Fluorescence Imaging Analysis: The Case of Calcium Transients.

Christophe Pouzat

Mathématiques Appliquées à Paris 5 (MAP5)

Université Paris-Descartes and CNRS UMR 8145

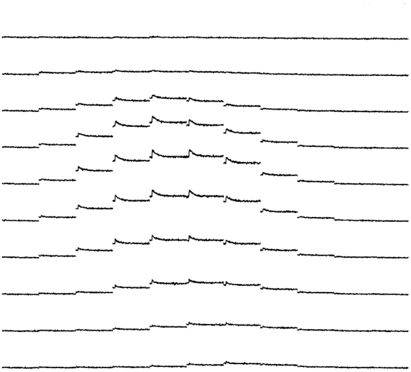
christophe.pouzat@parisdescartes.fr

ENP Course: June 17 2015

Outline

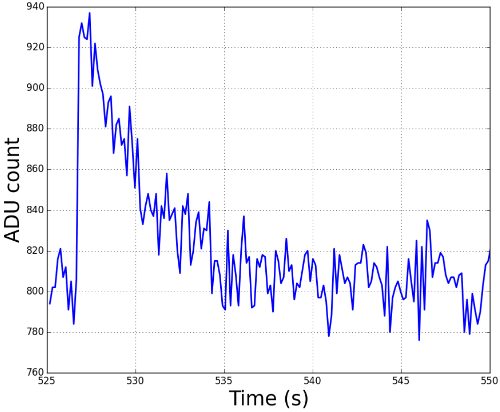
Where are we ?

The variability inherent to fluorescence imaging data (1)



ADU counts (raw data) from Fura-2 excited at 340 nm. Each square corresponds to a pixel. 25.05 s of data are shown. Same scale on each sub-plot. Data recorded by Andreas Pippow (Kloppenburg Lab. Cologne University).

The variability inherent to fluorescence imaging data (2)



One of the central pixels of the previous figure.

What do we want? (1)

Given the data set illustrated on the last two slides we might want to estimate parameters like:

- the peak amplitude
- the decay time constant(s)
- the baseline level
- the whole time course (strictly speaking, a function).

What do we want? (2)

If we have a model linking the calcium dynamics—the time course of the free calcium concentration in the cell—to the fluorescence intensity like:

$$\frac{\mathrm{d}Ca_t}{\mathrm{dt}} \left(1 + \kappa_F(Ca_t) + \kappa_E(Ca_t)\right) + \frac{j(Ca_t)}{v} = 0\,,$$

where Ca_t stands for $[Ca^{2+}]_{free}$ at time t, v is the volume of the neurite—within which diffusion effects can be neglected—and

$$j(Ca_t) \equiv \gamma(Ca_t - Ca_{steady})\,,$$

is the model of calcium extrusion—Ca_{steady} is the steady state $[Ca^{2+}]_{free}$—

$$\kappa_F(Ca_t) \equiv \frac{F_{total}\,K_F}{(K_F + Ca_t)^2} \quad \text{and} \quad \kappa_E(Ca_t) \equiv \frac{E_{total}\,K_E}{(K_E + Ca_t)^2}\,,$$

where F stands for the fluorophore en E for the *endogenous* buffer.

What do we want? (3)

In the previous slide, assuming that the fluorophore (Fura) parameters: F_{total} and K_F have been calibrated, we might want to estimate:

- the extrusion parameter: γ
- the endogenous buffer parameters: E_{total} and K_E

using an equation relating measured fluorescence to calcium:

$$Ca_t = K_F \, \frac{S_t - S_{min}}{S_{max} - S_t} \, ,$$

where S_t is the fluorescence (signal) measured at time t, S_{min} and S_{max} are *calibrated* parameters corresponding respectively to the fluorescence in the absence of calcium and with saturating $[Ca^{2+}]$ (for the fluorophore).

What do we want? (4)

- The variability of our signal—meaning that under replication of our measurements *under the exact same conditions* we wont get the exact same signal—implies that our estimated parameters will also fluctuate upon replication.
- Formally our parameters are modeled as *random variables* and it is not enough to summarize a random variable by a single number.
- If we cannot get the full distribution function for our parameters, we want to give at least ranges within which the true value of the parameter should be found with a given probability.
- In other words: an analysis without confidence intervals is not an analysis, it is strictly speaking useless since it can't be reproduced—if I say that my time constant is 25.76 ms the probability that upon replication I get again 25.76 is essentially 0; if I say that the actual time constant has a 0.95 probability to be in the interval [24,26.5], I can make a comparison with replications.

A proper handling of the "variability" matters (1)

Let us consider a simple data generation model:

$$Y_i \sim \mathcal{P}(f_i)\,, \quad i = 0, 1, \ldots, K \ ,$$

where $\mathcal{P}(f_i)$ stands for the *Poisson distribution* with parameter f_i :

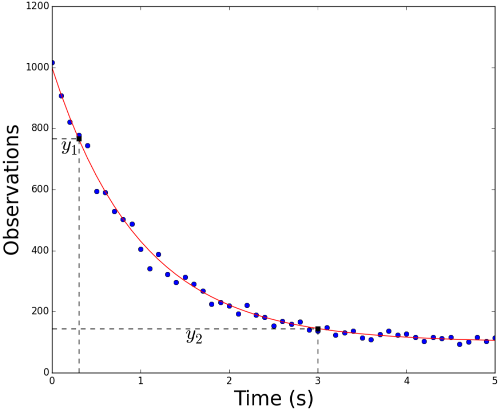
$$\mathrm{Pr}\{Y_i = n\} = \frac{(f_i)^n}{n!} \exp(-f_i)\,, \quad \text{for} \quad n = 0, 1, 2, \ldots$$

and

$$f_i = f(\delta i | f_\infty, \Delta, \beta) = f_\infty + \Delta \exp(-\beta\, \delta i) \ ,$$

δ is a time step and f_∞, Δ and β are model parameters.

