

**Outils
pour
la prévision des performances
et
la conception optimales des expériences**

8 Octobre 2004

C.Gervaise

Plan de la présentation

I] Introduction

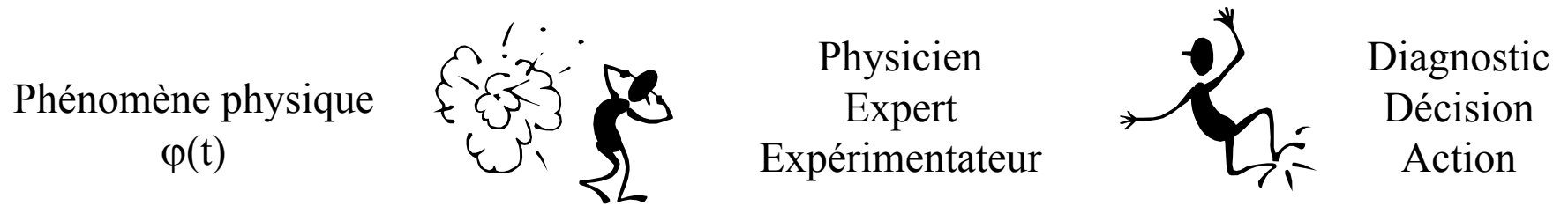
II] Prévisions de Performances

III] Conception Optimale des Expériences

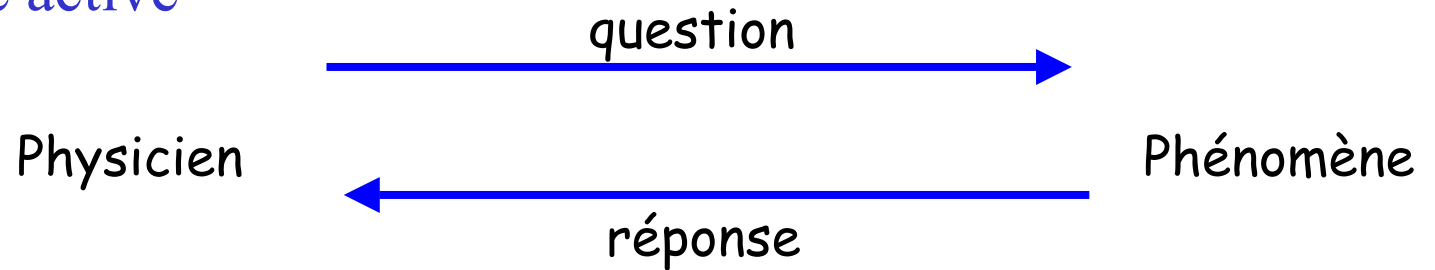
IV] Exemples d'applications passées et à venir

Introduction

Problématique de la caractérisation d'un phénomène physique inconnu



Approche active

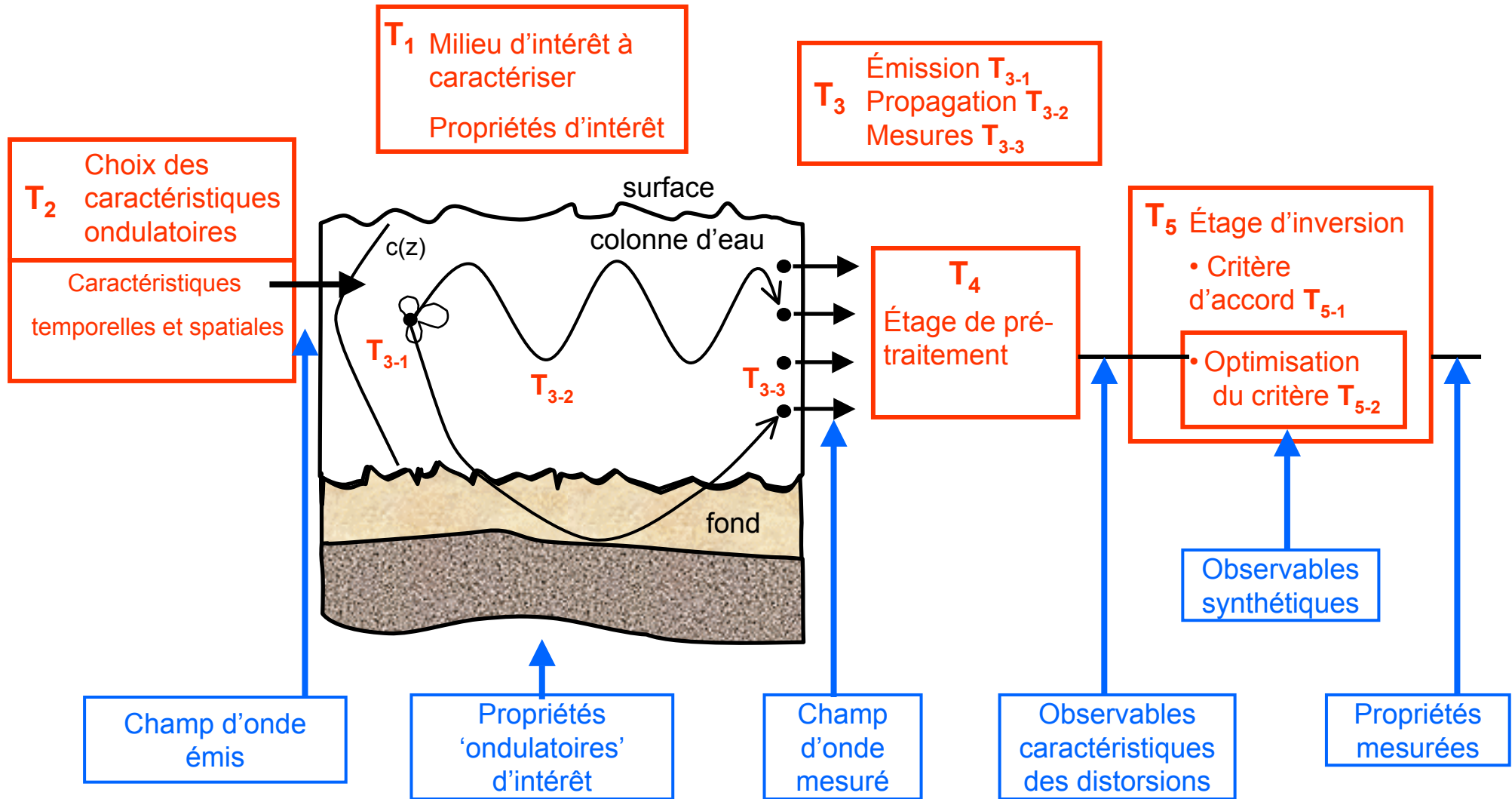


Approche passive



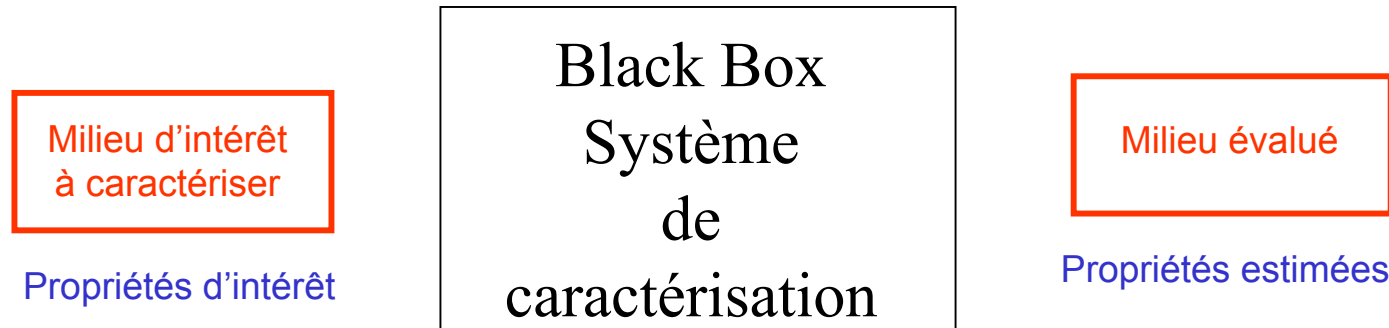
Introduction

Problématique de la caractérisation d'un environnement inconnu



Introduction

Problématique de la caractérisation d'un environnement inconnu – Un bon système ?



