il picco principale cade esattamente su un campione ($f_0 = 1/128$), per cui tutti i restanti vengono estratti in corrispondenza degli zeri del valore atteso (3.2.2); risulta allora evidentemente valida l'ipotesi di esponenzialità dei suddetti campioni, e con essa la relazione (3.2.16). Nel secondo, invece, il lobo principale interessa due campioni consecutivi ($f_0 = 1/256$),

mentre tutti gli altri cadono in corrispondenza dei massimi dei lobi secondari; conseguentemente viene meno la validità dell'approssimazione esponenziale e della relazione (3.2.16).

Per eliminare il problema dei lobi laterali (in modo da garantire l'indipendenza della P_{FA} dalla posizione del picco della ml-HAF) si può ricorrere alla pesatura dell'HIM attraverso una opportuna finestra. Detta:

$$(3.2.18) \quad N_{HIM} = N - 2 \cdot \Sigma \tau$$

la lunghezza dell'HIM $x_M(n ; \underline{\tau}_{M-1})$, la finestra è data da:

$$(3.2.19) \quad W(n) = k + (1-k) \cdot \cos^2 [\pi \cdot n / (N_{HIM}-1)] \quad |n| \leq (N_{HIM}-1)/2$$

L'HIM pesato e la HAF pesata sono definite come segue:

$$(3.2.20) \quad x_M'(n ; \underline{\tau}_{M-1}) = W(n) \cdot x_M(n ; \underline{\tau}_{M-1})$$

$$(3.2.21) \quad X_M'(f ; \underline{\tau}_{M-1}) = \Sigma x_M'(n ; \underline{\tau}_{M-1}) \cdot e^{-j2\pi fn}$$

La suddetta operazione consente di abbattere i lobi laterali, ma viene pagata con una perdita di risoluzione (ossia con l'allargamento del lobo principale); ciò non provoca comunque alcun tipo di problema, avendo eliminato dal computo della soglia i campioni adiacenti a quello centrale.

In modo analogo a quanto avviene nel caso di assenza di pesatura, anche stavolta si può scrivere (per alti SNR):

$$(3.2.22) \quad X_M'(f ; \underline{\tau}_{M-1}) \cong S_M'(f ; \underline{\tau}_{M-1}) + \delta S_M'(f ; \underline{\tau}_{M-1})$$

e si può ancora assumere che $\delta S_M'(f ; \underline{\tau}_{M-1})$ sia Gaussiano a valore atteso nullo. La varianza di (3.2.22) sarà senz'altro diversa da quella calcolata a suo tempo (ed in particolare non è detto che nel caso $M = 2$ risulti costante). Si può comunque evitare di calcolarne l'espressione visto che non dovrebbe influire eccessivamente sul valore della P_{FA} ottenibile ($P_{FA} \cong (G+1)^{-p}$). Negli esempi seguenti sono riportati i risultati di alcune simulazioni effettuate adottando la pesatura dell'ml-HIM.

Esempio 2:

Sia: $x_1(n) = e^{j \cdot 2\pi \cdot (0.1 + 0.2 \cdot n + n^2 / 32768)} + w(n)$
 $x_2(n) = e^{j \cdot 2\pi \cdot (0.1 + 0.2 \cdot n + n^2 / 65536)} + w(n)$

dove:

$$N = 256 \quad L = 1 \quad \tau = 64$$

$$G = 1 \quad p = 4$$

In entrambi i casi la P_{FA} teorica dovrebbe assumere il valore $(G+1)^{-p} = 0.0625$. Nella seguente tabella sono riportati i risultati ottenuti per simulazione nei due casi di assenza di pesatura ($k = 1$) e di pesatura con $k = 0.08$; per ciascuno di essi viene effettuato il calcolo con i due differenti valori del coefficiente di grado massimo:

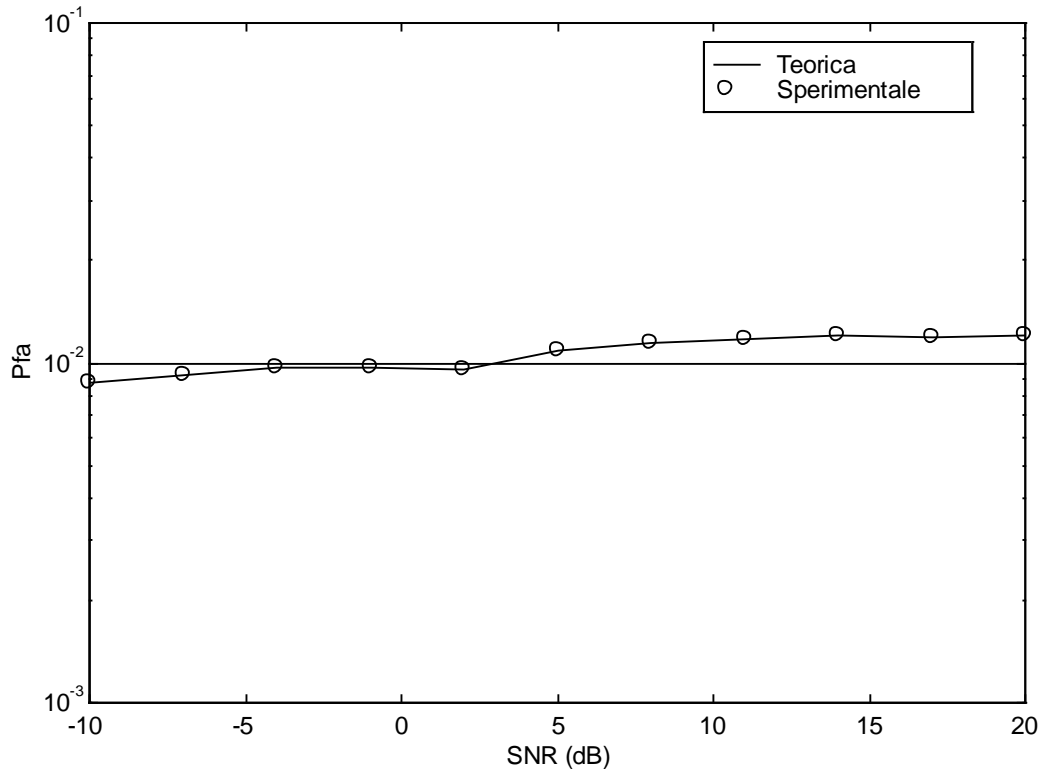
CNR (dB)	P_{FA}		P_{FA}	
	$a_2=1/32768$ $k=1$	$a_2=1/65536$ $k=1$	$a_2=1/32768$ $k=0.08$	$a_2=1/65536$ $k=0.08$
0	0.0596	0.0647	0.0697	0.0698
5	0.0610	0.0642	0.0701	0.0700
10	0.0602	0.0611	0.0701	0.0698
15	0.0604	0.0558	0.0703	0.0684
20	0.0611	0.0450	0.0702	0.0670
25	0.0602	0.0262	0.0705	0.0619
30	0.0607	0.0107	0.0701	0.0521

Come si può notare la P_{FA} ottenuta con adozione della pesatura è praticamente indipendente da a_2 , ossia dalla posizione del picco, e si conserva approssimativamente costante per valori di rapporto SNR che variano da 0 dB a 30 dB. Da osservare però che il suo valore è leggermente diverso da quello teorico; questo effetto potrebbe essere dovuto al fatto che stavolta la varianza della ml-HAF non è costante, per cui la relazione $P_{FA} = (G+1)^{-p}$ è valida solo con una certa approssimazione.

L'autogate sembra funzionare bene anche nel caso $M = 3$, come testimonia il seguente esempio.

Esempio 3:

Sia dato un segnale PPS mono-componente di ordine $M = 3$ immerso in AWGN; sia poi $N=192$, $L=1$, $\tau_1=\tau_2=32$, $p=4$, $G=100^{1/4}-1 \cong 2.1623$ (per cui P_{FA} dovrebbe assumere il valore 10^{-2}). I risultati ottenuti per simulazione sono riportati nel seguente grafico:



Nonostante la (3.2.16) sia da intendersi valida solo in modo approssimato per $M = 3$ si può notare una buona coincidenza fra risultati sperimentali e teorici.

All'aumentare del numero di prese p il comportamento dell'autogate tende a peggiorare leggermente; infatti la P_{FA} ottenuta si mantiene ancora all'incirca costante, ma la validità della relazione (3.2.16) sembra diminuire.

Esempio 4:

Sia:

$$M=2 \quad N=128 \quad L=1 \quad \tau=32 \quad p=16 \quad G=0.55$$

La P_{FA} dovrebbe assumere il valore $9.009 \cdot 10^{-4}$. I risultati ottenuti per simulazione sono riportati nella seguente tabella:

<i>SNR (dB)</i>	<i>-10</i>	<i>-5</i>	<i>0</i>	<i>5</i>	<i>10</i>	<i>15</i>	<i>20</i>
P_{FA}	$1.275 \cdot 10^{-3}$	$1.100 \cdot 10^{-3}$	$1.054 \cdot 10^{-3}$	$1.118 \cdot 10^{-3}$	$1.128 \cdot 10^{-3}$	$1.130 \cdot 10^{-3}$	$9.943 \cdot 10^{-4}$

Il valore della P_{FA} ottenuta differisce leggermente da quello teorico (più di quanto non accada nel caso $p=4$).

Ciò può essere spiegato osservando che la relazione (3.2.16) presuppone implicitamente una serie di ipotesi che risultano in realtà valide solo in modo approssimato; tra queste occorre ricordare in particolare quella di esponenzialità dei campioni del modulo quadro della ml-HAF. L'espressione $P_{FA} = (1+G)^{-p}$ è stata ottenuta sfruttando il fatto che la somma di p variabili aleatorie esponenziali ha distribuzione Gamma; se però le suddette variabili sono “quasi” esponenziali, la loro somma ammetterà una densità di probabilità solo approssimativamente del tipo Gamma, ed intuitivamente l'approssimazione sarà tanto peggiore quanto maggiore è il numero di variabili p che vengono sommate.

3.3 Analisi delle prestazioni ($L=1$).

Nel paragrafo precedente è stato introdotto il principio di funzionamento dell'autogate ed è stato esaminato il legame intercorrente fra i parametri di quest'ultimo (guadagno G e numero di prese p) e la corrispondente P_{FA} ottenibile. Occorre a questo punto esaminare le prestazioni fornite dal dispositivo, mostrando l'andamento della probabilità di detezione P_D al variare del rapporto segnale-rumore (per un prefissato valore della probabilità di falso allarme P_{FA}). Il calcolo teorico di un'espressione del tipo $P_D = P_D (SNR, P_{FA})$ può essere impostato in modo analogo a quanto fatto in precedenza per la valutazione della P_{FA} . Facciamo per il momento alcune ipotesi semplificative:

1. $M=2$ (segnali del secondo ordine \Rightarrow varianza costante);
2. $L=1$;
3. $k=1$ (assenza di pesatura);
4. campionamento sul picco utile della ml-HAF (ossia in $f = f_0$).

Di conseguenza il campione della ml-HAF corrispondente al lobo principale può essere assunto Gaussiano complesso, con valore atteso:

$$(3.3.1) \quad \mu = S_2(f_0; \tau) = A^2 \cdot [N - 2 \cdot \tau] \cdot e^{j \cdot 2\pi \cdot P_M \cdot a_{M-1}}$$

e varianza:

$$(3.3.2) \quad \sigma_2^2 = 2 \cdot \frac{A^4}{\text{SNR}} \cdot [N - 2 \cdot \tau]$$

La densità di probabilità del suo modulo (indicato con u) è del tipo Rice:

$$(3.3.3) \quad p_U(u) = \frac{2 \cdot u}{\sigma_2^2} \cdot e^{-\frac{u^2 + |\mu|^2}{\sigma_2^2}} \cdot I_0 \left[\frac{2 \cdot |\mu|}{\sigma_2^2} \cdot u \right]$$

In condizioni di elevato SNR la varianza (3.3.2) è piccola rispetto al modulo quadro del valor medio (3.3.1), per cui la (3.3.3) può essere adeguatamente approssimata da una Gaussiana. La condizione di validità dell'approssimazione è la seguente:

$$(3.3.4) \quad \frac{\sigma_2^2}{|\mu|^2} \ll 1$$

da cui:

$$(3.3.5) \quad \text{SNR} \gg \frac{2}{N - 2 \cdot \tau}$$

Ad esempio, scegliendo $N = 128$ e $\tau = 128/4 = 32$ si ha:

$$(3.3.6) \quad \text{SNR} \gg 1/32 \cong -15 \text{ dB}$$

La densità di probabilità (3.3.3) risulta quindi approssimabile con una Gaussiana già a partire da valori di rapporto segnale-rumore estremamente bassi. Si ha allora:

¹⁰ $I_0(x)$ indica la funzione di Bessel di ordine 0, definita come segue:

$$I_0(x) = \frac{1}{2\pi} \cdot \int_0^{2\pi} e^{x \cdot \cos \vartheta} \cdot d\vartheta$$

$$(3.3.7) \quad p_U(u) \cong \frac{1}{\sqrt{2\pi} \cdot \sigma_{\text{Rice}}} \cdot e^{-\frac{(u-\mu_{\text{Rice}})^2}{2 \cdot \sigma_{\text{Rice}}^2}}$$

dove:

$$(3.3.8) \quad \mu_{\text{Rice}} = E[u] = \frac{\sigma_2 \cdot \sqrt{\pi}}{2} \cdot L_{\frac{1}{2}}^0 \left[-\frac{|\mu|^2}{\sigma_2^2} \right] \quad 11$$

$$(3.3.9) \quad \sigma_{\text{Rice}}^2 = E[u^2] - E^2[u] = |\mu|^2 + \sigma_2^2 - \mu_{\text{Rice}}^2$$

La probabilità di detezione risulta data da un'espressione analoga alla (3.2.11) del paragrafo precedente:

$$(3.3.10) \quad P_D = \text{Prob} \{ q > T \} = \text{Prob} \{ T - q < 0 \} = F_Z(0)$$

nella quale però si assume che q rappresenti il modulo quadro della ml-HAF campionata in corrispondenza del picco principale ($q = |X_2(f_0; \tau)|^2 = u^2$). Si ha quindi:

$$(3.3.11) \quad F_Z(z) = \int_{-\infty}^{+\infty} p_T(T) \cdot [1 - F_Q(T - z)] \cdot dT$$

dove:

$$(3.3.12) \quad \begin{aligned} [1 - F_Q(T - z)] &= \int_{T-z}^{+\infty} p_Q(q) \cdot dq = \text{Prob}\{q > T - z\} = \\ &= \text{Prob}\{u > \sqrt{T - z}\} = \int_{\sqrt{T-z}}^{+\infty} p_U(u) \cdot du \cong \frac{1}{2} \cdot \text{erfc} \left[\frac{\sqrt{T - z} - \mu_{\text{Rice}}}{\sqrt{2} \cdot \sigma_{\text{Rice}}} \right] \end{aligned}$$

Tenendo conto delle (3.3.10) e (3.3.11) si ottiene:

¹¹ $L_n^a[x]$ e' il polinomio di Laguerre definito come soluzione dell'equazione differenziale:
 $x \cdot y'' + (\alpha + 1 - x) \cdot y' + n \cdot y = 0$

$$\begin{aligned}
P_D &= F_Z(0) = \int_{-\infty}^{+\infty} p_T(T) \cdot \frac{1}{2} \cdot \operatorname{erfc} \left[\frac{\sqrt{T} - \mu_{\text{Rice}}}{\sqrt{2} \cdot \sigma_{\text{Rice}}} \right] \cdot dT = \\
(3.3.13) \quad &= \frac{1}{2} \cdot \int_0^{+\infty} \frac{1}{(G \cdot \sigma_2^2)^p \cdot \Gamma(p)} \cdot T^{p-1} \cdot e^{-\frac{T}{G \cdot \sigma_2^2}} \cdot \operatorname{erfc} \left[\frac{\sqrt{T} - \mu_{\text{Rice}}}{\sqrt{2} \cdot \sigma_{\text{Rice}}} \right] \cdot dT = \\
&= \frac{1}{2} - \frac{1}{2} \cdot \int_0^{+\infty} \frac{x^{p-1} \cdot e^{-x}}{(p-1)!} \cdot \operatorname{erf} \left[\sqrt{\frac{G \cdot \sigma_2^2 \cdot x}{2\sigma_{\text{Rice}}^2}} - \frac{\mu_{\text{Rice}}}{\sqrt{2} \cdot \sigma_{\text{Rice}}} \right] \cdot dx
\end{aligned}$$

L'uguaglianza fra penultimo ed ultimo membro della (3.3.13) discende dalla sostituzione $x = T/(G \cdot \sigma_2^2)$. L'integrale precedente non ammette un'espressione in forma chiusa, per cui non risulta possibile determinare una relazione teorica che fornisca P_D in funzione di SNR e di P_{FA} . L'unica strada percorribile rimane pertanto quella delle simulazioni al calcolatore. Di seguito vengono presentati i risultati ottenuti in alcuni casi significativi, presentandoli sia sotto forma di tabelle che di grafici (curve di rivelazione o detezione).

Esempio 1:

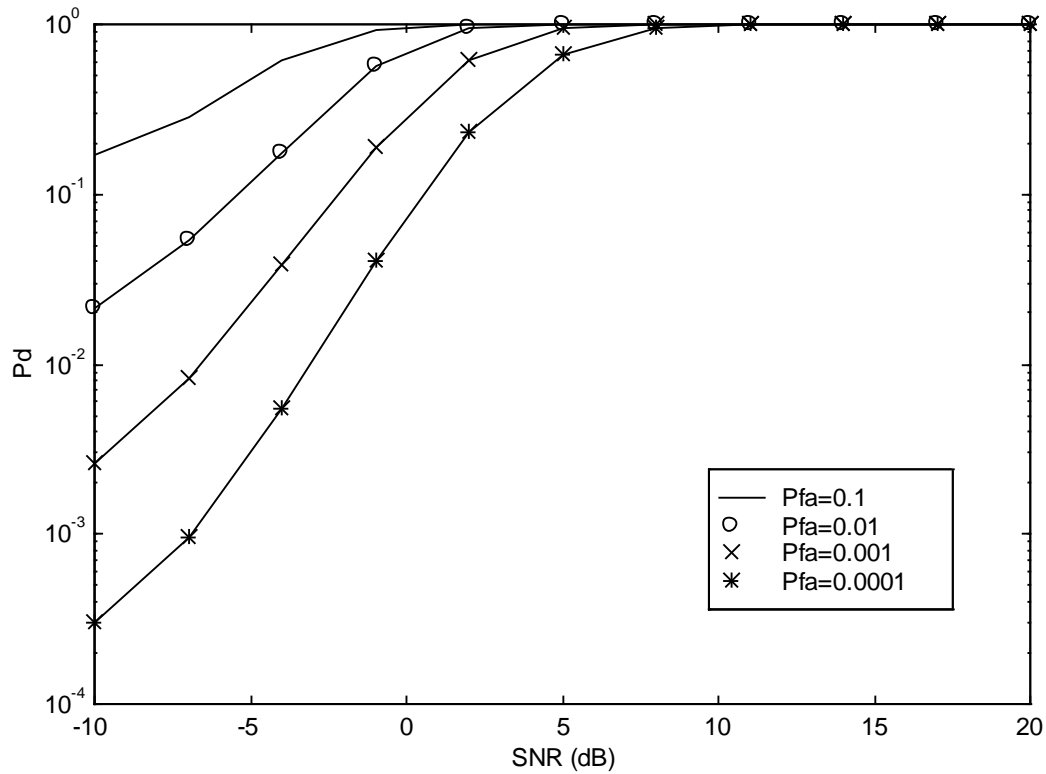
Si consideri un segnale mono-component PPS del secondo ordine ($M = 2$), immerso in AWGN e risulti $N = 128$. Si ipotizzi ulteriormente di avere $L = 1$, $\tau = 32$, $k = 0.08$ (parametro di pesatura) e di utilizzare un autogate a 4 prese ($p = 4$). Nella tabella seguente sono riportati i valori del guadagno G utilizzati nelle simulazioni, con i corrispondenti valori teorici e sperimentali della P_{FA} :

G	P_{FA} teorica	P_{FA} sperimentale
$10^{1/4}-1$	10^{-1}	$1.0084 \cdot 10^{-1}$
$100^{1/4}-1$	10^{-2}	$1.1973 \cdot 10^{-2}$
$1000^{1/4}-1$	10^{-3}	$1.3677 \cdot 10^{-3}$
$10000^{1/4}-1$	10^{-4}	$1.4921 \cdot 10^{-4}$

La successiva tabella mostra invece i valori della P_D al variare di SNR e di G (ossia di P_{FA}):

CNR (dB)	$G=10^{1/4}-1$	$G=100^{1/4}-1$	$G=1000^{1/4}-1$	$G=10000^{1/4}-1$
-10	0.1700	0.0215	0.0026	0.0003
-7	0.2820	0.0540	0.0082	0.0009
-4	0.6070	0.1735	0.0387	0.0054
-1	0.9230	0.5740	0.1870	0.0402
2	1	0.9400	0.6144	0.2317
5	1	1	0.9495	0.6557
8	1	1	0.9994	0.9492
11	1	1	1	0.9990
14	1	1	1	1
17	1	1	1	1
20	1	1	1	1

Le corrispondenti curve di rivelazione sono riportate di seguito:



Come era lecito attendersi, a parità di SNR la probabilità di detezione diminuisce all'aumentare del guadagno G (ossia al diminuire di P_{FA}).

