Metti 5

Lecture L2 : Basics for linear inversion, the 'white box' case

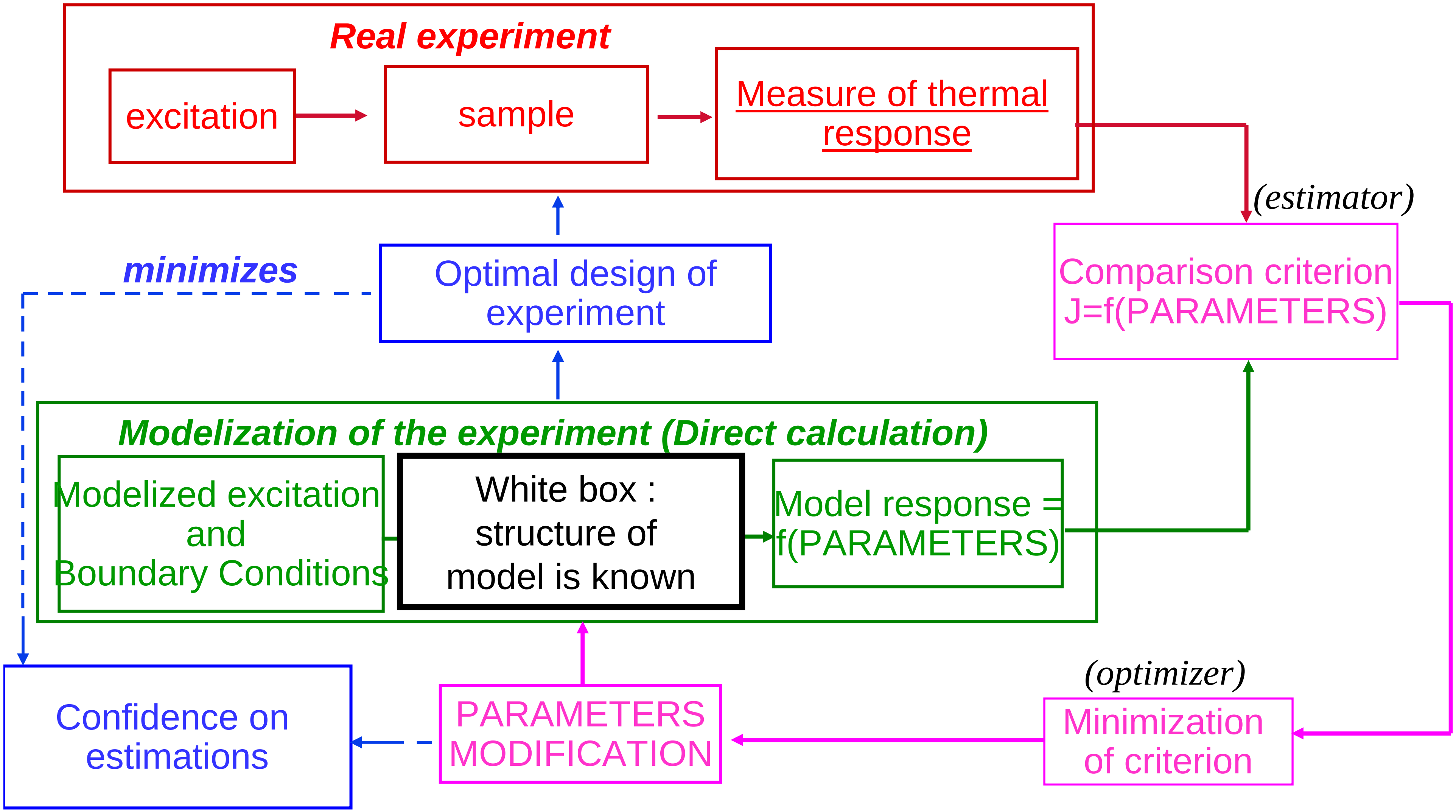
Fabrice Rigollet, Christophe Le Niliot
IUSTI UMR CNRS 6595
Marseille, France

Olivier Fudym
RAPSODEE FRE 3213 CNRS
Ecole des Mines d'Albi, France

Daniel Petit
PPRIME
Poitiers, France

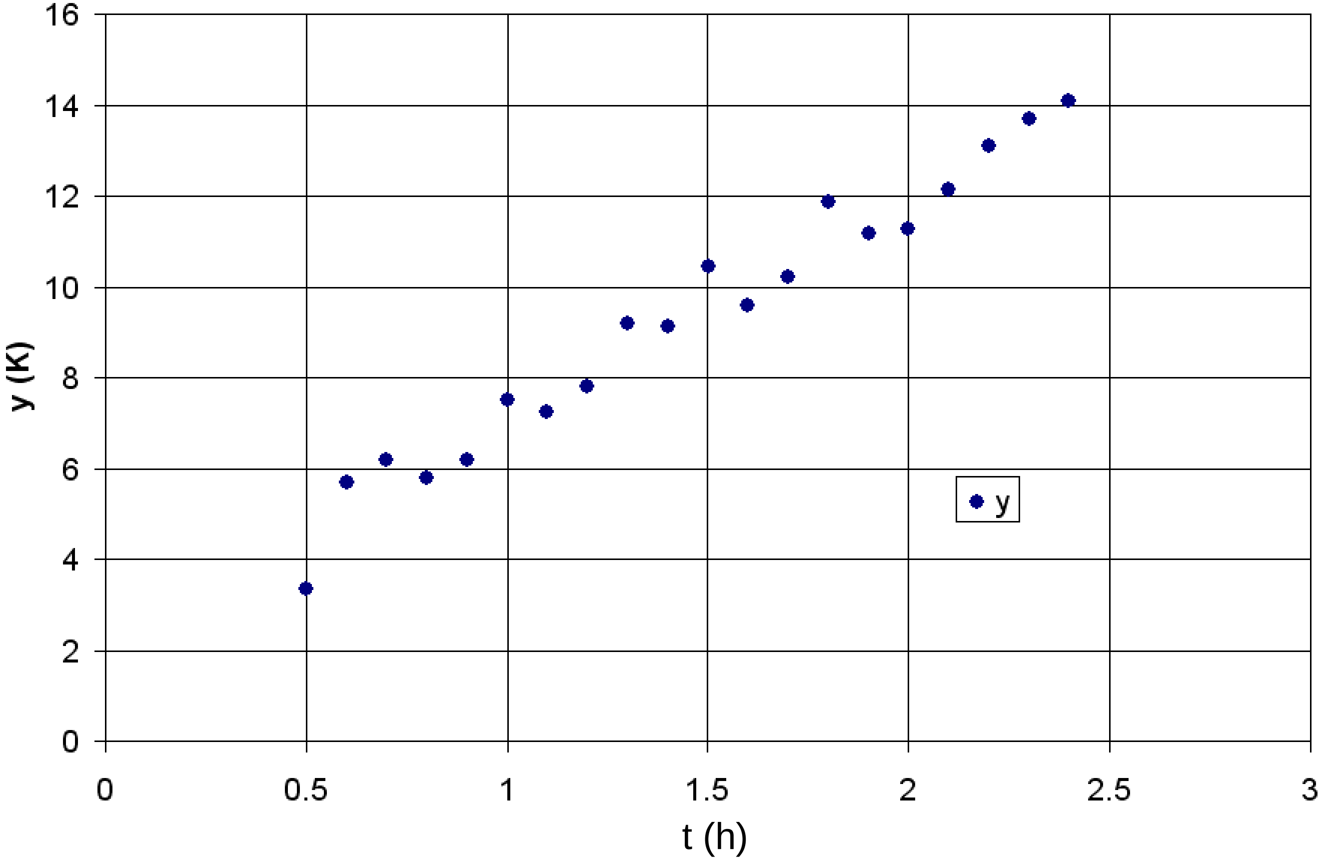
Denis Maillet
LEMTA UMR CNRS
Nancy, France

General context, white box case



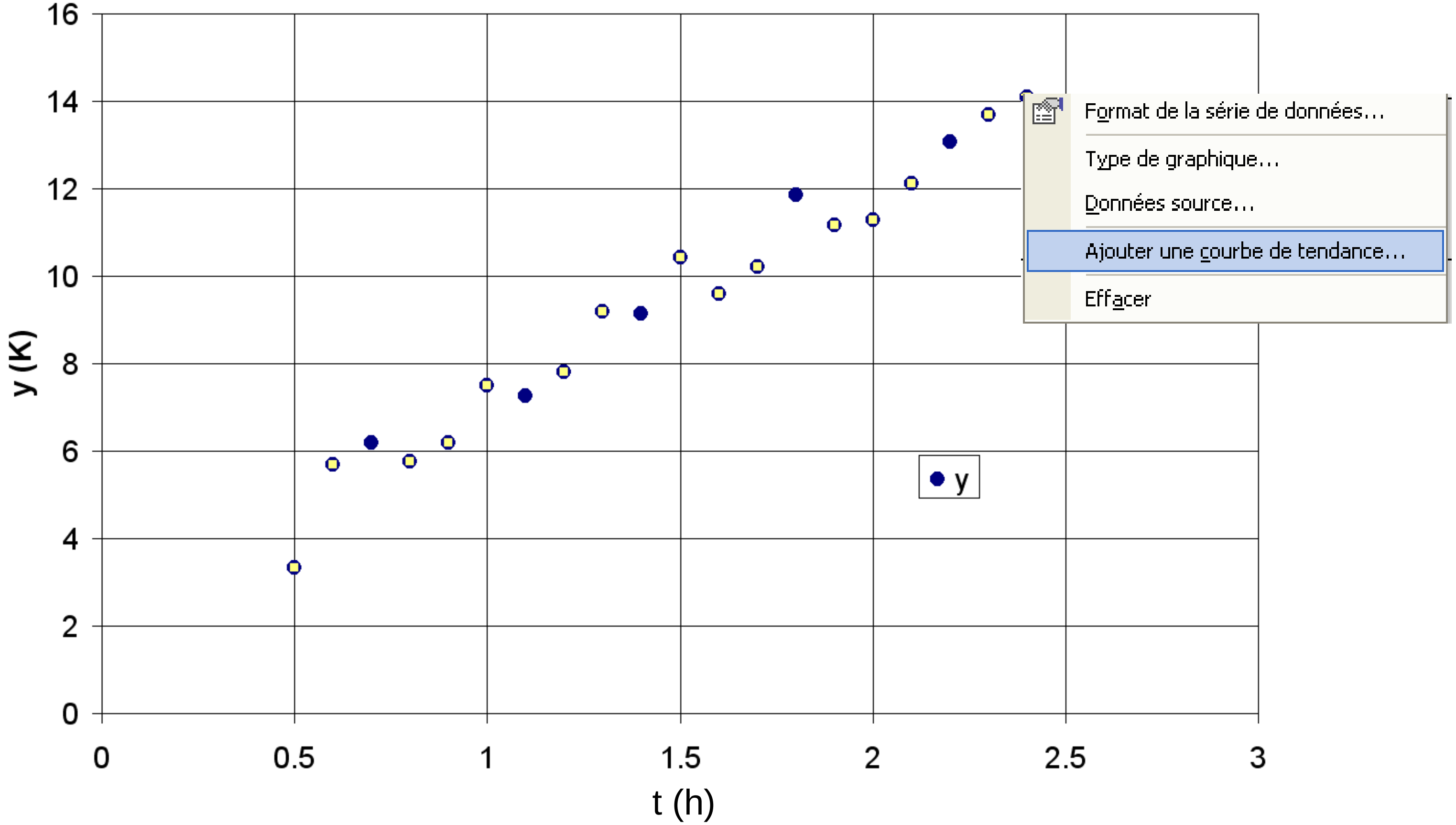
First : Beyong the magic of 'trendline tool' ('courbe de tendance' in français...)

20 measurements y_i at 20 times t_i, i=1 to 20



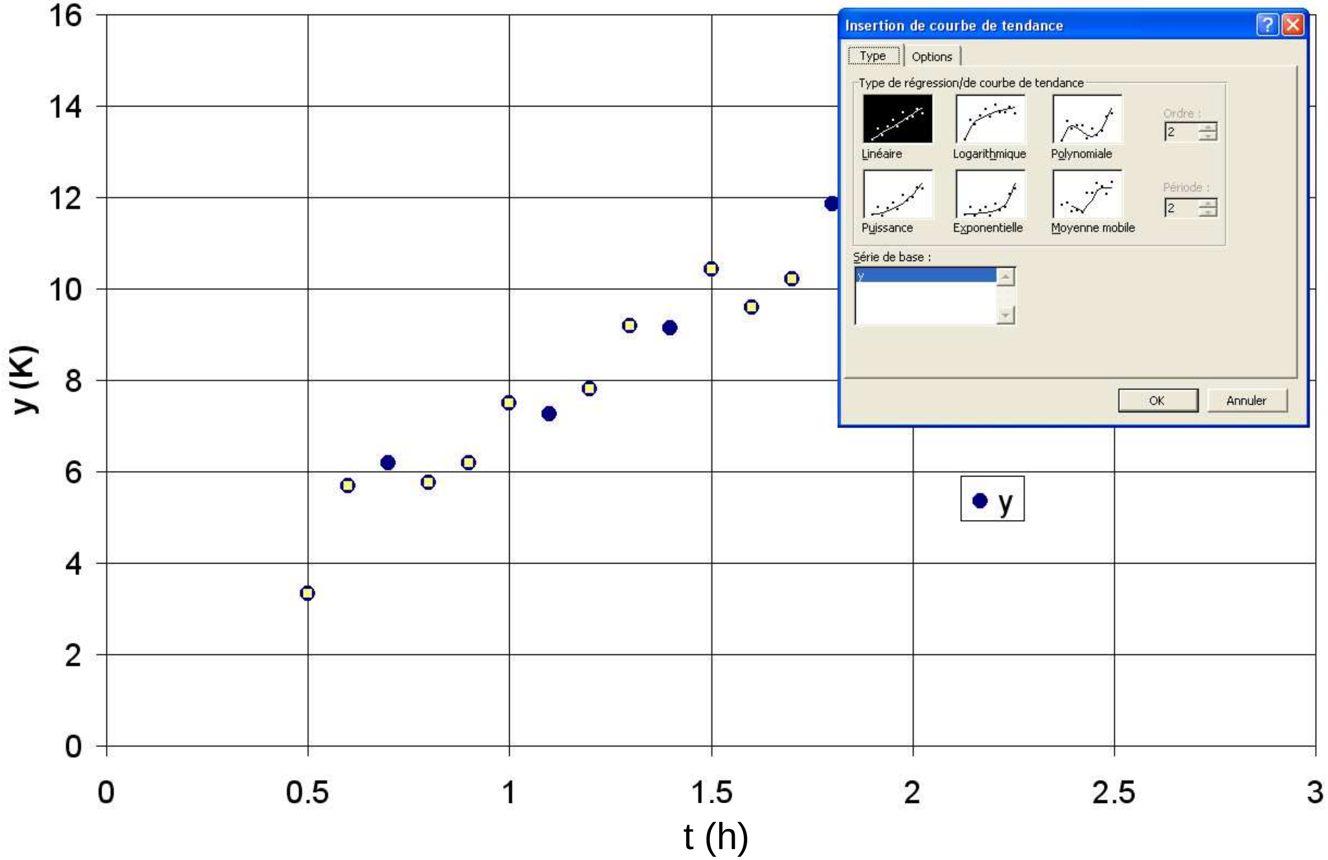
Beyong the magic of 'trendline tool'

