

Université du Québec
Institut national de la recherche scientifique
Énergie Matériaux Télécommunications

Nouvelles bornes et techniques de synchronisation temporelle pour les futurs systèmes de communications sans fil

Par
Ahmed Masmoudi

Mémoire pour l'obtention du grade de Maître ès sciences (M.Sc.)
en télécommunications

Jury d'évaluation

Directeur de recherche
Examineur interne
Examineur externe

Sofène Affes, INRS-ÉMT
Serge Tatu, INRS-ÉMT
Paul Fortier,
Département de génie électrique et de génie informatique
Université Laval

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Avant-Propos

Je souhaite remercier Pr. Sofiene Affes pour m'avoir accepté dans son équipe et guidé tout au long de ma maîtrise. Ses remarques et critiques pertinentes m'ont conduit vers la bonne voie. Un grand merci aussi à M. Faouzi Bellili. C'était un vrai plaisir de travailler avec lui.

Je remercie aussi mes amis Mouhamed, Hadhami, Foued, Ghaya et mes camarades de l'INRS pour leurs encouragements et leurs soutiens tout au long de ces deux années ainsi que Jigo et Boss pour leur APS.

Ces avant-propos seraient incomplets sans un remerciement adressé aux membres de ma famille, en particulier mes parents, mon frère Mouhamed et ma sœur Ines. Ce travail leur appartient à tous.

شُكْرًا بَاعَثَ الْقَنَاءَةَ.

Résumé

Ce mémoire est consacré à la synchronisation temporelle (ou estimation du délai) pour les récepteurs numériques. Les principales contributions incluent deux grands volets. Le premier consiste à estimer le retard où les nouvelles techniques d'estimation développées sont basées sur le critère du maximum de vraisemblance. D'une part, nous développons un estimateur à maximum de vraisemblance par la méthode "importance sampling" pour les signaux numériques linéairement modulés où l'on considère un seul trajet de propagation, et donc un seul paramètre à estimer. Dans cette configuration, nous supposons que les données transmises sont totalement inconnues au niveau du récepteur. Le délai reste constant sur l'intervalle d'observation et le bruit est supposé blanc. Nous appliquons aussi la méthode proposée dans le cas d'un seul trajet au cas de plusieurs trajets et donc le nombre de paramètres à estimer augmente avec le nombre de trajets détectés. Nous signalons que dans ce cas le signal transmis est connu par le récepteur. Cette méthode peut s'appliquer dans les radars et les systèmes de localisations. En plus nous nous sommes intéressés à la synchronisation temporelle pour les systèmes CDMA en développant deux estimateurs à maximum de vraisemblance. Le premier se base sur la méthode "importance sampling" et l'autre sur l'algorithme itératif "expectation maximization".

D'autre part, nous dérivons aussi les expressions analytiques des bornes de Cramér-Rao pour les estimateurs non biaisés du retard dans le cas des signaux QAM carrés et des systèmes CDMA.

Table des matières

Avant-Propos	i
Résumé	ii
List des figures	vi
Introduction	2
1 La synchronisation temporelle pour les signaux modulés	3
1.1 Introduction	4
1.2 L'importance de la synchronisation temporelle	4
1.3 Estimateur à maximum de vraisemblance	6
1.4 Estimateurs du retard à maximum de vraisemblance	7
1.4.1 Estimateur pour Faible RSB	7
1.4.2 Estimateur à Maximum de Vraisemblance Conditionnel	9
1.5 Limites de performance des estimateurs	9
1.5.1 Les Bornes de Cramér-Rao	9
1.5.2 Les Bornes de Cramér-Rao Modifiées	10
1.6 Conclusion	12
2 A Non-Data-Aided Maximum Likelihood Time Delay Estimator using Importance Sampling	15
2.1 Introduction	16
2.2 Discrete-Time Signal Model	18
2.3 Likelihood function	19
2.4 Global Maximization of the Compressed Likelihood Function	21
2.5 The Importance Sampling Technique	22
2.6 Estimation of the Time Delay	25
2.6.1 First scenario : $\tau^* \in [0, PT]$	25
2.6.2 Second scenario : $\tau^* \in [0, T]$	28
2.6.3 Summary of steps	29
2.7 Simulation Results	30

2.8	Conclusion	34
3	Closed-Form Expressions for the Exact Cramér-Rao Bounds of Timing Recovery Estimators from BPSK, MSK and Square-QAM Transmissions	41
3.1	Introduction	43
3.2	System Model	44
3.3	Time Delay CRLB for Square QAM-Modulated Signals	46
3.4	CRLB for BPSK and MSK Modulated Signals	52
3.5	Graphical Representations	53
3.6	Conclusion	56
4	A Maximum Likelihood Time Delay Estimator in a Multipath Environment Using Importance Sampling	65
4.1	Introduction	66
4.2	System Model and Compressed Likelihood Function	68
4.3	Global Maximization of the Compressed Likelihood Function	69
4.4	The Importance Sampling Based Time Delays Estimation	71
4.4.1	IS Concept	71
4.4.2	Time Delays Estimation in Active Systems	75
4.4.3	Time Delays Estimation in Passive Systems	77
4.5	Simulation Results	79
4.6	Conclusion	82
5	Maximum Likelihood Time Delay Estimation for Direct-Sequence CDMA Multipath Transmissions	87
5.1	Introduction	88
5.2	System Model and Background	91
5.3	The Cramèr-Rao Lower Bound	93
5.4	Expectation Maximization Algorithm	95
5.5	The Importance Sampling Technique	98
5.6	Extension To MC-CDMA Systems	103
5.7	Simulation Results	105
5.8	Conclusion	108
	Conclusion	118

Table des figures

1.1	Diagramme de l'œil pour un filtre à cosinus surélevé avec un coefficient de retombée de 20%	5
1.2	Diagramme de l'œil pour un filtre à cosinus surélevé avec un coefficient de retombée de 100%	5
2.1	Plot of $g'_{\rho'_1}(\cdot)$ for $\rho'_1 = 20$ and $\rho'_1 = 10$ using a root-raised cosine pulse and for $K = 100$	24
2.2	Plot of the cumulative distribution function (CDF), $G'(\tau)$ whose pdf is $g'(\tau)$, SNR = 5 dB.	26
2.3	Performance versus ρ'_1 for SNR=10 dB.	30
2.4	Estimation performance considering $L'_{c,\rho_0}(\tau)/g'(\tau)$ and setting $L'_{c,\rho_0}(\tau)/g'(\tau)$ to 1 with QPSK modulation.	31
2.5	Comparison between the estimation performance of the IS algorithm using the two scenarios and the tracking performance of the CML-TED using QPSK modulation.	32
2.6	Comparison between the estimation performance of IS algorithm and the tracking performance of CML TED using QPSK modulation and for a roll-off factor of 0.2.	33
2.7	Normalized MSE of the time delay estimate for different QAM modulation order, using a root-raised cosine filter with a roll-off factor 0.5.	34
2.8	Comparison of the estimation performance with and without forced symbols using 16-QAM modulation and for a roll-off factor of 0.5.	34
2.9	Plot of $g_0(\cdot)$, $g_1(\cdot)$ and $g_2(\cdot)$	36
3.1	Comparison between the empirical CRLB and the analytical expression in (3.44) for different modulation orders using $K = 100$ and a raised-cosine pulse with rolloff factor of 0.2.	53
3.2	CRLB/MCRLB ratio vs. SNR for different modulation orders using $K = 100$ and a raised-cosine pulse with rolloff factor of 0.2 and 1.	54
3.3	CRLB vs. SNR for different modulation orders using $K = 100$ and a raised-cosine pulse with rolloff factor of 0.2.	55

3.4	CRLB/MCRLB ratio vs. SNR for different modulations and a time-limited shaping pulse.	55
4.1	Complementary cumulative distribution function of the ratio $K(\tau_{m;n})$. . .	73
4.2	Plot of $\bar{g}_{\rho'_1}(\cdot)$ at SNR = 10 for (a) $\rho'_1 = 1$ and (b) $\rho'_1 = 6$	75
4.3	Estimation performance as a function of ρ'_1	80
4.4	Estimation performance of the IS-based, EM ML and the MUSIC-type algorithms in an active system vs. SNR.	81
4.5	Estimation performance of the IS-based, EM ML and the MUSIC-type algorithms in a passive system vs. SNR.	82
4.6	Estimation performance of the IS-based, EM ML and the MUSIC-type algorithms in a passive system as a function of $\Delta\tau$	83
4.7	Estimation performance of the IS-based, EM ML and the MUSIC-type algorithms vs. signal bandwidth at SNR = 10 dB in an active system. . .	83
4.8	Estimation performance of the IS-based and the MUSIC-type algorithms vs. Doppler shift at SNR = 10 dB.	84
5.1	Estimation performance of the IS-based, the EM-based and the Root-MUSIC algorithms for closely-separated delays, $M = 4$	106
5.2	Estimation performance of the IS-based, the EM-based and the Root-MUSIC for closely-separated delays, $M = 1$	107
5.3	MSE vs. number of antenna branches M for $N = 1$ symbol, $K = 1$ subcarrier, and SNR = 10 dB.	108
5.4	MSE vs. number of received symbols N for $M = 1$ antenna branch, $K = 1$ subcarrier, and SNR = 10 dB.	108
5.5	MSE vs. number of subcarrier with $M = 1$ and $N = 1$	109
5.6	Complementary cumulative distribution function of the ratio $F(\Delta\tau_{m;n})$. .	111

Introduction

La synchronisation est une tâche essentielle pour n'importe quel système de communications numérique. Souvent les performances d'un système de transmission sont dictées par la fiabilité de la fonction de synchronisation. En effet, le signal reçu est complètement connu à l'exception des données transmises et des paramètres introduits par le canal (délai, phase, décalage fréquentiel...). Même si la fonction primordiale du récepteur est de reconstruire les données transmises, ceci ne peut se faire qu'en connaissant les paramètres introduits par le canal. Dans ce travail, nous nous intéressons à la synchronisation temporelle, ou en d'autres termes à l'estimation du délai de propagation. Dans ce contexte, bien que plusieurs estimateurs du délai aient été développés durant les dernières décennies, ce problème suscite encore une grande attention surtout après le succès des systèmes de communication sans fil et les avancées en microélectronique qui offrent plus de possibilité d'implémentation.

Les estimateurs du retard peuvent être classés en plusieurs sous-classes : supervisés, semi-aveugles ou aveugles. Les techniques supervisées exploitent la connaissance des symboles transmis dans les blocs de synchronisation pour faciliter la procédure d'estimation. Les méthodes semi-aveugles s'inspirent des techniques supervisées puisque les symboles inconnus sont d'abord estimés puis utilisés dans la synchronisation. Bien qu'elles requièrent la transmission de moins de symboles connus, elles souffrent d'erreurs de détections qui dégradent les performances du système. Dans la suite, nous nous intéressons au cas où les données transmises sont inconnues.

Dans ce contexte, plusieurs estimateurs ont été rapportés dans la littérature pour offrir les meilleures performances possibles. Il est connu que l'estimateur à maximum de vraisemblance est un estimateur asymptotiquement efficace et qu'il réalise la meilleure performance à des valeurs relativement élevées du rapport signal sur bruit (RSB), même sur de courts intervalles d'observation. Par conséquent, il a été l'objet de recherches intensives. Dans le cas où les symboles transmis sont connus, une expression analytique du maximum global de la fonction de vraisemblance peut être obtenue. Toutefois, lorsque les données transmises sont totalement inconnues (c'est-à-dire le paramètre d'intérêt doit être estimé d'une manière aveugle), la fonction de vraisemblance devient une fonction non-linéaire et il est difficile, même impossible, de trouver une expression analytique du maximum global de la fonction de vraisemblance. Dans ce cas, des méthodes de résolution

numériques doivent être envisagées. Le travail présenté dans ce mémoire s'inscrit dans le cadre de développement de méthodes pour trouver le maximum global de la fonction de vraisemblance. L'étape primordiale pour le développement de tels estimateurs est d'exprimer le problème sous la forme d'un modèle linéaire généralisé. Pour ce faire nous traitons les modèles les plus répandus en communication numérique, à savoir le cas d'une transmission d'un signal modulé sur un seul trajet, une transmission sur plusieurs trajets et le cas des systèmes CDMA. Une telle distinction est indispensable puisque les estimateurs développés par la suite changent d'un modèle à un autre.

Une fois que les estimateurs sont développés, nous évaluons leurs performances en termes de variance de l'estimé comme une mesure de performances du système. Pour ce faire, il est évident qu'il faut les comparer aux autres estimateurs mais aussi par rapport à une borne inférieure de la variance de tout estimateur. Dans ce contexte, les bornes de Cramér-Rao sont connues comme des bornes inférieures contre lesquelles les performances des estimateurs sont comparées. Elles indiquent la limite inférieure de la précision d'estimation qui peut être atteinte. Notre revue de la littérature nous a révélé que ces bornes n'ont pas encore été dérivées analytiquement en estimation aveugle. Elles n'ont été calculées que de façon numérique à partir d'expressions très complexes. La dérivation de ces expressions permet de mieux caractériser et analyser les performances du système. Dans ce travail, nous dérivons les expressions analytiques de ces bornes pour les estimateurs de délai en présence de la plupart des constellations utilisées couramment et pour les systèmes CDMA.

Ce rapport est structuré comme suit. Dans le chapitre 1, nous présentons une brève introduction sur l'estimation du retard en communication sans fil. Les contributions effectuées sont présentées dans les chapitres suivants. Le chapitre 2 décrit le nouvel estimateur pour les signaux modulés. Ensuite, dans le chapitre 3, nous dérivons les expressions exactes des bornes de Cramér-Rao pour l'estimation aveugle. Dans le chapitre 4, nous développons un estimateur de délai dans le cas d'un canal à plusieurs trajets. Et dans le dernier chapitre nous nous intéressons aux systèmes CDMA où nous développons deux estimateurs à maximum de vraisemblance et les bornes de Cramér-Rao correspondantes.

Chapitre 1

La synchronisation temporelle pour les signaux modulés

1.1 Introduction

La synchronisation est une tâche essentielle pour n'importe quel système de communications numérique. La synchronisation temporelle, appelée aussi estimation du délai de propagation, est un problème fréquemment rencontré dans la synchronisation. Le but est de s'assurer que les échantillons pris du signal reçu concordent avec la valeur optimale pour reconstruire les données d'une manière fiable. Autrement dit, le délai introduit par le canal de propagation doit être pris en considération au niveau du récepteur. Une des solutions pour estimer ce paramètre est d'envoyer un signal connu au récepteur. La méthode classique consiste à faire l'auto-corrélation de l'observation. Cependant, cette approche est coûteuse en termes d'énergie et de bande passante puisque le signal transmis ne transporte pas d'information utile. C'est pour cette raison que plusieurs travaux ont traité le problème de la synchronisation temporelle en utilisant directement le signal reçu. Les méthodes résultantes sont classées comme aveugle (non-data-aided) où le récepteur n'a pas besoin de connaître les données transmises. Une étape importante dans le processus d'estimation est de déterminer une fonction objective, calculée à partir du signal reçu, de telle façon qu'une estimée du délai de propagation peut être obtenu. En se basant sur cette fonction, les estimateurs du délai sont classés comme suit [1] : à erreur quadratique moyenne minimale (minimum mean square error), forçage à zéro (zero forcing), early-late gate, maximum de vraisemblance, etc. Dans ce qui suit, nous nous intéressons au critère du maximum de vraisemblance.

1.2 L'importance de la synchronisation temporelle

Dans plusieurs systèmes de communications, les performances sont étroitement liées à la marge temporelle allouée. La marge temporelle est l'erreur de synchronisation maximale que le récepteur tolère sans provoquer des dégradations en performances. Elle peut être jugée en examinant le diagramme de l'œil du signal à l'entrée du bloc de décision au récepteur. Le diagramme de l'œil est obtenu en superposant plusieurs répliques du signal reçu. Le nom de ce diagramme vient du fait que la figure obtenue ressemble à un œil humain. Pour évaluer la marge de manœuvre d'un système par rapport aux distorsions que subit le système, par exemple à l'erreur de synchronisation, à l'interférence inter-symbole et au bruit, nous étudions la forme et l'ouverture du diagramme de l'œil.

Les figures 1.1 et 1.2 illustrent les diagrammes de l'œil en absence de bruit et en utilisant un filtre à cosinus surélevé avec un coefficient de retombé de 20% et 100%, respectivement. L'instant auquel l'ouverture du diagramme est maximale correspond à l'instant optimal d'échantillonnage. Comme nous pouvons le voir, l'ouverture de l'œil diminue de plus en plus que nous nous éloignons de l'instant $t = 0$. En effet, l'interférence inter-symbole augmente d'autant plus que l'instant d'échantillonnage s'éloigne de l'instant op-

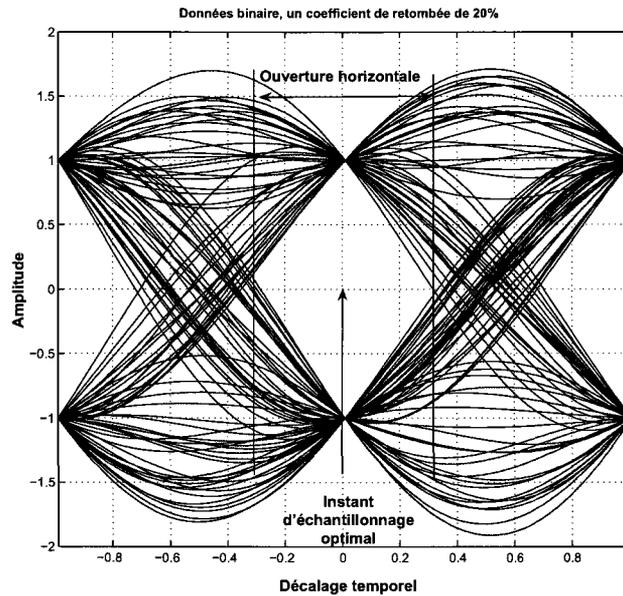


FIGURE 1.1 – Diagramme de l’œil pour un filtre à cosinus surélevé avec un coefficient de retombée de 20%

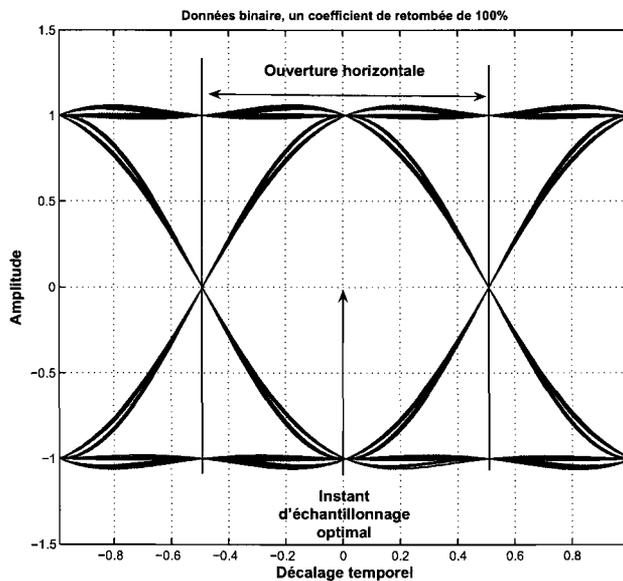


FIGURE 1.2 – Diagramme de l’œil pour un filtre à cosinus surélevé avec un coefficient de retombée de 100%

timal, et par conséquent, la marge de bruit diminue. En absence de bruit, les données peuvent être parfaitement détectées tant que l’instant d’échantillonnage est à l’intérieur de la zone d’ouverture de l’œil. Une synchronisation parfaite qui satisfait le critère de Nyquist minimise ou annule complètement l’interférence inter-symbole. Cependant, plus le

rapport signal sur bruit (RSB) diminue, plus l'intervalle d'échantillonnage fiable diminue et le système devient plus sensible aux erreurs de synchronisation.

1.3 Estimateur à maximum de vraisemblance

L'estimateur à maximum de vraisemblance (MLE) est un estimateur dit à efficacité asymptotique. Il est défini comme la valeur du paramètre qui maximise la fonction de vraisemblance. En général, il a été prouvé en [2] que le MLE est asymptotiquement non-biaisé, atteint la borne de Cramér-Rao (CRLB) et son erreur possède une Distribution Gaussienne. Concernant notre problème, le MLE devrait atteindre la CRLB pour de grandes valeurs du rapport signal sur bruit. Pour cette raison, la plupart des algorithmes de synchronisation sont basés sur le critère de maximum de vraisemblance.

La formulation du problème d'estimation varie suivant que nous considérons un signal à temps continu ou un signal à temps discret. La première approche semble être la plus appropriée à cause de la nature physique du signal, mais les récepteurs numériques opèrent sur des séquences échantillonnées.

Dans cette partie, nous considérons d'abord une formulation à temps continue pour étendre, dans les chapitres qui suivent, l'approche au temps discret. Nous notons par γ l'ensemble des paramètres inconnus qui inclut la fréquence porteuse, l'offset de phase, le retard introduit par le canal et les symboles transmis dans le cas d'une estimation aveugle. Nous adoptons la notation $x(t, \gamma)$ pour le signal reçu en absence de bruit qui met en évidence la dépendance en γ . Le modèle en bande de base est :

$$y(t) = x(t, \gamma) + w(t), \quad (1.1)$$

où $w(t)$ est le bruit additif complexe. On considère que $y(t)$ est une réalisation d'un processus aléatoire $y(t)$ pour une valeur donnée de $\tilde{\gamma} = \gamma$. En effet, une réalisation de $y(t)$ a un certain degré de ressemblance avec $y(t)$ dépendamment de la ressemblance entre $x(t, \tilde{\gamma})$ et $x(t, \gamma)$, en d'autres termes, la distance entre $\tilde{\gamma}$ et γ . L'estimateur à maximum de vraisemblance est basé sur le calcul de $\tilde{\gamma}$ de sorte que la ressemblance entre $y(t)$ et la réalisation $y(t)$ soit maximale. En termes de probabilité, nous appelons $p(y(t)|\tilde{\gamma})$ la densité de probabilité de $y(t)$ conditionnée par $\tilde{\gamma}$. Supposons que pour deux réalisations de $\tilde{\gamma}$, notées par $\tilde{\gamma}_1$ et $\tilde{\gamma}_2$, nous avons :

$$p(y(t) = y(t)|\tilde{\gamma}_1) < p(y(t) = y(t)|\tilde{\gamma}_2), \quad (1.2)$$

alors $\tilde{\gamma}_2$ est dit plus vraisemblable que $\tilde{\gamma}_1$.

Comme nous l'avons mentionné, le but est de maximiser $p(y(t) = y(t)|\tilde{\gamma})$ par rapport à $\tilde{\gamma}$. La position du maximum est appelée estimé à maximum de vraisemblance et est donnée

par :

$$\hat{\gamma}_{ML} = \arg \max_{\tilde{\gamma}} \{p(y(t) = y(t)|\tilde{\gamma})\}. \quad (1.3)$$

Cependant, $\tilde{\gamma}$ inclut, en plus des paramètres de synchronisation, les symboles inconnus qu'on ne cherche pas à estimer à ce niveau. C'est pour cette raison que les paramètres à estimer sont rassemblés dans λ et les autres paramètres, appelés paramètres de nuisance, sont rassemblés dans n . Maintenant, nous devons reformuler l'estimateur en (1.3) pour tenir compte de n . Soient $\tilde{\lambda}$ et \tilde{n} deux valeurs hypothétiques de λ et n , respectivement. En modélisant n comme un vecteur aléatoire de densité de probabilité $p(n)$, la formule des probabilités totales permet d'écrire :

$$p(y(t) = y(t)|\tilde{\lambda}) = \int_{-\infty}^{+\infty} p(y(t) = y(t)|\tilde{\gamma}) p(\tilde{n}) d\tilde{n}. \quad (1.4)$$

Puis, l'estimateur à maximum de vraisemblance de λ est :

$$\hat{\lambda}_{ML} = \arg \max_{\tilde{\lambda}} \{p(y(t) = y(t)|\tilde{\lambda})\}. \quad (1.5)$$

Dans la suite de ce chapitre, ainsi que dans les chapitres 2 et 3, nous nous concentrons sur le cas où λ est un scalaire (le délai τ) et nous traitons le cas où λ est un vecteur dans les chapitres 4 et 5.

1.4 Estimateurs du retard à maximum de vraisemblance

1.4.1 Estimateur pour Faible RSB

Dans cette section, nous nous intéressons au cas aveugle (symboles non connus). Nous présentons le traditionnel estimateur pour faible valeurs de RSB développé en [3]. La fonction de vraisemblance est :

$$\Lambda(y|\tau, \mathbf{c}) = \exp \left\{ \frac{2}{N_0} \int_0^T y(t)x(t)dt - \frac{1}{N_0} \int_0^T x^2(t)dt \right\}, \quad (1.6)$$

où N_0 est la puissance du bruit, $\mathbf{c} = \{c_0, c_1, \dots, c_L\}$ représente la séquence de données inconnue et $x(t)$ est définie comme suit :

$$x(t) = \sum_{i=0}^L c_i h(t - iT - \tau). \quad (1.7)$$

L'objectif est d'estimer le retard τ sans avoir une connaissance *a priori* des symboles. La méthode la plus directe est de considérer les données \mathbf{c} comme un vecteur aléatoire. La fonction de vraisemblance est moyennée par rapport aux données transmises pour obtenir la fonction de vraisemblance inconditionnelle :

$$\Lambda(y|\tau) = E_{\mathbf{c}} \{ \Lambda(y|\tau, \mathbf{c}) \}. \quad (1.8)$$

Malheureusement, l'expression analytique de la fonction de vraisemblance inconditionnelle est difficile à évaluer. C'est pour cette raison que quelques approximations sont utilisées. Premièrement, le second terme dans l'expression de la fonction de vraisemblance en (1.6) est ignoré. Cette simplification entrainera une dégradation des performances d'estimation. Ensuite, une approximation de la fonction résultante peut être envisagée en utilisant le développement de Taylor sous l'hypothèse de faibles valeurs de RSB. Alors nous obtenons [3] :

$$\Lambda(y|\tau, \mathbf{c}) \approx 1 + \frac{2}{N_0} \int_0^T y(t)x(t)dt + \frac{2}{N_0^2} \left[\int_0^T y(t)x(t)dt \right]^2. \quad (1.9)$$

Considérant la définition (1.7), nous avons :

$$\int_0^T y(t)x(t)dt = \sum_{i=0}^{L-1} c_i \int_0^T y(t)h(t - iT - \tau)dt. \quad (1.10)$$

Troisième approximation, puisque les coefficients de $h(t)$ sont négligeables pour $t \notin [0, T]$, les limites de l'intégrale dans (1.10) peuvent être élargies à l'infini pour obtenir :

$$\int_0^T y(t)x(t)dt \approx \sum_{i=0}^{L-1} c_i r(iT + \tau), \quad (1.11)$$

avec

$$r(t) = \int_{-\infty}^{+\infty} y(v)h(v - t)dv. \quad (1.12)$$

Maintenant, nous substituons (1.11) dans (1.9) et tenant compte de l'hypothèse $E\{c_i\} = 0$, la fonction de vraisemblance inconditionnelle s'écrit comme suit :

$$\Lambda(y|\tau) = \sum_{i=0}^{L-1} r^2(iT + \tau), \quad (1.13)$$

où les termes constants sont ignorés. De cette façon, une expression plus pratique de la fonction de vraisemblance inconditionnelle est trouvée. C'est une expression quadratique qui implique la réponse du filtre adapté au signal reçu. A partir de (1.13), la valeur optimale du retard τ maximise l'énergie de la séquence $\{r(iT + \tau)\}$. Pour trouver cette valeur optimale, il faut annuler la dérivée de $\Lambda(y|\tau)$ par rapport à τ :

$$\Lambda'(y|\tau) \approx 2 \sum_{i=0}^{L-1} r(iT + \tau)r'(iT + \tau). \quad (1.14)$$

Pour $\tau = \tau_k$, la dérivée de la fonction de vraisemblance inconditionnelle est considérée comme une erreur d'estimation et utilisée pour estimer le retard de façon itérative :

$$\tau_{k+1} = \tau_k + \mu e(k), \quad (1.15)$$

avec

$$e(k) = r(kT + \tau_k)r'(kT + \tau_k), \quad (1.16)$$

et μ correspond au pas d'adaptation.

1.4.2 Estimateur à Maximum de Vraisemblance Conditionnel

Nous avons vu que suite à plusieurs approximations, le critère du maximum de vraisemblance se formule d'une manière pratique. Cependant, l'algorithme est dérivé sous l'hypothèse d'un faible RSB. Cette hypothèse limite les performances de l'estimateur pour les larges valeurs du RSB de telle façon que la différence entre la borne de Cramer-Rao et les performances d'estimation est d'autant plus importante que le RSB augmente. En effet, l'estimateur à faible RSB est une approximation de l'estimateur à maximum de vraisemblance. À grand RSB, cette approximation ne reflète plus la vraie fonction de vraisemblance et les performances de l'estimateur ne sont plus optimales. C'est pour cette raison qu'un autre algorithme, connu sous le nom de maximum de vraisemblance conditionnelle (MVC), a été développé [4]. Afin d'éviter les approximations, une approche différente est utilisée pour calculer la fonction de vraisemblance. Contrairement à la première méthode présentée en 1.4.1, les symboles reçus sont modélisés comme déterministes et inconnus. Les paramètres de nuisance (données, phase) sont exprimés en fonction du paramètre à estimer et du signal reçu. De cette façon, nous obtenons une fonction qui dépend seulement du paramètre inconnu qui est le retard à estimer. À partir de cette fonction, un signal d'erreur est calculé puis utilisé pour mettre à jour l'estimé du retard.

Nous présentons plus de détails sur ce point dans chapitre 2.

1.5 Limites de performance des estimateurs

1.5.1 Les Bornes de Cramér-Rao

Considérons n'importe quelle méthode d'estimation de τ et notons par $\hat{\tau}$ l'estimé correspondant. Vu que $\hat{\tau}$ dépend de l'observation y , différentes observations engendrent différents estimés. Dans ce cas $\hat{\tau}$ est une variable aléatoire dont la moyenne peut coïncider avec les vraies valeurs de τ . Dans ce cas, l'estimateur est dit non biaisé. Cette propriété est un caractère d'évaluation des performances d'estimation puisque, en moyenne, l'estimateur fournit la vraie valeur du paramètre. Cependant, l'erreur d'estimation $\hat{\tau} - \tau$ est aussi une mesure importante, d'où la nécessité de minimiser cette erreur, ou encore sa variance. Alors quand est ce qu'on peut dire que l'erreur d'estimation est acceptable ?

Dans ce contexte, la borne de Cramér-Rao est une limite théorique qui fournit une borne inférieure pour la variance de tout estimateur non-biaisé [2] :

$$\text{var}\{\hat{\tau} - \tau\} \geq \text{CRLB}(\tau), \quad (1.17)$$

avec

$$\begin{aligned} \text{CRLB}(\tau) &= -\frac{1}{\text{E}\left\{\frac{\partial^2 \ln(\Lambda(r|\tau))}{\partial \tau^2}\right\}} \\ &= \frac{1}{\text{E}\left\{\left(\frac{\partial \ln(\Lambda(r|\tau))}{\partial \tau}\right)^2\right\}}. \end{aligned} \quad (1.18)$$

Dans les problèmes de synchronisation, l'application de cette borne est difficile à cause de la complexité de $\Lambda(r|\tau)$. En effet, nous devons moyenner $\Lambda(r|\tau, \mathbf{u})$ par rapport aux paramètres de nuisance \mathbf{u} :

$$\Lambda(r|\tau) = \int_{-\infty}^{+\infty} \Lambda(r|\tau, \mathbf{u})p(\mathbf{u})d\mathbf{u} \quad (1.19)$$

qui, jusqu'à présent, n'a pas d'expression analytique. Nous présentons dans le Chapitre 3 une méthode pour dériver l'expression analytique de la borne de Cramér-Rao pour l'estimation du délai et ceci pour une large gamme de modulations.

Une alternative à la borne de Cramér-Rao, connue sous le nom de borne de Cramér-Rao modifiée, est présentée en [5] pour contourner les problèmes de calcul.

1.5.2 Les Bornes de Cramér-Rao Modifiées

Une approche modifiée, proposée par D'Andrea, Mengali et Reggiannini [5] est souvent utilisée. Il s'agit de la borne de Cramer Rao modifiée (MCRLB pour Modified Cramer Rao Bound). Cette nouvelle borne s'écrit :

$$\text{MCRLB} = \frac{N_0}{\text{E}_{\mathbf{u}}\left\{\int_0^{KT} \left\|\frac{\partial x(t, \tau, \mathbf{u})}{\partial \tau}\right\|^2 dt\right\}}. \quad (1.20)$$

Dans (1.20), la moyenne $\text{E}_{\mathbf{u}}\{\cdot\}$ est effectuée sur les paramètres de nuisance \mathbf{u} et K est le nombre de symboles dans l'intervalle d'observation. La notation $x(t, \tau, \mathbf{u})$ est introduite afin de séparer le paramètre à estimer τ des paramètres de nuisance \mathbf{u} .

Deux remarques doivent être soulignées à ce niveau sur la MCRLB. Premièrement, dans les étapes de dérivation de la borne, nous supposons que :

- les densités de probabilité de la phase et de la fréquence sont connues ;
- les symboles $\{c_i\}$ sont des variables aléatoires indépendantes, de moyenne nulle et $\text{E}\{(c_i)^2\} = C$.

De plus, D'Andrea a montré que la borne de Cramér-Rao est supérieure à la borne modifiée [5]. Il est évident qu'il y a égalité si le vecteur \mathbf{u} est parfaitement connu. Notez que cette remarque reste valide quelque soit le paramètre à estimer.

La dérivation de la MCRLB peut être effectuée de la façon suivante :

$$\mathbf{E}_{\mathbf{u}} \left\{ \int_0^{KT} \left\| \frac{\partial x(t, \tau, \mathbf{u})}{\partial \tau} \right\|^2 dt \right\} = \int_0^{KT} \mathbf{E}_{\mathbf{u}} \left\{ \|m'(t)\|^2 \right\} dt, \quad (1.21)$$

avec

$$m'(t) = \sum_i c_i p(t - iT - \tau), \quad (1.22)$$

et $p(t) = dh(t)/dt$. Comme $m'(t)$ est indépendante de f_c et θ , le moyennage dans (1.21) se limite au moyennage sur les symboles émis. Par suite, nous obtenons :

$$\begin{aligned} \mathbf{E}_{\mathbf{u}} \{ \|m'(t)\|^2 \} &= C \sum_i p(t - iT - \tau) \\ &= \frac{C}{T} \sum_n P_2 \left(\frac{n}{T} \right) e^{j2\pi n(t-\tau)/T}, \end{aligned} \quad (1.23)$$

où $P_2(f)$ représente la transformée de Fourier de $p^2(t)$. La dernière égalité dans (1.23) est vérifiée par la formule de Poisson. Ensuite, considérant (1.21) et (1.23) et notant que $\int_0^{KT} e^{j2\pi n(t-\tau)/T} dt = \delta(n)KT$, nous obtenons :

$$\begin{aligned} \mathbf{E}_{\mathbf{u}} \left\{ \int_0^{KT} \left\| \frac{\partial x(t, \tau, \mathbf{u})}{\partial \tau} \right\|^2 dt \right\} &= \frac{C}{T} \sum_n P_2 \left(\frac{n}{T} \right) \int_0^{KT} e^{j2\pi n(t-\tau)/T} dt \\ &= KCP_2(0). \end{aligned} \quad (1.24)$$

Cependant, $P_2(0)$ est en relation directe avec $H(f)$, la transformée de Fourier de $h(t)$:

$$P_2(0) = \int_{-\infty}^{+\infty} \left(\frac{d(h(t))}{dt} \right)^2 = 4\pi \int_{-\infty}^{+\infty} f^2 |H(f)|^2 df. \quad (1.25)$$

Finalement, en intégrant (1.24) dans (1.20), nous obtenons le résultat désiré :

$$\text{MCRLB}(\tau) = \frac{T^2}{8\pi^2 C_h K} N_0 E_s, \quad (1.26)$$

où E_s est l'énergie moyenne du signal émise par symbole et C_h est le carré de la largeur de bande moyenne normalisée qui dépend de la fonction de mise en forme et sont donnés, respectivement, par :

$$E_s = \frac{C}{2} \int_{-\infty}^{+\infty} |H(f)|^2 df, \quad (1.27)$$

$$C_h = T^2 \frac{\int_{-\infty}^{+\infty} f^2 |H(f)|^2 df}{\int_{-\infty}^{+\infty} |H(f)|^2 df}. \quad (1.28)$$

L'expression analytique en (1.26) montre que la MCRLB est inversement proportionnelle au rapport signal sur bruit E_s/N_0 et à la taille de l'observation. De plus, elle est inversement proportionnelle à C_h , le carré de la largeur de bande de $H(f)$. Ceci veut dire que l'estimation du retard est plus facile à effectuer avec des signaux à large bande. Ceci peut s'expliquer intuitivement par le fait que les filtres de mise en forme à large bande ont relativement une petite durée temporelle et donc peuvent être mieux détectés en présence de bruit.

1.6 Conclusion

Dans ce chapitre, nous avons présenté les caractéristiques de l'estimateur à maximum de vraisemblance. L'absence d'une expression analytique de l'estimé a favorisé le développement de plusieurs algorithmes d'estimation avec des complexités et des performances variées. Dans ce contexte, la dérivation de bornes de précision d'estimation est importante puisque celles-ci représentent une limite théorique à la performance des estimateurs. Les bornes de Cramér-Rao indiquent les limites inférieures pour la variance de l'erreur d'estimation. Cependant, leur application au problème de synchronisation résulte en de sérieuses difficultés mathématiques. Les bornes de Cramér-Rao modifiées sont beaucoup plus faciles à calculer, mais elles se détachent des vraies bornes de Cramér-Rao de façon imprévisible et donc ne reflètent pas les vraies performances possibles.

Bibliographie

- [1] J. W. M. Bergmans, *Digital Baseband Transmission and Recording*, Boston : Kluwer Academic Publishers, 1996.
- [2] S. Kay, *Fundamentals of Statistical Signal Processing, Estimation theory*, Prentice Hall, Englewood Cliffs, NJ, 1993.
- [3] U. Mengali and A. N. D'Andrea, "Synchronization Techniques for Digital Receivers", New York : Plenum, 1997.
- [4] J. Riba, J. Sala, and G. Vazquez, "Conditional Maximum Likelihood Timing Recovery : Estimators and Bounds," in *IEEE Trans. on Sig. Process*, vol. 49, pp. 835-850, Apr. 2001.
- [5] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its applications to synchronization problems," *IEEE Trans. Comm.*, vol. 24, pp. 1391-1399, Feb.-Apr. 1994.

Chapitre 2

A Non-Data-Aided Maximum Likelihood Time Delay Estimator using Importance Sampling

¹A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, “A Non-Data-Aided Maximum Likelihood Time Delay Estimator Using Importance Sampling”, *IEEE Trans. on Sign. Process.*, vol. 59, no. 10, pp. 4505-4515, October 2011. IEEE (c)

DOI: 10.1109/TSP.2011.2161293

Abstract

Dans cet article, nous présentons un nouvel estimateur à maximum de vraisemblance pour le retard de propagation basé sur "importance sampling" (IS). Nous montrons que la recherche exhaustive et les problèmes de convergence dont souffrent les méthodes itératives peuvent être contournés. Les données transmises sont supposés inconnues. Le retard reste constant sur l'intervalle d'observation et le signal est entaché par un bruit blanc gaussien. Nous utilisons IS pour trouver le maximum global de la fonction de vraisemblance. L'idée du nouvel estimateur est de générer des réalisations à partir d'une version simplifiée de la fonction de vraisemblance. Nous verrons que les paramètres de l'algorithme affectent les performances d'estimation et qu'avec un choix approprié de ces paramètres, le retard peut être estimé de façon précise.

In this paper, we present a new time delay maximum likelihood estimator based on importance sampling (IS). We show that a grid search and lack of convergence from which most iterative estimators suffer can be avoided. It is assumed that the transmitted data are completely unknown at the receiver. Moreover the carrier phase is considered as an unknown nuisance parameter. The time delay remains constant over the observation interval and the received signal is corrupted by additive white Gaussian noise (AWGN). We use importance sampling to find the global maximum of the compressed likelihood function. Based on a global optimization procedure, the main idea of the new estimator is to generate realizations of a random variable using an importance function, which approximates the actual compressed likelihood function. We will see that the algorithm parameters affect the estimation performance and that with an appropriate parameter choice, even over a small observation interval, the time delay can be accurately estimated at far lower computational cost than with classical iterative methods.