'Précision'

Bon conditionnement
au sens de Hadamard

Existence

Unicité

Stabilité face au bruit

'Robustesse'

Capacité du système à
fonctionner en présence
d'écarts 'imprévus'

Incertain est modélisable par
un sur-paramétrage du
problème

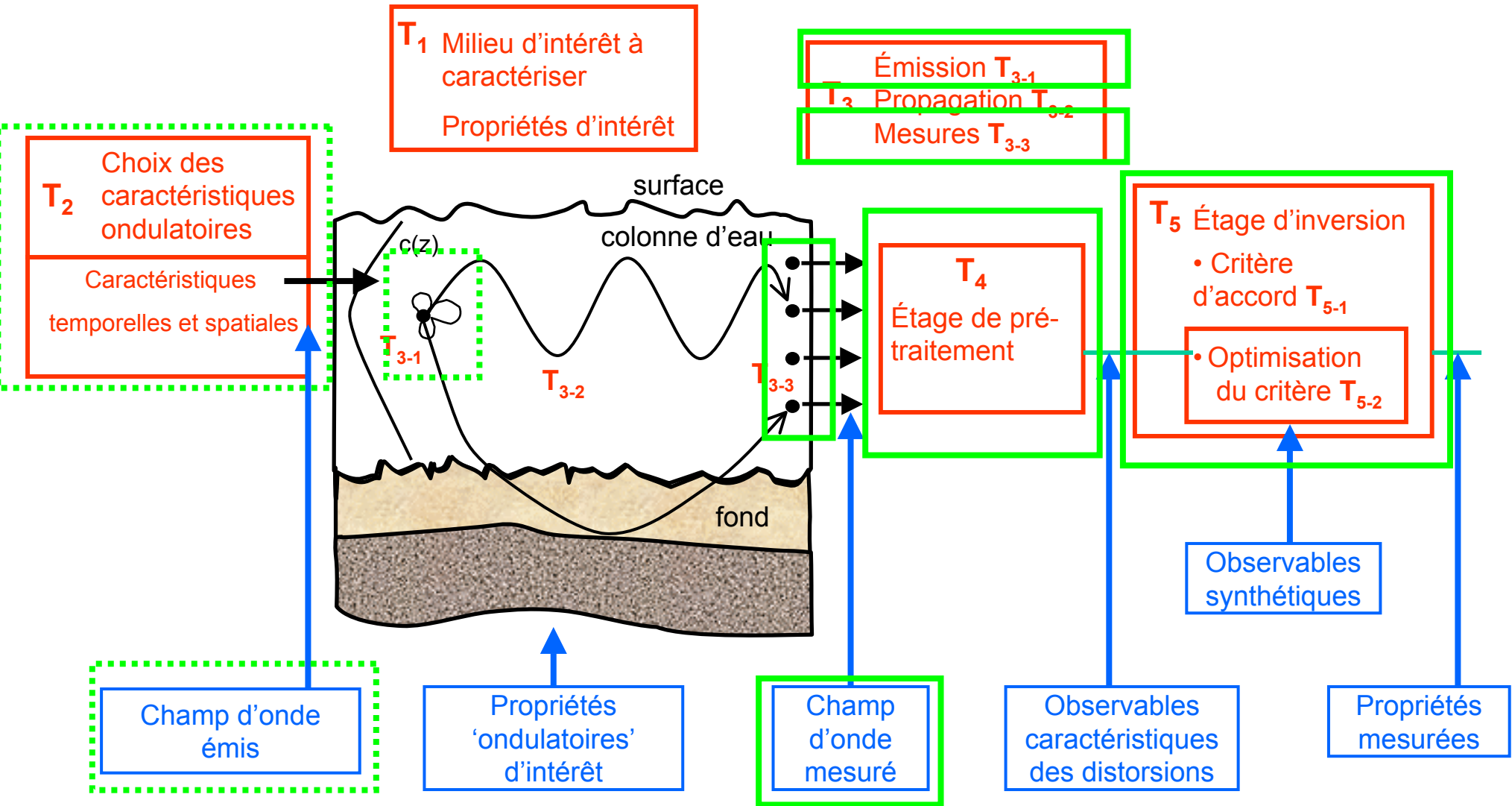
'Complexité'

Complexité algorithmique

Complexité de la
géométrie expérimentale

Introduction

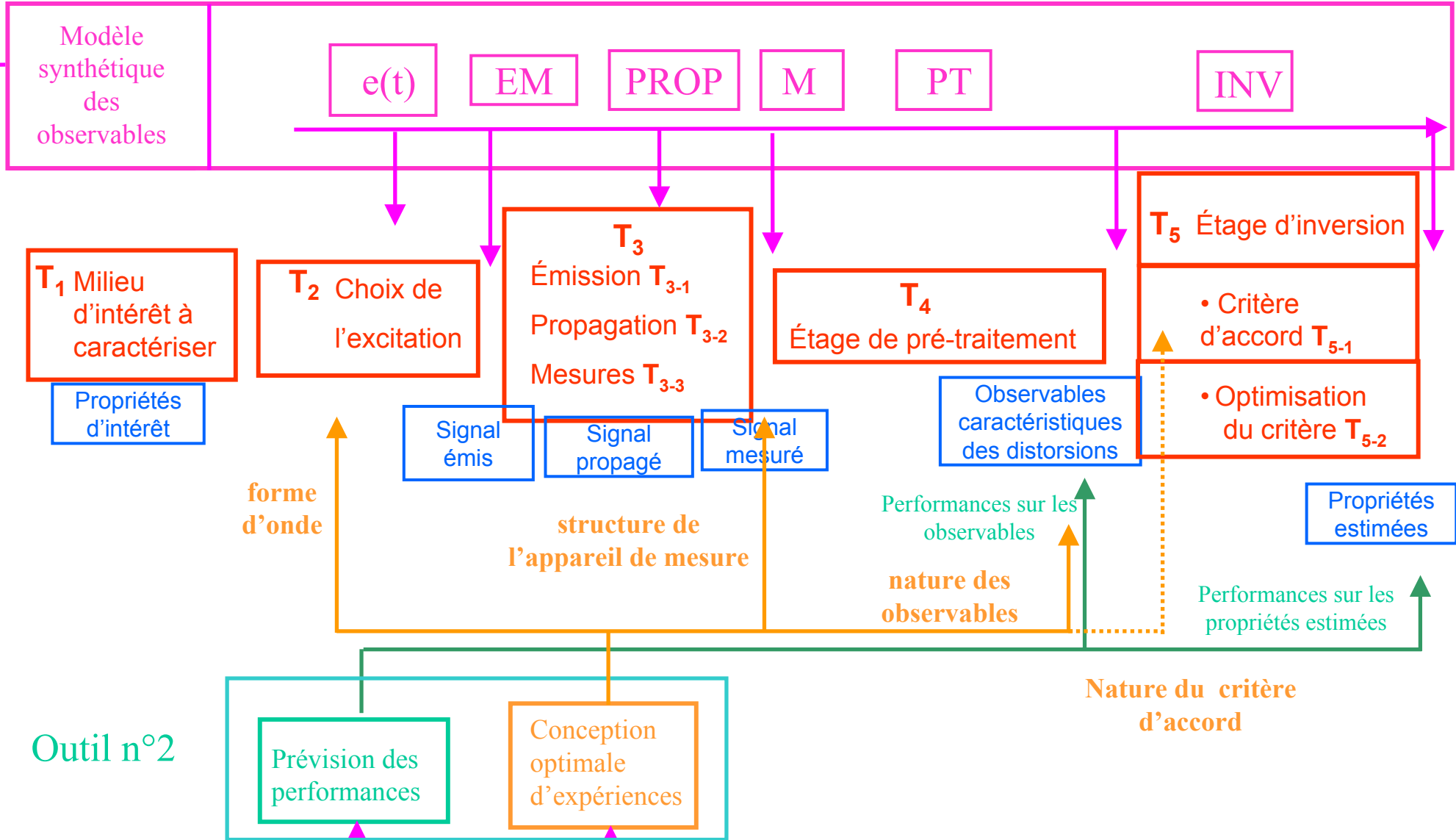
Problématique de la caractérisation d'un environnement inconnu