20 measurements y_i at 20 times t_i, i=1 to 20



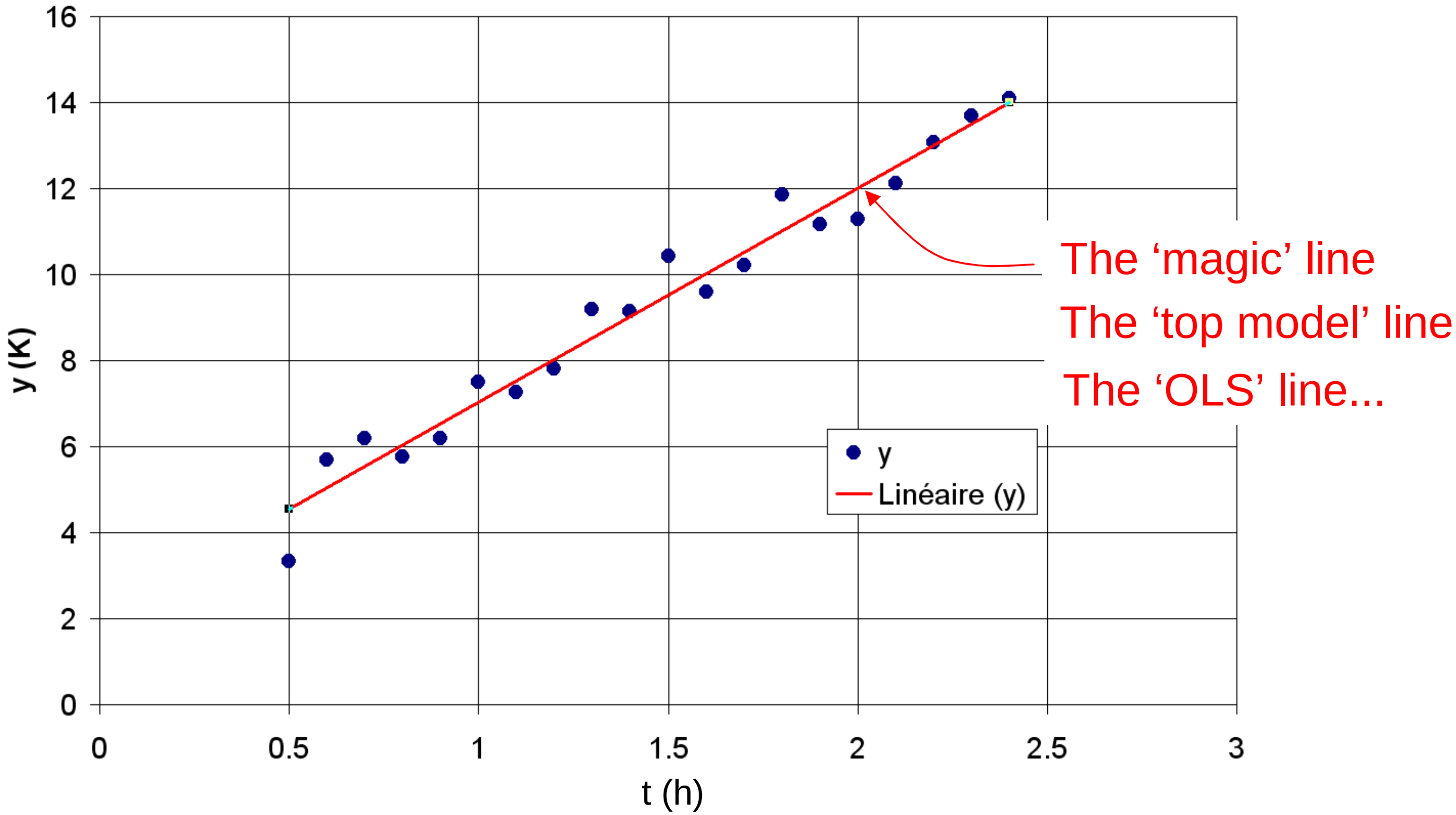
Beyong the magic of ‘trendline tool’

20 measurements y_i at 20 times t_i, i=1 to 20



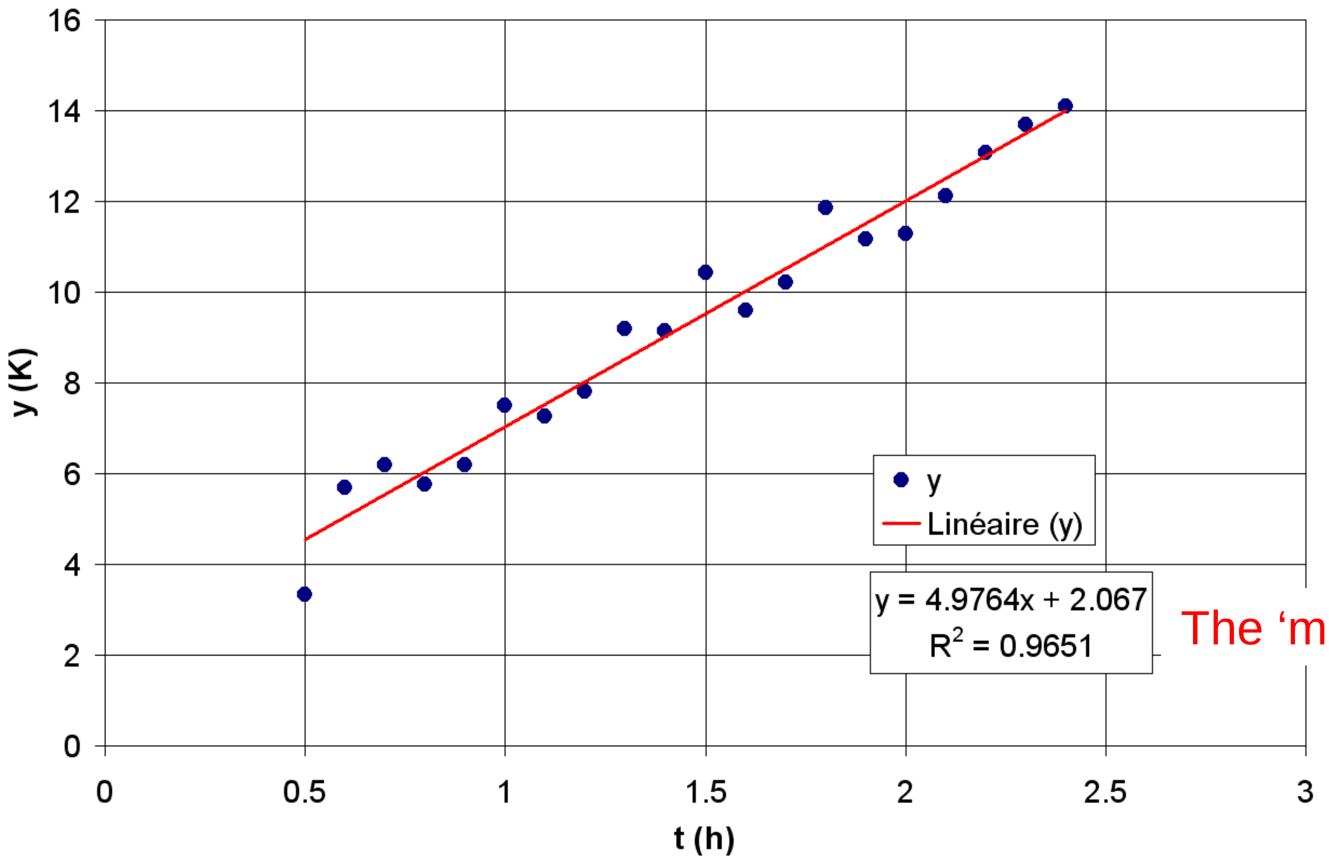
Beyong the magic of ‘trendline tool’

20 measurements y_i at 20 times t_i, i=1 to 20



Beyong the magic of 'trendline tool'

20 measurements y_i at 20 times t_i, i=1 to 20

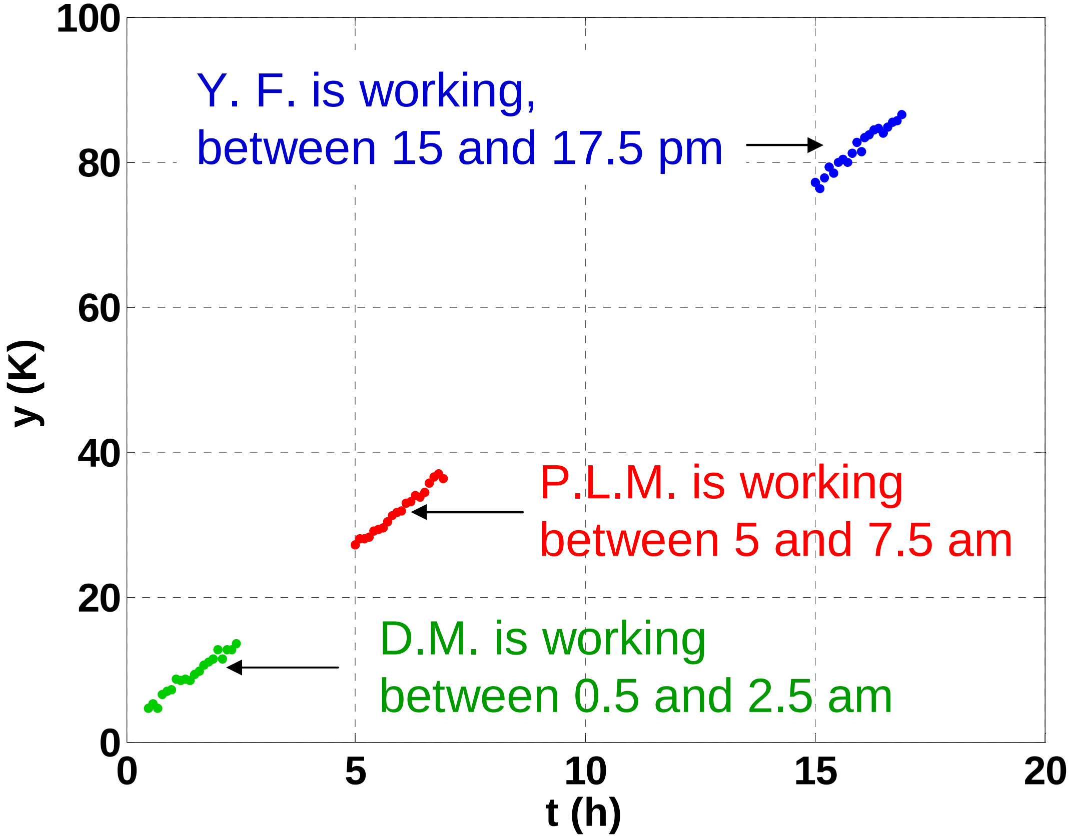


The 'magic' model

Let's play with the trendline tool, and let's observe what happens...

Let's play with the 'trendline tool'

A researcher (names Y. J.) works with three students, on an experiment that begins at 0.00 o'clock and that is during about 20h. Each student performs m=20 measurements y_i. Each measurement is done with the same accuracy.



