QUATORZIEME COLLOQUE GRETSI - JUAN-LES-PINS - DU 13 AU 16 SEPTEMBRE 1993

ESTIMATION EN PRÉSENCE DE MODÈLES INCERTAINS: SÉLECTION DE FORMULATIONS ENSEMBLISTES

P. L. Combettes[†] *et T. J. Chaussalet*[‡]

[†]Department of Electrical Engineering, City College and Graduate School,
City University of New York, New York, NY 10031, USA.
[‡]Division of Mathematical and Decision Sciences,
School of Computer Science and Information Systems Engineering,
University of Westminster, London, W1M 8JS, UK.

RÉSUMÉ

Les problèmes d'estimation ensembliste avec modèles incertains sont souvent accompagnés d'informations probabilistes décrivant l'incertitude. Idéalement, on devrait sélectionner parmi ces informations celles dont les ensembles associés ont une intersection aussi petite que possible, tout en contenant la vraie solution avec un niveau de confiance raisonnable. Dans cet article, on analyse l'utilisation rationnelle des propriétés probabilistes de l'incertitude en vue de la sélection de formulations ensemblistes.

ABSTRACT

Set theoretic estimation problems with uncertain models are often accompanied with probabilistic information describing the uncertainty. Ideally, one should select those pieces of information whose associated sets have the smallest possible intersection, while containing the true solution to within some reasonnable confidence level. In this paper, we analyze the rational utilization of the probabilistic properties of the uncertainty with a view towards the selection of set theoretic formulations.

INTRODUCTION

Considérons le problème d'estimer un objet h appartenant à un espace Ξ. En estimation ponctuelle conventionnelle, on cherche à trouver un estimateur a de h qui soit optimal par rapport à un objectif donné, par exemple vraisemblance, entropie, erreur quadratique moyenne, ou coût bayésien. La solution du problème est alors souvent un point unique, dit optimal.

Dans cet article, on se placera dans le cadre de l'estimation ensembliste, dont les principes de base sont les suivants (on se rapportera à [2] pour un exposé détaillé des fondements et des applications en traitement du signal et en automatique). Soit $(\Psi_\iota)_{\iota\in I}$ la famille de propositions représentant toutes les informations disponibles sur le problème, qu'elles proviennent de connaissances à priori ou bien des observations. On construit alors les sous-ensembles de Ξ

$$(\forall \iota \in I)\ \ S_\iota = \{a \in \Xi \mid a \ \text{ satisfait } \ \Psi_\iota\}. \qquad (1)$$

$(\Xi, (S_\iota)_{\iota\in I})$ est la formulation ensembliste du problème et sa solution est l'ensemble des points qui s'accordent avec toute l'information connue, à savoir $S = \bigcap_{\iota\in I} S_\iota$.

En traitement du signal, les observations comprennent généralement une composante incertaine due, par exemple, aux erreurs de modélisation ou au bruit. Certaines propriétés probabilistes de cette composante sont souvent connues à priori et elles permettent de construire un grand nombre de sous-ensembles. Pour sélectionner les sous-ensembles à utiliser, on doit considérer trois facteurs. Tout d'abord, puisque la complexité des algorithmes de calcul de solutions ensemblistes augmente avec le nombre d'ensembles, il est souhaitable de conserver un nombre minimal d'ensembles. On disposera donc de ceux qui ne contribuent pas à la réduction de l'ensemble solution [2]. Le second objectif est d'obtenir l'ensemble solution de plus petite mensuration[1] (taille) possible afin de minimiser la dispersion des solutions ensemblistes et de garantir ainsi leur qualité. Si Ψ_ι est une propriété probabiliste alors l'appartenance de h à S_ι ne peut être garantie au delà d'un niveau de confiance donné. Le troisième objectif sera donc d'assurer la fiabilité des solutions ensemblistes en maintenant au dessus d'un certain seuil la probabilité que l'ensemble solution contient h. Cette probabilité s'appellera la justesse de la formulation ensembliste.