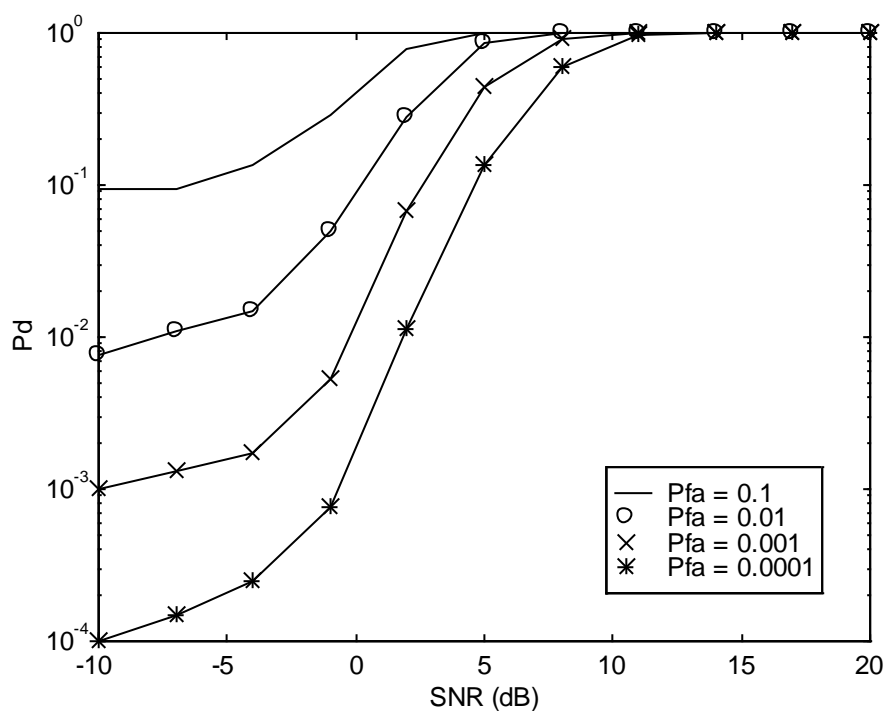
Delle curve analoghe vengono ottenute nel caso $M = 3$ (segnali PPS del terzo ordine).

Esempio 2:

I parametri siano gli stessi dell'esempio precedente, ad eccezione di M ($M=3$) ed N ($N=192$). La tabella con i valori della probabilità di detezione è la seguente:

CNR (dB)	$G=10^{1/4}-1$	$G=100^{1/4}-1$	$G=1000^{1/4}-1$	$G=10000^{1/4}-1$
-10	0.0950	0.0075	0.0010	0.0001
-7	0.0940	0.0110	0.0013	0.0001
-4	0.1340	0.0150	0.0017	0.0002
-1	0.2870	0.0500	0.0053	0.0008
2	0.7850	0.2780	0.0679	0.0114
5	0.9960	0.8550	0.4414	0.1340
8	1	0.9985	0.9260	0.6028
11	1	1	0.9994	0.9596
14	1	1	1	0.9998
17	1	1	1	1
20	1	1	1	1

a cui corrispondono le seguenti curve:



L'ultimo esempio riportato fa invece riferimento ad una simulazione con un autogate a 16 prese ($p = 16$).

Esempio 3:

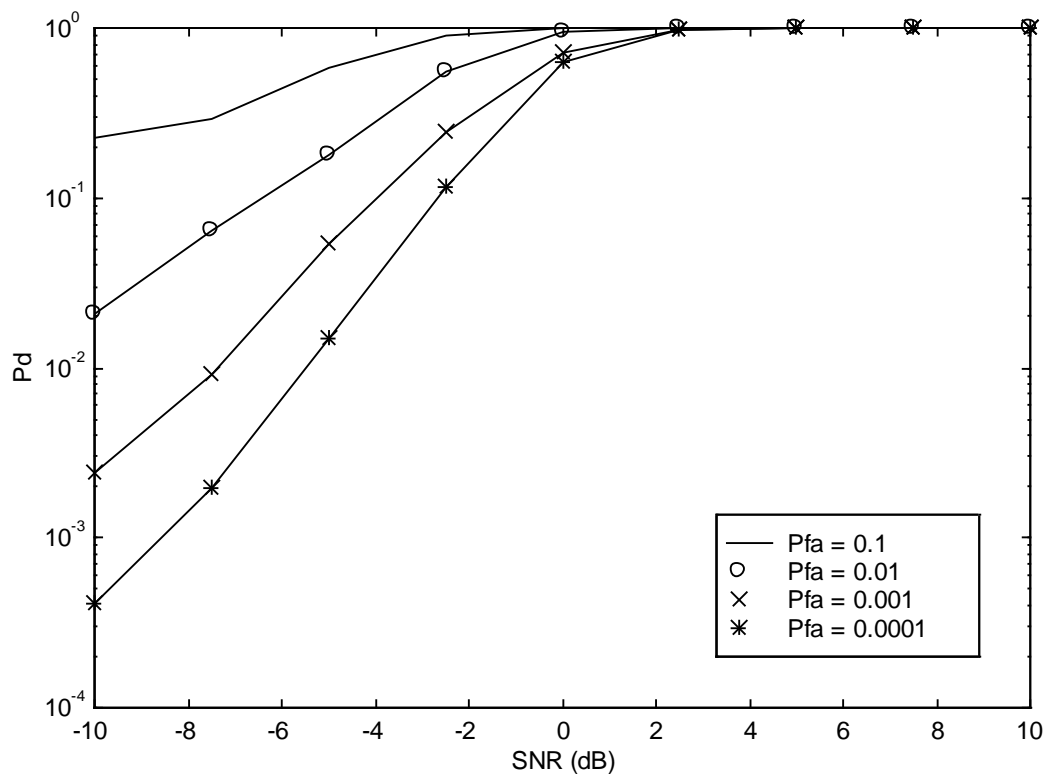
Sia:

$$M=2 \quad N=128 \quad L=1 \quad \tau=32 \quad p=16$$

Nella tabella che segue vengono riportati i valori di guadagno utilizzati nelle simulazioni, con le rispettive probabilità di falso allarme (teoriche e sperimentali):

G	P_{FA} teorica	P_{FA} sperimentale
0.1400	$1.2289 \cdot 10^{-1}$	$9.99 \cdot 10^{-2} \cong 10^{-1}$
0.3335	$1.0000 \cdot 10^{-2}$	$9.57 \cdot 10^{-3} \cong 10^{-2}$
0.5500	$9.0093 \cdot 10^{-4}$	$1.11 \cdot 10^{-3} \cong 10^{-3}$
0.8500	$5.3119 \cdot 10^{-5}$	$9.65 \cdot 10^{-5} \cong 10^{-4}$

Le curve di detezione sono riportate di seguito:



Si noti che per ottenere gli opportuni valori sperimentali di P_{FA} è stato necessario utilizzare dei valori di guadagno G leggermente diversi da quelli teorici previsti dalla (3.2.16); ciò in accordo con le osservazioni fatte al termine del paragrafo precedente, secondo le quali la validità della suddetta relazione diminuisce all'aumentare del numero di prese p dell'autogate.

3.4 Estensione al caso $L \geq 2$.

Lo studio condotto fino a questo punto ha avuto lo scopo di illustrare il principio di funzionamento dell'autogate nel caso in cui si adottasse un solo set di lags nel calcolo della PHAF ($L = 1$); in base alle osservazioni effettuate nel capitolo 1 risulta tuttavia lecito attendersi un miglioramento delle prestazioni all'aumentare del suddetto parametro.¹² Lo scopo del presente paragrafo è proprio quello di estendere le considerazioni precedentemente effettuate al caso $L \geq 1$.

Il primo problema che occorre affrontare è quello della determinazione del legame intercorrente fra la P_{FA} ottenuta ed i parametri G e p dell'autogate. Purtroppo nel caso in esame non risulta possibile individuare una semplice espressione in forma chiusa analoga alla (3.2.16). La PHAF è infatti definita nel modo seguente:

$$(3.4.1) \quad X_M^L \left(f; \underline{T}_{M-1}^L \right) = \prod_{h=1}^L X_M \left(\alpha_h \cdot f; \underline{\tau}_{M-1}^{(h)} \right)$$

Nell'ipotesi di elevato rapporto SNR le $X_M(\alpha_h \cdot f; \underline{\tau}_{M-1}^{(h)})$ possono essere ritenute gaussiane a valore atteso nullo (al di fuori del lobo principale), per cui come già visto i corrispondenti campioni del modulo quadro assumono una distribuzione esponenziale; essendo il modulo quadro della PHAF dato dai prodotti dei moduli quadri delle ml-HAF componenti, i campioni di $| X_M^L(f; \underline{T}_{M-1}^L)|^2$ risultano avere una distribuzione ottenibile come prodotto di L esponenziali. Ad esempio per $L = 2$ risulta:

$$(3.4.2) \quad p_{w_i}(w_i) = 2 \cdot \gamma_i \cdot K_0(2\sqrt{\gamma_i \cdot w_i})$$

dove:

$$(3.4.3) \quad w_i = | X_M^2(f_i; \underline{T}_{M-1}^2)|^2$$

$$(3.4.4) \quad K_0(z) = \int_0^{+\infty} e^{-z \cdot \cosh \xi} \cdot d\xi \quad (z > 0 - \text{fz. di Bessel modificata})$$

¹²Nel paragrafo 1.3 si è infatti osservato che all'aumento di L corrisponde, fra le altre cose, una diminuzione del rumore di sottofondo della PHAF. Intuitivamente ciò comporta una diminuzione della P_{FA} a parità di P_D , ovvero un aumento di P_D a parità di P_{FA} .

$$(3.4.5) \quad \gamma_i = E [|X_M^2(f_i; \underline{T}_{M-1}^2)|^2]^{-1}$$

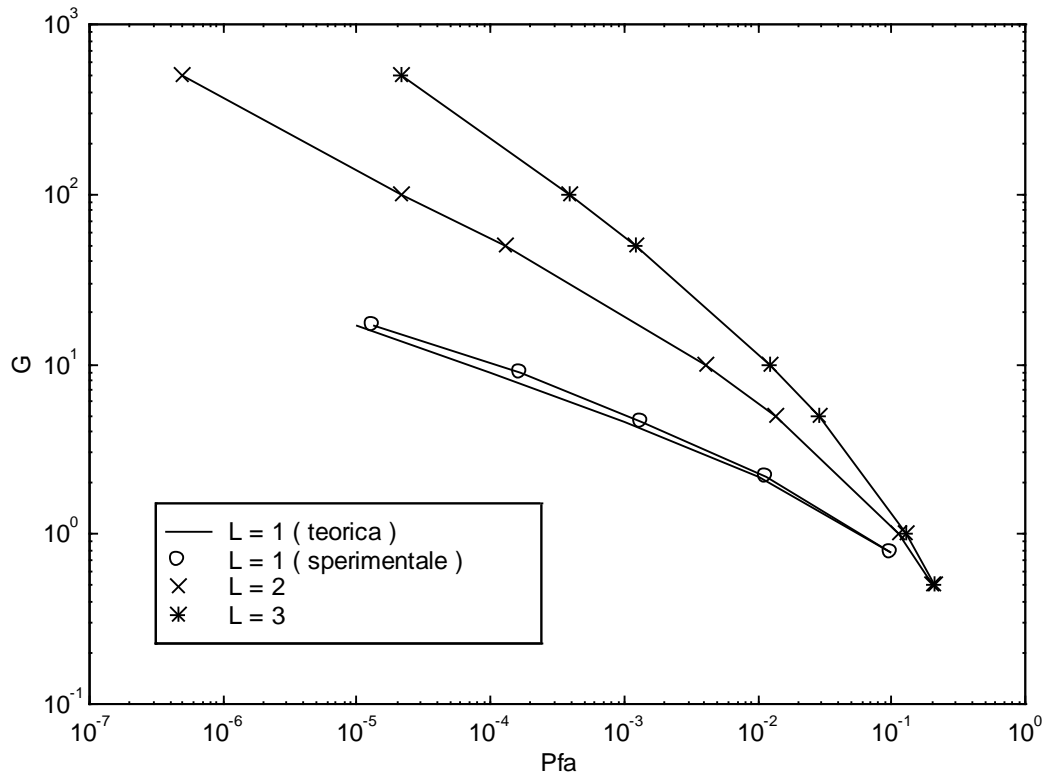
In modo analogo a quanto fatto nel caso $L = 1$ si definisce la soglia T nel modo seguente:

$$(3.4.6) \quad T = G \cdot X = G \cdot \sum_{i=1}^p w_i$$

A questo punto occorrerebbe trovare la densità di probabilità di X . Nell'ipotesi di indipendenza statistica fra le w_i (vera con ottima approssimazione) si avrebbe:

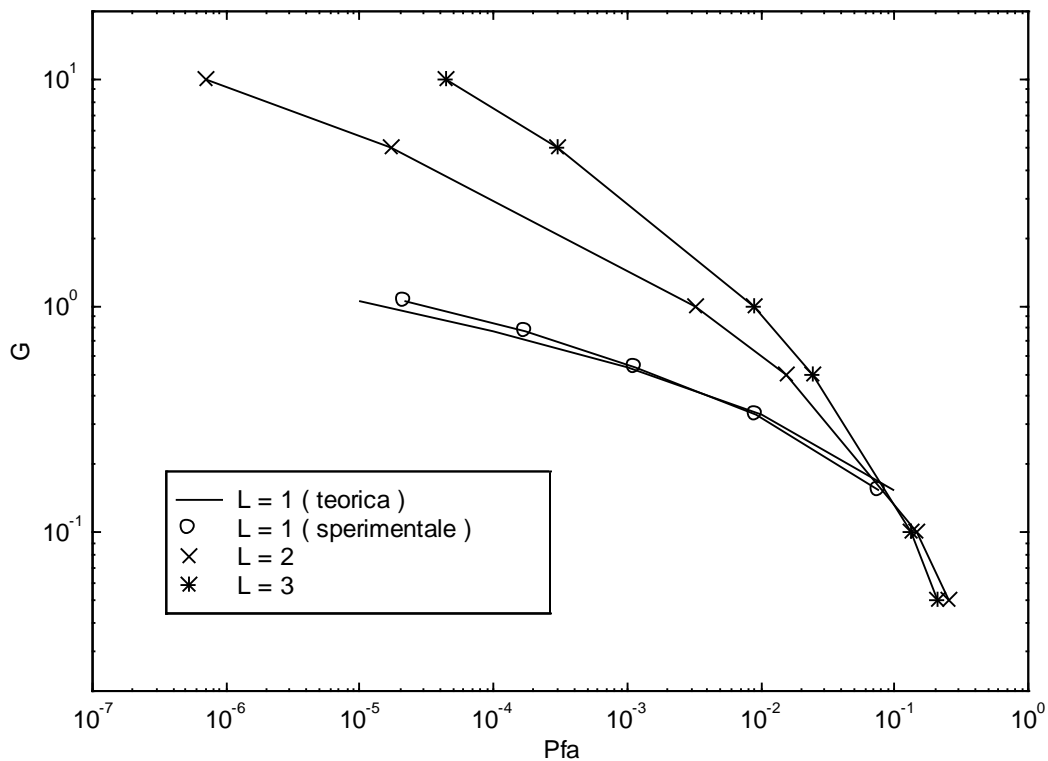
$$(3.4.7) \quad p_x(X) = p_{w1}(X) * p_{w2}(X) * \dots * p_{wp}(X)$$

Il problema discende dal fatto che non è possibile calcolare in forma chiusa l'espressione precedente, né direttamente, né attraverso il prodotto delle funzioni caratteristiche (non risulta nemmeno possibile determinare un'espressione per queste ultime). Anche procedere numericamente presenta dei problemi, sia a causa della dipendenza della (3.4.7) dai vari γ_i ($i=1,2,\dots,p$), sia a causa della singolarità nell'origine della (3.4.2). Non rimane quindi che l'approccio sperimentale; nei grafici seguenti sono riportate le curve del tipo $G = G(P_{FA})$ per un autogate a 4 prese nei casi $L = 1,2,3$.



$G = G(P_{FA})$ per $p = 4$

La curva continua, riferita al caso $L = 1$, riporta l'andamento teorico della P_{FA} in funzione di G (ottenuto dall'inversione della (3.2.16)); si noti il buon accordo fra quest'ultimo e l'andamento sperimentale. Curve analoghe si ottengono per un autogate a 16 prese, come appare evidente dal seguente diagramma:



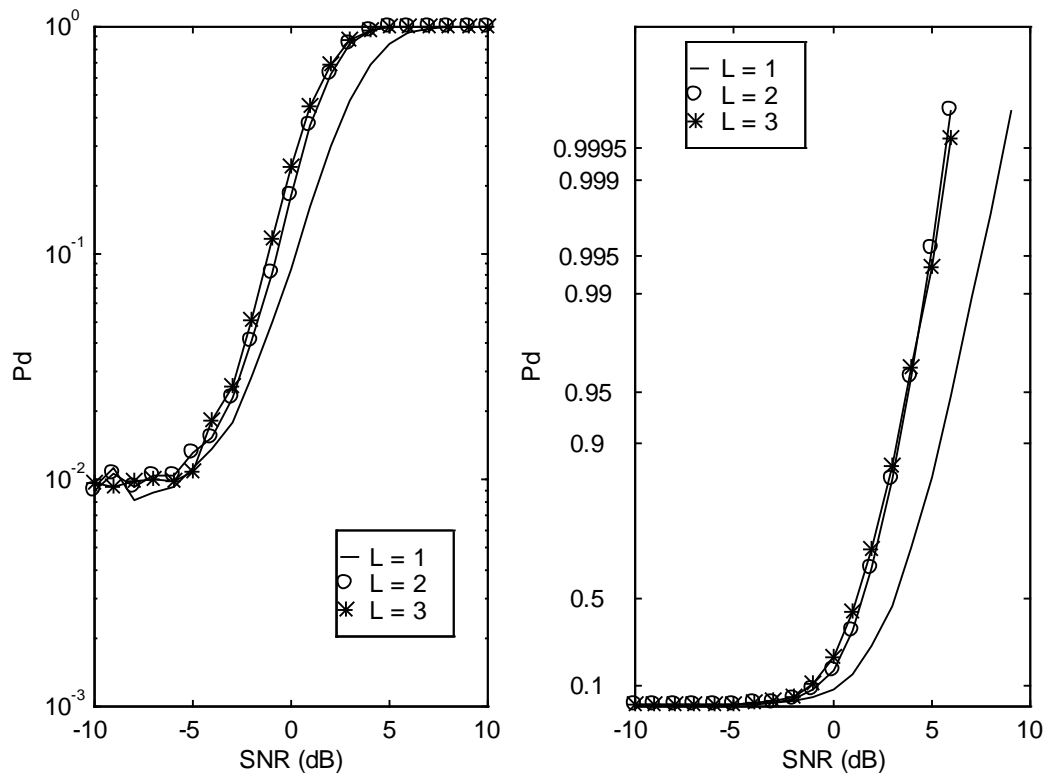
$G = G(P_{FA})$ per $p = 16$

Dopo aver stabilito il legame intercorrente fra i parametri dell'autogate e la corrispondente probabilità di falso allarme risulta possibile effettuare dei confronti fra le curve di P_D . A titolo di esempio nella tabella seguente vengono riportati i valori di guadagno G da utilizzare per conseguire P_{FA} pari rispettivamente a 10^{-2} , 10^{-3} , 10^{-4} (si suppone $p = 4$):

P_{FA}	$G (L=1)$	$G (L=2)$	$G (L=3)$
10^{-2}	$100^{1/4}-1 = 2.1623$	6	11
10^{-3}	$1000^{1/4}-1 = 4.6234$	20	60
10^{-4}	$10000^{1/4}-1 = 9.0000$	55	200