2.1 Introduction

Parameter estimation is a crucial operation for any digital receiver ; in particular the recovery of time delay introduced by the channel. Typically, in network communications, the time delay is usually assumed to be confined within the symbol duration [1]. Particularly, symbol timing recovery allows for sampling the signal at accurate time instants in order to achieve satisfactory performances. The key task of timing recovery consists in determining the time instants at which the received signal should be sampled in order to perform reliable data recovery. However, in many other applications such as radar or sonar systems [2] [3], where it can exceed the symbol duration, the time delay is used to localize targets. During the last few decades, many time-delay estimators have been developed trying to achieve the well-known Cramér-Rao lower bound (CRLB). A key step in time recovery

schemes is the determination of an objective function from the statistics of the received signal from which an estimate of the time delay can be extracted. In this sense, it is known that the maximum likelihood estimator is an asymptotically efficient estimator, and that it performs close to the CRLB at relatively high SNR values [7], even for short data records. Therefore, it has been subject to intense research. In the case of data-aided transmissions, where the transmitted data are *a priori* completely known, an expression for the global maximum of the log-likelihood function is analytically tractable. However, when the transmitted data are completely unknown (i.e., the parameter of interest should be blindly estimated), the log-likelihood function becomes extremely non-linear and it is difficult to analytically find its global maximum. In this case, maximum likelihood (ML) solutions must be numerically tackled. The grid search technique is the most basic alternative to numerically find the maximum of the non-linear likelihood function. Unfortunately, this technique can be used only if the range of the parameter is confined to a finite interval, otherwise, iterative maximization procedures must be envisaged. The most famous iterative procedures are the Newton-Raphson method [5] and the expectation-maximization algorithm [6]. However, these two prominent methods are known to converge to the ML solution only if the initial guess is close enough to the true unknown parameter value. If not, these iterative algorithms may converge to a local maximum of the likelihood function, or even diverge. To circumvent this problem, these algorithms may use many initial values to improve their performance. But this increases in counterpart their computational complexity without even ultimately warranting their convergence to the global maximum. In this work, we resort to an entirely different approach for the estimation of the time delay parameter. The compressed likelihood function is derived considering the transmitted symbols as unknown but deterministic. Based on this function, an iterative algorithm earlier implemented in [8] performs better in the high SNR region than the low-SNR unconditional ML (UML) timing error detectors (TEDs) [1], but its performance still depends on the initialization value making it therefore prone to severe degradation due convergence uncertainty.

Motivated by these facts, we develop in this paper a new non-iterative approach to find the time delay conditional maximum likelihood (CML) estimates. We implement the CML algorithm in a non-iterative way. We avoid the grid search, essential in traditional iterative approaches, by using the importance sampling technique which has been shown to be a powerful tool in performing NDA ML estimation. In fact, this method was successfully applied to estimate other crucial parameters such as the direction of arrival (DOA) [9], the carrier frequency [10] or the joint DOA-Doppler frequency [11]. The importance sampling technique is used in this paper in the context of time delay estimation. Moreover, we adopt the discrete-time model widely used in the field of sensors array processing [12] and more recently formulated in the context of time-delay estimation [8]. The resulting IS-based estimator attains the modified CRLB (MCRLB) over both the medium and high

SNR regions, whereas the traditional UML TED, being derived under the assumption of low SNR, does not approach the MCRLB at the high SNR region.

The remainder of this paper is organized as follows. In section II, we present the discrete-time signal model that will be used throughout this article. We derive the compressed likelihood function in section III. In section IV, we introduce the importance sampling method that will be used in this article to find the global maximum of the compressed likelihood function. Section V deals with the choice of the importance function and discusses the impact of some parameters on the estimator performance. The newly proposed algorithm is developed in section VI. Simulation results are discussed in section VII and, finally, some concluding remarks are drawn out in section VIII.

2.2 Discrete-Time Signal Model

First, we present a list of notations and definitions that will be used in this article.

$E_x\{\cdot\}$:	the expectation with respect to x .
$\ \cdot\ $:	Euclidean norm.
$(\cdot)^T, (\cdot)^H$:	transposition and conjugate transposition.
SNR :	signal to noise ratio.
IS :	importance sampling.
ML :	maximum likelihood.
CML :	conditional maximum likelihood.
MCRLB :	Modified Cramér-Rao lower bound.
QAM :	quadrature amplitude modulation.
PAM :	pulse-amplitude modulation.

Consider a traditional communication system where on one hand the channel delays the transmitted signal and on the other hand an AWGN with an overall power of N_0 corrupts the received signal as follows :

$$y(t) = \sqrt{E_s} x(t - \tau^*) e^{j\theta} + w(t), \quad (2.1)$$

where τ^* is the unknown time delay to be estimated, θ is the unknown but deterministic channel distortion phase, $w(t)$ is an additive white Gaussian noise (AWGN) with independent real and imaginary parts, each of variance $N_0/2$ and $\sqrt{E_s}$ is the signal amplitude. The unknown transmitted signal $x(t)$ is modeled as follows :

$$x(t) = \sum_{i=0}^{K-1} c_i h(t - iT), \quad (2.2)$$

where K is the number of transmitted symbols in the observation interval, $\{c_i\}_{i=0}^{K-1}$ are the unknown complex-valued symbols, $h(t)$ is the shaping pulse of energy E_h and T is the

symbol's duration.

In the sequel, we outline the discrete-time signal model which was proposed for the first time in [8] to derive an iterative CML timing recovery algorithm. The received signal $y(t)$ is passed through an ideal lowpass filter of bandwidth $F_s/2$ and sampled at a frequency $F_s = 1/T_s = k/T$, where k is a given integer which guarantees that F_s is above the Nyquist rate. Then, the received samples $\mathbf{y} = [y(0), y(T_s), y(2T_s), \dots, y((M-1)T_s)]^T$ can be written in a matrix form as follows :

$$\mathbf{y} = \mathbf{A}_{\tau^*} \mathbf{x} + \mathbf{w}, \quad (2.3)$$

where M is the number of samples of $y(t)$ and \mathbf{w} and \mathbf{A}_{τ^*} are defined as follows :

$$\mathbf{w} = [w(0), w(T_s), \dots, w((M-1)T_s)]^T, \quad (2.4)$$

$$\mathbf{A}_{\tau^*} = [\mathbf{a}_0(\tau^*), \mathbf{a}_1(\tau^*), \dots, \mathbf{a}_{K-1}(\tau^*)], \quad (2.5)$$

with

$$\mathbf{a}_i(\tau^*) = [h(-iT - \tau^*), h(T_s - iT - \tau^*), \dots, h((M-1)T_s - iT - \tau^*)]^T. \quad (2.6)$$

In (2.3), \mathbf{x} is the set of unknown data and signal phase which is given by :

$$\mathbf{x} = \mathbf{c}e^{j\theta} = [c_0, c_1 \dots c_{K-1}]^T e^{j\theta}. \quad (2.7)$$

Moreover, the covariance matrix of \mathbf{w} is given by :

$$\mathbf{C}_w = 2\sigma^2 \mathbf{I}_M = N_0 F_s \mathbf{I}_M, \quad (2.8)$$

where \mathbf{I}_M refers to the $(M \times M)$ identity matrix and $2\sigma^2 = N_0 F_s$. The sampled data \mathbf{y} is a linear function of the vector \mathbf{x} but depends non-linearly on the time delay τ^* . We mention that the model of Eq. (2.3) presented in [8] is inspired from the model widely used in array signal processing where each column of the transfer matrix is a function of a different parameter, usually the direction-of-arrival or the frequency of each incoming signal. In the context of time delay estimation, the entire matrix \mathbf{A}_{τ^*} depends on the same parameter τ^* .

2.3 Likelihood function

The conditional likelihood function of the observed data \mathbf{y} is given by :

$$\Lambda(\mathbf{y}|\mathbf{x}; \tau) \propto p(\mathbf{y}|\mathbf{x}; \tau) = C \exp \left\{ -\frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{A}_\tau \mathbf{x}\|^2 \right\}, \quad (2.9)$$

where $p(\mathbf{y}|\mathbf{x}; \tau)$ is the probability density function (pdf) of \mathbf{y} conditioned on \mathbf{x} and parameterized by τ , and C is a positive constant which does not depend on the time delay

and therefore will be dropped, without loss of generality. Note here that τ is any possible value of the time delay parameter τ^* and that $\Lambda(\mathbf{y}|\mathbf{x}; \tau)$ attains its maximum at $\tau = \tau^*$, i.e., $\tau^* = \arg \max_{\tau} \Lambda(\mathbf{y}|\mathbf{x}; \tau)$.

Actually, one needs to maximize $\Lambda(\mathbf{y}|\mathbf{x}; \tau)$ with respect to τ in order to find the ML solution $\hat{\tau}^*$. However, (2.9) imposes a joint estimation of \mathbf{x} and τ^* , which is very difficult to perform. Therefore, two principal approaches are developed in the literature in order to obtain a likelihood function that depends only on τ . On one side, the unconditional maximum likelihood (UML) estimator introduced in [1] considers the data symbols as random and hence averages the joint likelihood function over \mathbf{x} to obtain a function that depends only on the time delay as follows :

$$\Lambda(\mathbf{y}|\tau) = E_{\mathbf{x}}\{\Lambda(\mathbf{y}|\mathbf{x}; \tau)\}. \quad (2.10)$$

On the other side, the data symbols are modeled as unknown but deterministic in the formulation of the conditional likelihood function. Therefore, $\hat{\mathbf{x}}$, the solution that maximizes (2.9) with respect to \mathbf{x} , for a given τ is used in (2.9) as a substitute of \mathbf{x} . Actually, $\hat{\mathbf{x}}$ which maximizes $\Lambda(\mathbf{y}; \mathbf{x}, \tau)$ also maximizes the log-likelihood function given by :

$$L(\mathbf{y}; \mathbf{x}, \tau) = -\frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{A}_{\tau}\mathbf{x}\|^2. \quad (2.11)$$

Therefore, taking the gradient of $L(\mathbf{y}|\mathbf{x}; \tau)$ with respect to \mathbf{x} and setting it to zero :

$$\frac{\partial L(\mathbf{y}; \mathbf{x}, \tau)}{\partial \mathbf{x}} = -\frac{1}{\sigma^2}(\mathbf{A}_{\tau}^T \mathbf{y} - \mathbf{A}_{\tau}^T \mathbf{A}_{\tau} \mathbf{x}) = 0, \quad (2.12)$$

yields the following result :

$$\begin{aligned} \hat{\mathbf{x}} &= (\mathbf{A}_{\tau}^T \mathbf{A}_{\tau})^{-1} \mathbf{A}_{\tau}^T \mathbf{y} \\ &= \mathbf{A}_{\tau}^{\#} \mathbf{y}, \end{aligned} \quad (2.13)$$

where $\mathbf{A}_{\tau}^{\#} = (\mathbf{A}_{\tau}^T \mathbf{A}_{\tau})^{-1} \mathbf{A}_{\tau}^T$ is the pseudo-inverse of the matrix \mathbf{A}_{τ} . Substituting $\hat{\mathbf{x}}$ into (2.11), one obtains the so-called compressed likelihood function, that depends only on the unknown time delay parameter :

$$L(\mathbf{y}; \tau, \hat{\mathbf{x}}) = -\frac{1}{2\sigma^2} \mathbf{y}^H (\mathbf{I}_K - \mathbf{A}_{\tau} \mathbf{A}_{\tau}^{\#}) \mathbf{y}, \quad (2.14)$$

which can be further simplified by dropping the constant terms to obtain the useful compressed likelihood function denoted by $L_c(\mathbf{y}; \tau)$ as follows :

$$L_c(\mathbf{y}; \tau) = \mathbf{y}^H \mathbf{A}_{\tau} (\mathbf{A}_{\tau}^T \mathbf{A}_{\tau})^{-1} \mathbf{A}_{\tau}^T \mathbf{y}. \quad (2.15)$$

Note that the expression in (2.15) represents the cross-energy between the pseudo-inverse filter $\mathbf{A}_{\tau}^{\#}$ and the sampled matched filter \mathbf{A}_{τ}^T . For τ equal to the timing parameter to be estimated, the filter $\mathbf{A}_{\tau}^{\#}$ becomes a zero-forcing equalizer since the components of $\mathbf{A}_{\tau}^{\#} \mathbf{y}$ are intersymbol interference (ISI)-free (i.e., $\mathbf{A}_{\tau}^{\#} \mathbf{y} = \mathbf{x} + \mathbf{A}_{\tau}^{\#} \mathbf{w}$, see [8]).

2.4 Global Maximization of the Compressed Likelihood Function

To perform maximum likelihood estimation, we have to maximize (2.15) with respect to τ . Unfortunately, a closed-form expression for this optimization problem is not analytically tractable since the objective function in (2.15) is extremely non-linear with respect to τ . Therefore, many methods have been developed to numerically find the maximum, but most of them are iterative [1-8]. We cannot deny that these methods provide good performance in terms of error variance, but unfortunately they require, in counterpart, a sufficiently close initial guess to converge to the global maximum of the likelihood function. Otherwise, the result may be a local maximum, which does not correspond to the true time delay value. This is why a suboptimal algorithm needs to be applied firstly and then its output is considered as an initial value for any iterative technique.

To avoid this challenging drawback of iterative methods, we propose in this paper an entirely different technique which does not claim any initial guess of the time delay parameter. We apply the global maximization method earlier proposed by Pincus [13] which provides a powerful tool for accomplishing nonlinear optimization and guarantees finding the global maximum without any initialization concerns. In fact, the theorem of Pincus states that the maximum of $L_c(\mathbf{y}; \tau)$ is given by :

$$\hat{\tau}^* = \lim_{\rho \rightarrow \infty} \int_J \tau L'_{c,\rho}(\tau) d\tau, \quad (2.16)$$

where

$$L'_{c,\rho}(\tau) = \frac{\exp\{\rho L_c(\mathbf{y}; \tau)\}}{\int_J \exp\{\rho L_c(\mathbf{y}; \tau)\} d\tau}, \quad (2.17)$$

can be viewed as the normalized function of $\exp\{\rho L_c(\mathbf{y}; \tau)\}$. Note that in (2.16) and (2.17), J is the integration interval in which τ is supposed to be confined. In a certain way, $L'_{c,\rho}(\tau)$ can be viewed as a pdf (since it verifies all the properties of a pdf), but since τ is actually deterministic, $L'_{c,\rho}(\tau)$ is more conveniently called a pseudo-pdf [9]. It is also worth noting that, as $\rho \rightarrow \infty$, $L'_{c,\rho}(\tau)$ becomes a Dirac delta function centered at the location of its original maximum. We leave broad details on this point in Appendix A.

The ML estimator for the time delay parameter, obtained from the location of the global maximum of $L_c(\mathbf{y}; \tau)$ is given, for a large value of ρ_0 , by :

$$\hat{\tau}^* = \int_J \tau L'_{c,\rho_0}(\tau) d\tau. \quad (2.18)$$

Now, we need to evaluate the integral given in (2.18), although a direct integration remains always difficult if not impossible. However, this integral is in a way the mean value of a random variable distributed according to $L'_{c,\rho_0}(\cdot)$. It was shown in [14] that this type of

integral can be efficiently evaluated using Monte-Carlo simulations as follows :

$$\hat{\tau}^* = \frac{1}{R} \sum_{k=1}^R \tau_k, \quad (2.19)$$

where $\{\tau_k\}_{k=1}^R$ are realizations of τ distributed according to the pseudo-pdf, $L'_{c,\rho_0}(\tau)$, and hence the global maximization problem reduces simply to a generation of random variables. Yet, since it is a non-linear function of τ , the direct generation of realizations according to $L'_{c,\rho_0}(\tau)$ is computationally hard. Thus, instead of pursuing a fruitless path, we use the importance sampling technique, as done in [9], [10] and [11] for the estimation of the signal directions of arrival, the carrier frequency and the Doppler frequencies, instead of directly using (2.17).

2.5 The Importance Sampling Technique

It has been shown that the importance sampling technique is a powerful tool to compute multiple integrals ; in particular the one given in (2.18). In fact, it can be easily seen that for any function² $f(\cdot)$:

$$\int_J f(\tau) L'_{c,\rho_0}(\tau) d\tau = \int_J f(\tau) \frac{L'_{c,\rho_0}(\tau)}{g'(\tau)} g'(\tau) d\tau, \quad (2.20)$$

where $g'(\cdot)$, called the normalized importance function, is another pseudo-pdf which must be chosen as a simple function of τ so that realizations distributed according to $g'(\cdot)$ can be easily generated. Then, the Monte-Carlo method is used to empirically compute the integral in (2.20) simply via the following summation :

$$\int_J f(\tau) L'_{c,\rho_0}(\tau) d\tau = \frac{1}{R} \sum_{k=1}^R f(\tau_k) \frac{L'_{c,\rho_0}(\tau_k)}{g'(\tau_k)}, \quad (2.21)$$

where τ_k is the k th realization of τ according to the normalized importance function $g'(\cdot)$ and R is the number of realizations. Typically, $g'(\cdot)$ and $L'_{c,\rho_0}(\cdot)$ should be very similar to reduce the variance of the estimates. However, $L'_{c,\rho_0}(\cdot)$ remains a complex function and in counterpart $g'(\cdot)$ needs to be as simple as possible. Therefore, some trade-offs must be found in the construction of the importance function. In fact, the inverse matrix $(\mathbf{A}_\tau^T \mathbf{A}_\tau)^{-1}$ in the actual compressed likelihood function, $L_c(\mathbf{y}; \tau)$ (or equivalently $L'_{c,\rho_0}(\cdot)$), is very non-linear with respect to τ . Intuitively, one can replace this inverse matrix by the diagonal matrix $\frac{T_s}{E_h} \mathbf{I}_K$. Hence, a reasonable approximation of the compressed likelihood function is :

$$L_c(\mathbf{y}; \tau) \approx \frac{T_s}{E_h} \mathbf{y}^H \mathbf{A}_\tau \mathbf{A}_\tau^T \mathbf{y}. \quad (2.22)$$

²In our case, we have $f(\tau) = \tau$.

The approximation of $(\mathbf{A}_\tau^T \mathbf{A}_\tau)^{-1}$ with $\frac{T_s}{E_h} \mathbf{I}_K$ is very reasonable for most of the conventional pulse shaping functions. For instance, it can be verified that for the widely used square root-raised cosine pulse, the diagonal elements of $(\mathbf{A}_\tau^T \mathbf{A}_\tau)^{-1}$ are dominant compared to its off-diagonal ones. In fact, as defined in (2.5), the columns of \mathbf{A}_τ are built upon shifted versions of the shaping pulse $h(\cdot)$, therefore every element of $\mathbf{A}_\tau^T \mathbf{A}_\tau$ can be seen as the convolution of two shifted versions of $h(\cdot)$ (the shift being an integer multiple of T), which value is maximum when the shift is the same, i.e., in the diagonal elements. Whereas, when the shift is not the same, the value of the convolution is very low. See Appendix B for more details about this observation. In the particular case where the pulse shape does not generate inter-symbol interference, the approximation becomes strict equality and (2.22) yields the exact compressed likelihood function. Then, a reasonable importance function is given by :

$$g_{\rho_1}(\tau) = \exp \left\{ \rho_1 \sum_{k=0}^{K-1} \frac{T_s}{E_h} \left| \sum_{i=0}^{M-1} y^*(iT_s) h(iT_s - kT - \tau) \right|^2 \right\}, \quad (2.23)$$

where ρ_1 is another constant different³ from ρ_0 . Note that the normalization of $g_{\rho_1}(\cdot)$ by $\int_J g_{\rho_1}(x) dx$ yields the normalized importance function $g'_{\rho_1}(\cdot)$ (i.e., $g'_{\rho_1}(\tau) = \frac{g_{\rho_1}(\tau)}{\int_J g_{\rho_1}(x) dx}$). But since the periodogram of the data evaluated at the time delay τ , $I_k(\tau)$, is given by :

$$I_k(\tau) = \left| \sum_{i=0}^{M-1} y^*(iT_s) h(iT_s - kT - \tau) \right|^2, \quad k = 0, 1, 2, \dots, K-1, \quad (2.24)$$

then, we rewrite the importance function as follows :

$$g_{\rho_1}(\tau) = \prod_{k=0}^{K-1} \exp \left\{ \rho'_1 I_k(\tau) \right\}, \quad (2.25)$$

with

$$\rho'_1 = \frac{T_s}{E_h} \rho_1. \quad (2.26)$$

The normalization of (2.25) leads to the pseudo-pdf $g'(\cdot)$ which will be used, hereafter, to generate the realizations involved in (21) :

$$g'_{\rho'_1}(\tau) = \frac{\prod_{k=0}^{K-1} \exp\{\rho'_1 I_k(\tau)\}}{\int_J \prod_{k=0}^{K-1} \exp\{\rho'_1 I_k(v)\} dv}. \quad (2.27)$$

It is also worth noting that the performance of the new maximum-likelihood estimator

³In our case, ρ_1 can be equal to ρ_0 , unlike for the multiple parameters estimation where ρ_1 should be different from ρ_0 .

depends on the choice of ρ'_1 . In fact, our ultimate goal is to find the global maximum of the function $L_c(\mathbf{y}; \tau) = \mathbf{y}^H \mathbf{A}_\tau (\mathbf{A}_\tau^T \mathbf{A}_\tau)^{-1} \mathbf{A}_\tau^T \mathbf{y}$. However, this function exhibits many local maxima even in the total absence of noise, and it is often difficult to distinguish between the global and a local maximum. For this purpose, ρ'_1 is chosen to render the objective function in (2.27) more peaked around its global maximum which will have a relatively higher peak compared to the local maxima. This behavior is illustrated in Fig. 2.1, which plots the function $g'_{\rho'_1}(\cdot)$ for $\rho'_1 = 10$ and $\rho'_1 = 20$, in the total absence of the additive noise. Moreover, we show in Appendix C how this parameter renders $g'_{\rho'_1}(\cdot)$ more peaked around its global maximum.

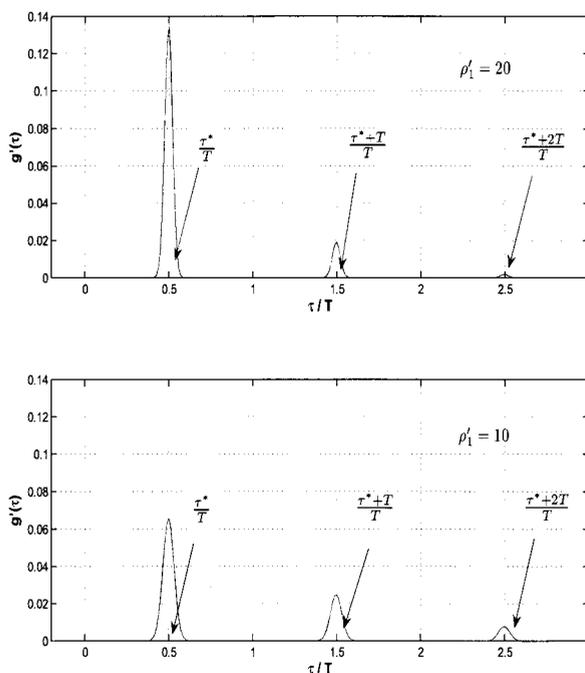


FIGURE 2.1 – Plot of $g'_{\rho'_1}(\cdot)$ for $\rho'_1 = 20$ and $\rho'_1 = 10$ using a root-raised cosine pulse and for $K = 100$.

Based on this fact, it can be stated, a priori, that it is better to arbitrarily increase ρ'_1 in order to achieve better performance. But, this is much easier said than done since, in practice, this leads to numerical overflows. Actually, the best value of ρ'_1 is the highest possible without resulting in any overflow in the computation of $g'_{\rho'_1}(\cdot)$ as it will be seen in Section VII. Same argument is valid for ρ_0 . In fact, the approximation of $(\mathbf{A}_\tau^T \mathbf{A}_\tau)^{-1}$ by $\frac{T_s}{E_h} \mathbf{I}_K$ means that the quantity $L_c(\mathbf{y}; \tau) = \mathbf{y}^H \mathbf{A}_\tau (\mathbf{A}_\tau^T \mathbf{A}_\tau)^{-1} \mathbf{A}_\tau^T \mathbf{y}$ is almost equal to $\frac{T_s}{E_h} L_c^{(a)}(\tau)$ where $L_c^{(a)}(\tau) = \mathbf{y}^H \mathbf{A}_\tau \mathbf{A}_\tau^T \mathbf{y}$ (note that the superscript (a) in $L_c^{(a)}(\tau)$ refers to the word "approximate" where we approximate $L_c(\mathbf{y}; \tau)$ with $L_c^{(a)}(\tau)$ by replacing $(\mathbf{A}_\tau^T \mathbf{A}_\tau)^{-1}$ by $\frac{T_s}{E_h} \mathbf{I}_K$). This means that ρ_0 will result in an overflow in $\exp\{\rho_0 L_c(\mathbf{y}; \tau)\}$ as far as ρ_1

results in an overflow in $\exp\left\{\rho_1 \frac{T_s}{E_h} L_c^{(a)}(\tau)\right\}$. Therefore, the optimal value, ρ_0^{opt} , of ρ_0 verifies $\rho_0^{opt} = \rho_1^{opt}$ where ρ_1^{opt} is again the optimal value of $\rho_1 = \frac{E_h}{T_s} \rho_1'$.

We also note that the matrix \mathbf{A}_τ in (2.22) exhibits an interesting structure since its columns are simply shifted versions of the pulse shape. Hence, the matrix-by-vector operation $\mathbf{A}_\tau^T \mathbf{y}$ can be viewed as a filtering operation of the received samples \mathbf{y} with a filter $h(\cdot)$ whose coefficients are the central row of \mathbf{A}_τ^T . Therefore, the computation of (2.27) is quite simple and realizations distributed according to $g'_{\rho_1'}(\cdot)$ in (2.27) can be easier generated than according to $L'_{c,\rho_0}(\cdot)$ in (2.17).

2.6 Estimation of the Time Delay

In the sequel, two scenarios will be proposed depending on the range of the unknown time delay parameter. In the first scenario, we assume that the time delay parameter takes values within $[0, PT]$, where P is a strictly positive integer (i.e., $P \geq 1$). In the second scenario, we assume that the time delay parameter takes values within $[0, T]$.

2.6.1 First scenario : $\tau^* \in [0, PT]$

As already mentioned in the introduction, in many applications such as radar or sonar transmissions, the actual time delay introduced by the channel may exceed the symbol's duration. In this subsection, we assume however that the time delay does not exceed PT , where P is a given strictly positive integer, i.e., $\tau^* \in [0, PT]$. This upper limitation of the interval is justified since, in each communication system, we always have an *a priori* idea about the maximum range⁴ of τ . As we have seen in the previous section, the maxima of $g'_{\rho_1'}(\cdot)$ are periodic, with period T . Therefore, many secondary peaks may appear which ultimately affects the estimate $\hat{\tau}^*$ of τ^* . In fact, to obtain unbiased estimates of τ^* , the expected value of the estimation error $\tau_e = \tau^* - \hat{\tau}^*$ should be equal to zero, i.e. :

$$\mathbb{E}\{\tau_e\} = \mathbb{E}\{\tau^* - \hat{\tau}^*\} = 0. \quad (2.28)$$

However, it may occur that the difference between τ^* and $\hat{\tau}^*$ is very important. In fact, to simplify, assume that $g'_{\rho_1'}(\cdot)$ has only 2 peaks and neglect the others. Then the generated values will take values around τ^* and $T + \tau^*$, with higher probability around τ^* where the highest peak is located. If we denote by C_1 the set of realizations taking values near τ^* and C_2 the set of realizations taking values near $\tau^* + T$, then from (2.21) the estimated, $\hat{\tau}^*$, can be approximated by :

$$\hat{\tau}^* = \frac{\text{card}(C_1)}{R} \tau^* + \frac{\text{card}(C_2)}{R} (\tau^* + T), \quad (2.29)$$

⁴Note that P can be always chosen as large as desired to ensure that $\tau^* \in [0, PT]$

with $\text{card}(C)$ denoting the cardinal of C , i.e., the number of elements of C . Therefore $\hat{\tau}^* \neq \tau^*$ since $R = \text{card}(C_1) + \text{card}(C_2)$ and $\text{card}(C_2)$ is always not equal to zero. Moreover, the bias is larger at low SNRs and/or short data records. This property was also previously observed in the case of frequency estimation in [15].

To circumvent this problem, the pseudo-pdf, $g'_{\rho'_1}(\cdot)$, must be centered around τ^* . To that end, two intuitive methods may be envisioned. We may either eliminate the secondary peaks to keep only the principal one, or we can generate other peaks in a way that the number of secondary lobes on either side of the principal lobe is the same. The first idea seems to be the most efficient, but it is unfortunately unrealizable and we opt for the second alternative. Indeed, as we have seen, the estimation bias stems from the peaks taking place after the principal lobe. Thus we have to modify $g'_{\rho'_1}(\cdot)$ so that it becomes quasi-symmetric around τ^* . To that end, the simplest way is to suppose, virtually, that τ takes negative values although τ is always positive. We extend the interval of definition of $g'_{\rho'_1}(\cdot)$ from $[0, PT]$ to $[-QT, PT]$, where Q is a positive integer smaller than P . In this way, virtual secondary lobes appear before as well as after τ^* . Moreover, as it can be seen from Fig. 2.2, the probability of generating realizations around $\tau^* + iT$ is almost the same as the one of generating realizations around $\tau^* - iT$. Hence, the estimator becomes unbiased and the estimate $\hat{\tau}^*$ is more accurate. So far, we have established an unbiased

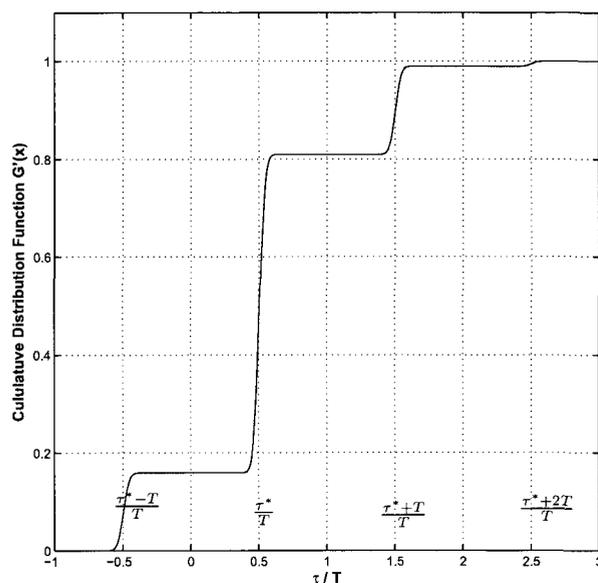


FIGURE 2.2 – Plot of the cumulative distribution function (CDF), $G'(\tau)$ whose pdf is $g'(\tau)$, SNR = 5 dB.

estimator based on a linear average of the generated realizations. But through simulations, we noted a performance change according to the constellation type. In fact, for a constant-

envelope constellation such as phase-shift keying (PSK), the estimator works perfectly. However, its performance degrades dramatically for non-constant-modulus constellations such as pulse-amplitude modulation (PAM), quadrature amplitude modulation (QAM), etc. In fact, as we have previously seen, the main problem that faces the new ML estimator is the presence of secondary peaks. Although we have explained, in section V, how to reduce the adverse effects of these peaks, they generate irreversible errors in the case of non-constant-envelope constellations. To simplify the problem, without loss of generality, we suppose again that $g'_{\rho'_1}(\cdot)$ exhibits only two peaks, one located at τ^* and the other located at $\tau^* + T$. From (2.25), $g_{\rho'_1}(\tau^*)$ and $g_{\rho'_1}(\tau^* + T)$ can be written as follows :

$$g_{\rho'_1}(\tau^*) = \exp \left\{ \rho'_1 \sum_{k=1}^{K-1} \left| \sum_{i=0}^{M-1} y^*(iT_s) h(iT_s - kT - \tau^*) \right|^2 \right\} \exp \left\{ \rho'_1 \left| \sum_{i=0}^{M-1} y^*(iT_s) h(iT_s - \tau^*) \right|^2 \right\}, \quad (2.30)$$

and

$$g_{\rho'_1}(\tau^* + T) = \exp \left\{ \rho'_1 \sum_{k=1}^{K-1} \left| \sum_{i=0}^{M-1} y^*(iT_s) h(iT_s - kT - \tau^*) \right|^2 \right\} \times \exp \left\{ \rho'_1 \left| \sum_{i=0}^{M-1} y^*(iT_s) h(iT_s - KT - \tau^*) \right|^2 \right\}. \quad (2.31)$$

Noting that the samples of the received signal are limited in time to $[0, KT]$ and the magnitudes of $h(t - KT - \tau)$ for $t \in [0, KT]$ are very small compared to those of $h(t - KT - \tau)$ for $t \in [KT, (K+1)T]$, then, the term $\sum_{i=0}^{M-1} y^*(iT_s) h(iT_s - KT - \tau)$ can be neglected. Considering this result, we express $g_{\rho'_1}(\tau^*)$ as a function of $g_{\rho'_1}(\tau^* + T)$:

$$g_{\rho'_1}(\tau^*) \approx g_{\rho'_1}(\tau^* + T) \exp\{\rho'_1 I_0(\tau^*)\}, \quad (2.32)$$

where $I_0(\tau^*)$ depends mainly on the amplitude of the first symbol. In practice, it may occur that the amplitude of the first transmitted symbol is the smallest one. In this case, the contribution of $\exp\{\rho'_1 I_0(\tau^*)\}$ in $g_{\rho'_1}(\tau^*)$ is far less important than the other terms, i.e., $\exp\{\rho'_1 I_0(\tau^*)\} \ll \exp\{\rho'_1 I_i(\tau^*)\}$ for $i = 1, 2, \dots, K-1$. As a result, $g_{\rho'_1}(\tau^*)$ will be closer to $g_{\rho'_1}(\tau^* + T)$, which is a local maximum making the estimate $\hat{\tau}^*$ shift toward $\tau^* + T$. The same problem occurs when the last transmitted symbol has the smallest amplitude with the only difference that the shift will be toward $\tau^* - T$.

To avoid these problems, $I_0(\tau^*)$ must be as large as possible. To that end, we slightly modify the algorithm in the case of non-constant-modulus constellations by sending two *a priori* known symbols : one at the beginning and one at the end of the frame. Moreover, these two known symbols must be of highest energy among the constellation points. Then,

$I_0(\tau^*)$ is no longer negligible compared to $I_i(\tau^*)$, for $i \neq 0$, and the difference between the magnitude of $g_{\rho_1}(\tau^*)$ and $g_{\rho_1}(\tau^* + T)$ is large enough to avoid an important detection error. The same thing holds for $g_{\rho_1}(\tau^*)$ and $g_{\rho_1}(\tau^* - T)$.

2.6.2 Second scenario : $\tau^* \in [0, T]$

In many cases, the time delay does not exceed the symbol duration T . Therefore, we must look for the global maximum only in $[0, T]$. As previously explained, the maxima of the importance function are periodically located, with a period equal to T . Moreover, since we know *a priori* that τ does not exceed T , then we can more conveniently use the circular⁵ (instead of the linear) mean to evaluate the mean in (18). It will be seen in section VII that the use of the circular mean provides considerable performance enhancements in the low-SNR region. As it will be explained later, the use of the circular mean considerably reduces the computational cost.

To introduce the concept of a circular mean, consider a circular random variable which takes values in a finite interval that can be mapped into the unit circle. For instance, let α be a random variable defined in $[0, 1]$ with pdf $P(\alpha)$. Then, the circular mean of α is defined as :

$$E_c\{\alpha\} = \frac{1}{2\pi} \angle \int \exp\{j2\pi\alpha\} P(\alpha) d\alpha, \quad (2.33)$$

where \angle denotes the angle in radians. Having R realizations of α , its circular mean is [16] :

$$\hat{\alpha} = \frac{1}{2\pi} \angle \frac{1}{R} \sum_{r=1}^R \exp\{j2\pi\alpha_r\}. \quad (2.34)$$

In our case, if the time delay is not confined within the interval $[0, 1]$, it can be easily transposed into this interval by normalizing τ^* by T . Then, the resulting transposed estimate is inversed to obtain an estimate in the original interval. Hence, the IS estimate of τ^* using (2.34) and (2.21) is :

$$\hat{\tau}^* = \frac{T}{2\pi} \angle \frac{1}{R} \sum_{k=1}^R \frac{L'_c(\tau_k)}{g'(\tau_k)} \exp\left\{j \frac{2\pi\tau_k}{T}\right\}, \quad (2.35)$$

or finally :

$$\hat{\tau}^* = \frac{T}{2\pi} \angle \frac{1}{R} \sum_{i=1}^R F(\tau_i) \exp\left\{j \frac{2\pi\tau_i}{T}\right\}, \quad (2.36)$$

where

$$F(x) = \frac{L'_{c,\rho_0}(x)}{g'_{\rho_1}(x)}. \quad (2.37)$$

⁵Note that the circular mean cannot be used in the first scenario when τ may exceed T since it always returns an estimate in $[0, T]$ by virtually bringing, into this interval, all the secondary lobes of the normalized importance function.

Note that we need to find the angle of a complex number, and thus, we can remove any positive real factor taking place in (2.36) without affecting the final result. This means that the two strictly positive normalization constants $\int_J L'(v)dv$ and $\int_J g'(v)dv$ can be simply dropped. Moreover, an overflow may occur since both the numerator and the denominator are exponentials. To circumvent this problem, we replace⁶ $F(\tau_i)$ by $F'(\tau_i)$:

$$F'(\tau_i) = \exp \left\{ \rho_0 L_c(\tau_i) - \rho'_1 \sum_{k=0}^{K-1} I_k(\tau_i) - \max_{1 \leq l \leq R} \left(\rho_0 L_c(\tau_l) - \rho'_1 \sum_{k=0}^{K-1} I_k(\tau_l) \right) \right\} \quad (2.38)$$

where we multiply $F(\tau_i)$ by a real scalar factor.

2.6.3 Summary of steps

In the following, we summarize the steps of the new algorithm for the two considered scenarios.

1. Based on the sampled data $y(iT)$, $i = 0, 1, \dots, M - 1$, evaluate the periodogram $I_k(\tau)$ according to (2.24).
2. Compute the normalized importance function in (2.27). Note that, in practice, we use a discrete model by substituting the integration in the denominator of (2.27) by a summation as follows :

$$g'_{\rho'_1}(\tau) = \frac{\prod_{k=0}^{K-1} \exp\{\rho'_1 I_k(\tau)\}}{\sum_{n=1}^N \prod_{k=0}^{K-1} \exp\{\rho'_1 I_k(\tau_n)\}}, \quad (2.39)$$

where N is the total number of points in the time delay interval.

3. Generate R realizations of the parameter $\{\tau_i\}_{i=1}^R$ using the inverse probability integration as detailed in Appendix D.
4. Evaluate the weight coefficient $F(\tau_i)$ defined in (2.37) (or $F'(\tau_i)$ defined in (2.38) if we consider that τ is in $[0, T)$) for each generated value τ_i .
5. Compute the mean of the generated variables multiplied by the weight coefficients to find the ML estimate of the time delay.

⁶Note that the same simplifications have been used in [9] to estimate the signal DOA.