Introduction

Problématique de la caractérisation d'un environnement inconnu – 2 outils complémentaires

Outil n°1



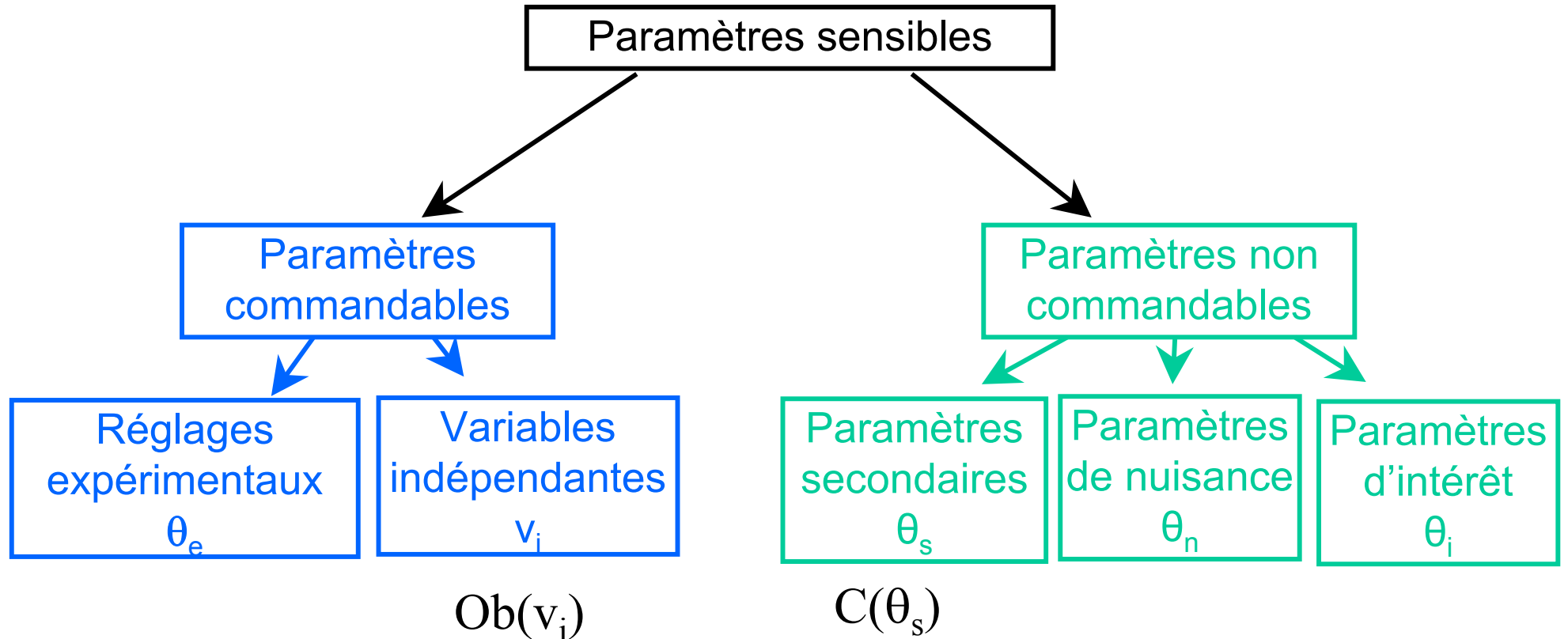
Outil n°2

Prévision des performances

Conception optimale d'expériences

Introduction

Problématique de la caractérisation d'un environnement inconnu – paramétrage



Introduction

Problématique de la caractérisation d'un environnement inconnu – Modèle des observables

Les outils proposés requièrent la connaissance du lien reliant les observables et les paramètres du problème : Modèle de données

$$\mathbf{O}_b(v_i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n) \# f_{Ob}(\alpha; v_i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n)$$

$$\mathbf{O}_b(v_i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n) \# N(\mu(v_i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n), COV(v_i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n))$$

$$\tilde{\mathbf{O}}_b(v_i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n) = \mathbf{f}(v_i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n) + \tilde{\mathbf{D}}(v_i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n)$$

$$\mathbf{f}(i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n) = PT\{M\{PROP\{EM\{e(t)\}\}\}\}$$

Introduction

Problématique de la caractérisation d'un environnement inconnu – Formulations

Formulation du problème d'estimation

Des **observables** sont recueillies en fonction **des variables indépendantes** \mathbf{v}_i comme le résultat d'une expérimentation effectuée par un appareil de mesures réglé avec les **réglages expérimentaux** θ_e . Ces Observables sont utilisées pour retrouver **les paramètres d'intérêt** θ_i en présence de **paramètres secondaires** θ_s et de **paramètres de nuisance** θ_n .

Si un modèle théorique d'observables est connu

Formulation du problème de prévision de performance

La **prévision de performances** consiste à calculer **les propriétés statistiques** comme le biais et la matrice de covariance **des paramètres d'intérêt** θ_i pour une **série d'observables recueillies en fonction de** \mathbf{v}_i avec un **appareil de mesure réglé par** θ_e en présence de **paramètres secondaires** θ_s et de **paramètres de nuisance** θ_n . Ces propriétés statistiques apparaissent comme une fonction de \mathbf{v}_i , θ_e , θ_i , θ_s et θ_n .

Formulation du problème de conception optimale des expériences

La **conception optimale d'expériences** consiste à choisir les valeurs des **variables indépendantes** \mathbf{v}_i et des **réglages expérimentaux** θ_e qui **optimisent** les performances d'estimation **des paramètres d'intérêt** θ_i afin que les **observables** apportent **le maximum d'information sur ces paramètres** et soient **peu sensibles aux paramètres secondaires** θ_s et aux **paramètres de nuisance** θ_n .