A proper handling of the "variability" matters (2)



Data simulated according to the previous model. We are going to assume that f_∞ and Δ are known and that (t_1, y_1) and (t_2, y_2) are given. We want to estimate β.

Two estimators (1)

We are going to consider two estimators for β:

- The "classical" least square estimator:

$$\tilde{\beta} = \arg\min \tilde{L}(\beta) \, ,$$

where

$$\tilde{L}(\beta) = \sum_j \left(y_j - f(t_j \mid \beta)\right)^2 \, .$$

- The least square estimator applied to the *square root* of the data:

$$\hat{\beta} = \arg\min \hat{L}(\beta) \, ,$$

where

$$\hat{L}(\beta) = \sum_j \left(\sqrt{y_j} - \sqrt{f(t_j \mid \beta)}\right)^2 \, .$$

Two estimators (2)

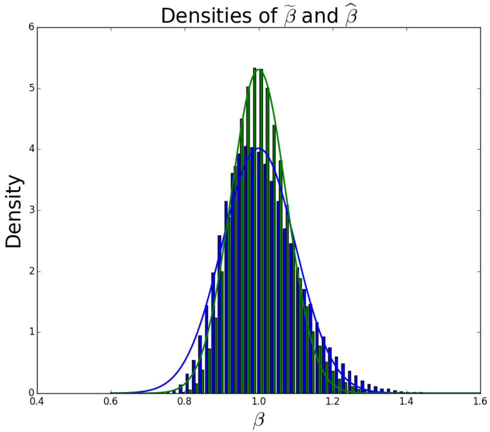
We perform an empirical study as follows:

- We simulate 100,000 experiments such that:

  $$(Y_1, Y_2) \sim \left(\mathcal{P}(f(0.3|\beta_0), \mathcal{P}(f(3|\beta_0)\right) ,$$

  with $\beta_0 = 1$.
- For each simulated pair, $(y_1, y_2)^{[k]}$ $(k = 1, \ldots, 10^5)$, we minimize $\tilde{L}(\beta)$ and $\hat{L}(\beta)$ to obtain: $(\tilde{\beta}^{[k]}, \hat{\beta}^{[k]})$.
- We build histograms for $\tilde{\beta}^{[k]}$ and $\hat{\beta}^{[k]}$ as density estimators of our estimators.

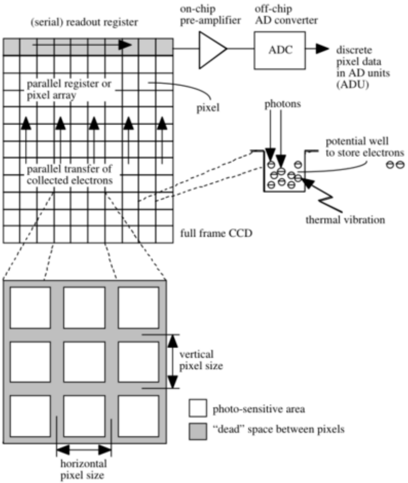
Two estimators (3)



Both histograms are built with 50 bins. $\hat{\beta}$ is clearly better than $\tilde{\beta}$ since its variance is smaller. The derivation of the theoretical (large sample) densities is given in Joucla et al (2010).

Where are we ?

CCD basics



Source: L. van Vliet et col. (1998) Digital Fluorescence Imaging Using Cooled CCD Array Cameras (figure 3).

"Noise" sources in CCD (1)

- The "Photon noise" or "shot noise" arises from the fact the measuring a fluorescence intensity, λ, implies counting photons—unless one changes the laws of Physics there is nothing one can do to eliminate this source of variability (improperly called "noise")—:

$$\mathrm{Pr}\{N = n\} = \frac{\lambda^n}{n!} \exp -\lambda \,, \quad n = 0, 1, \ldots \,, \quad \lambda > 0 \,.$$

- The "thermal noise" arises from thermal agitation which "dumps" electrons in potential wells; this "noise" also follows a Poisson distribution but it can be made negligible by *cooling down* the camera.

"Noise" sources in CCD (2)

- The "read out noise" arises from the conversion of the number of photo-electrons into an equivalent tension; it follows a normal distribution whose variance is independent of the mean (as long as reading is not done at too high a frequency).
- The "digitization noise" arises from the mapping of a continuous value, the tension, onto a grid; it is negligible as soon as more than 8 bit are used.

A simple CCD model (1)

- We can easily obtain a simple CCD model taking into account the two main "noise" sources (photon and read-out).
- To get this model we are going the fact (a theorem) that when a large number of photon are detected, the Poisson distribution is well approximated by (converges in distribution to) a normal distribution with identical mean and variance:

$$\mathrm{Pr}\{N = n\} = \frac{\lambda^n}{n!} \exp -\lambda \approx \mathcal{N}(\lambda, \lambda) \; .$$

- In other words:

$$N \approx \lambda + \sqrt{\lambda}\,\epsilon \; ,$$

where $\epsilon \sim \mathcal{N}(0, 1)$ (follows a standard normal distribution).

A simple CCD model (2)

- A read-out noise is added next following a normal distribution with 0 mean and variance σ_R^2.
- We are therefore adding to the random variable N a new independent random variable $R \sim \mathcal{N}(0, \sigma_R^2)$ giving:

$$M \equiv N + R \approx \lambda + \sqrt{\lambda + \sigma_R^2}\,\epsilon \;,$$

where the fact that the sum of two independent normal random variables is a normal random variable whose mean is the sum of the mean and whose variance is the sum of the variances has been used.