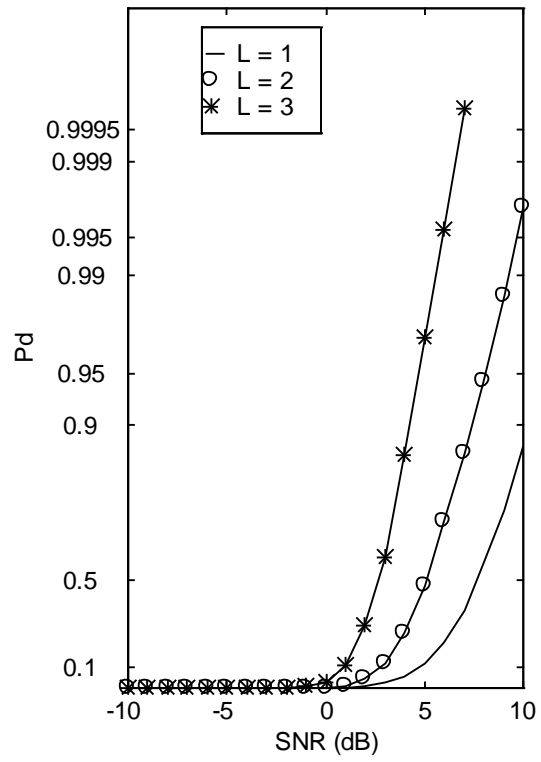
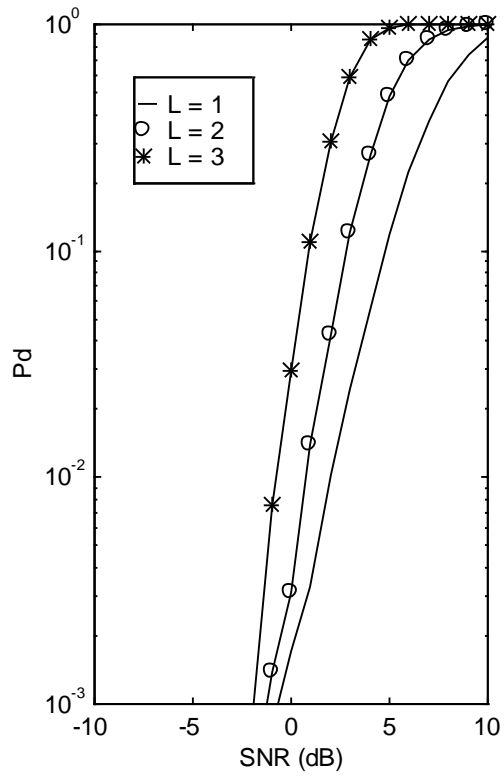
Nei grafici seguenti viene effettuato un confronto fra le curve di rivelazione all'aumentare di L . La prima figura riporta l'andamento della P_D al variare di SNR per una P_{FA} pari a 10^{-2} ; le tre curve sono relative ai casi $L = 1,2,3$. Allo scopo di evidenziare meglio il miglioramento conseguito, nel diagramma di destra le stesse curve vengono riportate nella forma normalmente adottata nella presentazione di risultati di questo tipo; essa viene ottenuta graficando la funzione $\text{erf}^{-1}(P_D)$, in modo da espandere la scala dell'asse verticale in prossimità dei valori unitari di P_D e da permettere quindi di quantificare in dB l'incremento di prestazioni ottenuto. Nel caso in esame il guadagno

rispetto alla curva $L = 1$ risulta di circa 3 dB, sia per $L = 2$ che per $L = 3$; si noti infatti la coincidenza fra le P_D ottenute in questi ultimi due casi.



$$P_{FA}=10^{-2}$$

Considerazioni del tutto analoghe valgono per una P_{FA} pari a 10^{-4} . Le curve corrispondenti sono riportate di seguito. Questa volta i guadagni ottenuti risultano decisamente più consistenti: si ha infatti un miglioramento di circa 4 dB passando da $L = 1$ ad $L = 2$, ed addirittura di 7 dB con $L = 3$ (ora, a differenza del caso precedente, si ottiene effettivamente un guadagno nel passaggio da $L = 2$ ad $L = 3$). Il fatto che l'incremento di prestazioni aumenti al diminuire di P_{FA} (ossia all'aumentare del guadagno G) non è casuale; al crescere di quest'ultimo parametro si ha infatti un innalzamento della soglia, che consente di sfruttare appieno la proprietà della PHAF di esaltare il picco principale rispetto a quelli spuri dovuti alla presenza del rumore.



$P_{FA}=10^{-4}$

Parte II

Applicazione dei segnali a fase polinomiale

(modulazione numerica PMSK)

4. Modulazione numerica PMSK

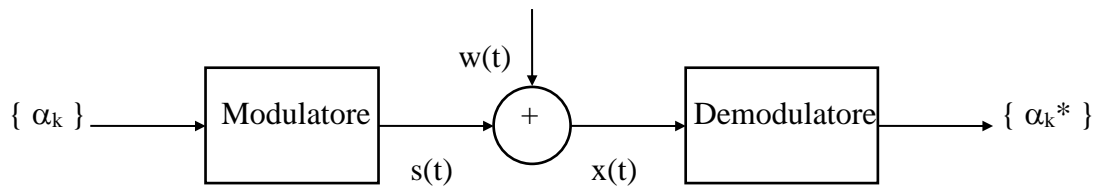
4.1 Introduzione.

Nella prima parte del presente lavoro è stata introdotta la nozione di segnale a fase polinomiale e sono stati quindi affrontati alcuni aspetti di carattere teorico ad essa relativi; si è discusso in modo particolare dei problemi della rivelazione e della stima dei parametri di segnali PPS immersi in rumore additivo Gaussiano bianco. Tali argomenti sono stati esaminati da un punto di vista prettamente teorico, prescindendo cioè dalle possibili applicazioni pratiche degli stessi. In questa seconda parte, invece, si passa ad affrontare con maggiore attenzione una fra le suddette applicazioni, relativa ad una forma di modulazione numerica indicata in letteratura (ad esempio in [7]) con la sigla PMSK (Polynomial Minimum Shift Keying). Il capitolo 4 è strutturato come segue: nel paragrafo 4.2 vengono presentati dei brevi richiami sulle modulazioni CPM in generale e sul PMSK in particolare. Nel paragrafo 4.3 viene illustrato il metodo di demodulazione basato sulla PHAF; nel paragrafo 4.4 viene quindi introdotta una variante del suddetto metodo, basata sull'applicazione dell'algoritmo di Viterbi. Infine il paragrafo 4.5 conclude il discorso confrontando le prestazioni ottenute con quelle dei metodi preesistenti.

4.2 Richiami sulle modulazioni numeriche con memoria: CPM e PMSK.

Le tecniche di modulazione numerica si prefiggono lo scopo di associare biunivocamente sequenze numeriche, costituite da simboli estratti da un alfabeto finito, a forme d'onda di caratteristiche adatte alla trasmissione su un opportuno canale. La biunivocità è richiesta per consentire l'operazione inversa di demodulazione, tramite la quale si risale dalla forma d'onda (segnale modulato) alla sequenza originaria¹³. La situazione è schematizzata nella figura seguente:

¹³Nella operazione di demodulazione interviene anche un problema di carattere decisionale, in quanto il segnale a monte del demodulatore è in generale affetto da rumore, alterazioni parametriche (sfasamenti, ritardi, ecc.) o addirittura distorsioni di varia natura. Nel seguito si supporrà di lavorare in presenza del solo rumore additivo; con altra terminologia si farà riferimento ad un problema di segnale noto in rumore additivo.



I simboli trasmessi vengono indicati con α_k ed appartengono ad un alfabeto discreto di dimensione M ; il segnale modulato viene indicato con $s(t)$, mentre quello ricevuto all'ingresso del demodulatore¹⁴ è dato da:

$$(4.2.1) \quad x(t) = s(t) + w(t)$$

essendo $w(t)$ un rumore additivo che viene supposto Gaussiano bianco. Il segnale modulato possiede in genere un'espressione del tipo seguente:

$$(4.2.2) \quad s(t) = \sum_k v(t - k \cdot T; \alpha_k, \sigma_k)$$

dove T rappresenta il periodo di simbolo e σ_k lo stato del modulatore nel k -esimo intervallo di simbolo. Il segnale $v(t - k \cdot T; \alpha_k, \sigma_k)$ viene detto segnale elementare; esso è definito nell'intervallo $k \cdot T \leq t \leq (k+1) \cdot T$ ed è nullo al di fuori. Pertanto il segnale modulato risulta dato dalla concatenazione di tanti segnali elementari consecutivi, uno per ciascun intervallo di simbolo; ognuno di essi dipende dal simbolo attuale α_k ed eventualmente da quelli precedenti attraverso lo stato σ_k . A tale proposito una prima fondamentale distinzione suddivide le tecniche di modulazione numerica in due categorie:

1. modulazione senza memoria

In questo caso la dipendenza dallo stato σ_k non sussiste, per cui il segnale elementare in ciascun intervallo di simbolo dipende esclusivamente dal simbolo corrispondente:

¹⁴Nella figura precedente il canale di trasmissione non è stato indicato in quanto si è supposto che esso non introduca distorsioni e che le eventuali alterazioni parametriche sulla forma d'onda trasmessa siano da ritenersi note.

$$(4.2.3) \quad s(t) = \sum_k v(t - k \cdot T; \alpha_k, \sigma_k) = \sum_k v(t - k \cdot T; \alpha_k)$$

In questa classe rientrano tutte le forme più semplici di modulazione (PAM, PSK, FSK, e via dicendo).

2. modulazione con memoria

Per questa seconda categoria di segnali vale più in generale la (4.2.2), per cui il segnale modulato nel k-esimo intervallo di simbolo assume un andamento che dipende sia da α_k che dallo stato σ_k (legato ai valori assunti dai simboli trasmessi in precedenza). Ciò rende più complessa l'operazione di demodulazione, ma viene ripagato da altri vantaggi, che si concretizzano solitamente in un aumento dell'efficienza spettrale.

Le modulazioni a fase continua (CPM = Continuous Phase Modulation) rientrano nella seconda classe e sono pertanto delle modulazioni con memoria. Un generico segnale CPM assume l'espressione:

$$(4.2.4) \quad s(t) = A \cdot e^{j2\pi \cdot h_F \sum_k \alpha_k q(t - kT)}$$

dove:

$$(4.2.5) \quad q(t) = \int_{-\infty}^t g(\tau) \cdot d\tau$$

e $g(t)$ è una generica forma di durata finita che prende il nome di impulso di shaping. La sua durata è normalmente un multiplo intero del periodo di simbolo e viene quindi posta uguale ad LT . L viene indicata con il nome di lunghezza di correlazione; nel caso $L=1$ si parla di modulazione CPM a risposta piena, mentre nel caso contrario essa viene detta a risposta parziale. Una convenzione usualmente seguita per motivi di normalizzazione è quella di scegliere $g(t)$ in modo che risulti:

$$(4.2.6) \quad \int_0^{LT} g(t) \cdot dt = \frac{1}{2}$$

Si noti che l'operazione di integrazione (4.2.5) ha l'effetto di rendere $q(t)$ continua, per cui la fase istantanea del segnale $s(t)$ assume un andamento continuo (cfr. (4.2.4)). In particolare nell'intervallo $KT \leq t \leq (K+1)T$ essa vale:

$$\begin{aligned}
 (4.2.7) \quad \phi(t; \underline{\alpha}) &= 2\pi h_F \int_{-\infty}^t \sum_k \alpha_k g(\tau - kT) d\tau = \\
 &= 2\pi h_F \alpha_K q(t - KT) + 2\pi h_F \sum_{k=K-L+1}^{K-1} \alpha_k q(t - kT) + \\
 &\quad + \pi h_F \sum_{k=-\infty}^{K-L} \alpha_k
 \end{aligned}$$

Come si può notare la fase istantanea dipende dal simbolo attuale α_K , dagli $(L-1)$ simboli precedenti ($\alpha_{K-1}, \alpha_{K-2}, \dots, \alpha_{K-L+1}$), i quali definiscono il cosiddetto stato correlativo, ed infine dagli infiniti simboli $\alpha_{K-L}, \alpha_{K-L-1}, \dots$, che individuano lo stato di fase.¹⁵ La proprietà di continuità della fase istantanea del segnale $s(t)$ giustifica il nome attribuito alla tecnica di modulazione esaminata.

Le modulazioni CPM possiedono alcune proprietà molto importanti: in primo luogo il loro inviluppo risulta costante, e questo le rende particolarmente adatte all'impiego su di un canale radio¹⁶. Inoltre la continuità della fase istantanea permette di ottenere un rapido decadimento dello spettro al di fuori della banda passante, e questo consente a sua volta di "avvicinare" i vari canali in frequenza conseguendo una migliore efficienza spettrale¹⁷.

La modulazione a fase polinomiale (PMSK = Polynomial Minimum Shift Keying) è una modulazione CPM a risposta parziale, nella quale l'impulso di shaping risulta definito come segue:

$$(4.2.8) \quad g_m(t) = \frac{1}{2} \mathfrak{F}^{-1} [\text{sinc}^m(\pi fT)]$$

¹⁵Lo stato del modulatore σ_k risulta definito dalla concatenazione dello stato correlativo e dello stato di fase.

¹⁶Ciò per due motivi: il primo è che in questo tipo di applicazioni il fading da cammini multipli agisce prevalentemente sull'ampiezza del segnale ricevuto, per cui non risulta opportuno racchiudere l'informazione da trasmettere sul suddetto parametro; il secondo è che normalmente i trasmettitori radio lavorano a piena potenza, ossia in prossimità della saturazione, per cui eventuali variazioni del segnale trasmesso potrebbero introdurre notevoli distorsioni.

¹⁷E' infatti noto che la rapidità di decadimento dello spettro al di fuori della banda passante risulta direttamente proporzionale all'ordine massimo di derivazione della fase istantanea del segnale modulato.

dove \mathfrak{F}^{-1} indica l'operatore di antitrasformazione secondo Fourier ed m prende il nome di ordine del segnale PMSK (esso coincide anche con la lunghezza di correlazione dell'impulso di shaping, come risulta facile verificare). La (4.2.8) è una funzione continua, di tipo polinomiale a tratti, per cui nel generico intervallo di simbolo ($(k-0.5)T \leq t \leq (k+0.5)T$) il segnale trasmesso assume l'espressione:

$$(4.2.9) \quad v(t - kT; \alpha_k, \sigma_k) = A \cdot e^{j2\pi \sum_{i=0}^m \beta_{k,i} (t - kT)^i}$$

(è infatti facile verificare che la fase in tale intervallo è un polinomio di grado m). Si può inoltre dimostrare¹⁸ che sussiste un legame lineare fra gli m coefficienti di grado più elevato del suddetto polinomio ed una m -upla di simboli consecutivi estratti dalla sequenza modulante. Nel caso generale tale legame può essere espresso nella forma:

$$(4.2.10) \quad \underline{\beta}_k = \underline{M} \cdot \underline{\alpha}_k$$

dove \underline{M} rappresenta una matrice quadrata non singolare di ordine m opportunamente costruita. Con riferimento al caso $m = 3$ si ha in particolare¹⁹:

$$(4.2.11) \quad \begin{bmatrix} \beta_{k,1} \\ \beta_{k,2} \\ \beta_{k,3} \end{bmatrix} = \frac{h_F}{2T^3} \begin{bmatrix} \frac{T^2}{8} & \frac{3T^2}{4} & \frac{T^2}{8} \\ \frac{T}{4} & 0 & \frac{T}{4} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} \alpha_{k-1} \\ \alpha_k \\ \alpha_{k+1} \end{bmatrix}$$

La demodulazione di un segnale PMSK immerso in rumore additivo Gaussiano bianco può quindi essere eseguita stimando dapprima per ciascun intervallo di simbolo gli m coefficienti di grado più elevato di (4.2.9) ed invertendo poi la relazione (4.2.10). Si noti che così facendo si ottengono m stime diverse di ciascun simbolo; per giungere alla sequenza demodulata occorrerà quindi combinarle opportunamente fra di loro. Tale problema verrà trattato più diffusamente nel prossimo

¹⁸Si veda a tale scopo [7].

¹⁹Si sceglie $m = 3$ perché l'impulso di shaping così ottenuto approssima molto bene l'andamento di quello utilizzato nel GMSK.

paragrafo. Per ora ci si limita ad elencare alcuni vantaggi che giustificano l'introduzione del PMSK: il primo di essi è legato alla modalità di implementazione del metodo, che risulta alquanto semplice essendo di fatto basata su delle FFT. Inoltre il metodo risulta Doppler tolerant²⁰ e sotto alcuni vincoli di banda consente anche di identificare alla cieca i parametri di un canale affetto da cammini multipli (l'identificazione avviene cioè senza la necessità di trasmettere sequenze di preambolo). Tutto ciò viene pagato essenzialmente con la necessità di sovracampionare il segnale e con l'incremento dell'indice di modulazione richiesto per ottenere prestazioni soddisfacenti.

²⁰Cfr. [7] ed [8].

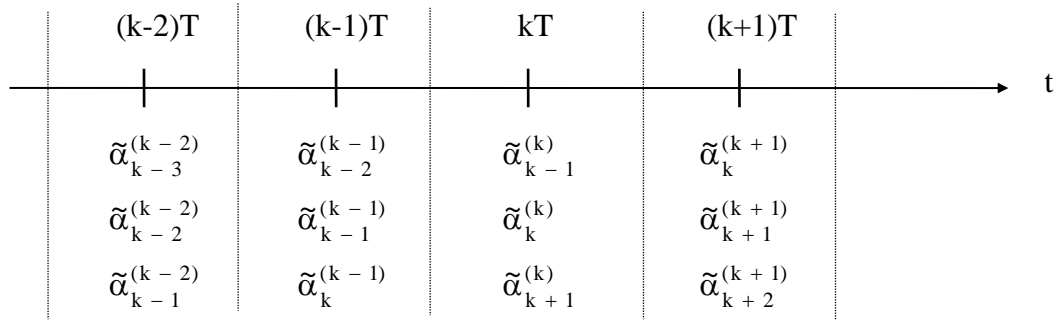
4.3 Demodulazione di segnali PMSK.

Nel paragrafo precedente è stata introdotta la modulazione PMSK; si è visto che essa appartiene alla classe delle modulazioni a fase continua, per le quali risultano già disponibili delle opportune tecniche di demodulazione ottima. Queste ultime sono tipicamente basate sull'applicazione del criterio MLSE (Maximum Likelihood Sequence Estimation) e conducono a delle strutture di demodulatori basati sull'algoritmo di Viterbi; il grande inconveniente di tali tecniche risiede nella notevole complessità realizzativa dei suddetti demodulatori, e nell'elevato costo che ne consegue. Nel presente paragrafo viene allora seguito un approccio completamente diverso al problema della demodulazione, che conduce a delle soluzioni di maggiore semplicità realizzativa, anche se sub-ottime dal punto di vista prestazionale.

Sia allora dato un segnale PMSK immerso in AWGN; per fissare le idee si faccia riferimento al caso di trasmissione di sequenze binarie ($\alpha_k = \pm 1$) e si supponga $m=3$, per cui:

$$(4.3.1) \quad x(t) = A e^{+j2\pi \sum_{i=0}^3 \beta_{k,i} (t-kT)^i} + w(t) \quad (k-0.5)T \leq t \leq (k+0.5)T$$

L'idea chiave alla base delle suddette tecniche sfrutta la caratteristica dei segnali PMSK di possedere una fase istantanea di tipo polinomiale a tratti; questa osservazione consente di adottare il metodo di stima basato sulla PHAF esaminato nella prima parte del presente lavoro. Si può quindi procedere nel modo seguente: per ciascun intervallo di simbolo si stimano i coefficienti del segnale (4.3.1), quindi si inverte la relazione (4.2.11) in modo da ottenere la corrispondente terna di simboli, e successivamente si combinano opportunamente le tre stime così ottenute di ciascun simbolo. Nella figura seguente è schematizzato tale procedimento:



La scrittura $\tilde{\alpha}_k^{(h)}$ indica la stima del k-esimo bit disponibile nell'h-esimo intervallo (quello centrato su hT); ovviamente $|k-h| \leq 1$. La stima dei coefficienti può essere effettuata adottando il metodo della PHAF; in base a quanto visto nella prima parte nell'ipotesi di SNR sufficientemente elevato lo stimatore risulta non polarizzato e possiede una matrice di covarianza approssimativamente coincidente con il limite inferiore di Rao-Cramer, dato da²¹:

$$(4.3.2) \quad \underline{J}^{-1} = \frac{1}{8\pi^2 \cdot h_F^2 \cdot \frac{E_b}{N_0}} \cdot \begin{bmatrix} 52320 & -21120 & 46560 \\ -21120 & 9120 & -21120 \\ 46560 & -21120 & 52320 \end{bmatrix}$$

Nella relazione precedente compare il rapporto E_b/N_0 (energia per bit diviso densità spettrale di rumore), che risulta legato al rapporto segnale/rumore sul singolo campione nel modo seguente:

$$(4.3.3) \quad E_b/N_0 = N \cdot A^2 / \sigma_w^2 = N \cdot \text{SNR}$$

dove N rappresenta il numero di campioni estratti in ciascun intervallo di bit di durata T ²². Lo scarto fra la matrice di covarianza effettiva ed il suddetto limite inferiore dipende da N e dal numero L di

²¹Il limite inferiore di Rao-Cramer (CRLB = Cramer-Rao Lower Bound) coincide con l'inversa della matrice di informazione di Fisher, definita come segue:

$$\underline{J} = \left\| \left\{ \mathbf{E} \left\{ \frac{\partial \log p(\underline{x}|\underline{\alpha})}{\partial \alpha_i} \cdot \frac{\partial \log p(\underline{x}|\underline{\alpha})}{\partial \alpha_j} \right\} \right\| \right\| \quad i, j = 1, 2, 3$$

La determinazione del limite inferiore di Rao-Cramer per segnali a fase polinomiale può essere trovata in [12].

