

Détection sous-optimale et quantifieurs stochastiques

Cédric DUCHÊNE, Steeve ZOZOR, Pierre Olivier AMBLARD

Laboratoire des Images et des signaux, CNRS UMR 5083
961 Rue de la Houille Blanche, B.P. 46, 38402 Grenoble Cedex, France

cedric.duchene@lis.inpg.fr, steeve.zozor@lis.inpg.fr, bidou.amblard@lis.inpg.fr

Résumé – L’implantation de la solution optimale d’un problème de détection peut être parfois difficile. Poor [1] a proposé en 1977 une solution sous-optimale fondée sur la quantification des données à traiter. Cependant, cette méthode, moins compliquée à implanter, reste néanmoins difficile à mettre en œuvre étant donné que le quantifieur optimal est solution d’un difficile problème d’optimisation. Par ailleurs, un quantifieur peut être considéré comme étant un réseau de seuils. Ces dispositifs possèdent des comportements étranges en présence de bruit [2, 3, 4]: ils sont capables d’améliorer un traitement de données par ajout de bruit sur les seuils. Nous montrerons que les quantifieurs stochastiques utilisés comme détecteurs présentent l’effet de détection améliorée par le bruit. De plus, une fois les seuils tirés, si les niveaux de quantifications sont adaptés aux statistiques des observations, alors les performances sont quasiment constantes quelle que soit la puissance des fluctuations.

Abstract – In many cases of binary detection problem the optimal detector cannot be implemented, notably for nonGaussian input noises. Even if optimal quantizer-detection systems [1] are easily implementable it is very difficult to determine their parameters because they are issued from an optimization problem and depend on the data at hand. Another point of view is to consider a quantizer as an array of threshold devices. These systems were studied in [2, 3, 4] for their strange behavior often named stochastic resonance: they are able to improve the processing of signals by adding noise prior to the nonlinearities. We propose to build quantizer-decision systems with stochastic thresholds which are able to use noise in a benefic way. Also, we will present the performance of detectors whose levels are fitted to the statistics of the datas, and we will remark these performance are constant whatever the power of the fluctuations.

1 Détection et quantifieurs

Considérons le problème de détection suivant :

$$\left. \begin{array}{l} H_0 : x_i = \xi_i + s_0 \\ H_1 : x_i = \xi_i + s_1 \end{array} \right\} i = 1, \dots, n \quad (1)$$

dans lequel $x_i \sim f_j(x) = f_\xi(x - s_j)$ avec f_ξ la densité de probabilité du bruit de canal, f_j la densité de probabilité des observations x_i conditionnée par l’hypothèse H_j . Nous nous placerons dans le cadre d’étude le plus simple : le bruit ξ est indépendant et identiquement distribué (iid) et les signaux s_0 et s_1 sont constants. Par conséquent, la séquence d’observation $\mathbf{x} = \{x_i\}_{i=0}^n$ est iid également. Selon les critères de Neyman Pearson, Bayes ou encore du Maximum *a posteriori* (MAP) [5], la meilleure stratégie de décision est le test du log-rapport-de-vraisemblance (LLRT pour log-likelihood ratio test) :

$$\Lambda(\mathbf{x}) = \log \frac{f_1(\mathbf{x})}{f_0(\mathbf{x})} = \sum_{i=1}^n \log l(x_i) \begin{array}{l} H_1 \\ \geq \eta_\Lambda \\ H_0 \end{array} \quad (2)$$

où $l(x_i) = f_1(x_i)/f_0(x_i)$ est le rapport de vraisemblance pour chaque échantillon x_i . Cependant, cette solution ne peut pas toujours être physiquement implantée dans une procédure de détection. En effet, il se peut que nous n’ayons pas accès aux densités f_j , ou que celles-ci n’aient pas d’expression analytique (cas des densités α -stable par exemple [6]). Ce sont les principales raisons qui ont poussé de nombreux auteurs [1, 7, 8, 9] à