A simple CCD model (3)

- Since the capacity of the photo-electron weels is finite (35000 for the camera used in the first slides) and since the number of photon-electrons will be digitized on 12 bit (4096 levels), a "gain" G smaller than one must be applied if we want to represent faithfully (without saturation) an almost full well.
- We therefore get:

$$Y \equiv G \cdot M \approx G\,\lambda + \sqrt{G^2\,(\lambda + \sigma_R^2)}\,\epsilon\,.$$

For completeness: Convergence in distribution of a Poisson toward a normal rv (1)

We use the moment-generating function and the following theorem (*e.g.* John Rice, 2007, *Mathematical Statistics and Data Analysis*, Chap. 5, Theorem A):

- If the moment-generating function of each element of the rv sequence X_n is $m_n(t)$,
- if the moment-generating function of the rv X is $m(t)$,
- if $m_n(t) \to m(t)$ when $n \to \infty$ for all $|t| \leq b$ where $b > 0$
- then $X_n \xrightarrow{D} X$.

For completeness: Convergence in distribution of a Poisson toward a normal rv (2)

Lets show that:

$$Y_n = \frac{X_n - n}{\sqrt{n}} \ ,$$

where X_n follows a Poisson distribution with parameter n, converges in distribution towards Z standard normal rv.
We have:

$$m_n(t) \equiv \mathrm{E}\left[\exp(Y_n t)\right] \ ,$$

therefore:

$$m_n(t) = \sum_{k=0}^{\infty} \exp\left(\frac{k-n}{\sqrt{n}} t\right) \frac{n^k}{k!} \exp(-n) \ ,$$

For completeness: Convergence in distribution of a Poisson toward a normal rv (3)

$$m_n(t) = \exp(-n)\exp(-\sqrt{n}t)\sum_{k=0}^{\infty}\frac{\left(n\exp\left(t/\sqrt{n}\right)\right)^k}{k!}$$

$$m_n(t) = \exp\left(-n - \sqrt{n}t + n\exp(t/\sqrt{n})\right)$$

$$m_n(t) = \exp\left(-n - \sqrt{n}t + n\sum_{k=0}^{\infty}\left(\frac{t}{\sqrt{n}}\right)^k\frac{1}{k!}\right)$$

$$m_n(t) = \exp\left(-n - \sqrt{n}t + n + \sqrt{n}t + \frac{t^2}{2} + n\sum_{k=3}^{\infty}\left(\frac{t}{\sqrt{n}}\right)^k\frac{1}{k!}\right)$$

$$m_n(t) = \exp\left(\frac{t^2}{2} + n\sum_{k=3}^{\infty}\left(\frac{t}{\sqrt{n}}\right)^k\frac{1}{k!}\right)$$

For completeness: Convergence in distribution of a Poisson toward a normal rv (4)

We must show:

$$n \sum_{k=3}^{\infty} \left(\frac{t}{\sqrt{n}} \right)^k \frac{1}{k!} \to_{n\to\infty} 0 \quad \forall\ |t| \leq b, \quad \text{where} \quad b > 0\,,$$

since $\exp(-t^2/2)$ is the moment-generating function of a standard normal rv. But

$$\left| n \sum_{k=3}^{\infty} \left(\frac{t}{\sqrt{n}} \right)^k \frac{1}{k!} \right| \to_{n\to\infty} 0 \quad \forall\ |t| \leq b, \quad \text{where} \quad b > 0$$

implies that since

$$- \left| n \sum_{k=3}^{\infty} \left(\frac{t}{\sqrt{n}} \right)^k \frac{1}{k!} \right| \leq n \sum_{k=3}^{\infty} \left(\frac{t}{\sqrt{n}} \right)^k \frac{1}{k!} \leq \left| n \sum_{k=3}^{\infty} \left(\frac{t}{\sqrt{n}} \right)^k \frac{1}{k!} \right| .$$

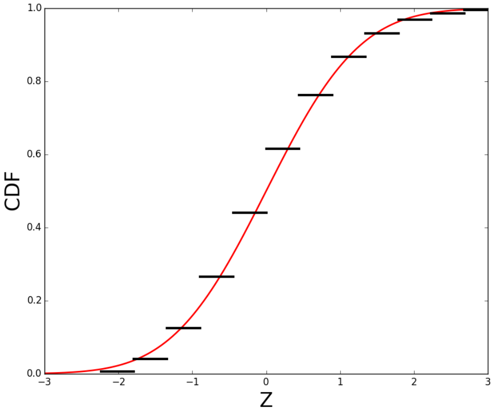
For completeness: Convergence in distribution of a Poisson toward a normal rv (5)

But for all $|t| \leq b$ where $b > 0$

$$
\begin{aligned}
0 \leq \left| n \sum_{k=3}^{\infty} \left(\frac{t}{\sqrt{n}}\right)^k \frac{1}{k!} \right| &\leq n \sum_{k=3}^{\infty} \left(\frac{|t|}{\sqrt{n}}\right)^k \frac{1}{k!} \\
&\leq \frac{|t|^3}{\sqrt{n}} \sum_{k=0}^{\infty} \left(\frac{|t|}{\sqrt{n}}\right)^k \frac{1}{(k+3)!} \\
&\leq \frac{|t|^3}{\sqrt{n}} \sum_{k=0}^{\infty} \left(\frac{|t|}{\sqrt{n}}\right)^k \frac{1}{k!} \\
&\leq \frac{|t|^3}{\sqrt{n}} \exp\left(\frac{|t|}{\sqrt{n}}\right) \rightarrow_{n\rightarrow\infty} 0\,,
\end{aligned}
$$

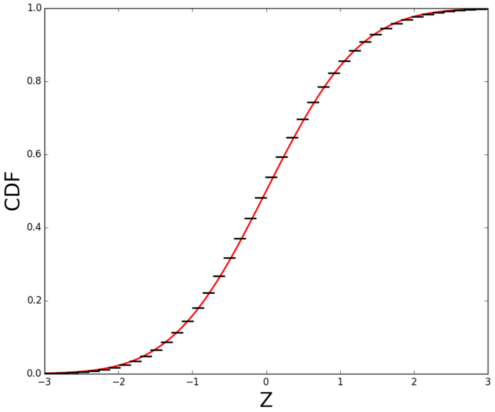
which completes the proof.