Le but de cet article est d'étudier la question de la sélection rationnelle de formulations ensemblistes en présence d'information probabilistes relatives à l'incertitude des modèles. On décrit tout d'abord les ensembles basés sur de telles informations. On introduit ensuite la notion de justesse et on étudie sa relation avec les niveaux de confiance individuels des sous-ensembles. Dans la section suivante, on s'intéresse au compromis entre justesse et mensuration. Les modèles à incertitude bornée sont traités dans la dernière section.

[1]Soit ν une fonction d'ensemble positive et monotone telle que $(\forall\iota\in I)\ \nu(S_\iota) > 0$. Alors $\nu(S)$ est la mensuration de S. [2].

ENSEMBLES BASÉS SUR LES PROPRIÉTÉS DE L'INCERTITUDE

Modèle. On se placera toujours sur un espace de probabilité sous-entendu $(\Omega, \mathcal{F}, \mathsf{P})$. L'objet à estimer, ou estimandum, est noté h et appartient à un espace vectoriel topologique Ξ. Le processus $(X_n)_{n\in\mathbb{Z}}$ des données observables est généré par le modèle

$$(\forall n \in \mathbb{Z})(\forall \omega \in \Omega)\ \ X_n(\omega) = T_n(h) + U_n(\omega), \tag{2}$$

où $(T_n)_{n\in\mathbb{Z}}$ est une suite connue de fonctionnelles sur Ξ représentant le système et $(U_n)_{n\in\mathbb{Z}}$ un processus aléatoire représentant l'incertitude sur le modèle. Ce processus peut représenter le bruit ambiant, comme dans [4], ou bien le bruit ambiant et la contribution d'une composante incertaine du système, comme dans [3]. Cette composante peut être induite par des erreurs de modélisation ou bien par des perturbations aléatoires affectant le système.

Méthode Générale de Construction. Soit Ψ_ι une propriété connue du processus d'incertitude $(U_n)_{n\in\mathbb{Z}}$. Pour chaque valeur estimée $a \in \Xi$, on peut définir un processus résiduel $(Y_n(a))_{n\in\mathbb{Z}} = (X_n - T_n(a))_{n\in\mathbb{Z}}$. Puisqu'en théorie $(Y_n(h))_{n\in\mathbb{Z}} = (U_n)_{n\in\mathbb{Z}}$, on peut alors construire le sous-ensemble

$$S_\iota = \{a \in \Xi \mid (Y_n(a))_{n\in\mathbb{Z}} \text{ satisfait } \Psi_\iota\}. \tag{3}$$

Il va de soi qu'un tel ensemble a peu de valeur pratique puisque seule une trajectoire finie $Y(a,\omega) = (X_n(\omega) - T_n(a))_{1\leq n\leq N}$ du processus résiduel est observée, pour une réalisation $\omega \in \Omega$. Soit $Q_\iota(a)$ une statistique de $(X_n - T_n(a))_{1\leq n\leq N}$ décrivant la propriété Ψ_ι. Alors, à partir de la loi de $Q_\iota(h)$, on peut déterminer un borélien r_ι tel que $\mathsf{P}\{\omega \in \Omega \mid Q_\iota(h,\omega) \in r_\iota\} \geq c_\iota$, où c_ι est le coefficient de confiance désiré. On substitue donc la propriété "$Q_\iota(h,\omega) \in r_\iota$" à la propriété exacte "$(Y_n(h))_{n\in\mathbb{Z}}$ satisfait Ψ_ι." Elle traduit le fait que h engendre une trajectoire résiduelle $Y(h,\omega)$ qui s'accorde avec Ψ_ι à un niveau de confiance c_ι. On aboutit ainsi au sous-ensemble

$$S_\iota(\omega) = \{a \in \Xi \mid Q_\iota(a,\omega) \in r_\iota\}, \tag{4}$$

dont le niveau de confiance est donné par

$$\mathsf{P}\{\omega \in \Omega \mid h \in S_\iota(\omega)\} \geq c_\iota. \tag{5}$$

Exemples. Donnons à présent quelques exemples d'ensembles (on renvoie le lecteur à [3] et [4] pour les détails). Le coefficient de confiance attaché à S_ι est c_ι.