2.7 Simulation Results

In this section, we will present numerical results to substantiate the performance of the new ML estimator as a function of the SNR. We will also refer to our new IS-based ML estimator as "IS algorithm". The normalized (by T^2) mean square error (NMSE), defined in (2.40), will be used as our performance measure :

$$\text{NMSE}(\tau) = \frac{E\{(\hat{\tau}^* - \tau^*)^2\}}{T^2}, \quad (2.40)$$

and computed over 1000 Monte-Carlo runs. The modified Cramér-Rao lower bound (MCRLB) is also normalized by T^2 and the total number of transmitted symbols, K , in the observation window is set to $K = 100$ and ρ'_1 is taken equal to 28. Unless specified otherwise, a root-raised cosine shaping pulse of roll-off factor of 0.5 is used. First, the effect of ρ_1 (or equivalently ρ'_1) on the performance of our IS-based ML estimator is shown in Fig. 2.3 at an SNR of 10 dB. As it could be predicted, the mean square error decreases as ρ_1 increases toward its optimal value and, for too large values, the performance deteriorates due to numerical overflows. In the implementation, ρ'_1 can be set as a function of the po-

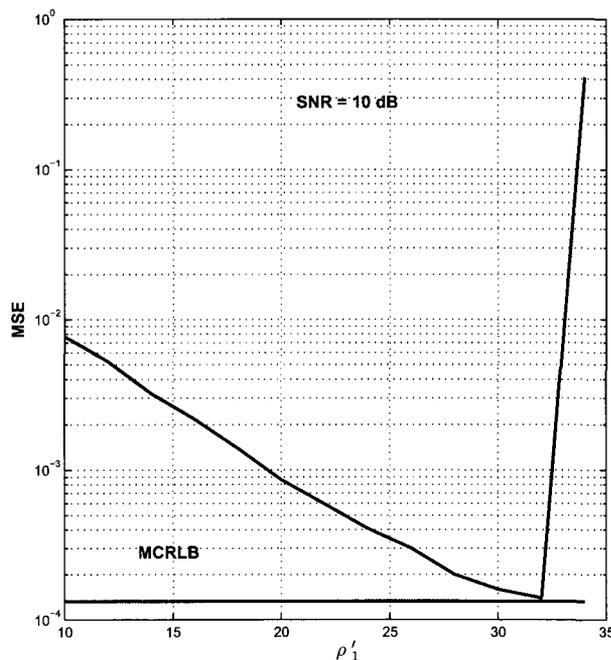


FIGURE 2.3 – Performance versus ρ'_1 for SNR=10 dB.

wer of the received samples \mathbf{y} . Moreover, using computer simulations, we verify that for a root-raised cosine filter the ratio $L'_{c,\rho_0}(\tau)/g'(\tau)$ is almost equal to 1. Then to reduce the computational complexity, we can set this ratio to 1 in (2.21) and (2.36). In fact, Fig. 2.4

shows the NMSE of the IS-based time delay ML estimator when this ratio is preserved in the importance function and when it is set to 1. As it can be seen, this simplification does not degrade the performance of the estimator while reducing the computational complexity considerably. This simplification is also valid for any linear modulation scheme. Therefore, in the following simulations we consider that :

$$\frac{L'_{c,\rho_0}(\tau)}{g'_{\rho'_1}(\tau)} = 1. \quad (2.41)$$

Note that this simplification remains valid when the intersymbol interference is not important (for high values of the roll-off factor). As the roll-off factor tends to 0, it appears necessary to consider the ratio in order to achieve better performance of the estimator. Moreover, we implement the iterative CML estimator, called CML-TED, proposed in [8]

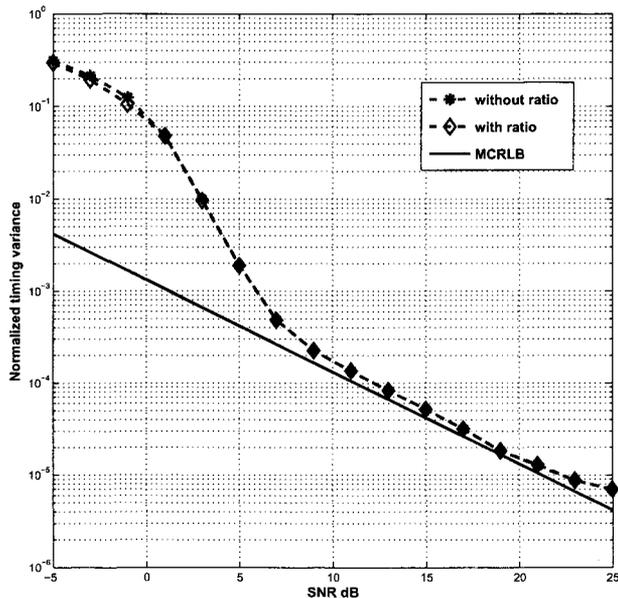


FIGURE 2.4 – Estimation performance considering $L'_{c,\rho_0}(\tau)/g'(\tau)$ and setting $L'_{c,\rho_0}(\tau)/g'(\tau)$ to 1 with QPSK modulation.

and compare its performance to the performance of our IS-based CML estimator. As far as we know, among all the existing synchronization techniques, the CML-TED algorithm achieves the best performance, but, as an iterative procedure, its performance depends strictly on the initial guess. To corroborate our claims, we consider in Fig. 5 two initial values of τ^* for the CML-TED, which should be seen as the result of another estimator. In Fig. 2.5, the small crosses represent the normalized variance where the initial value is very close to the true time delay value, i.e., verifies $|\tau_0^* - \tau^*| = T/10$, with τ^* being the true time delay value to be estimated and τ_0^* is the initial guess. As it can be seen from this figure,

even with a close-enough initial guess, our IS-based estimator outperforms the CML-TED estimator in the high SNR region although the CML-TED achieves better performance at low SNR values. We also note that, at high SNR values, the performance of our IS-based estimator is close to the MCRLB. This means that, in this region, our new time delay estimator exhibits performances equivalent to those that could be achieved if the transmitted data were perfectly known to the receiver. However, if we consider $|\tau_0^* - \tau^*| = T/2$, the performance of the CML-TED deteriorates considerably over the entire SNR region. This illustrates the fact that the CML-TED algorithm fails to estimate the time delay if the initial value is not appropriately chosen, while no initialization concerns are raised with our new IS-based CML estimator. Moreover, the second variant of the IS algorithm, namely

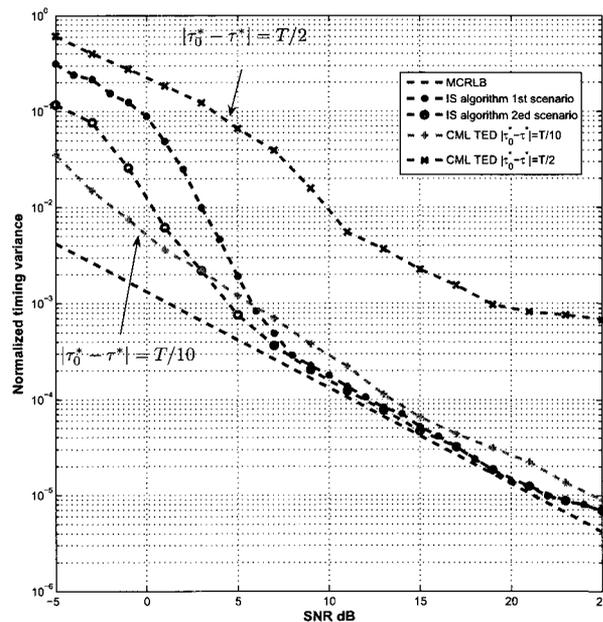


FIGURE 2.5 – Comparison between the estimation performance of the IS algorithm using the two scenarios and the tracking performance of the CML-TED using QPSK modulation.

considering the time delay as a circular variable, is also represented in Fig. 2.5. We see in this case that the variance error is reduced in the low SNR region. In addition, in both cases, starting from an SNR value of about 5 dB, our IS-based algorithm surpasses the iterative algorithm, even when assuming a sufficiently accurate initial guess.

Furthermore, in Fig. 2.6, the CML-TED algorithm exhibits a variance penalty for a roll-off equal to 0.2. This penalty is higher for smaller excess bandwidth. It has been shown in [8] that the CML-TED reaches the asymptotic compressed CRLB ($CRLB_c$), and the difference between the MCRLB and the $CRLB_c$ becomes more important as the roll-off factor decreases. Then the performance of the CML-TED cannot approach asymp-

totically the MCRLB for small roll-off factors. In contrast, the new IS-based algorithm always reaches the MCRLB in the high SNR range, irrespectively of the roll-off factor value.

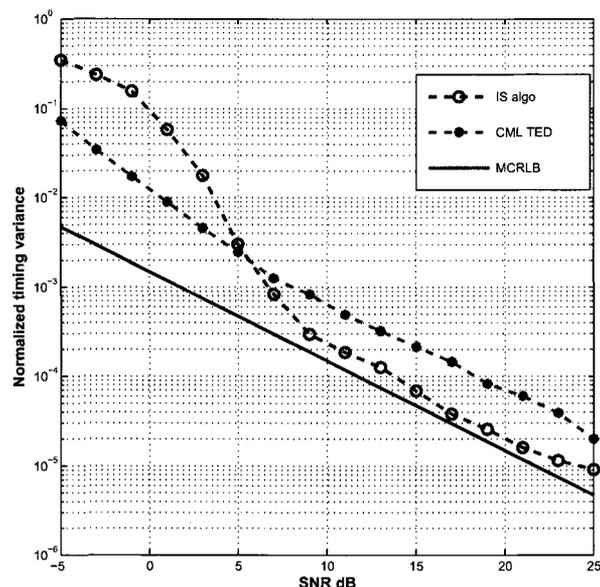


FIGURE 2.6 – Comparison between the estimation performance of IS algorithm and the tracking performance of CML TED using QPSK modulation and for a roll-off factor of 0.2.

In Fig. 2.7, performance curves are drawn for 16-QAM and 64-QAM, as examples for non-constant modulus constellations. As explained in section VI, we force the first and the last transmitted symbols to be of maximum energy. To illustrate the performance degradation in the case of higher-order modulations, we also plot the NMSE for QPSK. As we can see, the IS algorithm achieves close performance for the three modulations orders, with, however, a small improvement for the QPSK modulation. To illustrate the performance enhancement achieved by forcing the first and the last symbols to have maximum constellation magnitude, we plot, in Fig. 2.8 the performance of the new IS-based ML estimator without this constraint. As anticipated, the performance is strongly affected by the two edge symbols since the curve corresponding to the non-forced symbols does not approach the MCRLB. Therefore, for non-constant-envelope modulated constellations, it is essential to force the first and the last transmitted symbols to have maximum energy, as explained in section VI.

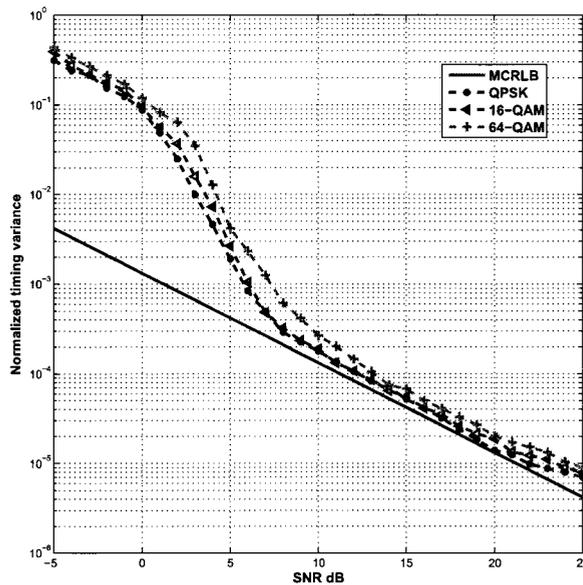


FIGURE 2.7 – Normalized MSE of the time delay estimate for different QAM modulation order, using a root-raised cosine filter with a roll-off factor 0.5.

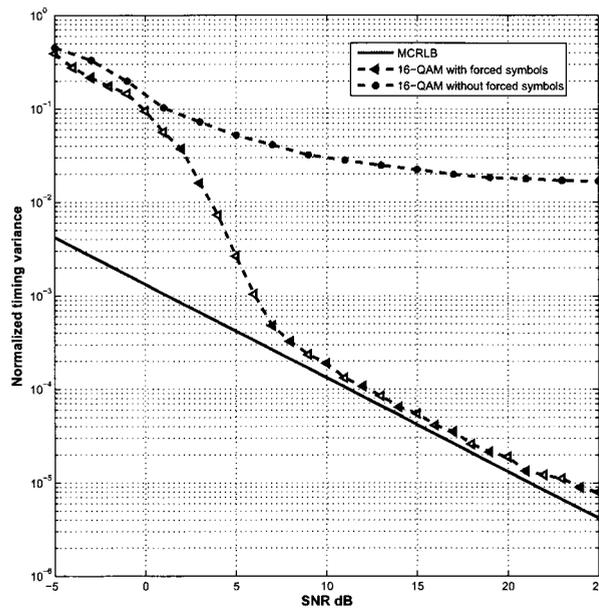


FIGURE 2.8 – Comparison of the estimation performance with and without forced symbols using 16-QAM modulation and for a roll-off factor of 0.5.

2.8 Conclusion

A computationally efficient technique has been developed to implement the CML estimator of the time delay parameter. Based on a discrete-time model, the transmitted symbols

are supposed to be unknown and no restriction on their distribution was assumed. To avoid iterative techniques and their drawbacks, the importance sampling method was used to find the ML solution. Its main advantage over the iterative procedures is that it does not require any initial guess of the time delay parameter and that it is far less computationally expensive while retaining good performances. Moreover, its convergence to the global maximum is guaranteed. Relative to other proposed methods such as the CML-TED, the IS-based estimator exhibits better performance at high SNR. In practice, the choice of the algorithm parameters ρ_0 and ρ'_1 is critical for the estimation performance and for a good choice of these parameters, a small number of generated realizations can be sufficient to achieve satisfactory performance and reduce the computation burden.

Appendix A

Proof of $\lim_{\rho_0 \rightarrow +\infty} L'_{c,\rho_0}(\tau) = \delta_\tau$

In the following, we prove that $L'_{c,\rho_0}(\tau)$ defined in (2.17) tends to a Dirac delta function centered at the location of its global maximum as $\rho_0 \rightarrow +\infty$. To do so, consider the general case where $f(x)$ is an integrable function having one global maximum, denoted a :

$$a = \arg \max_{x \in \mathbb{R}} f(x). \quad (2.42)$$

And denoting by $F(x)$ the following normalized function :

$$F(x) = \frac{\exp\{\rho_0 f(x)\}}{\int_I \exp\{\rho_0 f(u)\} du}, \quad (2.43)$$

where I is the definition domain of $F(\cdot)$. Then, for a given real number $b \neq a$, we have :

$$F(b) = \frac{\exp\{\rho_0 f(b)\}}{\int_I \exp\{\rho_0 f(x)\} dx} < \frac{\exp\{\rho_0 f(a)\}}{\int_I \exp\{\rho_0 f(x)\} dx}. \quad (2.44)$$

However, since $f(a)$ is the maximum value of the function $f(x)$, then $f(b) - f(a)$ is a negative number and, therefore, $\exp\{\rho_0(f(b) - f(a))\}$ tends to 0, as well as $F(b)$, when ρ_0 tends to ∞ . As a result :

$$\lim_{\rho_0 \rightarrow \infty} F(x) = 0, \quad (2.45)$$

for any real $x \neq a$. Moreover, if we consider that $\lim_{\rho_0 \rightarrow \infty} F(a) = 0$, then whatever $x \in \mathbb{R}$,

we have $\lim_{\rho_0 \rightarrow \infty} F(x) = 0$ and $\lim_{\rho_0 \rightarrow \infty} \int_{-\infty}^{+\infty} F(u) du = 0$, which is in conflict with the assumption that $\int_{-\infty}^{+\infty} F(u) du = 1$.

Finally, we conclude that $\lim_{\rho_0 \rightarrow \infty} F(a) \neq 0$ and since $\lim_{\rho_0 \rightarrow \infty} \int_{-\infty}^{+\infty} F(u) du = 1$, $F(x)$ becomes a Dirac delta function centered at a when ρ_0 tends to $+\infty$.

Appendix B

Justification of the approximation $\mathbf{A}_\tau^T \mathbf{A}_\tau \approx \frac{E_h}{T_s} \mathbf{I}_K$

The diagonal elements of $\mathbf{A}_\tau^T \mathbf{A}_\tau$ are the convolution of the same shifted version of $h(\cdot)$ ($[\mathbf{A}_\tau^T \mathbf{A}_\tau]_{i,i} = \mathbf{a}_i^T(\tau) \mathbf{a}_i(\tau)$ for $i = 0, 1, \dots, K-1$ where $\mathbf{a}_i(\tau)$ is defined in (6)). Whereas, when the shift is not the same (i.e., $[\mathbf{A}_\tau^T \mathbf{A}_\tau]_{i,j} = \mathbf{a}_i^T(\tau) \mathbf{a}_j(\tau)$, $i \neq j$), the value of the convolution, $\mathbf{a}_i^T(\tau) \mathbf{a}_j(\tau)$, is, according to the Nyquist criteria, equal to zero. However, since we take some samples of $h(\cdot)$, $[\mathbf{A}_\tau^T \mathbf{A}_\tau]_{i,j} = \mathbf{a}_i^T(\tau) \mathbf{a}_j(\tau)$, $i \neq j$ is not really equal to zero but still very small and negligible in front of $[\mathbf{A}_\tau^T \mathbf{A}_\tau]_{i,i}$. To clarify, we plotted in Fig. 2.9 the following three functions : $g_0(x) = h(x) \times h(x) = h(x)^2$, $g_1(x) = h(x) \times h(x-T)$ and $g_2(x) = h(x) \times h(x-2T)$ and we take for example $[\mathbf{A}_\tau^T \mathbf{A}_\tau]_{1,2} = \mathbf{a}_1^T(\tau) \mathbf{a}_2(\tau) = \sum_{m=0}^{M-1} g_1(mT_s)$. Then, notice from this figure that some samples of $g_1(\cdot)$ are negative, which will compensate for positive samples in $\sum_{m=0}^{M-1} g_1(mT_s)$ from the same function thereby resulting in very low convolution. However, the samples of $g_0(x)$ are always positive and there is no compensation effect in $[\mathbf{A}_\tau^T \mathbf{A}_\tau]_{1,1} = \sum_{m=0}^{M-1} g_0(mT_s)$ added to the fact that $g_0(x) \gg g_1(x)$, this results in the following conclusion : $[\mathbf{A}_\tau^T \mathbf{A}_\tau]_{1,1} \gg [\mathbf{A}_\tau^T \mathbf{A}_\tau]_{1,2}$. Using the same arguments for the other off-diagonal elements of $\mathbf{A}_\tau^T \mathbf{A}_\tau$ leads to the following approximation :

$$\mathbf{A}_\tau^T \mathbf{A}_\tau \approx \frac{E_h}{T_s} \mathbf{I}_K \quad (2.46)$$

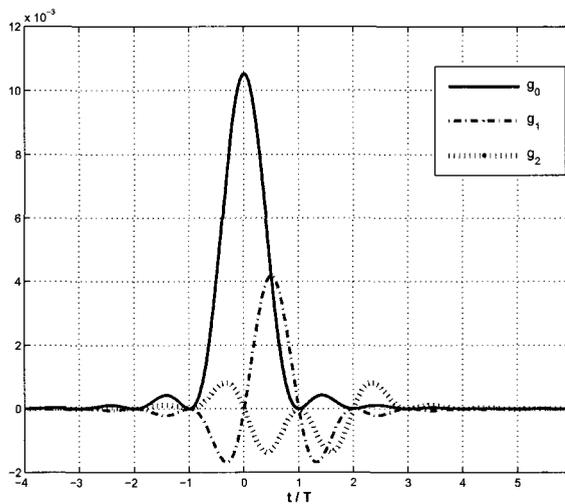


FIGURE 2.9 – Plot of $g_0(\cdot)$, $g_1(\cdot)$ and $g_2(\cdot)$.

Appendix C

Proof of the effect of ρ'_1 on $g'_{\rho'_1}(\cdot)$

In the following, we briefly show how ρ'_1 (or equivalently ρ_1) can render $g'_{\rho'_1}(\cdot)$ more peaked around its global maximum. Starting from $g'_{\rho'_1}(\tau^*)$, define the function $H_{\tau^*}(\rho'_1) = g'_{\rho'_1}(\tau^*)$:

$$H(\rho'_1) = \frac{\exp\{\rho'_1 L_c^{(a)}(\tau^*)\}}{\int_J \exp\{\rho'_1 L_c^{(a)}(v)\} dv}, \quad (2.47)$$

where τ^* is the true time delay value to be estimated and $L_c^{(a)}(\tau)$ is the approximation of $L_c(\mathbf{y}; \tau)$ defined in the right-hand side of (22), i.e., $L_c^{(a)}(\tau) = \mathbf{y}^H \mathbf{A}_\tau \mathbf{A}_\tau^T \mathbf{y}$. The first derivative of $H(\rho'_1)$ with respect to ρ'_1 is given by equation (2.48).

$$\begin{aligned} H'(\rho'_1) &= \frac{L_c^{(a)}(\tau^*) \exp\{\rho'_1 L_c^{(a)}(\tau^*)\} \int_J \exp\{\rho'_1 L_c^{(a)}(v)\} dv - \exp\{\rho'_1 L_c^{(a)}(\tau^*)\} \int_J L_c^{(a)}(v) \exp\{\rho'_1 L_c^{(a)}(v)\} dv}{\left(\int_J \exp\{\rho'_1 L_c^{(a)}(v)\} dv \right)^2} \\ &= \frac{\exp\{\rho'_1 L_c^{(a)}(\tau^*)\} \left(L_c^{(a)}(\tau^*) \int_J \exp\{\rho'_1 L_c^{(a)}(v)\} dv - \int_J L_c^{(a)}(v) \exp\{\rho'_1 L_c^{(a)}(v)\} dv \right)}{\left(\int_J \exp\{\rho'_1 L_c^{(a)}(v)\} dv \right)^2}. \end{aligned} \quad (2.48)$$

And noting that $\tau^* = \arg \max_\tau L_c^{(a)}(\tau)$, it follows that

$$\int_J L_c^{(a)}(v) \exp\{\rho'_1 L_c^{(a)}(v)\} dv < L_c^{(a)}(\tau^*) \int_J \exp\{\rho'_1 L_c^{(a)}(v)\} dv. \quad (2.49)$$

Therefore $H'_{\tau^*}(\rho'_1) > 0 \forall \rho'_1$. Hence, $H_{\tau^*}(\rho'_1)$ is an increasing function with respect to ρ'_1 , i.e., for every $\rho'_1{}^{(2)} > \rho'_1{}^{(1)}$ we have $H_{\tau^*}(\rho'_1{}^{(2)}) > H_{\tau^*}(\rho'_1{}^{(1)})$, which means $g'_{\rho'_1{}^{(2)}}(\tau^*) > g'_{\rho'_1{}^{(1)}}(\tau^*)$. We conclude that ρ'_1 renders the objective function more and more peaked around its global maximum.

Appendix D

Method to generate $[\mathcal{T}_i]_{i=1}^R$

In this appendix, we detail how to generate realizations according to $g'(\tau)$. First, we generate a vector $\mathbf{u} = [u_1, u_2, \dots, u_R]$ of R realizations uniformly distributed in $[0, 1]$.

Then, we search $\tau_i = G'^{-1}(u_i)$, where $G'^{-1}(\cdot)$ is the reciprocal function of the cumulative distribution function (CDF) $G'(x)$ of x defined as :

$$G'(x) = \int_0^x g'_{\rho_1}(v)dv. \quad (2.50)$$

Unfortunately, a closed-form expression of $G'^{-1}(x)$ is not analytically tractable. Moreover, since $G'(x)$ is a steep-slope function, a fine search to find τ_i as $\arg \min_{\tau} |u_i - G'(\tau)|$ is required and makes the process computationally intensive. However, since $G'(\tau)$ is an increasing function of τ , the function $S(\tau) = |u_i - G'(\tau)|$ is unimodal. This observation allows us to adopt the golden search [17] to find the location of the minimum of $S(\tau)$. The golden search is appropriate for this problem because it converges after a small number of iterations and requires only one function evaluation per iteration.

Bibliographie

- [1] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*, New York : Plenum, 1997.
- [2] M.I. Skolnik, *Introduction to Radar Systems*, third edition, McGraw-Hill, New York : 2001.
- [3] G.C. Carter, *Coherence and Time Delay Estimation : An Applied Tutorial for Research, Development, Test and Evaluation Engineers*, IEEE Press, New York, 1993.
- [4] J. W. M. Bergmans, *Digital Baseband Transmission and Recording*, Boston : Kluwer Academic Publishers, 1996.
- [5] M. Olsson, H. Johansson, and P. Lowenborg, "Time-Delay Estimation Using Farrow-Based Fractional-Delay FIR Filters : Filter Approximation vs. Estimation Errors" *EUSIPCO 2006*.
- [6] N. Noels *et al.*, "Turbo Synchronization : An EM Algorithm Interpretation." *IEEE Conf. Commun.*, Anchorage, AK, pp. 2933-2937, May 2003.
- [7] S. Kay, *Fundamentals of Statistical Signal Processing, Estimation theory*, Prentice Hall, Englewood Cliffs, NJ, 1993.
- [8] J. Riba, J. Sala, and G. Vazquez, "Conditional Maximum Likelihood Timing Recovery : Estimators and Bounds," *IEEE Trans. on Sig. Process.*, vol. 49, pp. 835-850, Apr. 2001.
- [9] S. Saha and S. Kay, "An Exact Maximum Likelihood Narrowband Direction of Arrival Estimator," *IEEE Trans. Sig. Process.*, vol. 56, pp. 1082-1092, Oct. 2008.
- [10] S. Saha and S. Kay, "Mean Likelihood Frequency Estimation," *IEEE Trans. Sig. Process.*, vol. 48, pp. 1937-1946, Jul. 2000.
- [11] H. Wang, and S. Kay, "A Maximum Likelihood Angle-Doppler Estimator using Importance Sampling," *IEEE Trans. Aerospace and Electro. Syst.*, vol. 46, Apr. 2010.
- [12] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Sig. Process.*, vol. 38, pp. 1783-1795, Oct. 1990.
- [13] M. Pincus, "A Closed Form Solution for Certain Programming Problems," *Oper. Research*, pp. 690-694, 1962.

- [14] L. Stewart, "Bayesian Analysis using Monte-Carlo Integration : a Powerful Methodology for Handling Some Difficult Problems," *American Statistician*, pp. 195-200, 1983.
- [15] S. Kay, "Comments on Frequency Estimation by Linear Prediction," *IEEE trans. on Acoustics, Speech and Sig. Process.*, pp. 198-199, Apr. 1979.
- [16] K. V. Mardia, *Statistics of Directional Data*, New York : Academic, 1972.
- [17] D. J. Wilde, *Globally Optimum Design*. Englewood Cliffs, NJ : Prentice-Hall, 1964.

Chapitre 3

Closed-Form Expressions for the Exact Cramér-Rao Bounds of Timing Recovery Estimators from BPSK, MSK and Square-QAM Transmissions

¹A. Masmoudi, F. Bellili, S. Affes, and A. Stephenne : “Closed-Form Expressions for the Exact Cramér-Rao Bounds of Timing Recovery Estimators from BPSK, MSK and Square-QAM Transmissions”, *IEEE Trans. on Sign. Process.*, vol. 59, no. 6, pp. 2474-2484, June 2011.

IEEE (c)

DOI: 10.1109/LSP.2008.921477

Abstract

Dans cet article, nous dérivons pour la première fois les expressions analytiques des bornes de Cramér-Rao pour l'estimation du retard pour les signaux à modulation par déplacement de phase, modulation à déplacement minimum et modulation d'amplitude en quadrature. Nous supposons que les données transmises sont inconnues au niveau du récepteur et que la fonction de mise en forme satisfait le critère de Nyquist. De plus, la phase et la fréquence porteuses sont considérées inconnues. Le retard reste constant sur l'intervalle d'observation et le signal reçu est entaché de bruit additif. Les nouvelles expressions montrent que les performances d'estimation ne dépendent pas de la vraie valeur du paramètre. De plus, ils concordent avec les résultats obtenus par calculs empiriques.

In this paper, we derive for the first time analytical expressions for the exact Cramér-Rao lower bounds (CRLB) for symbol timing recovery of binary phase shift keying (BPSK), minimum shift keying (MSK) and square QAM-modulated signals. It is assumed that the transmitted data are completely unknown at the receiver and that the shaping pulse verifies the first Nyquist criterion. Moreover the carrier phase and frequency are considered as unknown nuisance parameters. The time delay remains constant over the observation interval and the received signal is corrupted by additive white Gaussian noise (AWGN). Our new expressions prove that the achievable performance holds irrespectively of the true time delay value. Moreover, they corroborate previous attempts to empirically compute the considered bounds thereby enabling their immediate evaluation.

3.1 Introduction

In modern communication systems, the received signal is usually sampled once per-symbol interval to recover the transmitted information. But the unknown time delay, introduced by the channel, must be estimated *a priori* in order to sample the signal at the accurate sampling times. In this context, many time delay estimators have been developed to meet this requirement. These estimators can be mainly categorized into two major categories : data-aided (DA) and non-data-aided (NDA) estimators. In DA estimation, *a priori* known symbols are transmitted to assist the estimation process, although the transmission of a known sequence has the drawback of limiting the whole throughput of the system. Whereas, in the NDA mode, the required parameter is blindly estimated assuming the transmitted symbols to be completely unknown. In both cases, the performance of an estimator affects the performance of the entire system. In the case of an unbiased estimation, the variance of the timing error is usually used to evaluate the estimation accuracy. The CRLB is a lower bound on the variance of any unbiased estimator and is often used as a benchmark for the performance evaluation of actual estimators [1, 2]. The computation of this bound has been previously tackled by many authors, under different simplifying assumptions. For instance, assuming the transmitted data to be perfectly known and one can derive the DA CRLB. The modified CRLB (MCRLB), which is also easy to derive, has been introduced in [3, 4], but unfortunately it departs dramatically from the exact (stochastic) CRLB, especially at low signal-to-noise ratios (SNR).

Actually, the time delay stochastic CRLBs of higher-order modulations were empirically computed in previous works. Their analytical expressions were tackled only for specific SNR regions, i.e., very low or very high-SNR values and the derived bounds are referred to as ACRLBs (asymptotic CRLBs). In fact, in [5] the stochastic CRLB was tackled under the low-SNR assumption and an analytical expression of the considered bound (ACRLB) was derived for arbitrary PSK, QAM and PAM constellations. In this SNR region, the authors of [5] approximated the likelihood function by a truncated Taylor series expansion to obtain a relatively simple ACRLB expression. An analytical expression was also introduced in [6] under the high-SNR assumption. This high-SNR ACRLB coincides with the stochastic CRLB in this SNR region but unfortunately it cannot be used even for moderate (practical) SNR values. Another approach was later proposed in [7] and [8] to compute the NDA deterministic (or conditional) CRLBs, in which the symbols are considered as deterministic unknown parameters. Then the conditional CRLB is derived from the compressed likelihood function $f(\mathbf{y}; \boldsymbol{\theta}, \hat{\mathbf{x}})$ in which \mathbf{y} stands for the observed vector, $\boldsymbol{\theta}$ is the parameter vector of interest (including the unknown time delay) and $\hat{\mathbf{x}}$ is the maximum likelihood estimate of the transmitted symbols \mathbf{x} . However, it is widely known that the conditional CRLB does not provide the actual performance limit (unconditional or stochastic CRLBs). In an other works, the stochastic CRLB was empirically computed [9]

assuming perfect phase and frequency synchronization and a time-limited shaping pulse. Later in [10], its computation was tackled in the presence of unknown carrier phase and frequency and pulses that are unlimited in time. Both [9] and [10] simplified the expression of the bounds but ultimately resorted to empirical methods to evaluate the exact CRLB, without providing any closed-form expressions.

Motivated by these facts, in this work, we derive for the first time analytical expressions for the stochastic CRLBs of symbol timing recovery from BPSK, MSK and square QAM-modulated signals. We consider the general scenario as in [10] in which the carrier phase and frequency offsets are completely unknown at the receiver, and we show that this assumption does not actually affect the performance of a time delay estimator from perfectly frequency- and phase-synchronized received samples. The derivations assume an AWGN-corrupted received signal and a shaping pulse that verifies the first Nyquist criterion. The last assumption is verified in practice for most of the shaping pulses.

This paper is organized as follows. In section II, we introduce the system model that will be used throughout this article. In section III, we derive the analytical expression of the stochastic CRLB for any square QAM modulation. Then, in section IV, we outline the derivation steps of the CRLB in the cases of BPSK and MSK transmissions. Some graphical representations are presented in section V and, finally, some concluding remarks are drawn out in section VI.

3.2 System Model

Consider a traditional communication system where the channel delays the transmitted signal and a zero-mean *proper*² AWGN, with an overall power σ^2 , corrupts the received signal. In the case of imperfect frequency and phase synchronization, the received signal is expressed as :

$$y(t) = \sqrt{E_s} x(t - \tau) e^{j(2\pi f_c t + \theta)} + w(t), \quad (3.1)$$

where τ is the time delay, θ is the channel distortion phase, f_c is the carrier frequency offset and j is the complex number verifying $j^2 = -1$. The parameters τ , θ and f_c are assumed to be deterministic but unknown. They can be gathered in the following unknown parameter vector :

$$\boldsymbol{\nu} = [\tau, \theta, f_c]^T. \quad (3.2)$$

In (3.1), $w(t)$ is a *proper* complex Gaussian white noise with independent real and imaginary parts, each of variance $\sigma^2/2$, and $x(t)$ is the transmitted signal given by :

²A *proper* complex random process $v(t)$ satisfies $E\{v(t)^2\} = 0$.

$$x(t) = \sum_{i=1}^K a_i h(t - iT), \quad (3.3)$$

with $\{a_i\}_{i=1}^K$ being the sequence of K transmitted symbols drawn from a BPSK, an MSK or any square-QAM constellation and T is the symbol duration. The transmitted symbols are assumed to be statistically independent and equally likely, with normalized energy, i.e., $E\{|a_i|^2\} = 1$. Finally, $h(t)$ is a square-root Nyquist shaping pulse function with unit-energy which will be seen in sections III and IV, as would be expected, to have an important impact on the CRLB and therefore on the system's performance. The Nyquist pulse $g(t)$ obtained from $h(t)$ is defined as :

$$g(t) = \int_{-\infty}^{+\infty} h(x)h(t+x)dx, \quad (3.4)$$

and satisfies the first Nyquist criterion :

$$g(nT) = 0, \quad n \neq 0. \quad (3.5)$$

Suppose that we are able to produce unbiased estimates, $\hat{\nu}$, of the vector ν from the received signal. Then the CRLB, which verifies $E\{(\hat{\nu} - \nu)^2\} \geq \text{CRLB}(\nu)$, is defined as [1, 2] :

$$\text{CRLB}(\nu) = I^{-1}(\nu), \quad (3.6)$$

where $I(\nu)$ is the Fisher information matrix (FIM) whose entries are defined as :

$$[I(\nu)]_{i,j} = E \left\{ \frac{\partial L(\nu)}{\partial \nu_i} \frac{\partial L(\nu)}{\partial \nu_j} \right\}, \quad i, j = 1, 2, 3, \quad (3.7)$$

with $L(\nu)$ being the log-likelihood function of the parameters to be estimated and $\{\nu_i\}_{i=1}^3$ are the elements of the unknown parameter vector ν .

To begin with, we show in Appendix A that the problem of time delay estimation is disjoint from the problem of carrier phase and frequency estimation. Indeed, we show that the FIM is block-diagonal structured as follows :

$$I(\nu) = \begin{pmatrix} \text{CRLB}^{-1}(\tau) & \mathbf{0} \\ \mathbf{0} & I_2(\theta, f_c) \end{pmatrix}, \quad (3.8)$$

where $\mathbf{0} = [0, 0]^T$, $\text{CRLB}(\tau)$ is the CRLB of the time delay parameter and $I_2(\theta, f_c)$ is the (2×2) FIM pertaining to the joint estimation of f_c and θ . Hence, we prove analytically that we deal with two separable estimation problems ; on one hand time delay estimation and on the other hand carrier phase and frequency estimation. Actually, this conclusion has been already made in [10] but the authors resorted to empirical evaluations to find

that the elements $[\mathbf{I}(\boldsymbol{\nu})]_{1,2}$ and $[\mathbf{I}(\boldsymbol{\nu})]_{1,3}$ of the FIM are almost equal to zero. Now, since the parameters are decoupled, we only need to derive the first element of the global FIM, $[\mathbf{I}(\boldsymbol{\nu})]_{1,1}$ in order to find the CRLB for time delay estimation under imperfect frequency and phase synchronization. Therefore, in the following, we consider the virtually derotated received signal $\tilde{y}(t)$ given by :

$$\begin{aligned}\tilde{y}(t) &= y(t) e^{-j(2\pi f_c t + \theta)} \\ &= \sqrt{E_s} x(t - \tau) + \tilde{w}(t),\end{aligned}\quad (3.9)$$

where $\tilde{w}(t) = w(t)e^{-j(2\pi f_c t + \theta)}$ is also a *proper* AWGN with an overall power σ^2 since the nuisance parameters are assumed to be deterministic.

We mention that $|\cdot|$, $\Re\{\cdot\}$, $\Im\{\cdot\}$ and $\{\cdot\}^*$ return the magnitude, real, imaginary and conjugate of any complex number and $E\{\cdot\}$ is the statistical expectation. We also define the SNR of the system as $\rho = E_s/\sigma^2$.

3.3 Time Delay CRLB for Square QAM-Modulated Signals

In this section, we introduce the main contribution embodied by this paper which consists in deriving closed-form expressions for the stochastic CRLBs of time delay estimation when the transmitted data are unknown and drawn from any M -ary square QAM-constellation (i.e., $M = 2^{2p}$).

Before further development, it is important to emphasize that an exact representation of $\tilde{y}(t)$ requires an infinite-dimensional vector representation $\tilde{\mathbf{y}}$. But let us consider the N -dimensional truncated vectors $\tilde{\mathbf{y}}_N$, \mathbf{x}_N and $\tilde{\mathbf{w}}_N$, representing the projection, over an orthonormal basis of N dimensions, of $\tilde{y}(t)$, $x(t)$ and $\tilde{w}(t)$, respectively. Then, the pdf of $\tilde{\mathbf{y}}_N$ conditioned on the transmitted symbols \mathbf{a} and parameterized by τ is given by [4] :

$$P(\tilde{\mathbf{y}}_N|\mathbf{a}; \tau) = \prod_{i=1}^N \frac{1}{\pi\sigma^2} \exp\left\{-\frac{|\tilde{y}_k - x_k|^2}{\sigma^2}\right\}. \quad (3.10)$$

To derive the likelihood function which incorporates all the information contained in $\tilde{y}(t)$, we should make N tend to infinity to get $P(\tilde{\mathbf{y}}|\mathbf{a}; \tau)$. However, convergence problems appear. To overcome these problems, $P(\tilde{\mathbf{y}}_N|\mathbf{a}; \tau)$ is divided by $1/(\pi\sigma^2)^N \exp\left\{1/\sigma^2 \sum_{k=1}^N |\tilde{y}_k|^2\right\}$ to obtain :

$$\Lambda(\tilde{\mathbf{y}}|\mathbf{a}; \tau) = \exp\left\{\frac{2\sqrt{E_s}}{\sigma^2} \sum_{k=1}^N \Re\{\tilde{y}_k x_k^*\} - \frac{E_s}{\sigma^2} \sum_{k=1}^N |x_k|^2\right\}, \quad (3.11)$$

and as N tends to infinity, we obtain the conditional likelihood function :

$$\Lambda(\tilde{\mathbf{y}}|\mathbf{a}; \tau) = \exp\left\{\frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)x(t)^* dt\} - \frac{E_s}{\sigma^2} \int_{-\infty}^{+\infty} |x(t)|^2 dt\right\}. \quad (3.12)$$

To begin with, we note that since the transmitted symbols $\{a_i\}_{i=1}^K$ are equally likely, then the desired likelihood function of the derotated observation vector $\tilde{\mathbf{y}}$ can be written as :

$$\Lambda(\tilde{\mathbf{y}}; \tau) = E \left\{ \prod_{i=1}^K F(a_i, \tilde{y}(t)) \right\}, \quad (3.13)$$

where the expectation is performed with respect to the vector of transmitted symbols and

$$F(a_i, \tilde{y}(t)) = \exp \left\{ \frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)a_i^*\} h(t - iT - \tau) dt - \frac{E_s}{\sigma^2} |a_i|^2 \right\}. \quad (3.14)$$

It can be shown that (3.13) reduces simply to :

$$\Lambda(\tilde{\mathbf{y}}; \tau) = \frac{1}{M^K} \prod_{i=1}^K H_i(\tau), \quad (3.15)$$

where

$$H_i(\tau) = \sum_{c_k \in C} \exp \left\{ -\frac{E_s}{\sigma^2} |c_k|^2 + \frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)c_k^*\} h(t - iT - \tau) dt \right\}, \quad (3.16)$$

in which C is the constellation alphabet. Actually, the main difficulty in deriving an analytical expression for the stochastic CRLB stems from the complexity of the log-likelihood function. Therefore, we will manipulate the summation involved in (3.16). In fact, considering only square QAM-modulated signals, we are able, by exploiting the full symmetry of the constellation, to factorize $H_i(\tau)$ which in turn linearizes the global log-likelihood function and ultimately linearizes all the derivations.

