Astérisque

GABRIEL RUGET

Quelques occurences des grands écarts dans la littérature « électronique »

Astérisque, tome 68 (1979), p. 187-199

<http://www.numdam.org/item?id=AST_1979__68__187_0>

QUELQUES OCCURENCES DES GRANDS ÉCARTS DANS LA LITTÉRATURE "ÉLECTRONIQUE"

Gabriel RUGET.

Sous le nom de "Chernoff bound", la majoration exponentielle de la probabilité qu'une somme de n variables indépendantes dépasse un seuil na est aussi répandue dans les mathématiques de l'électronique que le théorème central limite ; pour voir un exemple très proche du centre de cet exposé, la transmission des données, on pourra examiner [18] puis [22]*.

Plus rares sont les utilisations (conscientes ou pas) de théorèmes de grands écarts d'apparence plus sophistiquée quoique équivalents (lois empiriques, marches aléatoires). Mais ceux qui soupèsent avec délicatesse la queue d'une queue (cf. [10]), pourraient utiliser les théorèmes de Borovkov (cf. exposé n° IV, Deshayes - Picard) qui, non seulement peuvent prendre la place de l'identité de Wald, mais permettent d'éviter le recours à [13]: une queue augmente d'une unité selon un processus de renouvellement (loi G de transformée de Cramer Γ) et diminue de même (loi F de transformée de Cramer φ, arrivées et départs indépendants) ; l'espérance de F est inférieure à celle de G, de sorte qu'en moyenne la queue se vide ; la probabilité d'une excursion, à partir de la taille d'équilibre 0, jusqu'à la taille n est asymptotiquement exponentielle et l'on

---

*Une suite de données (à nombre de niveaux fini) passe dans un filtre linéaire (interférence intersymboles), puis est bruitée ; suivant une idée de Forney, on veut la rétablir en filtrant de façon à réduire le nombre de symboles interférant, sans trop augmenter le bruit, puis en décodant séquentiellement (et on peut alors se permettre de décoder au maximum de vraisemblance en utilisant l'algorithme de Viterbi); la ligne de transmission (i.e. le premier filtre) n'étant pas stationnaire, le filtre à la réception est rendu adaptatif; il n'est donc jamais exactement ce qu'on voudrait, et cela provoque un bruit d'interférence intersymboles résiduel que l'on veut majorer pour évaluer les performances du décodeur.

cherche l'exposant. Heuristiquement : si C_i désignent les temps entre départs successifs, $P(C_1 + \ldots + C_k \simeq kx) \simeq e^{-k\phi(x)}$ donc $P(C_1 + \ldots + C_{\alpha T} \simeq T) \simeq e^{-\alpha T \phi(1/\alpha)}$; la probabilité que, pendant un temps T, il y ait α départs et β arrivées par unité de temps, vaut $\exp - T(\alpha\phi(1/\alpha) + \beta\Gamma(1/\beta))$; or le temps pour que, dans ces conditions, la taille de la queue augmente de n, vaut $\frac{n}{\beta-\alpha}$; l'exposant cherché est donc

$$\min_{\alpha,\beta} \frac{1}{\beta-\alpha} (\alpha\phi(1/\alpha) + \beta\Gamma(1/\beta))$$

En dérivant, on trouve les conditions

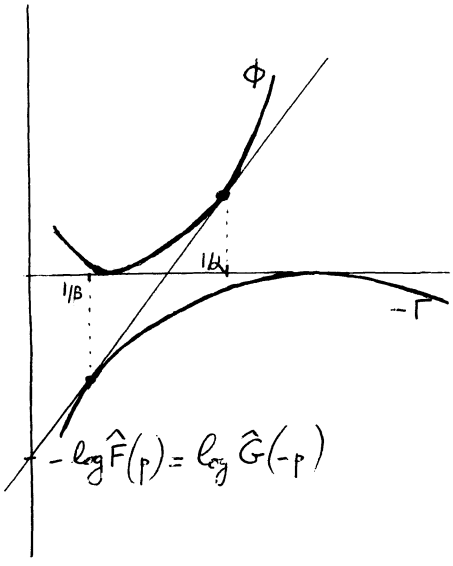
$$\phi(1/\alpha) + \Gamma(1/\beta) = (\frac{1}{\alpha} - \frac{1}{\beta})\,\phi'(1/\alpha) = (\frac{1}{\beta} - \frac{1}{\alpha})\,\Gamma'(1/\beta)$$