Supposons tout d'abord que les variables aléatoires $(U_n)_{n\in\mathbb{Z}}$ sont identiquement distribuées. Si l'on connait deux réels λ_1 et λ_2 tels que $\mathsf{P}\{\omega\in\Omega \mid \lambda_1 \leq U_0(\omega) \leq \lambda_2\} \geq c_\iota$ alors, pour chaque $n \in \{1, \cdots, N\}$, on obtient le sous-ensemble

$$S_n(\omega) = \{a \in \Xi \mid \lambda_1 \leq X_n(\omega) - T_n(a) \leq \lambda_2\}. \tag{6}$$

Considérons maintenant les moments. On sait que si $(U_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^{2k}(\Omega, \mathcal{F}, \mathsf{P})$ est suffisamment mélangé, la statistique $Q_\iota(h) = (1/N)\sum_{n=1}^N (X_n - T_n(h))^k$ est asymptotiquement normale avec moyenne $\mathsf{E}U_0^k$ et variance $\sigma_k^2 = (\mathsf{E}U_0^{2k} - \mathsf{E}^2 U_0^k)/N$. On arrive donc au sous-ensemble

$$S_k(\omega) = \{a\in\Xi \mid \mathsf{E}U_0^k - \alpha\sigma_k \leq \sum_{n=1}^N (X_n(\omega) - T_n(a))^k/N \leq \mathsf{E}U_0^k + \alpha\sigma_k\}, \tag{7}$$

où α est tiré des tables de la loi normale en fonction de c_ι. En particulier, si le moment d'ordre k impair est nul, $S_k(\omega)$ devient

$$S_k(\omega) = \{a\in\Xi \mid \Big|\sum_{n=1}^N (X_n(\omega) - T_n(a))^k\Big| \leq \alpha\sqrt{N\mathsf{E}U_0^{2k}}\}. \tag{8}$$

De même, pour les moments absolus, on arrive au sous-ensemble

$$S_p(\omega) = \{a\in\Xi \mid \mathsf{E}|U_0|^p - \alpha\sigma_p \leq \sum_{n=1}^N |X_n(\omega) - T_n(a)|^p/N \leq \mathsf{E}|U_0|^p + \alpha\sigma_p\}, \tag{9}$$

où $\sigma_p^2 = (\mathsf{E}|U_0|^{2p} - \mathsf{E}^2|U_0|^p)/N$. Notons que pour construire (7)-(9) seule la connaissance de 2 moments est nécessaire.

Supposons maintenant que $(U_n)_{n\in\mathbb{Z}}$ est fortement mélangé et stationnaire, avec fonction de cumulant sommable aux ordres 2 et 4, $\mathsf{E}U_0 = 0$, et densité spectrale g connue à la fréquence $\nu_l \in]0, 1/2[$. La loi asymptotique en χ_2^2 du périodogramme normalisé qui résulte conduit au sous-ensemble

$$S_l(\omega) = \{a \in \Xi \mid \Big|\sum_{n=1}^N (X_n(\omega) - T_n(a))e^{-\imath 2\pi\nu_l n}\Big|^2 \leq -Ng(\nu_l)\ln(1 - c_l)/2\}. \tag{10}$$

Il convient de souligner que dans certains cas l'information sur $(U_n)_{n\in\mathbb{Z}}$ est suffisante pour construire tous ces sous-ensembles. Ainsi, si $(U_n)_{n\in\mathbb{Z}}$ est blanc et gaussien, de puissance $\mathsf{E}|U_0|^2$ connue, on peut littéralement obtenir une infinité d'ensembles puisque tous les moments sont calculables.