se pencher sur le problème de conception de détecteurs sous-optimaux. Plusieurs stratégies, autres que celles suivies pour l’élaboration du LLRT, peuvent être envisagées pour la construction de détecteurs sous-optimaux. Ces stratégies doivent tenir compte des informations disponibles sur le problème à traiter. En revanche, comme nous l’avons déjà souligné ce sont souvent les informations non disponibles qui nous guident vers une stratégie sous-optimale. Dans tous les cas, utiliser un tel détecteur conduit bien évidemment à une perte de performances par rapport à celles qui seraient obtenues avec le LLRT, celles-ci pouvant être bornées par les expressions données par Orsak et Paris dans [7]. D’autres auteurs [1, 8, 9] ont restreint le problème en imposant la structure du détecteur à celle d’un quantifieur. Rappelons qu’un quantifieur scalaire est un dispositif qui convertit une entrée réelle en une grandeur prédéfinie sur M niveaux. A chaque niveau de quantification est associé un intervalle auquel chaque entrée est comparée. Un quantifieur Q est complètement caractérisé par le couple de vecteurs (\mathbf{t}, \mathbf{q}) et est défini par :

$$Q : \mathbb{R} \mapsto \{q_1, \dots, q_M\} \quad (3)$$

$$x \mapsto y = q_k \text{ si } x \in [t_{k-1}; t_k[$$

où t_k sont seuils et q_k les niveaux de quantification. Dans ce cas, la stratégie de décision basée sur le quantifieur Q est décrite par :

$$T_Q(\mathbf{x}) = \sum_{i=1}^n Q(x_i) = \sum_{k=1}^M Z_k \cdot q_k \begin{array}{l} H_1 \\ \geq \eta \\ H_0 \end{array} \quad (4)$$

où Z_k correspond au nombre de fois où une entrée x_i tombe dans l'intervalle $[t_{k-1}; t_k[$. Sous chaque hypothèse H_j , $j = 0, 1$, le vecteur $\mathbf{Z} = \{Z_1, \dots, Z_M\}$ est distribué selon une loi multinomiale que nous noterons :

$$\mathbf{Z} \sim \mathcal{M}(n, p_{1,j}(\mathbf{t}), \dots, p_{M,j}(\mathbf{t})) \quad (5)$$

avec $p_{k,j}(\mathbf{t}) = \Pr[x_i \in [t_{k-1}; t_k[\mid H_j]$. Poor et Thomas ont proposé dans [1] une méthode fondée sur la recherche des intervalles de quantification optimaux par maximisation d'une « distance » (f -divergence) entre les densités de probabilité post quantification $f_{j,Q}(\mathbf{x})$ pour $j = 0, 1$. Au préalable, on montre que pour \mathbf{t} fixé, les valeurs des niveaux de quantification \mathbf{q} optimaux sont les rapports de vraisemblance locaux pour chaque intervalle de quantification. Il reste ensuite à trouver la valeur des seuils \mathbf{t} optimaux maximisant les statistiques sous chaque hypothèse après quantification. Cette procédure se résume par :

$$q_k^{opt} = \log \frac{F_1(t_k) - F_1(t_{k-1})}{F_0(t_k) - F_0(t_{k-1})} \quad (6)$$

$$\mathbf{t}^{opt} = \underset{\mathbf{t}}{\text{Arg max}} \{d_f(f_{0,Q}(\mathbf{x}), f_{1,Q}(\mathbf{x}))\} \quad (7)$$

dans laquelle on note $F_j(t_k) = \int_{-\infty}^{t_k} f_j(x) dx$ la fonction de répartition des observations conditionnelle à H_j . Le quantifieur ainsi obtenu, peut être vu comme étant la meilleure approximation par escalier de la log-rapport-de-vraisemblance du problème.