²²Il periodo di campionamento è quindi dato da $T_s = T/N$. Essendo la banda del segnale PMSK pari a $B_{\text{PMSK}} = h_F/T$ è possibile definire una banda normalizzata (alla frequenza di campionamento) che vale:

set di lags utilizzati, e si aggira tipicamente nell'intorno di 2-3 dB. Per bassi SNR, invece, il metodo soffre del già citato inconveniente dovuto all'effetto soglia. Maggiori dettagli in proposito possono essere trovati in [5] e [6].

Dopo aver risolto il problema della stima dei parametri del segnale (4.3.1) ed avere di conseguenza ottenuto le terne di bit stimati sorge l'esigenza di combinare opportunamente tali stime. Esistono a tale proposito varie tecniche: l'operazione più semplice in assoluto consiste nel non combinarle affatto, limitandosi ad utilizzarne una sola fra le tre disponibili (ad esempio quella centrale, per la quale in base alla (4.3.2) è minima la varianza dell'errore di stima); è tuttavia evidente che così facendo le prestazioni ottenute non possono certamente essere ottimali. Una seconda alternativa è basata su un approccio più complesso; prima di illustrarla è però opportuno introdurre alcune definizioni preliminari. Siano allora rispettivamente:

$$(4.3.4) \quad \underline{\alpha}^{(k)} = \begin{bmatrix} \alpha_{k-1} \\ \alpha_k \\ \alpha_{k+1} \end{bmatrix}$$

$$(4.3.5) \quad \tilde{\underline{\alpha}}^{(k)} = \begin{bmatrix} \tilde{\alpha}_{k-1}^{(k)} \\ \tilde{\alpha}_k^{(k)} \\ \tilde{\alpha}_{k+1}^{(k)} \end{bmatrix}$$

$$(4.3.6) \quad \bar{\underline{\alpha}}^{(k)} = \begin{bmatrix} \bar{\alpha}_{k-1}^{(k)} \\ \bar{\alpha}_k^{(k)} \\ \bar{\alpha}_{k+1}^{(k)} \end{bmatrix}$$

la generica terna di bit trasmessi nel k-esimo intervallo di simbolo, la terna stimata attraverso il metodo della PHAF e l'inversione della (4.2.11), e la terna a minima distanza da quest'ultima, ossia quella per la quale risulta:

$$(4.3.7) \quad \bar{\underline{\alpha}}^{(k)} = \arg \min_{i=1..8} \sqrt{\left[\tilde{\underline{\alpha}}^{(k)} - \underline{\alpha}_i \right]^T \underline{C}_\alpha^{-1} \left[\tilde{\underline{\alpha}}^{(k)} - \underline{\alpha}_i \right]}$$

essendo \underline{C}_α la matrice di covarianza degli errori di stima. Il procedimento di combinazione in esame richiede che, in ciascun intervallo di simbolo, venga dapprima stimata la terna (4.3.5) e venga

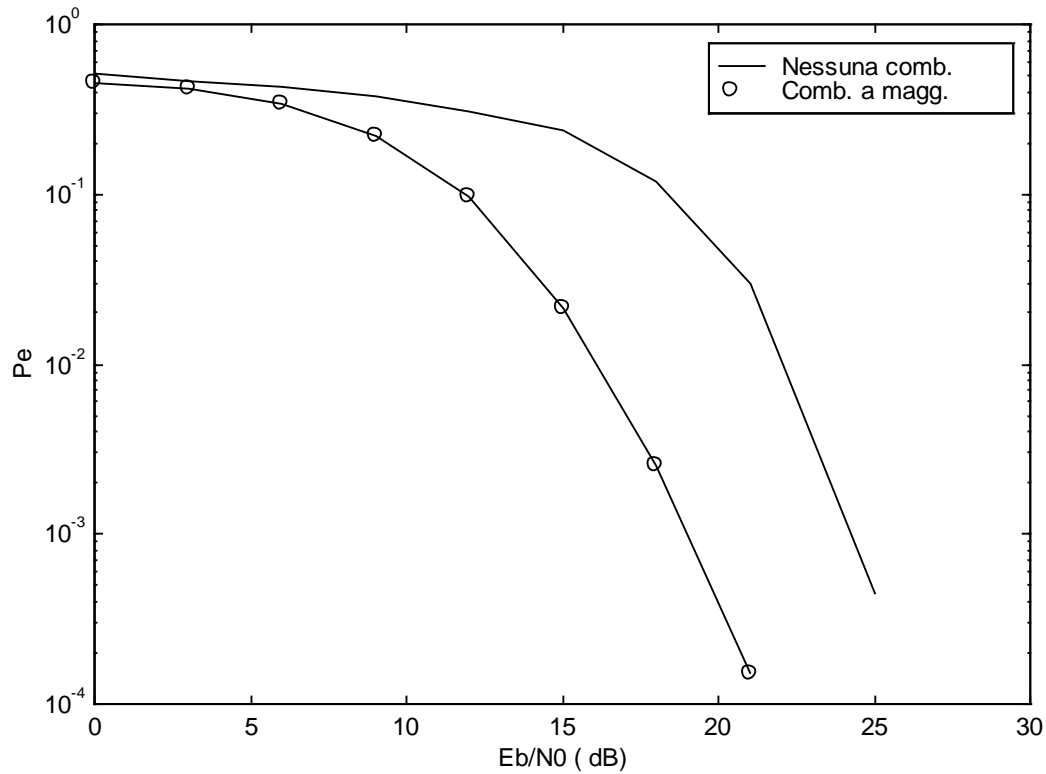
quindi ottenuta da essa la terna a minima distanza (4.3.6). Fatto ciò risulteranno disponibili tre stime del bit k -esimo (negli intervalli $(k-1)$ -esimo, k -esimo e $(k+1)$ -esimo); ciascuna di esse potrà valere $+1$ oppure -1 , e il bit demodulato verrà ottenuto semplicemente scegliendo fra i due valori quello che si presenta almeno due volte su tre. Nel seguito per semplicità si farà riferimento a questa tecnica con il nome di combinazione a maggioranza. Risulta lecito attendersi da essa delle prestazioni migliori rispetto a quelle del metodo precedente; ciò è confermato dai risultati delle simulazioni, come testimonia il seguente esempio.

Esempio 1:

Si prenda in esame un segnale PMSK del terzo ordine immerso in rumore additivo Gaussiano bianco. Risulti poi, con le notazioni introdotte in precedenza:

- numero di campioni per intervallo di bit $N = 16$
- banda normalizzata $B = 0.5$
- numero di set di lags nella PHAF $L = 1$

Nel grafico seguente sono riportate due curve; esse forniscono le probabilità di errore sul simbolo (che in questo caso coincidono con le probabilità di errore sul bit) al variare del rapporto E_b/N_0 .



Appare evidente il netto miglioramento del metodo di combinazione a maggioranza rispetto alla prima tecnica esaminata.

Nonostante il considerevole incremento di prestazioni conseguito attraverso la combinazione a maggioranza, questa non rappresenta ancora l'ottimo. Pur essendo infatti la tecnica di stima proposta "quasi" ottima nella stima all'interno di ciascun intervallo di bit, si ha una notevole perdita di prestazioni rispetto alla già citata tecnica MLSE nella combinazione su intervalli contigui. Un parziale recupero della suddetta perdita può però essere realizzato cambiando la modalità di combinazione delle terne stimate e ricorrendo ad un approccio basato su un algoritmo del tipo Viterbi, che verrà analizzato diffusamente nel paragrafo successivo.

4.4 Algoritmo di Viterbi per la combinazione delle stime su intervalli consecutivi.

Il problema che si vuole affrontare nel presente paragrafo è quello di migliorare le prestazioni del metodo di demodulazione di segnali PMSK esaminato nelle pagine precedenti; ciò verrà ottenuto attraverso l'adozione di una nuova tecnica di combinazione delle stime su intervalli di simbolo consecutivi basata sul noto algoritmo di Viterbi. Prima di procedere oltre è opportuno fornire alcune definizioni preliminari. Sia:

$$(4.4.1) \quad \underline{\alpha}' = \begin{bmatrix} -1 \\ -1 \\ \alpha_0 \\ \alpha_1 \\ \dots \\ \alpha_{M-1} \\ -1 \\ -1 \end{bmatrix}$$

la sequenza trasmessa (binaria); essa risulta composta da M+4 bit, dei quali gli M centrali costituiti da informazione utile. I primi 2 e gli ultimi 2 bit sono noti (nel caso in esame valgono -1) e vengono aggiunti alla sequenza da trasmettere per motivi che saranno chiari in seguito. Con riferimento ad una modulazione PMSK di ordine $m = 3$ possiamo allora definire lo stato del modulatore al passo k attraverso la coppia di bit²³:

$$(4.4.2) \quad \sigma_k = (\alpha_{k-1}, \alpha_k)$$

Esistono ovviamente 4 possibili stati, riportati nella tabella seguente:

²³Lo stato così definito coincide con quello che nel paragrafo 4.2 è stato definito stato correlativo. Nella definizione (4.4.2) non compare cioè lo stato di fase, legato agli infiniti simboli precedenti (che intervengono esclusivamente nel coefficiente di grado 0 del polinomio di fase, e che non vengono pertanto considerati nel metodo di stima esaminato nel paragrafo precedente).

Stato	(α_{k-1}, α_k)
$\sigma_k^{(0)}$	(-1, -1)
$\sigma_k^{(1)}$	(-1, +1)
$\sigma_k^{(2)}$	(+1, -1)
$\sigma_k^{(3)}$	(+1, +1)

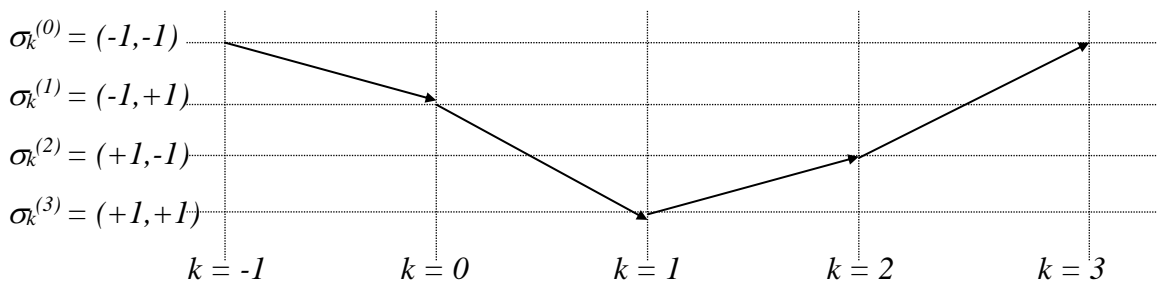
In corrispondenza dell'emissione del nuovo simbolo α_{k+1} si ha pertanto una transizione di stato:

$$(4.4.3) \quad \sigma_k^{(i)} = (\alpha_{k-1}, \alpha_k) \rightarrow \sigma_{k+1}^{(j)} = (\alpha_k, \alpha_{k+1})$$

L'emissione di una data sequenza del tipo (4.4.1) produce allora una serie di transizioni che possono essere rappresentate attraverso un opportuno diagramma a traliccio.

Esempio:

Nel diagramma seguente è rappresentato il cammino corrispondente all'emissione della sequenza $\underline{\alpha}' = [-1, -1, +1, +1, -1, -1]^T$:



Si noti che il cammino parte sempre dallo stato $\sigma_k^{(0)} = (-1, -1)$ e termina sempre nello stesso stato. Ciò è conseguenza delle due coppie di bit noti aggiunti in testa ed in coda alla sequenza utile. Si può anzi affermare che uno dei motivi della suddetta operazione è proprio quello di fare in modo che i cammini godano di questa proprietà, allo scopo di facilitare la successiva applicazione dell'algoritmo di demodulazione.

Si indichi ora con \underline{S}_K la seguente matrice di dimensioni $3 \cdot [K-2] \times K$:

$$(4.4.4) \quad \underline{S}_k = \begin{bmatrix} 1 & 0 & 0 & & & \\ 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \\ & 1 & 0 & 0 & & \\ & 0 & 1 & 0 & & \\ & 0 & 0 & 1 & & \\ & & & & 1 & 0 & 0 \\ & & & & 0 & 1 & 0 \\ & & & & 0 & 0 & 1 \end{bmatrix}$$

Premoltiplicando tale matrice per la sequenza (4.4.1) si ottiene il vettore:

$$(4.4.5) \quad \underline{\alpha} = \underline{S}_{M+4} \cdot \underline{\alpha}' = [-1 \ -1 \ \alpha_0 \ -1 \ \alpha_0 \ \alpha_1 \ \alpha_0 \ \alpha_1 \ \alpha_2 \ \dots \ \alpha_{M-1} \ -1 \ -1]^T =$$

$$= \begin{bmatrix} \underline{\alpha}^{(-1)} \\ \underline{\alpha}^{(0)} \\ \dots \\ \underline{\alpha}^{(M)} \end{bmatrix}$$

Esso ha lunghezza $3 \cdot [M+2]$, ed è formato nell'ordine da tutte le possibili terne di bit consecutivi che è possibile estrarre dalla sequenza trasmessa (4.4.1). Il vettore seguente è invece ottenuto concatenando le terne stimate su ciascun intervallo di bit:

$$(4.4.6) \quad \underline{\tilde{\alpha}} = \begin{bmatrix} \underline{\tilde{\alpha}}^{(-1)} \\ \underline{\tilde{\alpha}}^{(0)} \\ \dots \\ \underline{\tilde{\alpha}}^{(M)} \end{bmatrix}$$

Si indichi infine con:

$$(4.4.7) \quad \underline{\varepsilon}^{(k)} = \underline{\tilde{\alpha}}^{(k)} - \underline{\alpha}^{(k)}$$

la terna degli errori di stima compiuti al passo k, e con:

$$(4.4.8) \quad \underline{\varepsilon} = \begin{bmatrix} \varepsilon^{(-1)} \\ \varepsilon^{(0)} \\ \vdots \\ \varepsilon^{(M)} \end{bmatrix} = \underline{\tilde{\alpha}} - \underline{\alpha}$$

il vettore formato dagli errori di stima sull'intera sequenza. Ciò posto si può osservare che nell'ipotesi di alto SNR quest'ultimo può essere ritenuto approssimativamente Gaussiano, con valore atteso nullo e con matrice di covarianza data da:

$$(4.4.9) \quad \underline{C} = \begin{bmatrix} \underline{C}_\alpha & \cdot & \cdot & \cdot \\ \cdot & \underline{C}_\alpha & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \underline{C}_\alpha \end{bmatrix}$$

Essa è di ordine $3 \cdot [M+2]$ e risulta diagonale a blocchi in virtù dell'indipendenza statistica fra errori di stima compiuti su intervalli diversi; ciascun blocco è dato dalla matrice di covarianza per la singola terna, che può approssimativamente essere posta uguale al limite inferiore di Rao-Cramer (4.3.2) (ossia $\underline{C}_\alpha \cong \underline{J}^{-1}$). La densità di probabilità della sequenza stimata condizionata alla sequenza trasmessa è quindi data da:

$$(4.4.10) \quad p(\underline{\tilde{\alpha}} | \underline{\alpha}') = \frac{1}{\sqrt{(2\pi)^{3(M+2)} \det \underline{C}}} e^{-\frac{1}{2} [\underline{\tilde{\alpha}} - \underline{\alpha}]^T \underline{C}^{-1} [\underline{\tilde{\alpha}} - \underline{\alpha}]} =$$

$$= \frac{1}{\sqrt{(2\pi)^{3(M+2)} \det \underline{C}}} e^{-\frac{1}{2} [\underline{\tilde{\alpha}} - \underline{s}_{M+4} \cdot \underline{\alpha}']^T \underline{C}^{-1} [\underline{\tilde{\alpha}} - \underline{s}_{M+4} \cdot \underline{\alpha}']}$$

L'applicazione del criterio MV richiede l'individuazione di quella particolare sequenza $\underline{\alpha}'$ per la quale risulta massima la (4.4.10). Ciò equivale a minimizzare l'espressione all'esponente:

$$\begin{aligned}
 (4.4.11) \quad d^2 &= [\tilde{\underline{\alpha}} - \underline{\alpha}]^T \underline{C}^{-1} [\tilde{\underline{\alpha}} - \underline{\alpha}] = \\
 &= [\tilde{\underline{\alpha}} - \underline{S}_{M+4} \cdot \underline{\alpha}']^T \underline{C}^{-1} [\tilde{\underline{\alpha}} - \underline{S}_{M+4} \cdot \underline{\alpha}']
 \end{aligned}$$

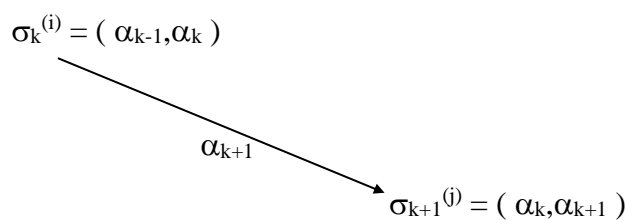
dove con d viene indicata la distanza fra $\tilde{\underline{\alpha}}$ e $\underline{\alpha}$ ²⁴ (la sequenza più verosimile è cioè quella a minima distanza). Introducendo nella precedente le definizioni (4.4.5), (4.4.6) e (4.4.9) si ricava:

$$(4.4.12) \quad d^2 = \sum_{k=-1}^M [\tilde{\underline{\alpha}}^{(k)} - \underline{\alpha}^{(k)}]^T \underline{C}_{\alpha}^{-1} [\tilde{\underline{\alpha}}^{(k)} - \underline{\alpha}^{(k)}]$$

Si noti che nella (4.4.12) non è possibile minimizzare separatamente i vari termini, i quali non risultano indipendenti a causa della particolare struttura del vettore $\underline{\alpha}$ ($\underline{\alpha} = \underline{S}_{M+4} \cdot \underline{\alpha}'$) che impone che le terne consecutive $\underline{\alpha}^{(k)}$ e $\underline{\alpha}^{(k+1)}$ possiedano due simboli in comune. Essa è tuttavia minimizzabile applicando l'algoritmo di Viterbi. A tale scopo si osservi dapprima che la transizione di stato (4.4.3):

$$(4.4.3) \quad \sigma_k^{(i)} = (\alpha_{k-1}, \alpha_k) \rightarrow \sigma_{k+1}^{(j)} = (\alpha_k, \alpha_{k+1})$$

viene rappresentata da un opportuno ramo nel diagramma a traliccio:



Supponendo di indicare con $\tilde{\underline{\alpha}}^{(k)} = [\tilde{\alpha}_{k-1}^{(k)} \quad \tilde{\alpha}_k^{(k)} \quad \tilde{\alpha}_{k+1}^{(k)}]^T$ la terna stimata al passo k è quindi possibile attribuire al suddetto ramo un costo pari al quadrato della distanza:

²⁴Trattasi di distanza in quanto la matrice \underline{C}^{-1} risulta definita positiva.

$$(4.4.13) \quad d_k^2 = \left(\underline{\tilde{\alpha}}^{(k)} - \begin{bmatrix} \alpha_{k-1} \\ \alpha_k \\ \alpha_{k+1} \end{bmatrix} \right)^T \underline{C}_{-\alpha}^{-1} \left(\underline{\tilde{\alpha}}^{(k)} - \begin{bmatrix} \alpha_{k-1} \\ \alpha_k \\ \alpha_{k+1} \end{bmatrix} \right)$$