Indeed, denoting by \tilde{C} the subset of the alphabet points with positive real and imaginary parts (i.e., $\tilde{C} = \{(2i-1)d_p + j(2k-1)d_p\}_{i,k=1,2,\dots,2^{p-1}}$), the constellation alphabet C is decomposed as follows :

$$C = \tilde{C} \cup \tilde{C}^* \cup (-\tilde{C}) \cup (-\tilde{C}^*). \quad (3.17)$$

Note that d_p is the inter-symbol distance derived under the assumption of a normalized-energy square QAM constellation as follows :

$$d_p = \frac{2^{p-1}}{\sqrt{2^p \sum_{k=1}^{2^{p-1}} (2k-1)^2}}. \quad (3.18)$$

Using (3.17), we rewrite (3.16) as :

$$\begin{aligned}
H_i(\tau) = \sum_{\tilde{c}_k \in \tilde{C}} \exp \left\{ -\frac{E_s}{\sigma^2} |\tilde{c}_k|^2 \right\} \times & \left(\exp \left\{ \frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)(-\tilde{c}_k^*)\} h(t - iT - \tau) dt \right\} \right. \\
& + \exp \left\{ \frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)(-\tilde{c}_k)\} h(t - iT - \tau) dt \right\} \\
& + \exp \left\{ \frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)\tilde{c}_k^*\} h(t - iT - \tau) dt \right\} \\
& \left. + \exp \left\{ \frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)\tilde{c}_k\} h(t - iT - \tau) dt \right\} \right). \quad (3.19)
\end{aligned}$$

Now using the hyperbolic cosine function defined by $2 \cosh(x) = e^x + e^{-x}$, (3.19) reduces simply to :

$$\begin{aligned}
H_i(\tau) = 2 \sum_{\tilde{c}_k \in \tilde{C}} \exp \left\{ -\frac{E_s}{\sigma^2} |\tilde{c}_k|^2 \right\} \left[\cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)\tilde{c}_k^*\} h(t - iT - \tau) dt \right) \right. \\
\left. + \cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)\tilde{c}_k\} h(t - iT - \tau) dt \right) \right]. \quad (3.20)
\end{aligned}$$

Moreover, using the fact that $\cosh(a) + \cosh(b) = 2 \cosh(\frac{a+b}{2}) \cosh(\frac{a-b}{2})$ and noting that $\tilde{c}_k + \tilde{c}_k^* = 2\Re\{\tilde{c}_k\}$ and $\tilde{c}_k - \tilde{c}_k^* = 2j\Im\{\tilde{c}_k\}$, we obtain :

$$\begin{aligned}
H_i(\tau) = 2 \sum_{\tilde{c}_k \in \tilde{C}} \exp \left\{ -\frac{E_s}{\sigma^2} |\tilde{c}_k|^2 \right\} \cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} \Re\{\tilde{c}_k\} \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)\} h(t - iT - \tau) dt \right) \times \\
\cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} \Im\{\tilde{c}_k\} \int_{-\infty}^{+\infty} \Im\{\tilde{y}(t)\} h(t - iT - \tau) dt \right). \quad (3.21)
\end{aligned}$$

Recall that $\tilde{C} = \{(2l-1)d_p + j(2m-1)d_p\}_{l,m=1,2,\dots,2^{p-1}}$ and hence the previous expression of $H_i(\tau)$ is rewritten as :

$$\begin{aligned}
H_i(\tau) = 4 \sum_{l=1}^{2^{p-1}} \sum_{m=1}^{2^{p-1}} \exp \left\{ -\frac{E_s ((2l-1)^2 + (2m-1)^2) d_p^2}{\sigma^2} \right\} \times \\
\cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} (2l-1) d_p \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)\} h(t - iT - \tau) dt \right) \times \\
\cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} (2m-1) d_p \int_{-\infty}^{+\infty} \Im\{\tilde{y}(t)\} h(t - iT - \tau) dt \right). \quad (3.22)
\end{aligned}$$

Then, splitting the two sums in (3.22), $H_i(\tau)$ is factorized as follows³ :

³Note that similar factorization was recently used to derive an analytical expression for the NDA SNR estimation [11], [12].

$$H_i(\tau) = 4 F(U_i(\tau)) F(V_i(\tau)), \quad (3.23)$$

where

$$U_i(\tau) = \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)\} h(t - iT - \tau) dt, \quad (3.24)$$

$$V_i(\tau) = \int_{-\infty}^{+\infty} \Im\{\tilde{y}(t)\} h(t - iT - \tau) dt, \quad (3.25)$$

and

$$F(x) = \sum_{k=1}^{2^{p-1}} \exp\left\{-\frac{E_s}{\sigma^2}(2k-1)^2 d_p^2\right\} \cosh\left(\frac{2\sqrt{E_s}}{\sigma^2}(2k-1)d_p x\right). \quad (3.26)$$

Now, injecting the expression of $H_i(\tau)$ in the likelihood function of the received signal (3.15), we obtain :

$$\Lambda(\tilde{\mathbf{y}}; \tau) = \left(\frac{4}{M}\right)^K \prod_{i=1}^K F(U_i(\tau)) F(V_i(\tau)). \quad (3.27)$$

Finally, the log-likelihood function of the received signal expands to :

$$L(\tau) = \sum_{i=1}^K \ln(F(U_i(\tau))) + \sum_{i=1}^K \ln(F(V_i(\tau))). \quad (3.28)$$

Note from (3.28) that due to the factorization of $H_i(\tau)$ in (3.19), the global log-likelihood function of interest in (3.28) involves the sum of two analogous terms. This reduces considerably the complexity of the stochastic CRLB derivation.

In fact, the first derivative of (3.28) with respect to τ is obtained as follows :

$$\frac{\partial L(\tau)}{\partial \tau} = \sum_{i=1}^K \frac{\dot{F}(U_i(\tau))}{F(U_i(\tau))} \frac{\partial U_i(\tau)}{\partial \tau} + \sum_{i=1}^K \frac{\dot{F}(V_i(\tau))}{F(V_i(\tau))} \frac{\partial V_i(\tau)}{\partial \tau}, \quad (3.29)$$

where $\dot{F}(x) = \frac{\partial F(x)}{\partial x}$ is given by :

$$\dot{F}(x) = \sum_{k=1}^{2^{p-1}} \exp\{-\rho(2k-1)^2 d_p^2\} \frac{2\sqrt{\rho}}{\sigma} (2k-1) d_p \sinh\left(\frac{2\sqrt{\rho}}{\sigma} (2k-1) d_p x\right). \quad (3.30)$$

Then the first diagonal element of the FIM matrix is expressed as :

$$\begin{aligned} [\mathbf{I}(\boldsymbol{\nu})]_{1;1} &= \mathbb{E} \left\{ \left(\frac{\partial L(\tau)}{\partial \tau} \right)^2 \right\} = \mathbb{E} \left\{ \sum_{i=1}^K \sum_{l=1}^K \frac{\dot{F}(U_i(\tau)) \dot{F}(U_l(\tau))}{F(U_i(\tau)) F(U_l(\tau))} \dot{U}_i(\tau) \dot{U}_l(\tau) \right\} + \\ &2 \mathbb{E} \left\{ \sum_{i=1}^K \sum_{l=1}^K \frac{\dot{F}(U_i(\tau)) \dot{F}(V_l(\tau))}{F(U_i(\tau)) F(V_l(\tau))} \dot{U}_i(\tau) \dot{V}_l(\tau) \right\} + \mathbb{E} \left\{ \sum_{i=1}^K \sum_{l=1}^K \frac{\dot{F}(V_i(\tau)) \dot{F}(V_l(\tau))}{F(V_i(\tau)) F(V_l(\tau))} \dot{V}_i(\tau) \dot{V}_l(\tau) \right\}, \end{aligned} \quad (3.31)$$

where $\dot{U}_l(\tau)$ and $\dot{V}_l(\tau)$ are the derivatives of $U_l(\tau)$ and $V_l(\tau)$ with respect to τ .

Starting from (3.31), the derivation of $[\mathbf{I}(\boldsymbol{\nu})]_{1;1}$ involves the evaluation of three expectations. However, it is easy to verify that the first and the last expectations in the right-hand side of (3.31) are performed with respect to two random processes having the same statistical properties and they are therefore identically equal. Moreover, as shown in Appendix B, the second expectation is equal to zero. Therefore, (3.31) reduces simply to :

$$[\mathbf{I}(\boldsymbol{\nu})]_{1;1} = 2 \sum_{i=1}^K \sum_{l=1}^K E \left\{ \frac{\dot{F}(U_i(\tau)) \dot{F}(U_l(\tau))}{F(U_i(\tau)) F(U_l(\tau))} \dot{U}_i(\tau) \dot{U}_l(\tau) \right\}, \quad (3.32)$$

First, we consider the case where $i = l$, and we show in Appendix C that $U_i(\tau)$ and $\dot{U}_i(\tau)$ are statistically independent. This results in :

$$E \left\{ \left(\frac{\dot{F}(U_i(\tau))}{F(U_i(\tau))} \right)^2 (\dot{U}_i(\tau))^2 \right\} = E \left\{ \left(\frac{\dot{F}(U_i(\tau))}{F(U_i(\tau))} \right)^2 \right\} E \left\{ (\dot{U}_i(\tau))^2 \right\}. \quad (3.33)$$

These two expectations involved in the right-hand side of (3.33) are easily evaluated as follows :

$$\begin{aligned} E \left\{ \left(\frac{\dot{F}(U_i(\tau))}{F(U_i(\tau))} \right)^2 \right\} &= \int_{-\infty}^{+\infty} \left(\frac{\dot{F}(U_i(\tau))}{F(U_i(\tau))} \right)^2 P(U_i(\tau)) dU_i(\tau) \\ &= \frac{1}{\sqrt{\pi\sigma^2}} \sqrt{\frac{4}{M}} \int_{-\infty}^{+\infty} \frac{\dot{F}^2(U_i(\tau))}{F(U_i(\tau))} \exp \left\{ -\frac{U_i^2(\tau)}{\sigma^2} \right\} dU_i(\tau) \end{aligned} \quad (3.34)$$

$$E \left\{ (\dot{U}_i(\tau))^2 \right\} = \frac{E_s}{2} \sum_{k=1}^K \dot{g}^2((i-k)T) - \frac{\sigma^2}{2} \ddot{g}(0). \quad (3.35)$$

where $\dot{g}(\cdot)$ and $\ddot{g}(\cdot)$ are the first and second derivative of $g(\cdot)$, respectively. We simplify (3.34) by changing $\sqrt{2}U_i(\tau)/\sigma$ by x and we obtain the following result :

$$E \left\{ \left(\frac{\dot{F}(U_i(\tau))}{F(U_i(\tau))} \right)^2 \right\} = \sqrt{\frac{2}{\pi M}} \frac{\rho}{\sigma^2} \int_{-\infty}^{+\infty} \frac{g_\rho^2(x)}{G_\rho(x)} e^{-\frac{x^2}{2}} dx, \quad (3.36)$$

where

$$g_\rho(x) = \sum_{k=1}^{2^{p-1}} \exp \left\{ -\rho(2k-1)^2 d_p^2 \right\} \sqrt{2}(2k-1) d_p \sinh \left(\sqrt{2\rho}(2k-1) d_p x \right), \quad (3.37)$$

$$G_\rho(x) = \sum_{k=1}^{2^{p-1}} \exp \left\{ -\rho(2k-1)^2 d_p^2 \right\} \cosh \left(\sqrt{2\rho}(2k-1) d_p x \right). \quad (3.38)$$

We now consider the case where $i \neq l$. The intersymbol interference results in a statistical dependence between $U_i(\tau)$ and the first derivatives $\dot{U}_i(\tau)$ and $\dot{U}_l(\tau)$ (likewise for $V_l(\tau)$ and the first derivatives $\dot{V}_i(\tau)$ and $\dot{V}_l(\tau)$). Thus using a standard probability approach to derive the expectations involved in (3.32), we first average by conditioning on $U_i(\tau)$ and $U_l(\tau)$, then average the resulting expression with respect to these two random variables. To that end, consider the expectation of $\dot{U}_i(\tau)$ and $\dot{U}_l(\tau)$ conditioned on $U_i(\tau)$ and $U_l(\tau)$:

$$E\{\dot{U}_i(\tau)|U_i(\tau), U_l(\tau)\} = U_l(\tau)\dot{g}((i-l)T), \quad (3.39)$$

$$E\{\dot{U}_l(\tau)|U_i(\tau), U_l(\tau)\} = U_i(\tau)\dot{g}((l-i)T). \quad (3.40)$$

Using (3.39) and (3.40), it follows that :

$$E\left\{\frac{\dot{F}(U_i(\tau))\dot{F}(U_l(\tau))}{F(U_i(\tau))F(U_l(\tau))}\dot{U}_i(\tau)\dot{U}_l(\tau)\middle|U_i(\tau), U_l(\tau)\right\} = \frac{\dot{F}(U_i(\tau))\dot{F}(U_l(\tau))}{F(U_i(\tau))F(U_l(\tau))} \times U_i(\tau)U_l(\tau)\dot{g}((i-l)T)\dot{g}((l-i)T)$$

and we obtain :

$$E\left\{\frac{\dot{F}(U_i(\tau))\dot{F}(U_l(\tau))}{F(U_i(\tau))F(U_l(\tau))}\dot{U}_i(\tau)\dot{U}_l(\tau)\right\} = -\left(E\left\{\frac{\dot{F}(U_i(\tau))}{F(U_i(\tau))}U_i(\tau)\right\}\right)^2 \dot{g}^2((i-l)T) \quad (3.42)$$

where the last equality follows from the statistical independence of $U_i(\tau)$ and $U_l(\tau)$ and :

$$E\left\{\frac{\dot{F}(U_i(\tau))}{F(U_i(\tau))}U_i(\tau)\right\} = \sqrt{\frac{2\rho}{\pi M}} \int_{-\infty}^{+\infty} xg_\rho(x)e^{-\frac{x^2}{2}} dx. \quad (3.43)$$

Finally, gathering all these results, we obtain the analytical expression of the stochastic CRLB for symbol timing estimation. From square QAM-modulated signals in the presence of carrier phase and frequency offsets as follows :

$$\begin{aligned} \text{CRLB}(\tau) = & \left[\left(2\rho^2 \sum_{m=1}^K \sum_{n=1}^K \dot{g}^2((m-n)T) - 2K\rho\ddot{g}(0) \right) \sqrt{\frac{2}{\pi M}} \int_{-\infty}^{+\infty} \frac{g_\rho^2(x)}{G_\rho(x)} e^{-\frac{x^2}{2}} dx - \right. \\ & \left. \frac{2\rho}{\pi M} \left(\int_{-\infty}^{+\infty} xg_\rho(x)e^{-\frac{x^2}{2}} dx \right)^2 \sum_{m=1}^K \sum_{n=1}^K \dot{g}^2((m-n)T) \right]^{-1}. \end{aligned} \quad (3.44)$$

Note that for large values of K , one can use the following accurate approximation [10] :

$$\sum_{m=1}^K \sum_{n=1}^K \dot{g}^2((m-n)T) \approx K \sum_{m=-\infty}^{+\infty} \dot{g}^2(mT). \quad (3.45)$$

It is worth mentioning that the new analytical expression in (3.44) allows the immediate evaluation of time delay stochastic CRLBs, contrarily to the empirical approaches presented in [9] and [10], and this is made possible for any square QAM modulation order.

Second, the shaping pulse is involved only via $\ddot{g}(0)$ and $\dot{g}^2((m-n)T)$, and is separate from the factors resulting from the modulation order. Moreover, to the best of our knowledge, we show here for the first time, through our new analytical expression, that the true value of the time delay parameter does not affect the actual achievable performance as intuitively expected, i.e., the variance of the estimation error holds irrespectively of the time delay value to be estimated.

3.4 CRLB for BPSK and MSK Modulated Signals

In this section, we consider the BPSK and MSK modulations. In BPSK transmissions, the data symbols take values in $\{-1, +1\}$ with equal probabilities. In MSK transmissions, the symbols are defined as $a_{k+1} = j a_k c_k$ where c_k is a sequence of BPSK symbols and a_0 is the original value drawn from the set $\{-1, -j, +1, +j\}$. For these two transmission schemes, the key derivation steps of the NDA CRLB will be briefly outlined in the following. All derivation details can be found in Appendix D.

First, the likelihood function of interest based on the received signal is :

$$\Lambda(\tilde{\mathbf{y}}; \tau) = \prod_{i=1}^K \cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{b_i^* \tilde{y}(t)\} h(t - iT - \tau) dt + \frac{E_s}{\sigma^2} |b_i|^2 \right), \quad (3.46)$$

where b_i is equal to 1 and $j^{i-1}a_0$ for BPSK and MSK, respectively. Therefore, we show that the useful log-likelihood function of $\tilde{\mathbf{y}}$ is given by :

$$L(\tau) = \sum_{i=1}^K \ln \left(\cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} \int_{-\infty}^{+\infty} \Re\{b_i^* \tilde{y}(t)\} h(t - iT - \tau) dt \right) \right). \quad (3.47)$$

Note that $\tilde{y}(t)$ is defined in (3.9). After some algebraic manipulations, detailed in Appendix D, it turns out that the analytical expression of the stochastic CRLB for time delay estimation is the same for BPSK and MSK modulations, and it is given by :

$$\begin{aligned} \text{CRLB} &= \left[4\rho \left\{ \left(1 - \frac{1}{\sqrt{2\pi}} e^{-\rho} \beta(\rho) \right) \left(\rho \sum_{m=1}^K \sum_{n=1}^K \dot{g}((m-n)T) - \frac{K}{2} \ddot{g}(0) \right) \right. \right. \\ &\quad \left. \left. - \rho \sum_{m=1}^K \sum_{n=1}^K \dot{g}((m-n)T) \right\} \right]^{-1}, \end{aligned} \quad (3.48)$$

where $\beta(\cdot)$ is defined as :

$$\beta(\rho) = \int_{-\infty}^{+\infty} \frac{e^{-\frac{x^2}{2}}}{\cosh(\sqrt{2\rho} x)} dx. \quad (3.49)$$

3.5 Graphical Representations

In this section, we provide graphical representations of the time delay CRLBs and the CRLB/MCRLB ratio for different modulation orders. First, we mention that the even integrand functions $\frac{g_\rho^2(x)}{G_\rho(x)}e^{-\frac{x^2}{2}}$, $xg_\rho(x)e^{-\frac{x^2}{2}}$ and $\frac{e^{-\frac{x^2}{2}}}{\cosh(\sqrt{2\rho}x)}$ involved in (3.44) and (3.49), respectively, decrease rapidly as $|t|$ increases. Therefore, the integrals over $[-\infty, +\infty]$ can be accurately approximated by a finite integral over an interval $[-A, A]$ and the Riemann integration method can be adequately used. In our simulations, we note that $A = 100$ and a summation step of 0.5 provided accurate values for the infinite integral.

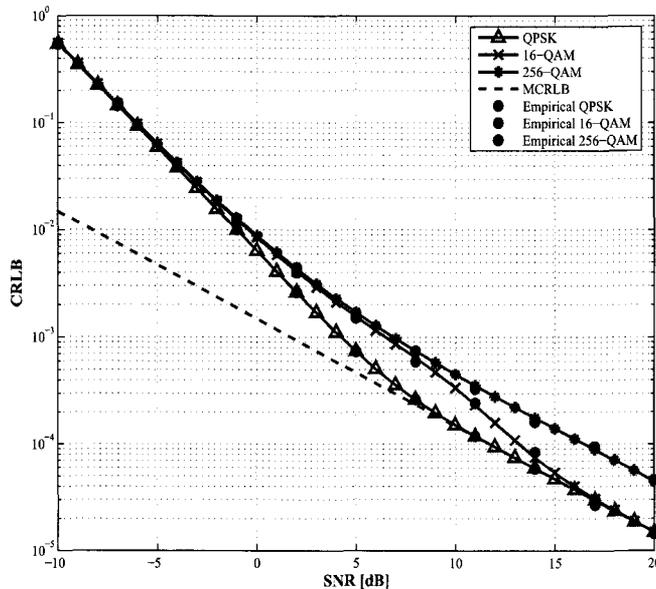


FIGURE 3.1 – Comparison between the empirical CRLB and the analytical expression in (3.44) for different modulation orders using $K = 100$ and a raised-cosine pulse with rolloff factor of 0.2.

First, we plot in Fig. 3.1 the CRLBs for different modulation orders and compare them to the ones previously obtained empirically in [10]. We see a good agreement between the two approaches thereby validating the developments above. Then, we confirm through Fig. 3.2 that, at low SNR values, the MCRLB is a looser bound compared to the exact CRLB. Indeed, this figure depicts the CRLB/MCRLB ratio as a function of the SNR. This ratio quantifies the performance degradation that arises from randomizing the transmitted data and it approaches 1 at high SNR values. Hence, in this SNR region, the MCRLB can be used as a benchmark to evaluate the performance of unbiased time delay estimators instead of the exact CRLB, since it is easier to evaluate. However, the gap between the two bounds becomes important as soon as the SNR drops below 7 dB, even for QPSK-modulated signals, where the stochastic CRLB quantifies the actual performance limit.

Moreover, we consider in this figure two values of the rolloff factor, 0.2 and 1, in order to illustrate the effect of the rolloff factor on timing estimation. Clearly, timing estimation is less accurate at a lower rolloff factor (larger intersymbol interference).

Moreover, we see from Fig. 3.3 that the different CRLBs tend to ultimately coincide with the MCRLB as long as the SNR gets increases. Actually, in the high SNR region, the achievable performance of NDA estimation of the signal time delay is equivalent to the one obtained when the received symbols are perfectly known since in this SNR range the MCRLB coincides with the DA CRLB.

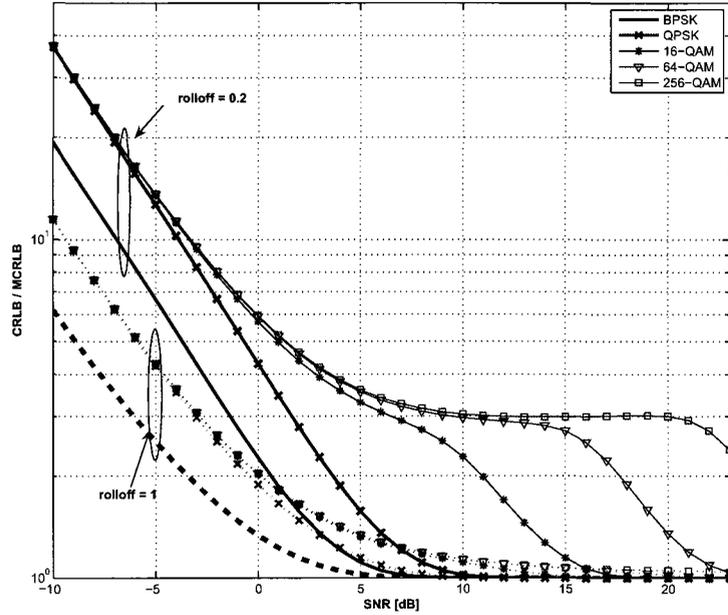


FIGURE 3.2 – CRLB/MCRLB ratio vs. SNR for different modulation orders using $K = 100$ and a raised-cosine pulse with rolloff factor of 0.2 and 1.

In the specific case where $h(t)$ is time limited to the symbol duration, the corresponding CRLB follows directly from the general expression in (3.44) by taking $\dot{g}(mT) = 0$ for all $m \in \mathbb{Z}$:

$$\text{CRLB}(\tau) = \left[-2K\rho\ddot{g}(0) \sqrt{\frac{2}{\pi M}} \int_{-\infty}^{+\infty} \frac{g_{\rho}^2(x)}{G_{\rho}(x)} e^{-\frac{x^2}{2}} dx \right]^{-1}. \quad (3.50)$$

Note from (3.50) that the resulting CRLB becomes the product of two separate terms ; one depending on the shaping pulse function and the other on the signal modulation. This special bound is plotted in Fig. 3.4. We see again a good agreement in this special case between the CRLBs obtained from our analytical expression in (3.50) and their empirical counterparts plotted in Fig. 1 of [9]. This particular expression still finds applications in

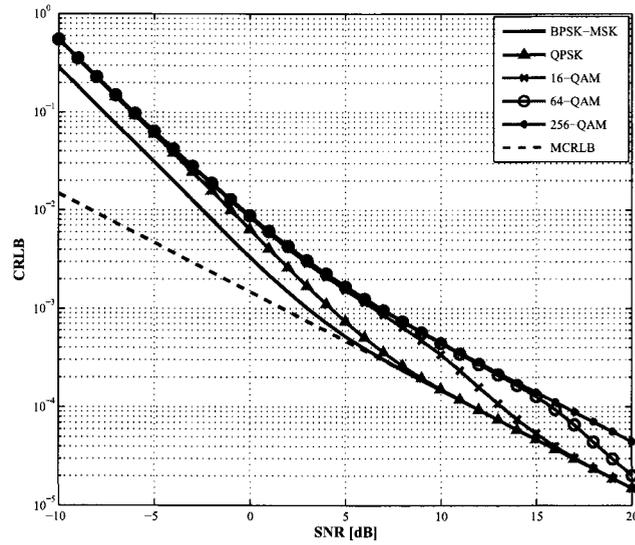


FIGURE 3.3 – CRLB vs. SNR for different modulation orders using $K = 100$ and a raised-cosine pulse with rolloff factor of 0.2.

many conventional systems and in the emerging impulse radio technology [13, 14] where, precisely, synchronization stands today as a very challenging issue.

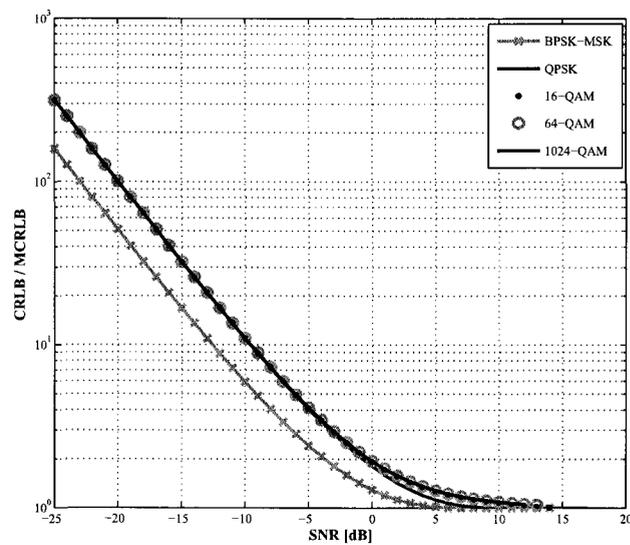


FIGURE 3.4 – CRLB/MCRLB ratio vs. SNR for different modulations and a time-limited shaping pulse.

3.6 Conclusion

In this paper, we derived, for the first time, analytical expressions of the Cramér-Rao lower bound for symbol timing estimation in the cases of BPSK, MSK and square-QAM modulations. We considered the stochastic CRLB where the transmitted data are unknown and randomly drawn. The carrier phase and frequency offsets are also supposed to be unknown (nuisance parameters). We showed that the knowledge of the phase and frequency does not bring any additional information to the time delay estimation problem and that the latter is decoupled from the joint estimation of the carrier frequency and phase offsets. Moreover, our analytical expressions for the CRLBs underline the fact that these bounds do not depend on the time delay value, which used to be stated only intuitively. We confirmed also that the modified CRLB is a valid approximation of the exact CRLB in the high SNR region and that it can be used as a benchmark since it is easier to evaluate. Furthermore, the derived analytical expressions corroborate previous works that empirically computed the stochastic CRLBs via Monte Carlo simulations, and hence provide a useful tool for a quick and easy evaluation of the CRLBs with BPSK, MSK and square-QAM modulations.

Appendix A

Proof of the Block-Diagonal Structure of the FIM

To show that τ and $\mathbf{u} = [f_c, \theta]^T$ are decoupled, we consider the actual received signal $y(t)$ instead of the virtually derotated signal $\tilde{y}(t)$. Then we follow the same derivation steps from (3.13) to (3.28) to retrieve the log-likelihood function parameterized by $\boldsymbol{\nu}$ as follows :

$$L(\boldsymbol{\nu}) = \sum_{i=1}^K \ln (F(U_i(\boldsymbol{\nu}))) + \sum_{i=1}^K \ln (F(V_i(\boldsymbol{\nu}))). \quad (3.51)$$

The first derivatives of this function with respect to the l^{th} element of \mathbf{u} , $\{u_l\}_{l=1}^{l=2}$, and τ are, respectively, given by :

$$\frac{\partial L(\boldsymbol{\nu})}{\partial u_l} = \sum_{i=1}^K \frac{\dot{F}(U_i(\boldsymbol{\nu}))}{F(U_i(\boldsymbol{\nu}))} \frac{\partial U_i(\boldsymbol{\nu})}{\partial u_l} + \frac{\dot{F}(V_i(\boldsymbol{\nu}))}{F(V_i(\boldsymbol{\nu}))} \frac{\partial V_i(\boldsymbol{\nu})}{\partial u_l}, \quad l = 1, 2, \quad (3.52)$$

and

$$\frac{\partial L(\boldsymbol{\nu})}{\partial \tau} = \sum_{i=1}^K \frac{\dot{F}(U_i(\boldsymbol{\nu}))}{F(U_i(\boldsymbol{\nu}))} \frac{\partial U_i(\boldsymbol{\nu})}{\partial \tau} + \frac{\dot{F}(V_i(\boldsymbol{\nu}))}{F(V_i(\boldsymbol{\nu}))} \frac{\partial V_i(\boldsymbol{\nu})}{\partial \tau}, \quad (3.53)$$

where $\dot{F}(\cdot)$ is defined in (3.30). Then we average $\frac{\partial L(\boldsymbol{\nu})}{\partial \tau} \frac{\partial L(\boldsymbol{\nu})}{\partial \mathbf{u}(l)}$ as in (3.31) to obtain the following result :

$$\begin{aligned}
[\mathbf{I}(\boldsymbol{\nu})]_{1;l+1} &= E \left\{ \frac{\partial L(\boldsymbol{\nu})}{\partial \tau} \frac{\partial L(\boldsymbol{\nu})}{\partial u_l} \right\} \\
&= 2 \sum_{i=1}^K \sum_{m=1}^K E \left\{ \frac{\dot{F}(U_i(\boldsymbol{\nu})) \dot{F}(U_m(\boldsymbol{\nu}))}{F(U_i(\boldsymbol{\nu})) F(U_m(\boldsymbol{\nu}))} \frac{\partial U_i(\boldsymbol{\nu})}{\partial \tau} \frac{\partial U_m(\boldsymbol{\nu})}{\partial u_l} \right\}. \quad (3.54)
\end{aligned}$$

In order to simplify the calculations, without loss of generality, we consider $l = 1$. To begin with, we first differentiate $U_m(\boldsymbol{\nu})$ with respect to f_c and we obtain :

$$\begin{aligned}
\frac{\partial U_m(\boldsymbol{\nu})}{\partial f_c} &= 2\pi \int_{-\infty}^{+\infty} \Im \{ y(t) e^{-j(2\pi f_c t + \theta)} \} h(t - mT - \tau) t dt \\
&= 2\pi \sum_{n=1}^K \Im \{ a_m \} \int_{-\infty}^{+\infty} h(t - nT - \tau) h(t - mT - \tau) t dt + \int_{-\infty}^{+\infty} \Im \{ \tilde{w}(t) \} h(t - mT - \tau) t dt.
\end{aligned} \quad (3.55)$$

$\frac{\partial U_m(\boldsymbol{\nu})}{\partial f_c}$ is a function of the imaginary part of the transmitted symbols and the derotated noise, which are mutually independent from the real part of the transmitted symbols and the derotated noise. As a result, $\frac{\partial U_m(\boldsymbol{\nu})}{\partial f_c}$ is independent from $U_i(\boldsymbol{\nu})$, $U_m(\boldsymbol{\nu})$ and $\frac{\partial U_i(\boldsymbol{\nu})}{\partial \tau}$. This allows us to split the expectations in (3.54) :

$$\begin{aligned}
E \left\{ \frac{\dot{F}(U_i(\boldsymbol{\nu})) \dot{F}(U_m(\boldsymbol{\nu}))}{F(U_i(\boldsymbol{\nu})) F(U_m(\boldsymbol{\nu}))} \frac{\partial U_i(\boldsymbol{\nu})}{\partial \tau} \frac{\partial U_m(\boldsymbol{\nu})}{\partial u_l} \right\} &= E \left\{ \frac{\dot{F}(U_i(\boldsymbol{\nu})) \dot{F}(U_m(\boldsymbol{\nu}))}{F(U_i(\boldsymbol{\nu})) F(U_m(\boldsymbol{\nu}))} \frac{\partial U_i(\boldsymbol{\nu})}{\partial \tau} \right\} \\
&E \left\{ \frac{\partial U_m(\boldsymbol{\nu})}{\partial u_l} \right\}. \quad (3.56)
\end{aligned}$$

Noting that the last expectation is equal to zero, it follows immediately that $[\mathbf{I}(\boldsymbol{\nu})]_{1;2}$ is also equal to zero. Thus, we show analytically that the two parameters τ and f_c are decoupled. The same manipulations are used to prove that τ and θ are also decoupled. Therefore, the FIM is block-diagonal structured as given by (3.8).

Appendix B

$$\text{Proof of } E \left\{ \frac{\dot{F}(U_i(\tau)) \dot{F}(V_l(\tau))}{F(U_i(\tau)) F(V_l(\tau))} \dot{U}_i(\tau) \dot{V}_l(\tau) \right\} = 0$$

In the following, we briefly show that $E \left\{ \frac{\dot{F}(U_i(\tau)) \dot{F}(V_l(\tau))}{F(U_i(\tau)) F(V_l(\tau))} \dot{U}_i(\tau) \dot{V}_l(\tau) \right\} = 0$. By definition, $U_i(\tau)$ depends on the real part of $\tilde{y}(t)$, while $V_l(\tau)$ involves the imaginary part of $\tilde{y}(t)$, which are statistically independent. It follows that $U_i(\tau)$ and $V_l(\tau)$ are independent. The same arguments hold to show the statistical independence of $\dot{U}_i(\tau)$ and $\dot{V}_l(\tau)$. Then, it immediately follows that :

$$E \left\{ \frac{\dot{F}(U_i(\tau)) \dot{F}(V_l(\tau))}{F(U_i(\tau)) F(V_l(\tau))} \dot{U}_i(\tau) \dot{V}_l(\tau) \right\} = E \left\{ \frac{\dot{F}(U_i(\tau))}{F(U_i(\tau))} \dot{U}_i(\tau) \right\} E \left\{ \frac{\dot{F}(V_l(\tau))}{F(V_l(\tau))} \dot{V}_l(\tau) \right\} \quad (3.57)$$

And since $\dot{U}_i(\tau)$ and $U_i(\tau)$ are statistically independent (see Appendix C), each with mean zero, we obtain :

$$E \left\{ \frac{\dot{F}(U_i(\tau)) \dot{F}(V_i(\tau))}{F(U_i(\tau)) F(V_i(\tau))} \dot{U}_i(\tau) \dot{V}_i(\tau) \right\} = 0. \quad (3.58)$$

Appendix C

C.1 - Pdfs of $U_i(\tau)$ and $V_i(\tau)$

In this Appendix, we establish the joint pdf of $U_i(\tau)$ and $V_i(\tau)$ defined, respectively, in (3.24) and (3.25). To that end, we define the *proper* complex random variable $Z_i(\tau) = \int_{-\infty}^{+\infty} \tilde{y}_i(t) h(t - iT - \tau) dt$. It can be easily seen that $Z_i(\tau) = U_i(\tau) + jV_i(\tau)$ and that $P(Z_i(\tau)) = P(U_i(\tau), V_i(\tau))$. Using the same algebraic manipulations from (3.16) through (3.23), we establish the pdf of $Z_i(\tau)$ as follows :

$$\begin{aligned} P(Z_i(\tau)) &= \frac{1}{M} \frac{1}{\pi\sigma^2} \exp \left\{ -\frac{|Z_i(\tau)|^2}{\sigma^2} \right\} H_i(\tau) \\ &= \frac{4}{M\pi\sigma^2} \exp \left\{ -\frac{U_i^2(\tau) + V_i^2(\tau)}{\sigma^2} \right\} F(U_i(\tau)) F(V_i(\tau)) \\ &= P(U_i(\tau)) P(V_i(\tau)), \end{aligned} \quad (3.59)$$

where

$$P(U_i(\tau)) = \frac{1}{\sqrt{\pi\sigma^2}} \sqrt{\frac{4}{M}} \exp \left\{ -\frac{U_i^2(\tau)}{\sigma^2} \right\} F(U_i(\tau)), \quad (3.60)$$

$$P(V_i(\tau)) = \frac{1}{\sqrt{\pi\sigma^2}} \sqrt{\frac{4}{M}} \exp \left\{ -\frac{V_i^2(\tau)}{\sigma^2} \right\} F(V_i(\tau)). \quad (3.61)$$

Note that the factorization of the joint pdf $P(U_i(\tau), V_i(\tau))$ of $U_i(\tau)$ and $V_i(\tau)$ to their elementary pdfs confirms that these are two independent random variables.

C.2 - Proof of Statistical Independence of $U_i(\tau)$ and $\dot{U}_i(\tau)$

First, note that, $U_i(\tau)$ can be written as :

$$U_i(\tau) = \sqrt{E_s} \Re\{a_i\} + \beta_i, \quad (3.62)$$

where

$$\beta_i = \int_{-\infty}^{+\infty} \Re\{\tilde{w}(t)\} h(t - iT - \tau) dt. \quad (3.63)$$

Therefore, $\dot{U}_i(\tau)$ is given by :

$$\begin{aligned}\dot{U}_i(\tau) &= \sum_{m=1}^K \Re\{a_m\} \dot{g}((i-m)T) - \int_{-\infty}^{+\infty} \Re\{\tilde{w}(t)\} \dot{h}(t-iT-\tau) dt \\ &= \sum_{m=1}^K \Re\{a_m\} \dot{g}((i-m)T) - \dot{\beta}_i.\end{aligned}\quad (3.64)$$

In addition, $\Re\{a_i\}$ and $\dot{\beta}_i$ are independent since the noise and the transmitted symbols are independent. Recall also that $\dot{g}(0) = 0$ (the maximum of $g(x)$ is located at 0). Then, $\sum_{m=1}^K \Re\{a_m\} \dot{g}((i-m)T)$ and $\Re\{a_i\}$ are also independent. Moreover, β_i and $\dot{\beta}_i$ are obtained by a linear transformation of the Gaussian process $\Re\{w(t)\}$. Hence they are also Gaussian processes. Then, since the cross-correlation of β_i and $\dot{\beta}_i$ is equal to zero, as shown below :

$$\begin{aligned}\mathbb{E}\{\beta_i \dot{\beta}_i\} &= \mathbb{E}\left\{ \int_{iT+\tau}^{(i+1)T+\tau} \Re\{w(t_1) e^{-j(2\pi f_c t_1 + \theta)}\} \Re\{w(t_2) e^{-j(2\pi f_c t_2 + \theta)}\} \times \right. \\ &\quad \left. - h(t_1 - iT - \tau) \dot{h}(t_2 - iT - \tau) dt_1 dt_2 \right\} \\ &= \frac{\sigma^2}{2} \iint_{-\infty}^{+\infty} \delta(t_1 - t_2) h(t_1) \dot{h}(t_2) dt_1 dt_2 \\ &= \frac{\sigma^2}{2} \dot{g}(0) \\ &= 0,\end{aligned}\quad (3.65)$$

then, β_i and $\dot{U}_i(\tau)$ are actually two uncorrelated Gaussian random processes and therefore they are independent. Thus, $U_i(\tau)$ and $\dot{U}_i(\tau)$ are independent.

Appendix D

Derivation of the Analytical Expressions for the CRLBs in Case of BPSK and MSK Modulations

Starting from the expression of the log-likelihood function given in (3.47), we will consider the two cases of BPSK and MSK separately. Starting with BPSK-modulated signals, we show that the log-likelihood function in (3.47) reduces to :

$$L(\tau) = \sum_{i=1}^K \ln \left(\cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} U_i(\tau) \right) \right), \quad (3.66)$$

where

$$U_i(\tau) = \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)\} h(t-iT-\tau) dt. \quad (3.67)$$

Then, the first derivative of the log-likelihood function with respect to the time delay parameter, τ , is given by :

$$\frac{\partial L(\tau)}{\partial \tau} = \frac{2\sqrt{E_s}}{\sigma^2} \sum_{i=1}^K \frac{\sinh\left(\frac{2\sqrt{E_s}}{\sigma^2} U_i(\tau)\right)}{\cosh\left(\frac{2\sqrt{E_s}}{\sigma^2} U_i(\tau)\right)} \dot{U}_i(\tau), \quad (3.68)$$

where $\dot{U}_i(\tau)$ denotes the first derivative of $U_i(\tau)$ with respect to τ . It is easy to see that :

$$\dot{U}_i(\tau) = \sqrt{E_s} \sum_{m=1}^K a_m \dot{g}((i-m)T) - \int_{-\infty}^{+\infty} \Re\{\tilde{w}_i(t)\} \dot{h}(t-iT-\tau) dt. \quad (3.69)$$

Now injecting (3.68) in (3.7), we obtain :

$$\begin{aligned} [\mathbf{I}]_{1;1} &= 4 \frac{E_s}{\sigma^4} \sum_{i=1}^K \sum_{l=1}^K E \left\{ \tanh\left(\frac{2\sqrt{E_s}}{\sigma^2} U_i(\tau)\right) \tanh\left(\frac{2\sqrt{E_s}}{\sigma^2} U_l(\tau)\right) \dot{U}_i(\tau) \dot{U}_l(\tau) \right\} \\ &= 4 \frac{E_s}{\sigma^4} \sum_{i=1}^K \sum_{l=1}^K E_{i,l}. \end{aligned} \quad (3.70)$$

Note that (3.70) is similar to (3.32) (obtained in the case of square QAM modulations). Thus, for the same reasons, it is more convenient to separate the cases when $i = l$ and $i \neq l$. Moreover, it can be shown that the pdf of $U_i(\tau)$ is given by :

$$P(U_i(\tau)) = \frac{1}{\sqrt{\pi\sigma^2}} \exp\left\{-\frac{U_i^2(\tau) + E_s}{\sigma^2}\right\} \cosh\left(\frac{2\sqrt{E_s}}{\sigma^2} U_i(\tau)\right). \quad (3.71)$$

Thus, it can be shown that, after some manipulations, the expectations involved in (3.70) reduce to :

$$\begin{aligned} E \left\{ \tanh^2\left(\frac{2\sqrt{E_s}}{\sigma^2}\right) \right\} &= \frac{\exp\left\{-\frac{E_s}{\sigma^2}\right\}}{\sqrt{\pi\sigma^2}} \int_{-\infty}^{+\infty} \frac{\sinh^2\left(\frac{2\sqrt{E_s}}{\sigma^2} U\right)}{\cosh\left(\frac{2\sqrt{E_s}}{\sigma^2} U\right)} \exp\left\{-\frac{U^2}{\sigma^2}\right\} dU \\ &= 1 - \frac{e^{-\rho}}{\sqrt{2\pi}} \beta(\rho), \end{aligned} \quad (3.72)$$

$$E \left\{ (U_i(\tau))^2 \right\} = E_s \sum_{m=1}^K \sum_{n=1}^K \dot{g}^2((m-n)T) - \frac{\sigma^2}{2} \ddot{g}(0), \quad (3.73)$$

$$\begin{aligned} E_{i,l} &= -\dot{g}^2((i-l)T) \left(\frac{e^{-\rho}}{\sqrt{\pi\sigma^2}} \int_{-\infty}^{+\infty} U \sinh\left(\frac{2\sqrt{E_s}}{\sigma^2} U\right) \exp\left\{-\frac{U^2}{\sigma^2}\right\} dU \right)^2 \\ &= -\rho \dot{g}^2((i-l)T). \end{aligned} \quad (3.74)$$

Finally, we obtain the closed-form expression for the stochastic CRLB of BPSK-modulated signals as follows :

$$\begin{aligned} \text{CRLB}_{\text{BPSK}} = & \left[4\rho \left[\left(1 - \frac{e^{-\rho}}{\sqrt{2\pi}} \beta(\rho) \right) \left(\rho \sum_{m=1}^K \sum_{n=1}^K \dot{g}^2((m-n)T) - \frac{K}{2} \ddot{g}(0) \right) \right. \right. \\ & \left. \left. - \rho \sum_{m=1}^K \sum_{n=1}^K \dot{g}((m-n)T) \right] \right]^{-1}, \end{aligned} \quad (3.75)$$

where $\beta(\rho)$ is defined in (3.49).