Y. J. suspects that every measurement can be explained by the simple model :

$$y_{mo} = x_1 t + x_2$$

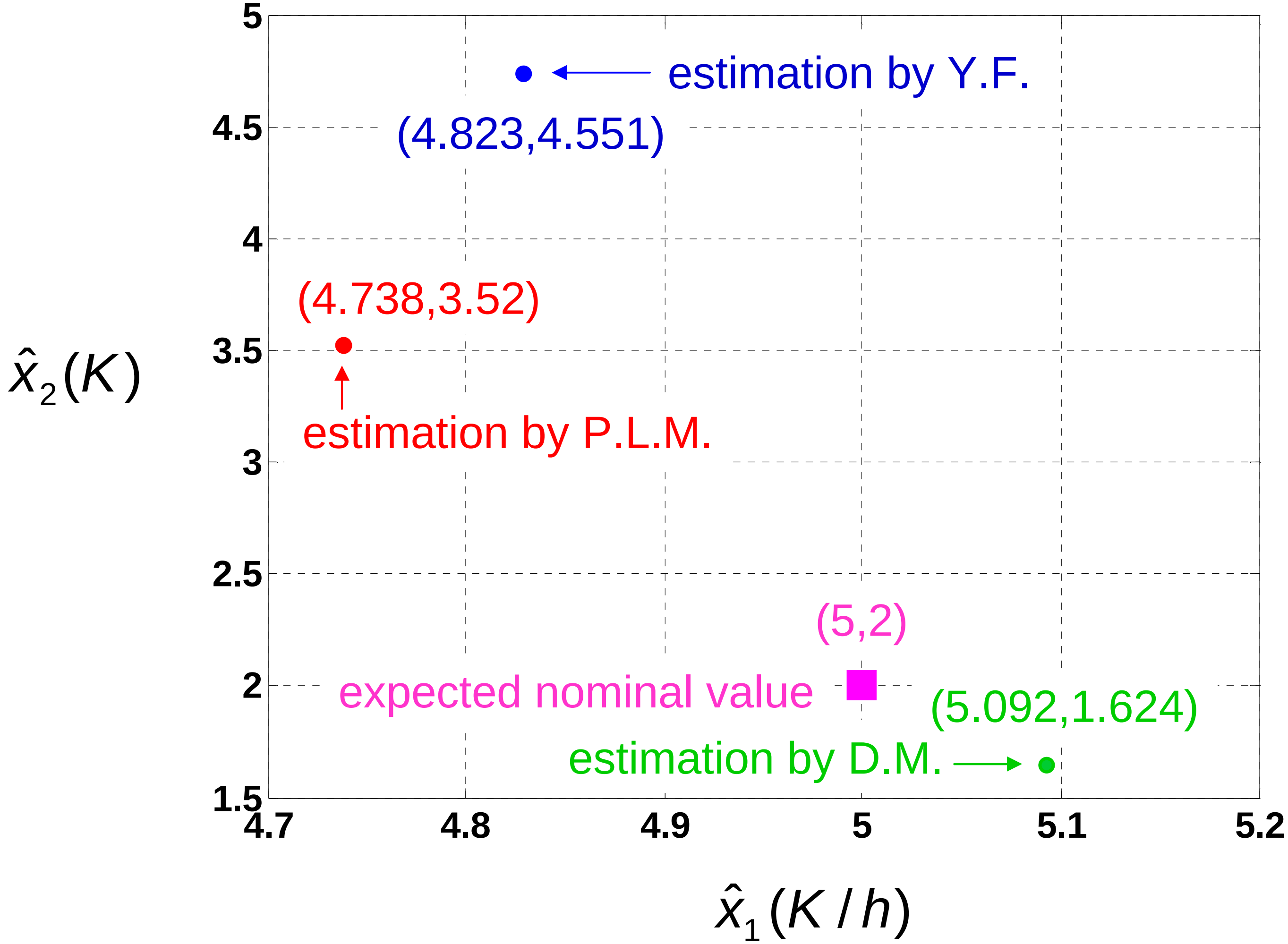
He also expects the values

$$x_1 = x_1^{nom} = 5\,K/h$$

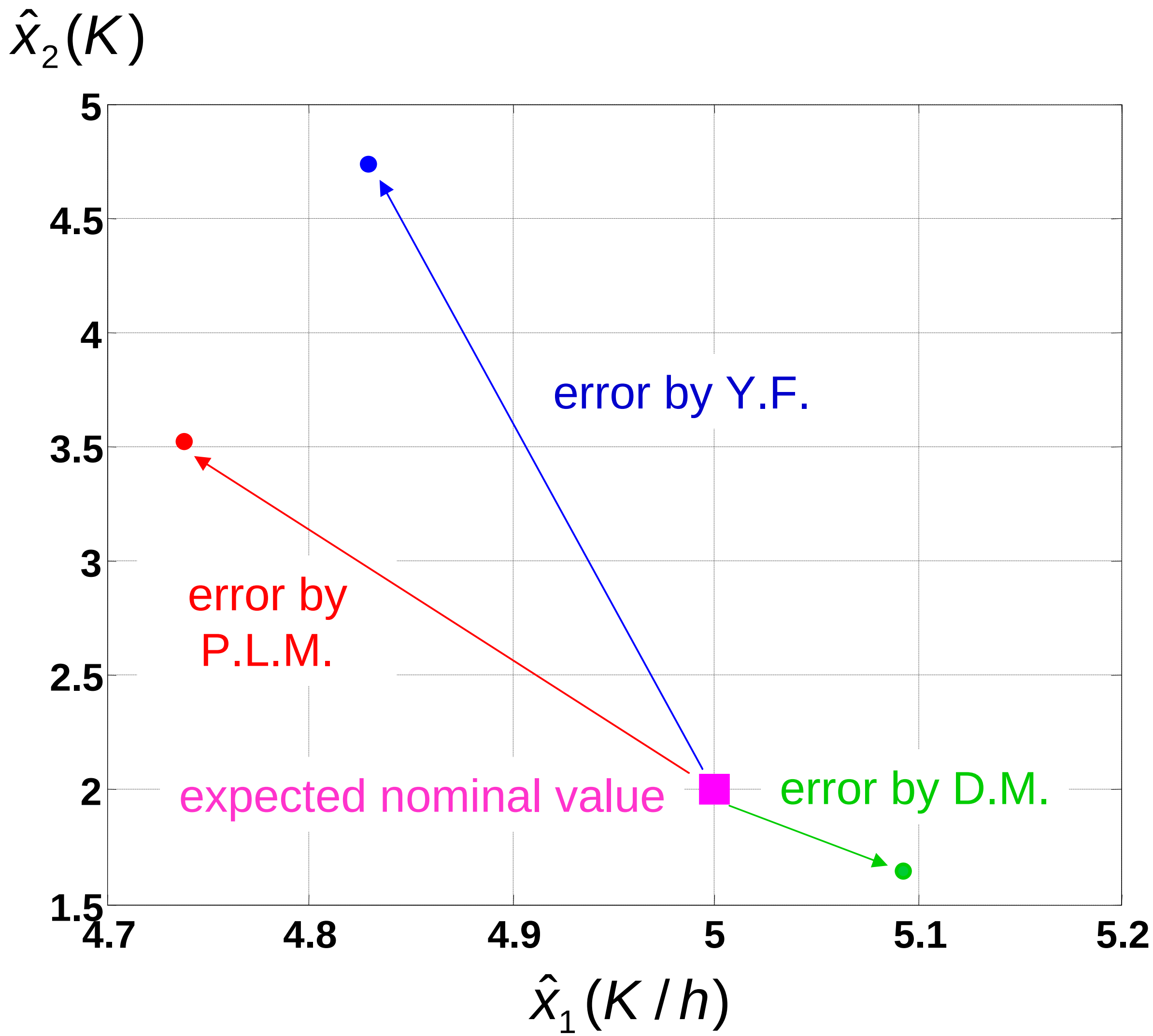
$$x_2 = x_2^{nom} = 2\,K$$

...but keep them secret... He asks each student to use the trendline tool on his own 20 measurements and give him the value of x_1 and x_2

Result of estimations in the 'estimation plane'



Result of estimations in the 'estimation plane'

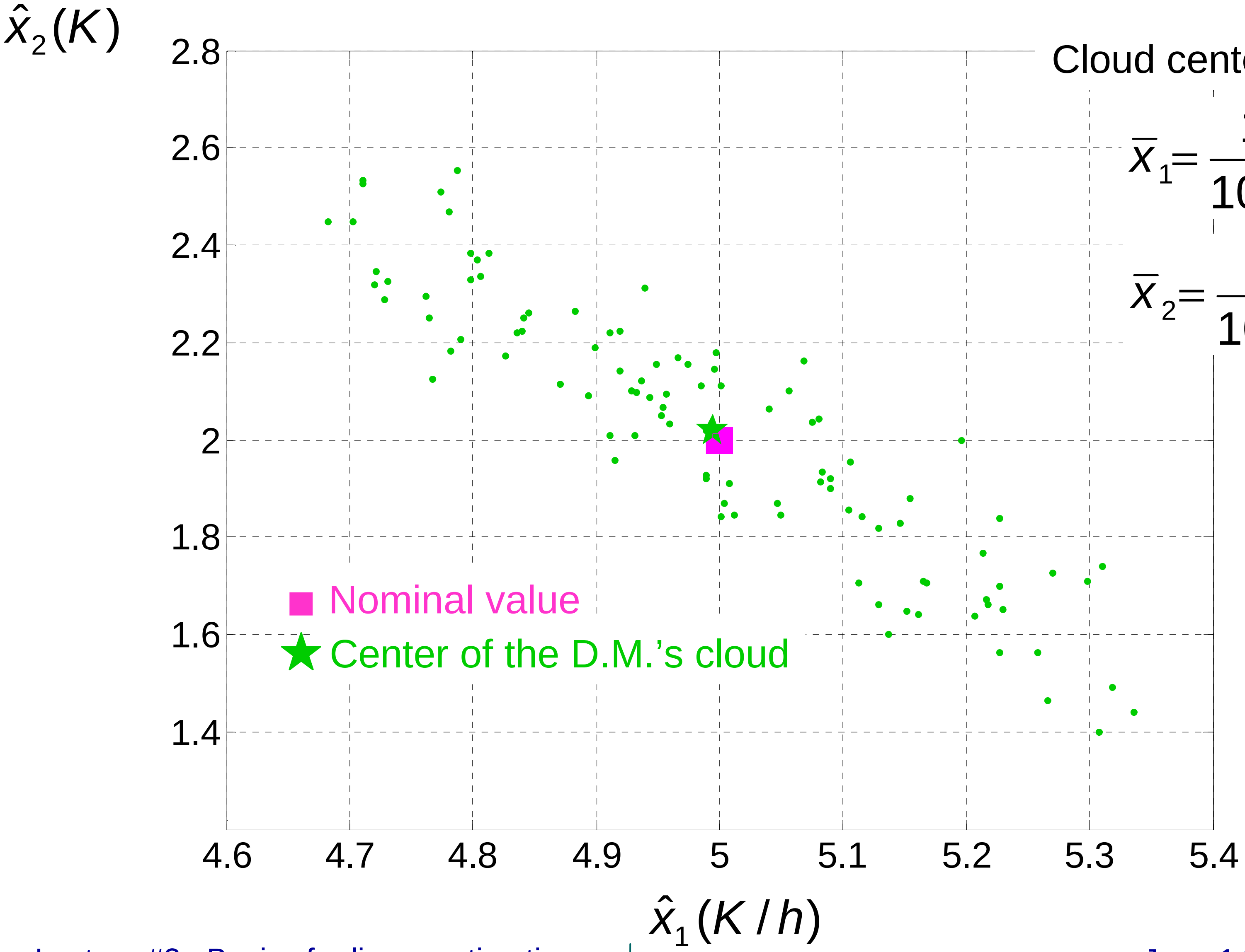


- each student gives a different result
- each student is at a different distance from the expected value

One way to be sure :
« Do experiment and parameter estimation again! And again, and again.... ! »

The experiment is the same... excepted the random part of it : the noise measurement

Cloud of 100 estimations $(\hat{x}_{1,i}, \hat{x}_{2,i})$ for D.M.

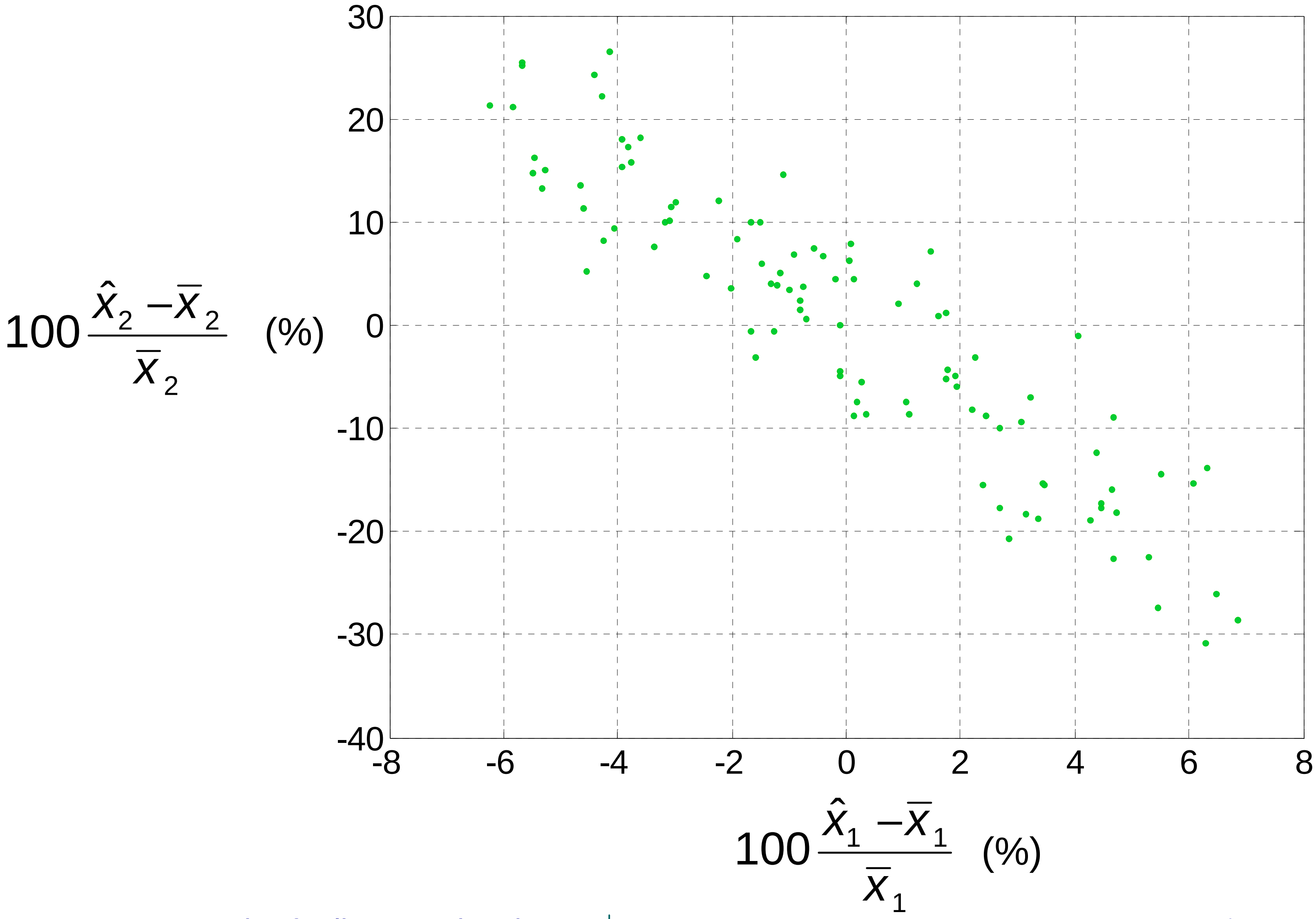


Cloud center coordinates :

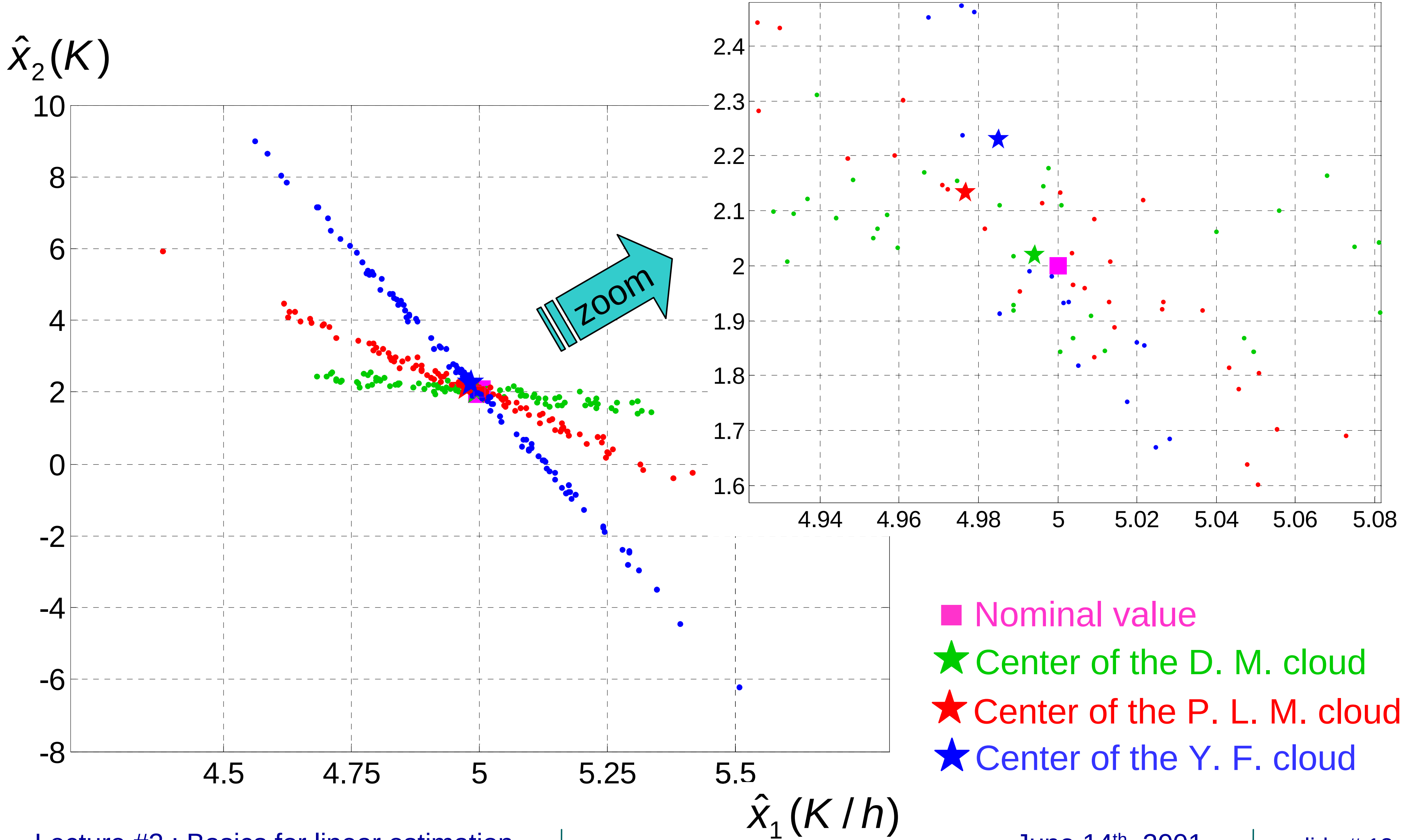
$$\bar{x}_1 = \frac{1}{100}\sum_{i=1}^{100}\hat{x}_{1,i}$$