For completeness: Convergence in distribution of a Poisson toward a normal rv (6)



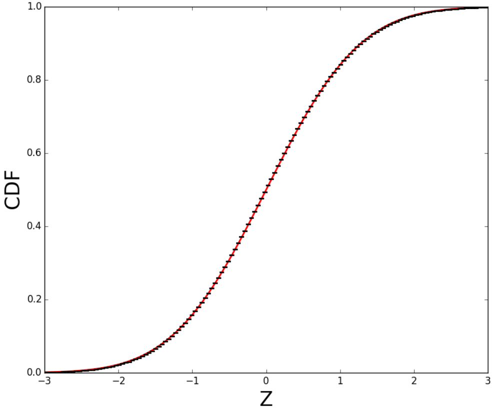
Cumulative distribution functions (CDF) of Y_5 (black) and Z a standard normal (red).

For completeness: Convergence in distribution of a Poisson toward a normal rv (7)



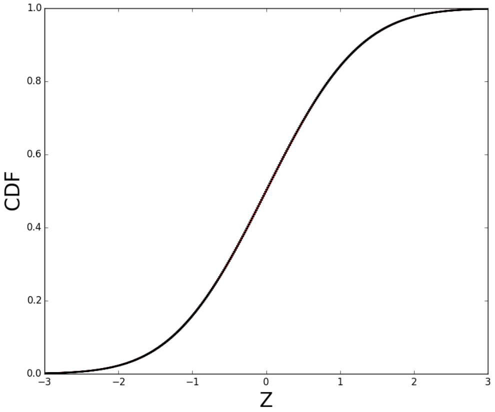
Cumulative distribution functions (CDF) of Y_{50} (black) and Z a standard normal (red).

For completeness: Convergence in distribution of a Poisson toward a normal rv (8)



Cumulative distribution functions (CDF) of Y_{500} (black) and Z a standard normal (red).

For completeness: Convergence in distribution of a Poisson toward a normal rv (9)



Cumulative distribution functions (CDF) of Y_{5000} (black) and Z a standard normal (red).

Where are we ?

CCD calibration (1)

If what I just exposed is correct, with the two (main) "noise" sources, the observations Y (from a CCD pixel) follow:

$$Y \sim G\,\lambda + \sqrt{G^2\,(\lambda + \sigma_R^2)}\,\epsilon\;,$$

where G is the camera gain, σ_R^2 is the read-out variance and ϵ is a standard normal rv. The values of G and σ_R^2 are specified by the manufacturer for each camera, but experience shows that manufacturers tend to be overoptimistic when it comes to their product performances—they can for instance give an underestimated σ_R^2. Its therefore a good idea to measure these parameters with calibration experiments. Such calibration experiments are also the occasion to check that our simple model is relevant.

CCD calibration (2)

- Our problem becomes: How to test $Y \sim G\,\lambda + \sqrt{G^2\,(\lambda + \sigma_R^2)}\,\epsilon$? Or how to set different values for λ?
- Let's consider a pixel of our CCD "looking" at a fixed volume of a fluorescein solution with a given (and stable) concentration. We have two ways of modifying λ :
  - Change the intensity i_e of the light source exciting the fluorophore.
  - Change the exposure time τ.

CCD calibration (3)

We can indeed write our λ as:

$$\lambda = \phi v c i_e \tau \,,$$

where

- v is the solution's volume "seen" by a given pixel,
- c is the fluorophore's concentration,
- ϕ is the quantum yield.

In practice it is easier to vary the exposure time τ and that's what was done in the experiments described next... Question: Can you guess what these experiments are?

CCD calibration (4)

Sebastien Joucla and myself asked our collaborators from the Kloppenburg lab (Cologne University) to:

- choose 10 exposure times,
- for each of the 10 times, perform 100 exposures,
- for each of the 10 x 100 exposures, record the value y_{ij} of the rv Y_{ij} of CCD's pixel i, j.

We introduce a rv Y_{ij} for each pixel because it is very difficult (impossible) to have a uniform intensity (i_e) and a uniform volume (v) and a uniform quantum yield (ϕ). We have therefore for each pixel:

$$Y_{i,j} \sim G\, p_{i,j}\tau + \sqrt{G^2\left(p_{i,j}\tau + \sigma_R^2\right)}\,\epsilon_{i,j} \ ,$$

where $p_{i,j} = c\phi_{i,j} v_{i,j} i_{e,i,j}$.

CCD calibration (5)

- If our model is correct we should have for each pixel i, j, for a given exposure time, a mean value:

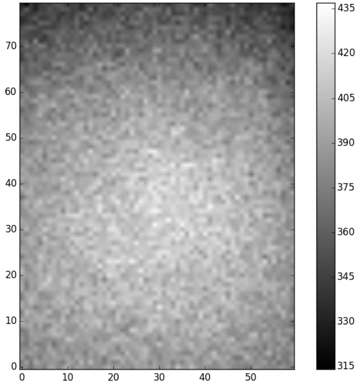
$$\bar{y}_{i,j} = \frac{1}{100}\sum_{k=1}^{100} y_{i,j,k} \approx G\, p_{i,j}\tau$$

- and a variance:

$$S^2_{i,j} = \frac{1}{99}\sum_{k=1}^{100}(y_{i,j,k} - \bar{y}_{i,j})^2 \approx G^2\left(p_{i,j}\tau + \sigma^2_R\right) .$$

- The graph of $S^2_{i,j}$ *vs* $\bar{y}_{i,j}$ should be a straight line with slope G ordinate at 0, $G^2\sigma^2_R$.

