Estimation de l'erreur d'un signal expérimental moyenné

V. CROQUETTE

Janvier 2009

Résumé

Le but de ce document est de vous faire réfléchir sur un cas concret de traitement du signal si possible avant le cours. Ce n'est pas un TD et il ne s'agit pas de faire un calcul noté (presque tout les calculs sont déjà fait), mais plutôt de vous confronter à un situation pratique où l'application du théorème de la limite centrale n'est pas directe et où il faut un peu comprendre les concepts avant de les appliquer. Nous nous proposons d'évaluer l'erreur sur un signal expérimental moyenné sur plusieurs échantillons. Le principe est évidemment d'utiliser le théorème de la limite centrale, ce qui nous donnera l'occasion d'en revoir ses fondements et les conditions de son applicabilité. Nous allons considérer deux types de signaux : l'un provenant d'une expérience, l'autre d'un générateur de bruit aléatoire.

1 Présentation des signaux

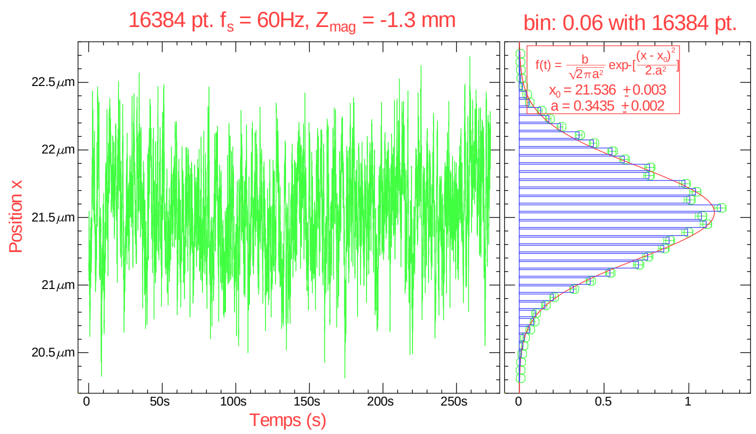

FIGURE 1 – *A gauche, signal expérimental : position $x(t)$ d'une bille soumise à une perturbation brownienne au cours du temps. Ce signal contient 16384 points, il est échantillonné à 60Hz. A droite, histogramme de ce signal indiquant que la distribution des positions est gaussienne.*

Le signal expérimental que nous utiliserons est caractérisé par un bruit important. Il correspond ici à la mesure de la position $x(t)$ d'une bille magnétique de $1\mu m$ de diamètre évoluant dans l'eau et attachée à une seule molécule d'ADN de $16\mu m$ de longueur. Elle est tirée par une force magnétique de quelques pN suivant la direction z. Ce système se comporte comme un pendule amorti, le mouvement Brownien fait bouger la bille de façon aléatoire tandis que la molécule d'ADN ramène la bille vers sa position d'équilibre. Cependant, ce que nous allons discuter sur ce signal sera valide pour la plupart des signaux expérimentaux. Notre signal expérimental a une valeur moyenne de $21.53\mu m$ et un écart type de $0.34\mu m$.

Nous nous intéresserons également à un signal simulé (Fig. 2) par un ordinateur présentant des caractéristiques analogues à notre signal expérimental généré en utilisant les fonctions aléatoires de distribution gaussienne de *numerical recipe*.

2 Moyenner pour diminuer le bruit

Très souvent on cherche à réduire le bruit d'une expérience en moyennant le signal. Nous nous proposons de moyenner les points de nos signaux par paquets et de prédire l'écart type de ce ces séries.

Si $x_0(i)$ est le signal orignal avec $i \in [0, N]$ et $\sigma_0 = 0.34\mu m$, notre série moyennée sur k points s'écrit : $x_k(i) = \sum_{ki}^{k(i+1)} x_0(j)/k$ avec $i \in [0, N/k]$, son écart type est noté σ_k.