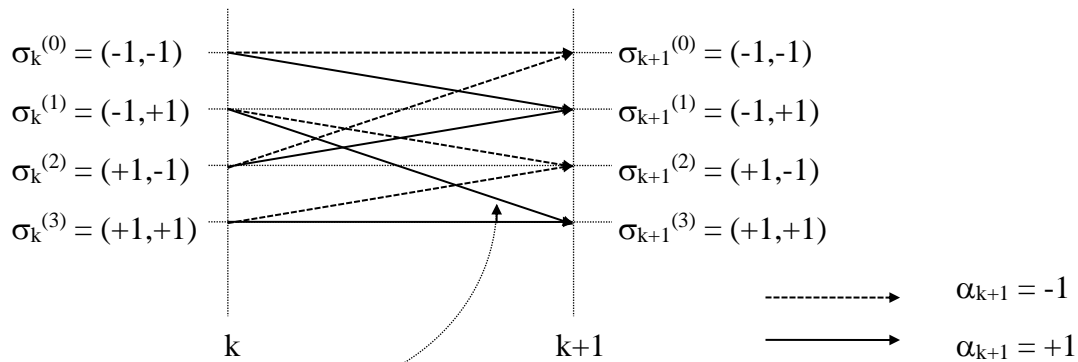
Un particolare cammino nel traliccio, dallo stato $\sigma_{-1}^{(0)}$ allo stato $\sigma_{h+1}^{(i)}$ ha pertanto un costo pari a:

$$(4.4.14) \quad d^2(h) = \sum_{k=-1}^h d_k^2$$

dove i singoli contributi sono dati dalla (4.4.13), avendo cura di scegliere ad ogni passo le terne $[\alpha_{k-1} \ \alpha_k \ \alpha_{k+1}]^T$ corrispondenti ai rami che compongono il percorso in esame. Poiché:

$$(4.4.15) \quad d^2 = d^2(M) = \sum_{k=-1}^M d_k^2$$

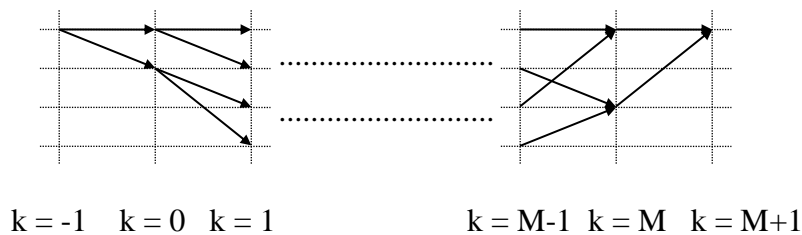
la minimizzazione della (4.4.15) equivale alla ricerca del percorso a costo minimo nel traliccio. L'algoritmo di Viterbi viene allora applicato nel modo seguente: al passo k sono memorizzati i 4 cammini a lunghezza minima che terminano in ciascuno dei 4 stati $\sigma_k^{(i)}$ ($i=0,1,2,3$), e con essi le rispettive lunghezze $l_i^{(k)}$. In corrispondenza della transizione $k \rightarrow k+1$ viene stimata la terna $\underline{\tilde{\alpha}}^{(k)}$ e vengono quindi attribuite le opportune metriche agli 8 rami associati a tutte le possibili transizioni del tipo suddetto; esse sono costituite da espressioni del tipo (4.4.13), nelle quali la terna $[\alpha_{k-1} \ \alpha_k \ \alpha_{k+1}]^T$ viene scelta in corrispondenza del ramo considerato:



Questo ramo possiede un costo dato da:

$$(\tilde{\alpha}^{(k)} - [-1 \ +1 \ +1]^T)^T \underline{C} \alpha^{-1} (\tilde{\alpha}^{(k)} - [-1 \ +1 \ +1]^T)$$

Dai quattro cammini sopravvissuti al passo k , di lunghezze $l_i^{(k)}$ ($i = 0,1,2,3$), vengono allora ottenuti otto nuovi cammini: le loro lunghezze sono determinate sommando opportunamente le metriche dei rami precedentemente calcolate alle lunghezze $l_i^{(k)}$ memorizzate fino al passo k . Gli otto cammini così ottenuti possono poi essere suddivisi in quattro gruppi da due; tra i due cammini di ciascun gruppo, che terminano nello stesso stato $\sigma_{k+1}^{(i)}$, viene quindi scelto quello di lunghezza minima, che viene memorizzato come cammino sopravvissuto assieme alla rispettiva lunghezza. Il procedimento si ripete quindi identico al passo $(k+1)$, e si procede in questo modo fino al termine. Da notare che tutti i cammini nel traliccio hanno inizio e termine nello stato 0, per cui i primi due e gli ultimi due passi vanno trattati in modo leggermente diverso (infatti nel primo e nell'ultimo passo si hanno esclusivamente due rami di transizione, mentre nel secondo e nel penultimo se ne hanno solo quattro):



Prima di passare all'esame delle prestazioni del metodo è opportuno fare alcune considerazioni. La prima riguarda la complessità dell'algoritmo: il numero totale di cammini è dato da:²⁵

$$(4.4.16) \quad N_{\text{CAMMINI}} = 2^M$$

²⁵Esso coincide infatti con il numero di diverse sequenze del tipo (4.4.1) che possono essere scelte.

per cui un procedimento di tipo esaustivo avrebbe un costo con dipendenza esponenziale dalla lunghezza della sequenza trasmessa. L'algoritmo di Viterbi possiede invece una complessità lineare. Assumendo infatti di attribuire un costo unitario alle operazioni di calcolo della lunghezza di un ramo e di confronto fra le lunghezze di due cammini, il costo globale dell'algoritmo risulta dato da:²⁶

$$(4.4.17) \quad C_{\text{Viterbi}} = 12M - 6$$

Una ulteriore precisazione va fatta in relazione alla particolare modalità di applicazione dell'algoritmo di Viterbi nel caso in esame. Esso viene infatti utilizzato esclusivamente per combinare in modo ottimo le stime dei bit ottenute tramite la PHAF su intervalli di simbolo consecutivi. Questo comporta una notevole semplificazione rispetto a quanto si otterrebbe applicando Viterbi nel modo classico. Si è già visto infatti che nel traliccio ottenuto compare esclusivamente lo stato correlativo;²⁷ se adottassimo il criterio MLSE classico, senza effettuare preliminarmente la stima attraverso la PHAF, avremmo anche la presenza dello stato di fase e ciò porterebbe inevitabilmente a dover lavorare su un traliccio molto più complesso. La semplificazione ottenuta è notevole, ma viene pagata da una perdita di prestazioni rispetto alle tecniche convenzionali. Questo ed altri aspetti saranno tuttavia affrontati nel paragrafo successivo, dedicato proprio allo studio delle prestazioni del metodo proposto.

²⁶Infatti si hanno 2 passi in cui si effettuano 2 calcoli di distanza ed 1 confronto (costo 3), 2 passi in cui si effettuano 4 calcoli di distanza e 2 confronti (costo 6) ed M-2 passi in cui si effettuano 8 calcoli di distanza e 4 confronti (costo 12). Sommando tutti i contributi si ottiene proprio la (4.4.17).

²⁷Questo è conseguenza del fatto che nella (4.2.9) l'unico coefficiente che dipende da tutti i bit trasmessi è quello di grado 0 ($\beta_{k,0}$). I restanti m, che sono quelli utilizzati nel metodo considerato, dipendono esclusivamente dal bit attuale e dagli (m-1) precedenti (vedi (4.2.10)).

4.5 Prestazioni.

L'esame delle prestazioni della tecnica di demodulazione esaminata nel paragrafo precedente non appare facilmente affrontabile da un punto di vista teorico, in quanto non risulta possibile esprimere in forma analitica chiusa la probabilità di errore in funzione del rapporto E_b/N_0 . Una valutazione approssimata del miglioramento conseguibile può essere tuttavia ottenuta esaminando il problema in modo indiretto, ossia ricorrendo alla nozione di distanza minima. Facendo riferimento in particolare al caso $m = 3$, e supponendo inoltre di operare in condizioni di elevato SNR, si considerino assegnate due sequenze di lunghezza N :

$$(4.5.1) \quad \underline{\alpha}_i' = [\alpha_1^{(i)}, \alpha_2^{(i)}, \dots, \alpha_N^{(i)}]^T$$

$$(4.5.2) \quad \underline{\alpha}_j' = [\alpha_1^{(j)}, \alpha_2^{(j)}, \dots, \alpha_N^{(j)}]^T \quad i, j = 1, 2, \dots, 2^N$$

Definiamo la loro distanza nel modo seguente:

$$\begin{aligned} (4.5.3) \quad d^2(i, j) &= [\underline{S}_N \cdot \underline{\alpha}_i' - \underline{S}_N \cdot \underline{\alpha}_j']^T \cdot \underline{C}^{-1} \cdot [\underline{S}_N \cdot \underline{\alpha}_i' - \underline{S}_N \cdot \underline{\alpha}_j'] = \\ &= [\underline{\alpha}_i' - \underline{\alpha}_j']^T \cdot \underline{S}_N^T \cdot \underline{C}^{-1} \cdot \underline{S}_N \cdot [\underline{\alpha}_i' - \underline{\alpha}_j'] = \\ &= [\underline{\alpha}_i' - \underline{\alpha}_j']^T \cdot \underline{C}_{tot} \cdot [\underline{\alpha}_i' - \underline{\alpha}_j'] \end{aligned}$$

dove:

$$(4.5.4) \quad \underline{C}_{tot} = \underline{S}_N^T \cdot \underline{C}^{-1} \cdot \underline{S}_N \quad (\text{matrice } N \times N)$$

e per \underline{S}_N e \underline{C} valgono le definizioni date nel paragrafo precedente. La distanza minima fra coppie di sequenze di lunghezza N è allora definita come segue:

$$(4.5.5) \quad d_{min}^2 = \min_{i \neq j} d^2(i, j)$$

Si osservi che in condizioni di elevato SNR si può porre:

$$(4.5.6) \quad \underline{C}_\alpha^{-1} \cong \underline{J} = 8\pi^2 \cdot h_F^2 \cdot \frac{E_b}{N_0} \cdot \begin{bmatrix} 1 & 11 & 31 \\ 1792 & 5040 & 80640 \\ 11 & 103 & 11 \\ 5040 & 10080 & 5040 \\ 31 & 11 & 1 \\ 80640 & 5040 & 1792 \end{bmatrix}$$

per cui, ricordando le (4.4.4) e (4.4.9) si ha:

$$(4.5.7) \quad \underline{C}_{\text{tot}} = \underline{S}_N^T \cdot \underline{C}^{-1} \cdot \underline{S}_N =$$

$$= \begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & \cdot & \cdot & \cdot & 0 \\ c_{12} & (c_{11} + c_{22}) & (c_{12} + c_{23}) & c_{13} & 0 & \cdot & \cdot & 0 \\ c_{13} & (c_{12} + c_{23}) & (c_{11} + c_{22} + c_{33}) & (c_{12} + c_{23}) & \cdot & \cdot & \cdot & \cdot \\ 0 & c_{13} & (c_{12} + c_{23}) & (c_{11} + c_{22} + c_{33}) & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & c_{13} & (c_{12} + c_{23}) & \cdot & \cdot & c_{13} & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & (c_{11} + c_{22} + c_{33}) & (c_{12} + c_{23}) & c_{13} \\ \cdot & \cdot & \cdot & \cdot & \cdot & (c_{12} + c_{23}) & (c_{11} + c_{22}) & c_{12} \\ 0 & 0 & \cdot & 0 & 0 & c_{13} & c_{12} & c_{11} \end{bmatrix}$$

(per semplicità nella precedente gli elementi della matrice $\underline{C}_\alpha^{-1}$ sono stati indicati con la notazione $c_{h,k}$). Sostituendo in essa i valori e calcolando la distanza minima secondo quanto detto in precedenza si ottiene:

$$(4.5.8) \quad d_{\min}^2 = \begin{cases} 8 \cdot \begin{pmatrix} c_{11} & -c_{13} \end{pmatrix} = \frac{1}{90} \cdot \pi^2 \cdot h_F^2 \cdot \frac{E_b}{N_0} & N = 3 \\ 4 \cdot c_{11} = \frac{1}{56} \cdot \pi^2 \cdot h_F^2 \cdot \frac{E_b}{N_0} & N \geq 4 \end{cases}$$

(nel caso $N = 3$ le coppie di sequenze a distanza minima sono affette da due errori di segno opposto sui due bit più esterni, mentre nel caso $N \geq 4$ esse differiscono per il primo oppure per l'ultimo bit). Il risultato precedente non sembra mostrare un grande guadagno nel passaggio da $N = 3$ (decisione sulla base del singola terna) ad $N \geq 4$ (decisione ottenuta combinando stime su intervalli consecutivi). Non si è però tenuto conto del fatto che per $N \geq 5$ (combinazione di almeno tre intervalli) si hanno meno di tre stime disponibili per i primi due e per gli ultimi due bit delle

sequenze $\underline{\alpha}_i'$ e $\underline{\alpha}_j'$. Volendo includere queste considerazioni nel calcolo della distanza minima occorre modificare la definizione delle suddette sequenze aggiungendo in testa ed in coda due coppie di bit noti (ad esempio -1)²⁸:

$$(4.5.9) \quad \underline{\alpha}_i' = [-1, -1, \alpha_1^{(i)}, \alpha_2^{(i)}, \dots, \alpha_N^{(i)}, -1, -1]^T$$

$$(4.5.10) \quad \underline{\alpha}_j' = [-1, -1, \alpha_1^{(j)}, \alpha_2^{(j)}, \dots, \alpha_N^{(j)}, -1, -1]^T \quad i, j = 1, 2, \dots, 2^N$$

Le due sequenze così ottenute hanno lunghezza $[N+4]$. La distanza minima viene quindi definita nel modo visto sopra, utilizzando però come espressione della matrice \underline{C}_{tot} la seguente:

$$(4.5.11) \quad \underline{C}_{tot} = \underline{S}_{N+4}^T \cdot \underline{C}^{-1} \cdot \underline{S}_{N+4} \quad (\text{matrice } [N+4] \times [N+4])$$

Poiché i primi e gli ultimi due bit valgono sempre -1, dal punto di vista del calcolo della distanza ciò equivale ad eliminare dalla matrice \underline{C}_{tot} le prime e le ultime due righe (e colonne). La matrice risultante, di dimensioni $N \times N$ ha pertanto la seguente struttura:

$$(4.5.12) \quad \underline{C}_{tot}' = \begin{bmatrix} (c_{11} + c_{22} + c_{33}) & (c_{12} + c_{23}) & c_{13} & \cdot \\ (c_{12} + c_{23}) & (c_{11} + c_{22} + c_{33}) & \cdot & \cdot \\ c_{13} & (c_{12} + c_{23}) & \cdot & (c_{12} + c_{23}) \\ \cdot & \cdot & (c_{12} + c_{23}) & (c_{11} + c_{22} + c_{33}) \end{bmatrix}$$

ed è una matrice di Toeplitz. Come conseguenza di ciò le coppie di sequenze a distanza minima sono quelle che differiscono per un solo bit, e per esse si ha:

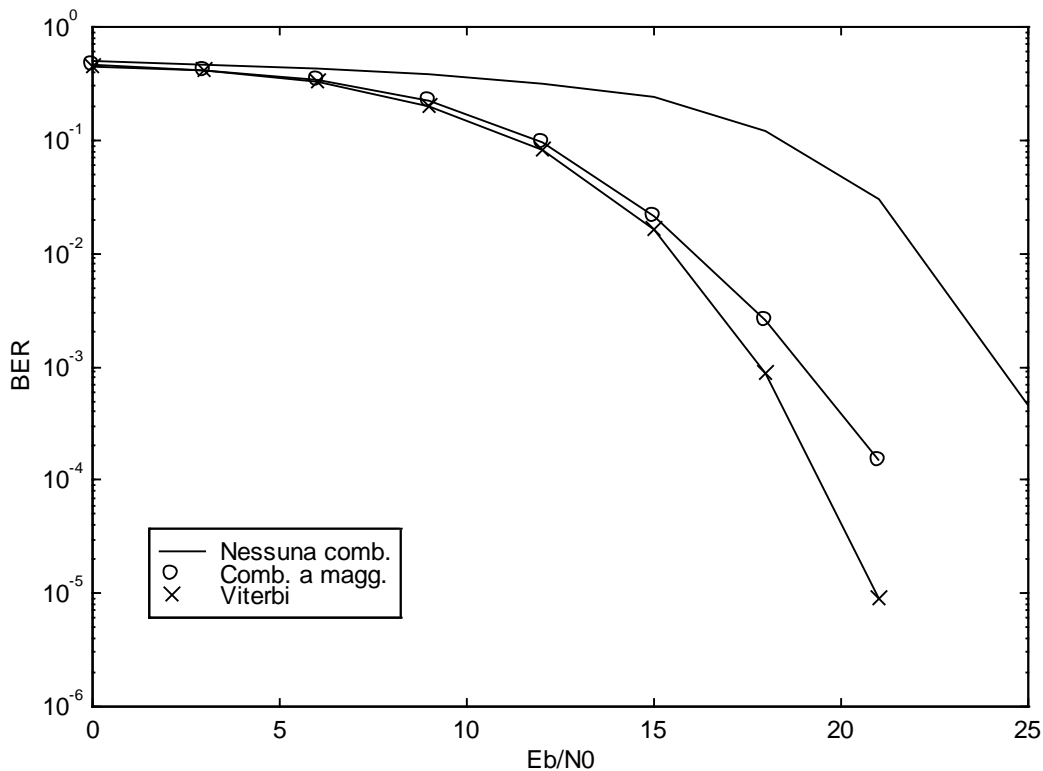
$$(4.5.13) \quad d_{min}^2 = 4 \cdot (c_{11} + c_{22} + c_{33}) = \frac{457}{1260} \cdot \pi^2 \cdot h_F^2 \cdot \frac{E_b}{N_0}$$

Come si può notare l'inclusione degli effetti transitori legati ai primi ed agli ultimi due bit consente di ottenere un valore di distanza minima notevolmente più elevato, che dovrebbe portare ad un miglioramento di prestazioni nell'intorno di $56 \cdot 457 / 1260 \cong 20 = 13$ dB. Purtroppo i risultati ottenuti per via sperimentale sembrano mostrare guadagni minori rispetto a quelli che sarebbe lecito attendersi in base alle considerazioni precedenti (si veda a tale proposito l'esempio seguente).

²⁸Risulta quindi chiarito anche il secondo motivo che porta alla introduzione delle due coppie di bit noti all'inizio ed alla fine della sequenza (cfr. def. (4.4.1)).

Esempio 2:

Si consideri nuovamente la situazione dell'esempio 1 (segnale PMSK di ordine $m = 3$ immerso in AWGN, con $N = 16$, $B = 0.5$ ed $L = 1$). Accanto alle curve già viste nell'esempio citato viene ora riportato l'andamento della probabilità di errore ottenuta tramite applicazione dell'algoritmo di Viterbi:



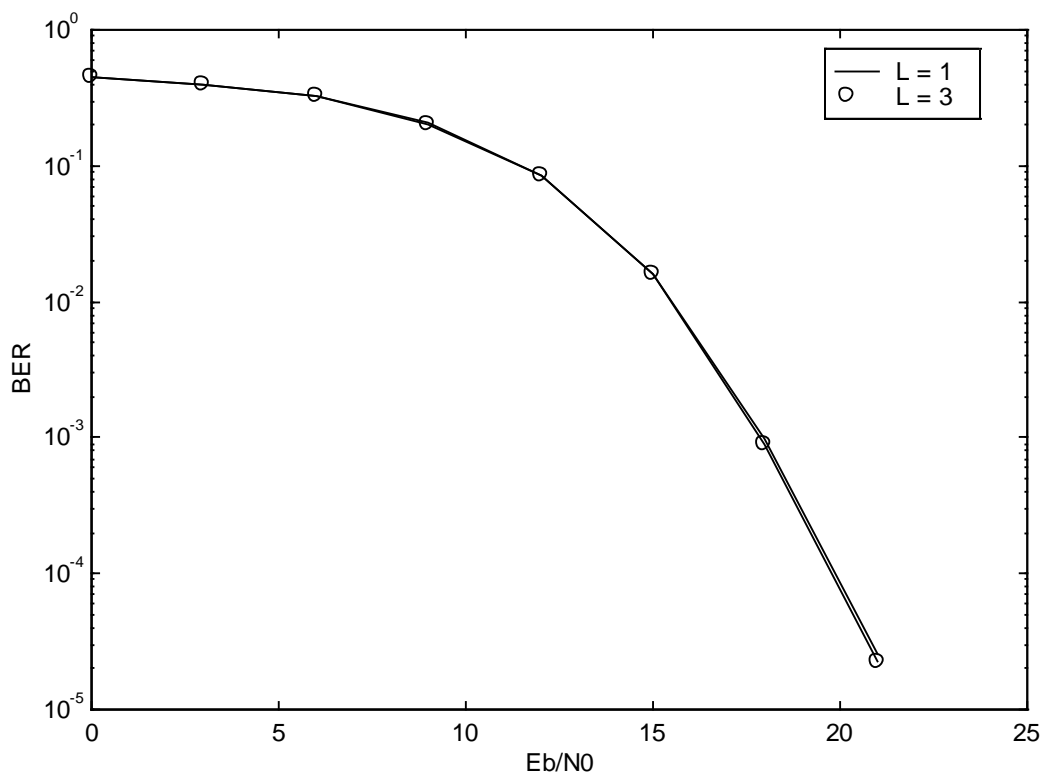
Dall'esame del grafico precedente si può dedurre che il miglioramento del suddetto algoritmo si presenta a partire da valori di rapporto E_b/N_0 abbastanza elevati (nel caso in esame da 17-18 dB in poi). Ciò è motivato dalla non linearità del metodo di stima e dal conseguente effetto soglia: si ricordi infatti che tutte le considerazioni effettuate allo scopo di introdurre l'algoritmo di Viterbi valgono nelle ipotesi in cui l'errore di stima sia Gaussiano, con vettore dei valori attesi nullo e matrice di covarianza approssimativamente coincidente con il limite inferiore di Rao-Cramer; tutto ciò risulta verificato solamente al di sopra della soglia, che nel caso in esame si presenta intorno ai 6-7 dB di rapporto SNR (equivalenti a 18-19 dB di rapporto E_b/N_0). Da osservare poi che l'incremento di prestazioni effettivamente conseguito sembra attestarsi sui 6-7 dB, e risulta pertanto notevolmente inferiore ai 13 dB ottenuti in precedenza. Anche questo fenomeno è imputabile al già citato effetto soglia, che limita la validità delle considerazioni precedenti. Per verificarne la bontà sarebbe quindi necessario considerare valori maggiori di E_b/N_0 ; purtroppo ciò

non risulta possibile per motivi di carattere pratico (infatti le probabilità di errore in gioco per tali valori di E_b/N_0 risultano molto piccole e richiedono quindi simulazioni su sequenze molto lunghe, e conseguentemente tempi di simulazione proibitivi).