$$\bar{x}_2 = \frac{1}{100}\sum_{i=1}^{100}\hat{x}_{2,i}$$

Relative scattering of D.M.'s cloud around its center



Three clouds of 100 estimations $(\hat{x}_{1,i}, \hat{x}_{2,i})$

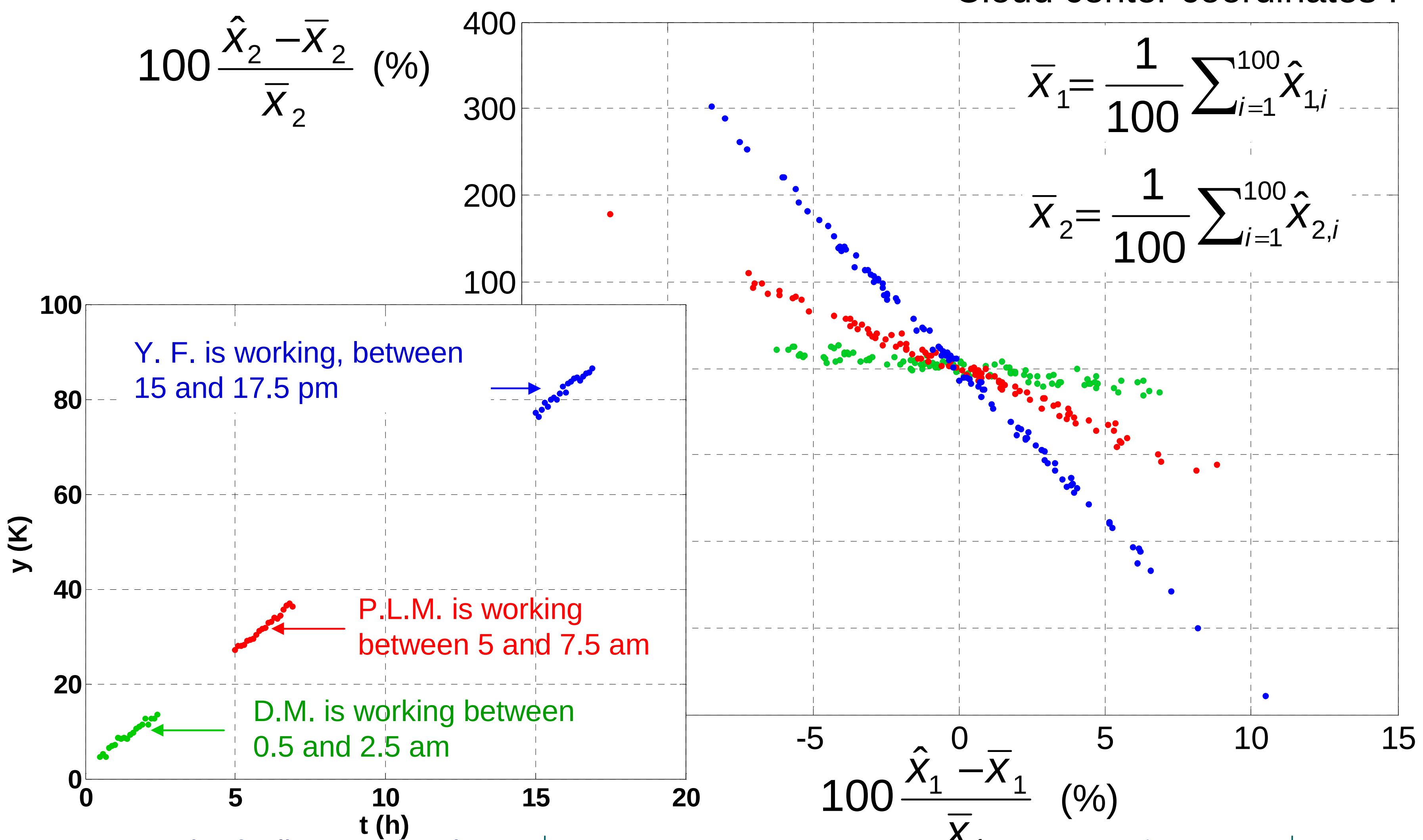


Relative scattering of each cloud around its center

Cloud center coordinates :

$$\bar{x}_1 = \frac{1}{100}\sum_{i=1}^{100}\hat{x}_{1,i}$$

$$\bar{x}_2 = \frac{1}{100}\sum_{i=1}^{100}\hat{x}_{2,i}$$



Results

Each student can annonce now :
- the central value of its cloud of 100 estimations
- a size of the region (absolute and relative) in which are located the majority of its estimations

Student	D.M.	P.L.M.	Y.F.
Time range (h)	0.5 h -2.5 h	5 h -7.5 h	15 h -17.5 h
Central value $\bar{x}_1$ (K/h)	4.994 K/h	4.738 K/h	4.985 K/h
Absolute interval (K/h)	±0.3 K/h	±0.35 K/h	±0.35 K/h
Relative interval (%)	± 6 %	± 7 %	± 7 %
Central value $\bar{x}_2$ (K)	2.019	3.52	2.223
Absolute interval (K)	±0.5 K	±1 K/h	±5.3 K/h
Relative interval (%)	± 10 %	± 28 %	± 106 %

First comments

- Finally we can say that to the question :

  'Find the x_1 and x_2 values of model $y_{mo} = x_1 t + x_2$,
  given m measurements y_i at t_i'

  the answer is not :

  'a unique point $(\hat{x}_1, \hat{x}_2)$'

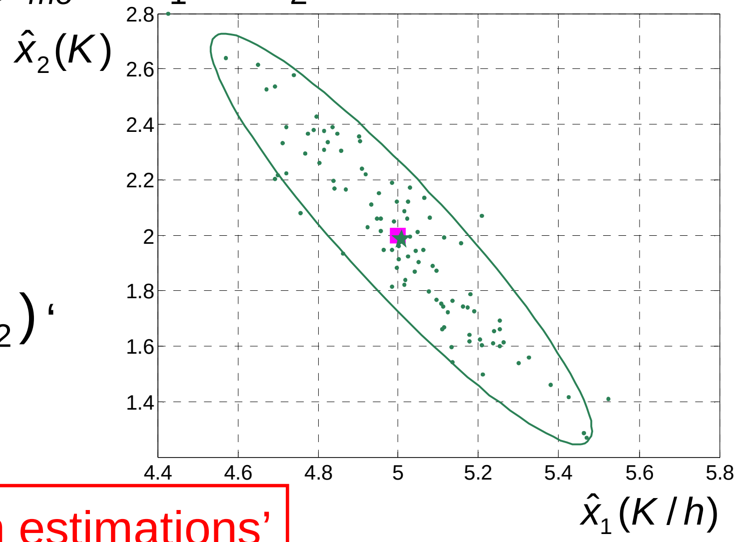
  but rather :

  'a SPOT (or a cloud) of points $(\hat{x}_1, \hat{x}_2)$'

  because of random noise measurement.

  In other words :

  'Blur on measurements gives blur on estimations'



- Here clouds of estimations have elliptical shapes with high density in the central region

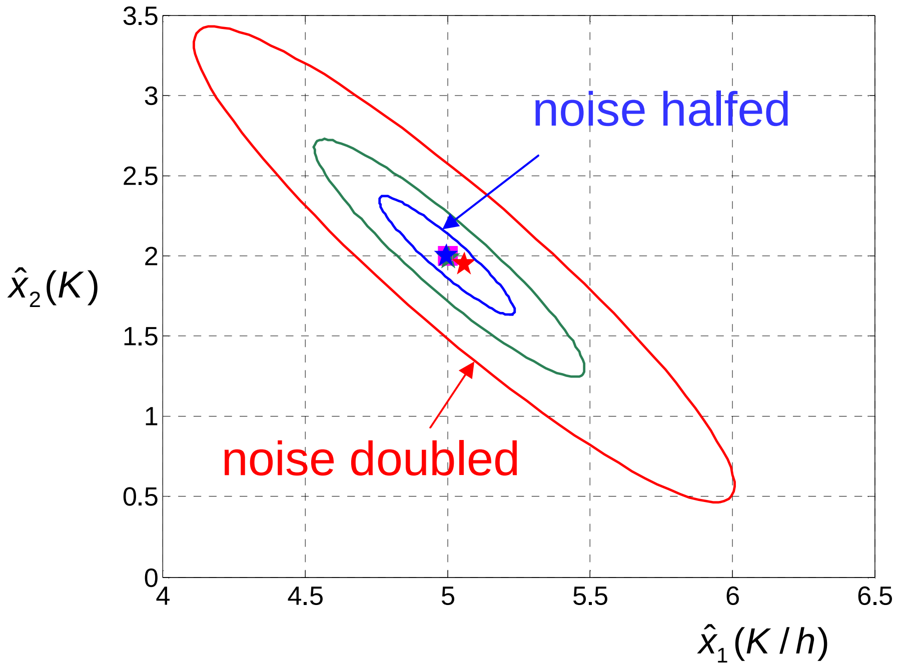
- The center of each spot is very close to the nominal value

- The ideal spot would be : - with the 'smallest' extension
  - centered on the nominal value

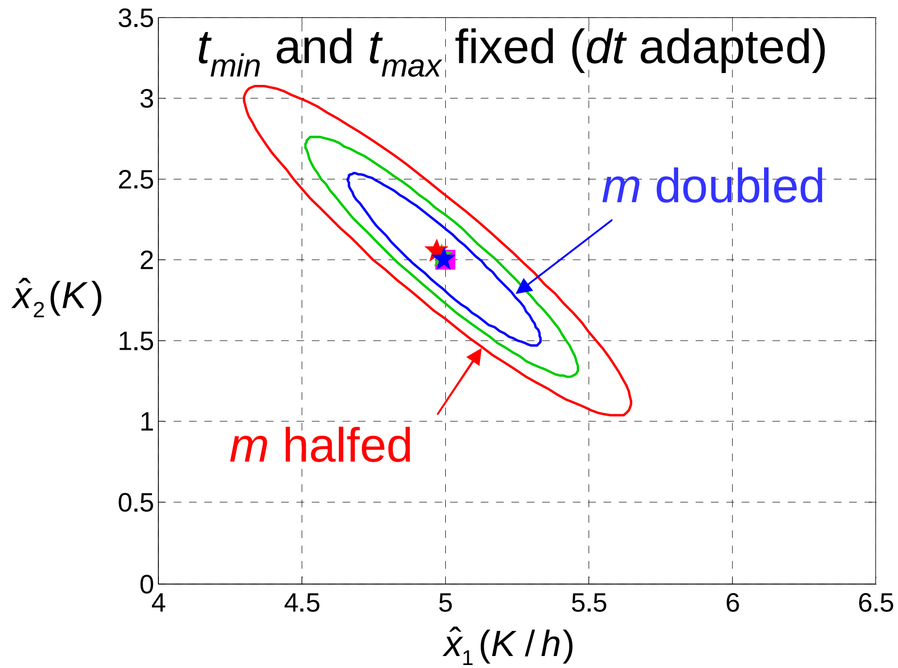
Next comments : influence of experimental conditions

- It seems that certain experimental conditions are better than others :
  - here, measurements have to be 'close' to t=0
- Suppose we know the equation of the elliptical solution spot (detailed later) :

- What happens if noise measurement magnitude changes ?



- What happens if number of measurements (m=20) changes ?

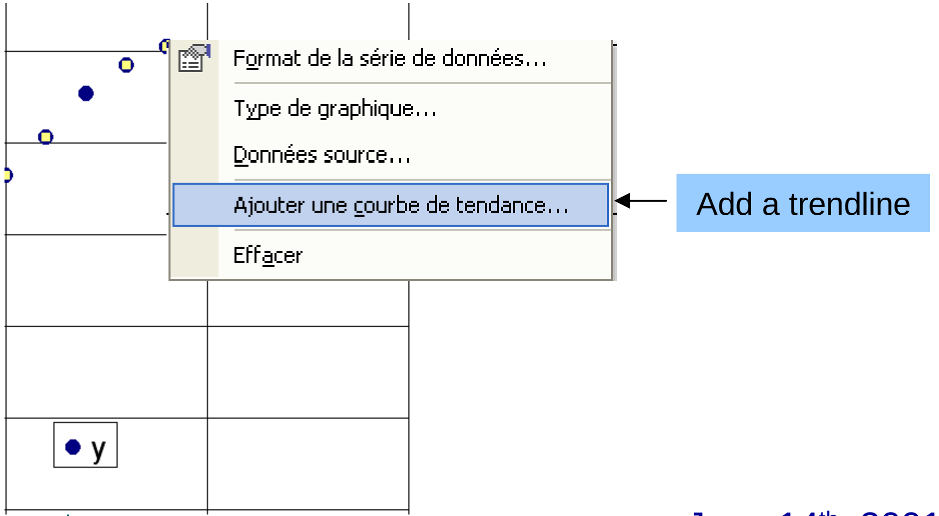


Back to the magic...

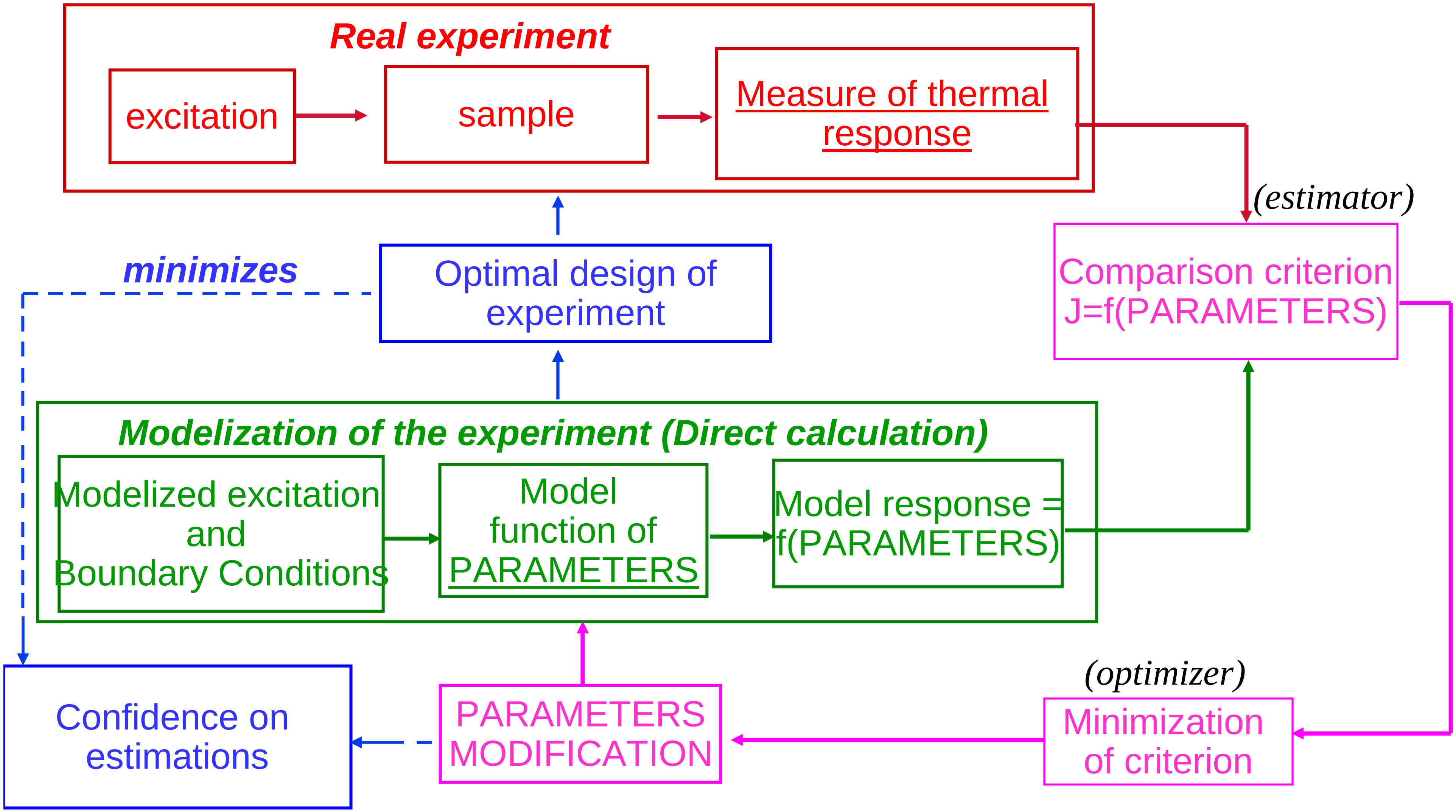
- So it would be very interesting to predict the performances of a parameter estimation method in term of 'spot (or cloud) of estimations' **without achieving 100 experiment/identifications!**

- We must try to predict the shape of the 'spot of estimations' (that will be called 'the confidence region'), associated to only one **experiment/identification realisation**

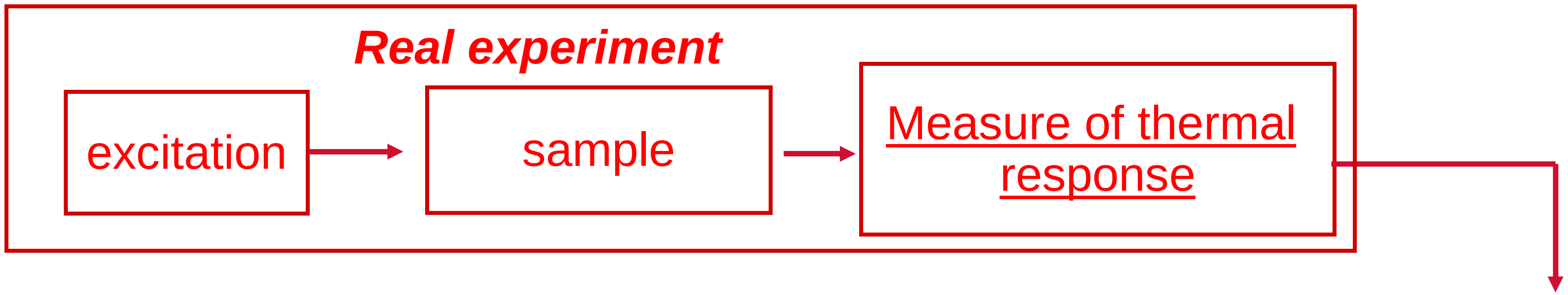
- But before, we have to reveal the secret of the 'magic/top model/OLS line'...