Now, consider an MSK-modulated signal. In order to find the derivative of (3.47) with respect to the time delay τ , we need to separate the cases where b_i is real or imaginary. To do so, we assume, without loss of generality, that K is an even number (i.e., $K = 2P$) and $a_0 = 1$. Using these assumptions, the log-likelihood function can be written as :

$$L(\tau) = \sum_{i=1}^P \ln \left(\cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} U_{2i-1}(\tau) \right) \right) + \ln \left(\cosh \left(\frac{2\sqrt{E_s}}{\sigma^2} V_{2i}(\tau) \right) \right), \quad (3.76)$$

where

$$U_i(\tau) = \int_{-\infty}^{+\infty} \Re\{\tilde{y}(t)\} h(t - iT - \tau) dt, \quad (3.77)$$

and

$$V_i(\tau) = \int_{-\infty}^{+\infty} \Im\{\tilde{y}(t)\} h(t - iT - \tau) dt. \quad (3.78)$$

Then, the first derivative of (3.76) with respect to τ is given by :

$$\frac{\partial L(\tau)}{\partial \tau} = \frac{2\sqrt{E_s}}{\sigma^2} \sum_{i=1}^P \left(\tanh \left(\frac{2\sqrt{E_s}}{\sigma^2} U_{2i-1}(\tau) \right) \dot{U}_{2i-1}(\tau) + \tanh \left(\frac{2\sqrt{E_s}}{\sigma^2} V_{2i}(\tau) \right) \dot{V}_{2i}(\tau) \right) \quad (3.79)$$

with $\dot{U}_{2i-1}(\tau)$ and $\dot{V}_{2i}(\tau)$ being the derivatives of $U_{2i-1}(\tau)$ and $V_{2i}(\tau)$ with respect to τ , respectively. Then, the first diagonal element of the FIM matrix is expressed as :

$$\begin{aligned} [\mathbf{I}(\tau)]_{1,1} = & E \left\{ \left(\frac{\partial L(\tau)}{\partial \tau} \right)^2 \right\} \\ = & \frac{4E_s}{\sigma^4} \sum_{i=1}^P \sum_{l=1}^P E \left\{ \tanh \left(\frac{2\sqrt{E_s}}{\sigma^2} U_{2i-1}(\tau) \right) \tanh \left(\frac{2\sqrt{E_s}}{\sigma^2} U_{2l-1}(\tau) \right) \dot{U}_{2i-1}(\tau) \dot{U}_{2l-1}(\tau) \right\} \\ & + 2 \sum_{i=1}^P \sum_{l=1}^P E \left\{ \tanh \left(\frac{2\sqrt{E_s}}{\sigma^2} U_{2i-1}(\tau) \right) \tanh \left(\frac{2\sqrt{E_s}}{\sigma^2} V_{2l}(\tau) \right) \dot{U}_{2i-1}(\tau) \dot{V}_{2l}(\tau) \right\} \\ & + \sum_{i=1}^P \sum_{l=1}^P E \left\{ \tanh \left(\frac{2\sqrt{E_s}}{\sigma^2} V_{2i}(\tau) \right) \tanh \left(\frac{2\sqrt{E_s}}{\sigma^2} V_{2l}(\tau) \right) \dot{V}_{2i}(\tau) \dot{V}_{2l}(\tau) \right\}. \end{aligned} \quad (3.80)$$

Note that (3.80) is equivalent to (3.31). Then for the same reasons, $[\mathbf{I}(\tau)]_{1;1}$ reduces simply to :

$$[\mathbf{I}(\tau)]_{1;1} = \frac{8E_s}{\sigma^2} \sum_{i=1}^P \sum_{l=1}^P E \left\{ \tanh \left(\frac{2\sqrt{\rho}}{\sigma^2} U_{2i-1} \right) \tanh \left(\frac{2\sqrt{\rho}}{\sigma^2} U_{2l-1} \right) \dot{U}_{2i-1} \dot{U}_{2l-1} \right\} \quad (3.81)$$

which is similar to (3.70) in the case of BPSK modulation. Thus we obtain the same expression for the stochastic CRLB in case of MSK and BPSK transmissions as given by (3.75).

Bibliographie

- [1] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, vol. 37, pp. 81-91, 1945.
- [2] S.M. Kay, *Fundamentals of Statistical Signal Processing : Estimation Theory*, Upper Saddle River, NJ : Prentice Hall, 1993.
- [3] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its applications to synchronization problems," *IEEE Trans. Comm.*, vol. 24, pp. 1391-1399, Feb.-Apr. 1994.
- [4] U. Mengali, and A.N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York : Plenum, 1997.
- [5] H. Steendam and M. Moeneclaey, "Low-SNR limit of the Cramer-Rao bound for estimating the time delay of a PSK, QAM, or PAM Waveform," *IEEE Comm. Letters*, vol. 5, pp. 31-33, Jan. 2001.
- [6] M. Moeneclaey, "On the true and the modified Cramer-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters," *IEEE Trans. Comm.*, vol. 46, pp. 1536-1544, Nov. 1998.
- [7] J. Riba, J. Sala, and G. Vázquez, "Conditional maximum likelihood timing recovery : Estimators and bounds", *IEEE Trans. Sig. Process.*, vol. 49, no. 4, pp. 835-850, 2001
- [8] G. Vázquez, and J. Riba, " Non-data-aided digital synchronization," *Signal Processing Advances in Wireless Communications*, Englewood Cliffs, NJ : Prentice-Hall, 2000, vol. 2, ch. 9.
- [9] I. Bergel and A.J. Weiss, "Cramér-Rao bound on timing recovery of linearly modulated signals with no ISI," *IEEE Trans. Comm.*, vol. 51, pp.134-641, Apr. 2003.
- [10] N. Noels, H. Wymeersch, H. Steendam, and M. Moeneclaey, "True Cramér-Rao bound for timing recovery from a bandlimited linearly modulated waveform with unknown carrier phase and frequency" *IEEE Trans. Comm.*, vol. 52, pp. 473-383, Mar. 2004.
- [11] F. Bellili, A. Stéphenne, and S. Affes, "Cramér-Rao bounds for NDA SNR estimates of square QAM modulated signals", *Proc. of IEEE WCNC'09*, Budapest, Hungary, Apr. 5-8, 2009.

- [12] F. Bellili, A. Stéphenne, and S. Affes, "Cramér-Rao lower bounds for non-data-aided SNR estimates of square QAM modulated transmissions", *IEEE Trans. Comm.*, vol. 58, pp. 3211-3218, Nov. 2010.
- [13] C.J. Le Martret and G.B. Giannakis, "All-digital PAM impulse radio for multiple access through frequency-selective multipath," in *Proc. IEEE GLOBECOM'2000*, pp. 77-81, 2000.
- [14] M.Z. Win and R.A. Scholtz, "Impulse radio : How it works," *IEEE Comm. Lett.*, vol. 2, no. 2, pp. 36-38, Feb. 1998.

Chapitre 4

A Maximum Likelihood Time Delay Estimator in a Multipath Environment Using Importance Sampling

¹A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, “A Maximum Likelihood Time Delay Estimator in a Multipath Environment Using Importance Sampling”, *submitted in IEEE Trans. on Sign. Process.*

IEEE (c)

DOI: 10.1109/TSP.2012.2222402

Abstract

Dans cet article, nous présentons une nouvelle implémentation du critère de maximum de vraisemblance pour l'estimation du délai de propagation dans un milieu multi-trajet, puis nous étendons la méthode proposée pour l'estimation de la différence du temps d'arrivée quand le signal émis est inconnu. La nouvelle technique implémente le concept de l'"importance sampling" (IS) pour trouver le maximum global de la fonction de vraisemblance. Nous évitons la traditionnelle recherche multidimensionnelle et les méthodes itératives pour maximiser la fonction de vraisemblance. Nous montrons par simulation que cette méthode permet d'estimer des délais très proches et offre de meilleures performances que les méthodes sous-optimales telles que MUSIC. Le principal avantage de cette méthode est que la convergence au maximum global est garanti contrairement aux algorithmes itératifs qui dépendent étroitement de l'initialisation.

In this paper, we present a new implementation of the maximum likelihood criterion for the estimation of the time delays in a multipath environment, and then we extend the proposed method to the estimation of the time difference of arrival when the transmitted signal is unknown. The new technique implements the concept of important sampling (IS) to find the global maximum of the compressed likelihood function in a modest computational manner. Thus we avoid the traditional multidimensional grid search or the iterative methods to maximize the compressed likelihood function. We show by simulations that the new technique allows the estimation of very close delays and surpasses suboptimal techniques such as the MUSIC algorithm. The main advantage of our method is the guaranteed convergence to the global maximum, contrarily to the popular iterative implementation of the maximum likelihood criterion by the well known *expectation maximization* algorithm.

4.1 Introduction

Time delay estimation is a well studied problem with applications in many areas such as radar [1], sonar [2], and wireless communication systems [3]. Typically, to allow estimation of the time delay, an *a priori* known waveform is transmitted through a multipath environment, which consists of several propagation paths, among which the dominant ones, relatively few, are considered. If the transmitted waveform is unknown, only the difference of arrival times can be estimated from the received signals at multiple separated sensors [4]. In what follows, we will treat the two cases.

These two time delay estimation problems have been extensively studied in previous years [5-7]. The maximum likelihood (ML) estimator is well known to be an optimal technique. For the problem at hand, the likelihood function depends on the time delays and on the complex channel coefficients making its solution intractable in a closed-form. A direct im-

plementation of this criterion requires a multidimensional grid search, whose complexity increases with the number of unknown delays. Therefore, many iterative methods, such as the *expectation maximization* (EM) algorithm, have been developed to achieve the well known Cramér-Rao lower bound (CRLB) at a lower cost. But their performances are closely linked to the initialization values and their convergence may take many complex iterative steps and therefore, a trade-off must be found between complexity and accuracy. Hence, there is yet a need for developing a non-grid-search-based and a non-iterative ML estimator. Alternatively, sub-optimal methods based on the eigen-decomposition of the sample covariance matrix, which initially gained much interest in the direction of arrival estimation, were later exploited in the context of time delay estimation [8-9]. While these suboptimal techniques offer an attractive reduced complexity compared to the grid search implementation of the ML criterion, they still suffer from heavy computation steps due to the eigenvalue decomposition. Moreover, their performances are relatively poor compared to the ML estimator, especially for closely spaced delays and/or few numbers of samples. Motivated by these facts, we derive, in this paper, a new non-iterative implementation of the ML time delays estimator which avoids the multidimensional grid search by applying :

- i) the global maximization theorem of Pincus proposed in [10] and
- ii) a powerful Monte Carlo technique called importance sampling (IS) offering thereby an efficient tool to find the global maximum of the likelihood function.

Note here that many other traditional Monte Carlo techniques (besides the IS method) can also be successfully applied. However, unlike the IS method, they often require a larger number of realizations that are, in addition, usually generated according to a complex probability density function (pdf). Hence they appear to be less attractive for practical considerations. In this sense, the importance sampling technique lends itself as a powerful alternative in which the required realizations are easily generated according to a simpler pdf. Additionally, it offers a way to process the obtained realizations in a more judicious manner [11]. This method has indeed been applied to the estimation of the direction of arrival (DOA) [13], the joint DOA-Doppler frequency [12] and more recently to the estimation of the time delay in the context of modulated signals and a single propagation path [14]. Based on the results of these works, the IS technique was shown to dramatically reduce the computational complexity of the ML estimates while still providing high accuracy.

The remainder of this paper is organized as follows. In section II, we present the system model for the active mode (i.e., known transmitted pulse) and derive the corresponding compressed likelihood function to be maximized. In section III, we detail the global maximization method applied to our problem. In section IV, the importance sampling technique is described and then applied to the estimation of the time delays both in active and passive (unknown transmitted pulse). Simulation results are discussed in section V and, finally, some concluding remarks are drawn out in section VI.

4.2 System Model and Compressed Likelihood Function

Consider an *a priori* known signal $x(t)$ transmitted through a multipath environment. The received signal is a superposition of multiple delayed replicas of the known transmitted waveform; modeled as follows :

$$y(t) = \sum_{i=1}^P \alpha_i x(t - \tau_i) + w(t), \quad (4.1)$$

where P is the total number of multipath components, $w(t)$ is an additive noise and $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_P]^T$ are the unknown complex path gains resulting from scattering and fading through the propagation medium. In addition, $\{\tau_i\}_{i=1}^P$ are the unknown time delays to be estimated and gathered in the vector $\tau = [\tau_1, \tau_2, \dots, \tau_P]^T$. If $F_s = 1/T_s$ is the sampling frequency, the resulting samples, taken at instances $\{nT_s\}_{n=1}^N$ are :

$$y(nT_s) = \sum_{i=1}^P \alpha_i x(nT_s - \tau_i) + w(nT_s), \quad n = 0, 1, \dots, N - 1, \quad (4.2)$$

where N stands for the total number of available samples.

In general, the IS principle is suitable for the estimation of non-linear parameters in the general linear models (GLM) described as :

$$\mathbf{y} = \Phi(\boldsymbol{\theta})\mathbf{s} + \mathbf{w} \quad (4.3)$$

where $\mathbf{y} = [y(0), y(T_s), \dots, y((N - 1)T_s)]^T$ is the received data vector which depends linearly on some nuisance unknown parameters \mathbf{s} and non-linearly on the delays $\boldsymbol{\theta}$. However, in contrast to the single-path scenario in [14], the formulation of the input-output relationship in (4.2) cannot be directly transformed into a GLM analogous to (4.3). Here, the received samples are transformed into the frequency domain where the model could be expressed in a matrix form. In fact, taking the discrete Fourier transform of (4.2), we obtain :

$$Y(k) = \sum_{i=1}^P \alpha_i X(k) e^{-\frac{j2\pi k \tau_i}{N}} + W(k), \quad k = 0, 1, \dots, N - 1, \quad (4.4)$$

where $\{Y(k)\}_{k=0}^{N-1}$, $\{X(k)\}_{k=0}^{N-1}$ and $\{W(k)\}_{k=0}^{N-1}$ are the discrete Fourier transforms (DFTs) of $y(nT_s)$, $x(nT_s)$ and $w(nT_s)$, respectively. Then, considering this transformation, the channel coefficients vector α and the time delays τ manifest themselves as the linear and non-linear unknown parameters, respectively. Hence we transform the basic model in (4.2) into the general form of (4.3), using a compact representation of (4.4) as follows :

$$\mathbf{Y} = \Phi_a(\boldsymbol{\tau})\boldsymbol{\alpha} + \mathbf{W}, \quad (4.5)$$

in which $\mathbf{Y} = [Y(0), Y(1), \dots, Y(N-1)]^T$ is viewed as the received vector, $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_P]^T$ and the matrix² $\Phi_a(\boldsymbol{\tau})$ depends only on the unknown delays gathered in the vector $\boldsymbol{\tau}$ and is given by :

$$\Phi_a(\boldsymbol{\tau}) = [\phi_a(\tau_1), \phi_a(\tau_2), \dots, \phi_a(\tau_P)], \quad (4.6)$$

with the columns $\{\phi_a(\tau_i)\}_{i=1}^P$ being defined as :

$$\phi_a(\tau_i) = [X(0), X(1)e^{-\frac{j2\pi\tau_i}{N}}, X(2)e^{-\frac{j2\pi 2\tau_i}{N}}, \dots, X(N-1)e^{-\frac{j2\pi(N-1)\tau_i}{N}}]^T, \quad (4.7)$$

and $\mathbf{X} = [X(0), X(1), \dots, X(N-1)]^T$ and $\mathbf{W} = [W(0), W(1), \dots, W(N-1)]^T$ are the $(N \times 1)$ -dimensional vectors containing the DFT coefficients of samples corresponding to the known transmitted pulse and the additive noise components, respectively.

First, we consider the active model where, in contrast to the passive model treated later in section 4.4.3, the transmitted signal $x(t)$ is known to the receiver. Now, following the same arguments of [15], the likelihood function of the active model (4.5) is given by :

$$\Lambda(\boldsymbol{\tau}, \boldsymbol{\alpha}) \propto p(\mathbf{Y}; \boldsymbol{\tau}, \boldsymbol{\alpha}) = \frac{1}{\pi^M \sigma^{2M}} \exp \left\{ -\frac{1}{\sigma^2} (\mathbf{Y} - \Phi_a(\boldsymbol{\tau})\boldsymbol{\alpha})^H (\mathbf{Y} - \Phi_a(\boldsymbol{\tau})\boldsymbol{\alpha}) \right\}, \quad (4.8)$$

where $p(\mathbf{Y}; \boldsymbol{\tau}, \boldsymbol{\alpha})$ is the probability density function (pdf) of \mathbf{Y} parameterized by $\boldsymbol{\tau}$ and $\boldsymbol{\alpha}$ and σ^2 is the spectrum power of the noise. Actually, the ML solution $\hat{\boldsymbol{\tau}}_{ML}$ is defined as the global maximum of the likelihood function in (4.8) with respect to $\boldsymbol{\tau}$. However, this formulation of the likelihood function imposes a joint estimation of $\boldsymbol{\tau}$ and $\boldsymbol{\alpha}$ which is computationally intensive. Therefore, it is of interest to obtain a likelihood function that depends only on $\boldsymbol{\tau}$ that can be more easily handled. Observing that $\Lambda(\boldsymbol{\tau}, \boldsymbol{\alpha})$ is quadratic with respect to $\boldsymbol{\alpha}$, we consider the nuisance parameter, $\boldsymbol{\alpha}$, as deterministic but unknown and substitute, in (4.8), $\boldsymbol{\alpha}$ by the solution $\hat{\boldsymbol{\alpha}}(\boldsymbol{\tau})$ which maximizes the log-likelihood function $L(\boldsymbol{\tau}, \boldsymbol{\alpha}) = \ln \{\Lambda(\boldsymbol{\tau}, \boldsymbol{\alpha})\}$ for a given $\boldsymbol{\tau}$. Indeed, it can be shown that $\hat{\boldsymbol{\alpha}}(\boldsymbol{\tau})$ is given by :

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\tau}) = (\Phi_a^H(\boldsymbol{\tau})\Phi_a(\boldsymbol{\tau}))^{-1} \Phi_a^H(\boldsymbol{\tau})\mathbf{Y}. \quad (4.9)$$

Replacing $\boldsymbol{\alpha}$ in (4.8) by $\hat{\boldsymbol{\alpha}}(\boldsymbol{\tau})$ and omitting the terms that do not interfere in the maximization with respect to $\boldsymbol{\tau}$, we obtain the so-called compressed likelihood function of the system as follows :

$$L_c(\boldsymbol{\tau}) = \frac{1}{\sigma^2} \mathbf{Y}^H \Phi_a(\boldsymbol{\tau}) \left(\Phi_a^H(\boldsymbol{\tau})\Phi_a(\boldsymbol{\tau}) \right)^{-1} \Phi_a^H(\boldsymbol{\tau})\mathbf{Y}. \quad (4.10)$$

4.3 Global Maximization of the Compressed Likelihood Function

To find the desired ML estimate, we need to maximize the compressed likelihood function in (4.10) over $\boldsymbol{\tau}$. Yet, $L_c(\boldsymbol{\tau})$ is non linear with respect to $\boldsymbol{\tau}$; hence, a closed-form solution

²Note that we index $\Phi_a(\boldsymbol{\tau})$ by a to refer to the active mode.

seems analytically intractable. It is quite common in the current literature to solve this maximization problem in an iterative way, as an alternative for the trivial multidimensional grid search. However, iterative approaches require an initial guess, usually taken from the output of another suboptimal algorithm. The *iterative quadratic* ML (IQML) [16], the *simulated annealing* technique [17] and the *expectation maximization* (EM) algorithm [5], taken as example in our simulation, are some of the most famous iterative implementations of the ML estimator. Naturally, the performances of these iterative algorithms depend severely on the available initial guess and may even converge to local maximum reflecting estimates which are completely different from the real values of the delays (corresponding to the global maximum).

In this context, the global maximization theorem proposed by Pincus [4.11] offers an alternative to find the global maximum of multidimensional functions, such as the one at hand in (4.10). Interestingly, it does not require any initialization and guarantees the convergence to the global maximum. The idea is very simple and claims that the solution is given by (4.11) :

$$\hat{\tau}_i = \lim_{\rho \rightarrow \infty} \frac{\int_J \dots \int_J \tau_i \exp \{ \rho L_c(\boldsymbol{\tau}) \} d\boldsymbol{\tau}}{\int_J \dots \int_J \exp \{ \rho L_c(\bar{\boldsymbol{\tau}}) \} d\bar{\boldsymbol{\tau}}}, \quad i = 1, 2, \dots, P, \quad (4.11)$$

where J is the interval in which the delays are confined. Defining the pseudo-pdf³ $L'_{c,\rho_0}(\boldsymbol{\tau})$, for some large value of ρ_0 , as :

$$L'_{c,\rho_0}(\boldsymbol{\tau}) = \frac{\exp \{ \rho_0 L_c(\boldsymbol{\tau}) \}}{\int_J \dots \int_J \exp \{ \rho_0 L_c(\bar{\boldsymbol{\tau}}) \} d\bar{\boldsymbol{\tau}}}, \quad (4.12)$$

then, according to (4.11), the optimal value of τ_i is simply given by :

$$\hat{\tau}_i = \int_J \dots \int_J \tau_i L'_{c,\rho_0}(\boldsymbol{\tau}) d\boldsymbol{\tau}, \quad i = 1, 2, \dots, P. \quad (4.13)$$

Intuitively, we can say that, as ρ_0 tends to infinity, the function $L'_{c,\rho_0}(\boldsymbol{\tau})$ becomes a P -dimensional Dirac-delta function centered at the location of the maximum of $L_c(\boldsymbol{\tau})$. Thus, the ML estimate is simply obtained from the evaluation of the P -dimensional integral in (4.13). Yet, this is a difficult task due to the complexity of the involved integrand function [the pseudo-pdf $L'_{c,\rho_0}(\cdot)$]. One solution is to exploit the fact that $L'_{c,\rho_0}(\cdot)$ is a pseudo-pdf and interpret $\hat{\tau}_i$ as the expected value of τ_i , the i^{th} element of a vector $\boldsymbol{\tau}$ distributed according to the multidimensional pseudo-pdf $L'_{c,\rho_0}(\cdot)$. Therefore, if one is able to generate R realizations of a random vector, $\{\boldsymbol{\tau}_k\}_{k=1}^R$ according to $L'_{c,\rho_0}(\boldsymbol{\tau})$, it is reasonable to approximate the expected value of $\boldsymbol{\tau}$ using Monte Carlo techniques [11] as follows :

$$\hat{\boldsymbol{\tau}} = \frac{1}{R} \sum_{k=1}^R \boldsymbol{\tau}_k. \quad (4.14)$$

³ $L'_{c,\rho_0}(\boldsymbol{\tau})$ is designated as a pseudo-pdf since it has all the properties of a pdf although $\boldsymbol{\tau}$ is not really a random variable.

Hence, we substitute the complex integration in (4.13) by a simple samples average. Clearly, as the number of generated values R increases, the variance of the sample mean becomes smaller and $\hat{\tau}$ gets closer to the global maximum of the compressed likelihood function. Yet a practical issue remains as how to easily generate realizations according to $L'_{c,\rho_0}(\tau)$. The proposed pseudo-pdf is a non-linear function of τ and needs to operate in a multidimensional space, which is not suitable for easy generation of realizations. One solution is to approximate the actual pseudo-pdf by a one-dimensional function and transpose the problem of generating a vector to the generation of P independent variables, then resort to the concept of IS as described in the next section.

4.4 The Importance Sampling Based Time Delays Estimation

4.4.1 IS Concept

Importance sampling is a Monte Carlo technique which makes use of an alternative distribution (carefully designed) to generate realizations. It is usually applied when the original distribution does not have a practical form, like $L'_{c,\rho_0}(\cdot)$ given in our problem.

The approach is based on the following simple observation on the integral involved in (4.13) :

$$\int_J \dots \int_J \tau L'_{c,\rho_0}(\tau) d\tau = \int_J \dots \int_J \tau \frac{L'_{c,\rho_0}(\tau)}{g'(\tau)} g'(\tau) d\tau, \quad (4.15)$$

where $g'(\tau)$ is also assumed to have all the properties of a pdf, called normalized importance function (IF). Then, the left-hand side of (4.15) is interpreted as the mean of $\tau_i \frac{L'_{c,\rho_0}(\tau)}{g'(\tau)}$ when τ is generated according to $g'(\tau)$. Unlike $L'_{c,\rho_0}(\cdot)$, it is of interest to choose $g'(\tau)$ to be a simple function of τ . Then, we use Monte-Carlo methods to numerically compute the expectation as done in (4.14) :

$$\int_J \dots \int_J \tau \frac{L'_{c,\rho_0}(\tau)}{g'(\tau)} g'(\tau) d\tau = \frac{1}{R} \sum_{k=1}^R \tau_k \frac{L'_{c,\rho_0}(\tau_k)}{g'(\tau_k)}, \quad (4.16)$$

where τ_k is now the k^{th} realization of τ according to $g'(\cdot)$.

Clearly, the choice of $g'(\cdot)$ affects the estimation performance. An inappropriate choice of $g'(\cdot)$ may need a large number of realizations R to reduce the estimation variance and result in a higher computational complexity. Therefore, the value of R depends on how much $g'(\cdot)$ resembles $L'_{c,\rho_0}(\cdot)$. In the ideal case, generations according to $g'(\cdot)$ are the same as if they were generated according to $L'_{c,\rho_0}(\cdot)$. Therefore, ideally, the shapes of the two functions $g'(\cdot)$ and $L'_{c,\rho_0}(\cdot)$ should be similar to reduce the variance of the estimator given by (4.16) [13]. On the other hand, we should keep in mind that $g'(\cdot)$ has to be simple enough so that realizations can be easily generated. Thus some tradeoffs are required to choose a

function as simple as possible yet similar to $L'_{c,\rho_0}(\cdot)$. In what follows, we will show that owing some simplifications of $L'_{c,\rho_0}(\cdot)$, we can build an appropriate function $g'(\cdot)$ to properly generate variables.

Now, coming back to the expression of the actual compressed likelihood function (4.10), the inverse matrix $(\Phi_a^H(\tau)\Phi_a(\tau))^{-1}$ makes the compressed likelihood function, and consequently the pseudo-pdf $L'_{c,\rho_0}(\cdot)$, very non-linear with respect to τ . One can approximate $\Phi_a^H(\tau)\Phi_a(\tau)$ by a diagonal matrix to avoid a heavy computation of the inverse. In fact, the diagonal elements of $\Phi_a^H(\tau)\Phi_a(\tau)$ are given by :

$$[(\Phi_a^H(\tau)\Phi_a(\tau))]_{l;l} = \sum_{k=0}^{N-1} |X(k)|^2, \quad l = 1, 2, \dots, P, \quad (4.17)$$

and its off-diagonal elements are :

$$[(\Phi_a^H(\tau)\Phi_a(\tau))]_{m;n} = \sum_{k=0}^{N-1} |X(k)|^2 \exp \left\{ \frac{j2\pi k(\tau_m - \tau_n)}{N} \right\}, \\ m, n = 1, 2, \dots, P, m \neq n. \quad (4.18)$$

It is easy to verify, for $\tau_m \neq \tau_n$, that :

$$[(\Phi_a^H(\tau)\Phi_a(\tau))]_{m;n} < [(\Phi_a^H(\tau)\Phi_a(\tau))]_{l;l}. \quad (4.19)$$

Although this inequality does not give sufficient condition to approximate $\Phi_a^H(\tau)\Phi_a(\tau)$ by a diagonal matrix, we verify statistically that this inequality holds with very high probability for almost all possible values of the delay difference $\tau_m - \tau_n$. To that end, we consider $\tau_{m;n} = \tau_m - \tau_n$ as a random variable uniformly distributed in⁴ $[-T, T]$ and we define $K(\tau_{m;n})$ the ratio (4.18)/(4.17) as follows :

$$K(\tau_{m;n}) = \frac{\sum_{k=0}^{N-1} |X(k)|^2 \exp \left\{ \frac{j2\pi k(\tau_m - \tau_n)}{N} \right\}}{\sum_{k=0}^{N-1} |X(k)|^2}. \quad (4.20)$$

Then, we plot in Fig. 4.1 the complementary cumulative distribution function of $K(\tau_{m;n})$ (randomized according to $\tau_{m;n}$), to verify that the diagonal elements of $\Phi_a^H(\tau)\Phi_a(\tau)$ are indeed dominant, with very high probability, compared to its off-diagonal elements. Therefore, we adopt the following well-justified approximation :

$$\Phi_a^H(\tau)\Phi_a(\tau) \approx \left(\sum_{k=1}^N |X(k)|^2 \right) \mathbf{I}_p, \quad (4.21)$$

where \mathbf{I}_p is the $p \times p$ identity matrix.

⁴We consider here that the delays do not exceed a given real value T (see section 4.4.2 for further details on this assumption).

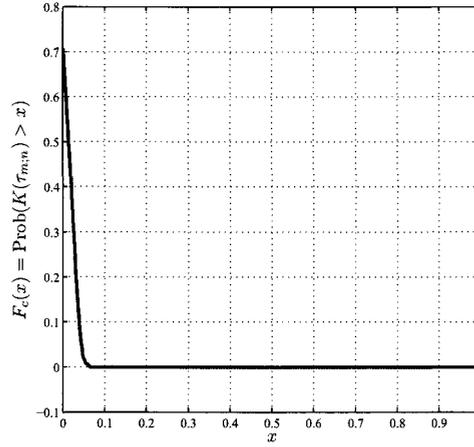


FIGURE 4.1 – Complementary cumulative distribution function of the ratio $K(\tau_{m;n})$.

Then, we define the importance function, $g_{\rho_1}(\cdot)$, without normalization, [i.e., $g'(\tau) = g_{\rho_1}(\tau) / \int g_{\rho_1}(\mathbf{u}) d\mathbf{u}$] in the active case as :

$$g_{\rho_1}(\tau) = \exp \left\{ \frac{\rho_1}{\sigma^2 \sum_{k=1}^N |X(k)|^2} \mathbf{Y}^H \Phi_a(\tau) \Phi_a^H(\tau) \mathbf{Y} \right\}, \quad (4.22)$$

where ρ_1 is another constant different from ρ_0 for some practical reasons. A further discussion on the appropriate choice of ρ_0 and ρ_1 is left to the end of this section.

After some easy algebraic manipulations, we express (4.22) as :

$$g_{\rho_1}(\tau) = \prod_{i=1}^P \exp \left\{ \frac{\rho_1}{\sigma^2 \sum_{k=1}^N |X(k)|^2} I(\tau_i) \right\}, \quad (4.23)$$

where $I(\tau_i)$ is the periodogram of the data in the frequency domain evaluated at each delay τ_i as follows :

$$I(\tau_i) = \left| \sum_{k=1}^N X(k)^* Y(k) \exp \left\{ \frac{j2\pi(k-1)\tau_i}{N} \right\} \right|^2. \quad (4.24)$$

Now, we comment on the advantage of this choice in (4.23) for the importance function (IF). First, we notice that the joint contribution of the different delays in $g_{\rho_1}(\cdot)$ is separable into the product of their individual contributions as seen from (4.23). Hence, we substitute the brute generation of realizations of the vector τ according to a multi-dimensional pdf to the generation of P independent scalar realizations (i.e., one realization for each entry of τ) using the elementary IF, $\bar{g}_{\rho_1}(\tau_i)$, defined as :

$$\bar{g}_{\rho_1}(\tau_i) = \frac{\exp \left\{ \frac{\rho_1}{\sigma^2 \sum_{k=1}^N |X(k)|^2} I(\tau_i) \right\}}{\int_J \exp \left\{ \frac{\rho_1}{\sigma^2 \sum_{k=1}^N |X(k)|^2} I(\bar{\tau}) \right\} d\bar{\tau}}. \quad (4.25)$$

Note here that, the multiplicative terms $X(k)$, $k = 1, 2, \dots, N$ act as weighting factors. They attenuate the contribution of the frequencies with low energy in the computation of $I(\tau_i)$ and hence emphasize the high-SNR frequencies. In fact, this property improves considerably the performance of the estimator compared to some other approaches where the received signal is divided, in the frequency domain, by the DFT of the known transmitted waveform [8]. Actually, this operation is not suitable for narrowband signals since it results in some harmful effects by amplifying the additive noise in the low-energy frequencies. It is suitable only for wideband signals, in contrast to our algorithm which is also adapted to narrowband signals.

Finally, the normalized IF is given by :

$$g'_{\rho'_1}(\boldsymbol{\tau}) = \frac{\prod_{i=1}^P \exp\{\rho'_1 I(\tau_i)\}}{\int_J \dots \int_J \prod_{i=1}^P \exp\{\rho'_1 I(u_i)\} du_i}, \quad (4.26)$$

with

$$\rho'_1 = \frac{\rho_1}{\sigma^2 \sum_{k=1}^N |X(k)|^2}. \quad (4.27)$$

We mention that the choice of the parameters ρ'_1 and ρ_0 are of great importance since its affects the performance of the new estimator. In fact, as already mentioned, $g'_{\rho'_1}(\boldsymbol{\tau})$ is separable as the product of P elementary IFs, $\bar{g}_{\rho'_1}(\cdot)$, corresponding to each delay τ_i (i.e., $g'_{\rho'_1}(\boldsymbol{\tau}) = \prod_{i=1}^P \bar{g}_{\rho'_1}(\tau_i)$). Hence, in practice, we use the same $\bar{g}_{\rho'_1}(\tau)$ P times to generate the P elements of the vector $\boldsymbol{\tau}$. Actually, for a noise-free observation, the function $\bar{g}_{\rho'_1}(\cdot)$ exhibits exactly P lobes centered at the locations of the true delays and at each run, a realization is generated from the vicinity of one of the P lobes. However, in the presence of additive noise, other secondary lobes appear and ultimately affect the generated values. For this reason, the parameter ρ'_1 should be increased to render the objective function $\bar{g}_{\rho'_1}(\cdot)$ more peaked around the actual delays $\{\tau_i\}_{i=1}^P$. This behavior is illustrated in Fig. 4.2 where we plot the function $\bar{g}_{\rho'_1}(\cdot)$ for two values of ρ'_1 .

Yet, we observe that ρ'_1 cannot be increased indefinitely. In fact, very large values of ρ'_1 will ultimately destroy some useful lobes and so useful realizations may not be generated. Obviously, proper choice of ρ'_1 is of great importance. Its optimal value is the highest one that makes at least⁵ P main lobes appear in $\bar{g}_{\rho'_1}(\cdot)$. Moreover, by attenuating the secondary lobes, we reduce the probability of generating undesired realizations. Consequently a good choice of ρ'_1 reduces the number of necessary realizations R and hence the complexity of the estimator.

Recall that the normalized IF $g'_{\rho'_1}(\boldsymbol{\tau})$ is built upon an approximation of the actual compressed likelihood function which results in biased estimates of the delays, especially at low SNR values. However, we emphasize here the fact that this bias can be reduced by the presence of the actual compressed likelihood function in the weighting factor $L'_{c,\rho_0}(\boldsymbol{\tau})/g'_{\rho'_1}(\boldsymbol{\tau})$

⁵ $\bar{g}_{\rho'_1}(\cdot)$ should have exactly P lobes, but the additive noise makes other relatively small secondary lobes appear.

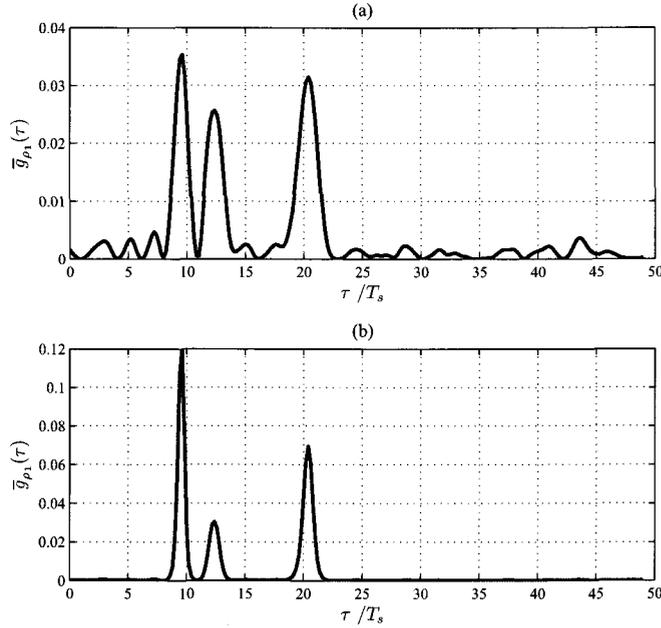


FIGURE 4.2 – Plot of $\bar{g}_{\rho'_1}(\cdot)$ at SNR = 10 for (a) $\rho'_1 = 1$ and (b) $\rho'_1 = 6$.

in (4.16). Thus, we can maximize the contribution of $L'_{c,\rho_0}(\cdot)$ in the weighting factor by choosing ρ'_1 smaller than ρ_0 .

4.4.2 Time Delays Estimation in Active Systems

The IS-based estimator requires the generation of realizations according to $\bar{g}_{\rho'_1}(\cdot)$ then evaluating the following mean values :

$$\hat{\tau}_i = \frac{1}{R} \sum_{k=1}^R \tau_k(i) \frac{L'_{c,\rho_0}(\tau_k)}{g'_{\rho'_1}(\tau_k)}, \quad (4.28)$$

where τ_k is the k^{th} generated vector and $\tau_k(i)$ refers to its i^{th} element.

Roughly speaking, the delays can actually take any positive value, but in practice, they are confined in the interval $[0, T]$ where T is any positive real value that can be chosen high enough⁶ so that $\tau_i \in [0, T]$ for $i = 1, 2, \dots, P$. Therefore, since the parameters are bounded from below and above, it is more convenient to use the circular mean instead of the linear mean in (4.28). The advantages of this operation will be discussed later.

To introduce the concept of circular mean, define a random variable X taking values in the finite interval $[0, 1]$ and denote by $G(X)$ its pdf. The circular mean of X is defined as

⁶In network communications, the delays are usually confined in the symbol duration, whereas for radar and sonar systems, the symbol duration does not really exist and the observation window must be longer than the largest delay.

[18] :

$$E_c\{X\} = \frac{1}{2\pi} \angle \int_0^1 \exp\{j2\pi x\} G(x) dx \quad (4.29)$$

where the operator $\angle(\cdot)$ returns the angle of its complex argument. Suppose that we have a set of x_1, \dots, x_R generated via the pdf $G(\cdot)$, then the circular mean in (4.29) is :

$$E_c\{X\} = \frac{1}{2\pi} \angle \frac{1}{R} \sum_{r=1}^R \exp\{j2\pi x_r\}. \quad (4.30)$$

In our time delays estimation problem, we first normalize the delays by T to transpose them into the interval $[0, 1]$. Then we apply directly the circular mean in (4.30). In this context, the alternative formulation of the IS-based estimator is given by :

$$\hat{\tau}_i = \frac{1}{2\pi T} \angle \frac{1}{R} \sum_{k=1}^R F(\tau_k) \exp\left\{j2\pi \frac{\tau_k(i)}{T}\right\}, \quad (4.31)$$

where $F(\tau_k)$ is the weighting factor defined by :

$$F(\tau_k) = \frac{L'_{c,\rho_0}(\tau_k)}{g'_{\rho'_1}(\tau_k)}. \quad (4.32)$$

From the formulation in (4.31), we only need to find the angle of a complex number. Therefore, any positive multiplicative term will not affect the final result. Thus, the two strictly positive constants $\int_J \dots \int_J \exp\{\rho_0 L_c(\mathbf{x})\} d\mathbf{x}$ and $\int_J \dots \int_J \prod_{i=1}^P \exp\{\rho'_1 I(u_i)\} du_i$, used in the normalization of $L'_{c,\rho_0}(\cdot)$ and $g'_{\rho'_1}(\mathbf{x})$, respectively, can be dropped. However, the exponential terms in both the numerator and the denominator of the weighting factor $F(\cdot)$ may result in an overflow in the computation. To circumvent this problem, $F(\cdot)$ is substituted by $F'(\cdot)$:

$$F'(\tau_k) = \exp\left\{\rho_0 L_c(\tau_k) - \rho'_1 \sum_{i=1}^P I(\tau_k(i)) - \max_{1 \leq l \leq R} \left(\rho_0 L_c(\tau_l) - \rho'_1 \sum_{i=1}^P I(\tau_l(i))\right)\right\} \quad (4.33)$$

Note from (4.33) that the arguments of the exponential terms are either negative or zero and that the values of the exponential cannot exceed one.