Può essere interessante a questo punto valutare la dipendenza della probabilità di errore dal numero L di set di lags adottato e dal numero di campioni N estratti in ogni intervallo di bit. Per quanto riguarda la dipendenza da L sarebbe lecito attendersi semplicemente un abbassamento della soglia del metodo proposto; purtroppo anche in questo caso, per motivi analoghi a quelli esposti precedentemente, tale fenomeno non trova riscontro nelle simulazioni, come testimonia il seguente esempio.

Esempio 3:

Nel diagramma successivo sono riportate due curve; esse rappresentano le probabilità di errore ottenute applicando l'algoritmo di Viterbi nei due casi $L = 1$ ed $L = 3$. Gli altri parametri di simulazione sono $m = 3$, $N = 16$ e $B = 0.5$.

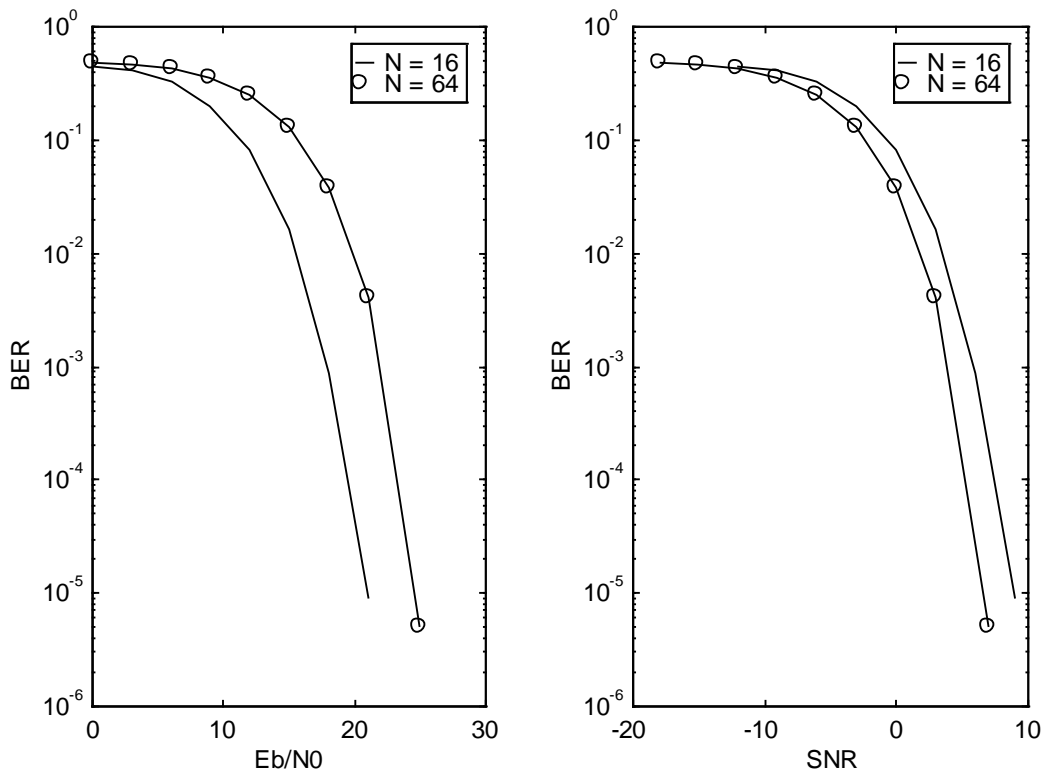


Come già preannunciato si può notare che fino a circa 20-21 dB le curve sono praticamente coincidenti (probabilmente si discostano a partire da valori superiori di rapporto E_b/N_0).

A proposito della dipendenza da N si può invece osservare quanto segue: la probabilità di errore in funzione di E_b/N_0 sembra aumentare al crescere di N , per cui un incremento di tale parametro appare a prima vista controproducente. In realtà effettuando il confronto a parità di rapporto SNR sul singolo campione la situazione si ribalta, come appare evidente dal seguente esempio.

Esempio 4:

Si supponga $m = 3$, $B = 0.5$ ed $L = 1$. Nei grafici seguenti vengono confrontate le probabilità di errore ottenute per simulazione nei due casi $N = 16$ ed $N = 64$; il confronto viene effettuato sia a parità di E_b/N_0 (grafico a sinistra) che a parità di SNR (grafico a destra)²⁹.

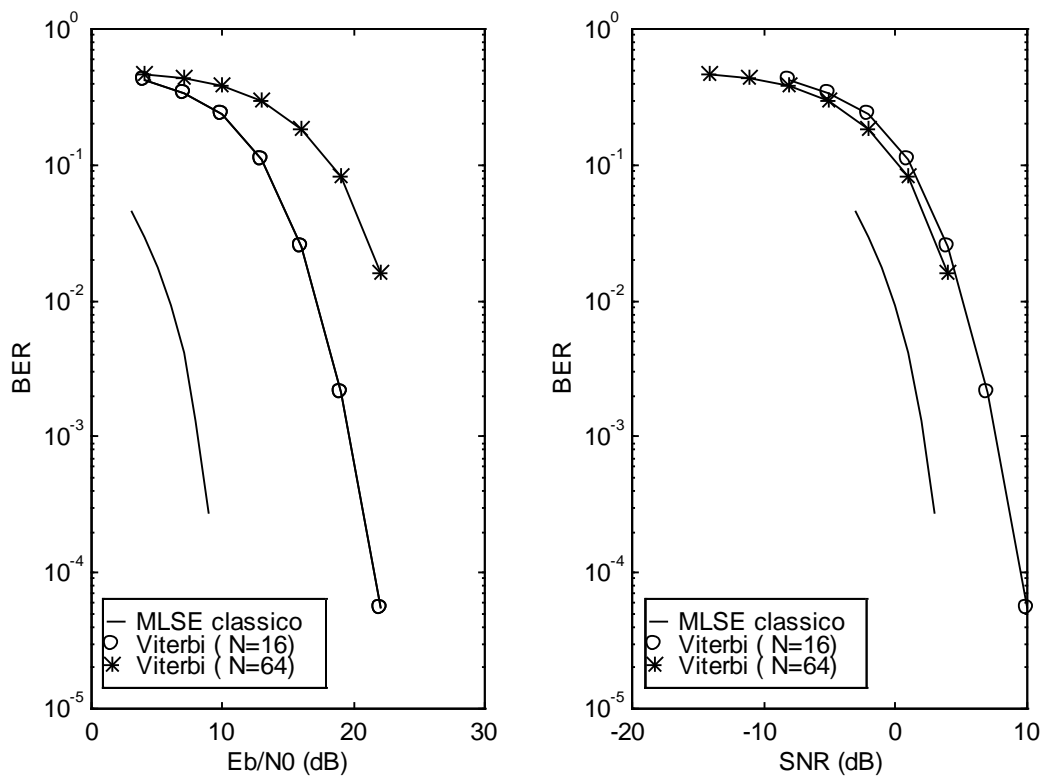


L'ultima osservazione da fare riguarda il confronto fra le prestazioni ottenute nel caso in esame con quelle ottenibili adottando tecniche convenzionali. Come già visto nel paragrafo precedente risulta lecito attendersi una certa perdita del metodo proposto rispetto al criterio MLSE classico. Ciò è confermato dall'esempio successivo.

²⁹Si ricordi che $E_b/N_0 = N \cdot \text{SNR}$.

Esempio 5:

Nei grafici seguenti sono riportate le curve di probabilità di errore per un segnale PMSK del terzo ordine con banda normalizzata $B = 0.25$. Nel diagramma di sinistra il confronto viene effettuato a parità di rapporto E_b/N_0 ; come si può notare la tecnica MLSE classica appare decisamente superiore a quella proposta, sia nel caso $N = 16$ che, a maggior ragione, in quello $N = 64$. L'entità dello scarto fra le curve riportate è in realtà dovuto alla somma di due contributi: il primo rappresenta l'effettiva perdita del metodo proposto rispetto alla tecnica MLSE ottima, mentre il secondo risulta legato al differente numero di campioni N adottato nei due casi ($N = 4$ nel caso MLSE, $N = 16$ oppure $N = 64$ nel caso di Viterbi). Allo scopo di eliminare quest'ultimo contributo e di valutare la perdita effettiva si può effettuare il confronto a parità di rapporto SNR sul singolo campione (grafico a destra). In tal modo la perdita effettiva si riduce a circa 5-6 dB.



Si noti che tale valore risulta poco influenzato dal numero di campioni estratti in ciascun intervallo di bit (N); un incremento di quest'ultimo parametro non comporta infatti grandi miglioramenti in termini di prestazioni.

A conclusione delle precedenti considerazioni possiamo osservare quanto segue; la tecnica MLSE convenzionale fornisce certamente prestazioni ottimali, ma viene pagata con la necessità di adottare strutture di demodulatori molto complesse dal punto di vista realizzativo. La tecnica

proposta nel presente capitolo, a fronte di una serie di svantaggi che si traducono sostanzialmente in una modesta perdita di prestazioni, possiede però alcuni innegabili vantaggi che la rendono molto interessante: il primo fra questi risulta dato dalla semplicità di implementazione circuitale, certamente non trascurabile in vista di possibili applicazioni in settori quali quello delle comunicazioni mobili di massa, per il quale tale tecnica risulta particolarmente indicata. Oltre a ciò occorre tenere conto del comportamento della suddetta tecnica in un ambiente radio reale, nel quale la presenza di cammini multipli o di una frequenza doppler influisce notevolmente sulle prestazioni ottenute: da questo punto di vista le tecniche classiche, basate sull'uso di correlatori, falliscono completamente. Il metodo di stima basato sulla PHAF risulta invece molto robusto: è infatti possibile dimostrare³⁰ che la presenza di una frequenza doppler conduce semplicemente ad uno shift dei simboli stimati, che non influisce eccessivamente sulle prestazioni ottenute. In ultimo occorre ricordare la possibilità di equalizzazione alla cieca che rende possibile, sotto opportuni vincoli di banda, la stima dei parametri di un canale reale senza la necessità di trasmettere sequenze di preambolo. Tutti questi vantaggi rendono sicuramente degna di interesse la tecnica proposta e ne fanno una valida alternativa ai metodi classici di demodulazione.

³⁰Si veda ad esempio [8].

Appendici

Appendice A1

Determinazione della varianza della ml-HAF in assenza di segnale.

Nel paragrafo 2.2 è stata definita la varianza della ml-HAF di una sequenza di campioni di rumore Gaussiano bianco; la definizione corrispondente viene qui riportata per comodità:

$$\begin{aligned}
 (A1.1) \quad \sigma_M^2(f) &= E [|X_M(f; \underline{\tau}_{M-1})|^2] = E [X_M(f; \underline{\tau}_{M-1}) \cdot X_M^*(f; \underline{\tau}_{M-1})] = \\
 &= \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [m-n]} \cdot E \left\{ \prod_{k=1}^{2^{M-1}} w^{*P(k)}[n+Z_M(k)] \cdot w^{*[-P(k)]}[m+Z_M(k)] \right\}
 \end{aligned}$$

Come già accennato la precedente non risulta esprimibile in forma chiusa. Nella presente appendice il calcolo viene quindi particolarizzato ai casi $M=2$ ed $M=3$.

1° caso: $M=2$

$$\begin{aligned}
 (A1.2) \quad \sigma_2^2(f) &= E [|X_2(f; \tau)|^2] = E [X_2(f; \tau) \cdot X_2^*(f; \tau)] = \\
 &= \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [m-n]} \cdot E [w(n+\tau) \cdot w^*(n-\tau) \cdot w^*(m+\tau) \cdot w(m-\tau)]
 \end{aligned}$$

Sfruttando le note proprietà sui momenti di una Gaussiana il valore atteso può essere espresso nel modo seguente:

$$\begin{aligned}
 (A1.3) \quad E [w(n+\tau) \cdot w^*(n-\tau) \cdot w^*(m+\tau) \cdot w(m-\tau)] &= \\
 &= R_w(2\tau) \cdot R_w(-2\tau) + R_w(n-m) \cdot R_w(m-n) = \sigma_w^4 \cdot \delta(n-m) \cdot \delta(m-n) = \\
 &= \sigma_w^4 \cdot \delta(n-m)
 \end{aligned}$$

Sostituendo (A1.3) in (A1.2) e tenendo presente che $\delta(n-m) = 1$ per $n=m$ e 0 altrove si ricava:

$$(A1.4) \quad \sigma_2^2(f) = \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [m-n]} \cdot \sigma_w^4 \cdot \delta(n-m) = \sigma_w^4 \cdot [N-2 \cdot \tau]$$

2° caso: M=3

Rispetto al caso precedente la complicazione analitica è maggiore, pur non essendoci differenze dal punto di vista concettuale:

$$(A1.5) \quad \begin{aligned} \sigma_3^2(f) &= E [|X_3(f; \tau_1, \tau_2)|^2] = E [X_3(f; \tau_1, \tau_2) \cdot X_3^*(f; \tau_1, \tau_2)] = \\ &= \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [m-n]} \cdot E [w(n+\tau_1+\tau_2) \cdot w(n-\tau_1-\tau_2) \cdot w^*(n+\tau_1-\tau_2) \cdot \\ &\quad \cdot w^*(n-\tau_1+\tau_2) \cdot w^*(m+\tau_1+\tau_2) \cdot w^*(m-\tau_1-\tau_2) \cdot w(m+\tau_1-\tau_2) \cdot w(n-\tau_1+\tau_2)] \end{aligned}$$

Quest'ultimo valore atteso può essere espanso nella somma di 105 prodotti, 101 dei quali si annullano³¹; i restanti 4 forniscono:

$$(A1.6) \quad \begin{aligned} \sigma_3^2(f) &= E [|X_3(f; \tau_1, \tau_2)|^2] = \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [m-n]} \cdot \\ &\quad \cdot \{ R_w^4[n-m] + R_w^2[n-m] \cdot R_w[n-m+2 \cdot (\tau_1-\tau_2)] \cdot R_w[n-m-2 \cdot (\tau_1-\tau_2)] + \\ &\quad + R_w^2[n-m] \cdot R_w[n-m+2 \cdot (\tau_1+\tau_2)] \cdot R_w[n-m-2 \cdot (\tau_1+\tau_2)] + \\ &\quad + R_w[n-m+2 \cdot (\tau_1+\tau_2)] \cdot R_w[n-m-2 \cdot (\tau_1+\tau_2)] \cdot R_w[n-m+2 \cdot (\tau_1-\tau_2)] \cdot \\ &\quad \cdot R_w[n-m-2 \cdot (\tau_1-\tau_2)] \} = \\ &= \sigma_w^8 \cdot \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [m-n]} \cdot \delta(n-m) \cdot [1 + \delta(\tau_1-\tau_2)] = \\ &= \sigma_w^8 \cdot [N-2 \cdot (\tau_1+\tau_2)] \cdot [1 + \delta(\tau_1-\tau_2)] \end{aligned}$$

³¹ Infatti 81 di questi 105 prodotti contengono fattori del tipo $E[w(\cdot) \cdot w(\cdot)]$ oppure $E[w^*(\cdot) \cdot w^*(\cdot)]$, che risultano essere nulli; altri 20 invece contengono termini del tipo $R_w(2\tau_1)$, $R_w(2\tau_2)$ oppure $R_w(\tau_1+\tau_2)$, che valgono 0 in conseguenza delle ipotesi fatte sulla sequenza di rumore $w(n)$.

Appendice A2

Indipendenza statistica fra ml-HAF ottenute con set di lags diversi

Lo scopo della presente appendice è quello di provare l'indipendenza statistica fra 2 ml-HAF calcolate sulla stessa sequenza di rumore adottando diversi set di lags. A tal fine è sufficiente verificare che sia nulla la covarianza incrociata fra $X_M(f; \underline{\tau}_{M-1})$ e $X_M(f; \underline{\nu}_{M-1})$ ³². Non essendo però possibile procedere in termini generali si esamineranno separatamente i due casi di interesse pratico ($M=2$ ed $M=3$).

1° caso: $M = 2$

La covarianza incrociata è data da:

$$\begin{aligned} (A2.1) \quad E [X_2^*(f; \tau) \cdot X_2(f; \nu)] &= \\ &= \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [n-m]} \cdot E [w^*(n+\tau) \cdot w(n-\tau) \cdot w(m+\nu) \cdot w^*(m-\nu)] = \\ &= \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [n-m]} \cdot \{ R_w(-2\tau) \cdot R_w(2\nu) + R_w(n-m+\Delta) \cdot R_w(m-n+\Delta) \} = \\ &= 0 \end{aligned}$$

se $\Delta = \nu - \tau \neq 0$ ($\Leftrightarrow \nu \neq \tau$). Da notare che se una delle due ml-HAF viene opportunamente scalata in frequenza, come accade ad esempio quando viene definita la PHAF, il risultato precedente non cambia; in altri termini, scelto $\alpha \neq 1$, anche $X_2(f; \tau)$ e $X_2(\alpha \cdot f; \nu)$ risultano indipendenti, come è facile verificare.

2° caso: $M = 3$

³² Infatti le due ml-HAF sono Gaussianhe, per cui la decorrelazione equivale all'indipendenza statistica.

L'indipendenza statistica fra $X_3(f; \tau_1, \tau_2)$ e $X_3(f; \upsilon_1, \upsilon_2)$ può essere provata procedendo in modo analogo. Si indichi con α il fattore di scala definito attraverso i rapporti dei lags. Si ha allora:

$$\begin{aligned}
 (A2.2) \quad E [X_3^*(f; \tau_1, \tau_2) \cdot X_3(\alpha \cdot f; \upsilon_1, \upsilon_2)] &= \\
 &= \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [n - \alpha \cdot m]} \cdot E [w^*(n + \tau_1 + \tau_2) \cdot w^*(n - \tau_1 - \tau_2) \cdot w(n + \tau_1 - \tau_2) \cdot w(n - \tau_1 + \tau_2) \cdot \\
 &\quad \cdot w(m + \upsilon_1 + \upsilon_2) \cdot w(m - \upsilon_1 - \upsilon_2) \cdot w^*(m + \upsilon_1 - \upsilon_2) \cdot w^*(m - \upsilon_1 + \upsilon_2)]
 \end{aligned}$$

Attraverso un procedimento analogo a quello utilizzato nell'appendice A1 per il calcolo della varianza³³ si giunge alla seguente espressione:

$$\begin{aligned}
 (A2.3) \quad E [X_3^*(f; \tau_1, \tau_2) \cdot X_3(\alpha \cdot f; \upsilon_1, \upsilon_2)] &= \\
 &= \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [n - \alpha \cdot m]} \cdot \{ R_w[(m-n) + (\upsilon_1 - \tau_1) + (\upsilon_2 - \tau_2)] \cdot \\
 &\quad \cdot R_w[(m-n) - (\upsilon_1 - \tau_1) - (\upsilon_2 - \tau_2)] \cdot R_w[(m-n) + (\upsilon_1 - \tau_1) - (\upsilon_2 - \tau_2)] \cdot \\
 &\quad \cdot R_w[(m-n) - (\upsilon_1 - \tau_1) + (\upsilon_2 - \tau_2)] + \\
 &\quad + R_w[(m-n) + (\upsilon_1 - \tau_1) + (\upsilon_2 - \tau_2)] \cdot R_w[(m-n) - (\upsilon_1 - \tau_1) - (\upsilon_2 - \tau_2)] \cdot \\
 &\quad \cdot R_w[-(m-n) + (\upsilon_1 + \tau_1) - (\upsilon_2 + \tau_2)] \cdot R_w[-(m-n) - (\upsilon_1 + \tau_1) + (\upsilon_2 + \tau_2)] + \\
 &\quad + R_w[(m-n) - (\upsilon_1 + \tau_1) - (\upsilon_2 + \tau_2)] \cdot R_w[(m-n) + (\upsilon_1 + \tau_1) + (\upsilon_2 + \tau_2)] \cdot \\
 &\quad \cdot R_w[-(m-n) - (\upsilon_1 - \tau_1) + (\upsilon_2 - \tau_2)] \cdot R_w[-(m-n) + (\upsilon_1 - \tau_1) - (\upsilon_2 - \tau_2)] + \\
 &\quad + R_w[(m-n) - (\upsilon_1 + \tau_1) - (\upsilon_2 + \tau_2)] \cdot R_w[(m-n) + (\upsilon_1 + \tau_1) + (\upsilon_2 + \tau_2)] \cdot \\
 &\quad \cdot R_w[(m-n) + (\upsilon_1 + \tau_1) - (\upsilon_2 + \tau_2)] \cdot R_w[(m-n) - (\upsilon_1 + \tau_1) + (\upsilon_2 + \tau_2)] \}
 \end{aligned}$$

Dall'analisi dei 4 addendi che compaiono nella (A2.3) segue che, nell'ipotesi in cui non sia vera nessuna delle due condizioni seguenti:

$$(A2.4) \quad \tau_1 = \upsilon_1 \quad e \quad \tau_2 = \upsilon_2$$

ovvero:

$$(A2.5) \quad \tau_1 = \upsilon_2 \quad e \quad \tau_2 = \upsilon_1$$

³³Il valore atteso viene espanso nella somma di 105 termini; la maggior parte di essi risultano nulli, mentre i restanti forniscono l'espressione (A2.3).

risulta :

$$(A2.6) \quad E [X_3^*(f; \tau_1, \tau_2) \cdot X_3(\alpha \cdot f; \nu_1, \nu_2)] = 0$$

Appendice A3

Determinazione della varianza della ml-HAF in presenza di segnale.