2 Quantifieurs stochastiques

Nous avons étudié deux quantifieurs stochastiques différents. Ceux-ci sont obtenus par fluctuations aléatoires des seuils de quantification t_k . On supposera les bruits ajoutés sur chaque seuil indépendant. Ainsi, la densité conjointe des seuils bruités \tilde{t}_k se factorise sous la forme $f_{\tilde{\mathbf{t}}}(\boldsymbol{\theta}) = \prod_{k=1}^{M-1} f_{\tilde{t}_k}(\theta_k)$. Pour garder une structure de quantifieur nous imposons aux seuils bruités la contrainte suivante : $\tilde{t}_0 = -\infty < \dots < \tilde{t}_M = \infty$.

- Le premier quantifieur porte le nom de quantifieur uniforme stochastique (QUS). Il est réalisé à partir d'un quantifieur déterministe dont les bornes \mathbf{t} et les niveaux \mathbf{q} de quantification sont répartis de façon uniforme et en faisant fluctuer \mathbf{t} .
- Le second, est dit adapté (QAS), car les niveaux \mathbf{q} ne sont plus uniformes mais pour chaque intervalle le niveau de quantification prend la valeur du rapport de vraisemblance local comme dans (6).

2.1 QUS et P_e minimale

Nous nous sommes intéressés à l'influence du seuil η et de la puissance des fluctuations $\sigma_{\tilde{\mathbf{t}}}^2$ pour un schéma de détection basé sur un QUS. Dans cette partie, les nouvelles bornes du quantifieur sont données par l'expression $\tilde{t}_k = t_k + \tau_k$ avec τ_k distribué uniformément sur l'intervalle $[0, 2\sqrt{3}\sigma_{\tilde{\mathbf{t}}}]$ et t_k la valeur de la borne sans fluctuation. Nous avons retenu la probabilité d'erreur comme

critère d'évaluation des performances de nos schémas de détection. Pour un quantifieur déterministe P_e est obtenue facilement en utilisant la propriété (5). Dans le cas d'un quantifieur stochastique, on écrit que la probabilité d'erreur conditionnelle à $\tilde{\mathbf{t}}$ est celle d'un quantifieur déterministe que l'on note $P_e|_{\tilde{\mathbf{t}}}$. Il suffit alors de moyenner suivant la densité conjointe des seuils bruités :

$$P_e = \int_{\mathbb{R}^{M-1}} P_e|_{\tilde{\mathbf{t}}=\boldsymbol{\theta}}(\boldsymbol{\theta}) \cdot f_{\tilde{\mathbf{t}}}(\boldsymbol{\theta}) \cdot d\boldsymbol{\theta}$$

pour obtenir la probabilité d'erreur du quantifieur stochastique. On obtient alors :

$$\begin{aligned} P_e &= \pi_0 P_{D_1|H_0} + \pi_1 P_{D_0|H_1} \\ &= \pi_1 + \int_{\mathbb{R}^{M-1}} \sum_{\mathbf{l} \in \mathcal{D}(n, \mathbf{q}, \eta)} \frac{n!}{l_1! \dots l_M!} \\ &\quad \times \left(\pi_0 \prod_{k=1}^M p_{k,1}^{l_k}(\boldsymbol{\theta}) - \pi_1 \prod_{k=1}^M p_{k,0}^{l_k}(\boldsymbol{\theta}) \right) f_{\tilde{\mathbf{t}}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \end{aligned} \quad (8)$$

dans laquelle $\mathcal{D}(n, \mathbf{q}, \eta) = \{\mathbf{l} \in \{0, \dots, n\}^M \text{ tel que } \sum_{k=1}^M l_k = n \text{ et } \sum_{k=1}^M l_k q_k > \eta\}$ et où $\mathcal{C}(\mathbf{l}) = \{0\} \times \{0, 1, \dots, l_2\} \times \dots \times \{0, 1, \dots, l_M\}$. π_j sont les probabilités *a priori* des hypothèses H_j . Pour un quantifieur uniforme, l'indépendance des t_k et quelques développements mathématiques conduisent à une expression de P_e dans laquelle seules des intégrales simples interviennent :