Plan de la présentation

I] Introduction

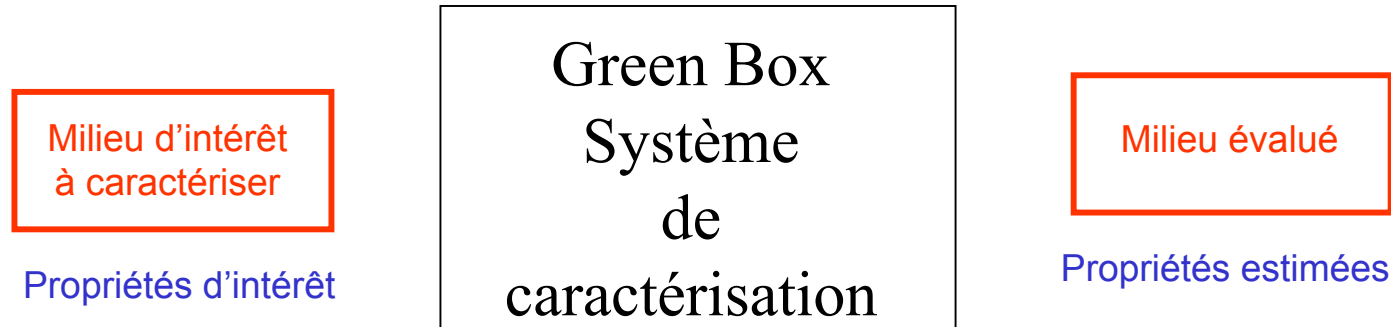
II] Prévisions de Performances

III] Conception Optimale des Expériences

IV] Exemples d'applications passées et à venir

Prévision de performance

Limitation du champs d'étude



'Précision'

Bon conditionnement
au sens de Hadamard

Existence

Unicité

Stabilité face au bruit

'Robustesse'

Capacité du système à
fonctionner en présence
d'écarts 'imprévus'

Incertain est modélisable par
un sur-paramétrage du
problème

'Complexité'