Roadmap for estimation #1 : measurements



Roadmap for estimation #1 : measurements



($m\times 1$) experimental measurements vector

$$\boldsymbol{y} = [y_1 \ldots y_i \ldots y_m]^{\mathrm{t}} \quad \text{with} \quad y_i = y(t_i)\,,\; t_i = t_{\min} + (i-1).dt \;,\; i=1,\ldots,m$$

($m\times 1$) time vector (explicative variable)

$$\boldsymbol{t} = [t_1 \ldots t_i \ldots t_m]^t$$

y_i

($m\times 1$) measurement errors vector

$$\boldsymbol{\varepsilon} = [\varepsilon_1 \ldots \varepsilon_i \ldots \varepsilon_n]^{\mathrm{t}}$$

ε_i be the (unknown) error associated to the measurement

Some assumptions have to be done on these measurement errors.

Roadmap for estimation #1 : measurements

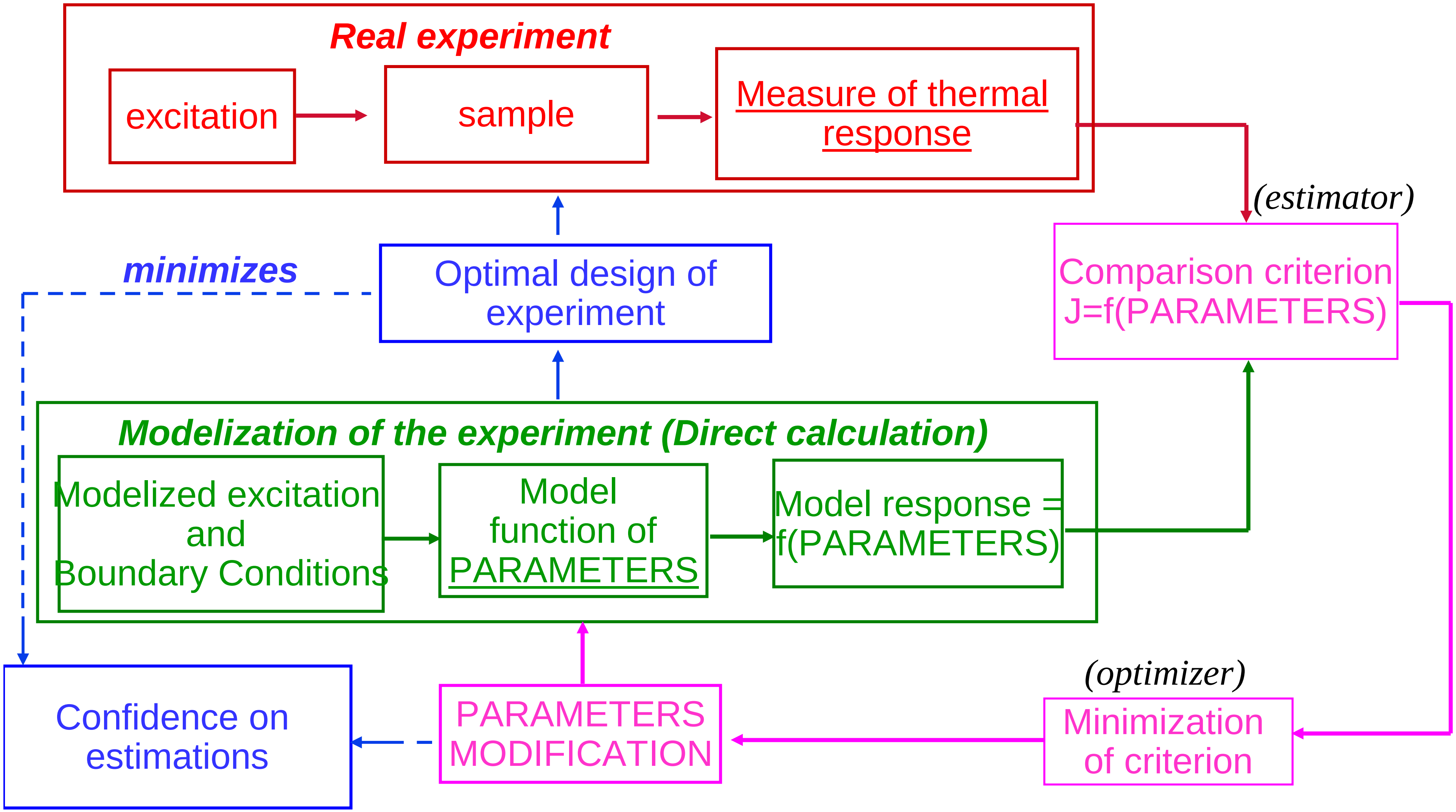
Number	Assumption on measurement errors	Explanation
1	Additive errors	$\boldsymbol{y} = \boldsymbol{y}_{perfect} + \boldsymbol{\varepsilon}$
2	Unbiased model	$\boldsymbol{y}_{perfect} = y_{mo}(x^{exact})$
3	Zero mean errors	$E[\boldsymbol{\varepsilon}] = 0$
4	Constant variance	$Var[\varepsilon] = \sigma_\varepsilon^2$
5	Uncorrelated errors	$Cov[\varepsilon_i \varepsilon_j] = 0$ for $i \neq j$
6	Normal probability distribution	
7	Known statistical parameters	
8	No error in the X_{ij}	**X** is not a random matrix
9	No prior information regarding the parameters	

$E[.]$ Is the expected value operator (representing the mean of a large number of realizations of the random variable)

Covariance Matrix of measurement errors

$$\boldsymbol{\psi} = E\left[(\boldsymbol{\varepsilon} - E[\boldsymbol{\varepsilon}])(\boldsymbol{\varepsilon} - E[\boldsymbol{\varepsilon}])^t\right] = E\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^t\right] = diag\left(\sigma_\varepsilon^2, \cdots, \sigma_\varepsilon^2, \cdots, \sigma_\varepsilon^2\right) = \boldsymbol{I}.\sigma_\varepsilon^2$$

Roadmap for estimation #1 : measurements



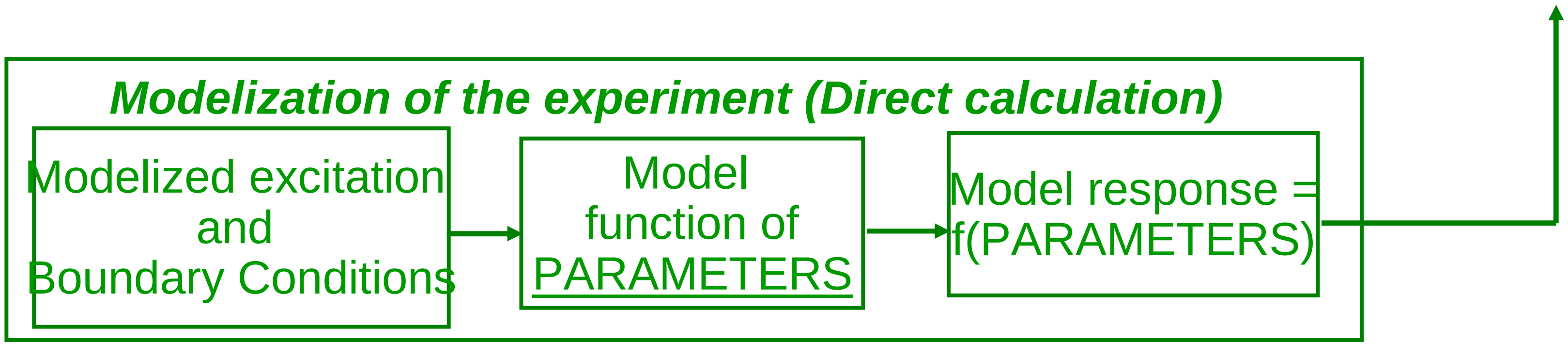
Roadmap for estimation #2 : model

($m\times 1$) experimental measurements vector

$$\boldsymbol{y}_{\boldsymbol{mo}}(\boldsymbol{t},\boldsymbol{x}) = \left[y_{mo,1}(t_1,\boldsymbol{x})\ldots y_{mo,i}(t_i,\boldsymbol{x})\ldots y_{mo,m}(t_m,\boldsymbol{x})\right]^{\mathbf{t}}$$

with $y_{mo}(t,\boldsymbol{x}) = \eta(t,\boldsymbol{x})$

parameters vector ($n\times 1$) : $\boldsymbol{x} = \left[x_1 \ \ldots \ x_n\right]^{\mathbf{t}}$



With here: $y_{mo}(t,\boldsymbol{x}) = x_1 t + x_2$

Roadmap for estimation #2 : model

$$y_{mo}(t,\boldsymbol{x}) = x_1 t + x_2$$

NB : that model is said 'linear' on the parameter estimation point of view because it is linear with respect to its parameter x_i. The following model

$$y_{mo}(t,\boldsymbol{x}) = x_1\sqrt{t} + x_2.erf(t)$$

is still linear with respect to its parameter x_i even it is not with respect to time. The following model

$$y_{mo}(t,\boldsymbol{x}) = x_1\sqrt{t} + \exp(-x_2 t)$$

is not linear with respect to x_2

Roadmap for estimation #2 : model

Writing the *m* model values for the *m* time values , the *m* resulting equations can be written in a matrix way as following :

$$\begin{bmatrix} y_{mo,1} \\ \vdots \\ y_{mo,i} \\ \vdots \\ y_{mo,m} \end{bmatrix} = \begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_i & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or

$$\boxed{\boldsymbol{y}_{mo} = \boldsymbol{S}\boldsymbol{x}}$$

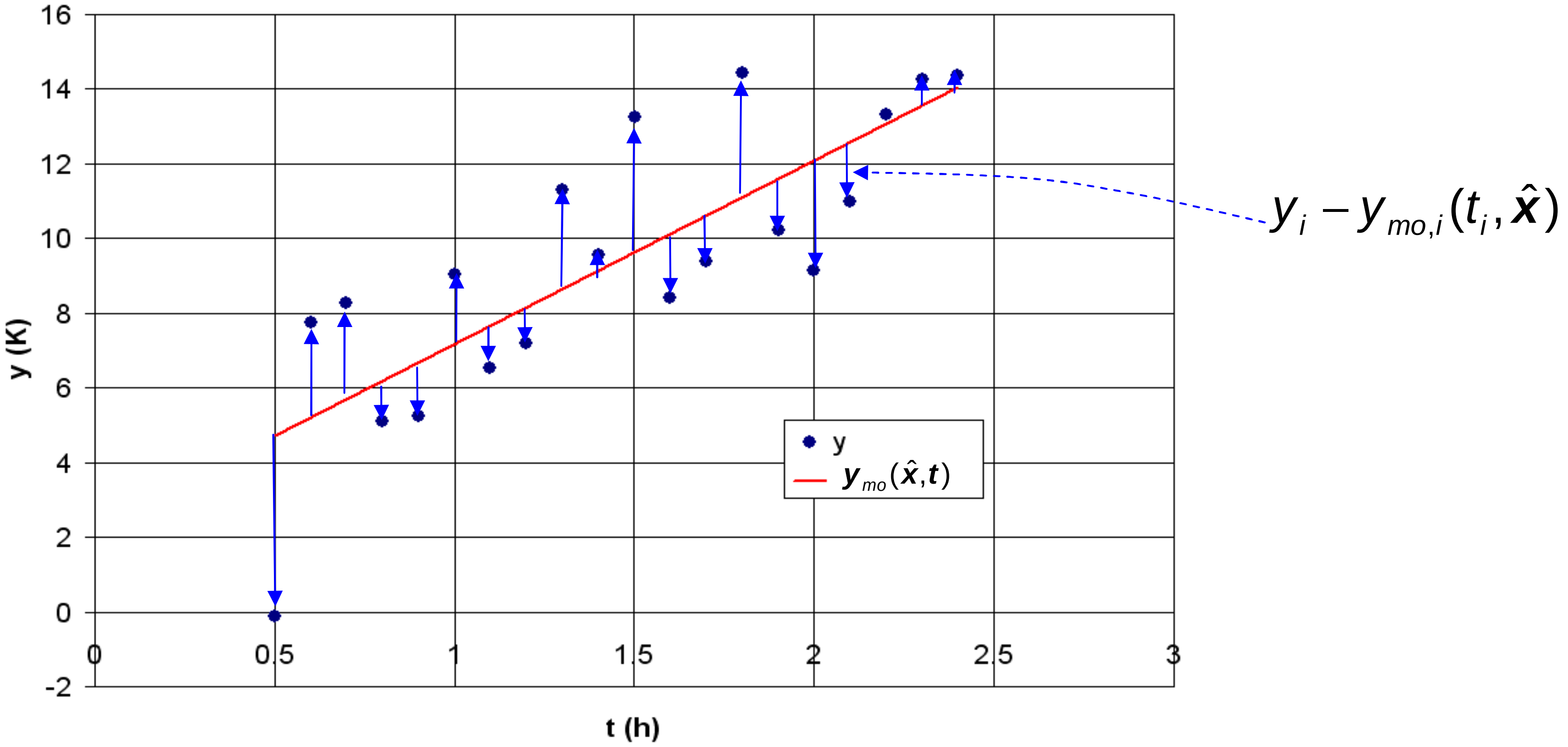
with

$$\boldsymbol{S} = \begin{bmatrix} S_1(t_1) & S_2(t_1) \\ \vdots & \vdots \\ S_1(t_i) & S_2(t_1) \\ \vdots & \vdots \\ S_1(t_m) & S_2(t_1) \end{bmatrix} = \begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_i & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix}$$

Sensitivity matrix

with $S_k(t,\boldsymbol{x}) = \left.\dfrac{\partial y_{mo}(t,\boldsymbol{x})}{\partial x_k}\right|_{t,x_j \text{ for } j \neq k}$: sensitivity coefficient relative to parameter x_k, $k=1,\ldots,n$