– Rappeler ce que nous indique le théorème de la limite centrale, et les conditions dans lesquels il est applicable.

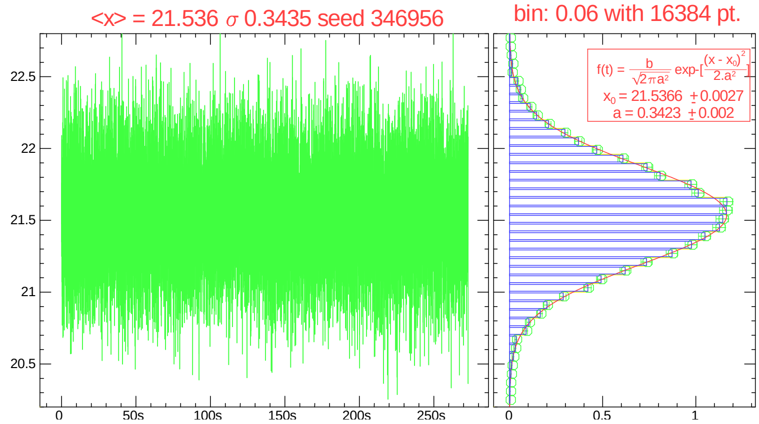

FIGURE 2 – *A gauche, signal simulé ayant la même valeur moyenne et le même écart type que le signal expérimental. A droite, histogramme du signal de gauche indiquant que la distribution des positions est gaussienne.*

– En utilisant le théorème de la limite centrale, indiquez la valeur de σ_k en particulier pour $k = 64$.

La mesure de σ_{64} faite sur la série des 256 points de $x_{64}(i)$ donne les résultats suivants : Pour le signal expérimental $\sigma_{E64} = 0.249\mu m$, pour le signal simulé $\sigma_{S64} = 0.0434\mu m$.

– Comprenez-vous ces nombres au regard du théorème de la limite centrale ? Celui-ci ne semble pas s'appliquer à notre signal expérimental, pourquoi ?

3 Comportement de σ_k en fonction de k

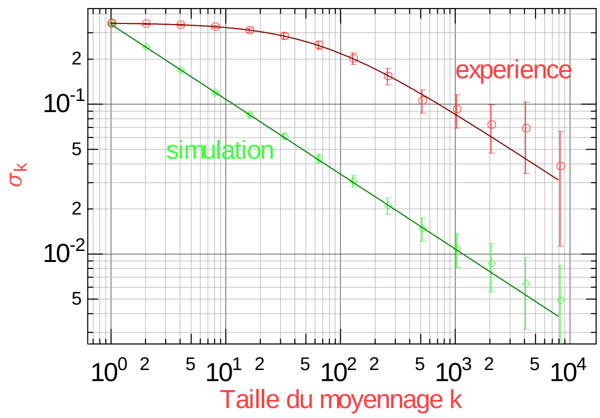

FIGURE 3 – *Evolution de l'écart type en fonction du nombre de points sur lequel est fait le moyennage. Les points correspondent au calcul de l'écart type pour les signaux expérimentaux et simulés. Le trait plein ajusté aux points du signal simulé correspond à* $\sigma_k = \sigma_0/\sqrt{k}$.

La figure 3 présente l'évolution de $\sigma_k(k)$ pour les deux types de signaux avec $k = 2^j$. Pour le signal simulé issu du générateur de nombres aléatoires, nous retrouvons le comportement prédit par le théorème de la limite centrale.

– Combien reste-t-il de points $x_k(i)$ pour la plus grande valeur de k ?
– Pour le signal expérimental, nous observons un changement de comportement lorsque $b = n_c \approx 64$, à quoi correspond ce nombre ?
– Que ce passe-t-il quand $k << nc$ et quand k devient grand pour le signal expérimental ?
– Proposer une interprétation de cette courbe.
– Quelle est la forme de la fonction qui décrit les points du signal expérimental ?
– Comprennez vous l'origine physique de n_c ?