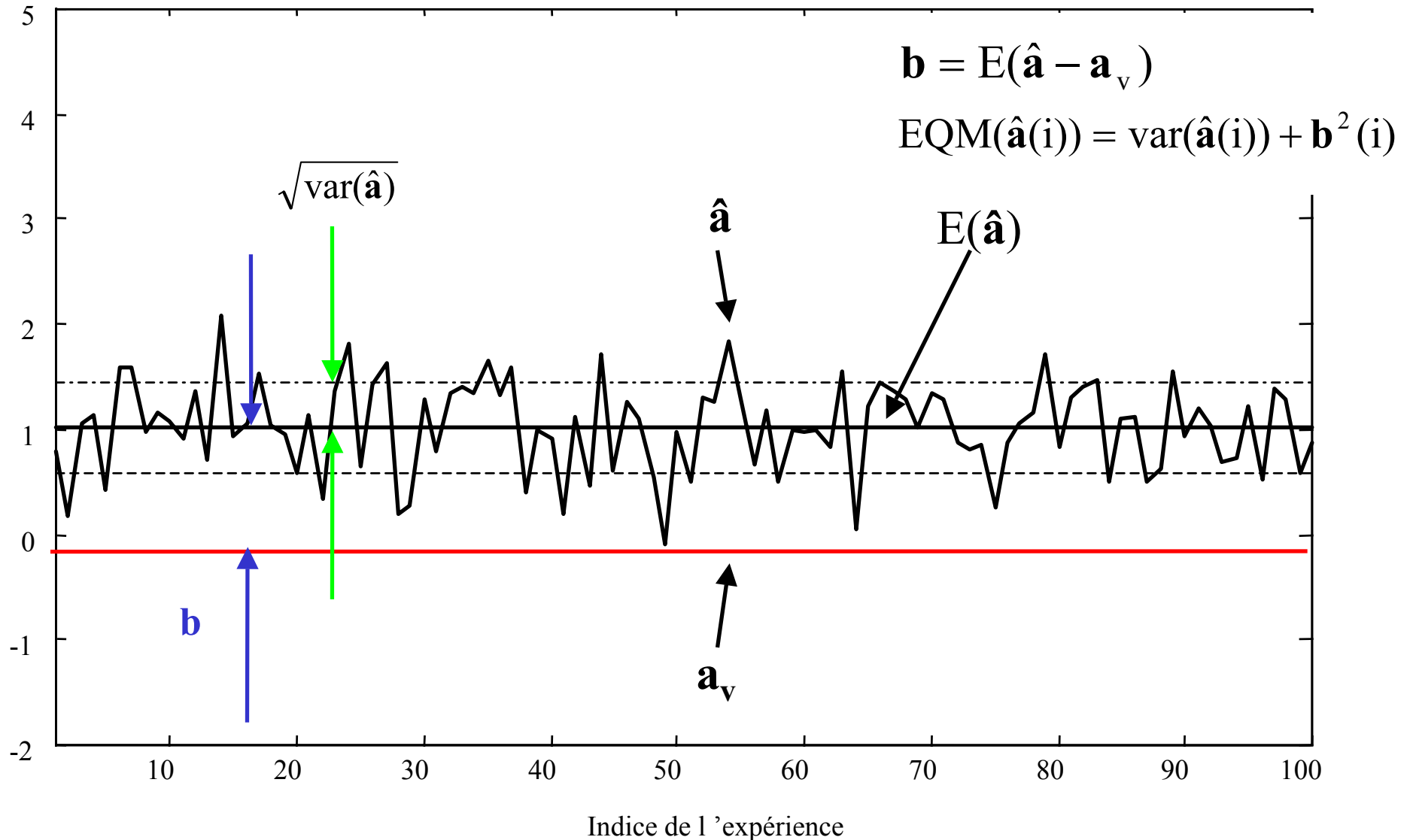
Complexité algorithmique

Complexité de la
géométrie expérimentale

Prévision de performance

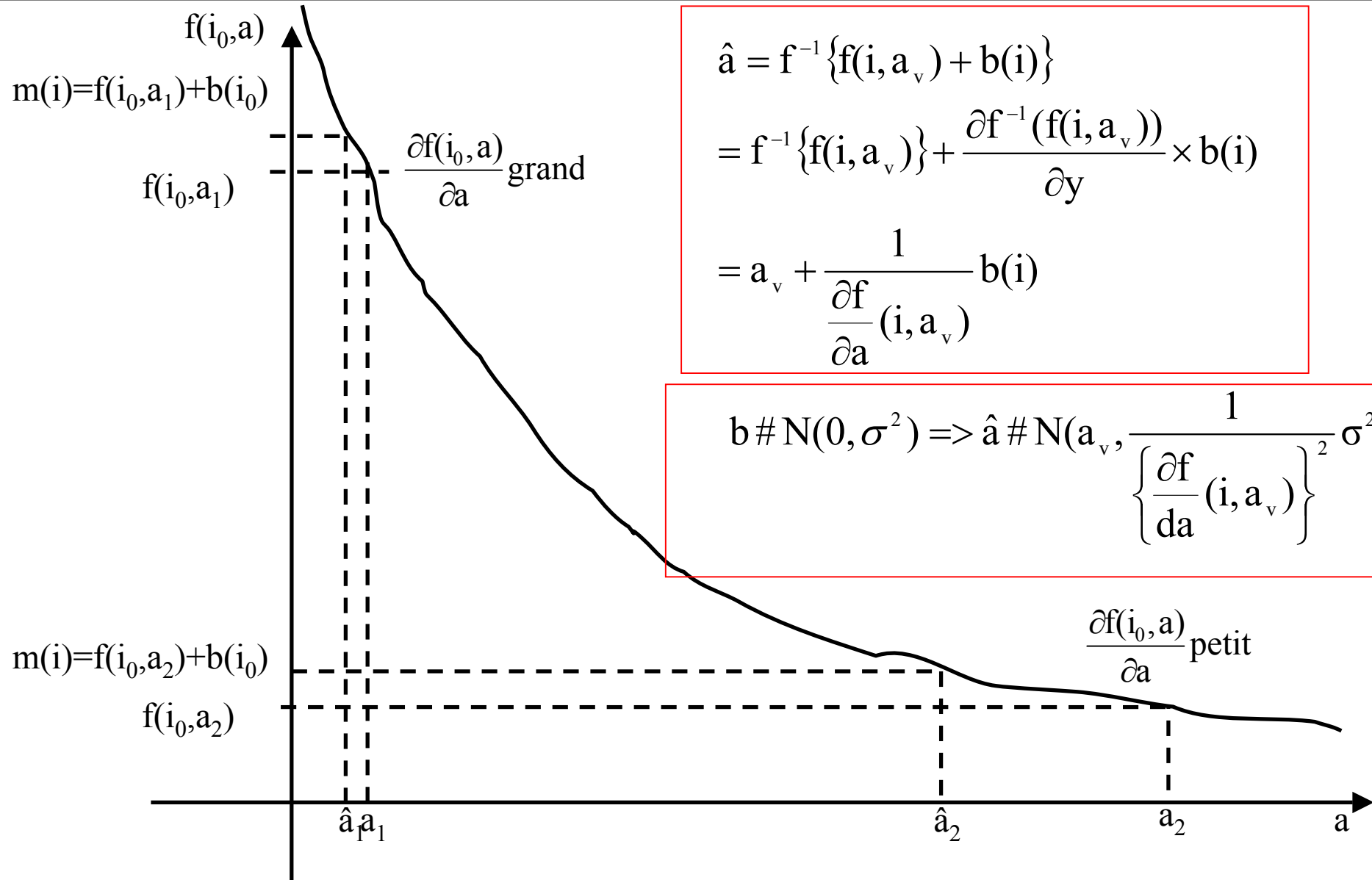
Limitation du champs d'étude

Stabilité face au bruit => estimateur non biaisé à minimum de variance



Prévision de performance

Approche intuitive 1 : cas de l'estimation d'un paramètre – influence de la quantité d'information de la mesure sur le paramètre



$$\hat{a} = f^{-1}\{f(i, a_v) + b(i)\}$$

$$= f^{-1}\{f(i, a_v)\} + \frac{\partial f^{-1}(f(i, a_v))}{\partial y} \times b(i)$$

$$= a_v + \frac{1}{\frac{\partial f}{\partial a}(i, a_v)} b(i)$$

$$b \# N(0, \sigma^2) \Rightarrow \hat{a} \# N\left(a_v, \frac{1}{\left\{\frac{\partial f}{\partial a}(i, a_v)\right\}^2} \sigma^2\right)$$

Prévision de performance

Approche intuitive 2 : cas de l'estimation de deux paramètres – influence de la corrélation entre les deux paramètres

$$m(1) = f(1, \alpha_1, \alpha_2) + b(1) = \alpha_1 + \alpha_2 + b(1)$$

$$m(2) = f(2, \alpha_1, \alpha_2) + b(2) = \alpha_1 + \alpha_2 + b(2)$$

$$\frac{\partial}{\partial \alpha_1}(\vec{f}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \frac{\partial}{\partial \alpha_2}(\vec{f}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix};$$

$$\Delta \vec{m} = \frac{\partial \vec{f}}{\partial \alpha_1}(\alpha_1^0) \Delta \alpha_1 + \frac{\partial \vec{f}}{\partial \alpha_2}(\alpha_2^0) \Delta \alpha_2 + \delta \vec{b}$$

une variation donnée de α_1 ou α_2 induit la même variation (colinéaire) sur le vecteur m .

Les paramètres α_1 ou α_2 sont dits corrélés à travers l'observation de m .

$$m(1) = f(1, \alpha_1, \alpha_2) + b(1) = \alpha_1 + \alpha_2 + b(1)$$

$$m(2) = f(2, \alpha_1, \alpha_2) + b(2) = \alpha_1 - \alpha_2 + b(2)$$

$$\frac{\partial}{\partial \alpha_1}(\vec{f}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \frac{\partial}{\partial \alpha_2}(\vec{f}) = \begin{bmatrix} 1 \\ -1 \end{bmatrix};$$

$$\Delta \vec{m} = \frac{\partial \vec{f}}{\partial \alpha_1}(\alpha_1^0) \Delta \alpha_1 + \frac{\partial \vec{f}}{\partial \alpha_2}(\alpha_2^0) \Delta \alpha_2 + \delta \vec{b}$$

une variation donnée de α_1 ou α_2 induit des variations différentes (orthogonales) sur le vecteur m .

Les paramètres α_1 ou α_2 sont dits décorrélés à travers l'observation de m .

Prévision de performance

Prévision de performances – paramètres d'intérêt – stabilité face au bruit

Théorème de Crao et Ramer

Soit \vec{O}_b un vecteur d'observations aléatoires fruit d'une expérience dépendant d'un vecteur de paramètres θ_i à travers le modèle de données :

$$\vec{O}_b \# P_{\vec{O}_b}(\vec{x}; \theta_i)$$

Si la loi de densité de probabilité suit la condition de régularité suivante :

$$E_{\vec{x}} \left\{ \frac{\partial P_{\vec{O}_b}(\vec{x}; \theta_i)}{\partial \theta} \right\} = \vec{0} \forall \theta_i$$

Alors quel que soit un estimateur non biaisé de θ_i , sa matrice de covariance vérifie :

$$C(\hat{\theta}_i) - I^{-1}(\hat{\theta}_i) \geq 0 \quad (C(\hat{\theta}_i) - I^{-1}(\hat{\theta}_i) \text{ est une matrice semi-définie positive})$$

Où $I(\theta_i)$ est la matrice d'information de Fisher-Price

$$I(\theta_i) = -E_{\vec{x}} \left(\frac{\partial^2 \ln(P_{\vec{O}_b}(\vec{x}, \theta_i))}{\partial \theta_i \partial \theta_i^t} \right) \Big|_{\theta_i = \theta_i \text{ vrai}}$$

Prévision de performance

Prévision de performances – paramètres d'intérêt – stabilité face au bruit

$$\mathbf{O}_b(v_i, \theta_e, \theta_i, \theta_s, \theta_n) \# N(\mu(v_i, \theta_e, \theta_i, \theta_s, \theta_n), COV(v_i, \theta_e, \theta_i, \theta_s, \theta_n))$$

$$\tilde{\mathbf{O}}_b(v_i, \theta_e, \theta_i, \theta_s, \theta_n) = \mathbf{f}(v_i, \theta_e, \theta_i, \theta_s, \theta_n) + \tilde{\mathbf{D}}(v_i, \theta_e, \theta_i, \theta_s, \theta_n)$$

$$\mathbf{f}(i, \theta_e, \theta_i, \theta_s, \theta_n) = PT \{M \{PROP \{EM \{e(t)\}\}\}\}$$

2 Formes de la matrice d'information de Fisher

$$\vec{\mathbf{O}}_b \# N(\vec{\mu}(\theta_i), C(\theta_i))$$

$$[I(\theta_i)]_{lm} = \frac{\partial \vec{\mu}}{\partial \theta_i(l)}{}^T C^{-1}(\theta_i) \frac{\partial \vec{\mu}}{\partial \theta_i(m)} + 0.5 \text{Tr} \left[C^{-1}(\theta_i) \frac{\partial C(\theta_i)}{\partial \theta_i(m)} C^{-1}(\theta_i) \frac{\partial C(\theta_i)}{\partial \theta_i(l)} \right]$$

$$\vec{\mathbf{O}}_b \# N(\vec{\mu}(\theta_i), \sigma^2 II)$$

$$[I(\theta_i)]_{lm} = \frac{1}{\sigma^2} \frac{\partial \vec{\mu}}{\partial \theta_i(l)}{}^T \frac{\partial \vec{\mu}}{\partial \theta_i(m)}$$