Problem : use the *m(=20)* measurements to estimate the *n (2)* unknown parameters :
Overdetermined problem transformed in a minimization problem :



Residual vector (m×1)

$$\mathbf{r}(\hat{\mathbf{x}}) = \mathbf{y} - \mathbf{y}_{mo}(\hat{\mathbf{x}}) = \left[y_1 - y_{mo,1}(t_1,\hat{\mathbf{x}}) \quad \dots \quad y_i - y_{mo,i}(t_i,\hat{\mathbf{x}}) \quad \dots \quad y_m - y_{mo,m}(t_m,\hat{\mathbf{x}}) \right]^t$$

Roadmap for estimation #3 : the cost function

Without any a priori information on the parameters and given the above assumptions for measurements errors, the square of the Euclidian norm of the residual vector is minimised :

$$J_{OLS}(\hat{\boldsymbol{x}}) = \left\| \boldsymbol{r}(\hat{\boldsymbol{x}}) \right\|^2 = \left\| \boldsymbol{y} - \boldsymbol{S}\hat{\boldsymbol{x}} \right\|^2$$

This scalar number is called the Ordinary Least Squares (OLS) cost function

$$J_{OLS}(\hat{\boldsymbol{x}}) = \sum_{i=1}^{m} r_i(\hat{\boldsymbol{x}})^2 = \sum_{i=1}^{m} \left(y_i - \sum_{j=1}^{n} S_j(t_i)\hat{x}_j) \right)^2 = \sum_{i=1}^{m} \left(y_i - y_{mo,i}(t_i, \hat{\boldsymbol{x}}) \right)^2$$

With a matrix formulation it gives :

$$J_{OLS}(\hat{\boldsymbol{x}}) = \left[\boldsymbol{y} - \boldsymbol{y}_{mo}(\hat{\boldsymbol{x}}) \right]^t \left[\boldsymbol{y} - \boldsymbol{y}_{mo}(\hat{\boldsymbol{x}}) \right]$$

$$J_{OLS}(\hat{\boldsymbol{x}}) = \left[\boldsymbol{y} - \boldsymbol{S}\hat{\boldsymbol{x}} \right]^t \left[\boldsymbol{y} - \boldsymbol{S}\hat{\boldsymbol{x}} \right]$$

The solution of the problem is then : $\hat{\boldsymbol{x}}_{OLS} = \arg\left[\min\left(J_{OLS}(\hat{\boldsymbol{x}}) \right) \right]$

Roadmap for estimation #4 : the OLS solution

The OLS estimator is the one that minimizes the scalar function $J_{OLS}(\hat{\boldsymbol{x}})$

$$\nabla_x J_{OLS}(\hat{\boldsymbol{x}}_{OLS}) = 0 \quad \text{with} \quad \nabla_x = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

$$\nabla_x J_{OLS}(\hat{\boldsymbol{x}}) = 2\left[\nabla_x \left[\boldsymbol{y} - \boldsymbol{y}_{mo}(\hat{\boldsymbol{x}})\right]\right]^t \left[\boldsymbol{y} - \boldsymbol{y}_{mo}(\hat{\boldsymbol{x}})\right]$$

Knowing that $\boldsymbol{S}^t = \left[\nabla_x \boldsymbol{y}_{mo}(\hat{\boldsymbol{x}})\right]^t$ and $\boldsymbol{y}_{mo}(\hat{\boldsymbol{x}}) = \boldsymbol{S}\hat{\boldsymbol{x}}$

$$\nabla_x J_{OLS}(\hat{\boldsymbol{x}}) = -2\boldsymbol{S}^t\left[\boldsymbol{y} - \boldsymbol{S}\hat{\boldsymbol{x}}\right]$$

Then $\hat{\boldsymbol{x}}_{OLS}$ is solution of : $\left[\boldsymbol{S}^t\boldsymbol{S}\right]\hat{\boldsymbol{x}}_{OLS} = \boldsymbol{S}^t\boldsymbol{y}$ (the Normal Equation)

Roadmap for estimation #4 : the solution

$$\hat{\boldsymbol{x}}_{OLS} = \left[\boldsymbol{S}^t \boldsymbol{S}\right]^{-1} \boldsymbol{S}^t \boldsymbol{y}$$

NB : $\left[\boldsymbol{S}^t \boldsymbol{S}\right]^{-1} \boldsymbol{S}^t$ (n×m) is the Moore Penrose matrix

If we distinguish parameters to be estimated $\boldsymbol{x_r}$ from parameters that will be fixed $\boldsymbol{x_c}$

$$\boldsymbol{y}_{\boldsymbol{mo}}(\boldsymbol{x}) = \boldsymbol{Sx} = \boldsymbol{S_r x_r} + \boldsymbol{S_c x}_c$$

$$\boldsymbol{S} = [\boldsymbol{S}_r \vdots \boldsymbol{S}_c] = \left[\begin{bmatrix} S_1(t_1) & \dots & S_r(t_1) \\ \vdots & \dots & \vdots \\ S_1(t_m) & \dots & S_r(t_m) \end{bmatrix} \vdots \begin{bmatrix} S_{r+1}(t_1) & \dots & S_q(t_1) \\ \vdots & \dots & \vdots \\ S_{r+1}(t_m) & \dots & S_q(t_m) \end{bmatrix}\right]$$

$$\hat{\boldsymbol{x}}_{OLS} = \left[\boldsymbol{S}_r^t \boldsymbol{S}_r\right]^{-1} \boldsymbol{S}_r^t (\boldsymbol{y} - \boldsymbol{S}_c \boldsymbol{x}_c)$$

Matrix $\boldsymbol{S}^t \boldsymbol{S}$ needs to be inverted

Roadmap for estimation #5 : the confidence through the covariance

$\boldsymbol{e}_r = \hat{\boldsymbol{x}}_{r,OLS}(\tilde{\boldsymbol{x}}_c) - \boldsymbol{x}_r^{exact}$: error on estimations

$\boldsymbol{e}_c = \tilde{\boldsymbol{x}}_c - \boldsymbol{x}_c^{exact}$: deterministic error (bias) on parameter

$\hat{\boldsymbol{x}}_{r,OLS}(\tilde{\boldsymbol{x}}_c) = \boldsymbol{A}_r(\boldsymbol{y} - \boldsymbol{S}_c\tilde{\boldsymbol{x}}_c)$ and $\boldsymbol{y} = \boldsymbol{y}_{mo}(\boldsymbol{x}^{exact}) + \boldsymbol{\varepsilon} = \boldsymbol{S}_r\boldsymbol{x}_r^{exact} + \boldsymbol{S}_c\boldsymbol{x}_c^{exact} + \boldsymbol{\varepsilon}$

$$\rightarrow \boldsymbol{e}_r = \hat{\boldsymbol{x}}_{r,OLS}(\tilde{\boldsymbol{x}}_c) - \boldsymbol{x}_r^{exact} = \boxed{\boldsymbol{A}_r\boldsymbol{\varepsilon}}\boxed{-\boldsymbol{A}_r\boldsymbol{S}_c\boldsymbol{e}_c} = \boldsymbol{e}_{r1} + \boldsymbol{e}_{r2} \quad , \quad \boldsymbol{A}_r = \left[\boldsymbol{S}_r^t\boldsymbol{S}_r\right]^{-1}\boldsymbol{S}_r^t$$

Random contribution due to random measurement errors

the non-random (deterministic) contribution to the total error vector due to the deterministic error on the fixed parameters

Covariance of estimations

$$\boldsymbol{C}_1 = \mathrm{cov}(\boldsymbol{e}_{r1}) = E[\boldsymbol{e}_{r1}\boldsymbol{e}_{r1}^t] = \boldsymbol{A}_r E[\boldsymbol{\varepsilon\varepsilon}^t]\boldsymbol{A}_r^t = \boldsymbol{A}_r \boldsymbol{\psi} \boldsymbol{A}_r^t = \left[\boldsymbol{S}_r^t \boldsymbol{S}_r\right]^{-1} \sigma_\varepsilon^2 = \boldsymbol{C}_1$$

→ $\left[\boldsymbol{S}_r^t \boldsymbol{S}_r\right]^{-1}$ is a matrix (r×r) that amplifies the noise measurements **(we have found the danger!)**

Bias of estimations

$$E\left[\boldsymbol{e}_{r2}\right] = -\boldsymbol{A}_r \boldsymbol{S}_c \boldsymbol{e}_c = \left[\boldsymbol{S}_r^t \boldsymbol{S}_r\right]^{-1} \boldsymbol{S}_r^t \boldsymbol{S}_c \boldsymbol{e}_c \neq 0$$

→ $\left[\boldsymbol{S}_r^t \boldsymbol{S}_r\right]^{-1} \boldsymbol{S}_r^t \boldsymbol{S}_c$ is a matrix (r×(n-r)) that amplifies the bias on fixed parameters **(we have found another danger!)**

Roadmap for estimation #5 : the confidence

For a fixed value of $\tilde{\boldsymbol{x}}_c$, the covariance matrix of estimations errors is

$$\boldsymbol{C}_r = \mathrm{cov}(\boldsymbol{e}_r) = E[(\boldsymbol{e}_r - E[\boldsymbol{e}_r])(\boldsymbol{e}_r - E[\boldsymbol{e}_r])^t] = E[\boldsymbol{e}_{r1}\boldsymbol{e}_{r1}^t] = \mathrm{cov}(\boldsymbol{e}_{r1}) = \boldsymbol{C}_1$$

The covariance matrix components are

$$\boldsymbol{C}_r = \begin{bmatrix} \sigma_1^2 & \mathrm{cov}(e_{r1}, e_{r2}) & \dots & \mathrm{cov}(e_{r1}, e_{rr}) \\ & \sigma_2^2 & & \mathrm{cov}(e_{r2}, e_{rr}) \\ & & \ddots & \vdots \\ sym & & & \sigma_r^2 \end{bmatrix}$$

Individual variances on the r estimations are on the diagonal

$$\boldsymbol{C}_1 = \left[\boldsymbol{S}_r^t \boldsymbol{S}_r\right]^{-1} \sigma_\varepsilon^2$$

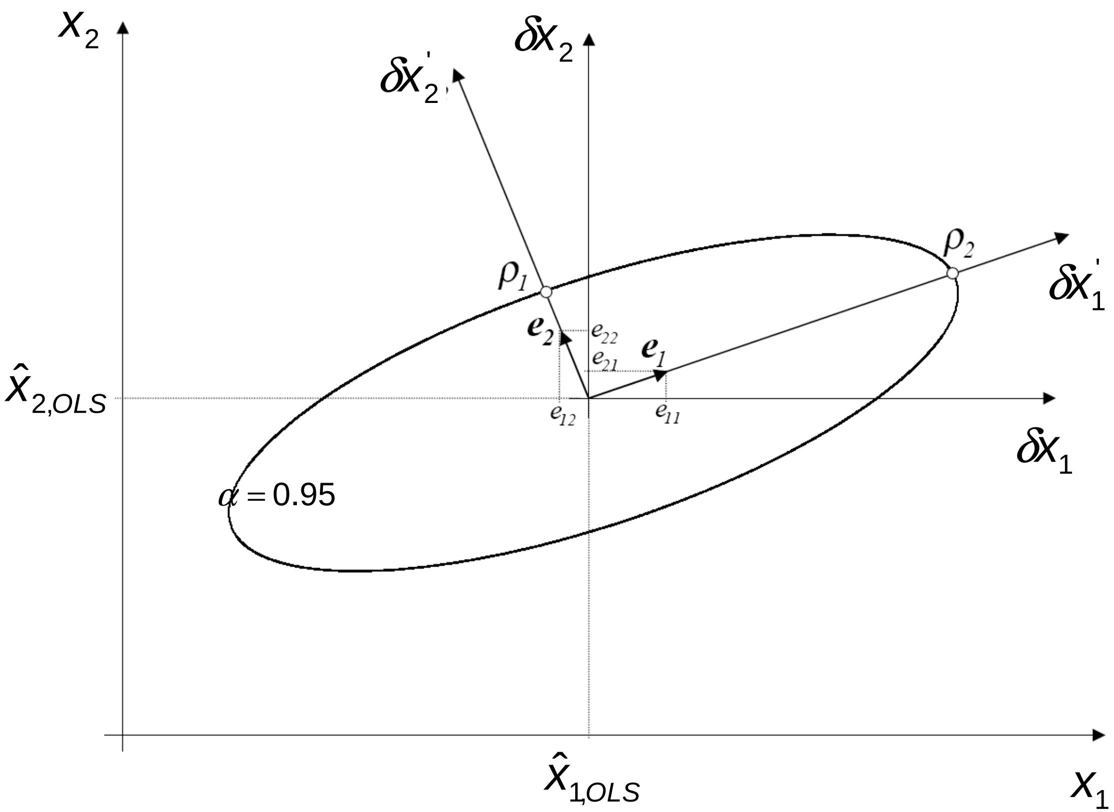
if σ_ε^2 is not measured before the experiment, an estimation of it may be obtained at the end of estimation thanks to the final value of the objective function :

$$J_{OLS}(\hat{\boldsymbol{x}}_{r,OLS}(\tilde{\boldsymbol{x}}_c)) = \sum_{i=1}^{m} r_i(\hat{\boldsymbol{x}}_{OLS}(\tilde{\boldsymbol{x}}_c))^2$$

a non biased estimation of σ_ε^2 for the estimation of *r* parameter from the use of *m* measurements is thus given by

$$\hat{\sigma}_\varepsilon^2 = \frac{J_{OLS}(\hat{\boldsymbol{x}}_{r,OLS}(\tilde{\boldsymbol{x}}_c))}{n-r}$$

Confidence ellipse at confidence level α



Equation in centered $(\delta x_1, \delta x_2)$ axes

$$\boldsymbol{\delta x}^t.\boldsymbol{S}^t\boldsymbol{S}.\boldsymbol{\delta x} = \Delta^2$$

$$\Delta^2 = \chi^2_{1-\alpha}(2)\sigma_\varepsilon^2$$

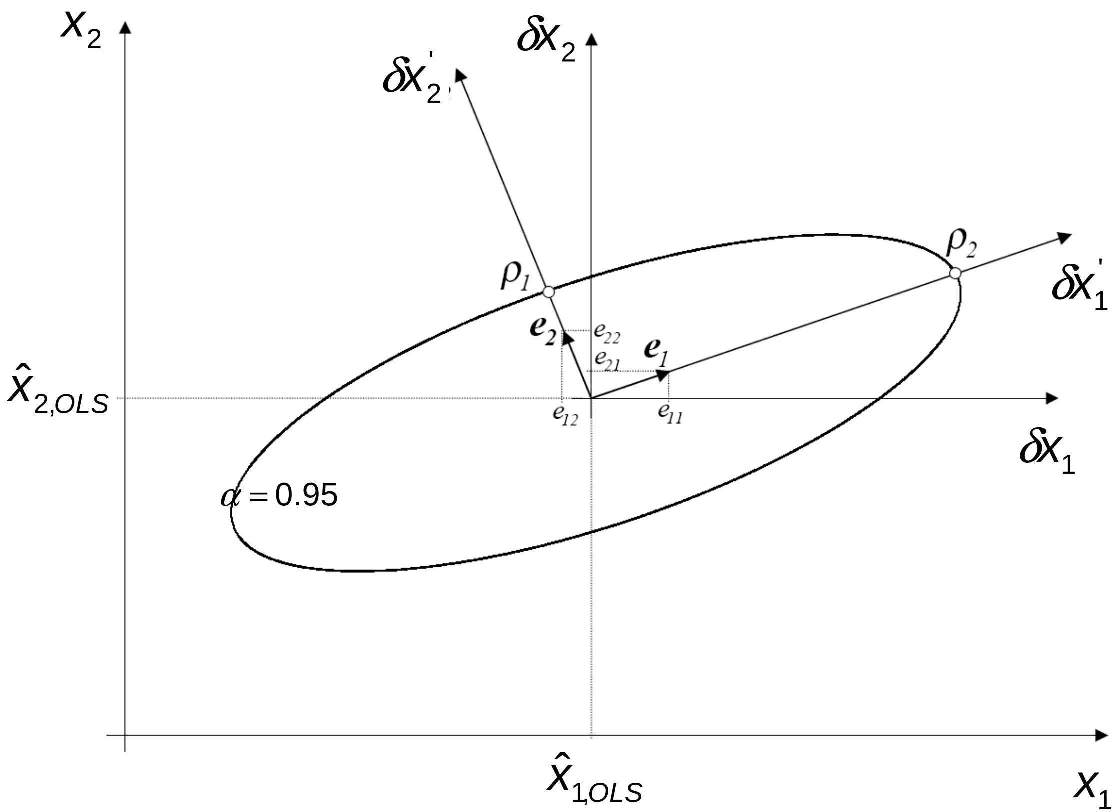
σ_ε^2 : variance of noise

$\chi^2_{1-\alpha}(2)$ *is computed by* chi2inv(alpha,2) *in MATLAB®* *for a confidence region at a level 95% (α=0.95)*