CCD calibration (6)

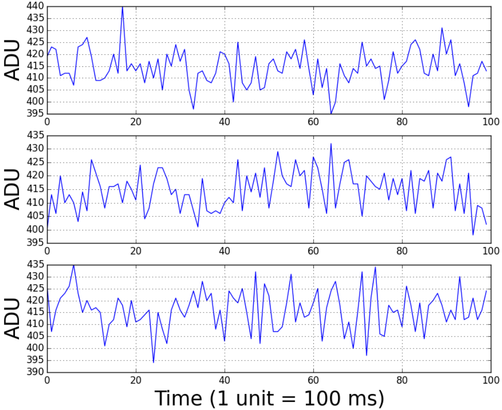


The first exposure of 10 ms (experiment performed by Andreas Pippow, Kloppenburg lag, Cologne University).

CCD calibration: Checking the assumptions (1)

- The data are going to be analyzed as if the $Y_{i,j,k}$ were IID, but they were sequentially recorded. It is therefore strongly recommended to check that the IID hypothesis is reasonable.
- The small example of the next figure shows that there are no (obvious) trends.
- We must also check the correlation function.

CCD calibration: Checking the assumptions (2)

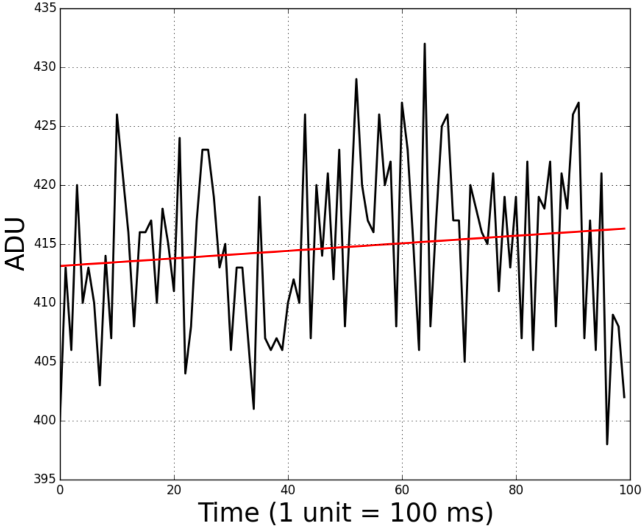


Counts time evolution for three neighboring pixels (10 ms exposure time).

CCD calibration: Checking the assumptions (3)

- If the $Y_{i,j,k}$ are not IID we expect a more or less linear trend—due to bleaching of the dye.
- Rather then looking at each individual pixel sequence like on the previous slide, we can fit the following linear model model to each pixel:
$$Y_{i,j,k} = \beta_0 + \beta_1 k + \sigma \epsilon_{i,j}$$
where the $\epsilon_{i,j} \overset{IID}{\sim} \mathcal{N}(0,1)$, and check if β_1 can be reasonably considered as null; while a trend due to bleaching would give a negative β_1.
- Without a trend, the theoretical distribution of $\hat{\beta}_1/\hat{\sigma}_{\beta_1}$—$\hat{\beta}_1$ is the estimate of β_1 and $\hat{\sigma}_{\beta_1}$ its estimated standard error—is a Student's t distribution with 98 degrees of freedom.
- Applying this idea to the central pixel of the previous slide we get...

CCD calibration: Checking the assumptions (4)



We get $\hat{\beta}_1 = 0.032$ and a 95 % conf. int. for it is: $[-0.018, 0.082]$.

CCD calibration: Checking the assumptions (5)

We can use the fact that, under the null hypothesis (no trend):

$$\hat{\beta}_1/\hat{\sigma}_{\beta_1} \sim t_{98}$$

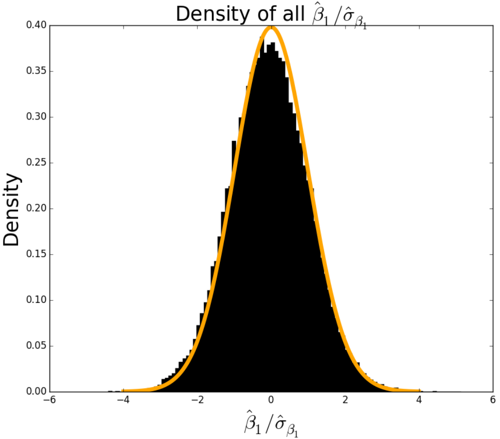
by constructing the empirical cumulative distribution function (ECDF) of the 60 x 80 pixels at each exposure time to get the maximal difference (in absolute value) with the theoretical CDF to apply a Kolmogorov test. The critical value of the latter for a 99% level and a sample size of 100 is 0.163. We get the following values:

100ms	10ms	20ms	30ms	40ms
0.09	0.089	0.116	0.058	0.135

50ms	60ms	70ms	80ms	90ms
0.209	0.041	0.178	0.153	0.07

The values at 50 and 70 ms are too large.

CCD calibration: Checking the assumptions (6)

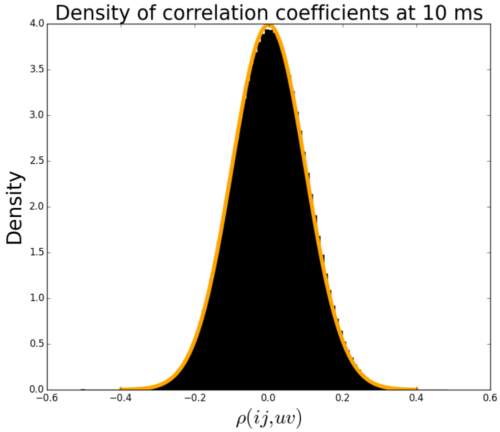


Empirical density in black, theoretical one (t with 98 df) in orange.