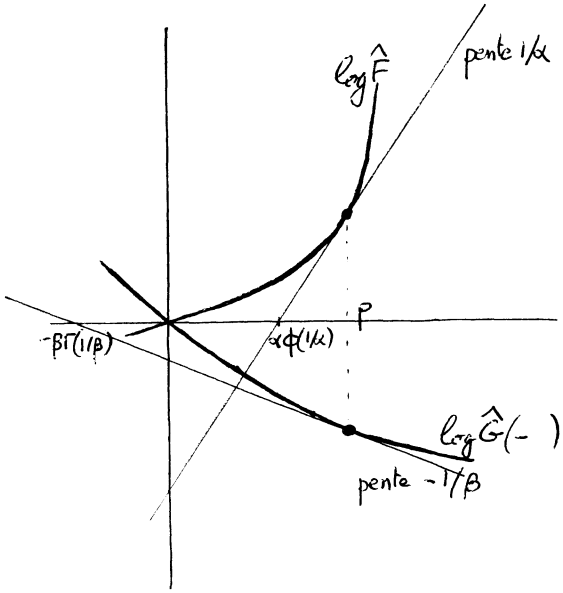
Posant $\phi'(1/\alpha) = -\Gamma'(1/\beta) = p$, au point solution, les conditions deviennent

$$\phi(1/\alpha) - p\frac{1}{\alpha} = -\Gamma(1/\beta) - p\frac{1}{\beta} \quad ,$$

ce qui exprime que les tangentes en $1/\alpha$ à ϕ et en $1/\beta$ à $-\Gamma$ sont confondues, d'où la condition déterminant p

$$\hat{F}(p)\,\hat{G}(-p) = 1$$





Un retour au graphe des transformées de Laplace montre qu'alors

$$\alpha\phi\ (1/\alpha) + \beta\Gamma\ (1/\beta) = -\beta \log \hat{G}\ (-p) - \alpha\ \log \hat{F}\ (p)$$

D'où l'exposant $\log \hat{F}\ (p)$.

Cette heuristique peut être transformée en analyse précise comme suit : on définit une nouvelle loi $\tilde{P}$ pour le processus en décrétant que les temps de service sont de loi $d\tilde{F} = e^{p\cdot}\ dF(.)/\hat{F}\ (p)$ et les temps entre arrivées de loi $d\tilde{G} = e^{-p\cdot}\ dG(.)/\hat{G}\ (p)$, pour le p déterminé ci-dessus. On peut évaluer $P\ (\Sigma)/\tilde{P}\ (\Sigma)$, où Σ désigne l'ensemble des trajectoires du processus issues de 0 (file vide) et atteignant la taille n sans se vider à aucun instant intermédiaire : on trouve

$$P\ (\Sigma)\ /\ \tilde{P}\ (\Sigma) = \frac{\alpha}{p}\ (\hat{F}\ (p) - 1)\ \hat{F}\ (p)^{-n}$$

$\tilde{P}\ (\Sigma)$, qui tend vers une limite non nulle lorsque $n \to \infty$, peut être évalué numériquement à peu de frais, ainsi que le temps moyen de retour à l'état vide, on dispose alors du temps moyen de premier passage à l'état n .

Bien entendu, cette approche permet (sans aboutir à un résultat explicite aussi simple) d'envisager le cas de plusieurs processus de renouvellement indépendants rythmant chacun un processus d'accroissements non plus déterministes, mais i.i.d.