Summary of steps

In the following, we recapitulate the different steps for the direct implementation of the new algorithm :

1. Compute the DFT $[Y(0), Y(1), \dots, Y(N-1)]$ of the received signal samples.
2. Use the Fourier transform coefficients to evaluate the periodogram according to (4.24).

3. Compute the samples of the one-dimensional pdf $\bar{g}_{\rho'_1}(\cdot)$, used for the generation of the required realizations, at K points as :

$$\bar{g}_{\rho'_1}(\tau_l) = \frac{\exp\{\rho'_1 I(\tau_l)\}}{\sum_{k=1}^K \exp\{\rho'_1 I(\tau_k)\}}, \quad l = 1, 2, \dots, K, \quad (4.34)$$

where K is the total number of points in the interval J . Note that we substitute the integration in the denominator of $\bar{g}_{\rho'_1}(\cdot)$ by a summation over the discrete points of the interval J .

4. Generate one realization of τ using $g'_{\rho'_1}(\cdot)$. To do so, we generate realizations according to $\bar{g}_{\rho'_1}(\cdot)$ P times to retrieve one realization of the P -dimensional vector τ . More details on this point are left to the Appendix. Repeat this step R times.
5. Evaluate the weighting factor $F'(\tau_i)$ for $i = 1, 2, \dots, R$ and compute the circular mean of the generated values balanced by the weighting factors to find the ML estimate of the multiple unknown delays. Note that we must evaluate the term $\rho_0 L_c(\tau_l) - \rho'_1 \sum_{m=1}^P I(\tau_l(m))$ for all generated vectors $\{\tau_l\}_{l=1}^R$ before computing $F'(\tau_i)$.

4.4.3 Time Delays Estimation in Passive Systems

In a passive system, the transmitted signal is considered to be unknown. In this case, only the time difference of arrival (TDOA) can be estimated from multiple received signals at spatially separated destinations [4]. In this section, we assume, without loss of generality, the presence of two separated sensors. The received signals at these two sensors are modeled as :

$$y_1(t) = \sum_{i=1}^{P_1} \alpha_{1;i} x(t - \tau_{1;i}) + w_1(t), \quad (4.35)$$

$$y_2(t) = \sum_{i=1}^{P_2} \alpha_{2;i} x(t - \tau_{2;i}) + w_2(t), \quad (4.36)$$

where $\{\tau_{n;i}\}_{i=1}^{P_n}$ and $\{\alpha_{n;i}\}_{i=1}^{P_n}$, for $n = 1, 2$, are the delays and the complex gains of the received signal at the n^{th} sensor and $\{P_n\}_{n=1}^2$ are the known numbers of multipath components. For the sake of simplicity, suppose that $y_1(t)$ has only one signal component ($P_1 = 1$). The received signal at this sensor is considered as a reference and hence it is assimilated to a noisy known signal. Then, similarly to (4.4), we express the sampled signals (4.35) and (4.36) in the frequency domain as :

$$Y_1(k) = \alpha_{1;1} X(k) \exp\left\{-\frac{j2\pi k \tau_{1;1}}{N}\right\} + W_1(k), \quad (4.37)$$

$$Y_2(k) = \sum_{i=1}^{P_2} \alpha_{2;i} X(k) \exp \left\{ -\frac{j2\pi k \tau_{2;i}}{N} \right\} + W_2(k),$$

$$k = 0, 1, \dots, N-1, \quad (4.38)$$

where $\{Y_1(k)\}_{k=0}^N$, $\{Y_2(k)\}_{k=0}^N$, $\{W_1(k)\}_{k=0}^N$ and $\{W_2(k)\}_{k=0}^N$ are N samples of the Fourier transform of samples of $y_1(t)$, $y_2(t)$, $w_1(t)$ and $w_2(t)$, respectively. As mentioned above, the TDOAs will be estimated by considering the received signal in the first sensor as a reference. This simplifies to the estimation of the P_2 delay differences $\Delta_\tau^{(i)} = \tau_{2;i} - \tau_{1;1}$ for $i = 1, 3, \dots, P_2$. Therefore, we rewrite (4.38) as follows :

$$Y_2(k) = \sum_{i=1}^{P_2} \beta_i Y_1(k) \exp \left\{ -\frac{j2\pi k \Delta_\tau^{(i)}}{N} \right\} + W_p(k), \quad (4.39)$$

in which

$$\beta_i = \frac{\alpha_{2;i}}{\alpha_{1;1}}, \quad i = 1, 2, \dots, P_2, \quad (4.40)$$

$$W_p(k) = W_2(k) - \sum_{i=1}^{P_2} \beta_i W_1(k) \exp \left\{ -\frac{j2\pi k \Delta_\tau^{(i)}}{N} \right\}. \quad (4.41)$$

Doing so, we highlight the parameters of interest in the expression of $Y_2(k)$. Moreover, there is an analogy between the formulation of the active case in (4.4) and the passive one in (4.39). More precisely, the major difference is in the reference signal (\mathbf{X} for the active case and \mathbf{Y}_1 for the passive case).

Then, gathering all the frequency samples, we obtain the following matrix representation :

$$\begin{aligned} \mathbf{Y}_2 &= [Y_2(0), Y_2(1), \dots, Y_2(N-1)]^T \\ &= \Phi_p(\Delta_\tau) \boldsymbol{\beta} + \mathbf{W}_p, \end{aligned} \quad (4.42)$$

where the matrix $\Phi_p(\Delta_\tau)$ is function of the TDOAs defined as :

$$\Phi_p(\Delta_\tau) = [\phi_p(\Delta_\tau^{(1)}), \phi_p(\Delta_\tau^{(2)}), \dots, \phi_p(\Delta_\tau^{(P_2)})], \quad (4.43)$$

$$\phi_p(\Delta_\tau^{(i)}) = \left[Y_1(0), Y_1(1) \exp \left\{ -\frac{j2\pi \Delta_\tau^{(i)}}{N} \right\}, \dots, Y_1(N-1) \exp \left\{ -\frac{j2\pi(N-1) \Delta_\tau^{(i)}}{N} \right\} \right]^T,$$

$$i = 1, \dots, P_2, \quad (4.44)$$

$$\boldsymbol{\beta} = \left[\frac{\alpha_{2;1}}{\alpha_{1;1}}, \frac{\alpha_{2;2}}{\alpha_{1;1}}, \dots, \frac{\alpha_{2;P_2}}{\alpha_{1;1}} \right]^T, \quad (4.45)$$

and

$$\Delta_\tau = [\Delta_\tau^{(1)}, \Delta_\tau^{(2)}, \dots, \Delta_\tau^{(P_2)}]^T, \quad (4.46)$$

is the vector of the TDOAs of interest. Considering these notations, it turns out that the estimation of the TDOAs can be performed using the same algorithm developed above for

the active system. We only have to substitute the vector $\phi_a(\tau)$ by $\phi_p(\Delta_\tau)$ and $X(k)$ by $Y_1(k)$ in the expression of the periodogram in (4.24). The remaining steps follow in the same way.

Now we rediscuss the estimation problem when P_1 is different from 1. The problem then consists of estimating $P_1 \times P_2$ different parameters. To that end, we refer again to the results of the active case. $P_1 \times P_2$ values are generated according to $\bar{g}_{\rho_1}(\cdot)$ by substituting $X(k)$ and $Y(k)$ in the expression of $I(\cdot)$ by $Y_1(\cdot)$ and $Y_2(\cdot)$, respectively. Then, the generated values are classified from the smallest to the highest and organized as follows :

$$\Delta_{\tau,k} = [\Delta_{\tau,k}^{(1)}, \Delta_{\tau,k}^{(2)}, \dots, \Delta_{\tau,k}^{(P_1)}], \quad (4.47)$$

where each vector $\{\Delta_{\tau,k}^{(i)}\}_{i=1}^{P_1}$ is formed from P_2 TDOAs. The final step consists of evaluating the following means as in (4.31) :

$$\Delta_{\tau,k}^{m,n} = \frac{1}{2\pi T} \angle \frac{1}{R} \sum_{r=1}^R F(\Delta_{\tau,k}^{(m)}) \exp \left\{ j2\pi \frac{\Delta_{\tau,k}^{(m)}(n)}{T} \right\}. \quad (4.48)$$

4.5 Simulation Results

To properly assess the performance of our new IS-based approach, we compare the performance of the proposed method to the *expectation maximization* (EM) algorithm [5] as one representative example of the iterative implementations of the ML criterion and to the MUSIC algorithm proposed in [4] as one representative example of suboptimal subspace-based methods. The estimation error of the three estimators is also compared to the Cramér-Rao lower bound (CRLB) which reflects the theoretical achievable performance taken as a benchmark for all the considered algorithms. In all simulations, the transmitted pulse is a chirp signal which is widely used in radar and sonar applications. The number of snapshots is set to $N = 70$. We consider 3 propagation paths with closely-spaced delays $[3T_s, 6T_s, 8T_s]$. The multipath gain is assumed to be equal for the three paths.

First, we study the influence of the parameters ρ_0 and ρ'_1 on the estimation performance of our new estimator. We verify that there is no dependence on ρ_0 as far as it is chosen to be higher than ρ'_1 (see section 4.4). However, as illustrated in Fig. 4.3, ρ'_1 affects seriously the estimation performance of the new IS-based algorithm. As already mentioned, small values of ρ'_1 may not reduce the effect of the additive noise involved in $\bar{g}_{\rho'_1}(\cdot)$, while too large values reduce the desired lobes revealing the actual delays in $\bar{g}_{\rho'_1}(\cdot)$ thereby preventing their generation. Therefore, an appropriate choice of ρ'_1 is necessary in order to obtain near-optimal performance. We see from Fig. 4.3 that for ρ'_1 taking values between 2 and 7, the performance is almost the same, and thus the optimal value of ρ'_1 can be freely selected from this relatively large range.

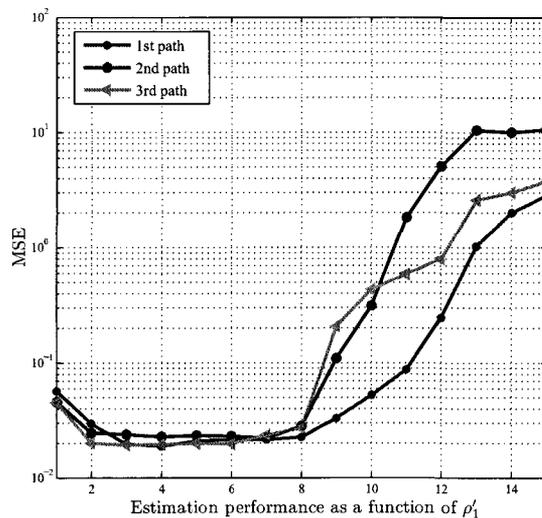


FIGURE 4.3 – Estimation performance as a function of ρ'_1

Now turning to the comparison of the different estimators, we recall that the EM and IS-based algorithms are two different implementations of the ML criterion. They are hence expected to exhibit the same performance since they both try to maximize the same objective function. Yet, it should be kept in mind that the IS-based and MUSIC algorithms do not require any initialization while the EM algorithm is iterative in nature. Consequently, we consider for the EM algorithm two scenarios in which the initial values are selected as random variables, centered at the real time delays and having a variance of $4T_s^2$ and $10T_s^2$ reflecting, respectively, relatively accurate and less accurate initializations. Fig. 4.4 depicts the performance of the three estimators. As expected, the two ML estimators perform better than the MUSIC estimator. However, for less accurate initialization, the performance of the EM algorithm deteriorates considerably over the entire SNR range. We see also that while the MUSIC technique approaches the CRLB only as far as the SNR is sufficiently high; the proposed algorithm performs close to the CRLB over the entire SNR range. This is hardly surprising since the IS-based estimator is far more accurate implementation of the ML criterion. Same conclusions hold for the passive case, also plotted in Fig. 4.6 for $P_1 = 1$ and $P_2 = 3$.

So far, comparisons have been performed as a function of the SNR. To study the resolution power of the different estimators, we consider two propagations paths and vary the delay separation $\Delta\tau = \tau_1 - \tau_2$ at an SNR value of 10 dB. The results are shown in Fig. 4.5. Clearly, as the difference between the delays is small, the estimate is less accurate for the three methods. The two ML-based estimators still perform better than the MUSIC algorithm. For well spaced delays, all the methods perform the same.

Another important point to study is the effect of the signal bandwidth on the estima-

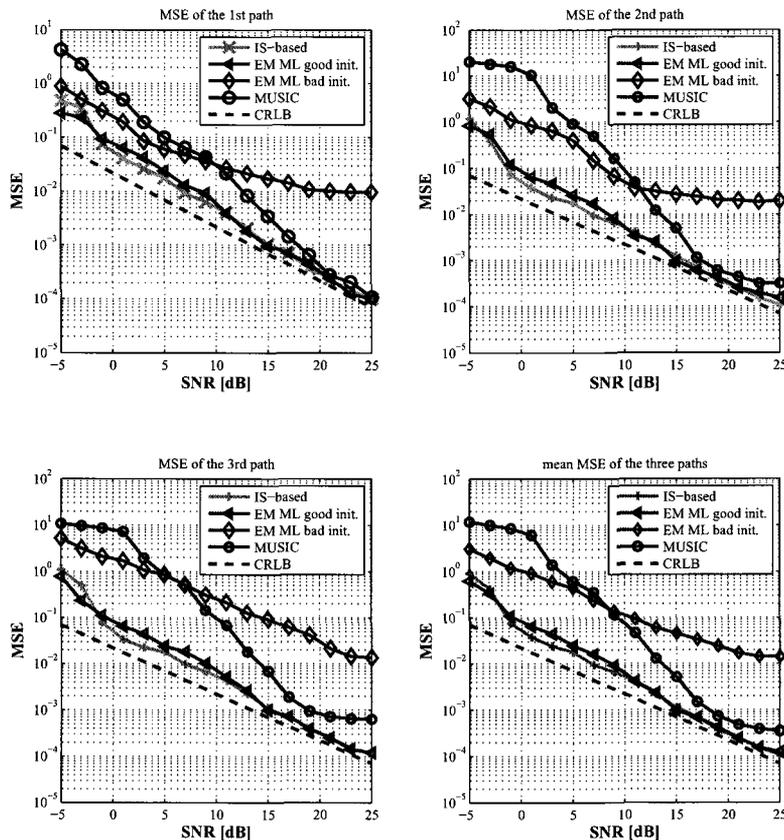


FIGURE 4.4 – Estimation performance of the IS-based, EM ML and the MUSIC-type algorithms in an active system vs. SNR.

tion performance. In fact, since all the derivations are made in the frequency domain, the signal bandwidth (defined in the given example here as the difference between the higher and the lower frequency in the chirp signal) is expected to have an impact on the estimation procedure. Therefore, we compare in Fig. 4.7 the three estimators under different signal bandwidths. Clearly, the proposed method outperforms the MUSIC algorithm over the entire bandwidth range, although the gap between the two methods decreases as the bandwidth increases. Note that the EM algorithm is also less sensitive to bandwidth variations. Same results hold for the passive system but the simulations were not included for the sake of conciseness.

Now we consider the case of time varying channels. While the proposed method is primarily developed under the assumption of constant paths gains, we verify through simulations that it is also robust to time variations and that the IS-based estimator outperforms MUSIC-type methods over relatively low Doppler frequency. Nonetheless, the performances of the two estimators degrade considerably as the Doppler factor increases. In

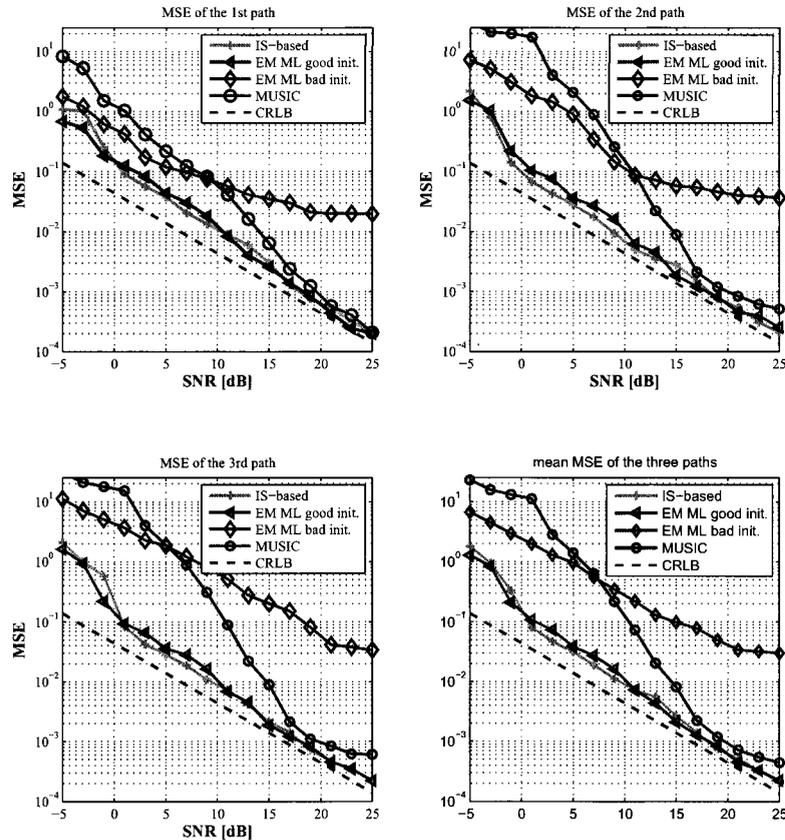


FIGURE 4.5 – Estimation performance of the IS-based, EM ML and the MUSIC-type algorithms in an passive system vs. SNR.

fact, the time variations of the channel coefficients are not taken into account when developing these algorithms, and it was shown in [19] that, in this case, the estimates become necessarily biased⁷. Note, in this case, that we are no longer able to obtain the estimates of the channel coefficients using (4.9). It is for this reason that the EM algorithm was omitted in this scenario since it is based, at each iteration, on an estimate of α , which cannot be performed for time varying channels.

4.6 Conclusion

In this paper, we developed a new implementation of the ML-based estimator for multiple time delays based on the concept of importance sampling (IS). We considered the

⁷In [19], the effect of the time varying envelope has been treated in the case of frequency estimation with the MUSIC and ESPRIT algorithms.

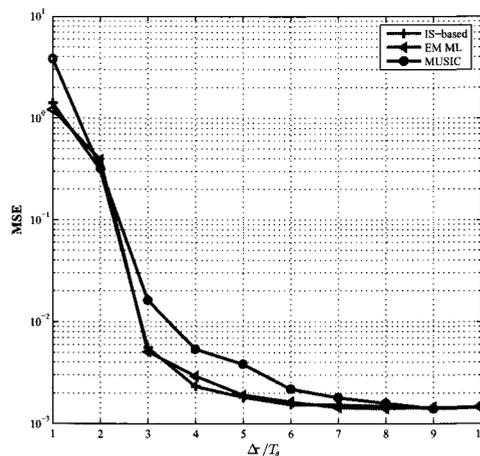


FIGURE 4.6 – Estimation performance of the IS-based, EM ML and the MUSIC-type algorithms in an passive system as a function of $\Delta\tau$.

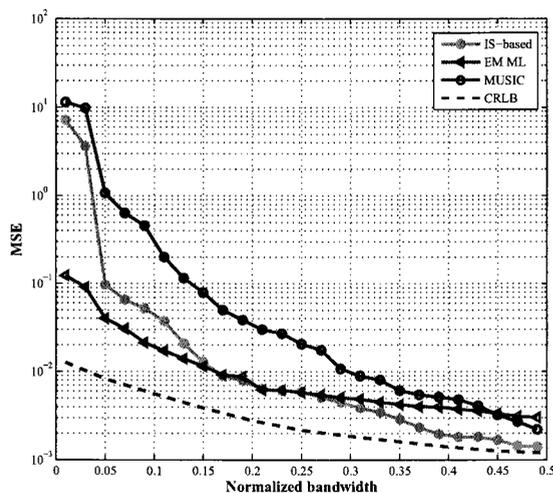


FIGURE 4.7 – Estimation performance of the IS-based, EM ML and the MUSIC-type algorithms vs. signal bandwidth at SNR = 10 dB in an active system.

two cases of active and passive systems. The new algorithm is far less expensive in terms of computational complexity than the traditional multidimensional grid search method. Moreover, unlike the iterative methods, the IS-based algorithm does not suffer from initialization drawbacks. It performs well over the entire SNR range since its convergence to the global maximum of the likelihood function is guaranteed. In addition, it avoids the computation burden of the eigen-decomposition operation that is widely encountered in classical subspace-based techniques in multiple parameters estimation. While these traditional methods perform well even for closely separated delays, at high SNR values, only

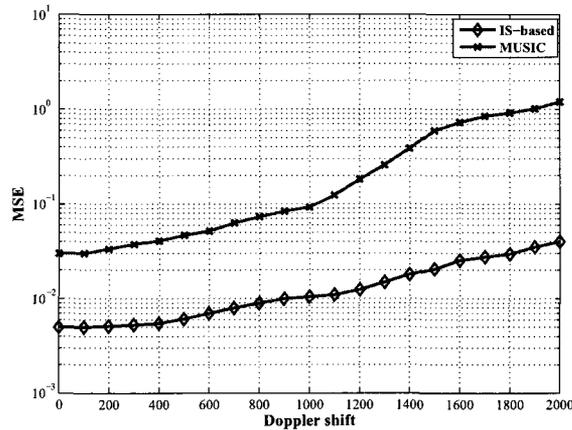


FIGURE 4.8 – Estimation performance of the IS-based and the MUSIC-type algorithms vs. Doppler shift at SNR = 10 dB.

the proposed IS-based technique provides accurate estimates at low SNRs and for challenging cases of more closely-spaced delays. In practice, an appropriate choice of the parameters ρ_0 and ρ'_1 can be performed to further optimize the estimation performance.

Appendix

Method to generate the vector τ

In this appendix, we present some practical hints to easily generate a single realization of the vector τ .

- First, define ξ as a discrete representation of the interval $[0, T]$ (i.e., $\xi = 0 : 1/s : T$ with $1/s$ being a given step for some s).
- Then, generate τ_1 according to $\bar{g}_{\rho'_1}(\cdot)$ using the inverse probability integration method. To do so, consider a vector \mathbf{u} of random variables uniformly distributed over $[0, 1]$. Then find $\tau_1 = \arg \max_{x \in \xi} G(x)$, where $G(x)$ is the cumulative distribution function associated to $\bar{g}_{\rho'_1}(\cdot)$ (for more details, see [14]).
- Then eliminate the generated value τ_1 from ξ so that it cannot be generated again.
- Repeat the last two steps $P-1$ times to generate $\tau_2, \tau_3, \dots, \tau_P$ and obtain one realization of the P -dimensional vector τ .

Bibliographie

- [1] M.I. Skolnik, *Introduction to radar systems*, third edition, McGraw-Hill, New York : 2001.
- [2] G. C. Carter, *Coherence and time delay estimation : An applied tutorial for research, development, test and evaluation engineers*, IEEE Press, New York, 1993.
- [3] M. C. Vanderveen, A.-J. van der Veen, and A. Paulraj, "Estimation of multipath parameters in wireless communications," *IEEE Trans. Sign. Process.*, vol. 46, pp. 682-690, March 1998.
- [4] F.-X. Ge, D. Shen, Y. Peng and V. O. K. Li, "Super-resolution time delay estimation in multipath environments," *IEEE Trans. Circuits Syst.*, vol. 54, no. 9, pp. 1977-1986, Sep. 2007.
- [5] M. Feder, and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm." *IEEE Trans. Acousr., Speech. Sign. Process.*, vol. 36, pp. 477-489, Apr. 1988.
- [6] P. J. Hahn, V. J. Mathews and T. D. Tran, "Adaptive realization of a maximum likelihood time delay estimator" *Conf. on Acoust., Speech, Sign. Process.*, vol. 6, pp. 3121-3124, May 1996.
- [7] K. Gedalyahu, and Y. C. Eldar, "Time delay estimation from low rate samples : A union of subspaces approach," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3017-3031, June 2010.
- [8] M. Pallas and G. Jourdain, "Active high resolution time delay estimation for large BT signals," *IEEE Trans. Sign. Process.*, vol. 39, no. 4, pp. 781-787, Apr. 1991.
- [9] A. M. Bruckstein, T. J. Shan and T. Kailath, "The resolution of overlapping echoes," *IEEE Trans. Acoust., Speech, Sign. Process.*, vol. 33, no. 6, pp. 1357-1367, Dec. 1985.
- [10] M. Pincus, "A closed form solution for certain programming problems," *Oper. Research*, pp. 690-694, 1962.
- [11] A. F. M. Smith, and A. Gelfand, "Bayesian statistics without tears : A sampling-resampling framework," *Amer. Statist.*, vol. 46, pp. 84-88, May 1992.
- [12] H. Wang and S. Kay, "A maximum likelihood angle-Doppler estimator using importance sampling," *IEEE Trans. Aerospace and Electro. Syst.*, vol. 46, pp. 610-622, Apr. 2010.
- [13] S. Saha and S. Kay, "An exact maximum likelihood narrowband direction of arrival estimator," *IEEE Trans. Sign. Process.*, vol. 56, pp. 1082-1092, Oct. 2008.

- [14] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, "A non-data-aided maximum likelihood time delay estimator using importance sampling," *IEEE Trans. Sign. Process.*, vol. 59, pp. 4505-4515, Oct. 2011.
- [15] L. Ng and Y. Bar-Shalom, "Multisensor multitarget time delay vector estimation," *IEEE Trans. Acoust., Speech, Sign. Process.*, vol. 34, pp. 669-677, Aug. 1986.
- [16] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Sign. Process.*, vol. ASSP-59, pp. 1081-1089, Oct. 1986.
- [17] K. Sharman, "Maximum likelihood parameter estimation by simulated annealing," *IEEE Int. Conf. Acoust., Speech, Sign. Process.*, 1988, pp. 2741-2744, Apr. 1988.
- [18] K. V. Mardia, *Statistics of directional data*, New York : Academic, 1972.
- [19] O. Besson and P. Stoica, "Analysis of MUSIC and ESPRIT frequency estimates for sinusoidal signals with lowpass envelopes," *IEEE Trans. Sign. Process.*, vol. 44, no. 9, pp. 2359-2364, Sep. 1996.

Chapitre 5

Maximum Likelihood Time Delay Estimation for Direct-Sequence CDMA Multipath Transmissions

¹A. Masmoudi, F. Bellili, and S. Affes, "Maximum Likelihood Time Delay Estimation for Direct-Sequence CDMA Multipath Transmissions", *to submit in IEEE Trans. on Sign. Process.*

IEEE (c)

Abstract

Dans cet article, nous considérons le problème de synchronisation temporelle pour le Direct-Sequence CDMA (DS-CDMA) dans un canal multi-trajet. Nous dérivons les expressions analytiques de la borne de Cramér-Rao pour l'estimation du retard dans les systèmes uni-porteuse DS-CDMA. Puis nous développons deux algorithmes d'estimation basés sur le critère du maximum de vraisemblance. Le premier reprend la méthode itérative "expectation maximization" (EM). Le second algorithme implémente le critère du maximum de vraisemblance d'une manière non-itérative et retourne le maximum global de la fonction de vraisemblance en utilisant l'IS. Nous généralisons aussi les deux algorithmes et la borne de Cramér-Rao pour les systèmes CDMA multi-porteuses. Les simulations montrent que l'algorithme EM est bien approprié pour les systèmes avec un grand nombre d'antennes alors que l'algorithme IS offre de meilleures performances en présence d'un petit nombre d'antennes.

In this paper, we address the problem of time delay estimation from Direct-Sequence CDMA (DS-CDMA) multipath transmissions. We derive for the first time a closed-form expression for the Cramer-Rao lower bound (CRLB) of multiple time delay estimation in single-carrier (SC) DS-CDMA systems. Then we develop two time delay estimators based on the ML criterion. The first one is based on the iterative expectation maximization (EM) algorithm and provides accurate estimates whenever a good initial guess of the parameters is available at the receiver. The second approach implements the ML criterion in a non-iterative way and finds the global maximum of the compressed likelihood function using the importance sampling technique. Unlike the EM-based algorithm, this non-iterative method does not require any initial guess of the parameters to be estimated. We also extend both the SC CRLB and the proposed SC algorithms to multicarrier (MC)-CDMA systems by exploiting the frequency gain over subcarriers. In this work, the estimation process can be performed using the channel estimate or directly from the received signal and thus we cover all possible cases. By an adequate formulation of the problem, we are able to exploit the time and frequency correlation if the channel estimate is used. We show by simulations that the EM-based algorithm is suitable for CDMA systems with large receive antenna arrays whereas the IS-based offers better performance for small array sizes.

5.1 Introduction

The most important challenge for wireless networks is the development of robust transceivers that are able to transmit at high data rates with a high bandwidth efficiency. Code-division multiple access (CDMA) systems can satisfy this requirement. CDMA has been adopted as the multiple access scheme for the third generation cellular mobile systems

because of its flexibility in cell planning, user capacity, support for different rates and robustness to multipath channel. One of the most important motivations behind the use of CDMA is to increase the number of simultaneous users (user capacity) with acceptable error performance. OFDM based CDMA systems, also called multicarrier (MC)-CDMA, is a promising multiple access for high speed communication system due to robustness against frequency selective fading channel and fully use the available bandwidth [1-2]. However, the performance of these systems is closely linked to synchronization. In the following, we focus on the CDMA array-receiver which has received much interest sequel to the performance potential it carries [3, 4, 5]. Roughly speaking, the post correlation model (PCM) of the despread data presents the signal in an interesting way to apply the researches done in the field of array processing. A suboptimal Root-MUSIC-based estimator was initially developed in [6] to recover the time delay and later refined in [1] to significantly reduce its complexity. This paper also investigates multiple time delay estimation, yet in an optimal way in which the ML criterion is adapted to the PCM.

The problem of high-resolution parameters estimation has been extensively studied in the past. In this context, it is well known that the ML technique always outperforms the other sub-optimal methods in the challenging cases of low signal-to-noise ratio (SNR) values or small number of available data snapshots. However, a direct implementation of the ML criterion requires a multi-dimensional grid search which is of course impractical. Alternatively, eigen-decomposition methods (which reduce the problem to one-dimensional grid search) [7, 8] have attracted much interest due to their simplicity and their high-resolution capacity. Yet, they are mainly based on the sample covariance matrix and require therefore a large number of data snapshots. However, as we will see later, the number of snapshots in CDMA systems is equivalent to the number of receiving antenna elements. Thus applying a traditional suboptimal technique would require a very large number of receiving antenna branches which is also impractical. Consequently, there is a need to derive an efficient implementation of the ML-based estimator that avoids the trivial multidimensional grid search approach. To that end, iterative methods are usually envisioned to find the ML estimates since a closed-form solution is, in most cases, deemed intractable. And to properly evaluate the performance of these estimators, we derive in the first part of this paper a closed-form expression of the Cramèr-Rao lower bound (CRLB).

Motivated by these facts, we develop in the second part of this paper an efficient scheme for the estimation of the delays based on the expectation maximization (EM) algorithm. While this method is widely used in multiple parameters estimation, especially for the estimation of multiple time delay from an incoming waveform [9], it has not been yet adapted to the context of CDMA systems. In few words, the EM method offers an interesting way to decompose the observed signal into different replicas, each one coming from one path, and then treat each component separately. Therefore, the multidimensional optimization task is interestingly transformed into multiple one-dimensional optimization problems, resulting

in tremendous numerical advantages. Under good initialization, the likelihood function of the estimated parameters is increased in each iteration and hence the algorithm converges to its global maximum.

Alternatively, when a good initialization is not available, we resort to the concept of importance sampling (IS) to derive, in the third part of this paper, another technique that finds the global maximum of the likelihood function, in a non-iterative way. In our case, the likelihood function depends on the time delay and the channel covariance matrix. To obtain a function that depends on the unknown delays only, the channel covariance matrix is replaced by its ML estimate (a function of the delays themselves). The resulting objective function, called compressed likelihood function, is then maximized with respect to the unknown time delay. To that end, we use the global maximization theorem introduced in [10] that provides an efficient tool of finding the global maximum of multidimensional functions. However, it still requires the computation of a multidimensional integral which is itself difficult to perform. Yet, this integral can always be tackled empirically using Monte Carlo (MC) methods [11]. Among these MC methods, the importance sampling technique, in particular, has been shown to be a powerful method that reduces considerably the computational complexity. Typically, it was successfully applied to the estimation of direction of arrival (DOA) [12], the joint DOA-Doppler frequency estimation [13] and, more recently, to the estimation of the time delay in the context of a single path and linearly-modulated signals [14]. Compared with the EM ML estimator, this method does not need any initialization and does not suffer from any convergence problem, yet it is computationally more intensive. The contributions of this work cover SC-CDMA and MC-CDMA as well. In the estimation process, we distinguish two major cases. In the first one, we estimate the delay directly from the received signal over the antenna array. In this case, we prove that time, frequency and space dimensions are mixed together to obtain one dimension which is the product of the three. On the other hand, a channel estimate can be obtained prior to time delay estimation [26] then we use the channel estimate to estimate the delays.

This paper is organized as follows. In section II, we briefly introduce the post-correlation model. Then, in section III, The analytical expression of the CRLB is derived. In section IV, we develop the new EM-based ML time delay estimator. In section V, derivation details of the new IS-based estimator are discussed. In section VI, we extend the presented work to the case of multi-carrier (MC)-CDMA. Section VII presents some simulation results that corroborate our findings and finally some concluding remarks are drawn out in section VIII.

5.2 System Model and Background

We consider a CDMA communication system where the receiver is equipped with M receiving antenna elements that capture signals travelling through a multipath propagation environment consisting of P different paths. The signals received on the M antennas are uncorrelated with the spreading code and sampled at the chip rate T_c . Denoting the processing gain by L (i.e., $L = T/T_c$ with T being the symbol duration), the resulting post-correlation data of the spatio-temporal observation of the n^{th} received symbol is modeled by the following matrix form [6] :

$$\mathbf{Z}_n = \mathbf{G}_n \mathbf{\Upsilon}_n \mathbf{D}^T(\boldsymbol{\tau}) s_n + \mathbf{N}_n, \quad (5.1)$$

where $s_n = b_n \psi_n$ is a function of the unknown transmitted symbol b_n and the square root, ψ_n , of the total received power ψ_n^2 and \mathbf{N}_n is the $(M \times L)$ -dimensional post-correlation noise matrix. The p^{th} column of the matrix $\mathbf{D}(\boldsymbol{\tau})$ that gathers the time delay parameters, $\tau_1, \tau_2, \dots, \tau_P$, is given by :

$$\mathbf{d}_p = [\rho_c(-\tau_p), \rho_c(T_c - \tau_p), \dots, \rho_c((L-1)T_c - \tau_p)]^T, \quad (5.2)$$

where $\rho_c(\cdot)$ is the correlation function of the spreading code. \mathbf{G}_n is the $M \times P$ spatial propagation matrix and $\mathbf{\Upsilon}_n$ is a $P \times P$ diagonal matrix representing the normalized power ratios over the different paths [i.e., $\text{trace}(\mathbf{\Upsilon}_n^2) = 1$] [6]. These two matrices can be further gathered in one single spatial-response matrix \mathbf{J}_n (i.e., $\mathbf{J}_n = \mathbf{G}_n \mathbf{\Upsilon}_n$) and by including the scalar term s_n in the matrix² \mathbf{J}_n , a more compact form of \mathbf{Z}_n is given by :

$$\mathbf{Z}_n^T = \mathbf{D}(\boldsymbol{\tau}) \mathbf{J}_n^T + \mathbf{N}_n^T. \quad (5.3)$$

Using the representation in (5.3), the original problem can be interpreted as the estimation of the time delay, involved in the matrix $\mathbf{D}(\boldsymbol{\tau})$, from M snapshots observed on L antenna branches. Each column of \mathbf{Z}_n^T represents an observation vector and the columns of \mathbf{J}_n^T are interpreted as the transmitted signals from P different sources. If we suppose that the delay vector $\boldsymbol{\tau}$ remains constant over N transmitted symbols, a compact representation of (5.3) over N symbols is given by :

$$\begin{aligned} \mathbf{Z} &= [\mathbf{Z}_1^T, \mathbf{Z}_2^T, \dots, \mathbf{Z}_N^T] \\ &= \mathbf{D}(\boldsymbol{\tau}) \mathbf{J}^T + \mathbf{N}^T, \end{aligned} \quad (5.4)$$

with $\mathbf{J}^T = [\mathbf{J}_1^T, \mathbf{J}_2^T, \dots, \mathbf{J}_N^T]$ and $\mathbf{N}^T = [\mathbf{N}_1^T, \mathbf{N}_2^T, \dots, \mathbf{N}_N^T]$.

Usually, high resolution methods (when applied to time delay estimation) transform the

²For the sake of simplicity, we keep the same notation \mathbf{J}_n for $\mathbf{J}_n s_n$. Hence, please note that the following formulation holds, unless specified otherwise, for both data-aided (i.e., s_n is a known reference signal) and non-data-aided transmissions.

problem into the frequency domain in order to obtain a formulation that is similar to the one encountered in frequency estimation [9], [17], [18], after which high-resolution methods, such as Root-MUSIC can be applied to estimate the delays (as in [6]). Following the same logic, we perform a column-by-column fast Fourier transform (FFT) of \mathbf{Z}^T to obtain :

$$\mathbf{Z} = \mathcal{D}(\boldsymbol{\tau})\mathbf{J}^T + \mathcal{N}, \quad (5.5)$$

where \mathcal{N} is the resulting transformed noise matrix and $\mathcal{D}(\boldsymbol{\tau})$ depends only on the unknown delays and is given by :

$$\mathcal{D}(\boldsymbol{\tau}) = [\mathbf{d}(\tau_1), \mathbf{d}(\tau_2), \dots, \mathbf{d}(\tau_P)], \quad (5.6)$$

where the columns $\{\mathbf{d}(\tau_i)\}_{i=1}^P$ are given by :

$$\mathbf{d}(\tau_p) = [c_0, c_1 e^{-\frac{j2\pi\tau_p}{L}}, \dots, c_{L-1} e^{-\frac{j2\pi(L-1)\tau_p}{L}}]^T, \quad (5.7)$$

and $\{c_l\}_{l=0}^{L-1}$ are the FFT coefficients of the spreading code correlation function. Note that, for CDMA systems, the correlation function of a perfect spreading code is a Dirac function, and hence the corresponding FFT coefficients are constant in this ideal case. This feature holds true as a very good approximation even with practical spreading codes [1], [6].

Before exposing the main contributions of this paper, we mention that the following development is applicable on the estimate of the channel coefficients matrix. Actually, considering again the formulation in (5.1), \mathbf{Z}_n can be written as follows :

$$\mathbf{Z}_n = \mathbf{H}_n s_n + \mathbf{N}_n, \quad (5.8)$$

in which $\mathbf{H}_n = \mathbf{J}_n \mathcal{D}^T(\boldsymbol{\tau})$ denotes the overall spatio-temporal propagation matrix. Interestingly, the model in (5.8) can be used to estimate the channel response, \mathbf{H}_n , in an efficient way. One can use any blind channel estimator to obtain an estimate, $\widehat{\mathbf{H}}_n$, of \mathbf{H}_n . More details on this subject can be found in [6]. Now, taking into account the estimation error, $\widehat{\mathbf{H}}_n$ is written as :

$$\widehat{\mathbf{H}}_n^T = \mathcal{D}(\boldsymbol{\tau})\mathbf{J}_n^T + \mathbf{E}_n^T, \quad (5.9)$$

where \mathbf{E}_n^T is the corresponding channel error matrix. The variance of the entries of \mathbf{E}_n^T depends of course on the noise variance in the received signal [15] and it has been stated that the power of \mathbf{E}_n is lower than the power of \mathbf{N}_n , which has the effect of increasing the SNR. Clearly, time delay could be estimated from the column-by-column FFT of $\widehat{\mathbf{H}}_n^T$ in (5.9) as well as from (5.5), with the only difference that the noise power is reduced in (5.9). In the remain of this paper, we consider the formulation in (5.5) since an estimate of the channel is not always available.

5.3 The Cramèr-Rao Lower Bound

Before we deal with the two estimators, we derive in this section a closed-form expression for the CRLB for the problem at hand, which will be used as a benchmark, in addition to the Root-MUSIC algorithm, against which we evaluate the performance of the new estimators.