CCD calibration: Checking the assumptions (7)

- We now look for potential correlations between recording from different pixels.
- We do that by computing the empirical correlation between pixels (i,j) and (u,v).
- We get the empirical mean at each pixel (for a given exposure time) that is: $\overline{Y}_{ij} = (1/K)\sum_{k=1}^{K} Y_{ijk}$.
- We get the empirical variance: $S_{ij}^2 = 1/(K-1)\sum_{k=1}^{K}(Y_{ijk} - \overline{Y}_{ij})^2$.
- We then obtain the normalized signal or *standard score*: $N_{ijk} = (Y_{ijk} - \overline{Y}_{ij})/\sqrt{S_{ij}^2}$.
- The correlation coefficient is then: $\rho(ij,uv) = 1/(K-1)\sum_{k=1}^{K} N_{ijk}N_{uvk}$.
- Under the null hypothesis, no correlation, $\rho(ij,uv) \sim \mathcal{N}(0, 1/K)$.

CCD calibration: Checking the assumptions (8)



Empirical density in black, theoretical one, $\mathcal{N}(0, 0.01)$, in orange.

CCD calibration: Checking the assumptions (9)

The empirical variance (x 100 and rounded to the third decimal) of the samples of correlation coefficients (1 sample per exposure duration) are:

100ms	10ms	20ms	30ms	40ms
1.009	1.01	1.009	1.01	1.01

50ms	60ms	70ms	80ms	90ms
1.01	1.009	1.01	1.009	1.009

Overall our IID modeling assumption is met with perhaps the exceptions of the 50 and 70 ms exposure times for the drift.

CCD calibration (5): again

We wrote previously :

- If our model is correct we should have for each pixel i, j, for a given exposure time, a mean value:

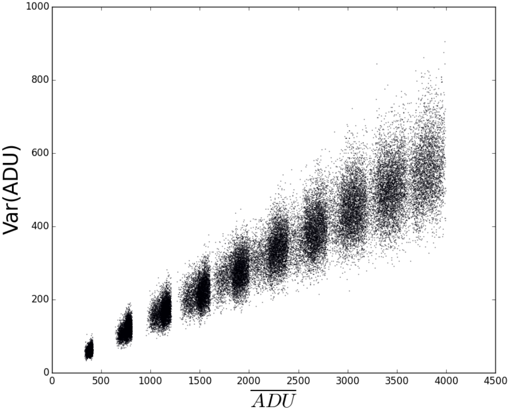
$$\bar{y}_{i,j} = \frac{1}{100} \sum_{k=1}^{100} y_{i,j,k} \approx G \, p_{i,j} \tau$$

- and a variance:

$$S_{i,j}^2 = \frac{1}{99} \sum_{k=1}^{100} (y_{i,j,k} - \bar{y}_{i,j})^2 \approx G^2 \left(p_{i,j}\tau + \sigma_R^2\right) .$$

- The graph of $S_{i,j}^2$ *vs* $\bar{y}_{i,j}$ should be a straight line with slope G ordinate at 0, $G^2\sigma_R^2$.

CCD calibration (7): $S^2_{i,j}$ vs $\bar{y}_{i,j}$



We do see the expected linear relation:
$\mathrm{Var}[ADU] = G^2\sigma_R^2 + G\mathrm{E}[ADU]$.

CCD calibration (8): Linear fit

The heteroscedasticity (inhomogeneous variance) visible on the graph is also expected since the variance of a variance for an IID sample of size K from a normal distribution with mean μ and variance σ^2 is:

$$\mathrm{Var}[S^2] = \frac{2\sigma^4}{(K-1)} \, .$$

- This means than when we do our linear fit,

  $$y_k = a + b x_k + \sigma_k \epsilon_k \, ,$$

  we should use weights.
- Here

  $$x_k = \overline{ADU}_k \quad y_k = \mathrm{Var}[ADU]_k \, ,$$
  $$b = G \quad a = G^2 \sigma_R^2 \, .$$

CCD calibration (9): Linear fit

- It's easy to show that the least square estimates are:

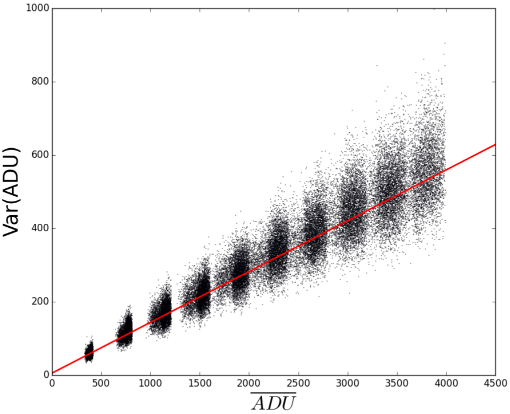
$$\hat{a} = \frac{1}{Z} \sum_k \frac{y_k - \hat{b} x_k}{\sigma_k^2} \quad \text{where} \quad Z = \sum_k \frac{1}{\sigma_k^2}$$

and

$$\hat{b} = \left(\sum_k \frac{x_k}{\sigma_k^2} \left(y_k - \frac{1}{Z} \sum_j \frac{y_j}{\sigma_j^2} \right) \right) / \left(\sum_k \frac{x_k}{\sigma_k^2} \left(x_k - \frac{1}{Z} \sum_j \frac{x_j}{\sigma_j^2} \right) \right) .$$