Mais le véritable avantage d'une approche consciente des situations de grands écarts, ce sont les outils disponibles lorsque les accroissements ne sont pas homogènes. Tel est le cas dans la modélisation du système ALOHA (transmission de trains de données de longueur fixe, synchronisés, sur un canal - satellite, fibre de verre ... - d'accès libre pour un grand nombre d'utilisateurs : ceux-ci, qui écoutent en permanence le canal, savent après chaque intervalle de temps si celui-ci a été utilisé par zéro, un, ou plusieurs terminaux ; dans ce dernier cas, tous les messages sont perdus, et doivent être réémis suivant une politique à déterminer . On désire calculer :

1°) Pour les politiques les plus simples, qui ne stabilisent pas le canal, le temps moyen avant déstabilisation (le processus $N(t)$ suivi est celui du nombre des messages à réémettre ; les politiques de réémission et la loi d'arrivée de nouveaux messages sont choisis de façon à le rendre markovien). Si $v(n) = E(N(t+1) - N(t) \mid N(t) = n)$ est $\geq \alpha > 0$ pour n assez grand, la chaîne est transiente ("instabilité") ; soit n_c le plus petit n tel que $v(n)$ soit > 0 pour $n \geq n_c$; c'est le temps du premier passage à n_c qu'on appelle temps de déstabilisation.

2°) Pour les politiques plus sophistiquées [7] qui stabilisent le canal (chaîne positive récurrente), le bon critère de comparaison est, non pas le temps moyen nécessaire pour faire passer un message,qui est faible, mais la probabilité que le temps d'attente dépasse un seuil (correspondant à une taille de mémoire tampon). On peut ramener ce problème à l'étude, pour une chaîne de Markov sur $\mathbb{N}$ à "petits" accroissements et petite probabilité d'extinction, de la probabilité de survie d'une trajectoire pendant un temps $> T$ (T grand).

Ces méthodes seront par exemple exposées dans [5] . Comme dans l'exemple précédent, une étude heuristique basée sur des théorèmes asymptotiques grossiers, fournit un bon changement de probabilité, après quoi on peut procéder à une simulation (c'est une technique d' "importance sampling") ou à un calcul analytique (approximation par une diffusion).

Terminons cet exposé par une revue rapide d'un rameau récent de la théorie de Shannon : il s'agit de borner des probabilités d'erreur - exponentiellement petites - lorsqu'on transmet sur un canal ou lorsqu'on code une source avec un critère de distorsion, les codes étant par blocs (ce qui tend vers l'infini et rend la probabilité d'erreur, resp. de dépasser un niveau de distorsion imposé, petite, est la taille des blocs). Précisément :

1° Problème : un canal bruité est caricaturé sous les traits d'une matrice $Q_{k/j}$ de probabilités de transition connue et fixe entre un alphabet d'entrée fini J et un alphabet de sortie fini K (on ne prend pas en compte ici

l'interférence intersymboles). Un code par blocs de longueur N et de rendement R est la donnée de $M = e^{NR}$ mots a_m de longueur N dans l'alphabet J, et d'une partition Y_m de l'ensemble des mots de longueur N de l'alphabet K, indexée par le vocabulaire d'entrée retenu. Outre son rendement, le code est apprécié en fonction de la probabilité d'erreur $\max_{a_m} P$ (la traduction de a_m par le canal ne tombe pas dans Y_m), ou d'une probabilité d'erreur moyenne, ce qui revient au même vu le peu de précision de ce qu'on peut dire de ces probabilités. On cherche une minoration de ces probabilités d'erreur, dont on démontre ensuite l'acuité par des majorations aussi voisines que possible (obtenues en exhibant des codes, même irréalistes, parce que la quantité de calculs à effectuer pour leur décodage serait prohibitive) ; le seul intérêt de ce travail est d'apprécier l'efficacité de codes utilisables. Au mieux, on obtient que la meilleure probabilité d'erreur accessible a un logarithme équivalent à $-NE(R)$ (lorsque $N \to \infty$; on précisera $E(R)$) ; la théorie, qui ne prend pas en compte les coûts de calcul et de transmission, ne dit pas si, pour atteindre une probabilité d'erreur donnée, il vaut mieux chercher à se rapprocher de l'optimum à N modérément grand, ou augmenter N, voire changer le canal (pour jouer sur $E(R)$) ; il est en effet difficile d'évaluer le coût en calcul sans vraiment particulariser les algorithmes, et la pondération entre les coûts de calcul et ceux de transmission dépend de la technologie, donc est très fluctuante ; il semble qu'actuellement, on puisse se permettre beaucoup de calculs, et des N grands.