$$\begin{aligned} P_e &= \pi_1 + \sum_{\mathbf{l} \in \mathcal{D}(n, \mathbf{q}, \eta), \mathbf{m} \in \mathcal{C}(\mathbf{l})} \frac{n! (-1)^{\sum_{k=1}^M m_k}}{\prod_{k=1}^M (l_k - m_k)! m_k!} \\ &\quad \times \left(\sum_{i=0}^1 (-1)^i \pi_i \prod_{k=1}^{M-1} \int_{\mathbb{R}} F_i(\theta)^{l_k - m_k + m_{k+1}} f_{\tilde{t}_k}(\theta) d\theta \right) \end{aligned} \quad (9)$$

La figure 1 représente l'évolution de la probabilité d'erreur d'un QUS en fonction de la puissance des fluctuations et du seuil de détection. Celle-ci est calculée par intégration numérique de (9). Le problème est asymétrique, les constantes à détecter sont : $s_0 = 0$ et $s_1 = 2$. Le quantifieur uniforme est de dimension $M = 4$. Ses seuils sans bruit sont $\mathbf{t} = [-\infty, -2, 0, 2, \infty]$ et ses niveaux de quantification sont $\mathbf{q} = [-3, -1, 1, 3]$. Notons que le quantifieur est symétrique sans fluctuation, mais dès l'application de fluctuations non centrées τ_k sur ces seuils t_k il perd cette propriété de symétrie. Il est intéressant de remarquer que l'influence du seuil η ne modifie pas les performances de détection de manière continue. Les courbes P_e à $\sigma_{\tilde{\mathbf{t}}}^2$ fixée sont constantes par paliers. Ceci s'explique par le fait que T_Q est une grandeur discrète définie dans un ensemble de taille finie. En effet, les quantifieurs en général, possèdent la propriété de robustesse vis-à-vis des fluctuations du seuil de détection. On peut également remarquer que pour un seuil η fixe, les fluctuations peuvent nous conduire à une amélioration des performances. C'est par exemple le cas pour $\eta = 1$, où la probabilité d'erreur P_e atteint un minimum en $\sigma = 0.23$. Le QUS possède un comportement de « détection améliorée par le bruit ». Ce comportement s'explique en partie par le fait que : sans fluctuation le quantifieur est centré, alors que le problème

est décentré. Si nous ne sommes pas en mesure de repositionner correctement le quantifieur, alors l'ajout de bruit de moyenne non nulle, peut corriger son mauvais positionnement vis-à-vis des constantes s_0 et s_1 . La probabilité d'erreur minimale ($P_e = 0.03$) est obtenue sans bruit, pour $\sigma_{\mathbf{t}} = 0$ et pour la plage $3 < \eta < 5$.

La figure 3 donne les variations de P_e (cercles) en fonction de $\sigma_{\mathbf{t}}^2$ pour le QUS avec un seuil $\eta = 1$, ainsi que le minimum de la probabilité d'erreur (croix) de ce même quantifieur. La première courbe correspond à une tranche du réseau tracé dans la figure 1, et la seconde est obtenue pour $\sigma_{\mathbf{t}}$ fixé et en recherchant le seuil de détection η donnant la plus faible probabilité d'erreur, c'est à dire $P_e^{\min} = \text{Arg min}_{\eta} \{P_e(\eta)\}$. La ligne en pointillé représente les performances obtenues avec un quantifieur optimal (sans bruit de seuil) fondé sur (6) et (7) et où la divergence utilisée est la divergence de Kullback.