La varianza della ml-HAF in presenza di segnale è definita nel modo seguente (cfr. paragrafo 2.3):

$$(A3.1) \quad \sigma_M^2(f) = E[|\delta S_M(f; \underline{\tau}_{M-1})|^2] = E[\delta S_M(f; \underline{\tau}_{M-1}) \cdot \delta S_M^*(f; \underline{\tau}_{M-1})]$$

La precedente può essere sviluppata ricordando la definizione del termine di perturbazione $\delta S_M(f; \underline{\tau}_{M-1})$. Si ha infatti:

$$(A3.2) \quad \begin{aligned} \delta S_M(f; \underline{\tau}_{M-1}) &= \sum_n \delta s_M(n; \underline{\tau}_{M-1}) \cdot e^{-j \cdot 2\pi \cdot f \cdot n} = \\ &= \sum_n s_M(n; \underline{\tau}_{M-1}) \cdot e^{-j \cdot 2\pi \cdot f \cdot n} \cdot \sum_{k=1}^{2^{M-1}} \frac{w^{*p_M(k)}[n + z_M(k)]}{s^{*p_M(k)}[n + z_M(k)]} \end{aligned}$$

da cui:

$$(A3.3) \quad \begin{aligned} \delta S_M^*(f; \underline{\tau}_{M-1}) &= \sum_m \delta s_M^*(m; \underline{\tau}_{M-1}) \cdot e^{+j \cdot 2\pi \cdot f \cdot m} = \\ &= \sum_m s_M^*(m; \underline{\tau}_{M-1}) \cdot e^{+j \cdot 2\pi \cdot f \cdot m} \cdot \sum_{h=1}^{2^{M-1}} \frac{w^{*[-p_M(h)]}[m + z_M(h)]}{s^{*[-p_M(h)]}[m + z_M(h)]} \end{aligned}$$

Sostituendo (A3.2) ed (A3.3) in (A3.1) si ricava:

$$(A3.4) \quad \sigma_M^2(f) = \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot [m-n]} \cdot s_M^*(m; \underline{\tau}_{M-1}) \cdot s_M(n; \underline{\tau}_{M-1}) \cdot \sum_{k=1}^{2^{M-1}} \sum_{h=1}^{2^{M-1}} \frac{E\left\{w^{*p_M(k)}[n+z_M(k)] \cdot w^{*[-p_M(h)]}[m+z_M(h)]\right\}}{s^{*p_M(k)}[n+z_M(k)] \cdot s^{*[-p_M(h)]}[m+z_M(h)]}$$

Ricordando che:

$$(A3.5) \quad s_M(n; \underline{\tau}_{M-1}) = A^{2^{M-1}} \cdot e^{j \cdot 2\pi \cdot P_M(a_{M-1} + a_M \cdot n)}$$

ed osservando inoltre che:

$$(A3.6) \quad E\left\{w^{*p_M(k)}[n+z_M(k)] \cdot w^{*[-p_M(h)]}[m+z_M(h)]\right\} = \sigma_W^2 \cdot \delta[n+z_M(k) - m - z_M(h)] \cdot \delta[p_M(k) - p_M(h)]$$

si ottiene:

$$(A3.7) \quad \sigma_M^2(f) = \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot (m-n)} \cdot s_M^*(m; \underline{\tau}_{M-1}) \cdot s_M(n; \underline{\tau}_{M-1}) \cdot \sum_{k=1}^{2^{M-1}} \sum_{h=1}^{2^{M-1}} \frac{E\left\{w^{*p_M(k)}[n+z_M(k)] \cdot w^{*[-p_M(h)]}[m+z_M(h)]\right\}}{s^{*p_M(k)}[n+z_M(k)] \cdot s^{*[-p_M(h)]}[m+z_M(h)]} = \sum_n \sum_m e^{+j \cdot 2\pi \cdot f \cdot (m-n)} \cdot A^{2^M} \cdot e^{+j \cdot 2\pi \cdot P_M \cdot a_M \cdot (n-m)} \cdot \sum_{k=1}^{2^{M-1}} \sum_{h=1}^{2^{M-1}} \frac{\sigma_W^2 \cdot \delta[n+z_M(k) - m - z_M(h)] \cdot \delta[p_M(k) - p_M(h)]}{A^2 \cdot e^{+j \cdot 2\pi \cdot p_M(k) \cdot \sum_i a_i \cdot [n+z_M(k)]^i} \cdot e^{-j \cdot 2\pi \cdot p_M(h) \cdot \sum_i a_i \cdot [m+z_M(h)]^i}} =$$

Nella precedente si pone quindi:

$$(A3.8) \quad f_0 = P_M \cdot a_M = 2^{M-1} \cdot M! \cdot \prod \tau_i \cdot a_M$$

$$(A3.9) \quad \text{SNR} = \frac{A^2}{\sigma_w^2}$$

per cui, sostituendo e semplificando si ottiene infine:

$$(A3.10) \quad \sigma_{M^2}(f) = \frac{A^{2M}}{\text{SNR}} \cdot \sum_n \sum_m e^{+j \cdot 2\pi \cdot (f - f_0) \cdot (m - n)} \cdot \sum_{k=1}^{2^{M-1}} \sum_{h=1}^{2^{M-1}} \delta[p_M(k) - p_M(h)] \cdot \delta[n - m + z_M(k) - z_M(h)] =$$

$$= \frac{A^{2M}}{\text{SNR}} \cdot \sum_{k=1}^{2^{M-1}} \sum_{h=1}^{2^{M-1}} \delta[p_M(k) - p_M(h)] \cdot \sum_n \sum_m e^{+j \cdot 2\pi \cdot (m - n) \cdot (f - f_0)} \cdot \delta[n - m + z_M(k) - z_M(h)]$$

La precedente espressione può essere particolareggiata nei casi $M = 2$ ed $M = 3$.

1° caso: $M = 2$.

Utilizzando le notazioni introdotte nel capitolo 1 si ha:

$$(A3.11) \quad \underline{C}_2 = \begin{pmatrix} +1 & -1 \end{pmatrix}$$

$$\underline{p}_2 = \begin{pmatrix} +1 & -1 \end{pmatrix}$$

$$\underline{z}_2 = \begin{pmatrix} +\tau & -\tau \end{pmatrix}$$

La varianza calcolata secondo la (A3.10) vale:

$$(A3.12) \quad \sigma_2^2(f) = \frac{A^4}{\text{SNR}} \cdot \sum_{k=1}^2 \sum_{h=1}^2 \delta[p_2(k) - p_2(h)] \cdot \sum_n \sum_m e^{+j \cdot 2\pi \cdot (m-n) \cdot (f-f_0)} \cdot \delta[n - m + z_2(k) - z_2(h)]$$

Nella tabella seguente sono riportati i valori assunti da $\delta[p_2(k) - p_2(h)]$ al variare di k ed h :

k	h	p₂(k)	p₂(h)	δ[p₂(k)-p₂(h)]
1	1	1	1	1
1	2	1	-1	0
2	1	-1	1	0
2	2	-1	-1	1

Considerando quindi i soli termini non nulli la (A3.12) diventa:

$$(A3.13) \quad \sigma_2^2(f) = \frac{A^4}{\text{SNR}} \cdot \left\{ \sum_n \sum_m e^{+j \cdot 2\pi \cdot (m-n) \cdot (f-f_0)} \cdot \delta[n - m + z_2(1) - z_2(1)] + \sum_n \sum_m e^{+j \cdot 2\pi \cdot (m-n) \cdot (f-f_0)} \cdot \delta[n - m + z_2(2) - z_2(2)] \right\} =$$

$$= 2 \cdot \frac{A^4}{\text{SNR}} \cdot \sum_n \sum_m e^{+j \cdot 2\pi \cdot (m-n) \cdot (f-f_0)} \cdot \delta[n - m] =$$

$$= 2 \cdot \frac{A^4}{\text{SNR}} \cdot \sum_n 1 = 2 \cdot \frac{A^4}{\text{SNR}} \cdot [N - 2\tau]$$

(si tenga presente che nella sommatoria si ha $-(N-1)/2 + \tau \leq n \leq (N-1)/2 - \tau$).

2° caso: M = 3.

Si procede in maniera analoga:

$$\begin{aligned}
 \underline{C}_3 &= \begin{pmatrix} +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \end{pmatrix} \\
 \text{(A3.14)} \quad \underline{p}_3 &= \begin{pmatrix} +1 & -1 & -1 & +1 \end{pmatrix} \\
 \underline{z}_3 &= \begin{pmatrix} \tau_1 + \tau_2 & -\tau_1 + \tau_2 & \tau_1 - \tau_2 & -\tau_1 - \tau_2 \end{pmatrix}
 \end{aligned}$$

Delle 16 possibili coppie del tipo (k,h) { k,h = 1,2,3,4} ne esistono solamente 8 per cui $\delta [p_3(k)-p_3(h)] = 1$; esse sono riportate nella tabella seguente accanto ai corrispondenti valori della differenza $z_3(k)-z_3(h)$:

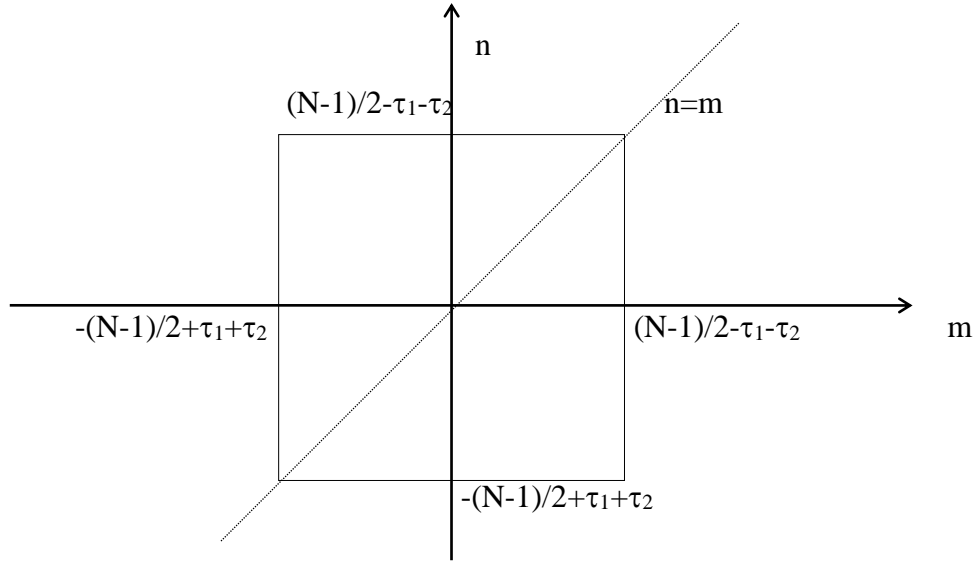
k	h	$\Delta z = z_3(k)-z_3(h)$
1	1	0
1	4	$2 \cdot (\tau_1 + \tau_2)$
2	2	0
2	3	$-2 \cdot (\tau_1 - \tau_2)$
3	2	$2 \cdot (\tau_1 - \tau_2)$
3	3	0
4	1	$-2 \cdot (\tau_1 + \tau_2)$
4	4	0

A questo punto conviene calcolare separatamente i contributi dei vari termini:

$$(1) \quad k = h \Rightarrow \Delta z = 0$$

$$\sigma_{3,1}^2(f) = \frac{A^8}{\text{SNR}} \cdot \sum_{k=1}^4 \sum_n \sum_m e^{+j \cdot 2\pi \cdot (m-n) \cdot (f-f_0)} \cdot \delta[n - m + \Delta z]$$

Nelle 2 sommatorie più interne n ed m variano all'interno di un quadrato centrato sull'origine (riferendosi per semplicità al caso simmetrico risulta $|n| \leq (N-1)/2 - \tau_1 - \tau_2$ ed $|m| \leq (N-1)/2 - \tau_1 - \tau_2$):



Di tutte le possibili coppie (m,n) le uniche che forniscono contributo non nullo nella sommatoria sono quelle per cui $n-m+\Delta z=0$; nel caso esaminato si ha $k=h$, per cui $\Delta z = z_3(k)-z_3(h)=0$. Tali coppie si trovano quindi sulla diagonale del quadrato ($n=m$) e sono in numero di $[N-2\cdot(\tau_1+\tau_2)]$; il primo contributo vale allora:

$$\sigma_{3,1}^2(f) = \frac{A^8}{\text{SNR}} \sum_{k=1}^4 \sum_n 1 = 4 \cdot \frac{A^8}{\text{SNR}} \cdot [N - 2 \cdot (\tau_1 + \tau_2)] \cdot u_{-1}[N - 2 \cdot (\tau_1 + \tau_2)]$$

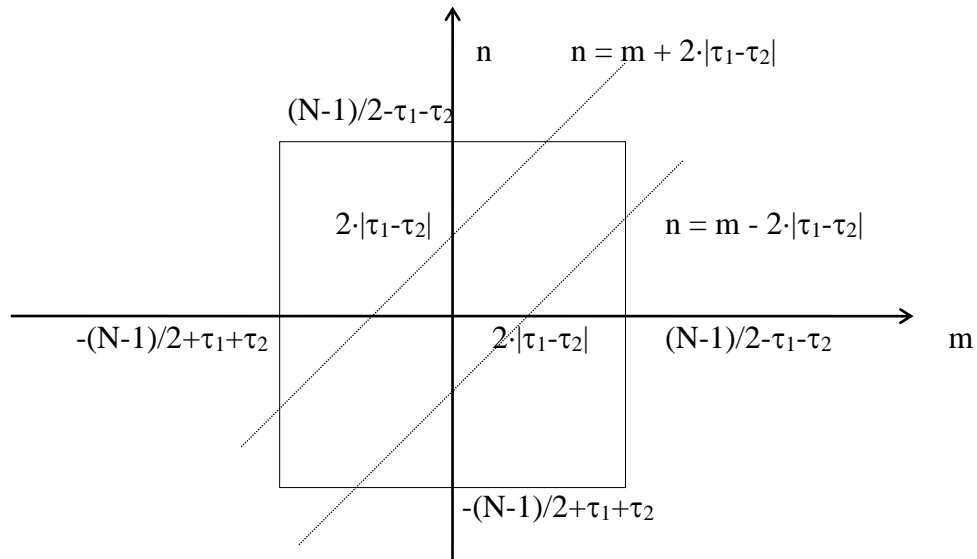
La funzione gradino $u_{-1}[N - 2 \cdot (\tau_1 + \tau_2)]$ vale 1 se $N - 2 \cdot (\tau_1 + \tau_2) \geq 0$ e 0 altrimenti; la sua presenza si spiega osservando che qualora dovesse risultare $N < 2 \cdot (\tau_1 + \tau_2)$ le due sommatorie scomparirebbero fornendo un contributo nullo.

$$(2) \quad k=2, h=3 \text{ e } k=3, h=2 \quad \Rightarrow \quad \Delta z = \pm 2 \cdot |\tau_1 - \tau_2|$$

Con ragionamenti analoghi a quelli precedenti si può scrivere:

$$\sigma_{3,2}^2(f) = \frac{A^8}{\text{SNR}} \cdot \left\{ \sum_n \sum_m e^{+j \cdot 2\pi \cdot (f-f_0) \cdot (m-n)} \cdot \delta[n - m + 2 \cdot |\tau_1 - \tau_2|] + \right. \\ \left. + \sum_n \sum_m e^{+j \cdot 2\pi \cdot (f-f_0) \cdot (m-n)} \cdot \delta[n - m - 2 \cdot |\tau_1 - \tau_2|] \right\}$$

Questa volta le coppie (m,n) per cui $\delta [n-m+\Delta z] = 1$ giacciono su due rette parallele alla diagonale principale del quadrato:



Per la retta superiore si ha:

$$n = m + 2 \cdot |\tau_1 - \tau_2|$$

e gli unici punti da considerare nella sommatoria sono quelli per cui m varia tra $-(N-1)/2 + \tau_1 + \tau_2$ ed $(N-1)/2 - \tau_1 - \tau_2 - 2 \cdot |\tau_1 - \tau_2|$. Analogamente per quella inferiore risulta:

$$m = n + 2 \cdot |\tau_1 - \tau_2|$$

e ci si limita a considerare i valori di n compresi fra $-(N-1)/2 + \tau_1 + \tau_2$ ed $(N-1)/2 - \tau_1 - \tau_2 - 2 \cdot |\tau_1 - \tau_2|$. Si ha allora:

$$\sigma_{3,2}^2(f) = \frac{A^8}{\text{SNR}} \cdot \left\{ \sum_{m = -\frac{N-1}{2} + \tau_1 + \tau_2}^{\frac{N-1}{2} - \tau_1 - \tau_2 - 2 \cdot |\tau_1 - \tau_2|} e^{-j \cdot 4\pi \cdot (f - f_0) \cdot |\tau_1 - \tau_2|} + \right.$$

$$\begin{aligned}
& \left. \sum_{n=-\frac{N-1}{2}+\tau_1+\tau_2}^{\frac{N-1}{2}-\tau_1-\tau_2-2\cdot|\tau_1-\tau_2|} e^{+j\cdot 4\pi\cdot(f-f_0)\cdot|\tau_1-\tau_2|} \right\} = \\
& = \frac{A^8}{\text{SNR}} \cdot \left[e^{-j\cdot 4\pi\cdot(f-f_0)\cdot|\tau_1-\tau_2|} + e^{+j\cdot 4\pi\cdot(f-f_0)\cdot|\tau_1-\tau_2|} \right] \cdot \\
& \cdot \sum_{n=-\frac{N-1}{2}+\tau_1+\tau_2}^{\frac{N-1}{2}-\tau_1-\tau_2-2\cdot|\tau_1-\tau_2|} 1 = \\
& = 2 \cdot \frac{A^8}{\text{SNR}} \cdot \cos \left[4\pi \cdot (f - f_0) \cdot (\tau_1 - \tau_2) \right] \cdot \\
& \cdot \left[N - 2 \cdot (\tau_1 + \tau_2) - 2 \cdot |\tau_1 - \tau_2| \right] \cdot u_{-1} \left[N - 2 \cdot (\tau_1 + \tau_2) - 2 \cdot |\tau_1 - \tau_2| \right] = \\
& = 2 \cdot \frac{A^8}{\text{SNR}} \cdot \cos \left[4\pi \cdot (f - f_0) \cdot (\tau_1 - \tau_2) \right] \cdot \\
& \cdot \left[N - 4 \cdot \max(\tau_1, \tau_2) \right] \cdot u_{-1} \left[N - 4 \cdot \max(\tau_1, \tau_2) \right]
\end{aligned}$$

La presenza del gradino è motivata dal fatto che se dovesse essere $N < 4 \cdot \max(\tau_1, \tau_2)$ le suddette rette cadrebbero al di fuori del quadrato.

$$(3) \quad k=1, h=4 \text{ e } k=4, h=1 \quad \Rightarrow \quad \Delta z = \pm 2 \cdot (\tau_1 + \tau_2)$$

Con procedimento analogo a quello precedente, qui non riportato per semplicità, si ottiene:

$$\begin{aligned}
\sigma_{3,3}^2(f) &= 2 \cdot \frac{A^8}{\text{SNR}} \cdot \cos \left[4\pi \cdot (f - f_0) \cdot (\tau_1 + \tau_2) \right] \cdot \\
& \cdot \left[N - 4 \cdot (\tau_1 + \tau_2) \right] \cdot u_{-1} \left[N - 4 \cdot (\tau_1 + \tau_2) \right]
\end{aligned}$$

Sommando i tre contributi calcolati nei punti (1), (2), (3) si ottiene infine l'espressione cercata:

$$(A3.15) \quad \sigma_3^2(f) = E[|\delta S_3(f; \underline{\tau}_2)|^2] = 4 \cdot \frac{A^8}{\text{SNR}} \cdot$$

$$\left\{ \begin{aligned} & \left[N - 2 \cdot (\tau_1 + \tau_2) \right] \cdot u_{-1} \left[N - 2 \cdot (\tau_1 + \tau_2) \right] + \\ & + \frac{1}{2} \cdot \cos \left[4\pi \cdot (f - f_0) \cdot (\tau_1 - \tau_2) \right] \cdot \left[N - 4 \cdot \max(\tau_1, \tau_2) \right] \cdot u_{-1} \left[N - 4 \cdot \max(\tau_1, \tau_2) \right] + \\ & + \frac{1}{2} \cdot \cos \left[4\pi \cdot (f - f_0) \cdot (\tau_1 + \tau_2) \right] \cdot \left[N - 4 \cdot (\tau_1 + \tau_2) \right] \cdot u_{-1} \left[N - 4 \cdot (\tau_1 + \tau_2) \right] \end{aligned} \right\}$$

Appendice A4

Determinazione della P_{FA} ottenuta tramite autogate nel caso di varianza non costante.