Résultats :

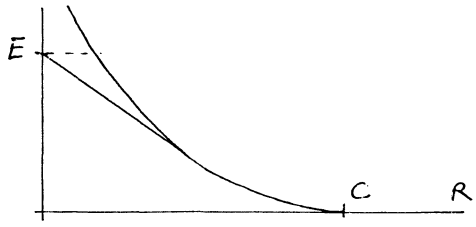
- Shannon s'était d'abord borné à dire que, pour $R < C = \sup_{p} \sum_{j,k} p_j \, Q_{k/j} \log \frac{Q_{k/j}}{\sum_j p_j \, Q_{k/j}}$, on peut réduire autant qu'on veut la probabilité d'erreur, et que c'est impossible pour $R > C$. Quelques articles (de [21] à [15]) étendent la notion de capacité à des canaux stationnaires, mais ayant une mémoire. Nous engageons plutôt le lecteur à consulter l'article de revue [25], où l'on décrit comment ce type de résultats a pu être étendu à des situations de communication plus générales, par exemple un canal portant deux

communications en sens contraire interférant, ou un codeur devant coder pour plusieurs canaux simultanément.(*)

- Ce sont les minorations exponentielles de probabilité d'erreur qui font appel aux grands écarts (pour les majorations, voir [8]). Dans [2], on trouvera une preuve (par réduction à un problème de test) de ce qu'on peut minorer l'erreur en prenant pour exposant

$$E(R) = \max_{s \geq 0} \max_{p} \left[- sR - \log \sum_{k} \left(\sum_{j} p_j \, Q_{k/j}^{1/1+s} \right)^{1+s} \right] .$$

Malheureusement, cette minoration n'est optimale que pour les rendements tels que le s maximisant les expressions ci-dessus soit ≤ 1. Le seul autre résultat dont on dispose est une minoration optimale aux rendements voisins de zéro [23]



$$E = \max_{p} - \sum_{ik} p_i \, p_k \log \sum_{j} \sqrt{Q_{j/i} \, Q_{j/k}}$$

L'enveloppe de ces deux résultats constitue presque l'état actuel de la minoration (voir tout de même [4].(*)

On trouvera dans [23] une minoration de la probabilité que le véritable mot d'entrée ne figure pas dans une liste de longueur donnée des mots les plus vraisemblables au regard de la sortie, résultat utilisé dans [14] pour analyser la principale faiblesse du décodage séquentiel (à rapprocher du point 2° évoqué à propos d'ALOHA). Enfin, dans [13], on trouvera des minorations valables pour tous les N, ce qui est peut-être plus intéressant dans d'autres domaines de la statistique que le codage.

2° Problème : Une source émet une suite infinie de lettres dans un alphabet fini J, de façon aléatoire, mais au moins stationnaire. On veut traduire cette source dans un (peut-être) autre alphabet K, et la qualité de cette traduction est sanctionnée par un critère de distorsion "context-free" d (j,k) . On additionne les distorsions sur les paires de lettres homologues du mot original et du mot traduit.

On fixe la vitesse R de la traduction : par exemple, si on code par blocs de longueur N grande, et si le vocabulaire traduction est constitué de M mots que les statistiques de la source conduisent à utiliser avec la même fréquence, $R = \frac{1}{N} \log M$; de façon plus générale - quoique n'englobant pas le cas précédant, si le procédé de traduction, et donc la traduction, sont stationnaires, c'est l'entropie de la traduction que l'on fixe, $\lim_{n \to \infty} \frac{1}{n} H(Y_n)$, où Y_n désigne le bloc de longueur n de la traduction - voir [20] pour comprendre la signification de l'entropie dans le cas ergodique ; en fait, si l'on code par blocs contigus (ou par bloc glissant) on fixe directement $R = \frac{1}{N} \log M$, car lorsqu'on voudra optimiser la distorsion, l'égalité des fréquences sera ipso facto vérifiée. On veut alors connaître ce qu'on peut faire de mieux comme probabilité qu'un mot source ne puisse être traduit avec une distorsion inférieure à D (par lettre, en moyenne) : dans le cas du codage par blocs, on veut minimiser

$$P\left(x \in J^N \;\middle|\; \min_{y \in \text{code}} \frac{1}{N} \sum d(x_i, y_i) > D\right) .$$