Remarquons que le QUS présente une courbe de performance qui tangente celle de probabilité d'erreur minimale pour une large plage de valeur de $\sigma_{\mathbf{t}}$. En terme de stratégie de décision, il suffirait donc d'optimiser la puissance des fluctuations pour obtenir un quantifieur assez performant lorsqu'il serait impossible d'ajuster le seuil η , ou encore d'éviter une phase coûteuse d'optimisation du seuil η . Notons également que les pertes de performance du QUS sans bruit ($\sigma_{\mathbf{t}} = 0$ et $3 < \eta < 5$) sont très faibles par rapport aux performances de détection du quantifieur optimal.

2.2 QAS versus QUS

Dans cette partie nous allons comparer les performances de deux quantifieurs stochastiques différents : le quantifieur uniforme et le quantifieur adapté. Dans les deux cas, l'expression des seuils bruités est donnée par $\tilde{t}_k(i) = t_k + \tau_k(i)$ avec $\tau_k(i)$ distribué uniformément sur l'intervalle $[0, 2\sqrt{3}\sigma_{\mathbf{t}}]$ et t_k la valeur du seuil sans fluctuation. Ainsi, cette étude sera menée pour deux modes de fonctionnement distinct des quantifieurs :

- Le mode variant dans le temps : où pour chaque échantillons x_i le vecteur $\tilde{\mathbf{t}}$ est tiré à nouveau.
- Le mode invariant dans le temps : où $\tilde{\mathbf{t}}$ reste le même pour toute la séquence \mathbf{x} , i.e. $\tau_k(i) = \tau_k$.

Comme précédemment, les performances de ces quatre quantifieurs sont évaluées à l'aide de leur probabilité d'erreur. Dans 2.1, à l'aide de (9), dans un cadre invariant dans le temps, nous avons réussi à simplifier P_e en une expression moins coûteuse en calcul. Dans le cadre du quantifieur adapté, cette tâche se complique fortement du fait que les niveaux \mathbf{q} de quantification sont des fonctions des seuils bruités $\tilde{\mathbf{t}}$. Dès que le quantifieur devient variant dans le temps, il n'est même plus possible de s'appuyer sur (5) pour le calcul de P_e . Donc, dans la majorité des cas, on s'aperçoit qu'il n'existe pas d'écriture simple de la probabilité d'erreur. Nous pouvons alors : soit avoir recours à une intégration numérique coûteuse pour son évaluation, ou encore estimer sa valeur par une méthode de Monte-Carlo. Pour tracer la figure 2 nous avons opté pour cette dernière démarche. Cette figure décrit donc le comportement des deux quantifieurs (uniforme et adapté) dans le

cas variant et invariant temporellement. Les performances sont évaluées en traçant l'évolution d'une estimation de la probabilité d'erreur P_e en fonction de la puissance des fluctuations pour un bruit de canal Laplacien. Comme sur la figure 3, la ligne en pointillé correspond aux performances obtenues avec un quantifieur optimal (sans bruit de seuil) fondé sur (6) et (7) et où la divergence utilisée est la divergence de Kullback.

On peut imaginer que dans le cas où nous n'avons pas accès aux densités f_j , il est judicieux d'implanter le quantifieur uniforme et d'optimiser la puissance des fluctuations afin d'obtenir les meilleures performances de détection. Cette stratégie peut aussi se révéler efficace en supposant que le quantifieur évolue dans un environnement intrinsèquement bruité. Ainsi, par ajout de bruit sur les seuils, on peut espérer compenser l'effet néfaste du bruit natif. Dans le cas où les densités sont connues, on peut alors utiliser le quantifieur adapté. Dans ce cas, nous devons calculer les valeurs des rapports de vraisemblance uniquement pour chaque borne de quantification et non pas pour chaque observation x_i comme on devrait le faire avec le LLRT. De même, on peut remarquer combien le quantifieur adapté est peu sensible aux variations de $\sigma_{\mathbf{t}}$. Les performances sont pratiquement semblables quelque soit la puissance des fluctuations. Nous pouvons alors imaginer une stratégie de conception de détecteur fondée sur celle de Poor, mais en s'affranchissant de l'étape d'optimisation des bornes t_k . En d'autres termes cette méthode de conception peut être vue comme une version stochastique de l'algorithme de Poor. Finalement, cette stratégie sous-optimale, beaucoup plus simple que le quantifieur optimal, possède néanmoins de bonnes performances. En effet, la probabilité d'erreur du quantifieur adapté est très proche de celle du quantifieur optimal de Poor.