In fact, the Cramèr-Rao lower Bound (CRLB) is a well known lower bound for the variance of unbiased estimators of an intended parameter. Many works have so far dealt with the evaluation of the CRLB for the time delay estimation problem but, as far as we know, no contributions have been made yet in the context of multipath time delay estimation in DS-CDMA systems. To that end, we assume that the multipath fading coefficients, gathered in \mathbf{J}_n^T , are random variables with unknown covariance matrix \mathbf{R}_J . Therefore, the vector of unknown parameters involved in the estimation process is :

$$\boldsymbol{\alpha} = [\boldsymbol{\tau}, \Re\{\mathbf{R}_J(m, n)\}_{m,n=1}^P, \Im\{\mathbf{R}_J(m, n)\}_{m,n=1}^P, \sigma^2]^T, \quad (5.10)$$

with $\boldsymbol{\tau} = [\tau_1, \tau_2, \dots, \tau_P]^T$. In the following, we suppose that the different columns of \mathbf{Z} , denoted \mathbf{Z}_i , are mutually independent and the columns of \mathcal{N} are also mutually independent and Gaussian distributed. Under these assumptions, the probability density function (pdf) of \mathbf{Z} , parameterized by $\boldsymbol{\tau}$ and \mathbf{R}_J (the covariance matrix of the columns of \mathbf{J}_n^T), is given by :

$$\begin{aligned} \bar{p}(\mathbf{Z}; \boldsymbol{\tau}, \mathbf{R}_J) &= \frac{1}{\pi^{MNL}} \frac{1}{(\det(\mathcal{D}(\boldsymbol{\tau})\mathbf{R}_J\mathcal{D}(\boldsymbol{\tau})^H + \sigma^2\mathbf{I}_L))^{MN}} \\ &\exp \left\{ - \sum_{i=1}^{MN} \mathbf{Z}_i^H (\mathcal{D}(\boldsymbol{\tau})\mathbf{R}_J\mathcal{D}(\boldsymbol{\tau})^H + \sigma^2\mathbf{I}_L)^{-1} \mathbf{Z}_i \right\}, \end{aligned} \quad (5.11)$$

where $\det(\cdot)$ returns the determinant of a given matrix and $\boldsymbol{\tau} = [\tau_1, \tau_2, \dots, \tau_P]^T$. Then the log-likelihood function, $L(\boldsymbol{\tau}, \mathbf{R}_J) = \ln(\bar{p}(\mathbf{Z}; \boldsymbol{\tau}, \mathbf{R}_J))$, reduces simply to :

$$\begin{aligned} L(\boldsymbol{\tau}, \mathbf{R}_J) &= - \ln(\det(\mathcal{D}(\boldsymbol{\tau})\mathbf{R}_J\mathcal{D}(\boldsymbol{\tau})^H + \sigma^2\mathbf{I}_L)) \\ &\quad - \frac{1}{MN} \sum_{i=1}^{MN} \mathbf{Z}_i (\mathcal{D}(\boldsymbol{\tau})\mathbf{R}_J\mathcal{D}(\boldsymbol{\tau})^H + \sigma^2\mathbf{I}_L)^{-1} \mathbf{Z}_i. \end{aligned} \quad (5.12)$$

The entries of the Fisher Information matrix (FIM), denoted here \mathbf{I} , are given by³ :

$$\mathbf{I}(m, n) = M \text{trace} \left\{ \mathbf{R}_Z^{-1} \frac{\partial \mathbf{R}_Z}{\partial \alpha(m)} \mathbf{R}_Z^{-1} \frac{\partial \mathbf{R}_Z}{\partial \alpha(n)} \right\}, \quad (5.13)$$

with \mathbf{R}_Z being the covariance matrix of \mathbf{Z}_i given by :

$$\mathbf{R}_Z = \mathcal{D}(\boldsymbol{\tau})\mathbf{R}_J\mathcal{D}(\boldsymbol{\tau})^H + \sigma^2\mathbf{I}_L. \quad (5.14)$$

³Note that the CRLB of the joint estimation of all parameters is simply the inverse of the FIM \mathbf{I} .

However, the derivation of the FIM starting for (5.13) appears to be intractable. Alternatively, the CRLB is asymptotically equivalent to the error covariance matrix, $\mathbf{C}_{\text{ML}}(\boldsymbol{\tau}) = E\{(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})^T\}$, of the maximum likelihood estimate, as M tends to infinity [25], which means that :

$$\mathbf{C}_{\text{ML}}(\boldsymbol{\tau}) = \text{CRLB}(\boldsymbol{\tau}). \quad (5.15)$$

Recall that the ML estimate $\hat{\boldsymbol{\tau}}$ of $\boldsymbol{\tau}$ verifies the following equation :

$$\frac{\partial L(\hat{\boldsymbol{\tau}}, \mathbf{R}_J)}{\partial \boldsymbol{\tau}} = \mathbf{0}, \quad (5.16)$$

where $\partial L(\cdot)/\partial \boldsymbol{\tau}$ is the gradient of $L(\cdot)$ with respect to $\boldsymbol{\tau}$. Applying the Taylor series expansion to the left-hand side of (5.16) and keeping only the first two terms of this development leads to :

$$\frac{\partial L(\boldsymbol{\tau}, \mathbf{R}_J)}{\partial \boldsymbol{\tau}} + \frac{\partial^2 L(\boldsymbol{\tau}, \mathbf{R}_J)}{\partial \boldsymbol{\tau}^2} (\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}) = \mathbf{0}, \quad (5.17)$$

where $\partial^2 L(\hat{\boldsymbol{\tau}}, \mathbf{R}_J)/\partial \boldsymbol{\tau}^2$ is a Hessian matrix. Hence, from (5.17) we obtain :

$$\hat{\boldsymbol{\tau}} - \boldsymbol{\tau} = - \left[\frac{\partial^2 L_0(\boldsymbol{\tau}, \mathbf{R}_J)}{\partial \boldsymbol{\tau}^2} \right]^{-1} \frac{\partial L(\boldsymbol{\tau}, \mathbf{R}_J)}{\partial \boldsymbol{\tau}}, \quad (5.18)$$

where $\partial^2 L_0(\boldsymbol{\tau}, \mathbf{R}_J)/\partial \boldsymbol{\tau}^2$ is a Hessian matrix when $\hat{\mathbf{R}}_{\mathbf{Z}} \rightarrow \mathbf{R}_{\mathbf{Z}}$ and $\hat{\sigma} \rightarrow \sigma$ as M tends to ∞ . It follows that :

$$\mathbf{C}_{\text{ML}} = \left[\frac{\partial^2 L_0(\boldsymbol{\tau}, \mathbf{R}_J)}{\partial \boldsymbol{\tau}^2} \right]^{-1} \left(\lim_{M \rightarrow \infty} E \left\{ \frac{\partial L(\boldsymbol{\tau}, \mathbf{R}_J)}{\partial \boldsymbol{\tau}} \frac{\partial L(\boldsymbol{\tau}, \mathbf{R}_J)^T}{\partial \boldsymbol{\tau}} \right\} \right) \left[\frac{\partial^2 L_0(\boldsymbol{\tau}, \mathbf{R}_J)}{\partial \boldsymbol{\tau}^2} \right]^{-1} \quad (5.19)$$

The analytical expressions of the gradient $\partial L(\boldsymbol{\tau}, \mathbf{R}_J)/\partial \boldsymbol{\tau}$ and the matrix $\partial^2 L_0(\boldsymbol{\tau}, \mathbf{R}_J)/\partial \boldsymbol{\tau}^2$ (derived in [25]) leads to the following analytical expression for the CRLB of time delay estimates :

$$\text{CRLB}(\boldsymbol{\tau}) = \frac{\sigma^2}{2MN} \left[\Re \left\{ (\mathbf{U}^H [\mathbf{I} - \boldsymbol{\Pi}] \mathbf{U}) * (\mathbf{R}_J \mathbf{D}^H(\boldsymbol{\tau}) \mathbf{R}_{\mathbf{Z}}^{-1} \mathbf{D}(\boldsymbol{\tau}) \mathbf{R}_J)^T \right\} \right]^{-1}, \quad (5.20)$$

where $*$ stands for the element-wise product, $\boldsymbol{\Pi}$ is an orthogonal projector matrix defined as $\boldsymbol{\Pi} = \mathcal{D}(\boldsymbol{\tau}) (\mathcal{D}^H(\boldsymbol{\tau}) \mathcal{D}(\boldsymbol{\tau}))^{-1} \mathcal{D}^H(\boldsymbol{\tau})$ and the matrix \mathbf{U} is defined as follows :

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_P], \quad (5.21)$$

$$\mathbf{u}_i = \frac{\partial \mathbf{d}(\tau_i)}{\partial \tau_i}. \quad (5.22)$$

5.4 Expectation Maximization Algorithm

The EM algorithm is a computational modest method to find the maximum likelihood estimate when a closed-form solution of this one is intractable. Rewrite the log-likelihood function in (5.12) in a more compact form as follow :

$$L(\boldsymbol{\tau}, \mathbf{R}_J) = -\ln(\det(\mathbf{R}_Z)) - \text{trace} \left\{ \mathbf{R}_Z^{-1} \widehat{\mathbf{R}}_Z \right\}, \quad (5.23)$$

where $\widehat{\mathbf{R}}_Z$ being an estimate of \mathbf{R}_Z computed from the columns of \mathbf{Z} as follows :

$$\widehat{\mathbf{R}}_Z = \frac{1}{MN} \sum_{i=1}^{MN} \mathbf{z}_i \mathbf{z}_i^H. \quad (5.24)$$

Note here that the log-likelihood function, $L(\boldsymbol{\tau}, \mathbf{R}_J)$, depends on the delays vector $\boldsymbol{\tau}$ of interest and the covariance matrix \mathbf{R}_J . Hence the problem can be formulated as follows : maximize $L(\boldsymbol{\tau}, \mathbf{R}_J)$ with respect to $\boldsymbol{\tau}$ and \mathbf{R}_J . We also mention that while σ^2 is usually unknown, it can be easily estimated either by averaging the $L - M$ smallest eigen-values of $\widehat{\mathbf{R}}_Z$ or simply by exploiting the estimated power carried out in a previous stage of the receiver [6].

Unfortunately, the above expression of the likelihood function cannot lead to a closed-form solution for its maxima. Thus, we resort as a first option to a well-known iterative algorithm, namely the EM algorithm [16], to resolve this problem numerically. The purpose is to decompose the observation, $\{\mathbf{z}_i\}_{i=1}^{MN}$ into P complete-data, then estimate the delays separately from each complete-data. This is equivalent to performing P parallel maximizations over a one-dimensional space. This method reduces considerably the computational complexity compared to the brute grid search solution. For this purpose, we define the set of complete data as :

$$\mathbf{z}^{(p)}(i) = \mathbf{J}^T(i, p) \mathbf{d}(\tau_p) + \mathbf{n}^{(p)}(i), \quad p = 1, 2, \dots, P, \quad i = 1, 2, \dots, MN. \quad (5.25)$$

where $\mathbf{z}^{(p)}(i)$ can be seen as the received signal on the i^{th} spatio-temporal snapshot from the p^{th} path. From (5.25), the covariance of $\mathbf{z}^{(p)}(i)$, $E \{ \mathbf{z}^{(p)}(i) \mathbf{z}^{(p)}(i)^H \}$ is given by :

$$\mathbf{R}_{\mathbf{z}^{(p)}} = \varepsilon_p^2 \mathbf{d}(\tau_p) \mathbf{d}(\tau_p)^H + \frac{\sigma^2}{P} \mathbf{I}_L, \quad (5.26)$$

with $\{\varepsilon_p^2\}_{p=1}^P$ being the diagonal elements of \mathbf{R}_J and $\mathbf{n}_i(k)$ is an arbitrary decomposition of the estimation error (i.e., $\mathcal{N}_i = \sum_{p=1}^P \mathbf{n}^{(p)}(i)$). From (5.25), any column, $\widehat{\mathbf{z}}_i$, of \mathbf{Z}_i can be written as a function of the complete data as follows :

$$\mathbf{z}_i = \sum_{p=1}^P \mathbf{z}^{(p)}(i). \quad (5.27)$$

Now, we are in a position to describe the Expectation-step (*E*-step) and the Maximization-step (*M*-step) of the EM algorithm. The *E*-step consists in finding the conditional expectations of the sample covariance matrices $\{\widehat{\mathbf{R}}_{\mathbf{z}^{(p)}}\}_{p=1}^P$ of the complete data defined as :

$$\widehat{\mathbf{R}}_{\mathbf{z}^{(p)}} = \frac{1}{MN} \sum_{i=1}^{MN} \mathbf{z}^{(p)}(i) (\mathbf{z}^{(p)}(i))^H. \quad (5.28)$$

Given $\mathbf{R}_J^{\{q-1\}}$ and $\boldsymbol{\tau}^{\{q-1\}}$ (the previous estimates of \mathbf{R}_J and $\boldsymbol{\tau}$ at iteration $(q-1)$) and $\widehat{\mathbf{R}}_{\mathbf{Z}}$, the expectation of $\widehat{\mathbf{R}}_{\mathbf{z}^{(p)}}$ can be computed from the classical formulas of the conditional expectation with Gaussian distributed random vectors to yield :

$$\begin{aligned} \widehat{\mathbf{R}}_{\mathbf{z}^{(p)}}^{\{q\}} &= E \left\{ \widehat{\mathbf{R}}_{\mathbf{z}^{(p)}} | \widehat{\mathbf{R}}_{\mathbf{Z}}; \mathbf{R}_J^{\{q-1\}}, \boldsymbol{\tau}^{\{q-1\}} \right\} \\ &= \mathbf{R}_{\mathbf{z}^{(p)}}^{\{q\}} \left(\mathbf{R}_{\mathbf{Z}}^{\{q\}} \right)^{-1} \widehat{\mathbf{R}}_{\mathbf{Z}} \left(\mathbf{R}_{\mathbf{Z}}^{\{q\}} \right)^{-1} \mathbf{R}_{\mathbf{z}^{(p)}}^{\{q\}} \mathbf{R}_{\mathbf{z}^{(p)}}^{\{q\}} - \mathbf{R}_{\mathbf{z}^{(p)}}^{\{q\}} \left(\mathbf{R}_{\mathbf{Z}}^{\{q\}} \right)^{-1} \mathbf{R}_{\mathbf{z}^{(p)}}^{\{q\}}, \end{aligned} \quad (5.29)$$

where the matrices $\mathbf{R}_{\mathbf{Z}}^{\{q\}}$ are computed at each iteration from the estimates $\tau_p^{\{q-1\}}$ and $\varepsilon_p^{2\{q-1\}}$ computed in the previous iteration. Now, turning to the estimation of $\mathbf{R}_{\mathbf{Z}}^{\{q\}}$, the procedure is different from the one used in previous EM algorithms in [9] and [19] to estimate the covariance matrix. In fact, in [9] and [19], the covariance matrix of the received signal is simply diagonal, which is not the case in our work. Therefore, we resort to another approach to estimate the covariance matrix $\mathbf{R}_{\mathbf{Z}}^{\{q\}}$. First, suppose that $\mathbf{R}_{\mathbf{Z}}^{\{q\}}$ is a Toeplitz matrix⁴ (case of stationary processes). We adopt the method proposed in [21] (briefly detailed next) to the estimation of Toeplitz covariance matrices, which is also based on the EM algorithm making it well suited to our algorithm. Then, we define the $N_s \times N_s$ circulant extended version of $\mathbf{R}_{\mathbf{Z}}$, denoted as \mathbf{R}_s . The matrix \mathbf{R}_s represents the covariance matrix of the extended vectors $\{\tilde{\mathbf{Z}}_i\}_{i=1}^{MN}$, where $\tilde{\mathbf{Z}}_i$ consists of the vector \mathbf{Z}_i augmented by $(N_s - L)$ -dimensional null vectors. The covariance matrix \mathbf{R}_s is characterized by its eigenvalues as follows :

$$\mathbf{R}_s = \mathbf{F}^H \mathbf{R}_C \mathbf{F}, \quad (5.30)$$

where \mathbf{F} is the standard $N_s \times N_s$ discrete Fourier transform (DFT) matrix and \mathbf{R}_C is a diagonal matrix constructed from the eigenvalues of \mathbf{R}_s . The DFT transform of $\tilde{\mathbf{Z}}_i$ results in the rotated vectors $\mathbf{C}_i = \mathbf{F} \tilde{\mathbf{Z}}_i$ for $i = 1, 2, \dots, MN$. Denoting by $\widehat{\mathbf{R}}_C$ the estimate of \mathbf{R}_C using $\{\mathbf{C}_i\}_{i=1}^M$ (i.e., $\widehat{\mathbf{R}}_C = 1/MN \sum_{i=1}^{MN} \mathbf{C}_i \mathbf{C}_i^H$), the expectation of $\widehat{\mathbf{R}}_C$ conditioned on $\mathbf{R}_{\mathbf{Z}}$ and $\widehat{\mathbf{R}}_{\mathbf{Z}}$ — applying the same formula used to find (5.29) — is given by :

$$E(\widehat{\mathbf{R}}_C | \mathbf{R}_{\mathbf{Z}}, \widehat{\mathbf{R}}_{\mathbf{Z}}) = \mathbf{R}_{\mathbf{Z}C} \left(\mathbf{R}_{\mathbf{Z}}^{-1} \widehat{\mathbf{R}}_{\mathbf{Z}} (\mathbf{R}_{\mathbf{Z}}^{-1})^H - \mathbf{R}_{\mathbf{Z}}^{-1} \right) \mathbf{R}_{\mathbf{Z}C} + \mathbf{R}_C, \quad (5.31)$$

where $\mathbf{R}_{\mathbf{Z}C}$ is the cross-covariance of \mathbf{C}_i and \mathbf{Z}_i . Noting that $\mathbf{Z}_i = \tilde{\mathbf{F}}^H \mathbf{C}_i$, with $\tilde{\mathbf{F}} = \mathbf{F} [\mathbf{I}_L \mathbf{0}]^T$ and $\mathbf{0}$ is the $(L \times N_s - L)$ null matrix, the cross-covariance matrix $\mathbf{R}_{\mathbf{Z}C}$ is equal

⁴This assumption implies that the covariance between \mathbf{Z}_m and \mathbf{Z}_n depends only on the difference between m and n , corresponding therefore to stationary processes.

to $\mathbf{R}_C \tilde{\mathbf{F}}$. Then the estimate of \mathbf{R}_C at iteration q is given by :

$$\mathbf{R}_C^{\{q\}} = \text{diag} \left(\mathbf{R}_C^{\{q-1\}} \tilde{\mathbf{F}} \left((\mathbf{R}_Z^{\{q-1\}})^{-1} \hat{\mathbf{R}}_Z (\mathbf{R}_Z^{\{q-1\}})^{-1} - (\mathbf{R}_Z^{\{q-1\}})^{-1} \right) \tilde{\mathbf{F}}^H \mathbf{R}_C^{\{q-1\}} + \mathbf{R}_C^{\{q-1\}} \right) \quad (5.32)$$

and \mathbf{R}_Z is obtained using the transformation $\mathbf{Z}_i = \tilde{\mathbf{F}}^H \mathbf{C}_i$ as follows :

$$\mathbf{R}_Z^{\{q\}} = \tilde{\mathbf{F}}^H \mathbf{R}_C^{\{q\}} \tilde{\mathbf{F}}. \quad (5.33)$$

Finally, it has been shown in [21] that the stable point of (5.33) is equal to the maximum likelihood estimate of \mathbf{R}_Z .

During the M -step of the EM algorithm, we aim to maximize the log-likelihood function of the complete-data with respect to the parameters of interest $\{\tau_i\}_{i=1}^P$. It is the same objective function given in (5.12), with the true expectation of the complete data replaced by the conditional expectation of $\mathbf{z}^{(p)}(i)$; in other words \mathbf{R}_Z substituted by $\mathbf{R}_{z^{(p)}}$ and $\hat{\mathbf{R}}_Z$ by $\hat{\mathbf{R}}_{z^{(p)}}^{\{q\}}$. Thus we obtain the log-likelihood function of the complete-data, $L_p(\tau_p, \mathbf{R}_{z^{(p)}})$, as follows :

$$L_p(\tau_p, \mathbf{R}_{z^{(p)}}) = -\ln(\det(\mathbf{R}_{z^{(p)}})) - \text{trace} \left\{ \hat{\mathbf{R}}_{z^{(p)}}^{\{q\}} \mathbf{R}_{z^{(p)}}^{-1} \right\}. \quad (5.34)$$

Then, at iteration q , the estimate $\tau_p^{\{q\}}$ and $\mathbf{R}_{z^{(p)}}$ are those which jointly maximize $L_p(\tau_p, \mathbf{R}_{z^{(p)}})$. Using the eigen-decomposition⁵ of $\mathbf{R}_{z^{(p)}}$, the log-likelihood function of the complete-data can be expressed as :

$$L_p(\tau_p, \mathbf{R}_{z^{(p)}}) = -\ln \left(\varepsilon_p^2 + \frac{\sigma^2}{P} \right) - (M-1) \ln \left(\frac{\sigma^2}{P} \right) - \left(\frac{1}{\varepsilon_p^2 + \frac{\sigma^2}{P}} - \frac{P}{\sigma^2} \right) \mathbf{d}(\tau_p)^H \hat{\mathbf{R}}_{z^{(p)}}^{\{q\}} \mathbf{d}(\tau_p) \frac{P}{\sigma^2} \text{trace} \left(\hat{\mathbf{R}}_{z^{(p)}}^{\{q\}} \right). \quad (5.35)$$

which emphasizes the dependence of $L_p(\tau_p, \mathbf{R}_{z^{(p)}})$ on τ_p and ε_p^2 . The closed-form expression of its maximum with respect to ε_p^2 , for a given $\tau_p^{\{q\}}$, is :

$$\varepsilon_p^{2\{q\}} = \mathbf{d}(\tau_p^{\{q\}})^H \hat{\mathbf{R}}_{z^{(p)}}^{\{q\}} \mathbf{d}(\tau_p^{\{q\}}) - \frac{\sigma^2}{P}. \quad (5.36)$$

Now, injecting this expression in (5.35) yields the following one-dimensional maximization problem :

$$\tau_p^{\{q\}} = \arg \max_{\tau_p} -\ln \left\{ \mathbf{d}(\tau_p)^H \hat{\mathbf{R}}_{z^{(p)}}^{\{q\}} \mathbf{d}(\tau_p) \right\} \frac{P}{\sigma^2} \mathbf{d}(\tau_p)^H \hat{\mathbf{R}}_{z^{(p)}}^{\{q\}} \mathbf{d}(\tau_p), \quad (5.37)$$

or simply the problem of maximizing $\mathbf{d}(\tau_p)^H \hat{\mathbf{R}}_{z^{(p)}}^{\{q\}} \mathbf{d}(\tau_p)$. This follows immediately from the fact that the function $f(x) = -\ln(x) + P/\sigma^2 x$ is monotonic.

So far, the ML estimate has been found in an iterative way. However, this method needs an initial guess of the parameters from which the algorithm starts operating. Alternatively, to avoid all initialization hurdles and issues, we develop in the next section a non-iterative algorithm to find the ML estimates without grid search based on importance sampling.

⁵Noting that the matrix $\varepsilon_p^2 \mathbf{d}(\tau_p) \mathbf{d}(\tau_p)^H$ is of rank one with $L-1$ null eigenvalues and one equal to ε_p^2 , the eigen-decomposition of $\mathbf{R}_{z^{(p)}}$ can be easily done.

5.5 The Importance Sampling Technique

Similar to the above algorithm, we start from the expression of the log-likelihood function in (5.12). A direct maximization of this function imposes joint maximization over τ and \mathbf{R}_J . Therefore, it will be of interest to formulate an objective function that depends on the time delay only. To that end, we first maximize the likelihood function with respect to the nuisance parameters \mathbf{R}_J . For this purpose, it can be shown that the value of \mathbf{R}_J that maximizes $L(\tau, \mathbf{R}_J)$ for a fixed vector τ is :

$$\widehat{\mathbf{R}}_J^{ML} = (\mathcal{D}^H(\tau)\mathcal{D}(\tau))^{-1} \mathcal{D}^H(\tau) \widehat{\mathbf{R}}_Z \mathcal{D}(\tau) (\mathcal{D}^H(\tau)\mathcal{D}(\tau))^{-1} - \sigma^2 (\mathcal{D}^H(\tau)\mathcal{D}(\tau))^{-1}. \quad (5.38)$$

Then, injecting $\widehat{\mathbf{R}}_J^{ML}$ in (5.12) yields the so-called compressed likelihood function of the system :

$$L_c(\tau) = \frac{1}{\sigma^2} \text{trace} \left(\mathbf{\Pi} \widehat{\mathbf{R}}_Z \right) - \ln \left(\det \left(\mathbf{\Pi} \widehat{\mathbf{R}}_Z \mathbf{\Pi} + \sigma^2 (\mathbf{I}_L - \mathbf{\Pi}) \right) \right), \quad (5.39)$$

where $\mathbf{\Pi}$, introduced in (5.20), is defined as $\mathbf{\Pi} = \mathcal{D}(\tau) (\mathcal{D}^H(\tau)\mathcal{D}(\tau))^{-1} \mathcal{D}^H(\tau)$. Now, the maximum likelihood estimates of the time delay are obtained by maximizing the obtained compressed likelihood function (again, the most obvious optimization technique that naturally comes to mind is to perform a P -dimensional grid search, whose complexity increases with the number of delays) $L_c(\tau)$ with respect to τ . In this section, as an alternative to the iterative method already presented in section 5.4, we implement here the ML criterion in a non-iterative way. We resort to the global maximization theorem of Pincus [10] in order to find the global maximum of the multi-dimensional function at hand. In fact, according to [10], the global maximum of $L_c(\tau)$ with respect to τ is given by :

$$\widehat{\tau}_p = \lim_{\rho \rightarrow \infty} \frac{\int_J \dots \int_J \tau_p \exp \{ \rho L_c(\tau) \} d\tau}{\int_J \dots \int_J \exp \{ \rho L_c(\tau) \} d\tau}, \quad (5.40)$$

with $J = [0, T]$ being the interval in which the unknown delays are supposed to be confined. Clearly, as ρ tends to infinity, the fraction $\frac{\int_J \dots \int_J \tau_p \exp \{ \rho L_c(\tau) \} d\tau}{\int_J \dots \int_J \exp \{ \rho L_c(\tau) \} d\tau}$ becomes a multidimensional Dirac function, centered at the global maximum of $L_c(\cdot)$. Therefore, if we define the pseudo-pdf $L'_{c,\rho}(\cdot)$ as :

$$L'_{c,\rho}(\tau) = \frac{\exp \{ \rho L_c(\tau) \}}{\int_J \dots \int_J \exp \{ \rho L_c(\tau) \} d\tau}, \quad (5.41)$$

the ML estimate of $\{\tau_p\}_{p=1}^P$, obtained by applying (5.40), can be reformulated as :

$$\widehat{\tau}_p = \int_J \dots \int_J \tau_p L'_{c,\rho_0}(\tau) d\tau, \quad i = 1, 2, \dots, P, \quad (5.42)$$

where ρ_0 is a sufficiently large number (whose optimal value is discussed later). The function $L'_{c,\rho_0}(\cdot)$ is called pseudo-pdf since it has all the properties of a pdf, although τ is not

truly a random variable. We note from (5.42) that the ML estimate requires the evaluation of the multi-dimensional integral, which is usually difficult to perform in practice. However, exploiting the fact that $L'_{c,\rho_0}(\cdot)$ is a pseudo-pdf, the involved integral can be simply interpreted as the mean value of τ_p , when the hole vector τ is distributed according to $L'_{c,\rho_0}(\cdot)$. Therefore, one can easily evaluate this mean — and hence the integrals in (5.42) — in order to obtain $\hat{\tau} = [\hat{\tau}_1, \hat{\tau}_2, \dots, \hat{\tau}_P]^T$ using Monte Carlo techniques [11]:

$$\hat{\tau} = \frac{1}{R} \sum_{k=1}^R \tau_k, \quad (5.43)$$

where $\{\tau_k\}_{k=1}^R$ are R realizations of τ , with τ being distributed according to $L'_{c,\rho_0}(\cdot)$. But another problem arises here: how to jointly generate $\{\tau_p\}_{p=1}^P$ for a multidimensional random variable. Actually, $L'_{c,\rho_0}(\cdot)$ is constructed using the actual compressed likelihood function in (5.39), which is a multi-dimensional function; making the generation of the vector τ a very difficult task if not impossible. Therefore, it is of interest to find another pseudo-pdf to generate the realizations instead of using $L'_{c,\rho_0}(\cdot)$. To do so, we resort to the concept of IS as detailed below.

First, we mention that IS is a powerful Monte Carlo technique [23] which allows generating realizations using another distribution that is simpler than the actual one. Through the IS technique, the generated samples are weighted and averaged in a judicious manner to obtain the desired estimates. This efficient weighting operation improves considerably the performance achieved by the IS method compared to other Monte Carlo techniques. The IS approach is based on the following simple observation:

$$\int_J \dots \int_J f(\tau) L'_{c,\rho_0}(\tau) d\tau = \int_J \dots \int_J f(\tau) \frac{L'_{c,\rho_0}(\tau)}{g'(\tau)} g'(\tau) d\tau, \quad (5.44)$$

where $g'(\cdot)$ is another pseudo-pdf called normalized importance function (IF), whose choice is discussed later and $f(\cdot)$ is any given parameter transformation. Now, the problem is recast as the computation of the expectation of $f(\tau) \frac{L'_{c,\rho_0}(\tau)}{g'(\tau)}$ with respect to the distribution $g'(\cdot)$; which is simply performed via Monte Carlo methods as follows:

$$\int_J \dots \int_J f(\tau) \frac{L'_{c,\rho_0}(\tau)}{g'(\tau)} g'(\tau) d\tau \approx \frac{1}{R} \sum_{k=1}^R f(\tau_k) \frac{L'_{c,\rho_0}(\tau_k)}{g'(\tau_k)}, \quad (5.45)$$

in which the realizations $\{\tau_k\}_{k=1}^R$ are now generated according to $g'(\cdot)$. Yet, a great attention should be given to the choice of $g'(\cdot)$. In fact, the accuracy of this method depends on the similarity of the shapes of $L'_{c,\rho_0}(\cdot)$ and $g'(\cdot)$. In the best cases, the global maxima of $L'_{c,\rho_0}(\cdot)$ and $g'(\cdot)$ are the same. Still $L'_{c,\rho_0}(\cdot)$ is a complicated function of τ and $g'(\cdot)$ must be as simple as possible to easily generate the required realizations. Therefore, some trade-offs must be found in the construction of the importance function. Moreover, an appropriate choice of $g'(\cdot)$ reduces the number R of realizations since the generated values

will appear as if they were generated according to the original pseudo-pdf $L'_{c,\rho_0}(\cdot)$ when $g'(\cdot)$ is faithful to $L'_{c,\rho_0}(\cdot)$. Next, we discuss the appropriate choice of $g'(\cdot)$.

First, as an alternative to the actual multidimensional compressed likelihood function, the importance function should be a separable function in the different delays, $\{\tau_p\}_{p=1}^P$, to reduce the generation of a P -dimensional random vector to the generation of P scalar random variables. We therefore simplify the expression of the compressed likelihood function $L_c(\cdot)$ to find $g'(\cdot)$. Indeed, it is seen from (5.39) that $L_c(\cdot)$ involves the sum of two independent terms. These two terms can be written as functions of the eigenvalues of $\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\mathbf{\Pi}$ as follows :

$$\begin{aligned} \ln\left(\det\left(\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\mathbf{\Pi} + \sigma^2(\mathbf{I}_L - \mathbf{\Pi})\right)\right) &= \ln\left((\sigma^2)^{L-P} \prod_{p=1}^P \lambda_p\right) \\ &= \sum_{i=1}^P \ln\left(\frac{\lambda_i}{\sigma^2}\right) + L \ln \sigma^2, \end{aligned} \quad (5.46)$$

and

$$\frac{1}{\sigma^2} \text{trace}\left(\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\right) = \frac{1}{\sigma^2} \text{trace}\left(\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\mathbf{\Pi}\right) = \sum_{p=1}^P \frac{\lambda_p}{\sigma^2}, \quad (5.47)$$

where $\lambda_1, \lambda_2, \dots, \lambda_P$ are the eigenvalues of $\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\mathbf{\Pi}$. Clearly, the term $\sum_{p=1}^P \frac{\lambda_p}{\sigma^2}$ is dominant compared to the term $\sum_{i=p}^P \ln\left(\frac{\lambda_p}{\sigma^2}\right)$. Consequently, in (5.39) we drop the term $L \ln \sigma^2$, independent on the delays, in $\ln\left(\det\left(\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\mathbf{\Pi} + \sigma^2(\mathbf{I}_L - \mathbf{\Pi})\right)\right)$ and it is reasonable neglect the term $\ln\left(\det\left(\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\mathbf{\Pi} + \sigma^2(\mathbf{I}_L - \mathbf{\Pi})\right)\right)$ with respect to $\frac{1}{\sigma^2} \text{trace}\left(\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\right)$. Moreover, one can approximate the matrix $\mathcal{D}^H(\tau)\mathcal{D}(\tau)$ by the diagonal matrix $\left(\sum_{l=0}^{L-1} |c_l|^2\right) \mathbf{I}_P$ to avoid the computation of the inverse involved in $\mathbf{\Pi}$. This approximation is well justified since the off-diagonal terms of the matrix $\mathcal{D}^H(\tau)\mathcal{D}(\tau)$ are negligible compared to its diagonal elements (see the Appendix for further details). Using this assumption, the term $\frac{1}{\sigma^2} \text{trace}\left(\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\right)$ is approximated by :

$$\frac{1}{\sigma^2} \text{trace}\left(\mathbf{\Pi}\hat{\mathbf{R}}_{\mathbf{Z}}\right) \approx \frac{1}{\sigma^2 \left(\sum_{l=0}^{L-1} |c_l|^2\right)} \text{trace}\left(\mathcal{D}^H(\tau)\hat{\mathbf{R}}_{\mathbf{Z}}\mathcal{D}(\tau)\right). \quad (5.48)$$

Lastly, considering all these observations, an approximation of the actual compressed likelihood function, $\bar{L}(\cdot)$, with unnecessary terms discarded, is given by :

$$\begin{aligned} \bar{L}(\tau) &= \frac{1}{\sigma^2} \text{trace}\left(\mathcal{D}(\tau)\mathcal{D}^H(\tau)\hat{\mathbf{R}}_{\mathbf{Z}}\right) \\ &= \frac{1}{MN\sigma^2} \sum_{k=1}^{MN} \sum_{p=1}^P \left| \sum_{l=1}^L c_{l-1} \exp\left\{-\frac{j2\pi(l-1)\tau_p}{L}\right\} \mathbf{z}(k, l) \right|^2 \\ &= \sum_{p=1}^P I(\tau_p), \end{aligned} \quad (5.49)$$

where

$$I(\tau) = \frac{1}{MN\sigma^2} \sum_{k=1}^{MN} \left| \sum_{l=1}^L c_{l-1} \exp \left\{ -\frac{j2\pi(l-1)\tau}{L} \right\} \mathbf{Z}(k, l) \right|, \quad (5.50)$$

can be evaluated using the Fast Fourier Transform (FFT). Hence, the normalized IF is selected as follows :

$$g'_{\rho_1}(\boldsymbol{\tau}) = \frac{\prod_{p=1}^P \exp \{ \rho_1 I(\tau_p) \}}{\left(\int_{\mathcal{J}} \exp \{ \rho_1 I(\tau) \} d\tau \right)^P}, \quad (5.51)$$

which is the product of P elementary functions, each of which depending on the delay of a given single path. Here, we succeed in making the different delays separable and distributed according to the same pdf $p(\cdot)$ given by :

$$p(\tau) = \frac{\exp \{ \rho_1 I(\tau) \}}{\int_{\mathcal{J}} \exp \{ \rho_1 I(\tau) \} d\tau}. \quad (5.52)$$

Hence, the joint pdf of the delays in $g'_{\rho_1}(\cdot)$ is split into the product of P individual pdfs, which transposes the problem of generating a P -dimensional random variable to the generation of P one-dimensional random variables according to a simpler common distribution. Moreover, the constant term ρ_1 in (5.51) and (5.52) is different from ρ_0 since it is more advantageous to use two different values as explained later. We mention here that the estimation performance depends on these two parameters. Actually, the pdf $p(\cdot)$ in (5.52) exhibits P lobes centered at the location of the true time delay. But the estimation error \mathcal{E}_n makes other undesired lobes appear which in turn biases the generated values not faithful (spurious values) to the true delays. For this reason, ρ_1 is increased to make the pdf $p(\cdot)$ more peaked around the actual delays $\{\tau_p\}_{p=1}^P$ so that the undesired lobes disappear. However, very large values of ρ_1 may also destroy some useful lobes and hence their corresponding delays will not be generated. Therefore, the optimal value of ρ_1 is the highest one for which the pdf $p(\cdot)$ still exhibits at least P main lobes. Moreover, one should keep in mind that the normalized IF in (5.51) is built upon an approximation of the actual compressed likelihood function, which we aim to maximize. Consequently, a bias will always appear in the mean of the values generated according to the normalized IF. Fortunately, this bias is alleviated by the weighting factor $L'_{c, \rho_0}(\cdot)/g'(\cdot)$ introduced by the concept of IS. Therefore, we maximize the contribution of the compressed likelihood in the weighting factor rather than its approximation by making ρ_0 higher than ρ_1 . Thus, an appropriate choice of these two parameters reduces the number, R , of required realizations and ultimately the computational complexity.

To summarize, the new IS-based ML estimator is given by :

$$\hat{\tau}_p = \frac{1}{R} \sum_{k=1}^R \tau_k(p) \frac{L'_{c, \rho_0}(\tau_k)}{g'_{\rho_1}(\tau_k)}, \quad (5.53)$$

where $\tau_k(p)$ is the p^{th} element of the vector τ_k . We also mention that in practice the delays are confined in the interval $[0, LT_c]$ [6]. This allows us to further use the *circular mean* instead of the *linear mean* given by (5.53), as detailed below.

In few words, to define the concept of the *circular mean*, consider a random variable X taking values in $[0, 1]$ according to a given distribution $G(\cdot)$. The circular mean of X is given by [24] :

$$E_c\{X\} = \frac{1}{2\pi} \angle \int_0^1 e^{j2\pi x} G(x) dx, \quad (5.54)$$

where the operator $\angle(\cdot)$ returns the argument of any complex number. Then, if we have a set of R realizations, x_1, \dots, x_R , drawn according to the pdf $G(\cdot)$, the circular mean in (5.54) is computed as :

$$E_c\{X\} = \frac{1}{2\pi} \angle \frac{1}{R} \sum_{k=1}^R e^{j2\pi x_k}. \quad (5.55)$$

Adopting the concept of *circular mean* in our problem, an alternative formulation of the estimate in (5.53) is given by :

$$\hat{\tau}_p = \frac{LT_c}{2\pi} \angle \frac{1}{R} \sum_{k=1}^R F(\tau_k) \exp \left\{ \frac{j2\pi \tau_k(p)}{LT_c} \right\}, \quad (5.56)$$

where the delays are transposed to the interval $[0, 1]$ after being normalized by LT_c and $F(\cdot)$ is the weighting factor defined as :

$$F(\tau) = \frac{L'_{c,\rho_0}(\tau)}{g'(\tau)}. \quad (5.57)$$

Note that the estimator in (5.56) relies on finding the angles of a complex number. Therefore, we no longer need to compute the two positive real normalization factors $\int_J \dots \int_J \exp \{ \rho L_c(\tau) \} d\tau$ and $(\int_J \exp \{ \rho_1 I(\tau) \} d\tau)^P$ since they can be dropped without ultimately affecting the final result. Moreover, during the computation of the weighting factor $F(\cdot)$, the exponential terms in the numerator and the denominator of $F(\cdot)$ may result in an overflow. To avoid this overflow, we substitute $F(\cdot)$ by $F'(\cdot)$ given by :

$$F'(\tau_k) = \exp \left\{ \rho_0 L_c(\tau_k) - \rho_1 \sum_{p=1}^P I(\tau_k(p)) - \max_{1 \leq l \leq R} \left(\rho_0 L_c(\tau_l) - \rho_1 \sum_{p=1}^P I(\tau_l(p)) \right) \right\} \quad (5.58)$$

by multiplying $F(\cdot)$ by a positive number. In such a way, the exponential argument in (5.58) no longer exceeds zero, alleviating thereby any computation difficulty.

Summary of steps

In the following, we summarize the entire steps of the new IS-based ML time delay estimator.

1. From the samples matrix, \mathcal{Z} , compute the periodogram $I(\cdot)$ expressed in (5.50) at discrete points of the interval $[0, LT_c]$. Then evaluate the elementary pdf as follows :

$$p(\tau_i) = \frac{\exp(\rho_1 I(\tau_i))}{\sum_{k=1}^K \exp(\rho_1 I(\tau_k))}, \quad i = 1, 2, \dots, K \quad (5.59)$$

where we substitute the integral in the denominator by a summation over all the discrete points in the integration interval.

2. Generate one realization of the vector $\boldsymbol{\tau}$ according to the pseudo-pdf $g'_{\rho_1}(\cdot)$. To simplify, we exploit the fact that the delays are separable in $g'_{\rho_1}(\cdot)$ and we generate P realizations $\{\tau_k(p)\}_{p=1}^P$ according to $p(\cdot)$ using the inverse probability integration [20]. It is important to make sure that the P generated entries $\tau_k(1), \tau_k(2), \dots, \tau_k(P)$ (in order to obtain one vector realization) are different. This condition is necessary since the delays of the paths are in practice different thereby ensuring that the matrix inverse $(\mathcal{D}^H(\boldsymbol{\tau})\mathcal{D}(\boldsymbol{\tau}))^{-1}$ in $L_c(\cdot)$ always exists.
3. Repeat *step 2*) $R - 1$ times then evaluate the weighting factors $F'(\tau_i)$ for $i = 1, \dots, R$.
4. Find the maximum likelihood estimate of the delays using the *circular mean* in (5.56).