Facilité de calcul si le modèle de données est connu

Prévision de performance

Prévision de performances – paramètres d'intérêt – retour sur les approches intuitives 1 et 2

Illustrons le calcul des Bornes de Cramer Rao pour l'estimation d'un vecteur de deux paramètres $\boldsymbol{\theta}_i = [\theta_1 \ \theta_2]^T$ et un jeu d'observables distribué suivant une loi normale de moyenne $\mathbf{f}(\boldsymbol{\theta}_i)$ et de matrice de covariance $\mathbf{C}_D = \sigma^2 \mathbf{I}$.

$$I(\boldsymbol{\theta}_i) = \frac{1}{\sigma^2} \mathbf{X}^T \mathbf{X} \quad \text{avec} \quad \mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2] = \begin{bmatrix} \frac{\partial \mathbf{f}(\boldsymbol{\theta}_i)}{\partial \theta_1^T} & \frac{\partial \mathbf{f}(\boldsymbol{\theta}_i)}{\partial \theta_2^T} \end{bmatrix}$$

$$I(\boldsymbol{\theta}_i) = \frac{1}{\sigma^2} \begin{bmatrix} \mathbf{X}_1^T \mathbf{X}_1 & \mathbf{X}_1^T \mathbf{X}_2 \\ \mathbf{X}_2^T \mathbf{X}_1 & \mathbf{X}_2^T \mathbf{X}_2 \end{bmatrix} \quad X_l^T X_m = \sum_{i=1}^N \frac{\partial f(i, \theta_i)}{\partial \theta_i(l)} \frac{\partial f(i, \theta_i)}{\partial \theta_i(m)}$$

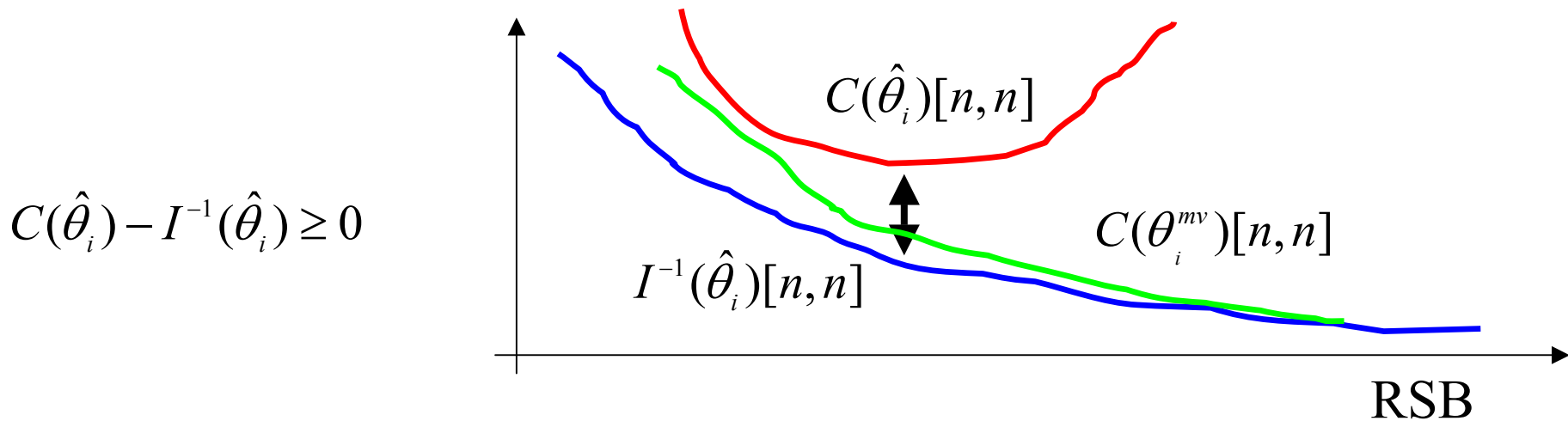
$\mathbf{X}_m^T \mathbf{X}_m$ = Energie de l'information apportée par les observables sur $\theta_i(m)$

$\mathbf{X}_m^T \mathbf{X}_n$ = Energie de l'information commune – la confusion apportée par les observables conjointement sur $\theta_i(m)$ et $\theta_i(n)$

$$I^{-1}(\boldsymbol{\theta}_i) = \sigma^2 \begin{bmatrix} \frac{1}{a(1-c_{\text{lim}}^2)} & -\frac{c_{\text{lim}}}{\sqrt{ab}(1-c_{\text{lim}}^2)} \\ -\frac{c_{\text{lim}}}{\sqrt{ab}(1-c_{\text{lim}}^2)} & \frac{1}{b(1-c_{\text{lim}}^2)} \end{bmatrix} \quad \begin{aligned} a &= \mathbf{X}_1^T \mathbf{X}_1 \\ b &= \mathbf{X}_2^T \mathbf{X}_2 \\ c &= \mathbf{X}_1^T \mathbf{X}_2 \end{aligned} \quad c_{\text{lim}} = -\frac{c}{\sqrt{ab}}$$

Prévision de performance

Prévision de performances – Borne inférieure à la matrice de covariance de tout estimateur non biaisé



Si une expérience fournit une réalisation des observables \vec{O}_b égale à $\vec{x}_{réal}$

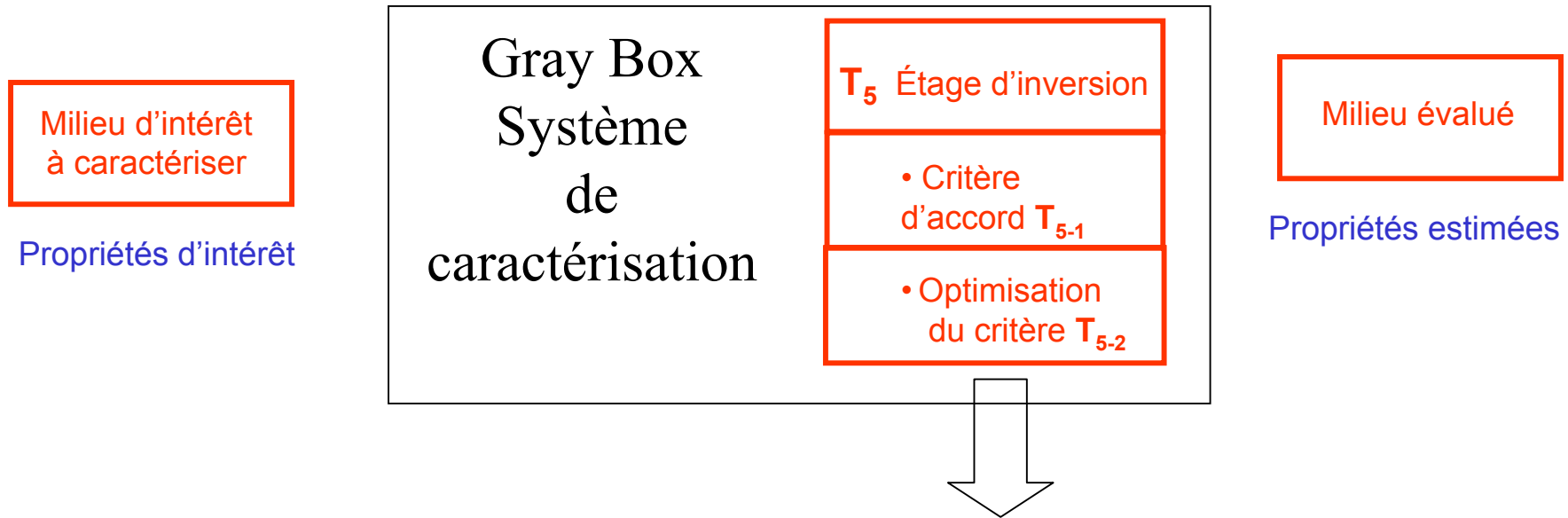
$$\theta_i^{mv} = \arg(\max_{\theta_i} (P_{\vec{O}_b}(\vec{x}_{réal}, \theta_i)))$$

est asymptotiquement non biaisé et efficace.