3 Conclusion

Nous avons proposé deux solutions sous-optimales basées sur l'utilisation de quantifieur stochastique. La première, le QUS, trouve un intérêt d'utilisation lorsqu'il est impossible d'avoir accès aux densités f_j . L'optimisation de la puissance des fluctuations $\sigma_{\mathbf{t}}^2$, bien que nécessaire, semble assez simple à mettre en œuvre car pour une large plage de valeur de $\sigma_{\mathbf{t}}^2$, le quantifieur donne ses meilleures performances. Quant au QAS, celui-ci peut être mis en pratique uniquement lorsque les densités f_j sont connues. Il est véritablement plus simple d'implanter le QAS plutôt que le LLRT. Le QAS donne des performances pratiquement équivalentes au quantifieur optimal de Poor. Nous avons vu que quelque soit la puissance des fluctuations $\sigma_{\mathbf{t}}^2$, ce quantifieur donne globalement les mêmes résultats de détection. Dans les deux cas, ce type de détection présente un intérêt pratique lorsque des bruits sont natifs dans l'implantation hard du quantifieur. Par ajout de bruit sur les seuils, on espère contrebalancer ces effets indésirables. Pour conclure, il est intéressant de remarquer à quel point ces schémas de détection sont semblables à certaines architectures neuronales (notamment l'architecture Pool).

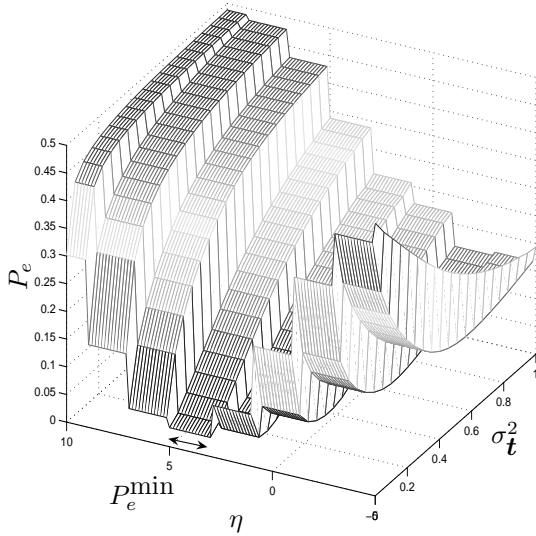


FIG. 1 – Probabilité d’erreur d’un QUS invariant temporellement en fonction de la puissance des fluctuations σ_t^2 et du seuil de détection η . Dimension $M = 4$, seuils de quantification $\mathbf{t} = [-\infty, -2, 0, 2, \infty]$, niveaux de quantification $\mathbf{q} = [-3, -1, 1, 3]$. Nombre d’échantillon $n = 5$, bruit de canal Laplacien centré et de variance unitaire $\sigma_\xi = 1$. Constantes à détecter $s_1 = 2$ et $s_0 = 0$.