Comme pour le premier problème, on trouve dans la littérature deux types de résultats : ou bien la source S est seulement supposée ergodique, voire stationnaire, et on peut donner une distorsion limite $D^S(R)$ au-dessus de laquelle la probabilité ci-dessus peut être rendue arbitrairement petite (en jouant sur N, et sur le code ; voir [6], [9], [12]). Ou bien on évalue asymptotiquement cette probabilité, mais après s'être placé dans un cas particulier : celui d'une source émettant des lettres i.i.d. (cf. [2], [19], [24]). Nous allons voir qu'il ne doit pas être difficile de faire un peu mieux.

Le résultat sur $D^S(R)$, ou plutôt sur la fonction réciproque $R^S(D)$, est le suivant : n donné, on considère les lois μ_n sur $J^n \times K^n$ de marginale $\pi_J \mu_n$ sur J^n la loi de la source, et telles que la distorsion moyenne $E_{\mu_n} \frac{1}{n} \sum d(x_i, y_i)$ soit inférieure à D. On regarde le minimum parmi ces lois de $\frac{1}{n} I(\mu_n)$, où

$$I(\mu_n) = H(\pi_J \mu_n) + H(\pi_K \mu_n) - H(\mu_n) .$$

Soit $R_n^S(D)$ ce minimum, alors $R^S(D) = \lim_{n \to \infty} R_n^S(D)$.

Les preuves de ce théorème commencent en examinant le cas où la source émet des lettres i.i.d., cas où on a en fait $\lim R_n^S(D) = R_1^S(D)$, et où l'on gagne à exprimer le problème de minimisation sur des probabilités de transition Q de J dans K plutôt que sur des lois jointes

$$R^p(D) = \inf_{Q \in \mathcal{Q}} \sum_{j,k} p_j Q_{jk} \log \frac{Q_{jk}}{\sum_j Q_{jk} p_j}$$

où $$\mathcal{Q} = \{Q \mid D(Q/p) = \sum_{j,k} p_j Q_{jk} d(j,k) \leq D\} .$$

Démontrons (en nous inspirant de [24]) que $R^p(D)$ est inférieur à l'expression proposée : cela nous donnera l'idée de la formule pour une source markovienne, voire n-markovienne (la minoration de $R^S(D)$ est claire, voir [12]).

Avec une probabilité aussi voisine de 1 que l'on veut (en augmentant la taille N des blocs sources codés), les proportions des lettres j dans les blocs sources sont arbitrairement proches des p_j. Nous allons recouvrir l'ensemble des blocs de longueur N ayant une telle "composition" p par des boules de rayon D (au sens de la distorsion) ; les centres de ces boules seront des mots tirés au hasard (lettres i.i.d. de loi q à choisir ; mots indépendants). Tout tient alors dans le

Lemme : étant donné un mot X de composition p, la probabilité que la boule tirée comme indiqué ci-dessus contienne X est, pour le q le plus favorable, "de l'ordre de" $\exp - NR^p(D)$.

Preuve : la loi de la distorsion est le mélange, dans les proportions p_j, des lois affectant à $d(j,k)$ le poids q_k ; son log Laplace est donc $\sum_j p_j \log \sum_k q_k \exp sd(j,k)$; d'où la transformée de Cramer

$$\sup_{s\ (\leq 0)} [sD - \sum_j p_j \log \sum_k q_k \exp sd(j,k)] = \inf_{D(Q/p) \leq D} \sum_{j,k} p_j Q_{jk} \log \frac{Q_{jk}}{q_k}$$

(pour comprendre cette dernière égalité, penser : en tirant sous q, quelle est la probabilité d'obtenir un "canal empirique" Q de distorsion $\leq D$?). La minimi-

sation en q donne alors le résultat énoncé.