Etant donné que l'on est capable de trouver un estimateur dont la matrice de covariance tend vers l'inverse de la matrice d'information de Fisher, celle-ci peut être considérée comme une bonne évaluation des performances statistiques de l'inversion

Prévision de performance

Prévision de performances – à la recherche d'un estimateur



- comment mettre en œuvre l'estimateur du maximum de vraisemblance
- rechercher un estimateur alternatif plus robuste et plus facile à mettre en œuvre que le maximum de vraisemblance mais gardant le même niveau de précision
- rechercher un estimateur plus précis à faible RSB

Non abordé ici, peut être pour un autre séminaire ?

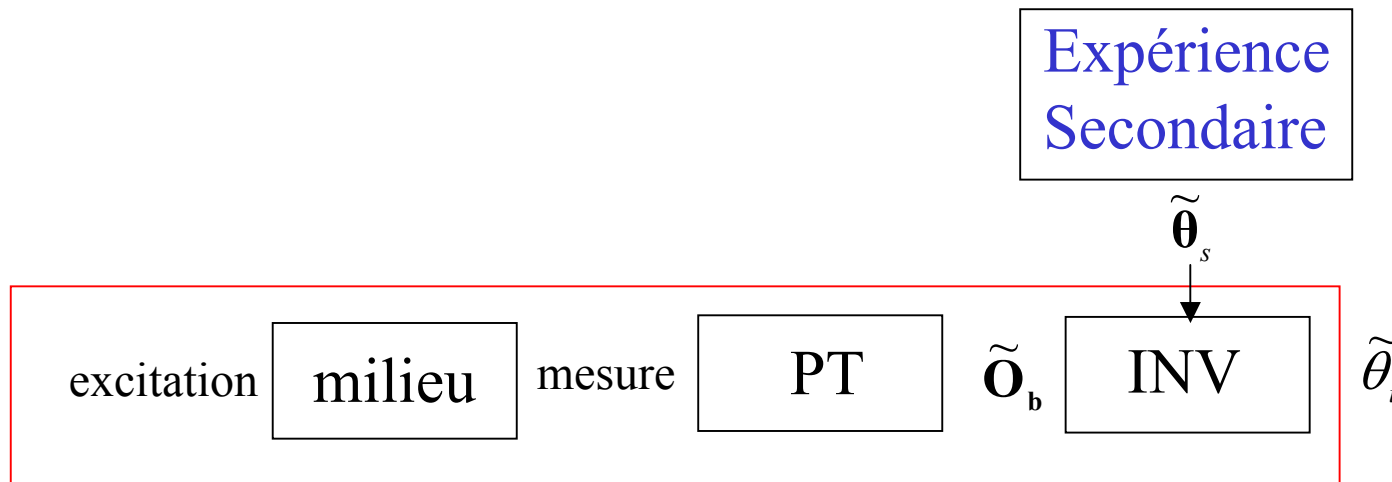
Prévision de performance

Prévision de performances – Prise en compte de la présence de paramètres secondaires mal connus

$$\tilde{\mathbf{O}}_b(v_i) = \mathbf{f}(v_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s) + \tilde{\mathbf{D}}(v_i, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s)$$

$\boldsymbol{\theta}_s$: paramètres secondaires, non commandables, connus avec des erreurs et modélisés comme des variables aléatoires normales centrées sur les vraies valeurs et de matrice de covariance $C(\boldsymbol{\theta}_s)$ connues, indépendant du bruit de mesure

A chaque réalisation d'une expérience fournissant un tirage des observables, il est supposé être effectuée une expérience secondaire fournissant une réalisation des paramètres secondaires.



Expérience
Primaire

Prévision de performance

Prévision de performances – Prise en compte de la présence de paramètres secondaires mal connus

Méthode utilisée pour prendre en compte la présence de paramètres secondaires et quantifier l'effet de leur présence sur les performances d'estimation des paramètres d'intérêt :

évaluer la matrice de covariance de l'estimateur de maximum de vraisemblance des paramètres θ_i sous l'hypothèse d'un modèle linéaire dans les paramètres d'intérêt et dans les paramètres secondaires.

Desolé ... on n'a pas pu faire mieux

Prévision de performance

Prévision de performances – Prise en compte de la présence de paramètres secondaires mal connus

Sous les hypothèses d'un modèle d'observables linéaire dans les paramètres d'intérêt et secondaires, on a :

$$\mathbf{O}_b = [\mathbf{X}_i \ \mathbf{X}_s][\boldsymbol{\theta}_i \ \boldsymbol{\theta}_s]^T + \mathbf{D} = \mathbf{X}_i\boldsymbol{\theta}_i + \mathbf{X}_s\boldsymbol{\theta}_s + \mathbf{D}$$

Soient une expérience primaire donnant une réalisation des observations \mathbf{O}_b^0 et une expérience secondaire fournissant une réalisation des paramètres secondaires $\boldsymbol{\theta}_s^0$, on peut former un nouveau type d'observables corrigé des paramètres secondaires :

$$\tilde{\mathbf{O}}_b^0 = \mathbf{O}_b^0 - \mathbf{X}_s\boldsymbol{\theta}_s^0$$

Possédant un modèle sous-jacent :

$$\tilde{\mathbf{O}}_b = \mathbf{O}_b - \mathbf{X}_s\boldsymbol{\theta}_s = \mathbf{X}_i\boldsymbol{\theta}_i + \mathbf{D}$$

L'estimateur du maximum de vraisemblance de $\boldsymbol{\theta}_i$ aux observables corrigées est défini par :

$$\boldsymbol{\theta}_i^{\text{mv}0} = (\mathbf{X}_i^T \mathbf{X}_i)^{-1} \mathbf{X}_i^T \tilde{\mathbf{O}}_b^0 = (\mathbf{X}_i^T \mathbf{X}_i)^{-1} \mathbf{X}_i^T (\mathbf{O}_b^0 - \mathbf{X}_s\boldsymbol{\theta}_s^0)$$

Le calcul de l'espérance mathématique et de la matrice de covariance de $\boldsymbol{\theta}_i^{\text{mv}}$ fournit les résultats suivants :

$$E(\boldsymbol{\theta}_i^{\text{mv}}) = \boldsymbol{\theta}_i$$

$$\mathbf{C}(\boldsymbol{\theta}_i^{\text{mv}}) = \sigma^2 (\mathbf{X}_i^T \mathbf{X}_i)^{-1} + (\mathbf{X}_i^T \mathbf{X}_i)^{-1} \mathbf{X}_i^T \mathbf{X}_s \mathbf{C}(\boldsymbol{\theta}_s) \mathbf{X}_s^T \mathbf{X}_i (\mathbf{X}_i^T \mathbf{X}_i)^{-1} = [\mathbf{I}(\boldsymbol{\theta}_i)]^{-1} + \mathbf{T}_1(\boldsymbol{\theta}_s)$$

La matrice de covariance de l'estimateur (la performance) est égale l'inverse de la matrice de Fisher obtenue en l'absence de paramètres secondaires augmentée d'un terme semi-défini positif $\mathbf{T}_1(\boldsymbol{\theta}_s)$

Prévision de performance

Prévision de performances – Prise en compte de la présence de paramètres de nuisance

Méthode utilisée pour prendre en compte la présence de paramètres de nuisance et quantifier l'effet de leur présence sur les performances d'estimation des paramètres d'intérêt :

Augmenter le vecteur de paramètres inconnus à estimer $\theta = [\theta_i, \theta_n]$ et appliquer le théorème de Cramer Rao sur ce vecteur de paramètres puis extraire de l'inverse de la matrice d'information de Fisher le sous bloc relatif à θ_i .