JUSTESSE DE L'ENSEMBLE SOLUTION

Une grandeur essentielle pour juger de la fiabilité d'une formulation ensembliste basée sur de l'information à caractère probabiliste est la probabilité que l'ensemble solution résultant contienne l'estimandum. On appellera cette grandeur la justesse de la formulation ensembliste.[2] Puisque les ensembles basés sur des informations déterministes n'ont aucune influence sur la justesse, considérons donc une formulation ensembliste $(\Xi, (S_\iota)_{\iota\in I})$, où les propositions $(\Psi_\iota)_{\iota\in I}$ se rapportent toutes au processus d'incertitude $(U_n)_{n\in\mathbb{Z}}$. On suppose que I est dénombrable.

[2] On rappelle qu'une formulation ensembliste $(\Xi, (S_\iota)_{\iota\in I})$ est dite juste si $h \in \bigcap_{\iota\in I} S_\iota$ [2].

L'ensemble asssocié à Ψ_ι pour une réalisation particulière $(X_n(\omega))_{1\leq n\leq N}$ des observations est défini par (4) et est doté du niveau de confiance (5). En outre, l'ensemble solution est $S(\omega) = \bigcap_{\iota\in I} S_\iota(\omega)$. Définissons les événements

$$\begin{aligned}(\forall \iota \in I)\ \ H_\iota &= \{\omega \in \Omega \mid Q_\iota(h,\omega) \in r_\iota\} \\ &= \{\omega \in \Omega \mid h \in S_\iota(\omega)\} \qquad (11)\end{aligned}$$

et

$$H = \{\omega \in \Omega \mid h \in S(\omega)\} = \bigcap_{\iota\in I} H_\iota. \qquad (12)$$

Alors, par définition, la justesse de $(\Xi, (S_\iota)_{\iota\in I})$ est $\mathsf{P}H$. On rappelle que $(\forall \iota \in I)\ \ \mathsf{P}H_\iota \geq c_\iota$. Le problème est de choisir les coefficients $(c_\iota)_{\iota\in I}$ en vue d'une valeur souhaitée de la justesse. Par exemple, dans les applications numériques, on a un nombre fini d'ensembles, disons $(S_\iota)_{1\leq\iota\leq m}$. Alors, si F désigne la fonction de distribution conjointe des statistiques associées $(Q_\iota)_{1\leq\iota\leq m}$, la justesse s'exprime

$$\mathsf{P}H = \int_{\times_{\iota=1}^{m} r_\iota} dF(q_1, \cdots, q_m). \qquad (13)$$

Cette formule exacte est souvent difficile à évaluer en raison de la complexité de F. Dans le cas général, la justesse peut être encadrée de la manière suivante

$$1 - \sum_{\iota\in I}(1 - c_\iota) \leq \mathsf{P}H \leq \min_{\iota\in I} \mathsf{P}H_\iota. \qquad (14)$$

L'inégalité de droite indique que la justesse de la formulation ensembliste ne peut excéder celle d'aucun sous-ensemble. L'inégalité de gauche donne une borne inférieure. Cette borne peut être améliorée en regroupant les statistiques mutuellement indépendantes (m.i.). En effet, soit $(J_k)_{k\in K}$ une famille de sous-ensembles finis et disjoints de I telle que pour chaque $k \in K$ les statistiques $(Q_\iota)_{\iota\in J_k}$ soient m.i. Posons

$$\begin{cases} (\forall k \in K)\ \ H^k = \cap_{\iota\in J_k} H_\iota, \\ H^- = \cap_{\iota\in I\setminus J} H_\iota, \ \text{ où } J = \cup_{k\in K} J_k. \end{cases} \qquad (15)$$

Alors, on a

$$\begin{aligned}\mathsf{P}H &= 1 - \mathsf{P} \bigcup_{k\in K} \complement H^k \bigcup \complement H^- \\ &\geq 1 - \sum_{k\in K}(1 - \mathsf{P}H^k) - \mathsf{P}\complement H^- \\ &\geq 1 - \sum_{k\in K}(1 - \prod_{\iota\in J_k} \mathsf{P}H_\iota) - \sum_{\iota\in I\setminus J}(1 - \mathsf{P}H_\iota) \\ &\geq 1 - \sum_{k\in K}(1 - \prod_{\iota\in J_k} c_\iota) - \sum_{\iota\in I\setminus J}(1 - c_\iota). \qquad (16)\end{aligned}$$