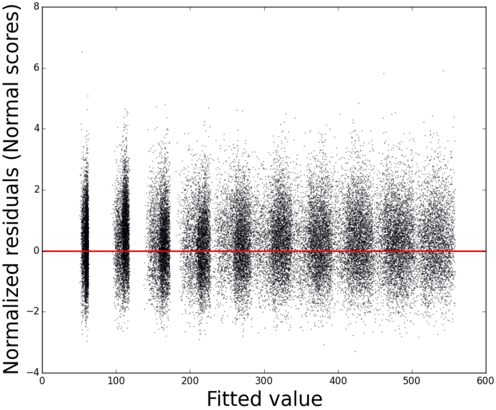
- We don't know σ_k but we have an estimation: $\hat{\sigma}_k^2 = \mathrm{Var}[S_k^2]$ we can "plug-in" this value to get our weights.

CCD calibration (10): Linear fit



We have here $\hat{G} = 0.14$ and $\hat{\sigma}_R^2 = 290$.

CCD calibration (11): Checking the fit



CCD calibration (12): Some remarks

- When we use a linear regression, we are (implicitly) assuming that the "independent" variable, here $\overline{ADU}_k$, is *exactly* known.
- This was clearly not the case here since $\overline{ADU}_k$ was measured (with an error).
- We could and will therefore refine our fit.

Where are we ?

Error propagation

- Let us consider two random variables: Y and Z such that:
- $Y \approx \mathcal{N}(\mu_Y, \sigma_Y^2)$ or $Y \approx \mu_Y + \sigma_Y \, \epsilon$
- $Z = f(Y)$, with f continuous and differentiable.
- Using a first order Taylor expansion we then have:

$$\begin{aligned} Z &\approx f(\mu_Y + \sigma_Y \, \epsilon) \\ &\approx f(\mu_Y) + \sigma_Y \, \epsilon \, \tfrac{df}{dY}(\mu_Y) \end{aligned}$$

- $\mathrm{E}Z \approx f(\mu_Y) = f(\mathrm{E}Y)$
- $\mathrm{Var}Z \equiv \mathrm{E}[(Z - \mathrm{E}Z)^2] \approx \sigma_Y^2 \, \tfrac{df}{dY}^2(\mu_Y)$
- $Z \approx f(\mu_Y) + \sigma_Y \left| \tfrac{df}{dY}(\mu_Y) \right| \epsilon$

Variance stabilization (1): Theory

- For our CCD model we have (for a given pixel):

$$Y \sim G\,\lambda + \sqrt{G^2\,(\lambda + \sigma_R^2)}\,\epsilon = \mu_Y + \sqrt{G\,\mu_Y + G^2\,\sigma_R^2}\,\epsilon\,.$$

- Then if $Z = f(Y)$ we get:

$$Z \approx f(\mu_Y) + \mid f'(\mu_Y) \mid G\sqrt{\mu_Y/G + \sigma_R^2}\,\epsilon$$

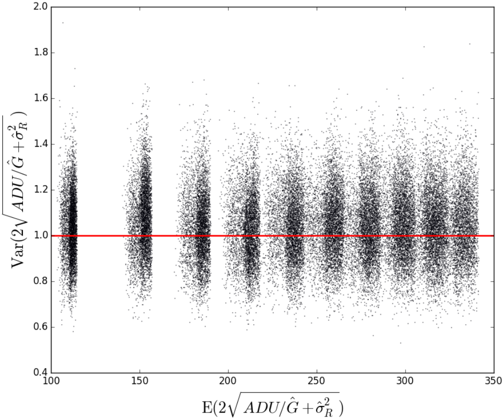
- What happens then if we take: $f(x) = 2\sqrt{x/G + \sigma_R^2}$?
- We have:

$$f'(x) = \frac{1}{G\sqrt{x/G + \sigma_R^2}}$$

- Leading to:

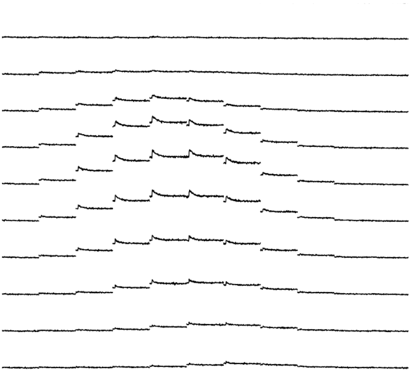
$$Z \approx 2\sqrt{\mu_Y/G + \sigma_R^2} + \epsilon$$

Variance stabilization (2): Example



Where are we ?

Back to where we started



ADU counts (raw data) from Fura-2 excited at 340 nm. Each square corresponds to a pixel. 25.05 s of data are shown. Same scale on each sub-plot. Data recorded by Andreas Pippow (Kloppenburg Lab. Cologne University).

Quick ROI detection (1): Motivation

- After variance stabilization: $Z_{i,j,k} = 2\sqrt{ADU_{i,j,k}/G + \sigma_R^2}$, the variance at each pixel (i,j) at each time, k, should be 1.
- If a pixel contains no dynamical signal—that is nothing more than a constant background signal—the following statistics:

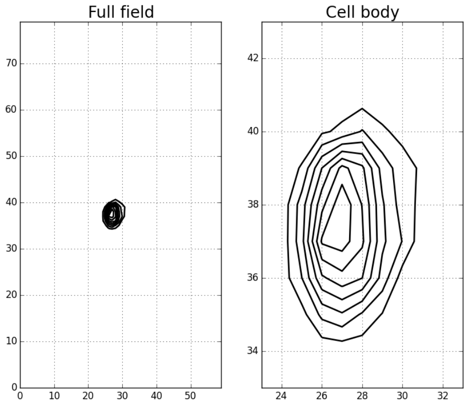
$$RSS_{i,j} \equiv \sum_{k=1}^{K} (Z_{i,j,k} - \overline{Z}_{i,j})^2 \quad \text{with} \quad \overline{Z}_{i,j} \equiv \frac{1}{K}\sum_{k=1}^{K} Z_{i,j,k}$$

  should follow a χ^2 distribution with $K-1$ degrees of freedom.
- We could therefore compute the values of the complementary cumulative distribution function of the theoretical χ^2_{K-1} distribution:

$$1 - F_{\chi^2_{K-1}}(RSS_{i,j})$$

  and look for very small values—that is very small probabilities—(using a log scale helps here).