On peut maintenant avancer sans difficulté une majoration pour $R^S(D)$ valable dans un cas plus général, disons, pour ne pas alourdir les notations, celui d'une source 1-markovienne. Soit c la loi jointe de deux lettres sources consécutives. Nous nous proposons de tirer des mots de code avec une probabilité de transition bien choisie. Appelons M^c l'ensemble des lois sur $J_1 \times J_2 \times K_1 \times K_2$ (avec $J_1 = J_2 = J$, $K_1 = K_2 = K$) de marginale c sur $J_1 \times J_2$, et de même marginale $\hat{\mu}$ sur $J_1 \times K_1$ et sur $J_2 \times K_2$; appelons $\bar{\mu}$ la marginale de μ sur $K_1 \times K_2$. Pour $\mu \in M^c$, posons

$$D(\mu) = \sum \hat{\mu}_{jk}\, d(j,k)$$

$$I(\mu) = \int d\mu \log \frac{d\mu\,(k_2/j_1, j_2, k_1)}{d\bar{\mu}\,(k_2/k_1)}$$

Alors

$$R^S(D) \leq \inf_{D(\mu) \leq D} I(\mu) \quad ,$$

et l'égalité parait plausible (sur ce sujet, voir [1] et [11]).

Revenons aux sources i.i.d. Il résulte du lemme qu'il existe un code (N,R) tel que le pourcentage des mots sources de composition p' codés avec une distorsion $\geq D$ soit

$$\leq (1 - \exp N(-R^{p'}(D) - O(N)))^{\exp NR}$$

$$\leq \exp - \exp NR \exp N(-R^{p'}(D) - O(N))$$

$$= \exp(-\exp N(R - R^{p'}(D) - O(N))) \ .$$

Toujours d'après [24], on minimise la probabilité de dépassement de la distorsion D' à rendement R donné en mélangeant les codes obtenus ci-dessus pour toutes les compositions p' telles que $R^{p'}(D) < R$ (le nombre de p' à envisager croît comme une puissance de N, ce qui est négligeable devant la taille de chaque code, qui est exponentielle en N). La probabilité étudiée est alors,

à un terme négligeable près, majorée par $\sum_{p' \mid R^{p'}(D) \geq R}$ Probabilité que un N-mot de la source soit de composition p', quantité dont -1/N le logarithme, que nous appellerons F (R,D), vaut

$$\inf_{R^{p'}(D) \geq R} \sum p'_j \log \frac{p'_j}{p_j}$$

Deux remarques :

1°) Nous n'avons guère été soigneux parce que nous savons que F (R,D) dépend continûment de R ; ceci tient à la convexité de F, qui ne se voit pas sur la forme indiquée, mais sous une autre expression due à Blahut [3]

$$F(R,D) = \sup_{Q} \inf_{\substack{I(Q/p') \geq R \\ D(Q,p') \geq D}} I(p'/p)$$

2°) Il est clair que la suite des arguments peut s'étendre aux sources markoviennes, puisque avec l'exposé n° VI , on sait calculer la probabilité qu'un N-mot source ait une loi jointe vérifisant certaines contraintes.

Pour terminer, indiquons comment l'on voit (d'après [20]) que exp - NF (R,D) minore, pour tous les codes (N,R) imaginables, la probabilité de dépassement de la distorsion D. On part du fait que, si q est une probabilité sur J, et $\mathcal{M}$ est l'ensemble des mots distordus par un code (N,R), si $R < R^q(D)$, alors $q(\mathcal{M}) \geq \text{Cste} > 0$. Utilisant toujours le même argument, pour tout p' tel que $R^{p'}(D) > R$, on a

$$p(\mathcal{M}) \geq p(\mathcal{M} \cap \text{composition} \simeq p')$$

$$\geq p'(\mathcal{M} \cap \text{composition} \simeq p') \exp(-N(I(p'/p) + O(N))$$

Puisque $p'(\mathcal{M})$ est minoré, le premier terme l'est aussi, d'où le résultat.