Bien sûr, pour $J = \emptyset$, on retrouve la borne inférieure de (14). A l'autre extrême, si I est fini et si toutes les statistiques $(Q_\iota)_{\iota\in I}$ sont m.i., alors $\mathsf{P}H = \prod_{\iota\in I} \mathsf{P}H_\iota \geq \prod_{\iota\in I} c_\iota$. Dans les exemples de la section précédente, des familles de sous-ensembles correspondant à des statistiques m.i. se présentent dans plusieurs cas. Ainsi on compte la famille $(S_n)_{1\leq n\leq N}$ de (6) si les variables aléatoires $(U_n)_{n\in\mathbb{Z}}$ sont m.i. (bruit blanc pur); la famille $(S_l)_{0<l<N/2}$ de (10) avec $\nu_l = l/N$ si $(U_n)_{n\in\mathbb{Z}}$ est un bruit blanc gaussien; certaines familles de moments pour des distributions particulières si $(U_n)_{n\in\mathbb{Z}}$ est un bruit blanc pur.

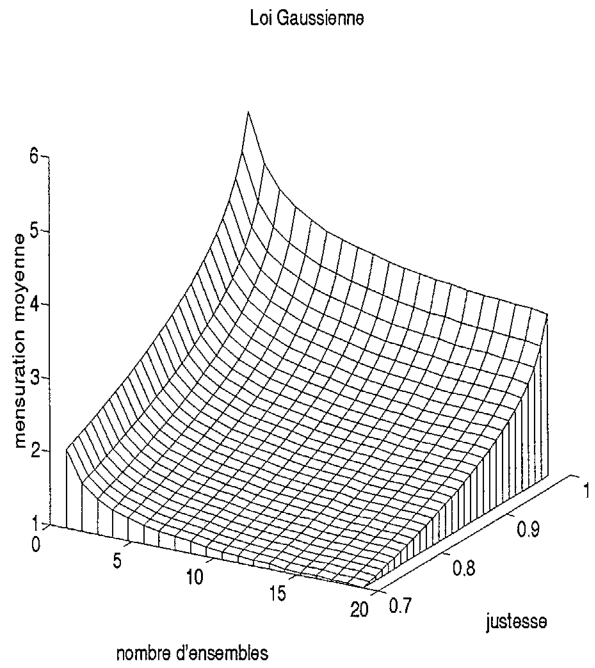

Figure 1: Mensuration Moyenne - Loi Gaussienne.

JUSTESSE ET MENSURATION

La justesse et la mensuration de S sont les deux grandeurs qui doivent guider la sélection d'une formulation ensembliste. Comme on l'a vu plus haut, la justesse représente la fiabilité de l'ensemble solution; la mensuration, quant à elle, donne sa résolution. L'objectif général est d'obtenir la plus petite mensuration avec le minimum d'ensembles pour une justesse donnée.

Une conception très répandue est qu'afin d'obtenir la meilleure solution ensembliste, il faut exploiter autant d'information que le permettent les connaissances à priori. Si toutes les informations disponibles $(\Psi_\iota)_{\iota\in I}$ sont de nature déterministe alors cette approche est correcte car seule la notion de mensuration intervient et on a donc intérêt à utiliser le plus grand nombre d'ensembles. Cependant, en présence d'informations probabilistes, les notions de justesse et de mensuration sont généralement antagonistes et une augmentation du nombre d'ensemble ne se traduit pas nécessairement par une meilleure mensuration pour une justesse donnée. En effet, d'après les résultats de la section précédente, augmenter la justesse revient à augmenter la confiance des sous-ensembles et donc à augmenter leur mensuration, ce qui aboutit à une mensuration plus grande de leur intersection. La question est donc de savoir si, pour obtenir la plus petite intersection, il faut mieux intersecter un grand nombre de grands ensembles ou bien un petit nombre de petits ensembles. La première tâche sera bien sûr d'éliminer les ensembles redondants: non seulement ils alourdissent la charge calculatoire mais ils dégradent la justesse sans améliorer la mensuration. Ensuite, faute de règle générale, il faudra se livrer à une analyse ensembliste des informations disponibles afin d'en déterminer les plus efficaces.