5.6 Extension To MC-CDMA Systems

In MC-CDMA transmitter, the original data are spread over different subcarriers using a spreading code. Therefore, it is possible to transmit several DS-SS waveforms in parallel. At time index n , the input information is first converted into $N_c = 2K + 1$ parallel sequences and modulated at rate $1/T_{MC}$, where $T_{MC} = N_c T$ is the symbol duration after serial/parallel conversion. Each of the parallel stream is then spreaded with a spreading code at rate $1/T_c$ and modulated by the inverse discrete Fourier transform (IDFT).

At the receiver, a reformulation of the post-correlation model for MC-CDMA of the spatio-temporal observation for the k^{th} subcarrier and the n^{th} observation is given by [26] :

$$\mathbf{Z}_{k,n} = s_{k,n} \mathbf{J}_{k,n} \mathbf{D}_k^T(\boldsymbol{\tau}) + \mathbf{N}_{k,n}, \quad (5.60)$$

where $s_{k,n}$ and $\mathbf{J}_{k,n}$ are the signal component and the spatial response matrix on the k^{th} subcarrier, respectively. The column of the time response matrix $\mathbf{D}_k(\boldsymbol{\tau}) = [\mathbf{d}_{k,1}, \mathbf{d}_{k,2}, \dots, \mathbf{d}_{k,P}]$ are given by :

$$\mathbf{d}_{k,p} = e^{-j2\pi k \lambda \frac{\tau_p}{T_{MC}}} [\rho_c(-\tau_p), \rho_c(T_c/k_s - \tau_p) e^{j2\pi k \frac{\lambda}{L k_s}}, \dots, \rho_c((L k_s - 1)T_c/k_s - \tau_p) e^{j2\pi k \frac{\lambda(L k_s - 1)}{L k_s}}]^T, \quad (5.61)$$

where λ determines the frequency spacing between two adjacent subcarrier⁶ ($f_k = \lambda k/T_{MC}$) and k_s is the oversampling ratio [26]. Note here that the propagation time-delays are supposed to be the same for all subcarriers. Unlike the single-carrier case, the received samples $Z_{k,n}$ cannot be directly used as an input to the algorithms. Therefore, we introduce the intermediate transformation of the samples matrix, denoted $Z_{k,n}^c$, given by :

$$\begin{aligned} Z_{k,n}^c &= Z_{k,n} * (\mathbf{a} \mathbf{1}_M^T) \\ &= s_{k,n} \mathbf{J}_{k,n} \mathbf{D}_k^{cT}(\boldsymbol{\tau}) + \mathbf{N}_{k,n}^c, \end{aligned} \quad (5.62)$$

where $\mathbf{1}_M = [1, \dots, 1]^T$ and $\mathbf{a} = [1, e^{-j2\pi\frac{\lambda}{Lk_s}}, \dots, e^{-j2\pi\frac{\lambda(Lk_s-1)}{Lk_s}}]^T$. The p^{th} column of $\mathbf{D}_k^c(\boldsymbol{\tau})$ is :

$$\begin{aligned} \mathbf{d}_{k,p}^c &= \mathbf{d}_{k,p} * \mathbf{a} \\ &= e^{-j2\pi k\lambda\frac{\tau_p}{T_{MC}}} [\rho_c(-\tau_p), \rho_c(T_c/k_s - \tau_p), \dots, \rho_c((Lk_s - 1)T_c/k_s - \tau_p)]^T. \end{aligned} \quad (5.63)$$

Hence we eliminate in $\mathbf{D}_k^c(\boldsymbol{\tau})$ the dependence row-wise of the phase slope of each column vector on k in the spectral domain. While the formulation in (5.62) seems to be adapted for the estimation process, the phase shift $e^{-j2\pi k\lambda\frac{\tau_p}{T_{MC}}}$ on each column of $\mathbf{D}_k^c(\boldsymbol{\tau})$ set against the direct use of the formulation in (5.62). To overcome this problem, we note that the matrix $\mathbf{D}_k^c(\boldsymbol{\tau})$ can be written as $\mathbf{D}_k^c(\boldsymbol{\tau}) = \mathbf{A}_k \mathbf{D}(\boldsymbol{\tau})$ where \mathbf{A}_k is a diagonal matrix which diagonal elements are $\left\{ e^{-j2\pi k\lambda\frac{\tau_p}{T_{MC}}} \right\}_{p=1}^P$. Then we insert the phase shift in the spatial-response matrix $\mathbf{J}_{k,n}$ to obtain a formulation similar to the one in (5.3). Finally, to exploit the frequency gain, we gather the transformed observation over the different subcarrier into the following compact representation :

$$\begin{aligned} \mathbf{Z}_n^c &= [\mathbf{Z}_{1,n}^{cT}, \mathbf{Z}_{2,n}^{cT}, \dots, \mathbf{Z}_{N_c,n}^{cT}] \\ &= \mathbf{D}(\boldsymbol{\tau}) \mathbf{J}_n^{cT} + \mathbf{N}_n^{cT}, \end{aligned} \quad (5.64)$$

where $\mathbf{J}_n^{cT} = [\mathbf{A}_1 \mathbf{J}_{1,n}^T, \mathbf{A}_2 \mathbf{J}_{2,n}^T, \dots, \mathbf{A}_{N_c} \mathbf{J}_{N_c,n}^T]$ and $\mathbf{N}_n^{cT} = [\mathbf{N}_{1,n}^{cT}, \mathbf{N}_{2,n}^{cT}, \dots, \mathbf{N}_{N_c,n}^{cT}]$. The interesting thing with the formulation in (5.64) is that it increases the number of observations proportionally to the number of subcarrier used in the system. Considering N transmitted symbols, we formulate a compact representation similar to (5.4) by concatenating the N observed symbols :

$$\begin{aligned} \mathbf{Z} &= [\mathbf{Z}_1^c, \mathbf{Z}_2^c, \dots, \mathbf{Z}_N^c] \\ &= \mathbf{D}(\boldsymbol{\tau}) \mathbf{J}^{cT} + \mathbf{N}^{cT}, \end{aligned} \quad (5.65)$$

⁶When λ is equal to 1, the transiver belongs to the class of multitone (MT)-CDMA, and if it is equal to L , the transiver belongs to the family of MC-DS-CDMA.

where $\mathbf{J}^{cT} = [\mathbf{J}_1^{cT}, \dots, \mathbf{J}_N^{cT}]$ and $\mathbf{N}^{cT} = [\mathbf{N}_1^{cT}, \dots, \mathbf{N}_N^{cT}]$.

Compared to a schema that estimates the delays over each subcarrier separately, the proposed model presents better performance. In addition, we are able to derive the corresponding CRLB following the same steps as presented in Section II. This reveals that the CRLB has a similar expression as in (5.20) by multiplying by a factor of $1/N_c$. If we denote by CRLB_0 the CRLB when $M = N = N_c = 1$, the resulting CRLB for the MC-CDMA is given by :

$$\text{CRLB} = \frac{1}{MNN_c} \text{CRLB}_0. \quad (5.66)$$

As a consequence, the time delay estimation merges space, time and frequency.

To obtain an expression of the CRLB by using the channel coefficients matrix to estimate the delays, some modifications to the model are needed and developed in Appendix III. In this case, we see that the resulting CRLB depends on both time and frequency correlation.

5.7 Simulation Results

In this section, we compare the performance of the two proposed maximum likelihood estimators against the popular Root-MUSIC algorithm and the CRLB. In all the simulations, we consider a multipath propagation environment with 3 propagations paths and we simulate a challenging scenario of closely-spaced delays equal to $0.12T$, $0.15T$ and $0.18T$. The mean square error (MSE) — used as performance measure — of the three estimators is compared to the CRLB. First, recall that the EM algorithm is iterative in nature ; hence initialization is a critical issue. Therefore, the initial values for this estimator are selected as random variables, centered at the real time delay and with a variance of $0.05T$. The processing gain is fixed at $L = 64$ and the optimal values of the parameters ρ_0 and ρ_1 for the IS-based technique are equal to 20 and 10, respectively. We also assume that the power is equally distributed between the three paths on average.

First, we consider a single carrier transceiver with $M = 4$ antenna branches at the receiver and one received samples ($N = 1$) and we compare the MSE of the two proposed ML algorithms to those of the Root-MUSIC in Fig. 5.1. We also plot the performance of the EM ML when the initial values have a variance of $0.18T$ reflecting less accurate initializations. Clearly, both the IS-based and the EM algorithms, for good initializations, outperform the Root-MUSIC technique over a large range of the SNR. While the two ML algorithms present almost the same performance, it is suggested in practice to use, in this configuration, the EM ML approach since it offers less computational complexity compared to the IS-based algorithm. Indeed, the EM ML estimator has the advantage of performing P parallel maximizations. Therefore, as the number of paths P increases, there is no additional noticeable computational time cost. On the other side, IS-based algorithm can guarantee robustness to the initial estimates, contrarily to the EM ML.

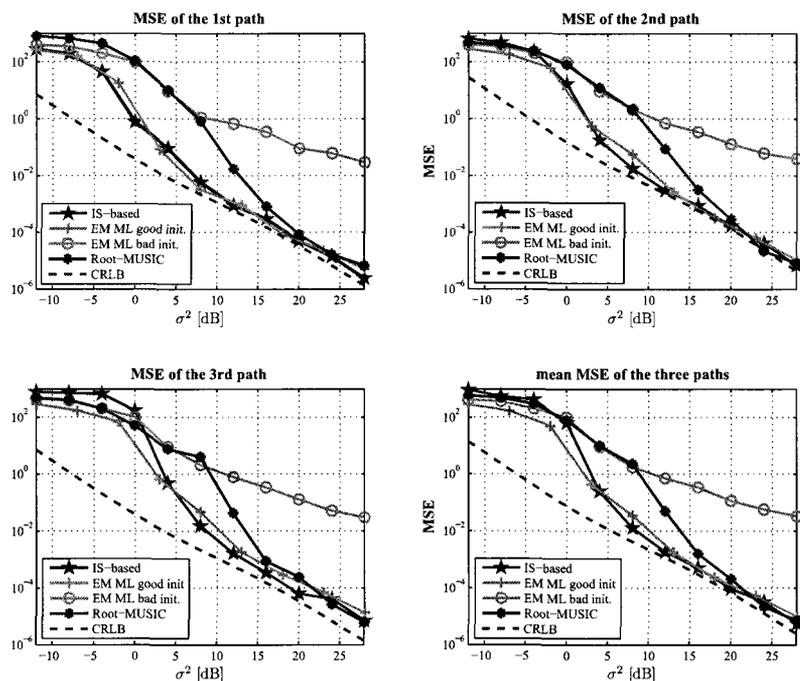


FIGURE 5.1 – Estimation performance of the IS-based, the EM-based and the Root-MUSIC algorithms for closely-separated delays, $M = 4$.

So far, all the methods exhibit good performance, with remarkable improvements for the two new ML estimators. However, a quick study of the EM algorithm and the Root-MUSIC reveals that they are based on an estimate of the covariance matrix of the received signal from the columns of the matrix \mathcal{Z} , and the accuracy of this estimate depends on the number of antennas (which plays the role of the number of samples). Therefore, we simulate the performance of all these algorithms considering only one receiving antenna element and keep the other simulation conditions the same as in Fig. 5.1. The results are shown in Fig. 5.2. Clearly, Root-MUSIC estimator is very sensitive to the number of receiving antenna branches. Its performance degrades considerably compared to the previous case. It fails completely in estimating the delays which, indeed, is due to the poor estimate of $\mathbf{R}_{\mathcal{Z}}$. On the other hand, the EM and the IS-based algorithms are less affected, in this challenging scenario. They still provide good estimates regardless of the challenging operating conditions based on short data snapshots. Fig. 5.2 actually suggests that the IS-based algorithm is even more robust than the EM ML estimator in this configuration.

To further investigate this issue, we fix the SNR value at 10 dB and vary the number of antenna branches M from 1 to 8 with $N = 1$. The MSE of the three algorithms versus M is plotted in Fig. 5.3. As expected, the IS-based methods attain the CRLB starting from a small value of M contrarily to the Root-MUSIC algorithm and the EM ML. This means that the IS-based estimator is well geared toward situations of reduced antenna array

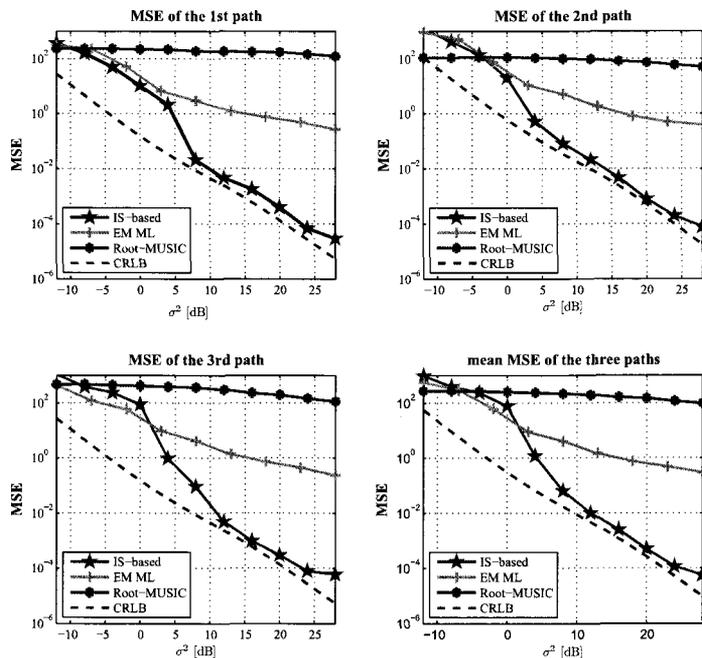


FIGURE 5.2 – Estimation performance of the IS-based, the EM-based and the Root-MUSIC for closely-separated delays, $M = 1$.

sizes. On the other hand, we plot in Fig. 5.4 the MSE versus N , considering one antenna branches in the receiver. Note here that to see the effect of the time channel variation, we use the channel estimate matrix instead of the received signal to estimate the delays. More details about the formulation used in the estimation process are presented in Appendix 2. Clearly, the performance of the three estimators is the same starting from $N = 5$. But it is not suitable to increase N since the values of the delays may change from one symbol to another. Usually, a tracking technique (as the one developed in [26]) is used to continue estimating the delays that is why we prefer to use small number of N in the estimation. In the region of small numbers of snapshots (N is less than 5), the ML-based methods perform better than the Root-MUSIC algorithm and the gap between these methods increases as N decreases.

To evaluate the impact of the frequency gain in the multicarrier systems, we plot in Fig. 5.5 the MSE versus the number of subcarriers N_c . We fix M at 1 and N at 1 to better illustrate the influence of the number of subcarriers N_c on the estimation performance. As N_c increases, the estimation performance improves to saturate for high value of N_c . This salutation can be explained by the increase of inter-carrier interference with N_c , due to the loss of orthogonality between subcarrier in a multipath environment. We should also mention the similarity between Fig. 5.3 - Fig. 5.5 which prove that the three dimensions time, space and frequency have the same impact on the estimation performance of the algorithms. This conclusion is further verified by the expression of the CRLB in (5.66)

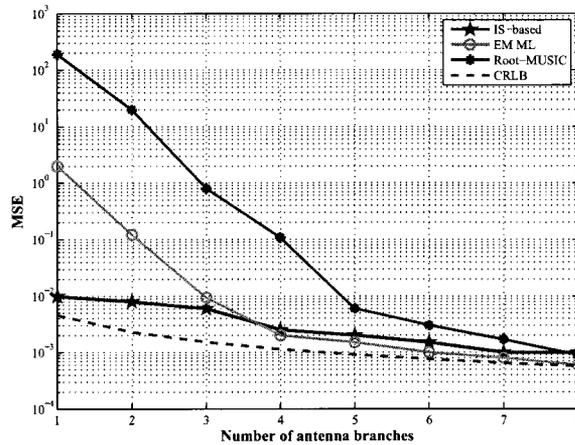


FIGURE 5.3 – MSE vs. number of antenna branches M for $N = 1$ symbol, $K = 1$ subcarrier, and $\text{SNR} = 10$ dB.

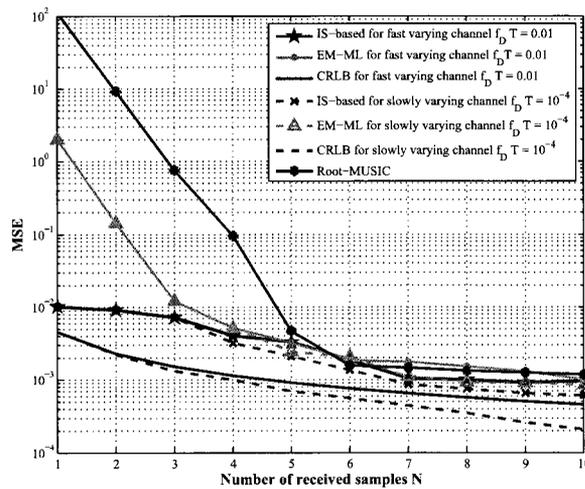


FIGURE 5.4 – MSE vs. number of received symbols N for $M = 1$ antenna branch, $K = 1$ subcarrier, and $\text{SNR} = 10$ dB.

which is inversely proportional to the number of receiving antenna, the number of symbols and the number of subcarriers.

5.8 Conclusion

In this paper, we developed two implementations of the ML criterion for the estimation of time delay both SC and MC air interface in multipath environment. We distinguish two estimations process : either using directly the received signal or using the channel estimate matrix. In the first option we show analytically, through the CRLB, and by simulations that

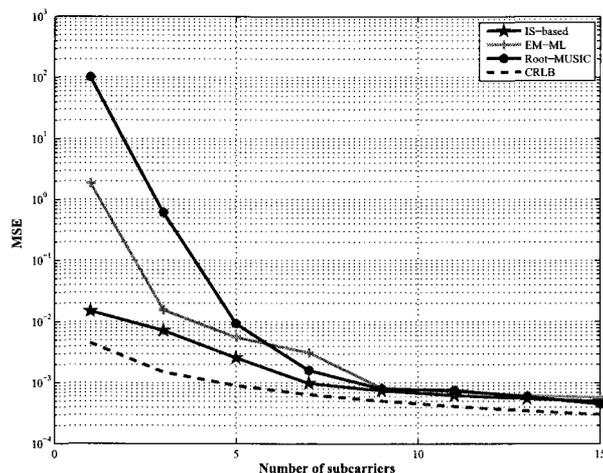


FIGURE 5.5 – MSE vs. number of subcarrier with $M = 1$ and $N = 1$.

the three dimensions : time, frequency and space have the same effect of the estimation performance. Whereas basing on the channel estimate matrix, we exploit the time and frequency correlation. While the two proposed methods are an implementation of the same criterion, each one has its attractive advantages. We also derived a closed-form expression for the corresponding CRLB in the context of DS-CDMA. The first estimator relies on the iterative EM algorithm with a moderate computation cost compared to the grid search technique since it transforms the problem of a multidimensional search into parallel easy searches over one-dimensional spaces. Compared to other eigen-based methods such as the Root-MUSIC, the EM approach exhibits better performance with a relatively good initialization, which is an important issue for this algorithm that affects its estimation performance.

The other algorithm is based on an entirely different approach. It relies on a global maximization theorem and the concept of importance sampling to directly find the global maximum. The IS-based approach also avoids the multidimensional grid search by approximating the actual compressed likelihood function ; splitting it thereby into separable one-dimensional functions of the delays. It does not require any initialization and hence it does not suffer from performance degradation. Its performance is almost equal to that of the EM algorithm (which requires good initialization), but at the expense of an increase in the computational burden. Moreover, only the IS-based algorithm produces accurate estimates for a small number of receiving antenna branches. The performance of the EM approach also degrades considerably in the specific case of single antenna (SISO configuration) where the Root-MUSIC algorithm fails completely to estimate the delays.

Appendix 1

Justification of the approximation $\mathcal{D}^H(\boldsymbol{\tau})\mathcal{D}(\boldsymbol{\tau}) \approx \sum_{l=0}^{L-1} |c_l|^2 \mathbf{I}_p$

The diagonal and off-diagonal elements of $\mathcal{D}^H(\boldsymbol{\tau})\mathcal{D}(\boldsymbol{\tau})$ are respectively given by :

$$[\mathcal{D}^H(\boldsymbol{\tau})\mathcal{D}(\boldsymbol{\tau})]_{m;m} = \sum_{l=0}^{L-1} |c_l|^2, \quad m = 1, 2, \dots, P, \quad (5.67)$$

$$[\mathcal{D}^H(\boldsymbol{\tau})\mathcal{D}(\boldsymbol{\tau})]_{m;n} = \sum_{l=0}^{L-1} |c_l|^2 \exp \left\{ \frac{j2\pi l(\tau_m - \tau_n)}{L} \right\},$$

$m, n = 1, 2, \dots, P, m \neq n.$ (5.68)

While it is easy to verify that :

$$[\mathcal{D}^H(\boldsymbol{\tau})\mathcal{D}(\boldsymbol{\tau})]_{m;n} < [\mathcal{D}^H(\boldsymbol{\tau})\mathcal{D}(\boldsymbol{\tau})]_{m;m}, \quad (5.69)$$

for $m \neq n$, the inequality in (5.69) does not guarantee that the diagonal elements are indeed dominant compared to the off-diagonal elements. To that end, we define $F(\cdot)$ the ratio between (5.68) and (5.69) as follows :

$$F(\Delta\tau_{m;n}) = \frac{\sum_{l=0}^{L-1} |c_l|^2 \exp \left\{ \frac{j2\pi l \Delta\tau_{m;n}}{L} \right\}}{\sum_{l=0}^{L-1} |c_l|^2}, \quad (5.70)$$

where $\Delta\tau_{m;n} = \tau_m - \tau_n$ is considered as a random variable uniformly distributed in $[-LT_c, LT_c]$. We plot in Fig. 5.6 the probability of having $F(\Delta\tau_{m;n}) \geq x$ for $x \in [0, 1]$ and we verify that the diagonal elements of $\mathcal{D}^H(\boldsymbol{\tau})\mathcal{D}(\boldsymbol{\tau})$ are dominant, with very high probability, compared to its off-diagonal elements. This justifies the following approximation :

$$\mathcal{D}^H(\boldsymbol{\tau})\mathcal{D}(\boldsymbol{\tau}) \approx \sum_{l=0}^{L-1} |c_l|^2 \mathbf{I}_p. \quad (5.71)$$

Appendix 2

Model used to estimate from the channel coefficients matrix

The development presented in the main body of the article is based on the observation matrix \mathbf{Z} . In this case, the columns of \mathbf{Z} are uncorrelated because of the presence of

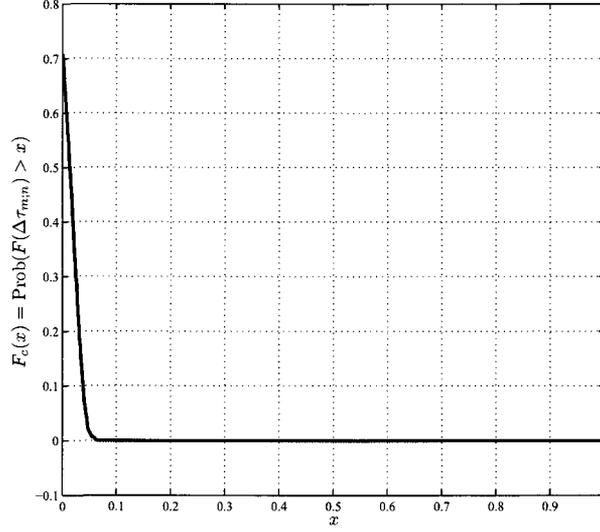


FIGURE 5.6 – Complementary cumulative distribution function of the ratio $F(\Delta\tau_{m;n})$.

uncorrelated transmitted symbols. But if we went to use the channel coefficients matrix $\widehat{\mathbf{H}} = [\widehat{\mathbf{H}}_1^T, \widehat{\mathbf{H}}_2^T, \dots, \widehat{\mathbf{H}}_N^T]$, these symbols are no longer presents and the columns of $\widehat{\mathbf{H}}$ are correlated. So we introduce a small modification on the formulation to remain the two algorithms valid. This problem can be solved if viewed in the proper way. First, we perform a column-by-column FFT of $\{\widehat{\mathbf{H}}_n\}_{n=1}^N$ to obtain :

$$\begin{aligned}\widehat{\mathbf{H}}_n &= \mathcal{D}(\tau)\mathbf{J}_n^T + \mathcal{E}_n \\ &= \mathcal{D}(\tau)[\mathbf{g}_n(1), \mathbf{g}_n(2), \dots, \mathbf{g}_n(M)] + \mathcal{E}_n,\end{aligned}\quad (5.72)$$

where the matrix \mathcal{E}_n is the resulting noise and $\mathbf{g}_n(i)$ is the i^{th} column of \mathbf{J}_n^T . From the formulation in (5.72), we bring together all the channel coefficient in one matrix $\widehat{\mathcal{H}}$ defined as :

$$\widehat{\mathcal{H}} = \mathcal{D}_{\text{modif}}[\mathbf{g}(1), \mathbf{g}(2), \dots, \mathbf{g}(M)] + \mathcal{E}, \quad (5.73)$$

in which $\mathbf{g}(i) = [\mathbf{g}_1(i)^T, \mathbf{g}_2(i)^T, \dots, \mathbf{g}_N(i)^T]^T$ and $\mathcal{D}_{\text{modif}} = \mathbf{I}_N \otimes \mathcal{D}$, where the operator \otimes stands for the Kronecker product. If we consider that the signal is transmitted through a Rayleigh channel and we denote the maximum Doppler frequency by f_D , the covariance matrix of $\mathbf{g}(i)$ is given by :

$$\mathbf{R}_g = \begin{pmatrix} J_0(0) & \dots & J_0(2\pi(N-1)f_D T) \\ \vdots & \ddots & \vdots \\ J_0(2\pi(N-1)f_D T) & \dots & J_0(0) \end{pmatrix} \otimes \mathbf{R}_J \quad (5.74)$$

Taking into account (5.73) and (5.74), the implementation of the two algorithms using the channel information matrix is straightforward.

Appendix 3

CRLB for MC-CDMA using the channel coefficients matrix

We use a similar developments as presented in Appendix II, with the difference that here we deal with time and frequency correlation. We define $\widehat{\mathcal{H}}_{k,n}^c$ as the column by column FFT of the channel response matrix $\widehat{\mathbf{H}}_{k,n}^c$ for the k^{th} subcarrier and n^{th} observation which is given by :

$$\begin{aligned}\widehat{\mathcal{H}}_{k,n}^c &= \mathcal{D}(\boldsymbol{\tau})\mathbf{J}_{k,n}^{cT} + \boldsymbol{\varepsilon}_{k,n} \\ &= \mathcal{D}(\boldsymbol{\tau})[\mathbf{g}_{k,n}(1), \mathbf{g}_{k,n}(2), \dots, \mathbf{g}_{k,n}(M)] + \boldsymbol{\varepsilon}_{k,n},\end{aligned}\quad (5.75)$$

where $\mathbf{g}_{k,n}(i)$ is the i^{th} column of $\mathbf{J}_{k,n}^{cT}$. Then we gather the channel coefficients in the matrix $\widehat{\mathcal{H}}^c$ given by :

$$\widehat{\mathcal{H}}^c = \overline{\mathcal{D}}[\overline{\mathbf{g}}(1), \overline{\mathbf{g}}(2), \dots, \overline{\mathbf{g}}(M)],\quad (5.76)$$

with $\overline{\mathbf{g}}(i) = [\mathbf{g}_{1,1}^T(i), \dots, \mathbf{g}_{M_c,1}^T(i), \mathbf{g}_{1,2}^T(i), \dots, \mathbf{g}_{M_c,2}^T(i), \dots, \mathbf{g}_{1,N}^T(i), \dots, \mathbf{g}_{M_c,N}^T(i)]$ and $\overline{\mathcal{D}} = \mathbf{I}_N \otimes \mathbf{I}_{M_c} \otimes \mathcal{D}$. Denote by $\phi(\Delta f, \Delta t)$ the autocorrelation of the channel transfer function (we suppose here uncorrelated scattering where the autocorrelation transfer function in frequency is a function of only the frequency difference Δf [27], the covariance matrix of $\mathbf{g}^c(i)$ is $\mathbf{R}_{g^c} = \Phi \otimes \mathbf{R}_J$ where the elements of Φ are function of $\phi(\Delta f, \Delta t)$. Injecting \mathbf{R}_{g^c} and $\overline{\mathcal{D}}$ in (5.20), the CRLB can be written in this alternative way :

$$\begin{aligned}(\text{CRLB}^{-1}(\boldsymbol{\tau}))_{i,j} &= \frac{2M}{\sigma^2} \Re \left\{ \text{trace} \left(\overline{\mathcal{D}}_j^H (1 - \mathbf{I}_N \otimes \mathbf{I}_{M_c} \otimes \Pi) \overline{\mathcal{D}}_i \mathbf{R}_{g^c} \overline{\mathcal{D}}^H \mathbf{R}_{\mathcal{H}^c}^{-1} \overline{\mathcal{D}} \mathbf{R}_{g^c} \right) \right\} \\ &= \frac{2M}{\sigma^2} \Re \left\{ \text{trace} \left(\mathbf{I}_N \otimes \mathbf{I}_{M_c} \otimes (\mathcal{D}_j^H (1 - \Pi) \mathcal{D}_i) \mathbf{R}_{g^c} \overline{\mathcal{D}}^H \mathbf{R}_{\mathcal{H}^c}^{-1} \overline{\mathcal{D}} \mathbf{R}_{g^c} \right) \right\},\end{aligned}\quad (5.77)$$

where $\mathbf{R}_{\mathcal{H}^c}$ is the covariance of \mathcal{H}^c and $\overline{\mathcal{D}}_i$ and \mathcal{D}_i are the derivative of $\overline{\mathcal{D}}$ and \mathcal{D} with respect to τ_i , respectively. If we denote by \mathbf{B}_k the k^{th} $P \times P$ block on the diagonal of $\mathbf{R}_{g^c} \overline{\mathcal{D}} \mathbf{R}_{\mathcal{H}^c}^{-1} \overline{\mathcal{D}} \mathbf{R}_{g^c}$, we get :

$$(\text{CRLB}^{-1}(\boldsymbol{\tau}))_{i,j} = \frac{2M}{\sigma^2} \Re \left\{ \sum_{k=1}^{NM_c} (\mathbf{u}_j^H (1 - \Pi) \mathbf{u}_i) \mathbf{B}_k(i, j) \right\},\quad (5.78)$$

then we obtain the following expression of the CRLB :

$$\text{CRLB}(\boldsymbol{\tau}) = \frac{\sigma^2}{2M} \left[\Re \left\{ \sum_{k=1}^{NM_c} \mathbf{U}^H (1 - \Pi) \mathbf{U} * \mathbf{B}_k \right\} \right]^{-1}.\quad (5.79)$$

Under certain conditions, Φ can be easily expressed. In fact, $\phi(\Delta f, \Delta t)$ is given by [27] :

$$\phi(\Delta f, \Delta t) = \int_{-\infty}^{+\infty} \phi_c(\tau, \Delta t) e^{-j2\pi\Delta f\tau} d\tau, \quad (5.80)$$

where $\phi_c(\tau, \Delta t)$ is the autocorrelation function of the channel impulse response. If the different paths have a Rayleigh distribution, $\phi_c(\tau, \Delta t)$ can be written as :

$$\begin{aligned} \phi_c(\tau, \Delta t) &= \int_{-\infty}^{+\infty} P(\tau) J_0(2\pi f_D \Delta t) e^{-j2\pi\Delta f\tau} d\tau \\ &= J_0(2\pi f_D \Delta t) F_f(\Delta f). \end{aligned} \quad (5.81)$$

Then $\Phi = \mathbf{J} \otimes \mathbf{F}_f$ where $F_f(i, j) = F_f((i - j)\lambda/T_{MC})$.

We verify that in the absence of time and frequency correlation, the expression of the CRLB obtained in (5.79) is the same as in (5.66).

Bibliographie

- [1] K. Cheikhrouhou, S. Affes and P. Mermelstein, "Impact of synchronization on performance of enhanced array-receivers in wideband CDMA networks", *IEEE J. Select. Areas in Comm.*, vol. 19, no 12, pp. 2462-2476, Dec. 2001.
- [2] I. Shakya, F. Ali, and E. Stipidis, "High User Capacity Collaborative CDMA," *IET Commun.*, vol. 5, no. 3, pp. 307-319, Feb. 2011.
- [3] B. Suard, A. Naguib, G. Xu and A. Paulraj, "Performance of a wireless MC-CDMA system with an antenna array in a fading channel : reverse link," *IEEE Trans. on commun.*, vol. 48, pp. 1257 -1261, Aug. 2000.
- [4] B. H. Khalaj, A. Paulraj and T. Kailath, "2D RAKE receivers for CDMA cellular systems," *Proc. GLOBECOM'94*, pp. 400-404, Nov. 1994.
- [5] "Performance analysis of multicarrier DS-SS-CDMA with antenna array in a fading channel," *Wireless Pers. Commun.* vol. 52, pp. 273-287, 2010.
- [6] S. Affes, P. Mermelstein, "A new receiver structure for asynchronous CDMA : STAR—the spatio-temporal array-receiver", *IEEE J. Select. Areas in Comm.*, vol. 16, no. 8, pp. 1411-1422, Oct. 1998.
- [7] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276-280, Mar. 1986.
- [8] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Sign. Process.*, vol. 37, no. 7, pp. 984-995, Jul. 1989.
- [9] M. Feder and E. Weinstein "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, Sign. Process.*, vol. 36, pp. 477-489, Apr. 1988.
- [10] M. Pincus, "A closed form solution for certain programming problems," *Oper. Research*, pp. 690-694, 1962.
- [11] A. F. M. Smith and A. Gelfand, "Bayesian statistics without tears : A sampling-resampling framework," *Amer. Statist.*, vol. 46, pp. 84-88, May 1992.

- [12] S. Saha and S. Kay, "An exact maximum likelihood narrowband direction of arrival estimator," *IEEE Trans. Sign. Process.*, vol. 56, pp. 1082-1092, Oct. 2008.
- [13] H. Wang and S. Kay, "A maximum likelihood angle-doppler estimator using importance sampling," *IEEE Trans. Aerospace and Electro. Syst.*, vol. 46, pp. 610-622, Apr. 2010.
- [14] A. Masmoudi, F. Bellili, S. Affes and A. Stéphenne, "A non-data-aided maximum likelihood time delay estimator using importance sampling," *IEEE Trans. Sign. Process.*, vol. 59, pp. 4505-4515, Oct. 2011.
- [15] S. Affes and P. Mermelstein, "Adaptive space-time processing for wireless CDMA," in *Adaptive Signal Processing : Application to Real-World Promblems*, J. Benesty and A.H. Huang, Eds., Springer, Berlin, Jan. 2003.
- [16] A. D. Dempster, N. M. Laird and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc.*, vol. B-39, pp. 1-37, 1977.
- [17] Z. Q. Hou and Z. D. Wu, "A new method for resolution estimation of time delay," *Proc. ICASSP*, pp. 420-423, 1982.
- [18] M. A. Hasan, M. R. Azimi-Sadjadi and G. J. Dobeck, "Separation of multiple time delays using new spectral estimation schemes," *IEEE Trans. Sign. Process.*, vol. 46, no. 6, pp. 1580-1590, June 1998.
- [19] M. I. Miller and D. R. Fuhrmann "Maximum-likelihood narrow-band direction finding and the EM algorithm," *IEEE Trans. Acoust., Speech, Sign. Process.*, vol. 38, pp. 1560-1577, Sept. 1990.
- [20] S. Kay, *Intuitive Probability and Random Processes Using MATLAB*, New York : Springer, 2005.
- [21] M. I. Miller and D. L. Snyder, "The role of likelihood and entropy in incomplete-data problems : Applications to estimating point-process intensities and Toeplitz constrained covariances," *Proceedings of the IEEE* , vol. 75, pp. 892-907, July 1987.
- [22] A. G. Jaffer, "Maximum likelihood direction finding for stochastic sources : A separable solution," in *Proc. IEEE Int. Conf. Acoust., Speech, Sign. Process.*, pp. 2893-2896, Apr. 1988.
- [23] W. B. Bishop and P. M. Djuric, "Model order selection of damped sinusoids in noise by predictive densities," *IEEE Trans. Sign. Process.*, vol. 44, pp. 611-619, Mar. 1996.
- [24] K. V. Mardia, *Statistics of Directional Data*, New York : Academic, 1972.
- [25] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. on Acoustics, Speech and Sign. Process.*, pp. 1783-1795, Oct. 1990.

- [26] B. Smida, S. Affes, L. Jun, and P. Mermelstein, "A spectrum-efficient multicarrier CDMA array-receiver with diversity-based enhanced time and frequency synchronization," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2315-2327, June. 2007.
- [27] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2001.

Conclusion

Dans ce mémoire, le problème de synchronisation temporelle en communication numérique a été traité. Nous avons développé des procédures d'estimation du délai de propagation basées sur le critère de maximum de vraisemblance pour divers systèmes de communication. Le premier estimateur est développé pour les signaux modulés où les symboles émis sont inconnus. La méthode "importance sampling" (IS) est adaptée pour trouver l'estimé à maximum de vraisemblance. Bien qu'il soit développé sous l'hypothèse d'un seul trajet de propagation, il trouve des applications dans plusieurs systèmes tels que les communications satellitaires. Une extension au cas multi-trajets est aussi proposée. Dans cette configuration, nous nous retrouvons face à plusieurs délais à estimer. Bien que la fonction de vraisemblance soit multidimensionnelle, l'IS offre une procédure attirante pour transformer le problème multidimensionnel en un problème unidimensionnel. Autre contribution, nous avons développé deux algorithmes de synchronisation pour les systèmes CDMA mono-porteuse et multi-porteuses. Le premier algorithme reprend le principe de l'IS au niveau du signal après désétalement. L'autre algorithme se base sur la méthode "expectation maximization" (EM) qui transforme le problème multidimensionnel en de simples opérations unidimensionnelles dont le nombre augmente linéairement avec le nombre de trajets détectés. Bien que les deux méthodes soient des implémentations du même critère, chacune possède ses propres points forts. La méthode EM estime des délais de propagation avec une complexité relativement faible comparée à d'autres algorithmes puisque, tel que démontré dans ce mémoire, l'estimation des différents délais peut se faire en parallèle, ce qui réduit le temps de calcul. Les simulations montrent que cet algorithme présente de meilleure performance que les techniques de sous-espace tel que le Root-MUSIC et ceci en utilisant de bonnes initialisations, qui doit être souligné comme un point faible de l'algorithme. La performance de l'algorithme EM dégrade considérablement dans le cas d'une seule antenne réceptrice alors que l'algorithme Root-MUSIC échoue complètement à estimer les délais. D'un autre côté, l'algorithme IS ne nécessite aucune initialisation et offre de bonne performance même en présence d'une seule antenne réceptrice, et ceci au prix d'une complexité accrue. Nous montrons aussi par simulations que les dimensions espace, temps et fréquence (pour les systèmes multi-porteuses) ont le même effet sur les performances de ces estimateurs.

L'autre volet traité dans ce rapport est la dérivation des expressions analytiques des bornes

de Cramér-Rao des estimateurs non biaisés du retard pour les modulation BPSK, MSK et QAM carrées en estimation aveugle. Ces expressions analytiques révèlent que les performances d'estimation du délai ne dépendent pas de la valeur du paramètre en question, chose qu'on ne pouvait pas confirmer auparavant en évaluant les bornes de Cramér-Rao par des méthodes empiriques et que l'estimation du délai est indépendante de l'estimation de la phase et de la fréquence. On dit que le délai est découplé de ces deux derniers paramètres. Nous avons aussi dérivé les expressions de ces bornes pour les systèmes SC- et MC-DS-CDMA. Dans ce cas, nous constatons que les trois dimensions : spatiale, temporelle et fréquentielle, agissent de la même façon sur les performances d'estimation, se qui confirme les résultats obtenus par simulations.

Cependant, plusieurs extensions restent à explorer. Dans ce mémoire, nous avons toujours supposé que le bruit additif était blanc et Gaussien. Ce sera intéressant de voir les modifications à faire, que ce soit au niveau des estimateurs de des CRLBs, si le bruit est coloré. De plus, le canal de propagation dans les chapitres 1 et 3 est supposé constant. L'adaptation d'un canal variable reflète mieux la réalité du canal de transmission. Aussi, la CRLB est développée dans le chapitre 2 sous l'hypothèse d'un seul trajet de propagation. La dérivation de la CRLB dans un canal à multi-trajets reste un bon sujet de recherche. En ce qui concerne le 5em chapitre, nous avons supposé que les antennes sont décorréliées. L'évaluation de la robustesse des estimateurs développés dans ce chapitre en cas d'antennes corrélées reste à venir et le développement d'estimateurs qui tiennent compte de cette corrélation serait peu être d'actualité si jamais nous remarquons une dégradation de performance dans ce cas.