Quick ROI detection (2)



Contour plots of $\log\left(1 - F_{\chi^2_{K-1}}(RSS_{i,j})\right)$

Pointwise time course estimation (1)

- We are going to be (very) conservative and keep as our ROI the pixels having an $\log\left(1 - F_{\chi^2_{K-1}}(RSS)\right) \leq -300$.
- We are then left with 12 pixels.
- We are going to model the fluorescence intensity of each of these pixels by:

  $$S_{i,j}(t) = \phi_{i,j}\, f(t) + b \ ,$$

  where $f(t)$ is a signal time course to all pixels of the ROI, $\phi_{i,j}$ is a pixel specific parameter and b is a background fluorescence assumed identical for each pixel.
- The time t is in fact a discrete variable, $t = \delta\, k$ (δ = 150 ms) and we are seeking a pointwise estimation: $\{f_1, f_2, \ldots, f_K\}$ (K = 168) where $f_k = f(\delta\, k)$.
- We end up with 12 ($\phi_{i,j}$) + 168 (f_k) + 1 (b) = 181 parameters for 12 x 168 = 2016 measurements.

Pointwise time course estimation (2)

- We need to add a constraint since with our model specification:

$$S_{i,j,k} = \phi_{i,j}\, f_k + b \,,$$

  we can multiply all the $\phi_{i,j}$ by 2 and divide all the f_k by 2 and get the same prediction.
- We are going to set $f_k = 1$ for the first 5 time points (the stimulation comes at the 11th) and our pointwise estimation relates to what is usually done with this type of data, $\Delta S(t)/S_0$ (where S_0 is a baseline average) through:

$$\Delta S(t)/S_0 = \frac{S(t) - S_0}{S_0} = f(t) - 1 + \mathrm{noise}\,.$$

- Notice that no independent background measurement is used.

Pointwise time course estimation (3)

- With variance stabilization we end up minimizing:

$$RSS\left(b, (\phi_{i,j}), (f_k)_{k=5,\ldots,168}\right) = \sum_{(i,j)\in \mathrm{ROI}} \sum_{k=1}^{168} \left(Z_{ijk} - F_{ijk}\right)^2 ,$$

where

$$Z_{ijk} = 2 * \sqrt{ADU_{ijk}/\hat{G} + \hat{\sigma}_R^2}$$

and

$$F_{ijk} = 2 * \sqrt{\phi_{i,j}\, f_k + b + \hat{\sigma}_R^2} .$$

- If our model is correct we should have:

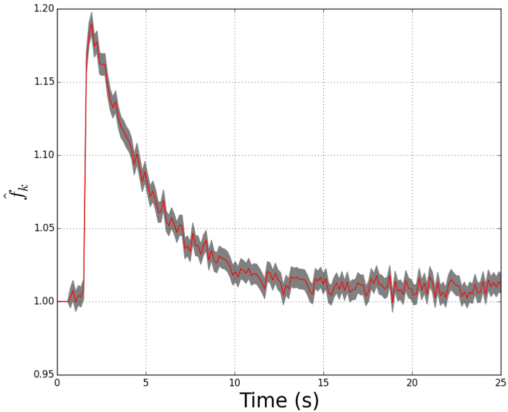
$$RSS\left(\hat{b}, (\hat{\phi}_{i,j}), (\hat{f}_k)_{k=5,\ldots,168}\right) \sim \chi^2_{12\times 168-176} .$$

- The method also generates confidence intervals for the estimated parameters.

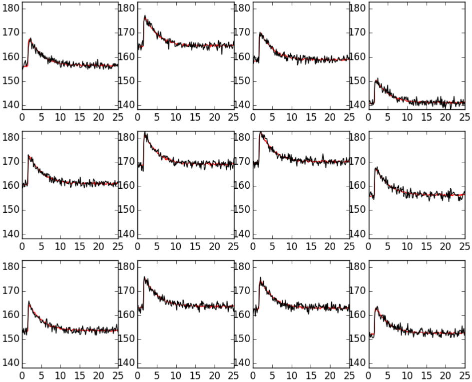
Pointwise time course estimation (4): Technical details

- To solve this 176 dimensional optimization problem in a reasonable time (< 10 s) in `Python` we use Newton's method with conjugate gradients for the inversion of the Hessian matrix.
- That means we have to define a function returning the gradient—vector of first derivatives—and the Hessian—matrix of second derivatives—of the `RSS` we just defined (that's a painful work).
- To improve numerical behavior, since all parameters are positive, we work with the log of the parameters.
- Giving all the details would be at least as long as the present talk, but they are fully disclosed in the source file of this talk that can be found on Github: `https://github.com/christophe-pouzat/ENP2015`.

Pointwise time course estimation (5): Time course estimate



Pointwise time course estimation (6): Data and fit



Data and fit after variance stabilization. The RSS is 1976 giving a probability of 0.986 (a bit large).

Thanks

This work was done in collaboration with:

- Sebastien Joucla
- Romain Franconville
- Andeas Pippow
- Peter Kloppenburg

Thank you for your attention!