Bibliographie :

[1] BERGER T. : Explicit bounds to R (D) for a binary symmetric Markov source, IEEE IT 23 , n° 1 (1977), p. 52

[2] BLAHUT R. : Hypothesis testing and information theory IEEE IT 20, n° 4 (1974), p. 405

[3] BLAHUT R. : Information bounds of the Fano-Kullback type IEEE IT 22, n° 4 (1976), p. 410

[4] BLAHUT R. : Composition bounds for channel block codes IEEE IT 23, n° 6 (1977), p. 656

[5] COTTRELL M., FORT J.C., MALGOUYRES G. 2nd Int. Conf. on Inf. and Syst. Theory, Patras 1979

[6] DAVISSON L., PURSLEY M. : A direct proof of the coding theorem for discrete sources with memory IEEE IT 21, n° 3 (1975), p. 310

[7] FAYOLLE G., GELENBE E., LABETOULLE J. : Stability and optimal control of the packet switching br adcast channel J.A.C.M., vol. 24, n° 3 (1977), p. 375

[8] GALLAGER R. : A simple derivation of the coding theorem and some applications IEEE IT 11, p. 3

[9] GALLAGER R. : Information theory and reliable communication, New York, Wiley 1968

[10] GOVER D., SHEDLER G. : Approximate models for processor utilization in multiprogrammed computer systems, Siam J. Comput. vol. 2, n° 3 (1973), p. 183

[11] GRAY R. : Information rates of autoregressive processes IEEE IT 16 (1970), p. 412

[12] GRAY R., NEUHOFF D., OMURA J. : Process definitions of distorsion rate functions and source coding theorems IEEE IT 21, n° 5 (1975), p. 524

[13] HAJI R., NEWELL G. : A relation between stationary queue and waiting time distributions
J. Appl. Probability 8 (1971), p. 617

[14] JACOBS I., BERLEKAMP E. : A lower bound to the distribution of computation for sequential decoding
IEEE IT 13, n° 2 (1967), p. 167

[15] KIEFFER J. : On sliding block coding for transmission of a source over a stationary nonanticipatory channel
Inf. and Control 35 (1977), p. 1

[16] KLEINROCK L. : Queuing systems , vol. 2 (computer applications)
John Wiley & Sons, 1976

[17] LAM S.S. : Packet switching in a multi-access broadcast channel with applications to satellite communication in a computer network,
Report UCLA-ENG 7429, mars 1974

[18] LUGANNANI R. : Intersymbol interference and probability of error in digital systems
IEEE IT 15, n° 6 (1969), p. 682

[19] MARTON K. : Error exponent for source coding with a fidelity criterion
IEEE IT 20, n° 2 (1974), p. 197

[20] ORNSTEIN D. : Ergodic theory, randomness, and dynamical systems
Yale Mathematical monographs n° 5 (1974)

[21] PARTHASARATHY K. : On the integral representation of the rate of transmission of a stationary channel,
Illinois J. Math. 5 (1961), p. 299

[22] QURESHI S., NEWHALL E. : An adaptive receiver for data transmission over time-dispersive channels
IEEE IT 19, n° 4 (1973), p. 448

[23] SHANNON C., GALLAGER R., BERLEKAMP E. : Lower bounds to error probability for coding on discrete memoryless channels I et II
Inf. and Control 10 (1967), p. 65 et p. 522

[24] SINGH S., KAMBO N. : Source code error bound in the excess rate region
IEEE IT 23, n° 1 (1977), p. 65

[25] VAN DER MEULEN E. : A survey of multi-way channels in formation theory : 1961-1976
IEEE IT 23, n° 1 (1977), p. 1

Gabriel RUGET

Mathématiques
E.R.A. CNRS 532 "Statistique Appliquée"
Université Paris-Sud
91405 ORSAY

(*) En deux ans, cet exposé a beaucoup vieilli. Il faut absolument prendre connaissance des méthodes "universelles" de codage-décodage dans Csiszar I. et J. Körner : Graph decomposition : a new key to coding theorem, IEEE International Symposium on Information theory, Grignano, 1979.

Sur les réseaux de communication, voir Csiszar I. et J. Körner : Towards a general theory of source networks, même référence, ou écrire au Math. Institute of the Hungarian Academy of Sciences.