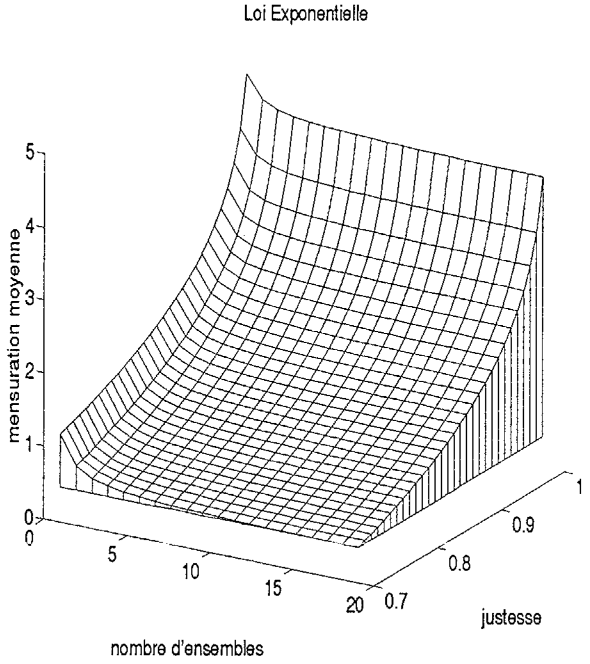

Figure 2: Mensuration Moyenne - Loi Exponentielle.

A titre d'illustration, considérons le cas suivant. Dans (2), soient $\Xi = \mathbb{R}$, $T_n(h) = h$, et $(U_n)_{n\in\mathbb{Z}}$ une suite de variables aléatoires indépendantes et identiquement distribuées. Considérons la formulation ensembliste $(\mathbb{R}, (S_n)_{1\leq n\leq N})$ où S_n est défini en (6). On prend le diamètre $\nu(S)$ de l'ensemble solution comme mensuration. La justesse est $\mathsf{P}H = c^N$, où c est la confiance sur chaque S_n. D'après (6), $S_n = [X_n - \lambda_2, X_n - \lambda_1]$ et on obtient

$$\mathsf{E}\nu(S) = \lambda_2 - \lambda_1 + \mathsf{E}\min_{1\leq n\leq N} X_n - \mathsf{E}\max_{1\leq n\leq N} X_n. \tag{17}$$

Les figures 1 et 2 montrent l'évolution de $\mathsf{E}\nu(S)$ en fonction de $\mathsf{P}H$ et de N pour les cas où U_0 est: 1) normale avec $\mathsf{E}U_0 = 0$ et $\mathsf{E}|U_0|^2 = 1$ et 2) exponentielle avec $\mathsf{E}U_0 = 1$ (voir [1] pour l'évaluation des espérances de (17)). On voit donc qu'au delà d'une dizaine d'ensembles il n'y a pratiquement plus d'amélioration de $\mathsf{E}\nu(S)$ pour $\mathsf{P}H$ fixée dans les deux cas. Notons bien que pour une formulation ensembliste quelconque il n'est pas à exclure que pour une justesse donnée la mensuration puisse même croître avec le nombre d'ensembles utilisés.