Equation in principal axes $(\delta x'_1, \delta x'_2)$: $\boldsymbol{\delta x'}^t.\boldsymbol{\Lambda}.\boldsymbol{\delta x'} = \Delta^2$

Where $\boldsymbol{\Lambda} = diag(\lambda_1, \lambda_2)$ contains the eigenvalues of $\boldsymbol{S}^t\boldsymbol{S}$

Confidence ellipse at confidence level α



Length of the two half axis are 'long' if eigenvalues are 'small' :

$$\rho_1 = \Delta / \sqrt{\lambda_1}$$

$$\rho_2 = \Delta / \sqrt{\lambda_2}$$

Notice : determinant of $\boldsymbol{S}^t\boldsymbol{S}$
Is given by

$$\det(\boldsymbol{S}^t\boldsymbol{S}) = \lambda_1\lambda_2$$

the area of the region inside the ellipse is given by

$$A = \pi.\rho_1.\rho_2 = \frac{\pi\chi^2_{1-\alpha}(2)\sigma^2_\varepsilon}{\sqrt{\det(\boldsymbol{S}^t\boldsymbol{S})}} = \frac{\pi\chi^2_{1-\alpha}(2)\sigma^2_\varepsilon}{\sqrt{\lambda_1\lambda_2}}$$

$$\boldsymbol{C}_1 = \left[\boldsymbol{S}_r^t \boldsymbol{S}_r\right]^{-1} \sigma_\varepsilon^2$$: 'absolute' covariance matrix of estimations

We can use $\boldsymbol{S}^* = \boldsymbol{S}.diag(\boldsymbol{x})$ instead of $\boldsymbol{S}$, whose the columns contain the reduced sensitivity coefficients (of same unit than model y_{mo})

$$S_k^*(t,\boldsymbol{x}) = x_k S_k(t,\boldsymbol{x}) = x_k \left.\frac{\partial y_{mo}(t,\boldsymbol{x})}{\partial x_k}\right|_{t,x_j \text{ for } j\neq k} = \left.\frac{\partial y_{mo}(t,\boldsymbol{x})}{\dfrac{\partial x_k}{x_k}}\right|_{t,x_j \text{ for } j\neq k}$$

$$\frac{\partial y_{mo}(t,\boldsymbol{x})}{\dfrac{\partial x_k}{x_k}}$$

Those reduced coefficients give the absolute variation of model due to relative variation on parameters

They can be compared between them and compared to the magnitude of noise

Roadmap for estimation #5 : the confidence

$$\boldsymbol{S}^* = \begin{bmatrix} S_1^*(t_1) & S_1^*(t_1) \\ \vdots & \vdots \\ S_i^*(t_i) & S_i^*(t_1) \\ \vdots & \vdots \\ S_m^*(t_m) & S_m^*(t_1) \end{bmatrix} = \begin{bmatrix} x_1 t_1 & x_2 \\ \vdots & \vdots \\ x_1 t_i & x_2 \\ \vdots & \vdots \\ x_1 t_m & x_2 \end{bmatrix}$$

With that reduced sensitivity matrix, we can build the relative covariance matrix

$$\boldsymbol{C}^* = \left[\boldsymbol{S}^{*t}_r \boldsymbol{S}^*_r\right]^{-1} \sigma_\varepsilon^2$$

$$\boldsymbol{C}^* = \begin{bmatrix} \left(\dfrac{\sigma_1}{\hat{x}_{1,OLS}}\right)^2 & \dfrac{\mathrm{cov}(e_{r1}, e_{r2})}{\hat{x}_{1,OLS}^2} & \cdots & \dfrac{\mathrm{cov}(e_{r1}, e_{rr})}{\hat{x}_{1,OLS}^2} \\ & \left(\dfrac{\sigma_2}{\hat{x}_{2,OLS}}\right)^2 & & \dfrac{\mathrm{cov}(e_{r2}, e_{rr})}{\hat{x}_{1,OLS}^2} \\ & & \ddots & \vdots \\ sym & & & \left(\dfrac{\sigma_r}{\hat{x}_{r,OLS}}\right)^2 \end{bmatrix}$$

Relative standard deviation of each estimation!

The same can be done with the ellipse equation

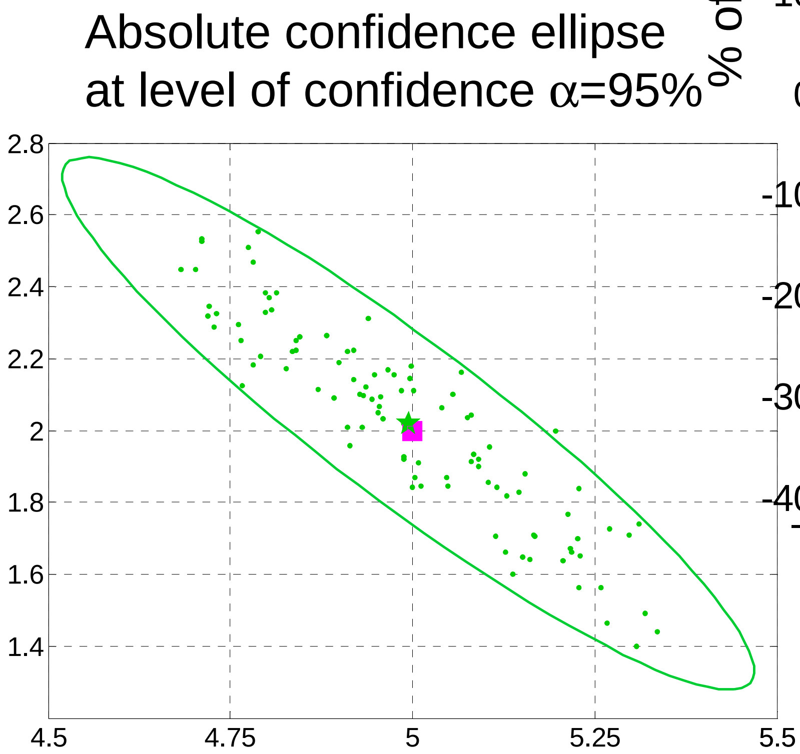
$$\boldsymbol{\delta x}^t.\boldsymbol{S}^t\boldsymbol{S}.\boldsymbol{\delta x} = \Delta^2$$

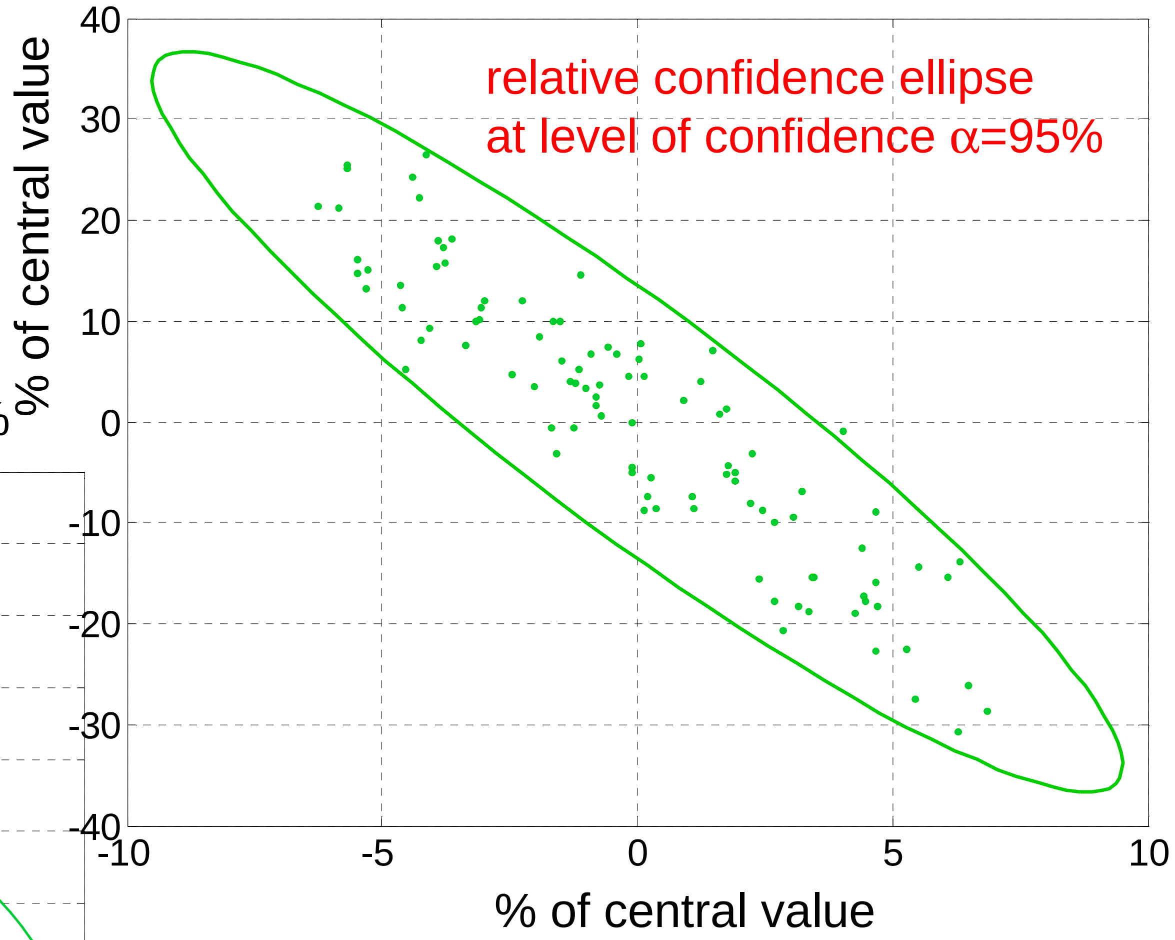
with $\boldsymbol{S}^* = \boldsymbol{S}.diag(\boldsymbol{x}^{nom}) \rightarrow \boldsymbol{S} = \boldsymbol{S}^* diag(\boldsymbol{x}^{nom})^{-1}$

gives $\underbrace{\left(diag(\boldsymbol{x}^{nom})^{-1}\boldsymbol{\delta x}\right)}_{\frac{\delta x}{x^{nom}}}{}^t.\boldsymbol{S}^t\boldsymbol{S}.diag(\boldsymbol{x}^{nom})^{-1}\boldsymbol{\delta x} = \Delta^2$

$\frac{\delta x}{x^{nom}} \longrightarrow$ Relative confidence ellipse (in %)

Cloud with absolute and relative Ellipses pour D.M.





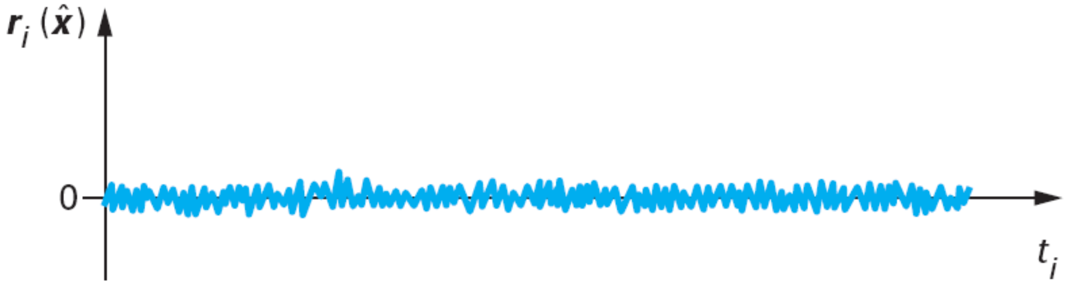
![Metti 5 ROSCOFF 2011]

Residuals analysis

Last, for qualifying the quality of estimation : the residuals analysis

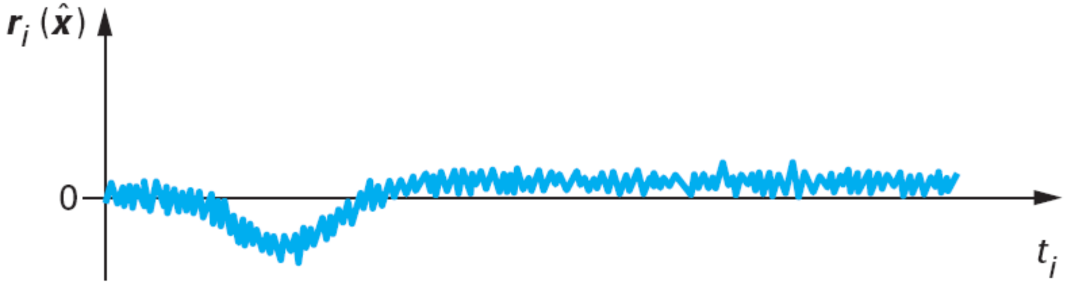
$$\boldsymbol{r}(\hat{\boldsymbol{x}}) = \boldsymbol{y} - \boldsymbol{y}_{mo}(\hat{\boldsymbol{x}}) = \left[y_1 - y_{mo,1}(t_1, \hat{\boldsymbol{x}}) \quad \ldots \quad y_i - y_{mo,i}(t_i, \hat{\boldsymbol{x}}) \quad \ldots \quad y_m - y_{mo,m}(t_m, \hat{\boldsymbol{x}}) \right]^t$$

Difference between measurements and model response with optimal parameters must 'look like' noise measurement : 'the right model with the right parameters must explain the measurements except its random part'



ⓐ Uncorrelated residuals

→ OK, see if the variances are not too large



ⓑ Signed residuals

→ Not OK : there may be a problem in the model or in the measurements

Roadmap for estimation : anticipating the performances

The danger has been identified : the inversion of $\mathbf{S}^t\mathbf{S}$ or $\mathbf{S}^{*t}\mathbf{S}^*$

It has been shown that the matrix $\mathbf{S}^t\mathbf{S}$ is fundamental in the processus of parameter estimation :

- it has to be inverted to achieve the OLS estimation

- it also has to be inverted to compute the covariance matrix. The inverse of $\mathbf{S}^t\mathbf{S}$ respectively $\mathbf{S}^{*t}\mathbf{S}^*$ plays the role of "noise amplification“, in absolute or, respectively, in relative values