Nel paragrafo 3.2 è stata ottenuta l'espressione della probabilità di falso allarme in funzione dei parametri dell'autogate:

$$(A4.1) \quad P_{FA} = \frac{1}{(G+1)^p}$$

Essa è stata ricavata nell'ipotesi che la varianza della ml-HAF ($\sigma_M^2(f)$) fosse costante; ciò accade ad esempio quando $M=2$. Nella presente appendice la relazione (A4.1) viene generalizzata eliminando la suddetta limitazione. Seguendo allora le notazioni già introdotte nel paragrafo citato indichiamo con q_i ($i=1,2,\dots,p$) e con q rispettivamente i moduli quadri delle ml-HAF sulle prese laterali e sulla presa centrale dell'autogate. In base a quanto riportato nel paragrafo 3.2 essi risultano esponenziali, con parametri indicati rispettivamente con λ_i ($i=1,2,\dots,p$) e λ . Alcuni fra i λ_i potrebbero risultare fra loro coincidenti; è quindi opportuno raggrupparli in $p' \leq p$ insiemi disgiunti, indicando poi con m_i ($i = 1,2 \dots p'$) le rispettive molteplicità:

Parametro	Molteplicità
λ_1	m_1
λ_2	m_2
..	..
..	..
$\lambda_{p'}$	$m_{p'}$
Totale	$p = \sum m_i$

Posto:

$$(A4.2) \quad X = \sum_{i=1}^p q_i$$

la densità di probabilità della variabile aleatoria X risulta data dalla convoluzione³⁴ delle densità (esponenziali) delle singole q_i . Indicando allora con $P_X(s)$ la trasformata di Laplace della densità di X e ricordando l'espressione della trasformata di un esponenziale si può scrivere:

$$(A4.3) \quad P_X(s) = \left(\frac{\lambda_1}{s + \lambda_1} \right)^{m_1} \cdot \left(\frac{\lambda_2}{s + \lambda_2} \right)^{m_2} \cdots \left(\frac{\lambda_{p'}}{s + \lambda_{p'}} \right)^{m_{p'}} =$$

$$= \sum_{i=1}^{p'} \sum_{j=1}^{m_i} \frac{c_{i,j}}{(s + \lambda_i)^j}$$

L'uguaglianza fra il penultimo e l'ultimo membro della (A4.3) discende dall'applicazione del teorema dei residui. Risulta poi:

$$(A4.4) \quad c_{i,j} = \frac{1}{(m_i - j)!} \cdot \left\{ \frac{d^{(m_i - j)}}{ds^{(m_i - j)}} \left[(s + \lambda_i)^{m_i} \cdot P_X(s) \right] \right\}_{s = -\lambda_i}$$

³⁴In virtù dell'ipotesi di indipendenza statistica fra i campioni q_i .

$$= \frac{\lambda_i^{m_i}}{(m_i - j)!} \cdot \left\{ \frac{d^{(m_i - j)}}{ds^{(m_i - j)}} \left[\prod_{\substack{k=1 \\ k \neq i}}^{p'} \left(\frac{\lambda_k}{s + \lambda_k} \right)^{m_k} \right] \right\}_{s = -\lambda_i}$$

Antitrasformando si ottiene la densità di probabilità cercata:

$$(A4.5) \quad p_X(x) = \sum_{i=1}^{p'} \sum_{j=1}^m \frac{c_{i,j}}{(j-1)!} \cdot x^{j-1} \cdot e^{-\lambda_i \cdot x} \cdot u_{-1}(x)$$

La soglia T è proporzionale ad X:

$$(A4.6) \quad T = G \cdot X$$

da cui:

$$(A4.7) \quad p_T(T) = \frac{1}{G} \cdot p_X\left(\frac{T}{G}\right) =$$

$$= \sum_{i=1}^{p'} \sum_{j=1}^m \frac{c_{i,j}}{(j-1)!} \cdot \frac{T^{j-1}}{G^j} \cdot e^{-\frac{\lambda_i}{G} \cdot T} \cdot u_{-1}(T)$$

La probabilità di falso allarme può essere determinata attraverso l'espressione seguente (cfr. paragrafo 3.2):

$$(A4.8) \quad P_{FA} = \int_0^{+\infty} p_T(T) \cdot [1 - F_Q(T)] \cdot dT$$

Sostituendo (A4.7) in (A4.8) e sviluppando si ottiene:

$$\begin{aligned}
 (A4.9) \quad P_{FA} &= \int_0^{+\infty} \sum_{i=1}^{p'} \sum_{j=1}^m \frac{c_{i,j}}{(j-1)!} \cdot \frac{T^{j-1}}{G^j} \cdot e^{-\left(\frac{\lambda_i}{G} + \lambda\right) \cdot T} \cdot dT = \\
 &= \sum_{i=1}^{p'} \sum_{j=1}^m \frac{c_{i,j}}{G^j \cdot (j-1)!} \cdot \int_0^{+\infty} T^{j-1} \cdot e^{-\frac{\lambda_i + \lambda \cdot G}{G} \cdot T} \cdot dT = \\
 &= \sum_{i=1}^{p'} \sum_{j=1}^m \frac{c_{i,j}}{G^j \cdot (j-1)!} \cdot \left(\frac{G}{\lambda_i + G \cdot \lambda}\right)^j \cdot (j-1)! = \\
 &= \sum_{i=1}^{p'} \sum_{j=1}^m \frac{c_{i,j}}{\left(\lambda_i + G \cdot \lambda\right)^j}
 \end{aligned}$$

Si consideri ora la (3.2.17), che viene qui riportata in forma leggermente diversa introducendo le molteplicità m_i :

$$(A4.10) \quad \frac{1}{\prod_{i=1}^{p'} \left(1 + G \cdot \frac{\lambda}{\lambda_i}\right)^{m_i}} = \prod_{i=1}^{p'} \left(\frac{\frac{\lambda_i}{\lambda}}{G + \frac{\lambda_i}{\lambda}}\right)^{m_i}$$

Espandendo la (A4.10) in frazioni parziali si ottiene:

$$(A4.11) \quad \prod_{i=1}^{p'} \left(\frac{\frac{\lambda_i}{\lambda}}{G + \frac{\lambda_i}{\lambda}}\right)^{m_i} = \sum_{i=1}^{p'} \sum_{j=1}^{m_i} \frac{B_{i,j}}{\left(G + \frac{\lambda_i}{\lambda}\right)^j}$$

dove:

(A4.12)

$$\begin{aligned}
 B_{i,j} &= \frac{1}{(m_i - j)!} \cdot \left\{ \frac{d^{(m_i - j)}}{dG^{(m_i - j)}} \left[\left(\frac{\lambda_i}{\lambda} \right)^{m_i} \cdot \prod_{\substack{k=1 \\ k \neq i}}^{p'} \left(\frac{\lambda_k}{G \cdot \lambda + \lambda_k} \right)^{m_k} \right] \right\}_{G = -\frac{\lambda_i}{\lambda}} \\
 &= \frac{\left(\frac{\lambda_i}{\lambda} \right)^{m_i}}{(m_i - j)!} \cdot \left\{ \frac{d^{(m_i - j)}}{ds^{(m_i - j)}} \left[\prod_{\substack{k=1 \\ k \neq i}}^{p'} \left(\frac{\lambda_k}{s + \lambda_k} \right)^{m_k} \right] \right\}_{s = -\lambda_i} \cdot \left[\frac{ds}{dG} \right]^{m_i - j}
 \end{aligned}$$

L'uguaglianza tra secondo e terzo membro discende dalla sostituzione:

$$(A4.13) \quad s = G \cdot \lambda$$

e dalla applicazione del teorema di derivazione delle funzioni composte. Osservando che:

$$(A4.14) \quad \left[\frac{ds}{dG} \right]^{m_i - j} = \lambda^{m_i - j}$$

e ricordando la (A4.4) si ottiene:

$$(A4.15) \quad B_{i,j} = \lambda^{m_i - j} \cdot \frac{1}{\lambda^{m_i}} \cdot c_{i,j} = \frac{c_{i,j}}{\lambda^j}$$

La P_{FA} è data dalla (A4.9); sostituendo in essa la (A4.15) e ricordando le (A4.10)-(A4.11) si ha:

$$\begin{aligned}
(A4.16) \quad P_{FA} &= \sum_{i=1}^{p'} \sum_{j=1}^m \frac{c_{i,j}}{\left(\lambda_i + G \cdot \lambda\right)^j} = \sum_{i=1}^{p'} \sum_{j=1}^m \frac{B_{i,j}}{\left(G + \frac{\lambda_i}{\lambda}\right)^j} = \\
&= \frac{1}{\prod_{i=1}^{p'} \left(1 + G \cdot \frac{\lambda}{\lambda_i}\right)^{m_i}}
\end{aligned}$$

che rappresenta l'espressione cercata. Da notare che nel caso particolare di varianza costante risulta:

$$\begin{aligned}
(A4.17) \quad p' &= 1 \\
m_1 &= p \\
\lambda_1 &= \lambda
\end{aligned}$$

per cui la (A4.16) si riduce alla già nota:

$$(A4.18) \quad P_{FA} = \frac{1}{(G+1)^p}$$

Appendice A5

Dipendenza della P_{FA} dalla frequenza.

Nel capitolo 3 si è visto che la probabilità di falso allarme ottenuta applicando un autogate sui campioni del modulo quadro della ml-HAF possiede in generale l'espressione:

$$(A5.1) \quad P_{FA} = \frac{1}{\prod_{i=1}^p \left[1 + G \cdot \frac{\lambda}{\lambda_i} \right]}$$

dove, seguendo le notazioni già introdotte, risulta:

$$(A5.2) \quad \lambda = 1/\sigma_M^2(f)$$

$$(A5.3) \quad \lambda_i = 1/\sigma_M^2(f_i)$$

Al variare di f (frequenza corrispondente al campione sulla presa centrale dell'autogate) e, conseguentemente delle f_i (frequenze dei campioni sulle prese laterali, che si trovano ad una distanza costante da f), varieranno sia λ che i vari λ_i ; in base alla (A5.1) dovrebbe pertanto variare anche la P_{FA} , che risulterebbe così dipendente dalla frequenza f . Fortunatamente tali variazioni sono molto contenute, per cui il rapporto λ/λ_i assume valore pressochè unitario: ciò consente di approssimare la (A5.1) con la:

$$(A5.4) \quad P_{FA} = \frac{1}{(G + 1)^p}$$

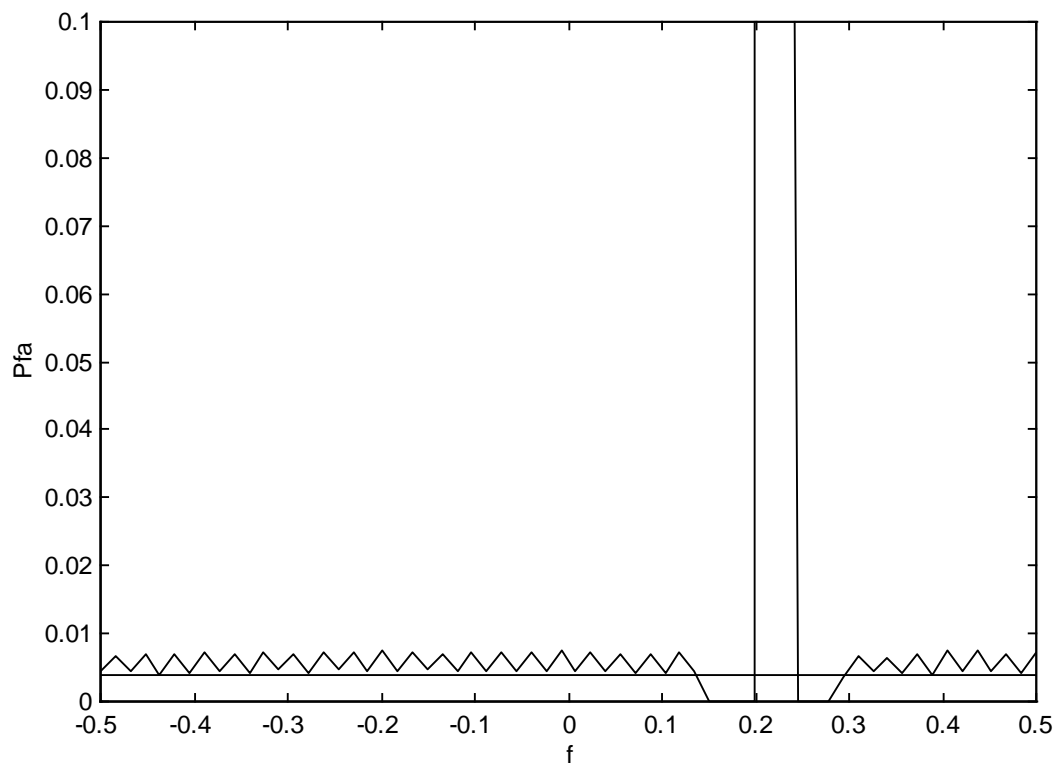
in cui tale dipendenza scompare. Le considerazioni precedenti sono confermate dai risultati sperimentali. L'esempio seguente è relativo ad una simulazione con i seguenti parametri:

$$N = 192 \quad M = 3 \quad L = 1 \quad \tau_1 = 40 \quad \tau_2 = 24 \quad G = 3 \quad p = 4 \quad k = 0.08 \quad \text{SNR} = 10 \text{ dB}$$

La varianza calcolata per via teorica vale:

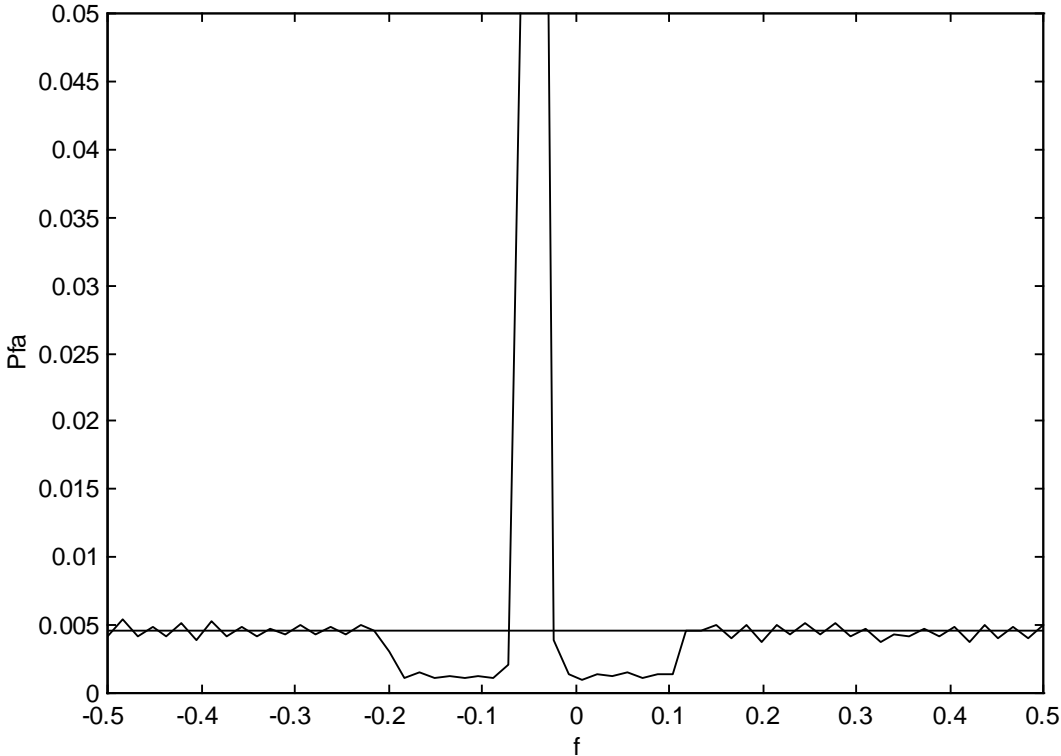
$$\sigma_3^2(f) = 25.6 + 6.4 \cdot \cos [64\pi(f-f_0)]$$

e dipende quindi dalla frequenza f . La probabilità di falso allarme teorica, valutata secondo la (A5.4) assume il valore $P_{FA} = 1/256 \cong 3.9 \cdot 10^{-3}$. Nella figura seguente accanto ad essa è riportato l'andamento della P_{FA} sperimentale ottenuta per simulazione.



Come si può notare la variabilità della probabilità di falso allarme con la frequenza è molto contenuta, per cui essa può essere ritenuta pressoché costante. Il picco che compare sulla destra è legato al lobo principale della ml-HAF, ossia alla presenza del segnale utile; si notino le caratteristiche zone d'ombra nell'intorno di tale picco, nelle quali la P_{FA} si annulla a causa dell'innalzamento della soglia da esso provocato. Evidentemente tale zona risulta tanto più larga quanto maggiore è il numero di prese dell'autogate, come risulta evidente dal seguente grafico,

relativo ad una simulazione con gli stessi parametri della precedente (ad eccezione di G e p che assumono rispettivamente i valori 0.4 e 16, per una P_{FA} teorica pari a $4.6 \cdot 10^{-3}$).



Bibliografia

- [1] S. Peleg e B. Porat, "Estimation and classification of signals with polynomial phase", *IEEE Trans. on Information Theory*, vol. 37, pp. 422-430, 1991.
- [2] S. Peleg e B. Friedlander, "The discrete polynomial phase transform", *IEEE Trans. on Signal Processing*, pp. 1901-1914, Agosto 1991.
- [3] S. Peleg, "Estimation and classification of signals with polynomial phase", Ph.D. Thesis, Univ. of California, Davis, CA, 1993.
- [4] S. Barbarossa, A. Porchia e A. Scaglione, "Multiplicative multilag higher-order ambiguity function", *Proc. of Int. Conference on Acoustics, Speech and Signal Processing*, Atlanta, GA, vol.5, pp. 3022-3026, 7-11 Maggio 1996.
- [5] S. Barbarossa e A. Scaglione, "Product high-order ambiguity function for multicomponent polynomial phase signal modelling", to be published on the *IEEE Transaction on Signal Processing*.
- [6] S. Barbarossa e A. Scaglione, "Statistical Analysis of the Product High-Order Ambiguity Function", submitted to the *IEEE Trans. on Information Theory*.
- [7] S. Barbarossa e A. Scaglione, "Polynomial Minimum Shift Keying Modulation: Simplified Decoding of CPM Signals", *Proc. of IEEE Workshop on Signal Processing Advances in Wireless Communications*, Paris, April 1997.
- [8] G. Agostini, "Demodulazione di segnali CPM mediante funzioni di ambiguità di ordine superiore", Tesi di laurea, Università "La Sapienza", Roma, AA. 1995/96.
- [9] A. Papoulis, "Probability, Random variables and Stochastic Processes", Mc. Graw Hill, Inc., International Edition 1991.
- [10] A. V. Oppenheim e R. W. Shafer, "Discrete time Signal Processing", Prentice Hall, Inc., Englewood Cliffs, NJ, 1989.
- [11] S. Benedetto, E. Biglieri, V. Castellani, "Digital Transmission Theory", Prentice Hall, Inc., Englewood Cliffs, NJ, 1987.
- [12] S. Peleg e B. Porat, "The Cramer-Rao Lower Bound for signals with Constant Amplitude and Polynomial Phase", *IEEE Trans. on Signal Processing*, 39, pp. 749-752, 1991b.