MODÈLES À INCERTITUDE BORNÉE

Dans cette section, on s'intéresse aux modèles pour lesquels le processus d'incertitude est uniformément borné, à savoir

$$(\forall n\in\mathbb{Z}) \quad |U_n| \leq \lambda \quad \mathsf{P}\text{-p.s.} \tag{18}$$

Ce type de modèle est fréquemment utilisé en traitement du signal et en analyse de systèmes [2], [5]. Grâce à (18), on peut garantir une confiance de 1 sur chaque ensemble de (6) en prenant $\lambda_2 = -\lambda_1 = \lambda$. Les N ensembles $(S_n)_{1\leq n\leq N}$ possibles doivent donc être retenus. Cette sélection produira la plus petite intersection sans pour autant nuire à la justesse puisque

$$\mathsf{P}\{\omega\in\Omega \mid h \in \bigcap_{n=1}^{N} S_n(\omega)\} = 1. \tag{19}$$

Penchons nous à présent sur la question d'améliorer la mensuration de l'ensemble solution en ajoutant d'autres sous-ensembles à ceux de (6) à partir de (18) et sans aucune connaissance des moments ou de la distribution spectrale de $(U_n)_{n\in\mathbb{Z}}$. Considérons tout d'abord l'ensemble (8), sous les hypothèses adéquates. Puisque $\mathsf{E}U_0^{2k} \leq \lambda^{2k}$, (8) est remplacé par

$$S_k'(\omega) = \{a\in\Xi \mid |\sum_{n=1}^{N}(X_n(\omega) - T_n(a))^k| \leq \alpha\lambda^k\sqrt{N}\}. \tag{20}$$

Un tel ensemble peut être utile car (18) donne seulement lieu à la contrainte $|\sum_{n=1}^{N}(X_n(\omega) - T_n(a))^k| \leq N\lambda^k$ et $S_k'(\omega)$ ne contient donc pas nécessairement $\bigcap_{n=1}^{N} S_n(\omega)$. La même démarche amène à remplacer (9) par

$$S_p'(\omega) = \{a\in\Xi \mid \sum_{n=1}^{N} |X_n(\omega) - T_n(a)|^p/N \leq \lambda^p(1+\alpha/\sqrt{N})\}. \tag{21}$$

Cependant cet ensemble ne présente aucun intérêt car d'après (18) il contient toujours $\bigcap_{n=1}^{N} S_n(\omega)$ et s'avère donc redondant. Finalement, considérons (10) avec en plus des hypothèses adéquates celle que $(U_n)_{n\in\mathbb{Z}}$ est un bruit blanc. En majorant sa puissance par λ^2, (10) devient

$$\begin{aligned} S_l'(\omega) = \{a\in\Xi \mid\ & |\sum_{n=1}^{N}(X_n(\omega) - T_n(a))e^{-\imath 2\pi\nu_l n}|^2 \\ & \leq -N\lambda^2\ln(1-c_l)\}. \end{aligned} \tag{22}$$

On remarque par ailleurs que (18) n'autorise que la contrainte $|\sum_{n=1}^{N}(X_n(\omega) - T_n(a))e^{-\imath 2\pi\nu_l n}|^2 \leq N^2\lambda^2$. $S_l'(\omega)$ n'est donc pas forcément redondant en présence de $\bigcap_{n=1}^{N} S_n(\omega)$ et il peut donc être utile.

RÉFÉRENCES

[1] N. Balakrishnan and A. Clifford-Cohen, *Order Statistics and Inference.* San Diego, CA: Academic Press, 1991.

[2] P. L. Combettes, "The Foundations of Set Theoretic Estimation," *Proceedings of the IEEE,* vol. 81, no. 2, pp. 182-208, February 1993.

[3] P. L. Combettes, M. Benidir, and B. Picinbono, "A General Framework for the Incorporation of Uncertainty in Set Theoretic Estimation," *ICASSP Proceedings,* vol. 3, pp. 349-352. San Francisco, CA, March 23-26, 1992.

[4] P. L. Combettes and H. J. Trussell, "The Use of Noise Properties in Set Theoretic Estimation," *IEEE Transactions on Signal Processing,* vol. 39, no. 7, pp. 1630-1641, July 1991.

[5] E. Walter (Guest Editor), "Parameter Identifications with Error Bounds," *Mathematics and Computers in Simulation,* vol. 32, no. 5-6, December 1990.