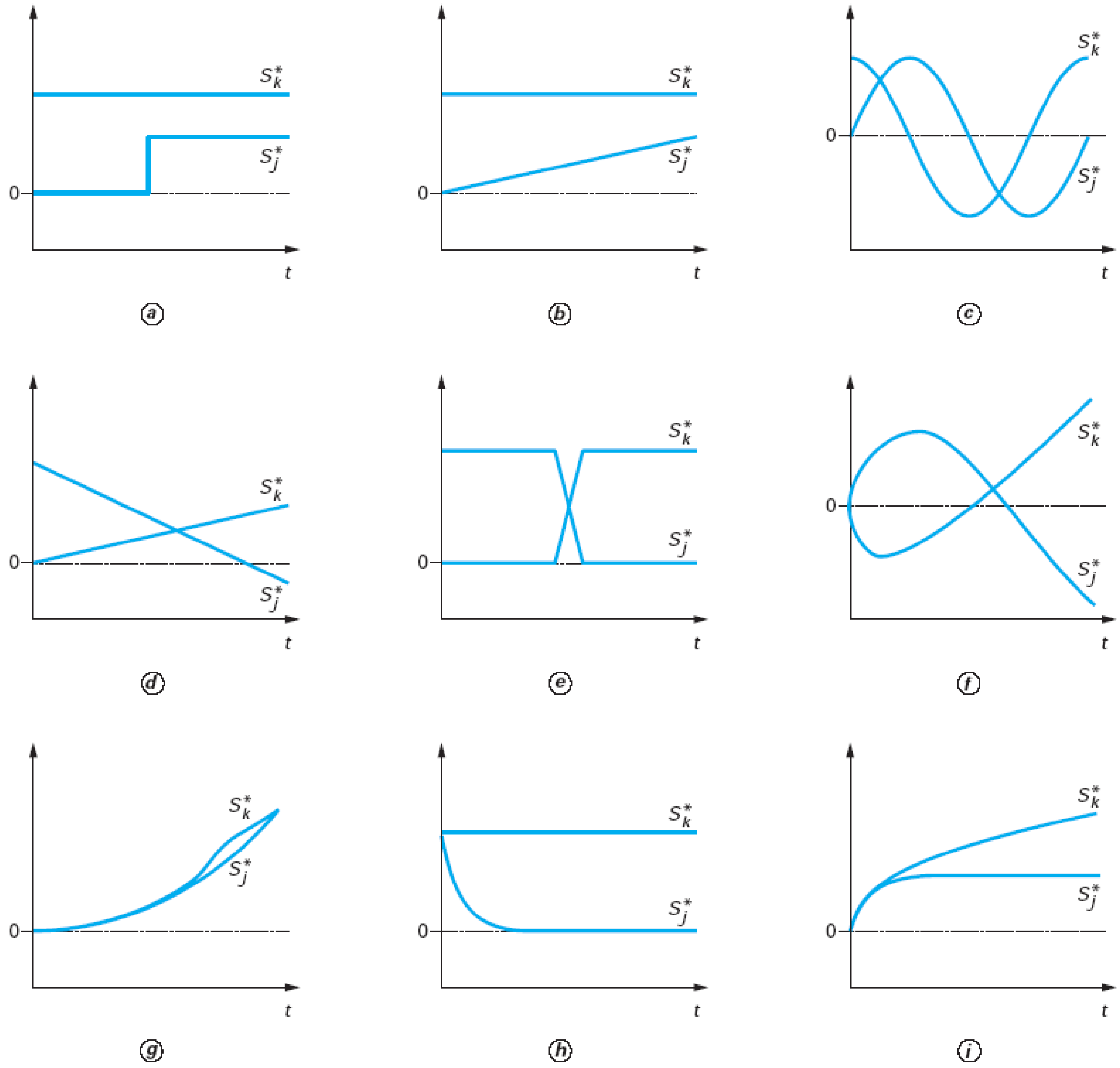
- the eigenvalues of $\mathbf{S}^t\mathbf{S}$ enable the calculation of the lengths of the half principal axis of the elliptical confidence region

- the determinant of $\mathbf{S}^t\mathbf{S}$ enables the calculation of the area of the elliptical confidence region

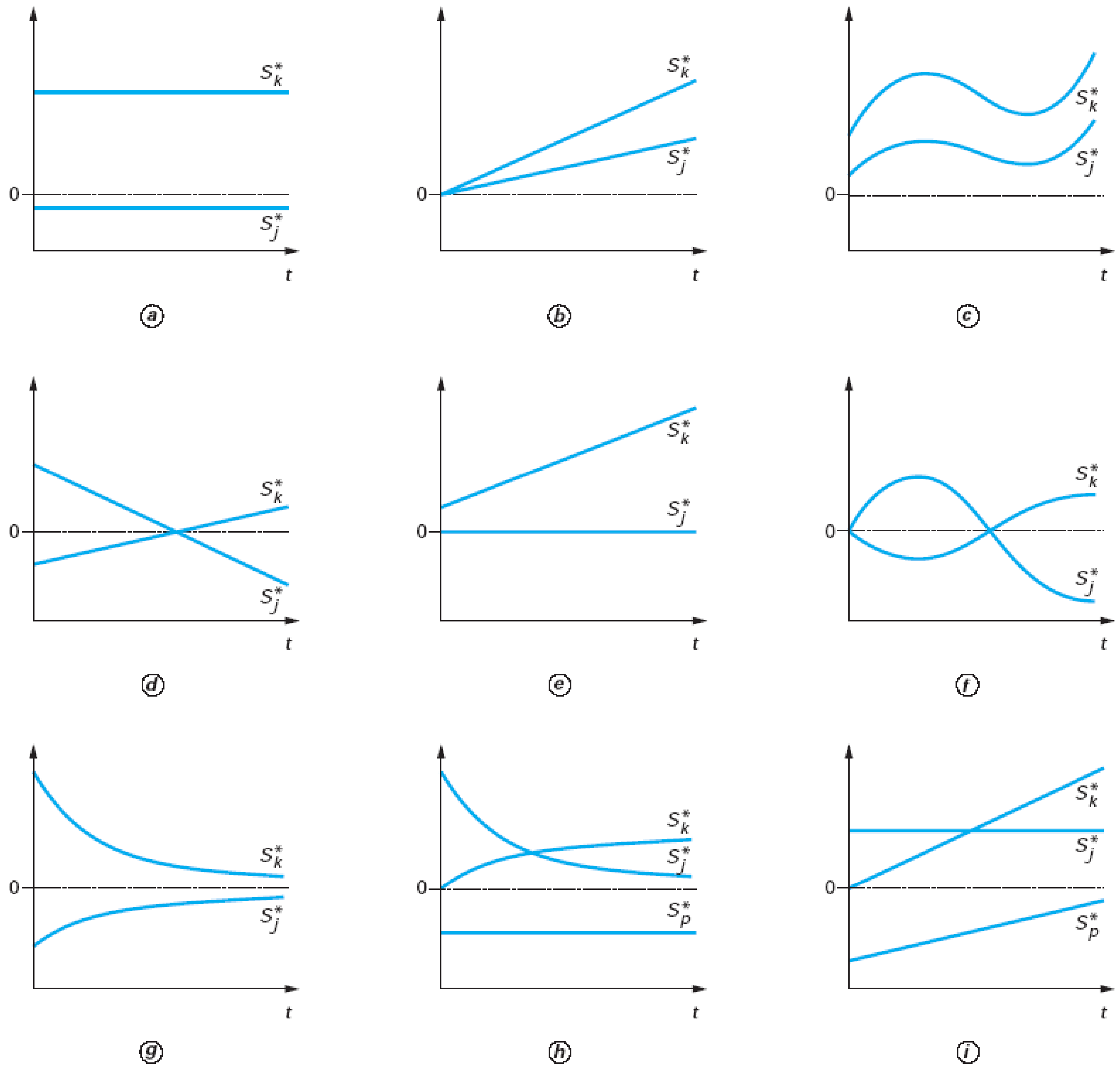
$\longrightarrow$ Illustration in our examples, using too the conditioning number of $\mathbf{S}^t\mathbf{S}$

SENSITIVITY COEFFICIENTS composing **S** MUST BE LINEARLY INDEPENDENT

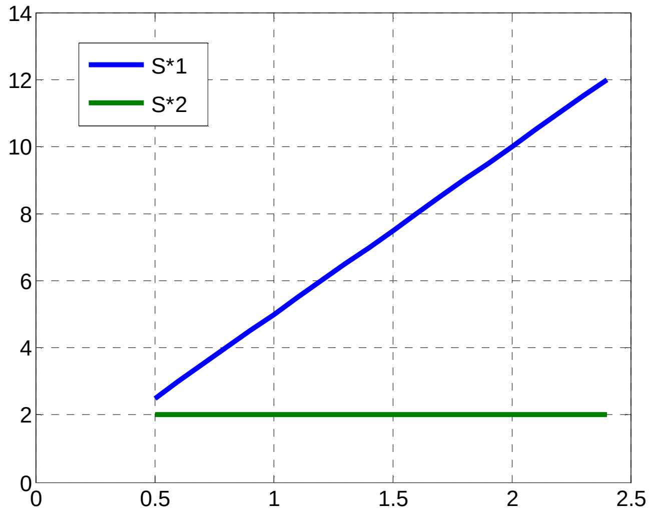
‘Graphical’ analysis of reduced sensitivity coefficients : independent case



‘Graphical’ analysis of reduced sensitivity coefficients : dependent case

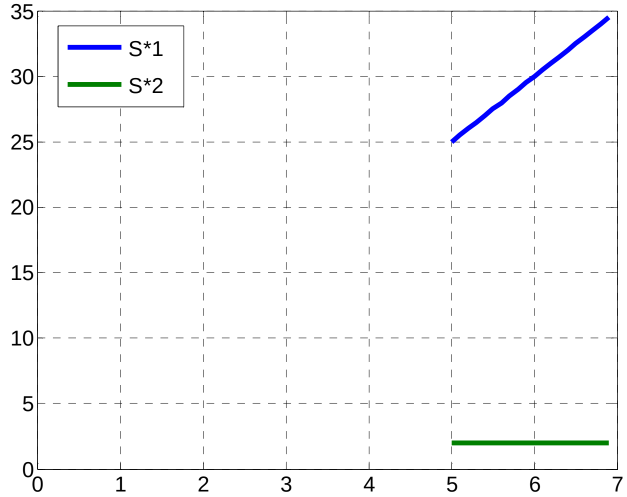


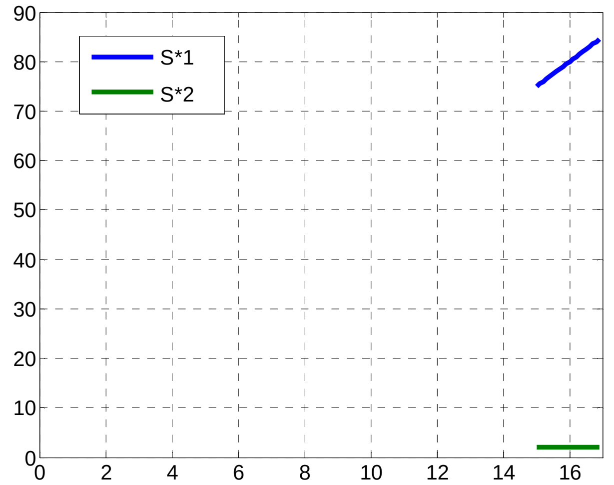
Analysis of sensitivity coefficients in our triple situation



Student D.M.

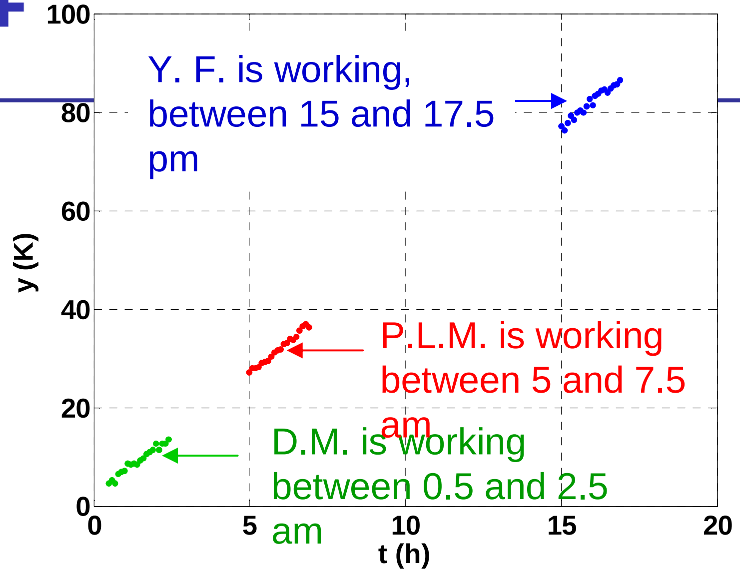
Student P.L.M.

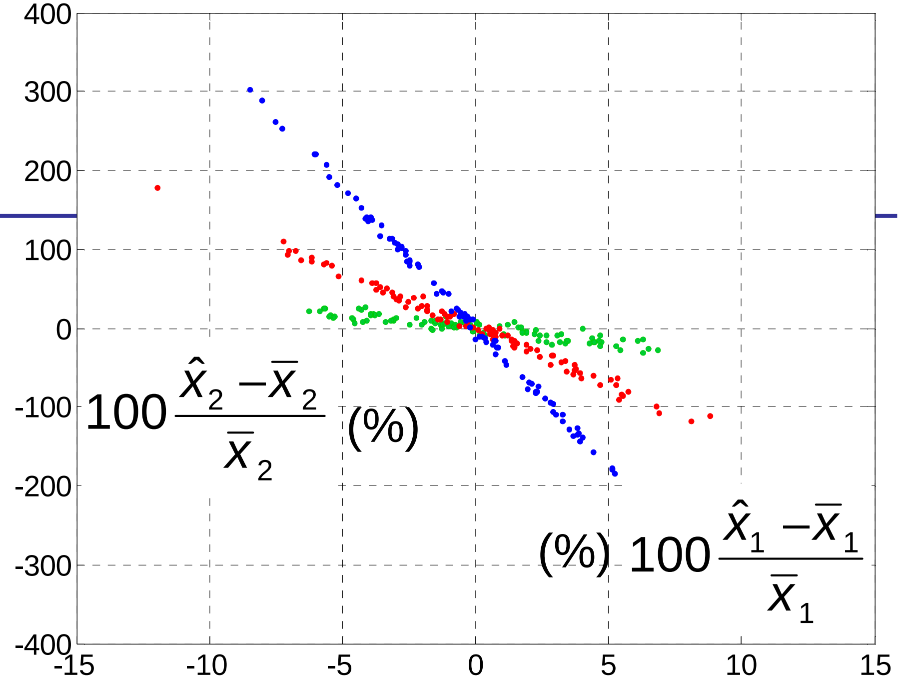




Student Y.F.

Summary of Indicators





Student	D.M.	P.L.M.	Y.F.
Time range (h)	0.5 h -2.5 h	5 h -7.5 h	15 h -17.5 h
λ_{min} of $S^{*t}S^*$ ↑	1.03e1	6.5e-1 ↓	6.2e-2 ↓
λ_{max} of $S^{*t}S^*$ ↓	1.29e3	1.8e4 ↑	1.3e5 ↑
$\det(S^{*t}S^*)$ ↑	1.34e4	1.18e4 ↓	8.0e3 ↓
Ellipse area ↓	3.52e-4	3.99e-4 ↑	5.9e-4 ↑
$\text{cond}(S^{*t}S^*) = \lambda_{max} / \lambda_{min}$ ↓	1.24e2	2.78e4 ↑↑	2.1e6 ↑↑
ρ_{12} ↓	-0.93	-0.995 ↑	-0.993 ↔

All indicators are ‘red’ : it is not a good design to have large times for simultaneous estimation of x_1 and x_2