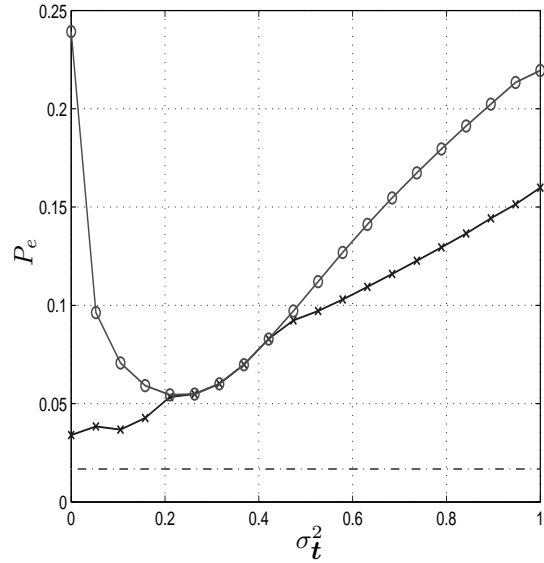


FIG. 3 – Cercles : Probabilité d’erreur d’un QUS invariant temporellement en fonction de la puissance des fluctuations σ_t^2 pour un seuil de détection $\eta = 1$ - Croix : Probabilité d’erreur minimale de ce même quantifieur. Les autres paramètres sont identiques à ceux de la Figure 1

Références

- [1] H. V. Poor and J. B. Thomas. Application of ali-silvey distance measures in the design of generalized quantizers for binary decision systems. *IEEE Transactions on Communication*, 25(9) :893–900, September 1977.
- [2] D. Rousseau, F. Duan, and F. Chapeau-Blondeau. Suprathreshold stochastic resonance and noise-enhanced fisher information in arrays of threshold devices. *Physical Review E*, 68, 2003.
- [3] N. G. Stocks. Suprathreshold stochastic resonance in multilevel threshold systems. *Physical Review Letters*, 84(11) :2310 – 2313, March 2000.
- [4] N. G. Stocks. Information transmission in parallel threshold arrays : Suprathreshold stochastic resonance. *Physical Review E*, 63, March 2001.
- [5] H. L. Van Trees. *Detection, Estimation, and Modulation Theory*. part 1. John Wiley & Sons, New York, 1968.
- [6] G. Samorodintsky and M. S. Taqqu. *Stable nonGaussian random processes. Stochastic models with infinite variance*. Chapman and Hall, New-York, 1994.
- [7] G. C. Orsak and B.-P. Paris. On the relationship between measures of discrimination and the performance of suboptimal detectors. *IEEE Transactions on Information Theory*, 41(1) :188–203, January 1995.
- [8] S. A. Kassam. Optimum quantization for signal detection. *IEEE Transactions on Communications*, 25(5) :479–484, May 1977.
- [9] B. Picinbono and P. Duvaut. Optimum quantization for detection. *IEEE Transaction on Communications*, 36(11) :1254, November 1988.

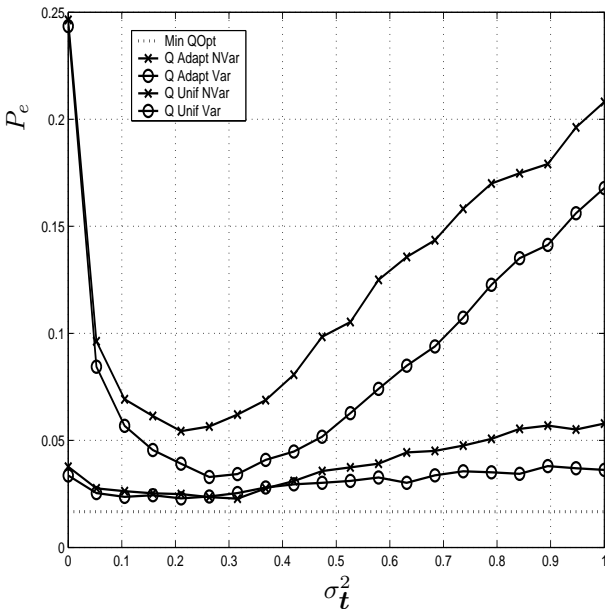


FIG. 2 – Probabilité d’erreur des QUS et QAS en fonction de la puissance des fluctuations σ_t^2 dans un cadre variant et invariant dans le temps. Les autres paramètres sont identiques à ceux de la Figure 1.