Prévision de performance

Prévision de performances – Prise en compte de la présence de paramètres de nuisance

Pour prendre en compte la présence des paramètres de nuisance inconnus,

$$\mathbf{O}_b = \mathbf{f}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_n) + \mathbf{D} \quad D \# N(\vec{0}, \sigma^2 \mathbf{I})$$

Soit $\mathbf{X} = [\mathbf{X}_i \ \mathbf{X}_n]$ la matrice des sensibilités (dérivées) du modèle par rapport aux paramètres,

$$\mathbf{X}_i = \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}_i^T} \quad \mathbf{X}_n = \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}_n^T}$$

L'application du théorème de Cramer Rao sur le vecteur $\theta = [\theta_i, \theta_n]$ permet d'obtenir

$$I^{-1}(\boldsymbol{\theta}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$$

Et si

$$I^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_3 \\ \mathbf{M}_3^T & \mathbf{M}_2 \end{bmatrix}$$

On a :

$$C(\tilde{\theta}_i) \geq \mathbf{M}_1$$

On va calculer \mathbf{M}_1 par l'application de deux lemmes d'inversion.

Prévision de performance

Prévision de performances – Prise en compte de la présence de paramètres de nuisance

Lemme 1

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \implies \mathbf{A}^{-1} = \begin{bmatrix} \left[\mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21} \right]^{-1} & -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \left[\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \right]^{-1} \\ -\mathbf{A}_{22}^{-1} \mathbf{A}_{21} \left[\mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21} \right]^{-1} & \left[\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \right]^{-1} \end{bmatrix}$$

Lemme 2

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{D} \mathbf{A}^{-1} \mathbf{B} + \mathbf{C}^{-1})^{-1} \mathbf{D} \mathbf{A}^{-1}$$

L'utilisation des 2 lemmes d'inversion permet d'obtenir finalement :

$$\text{cov}(\tilde{\boldsymbol{\theta}}_i) \geq \mathbf{M}_1 = \sigma^2 (\mathbf{X}_i^T \mathbf{X}_i)^{-1} + \mathbf{T}_1(\boldsymbol{\theta}_n) + \mathbf{T}_2(\boldsymbol{\theta}_n)$$

$$\mathbf{T}_1(\boldsymbol{\theta}_n) = [\mathbf{X}_i^T \mathbf{X}_i]^{-1} \mathbf{X}_i^T \mathbf{X}_n \sigma^2 (\mathbf{X}_n^T \mathbf{X}_n)^{-1} \mathbf{X}_n^T \mathbf{X}_i [\mathbf{X}_i^T \mathbf{X}_i]^{-1} = [\mathbf{X}_i^T \mathbf{X}_i]^{-1} \mathbf{X}_i^T \mathbf{X}_n \mathbf{C}(\boldsymbol{\theta}_n) \mathbf{X}_n^T \mathbf{X}_i [\mathbf{X}_i^T \mathbf{X}_i]^{-1}$$

$$\begin{aligned} \mathbf{T}_2(\boldsymbol{\theta}_n) &= \sigma^2 (\mathbf{X}_i^T \mathbf{X}_i)^{-1} \mathbf{X}_i^T \mathbf{X}_n (\mathbf{X}_n^T \mathbf{X}_n)^{-1} (\mathbf{X}_n^T \mathbf{X}_i) \\ &\quad \times [-\mathbf{X}_i^T \mathbf{X}_n (\mathbf{X}_n^T \mathbf{X}_n)^{-1} \mathbf{X}_n^T \mathbf{X}_i + \mathbf{X}_i^T \mathbf{X}_i]^{-1} \\ &\quad \times \mathbf{X}_i^T \mathbf{X}_n (\mathbf{X}_n^T \mathbf{X}_n)^{-1} \mathbf{X}_n^T \mathbf{X}_i (\mathbf{X}_i^T \mathbf{X}_i)^{-1} \end{aligned}$$

Une borne inférieure de la matrice de covariance des paramètres d'intérêt est égale l'inverse de la matrice de Fisher obtenue en l'absence de paramètres de nuisance augmentée de deux termes semi-définis positifs $\mathbf{T}_1(\boldsymbol{\theta}_n)$ et $\mathbf{T}_2(\boldsymbol{\theta}_n)$

Prévision de performance

Prévision de performances – 1^{ier} cas d'école : estimation en présence de paramètres de nuisance ou de paramètres secondaires

$$O_b(i) = a_0 + a_1 i + b(i); i \in \{1, \dots, N\} \quad a_0: \text{paramètre d'intérêt, } a_1 \text{ paramètre ?}$$

Si a_1 est parfaitement connu, on construit les nouvelles observables :

$$\tilde{O}_b(i) = O_b(i) - a_1 i = a_0 + b(i); i \in \{1, \dots, N\}$$

Et on applique le théorème de Cramer Rao dans ce cas pour obtenir :

$$C(a_0) = \frac{\sigma^2}{N}$$

Si a_1 est un paramètre secondaire # $N(a_1^v, \sigma_{a_1}^2)$

$$\begin{aligned} C(a_0^{mv}) &= \sigma^2 (\mathbf{X}_{a_0}^T \mathbf{X}_{a_0})^{-1} + (\mathbf{X}_{a_0}^T \mathbf{X}_{a_0})^{-1} \mathbf{X}_{a_0}^T \mathbf{X}_{a_1} (\sigma_{a_1}^2) \mathbf{X}_{a_1}^T \mathbf{X}_{a_0} (\mathbf{X}_{a_0}^T \mathbf{X}_{a_0})^{-1} = [I(a_0)]^{-1} + \mathbf{T}_1(a_1) \\ &= \frac{\sigma^2}{N} + \frac{1}{N} \times \sum_{i=1}^N i \times \sigma_{a_1}^2 \times \sum_{i=1}^N i \times \frac{1}{N} \\ &= \frac{\sigma^2}{N} + \frac{(N+1)^2 \sigma_{a_1}^2}{4} \end{aligned}$$

Si a_1 est un paramètre secondaire # $N(a_1^v, \sigma_{a_1}^2)$ et que l'expérience secondaire est la mesure de a_1 par régression linéaire sur le même nombre de points que l'expérience principale et le même niveau de bruit, on a :

$$\sigma_{a_1}^2 = \frac{6\sigma^2}{N^2(N+1)(2N+1)} \quad C(a_0^{mv}) = \frac{4N^2 + 5N + 3}{4N^3 + 2N^2} \sigma^2 \underset{N \rightarrow \infty}{\approx} \frac{\sigma^2}{N}$$

Prévision de performance

Prévision de performances – 1^{ier} cas d'école : estimation en présence de paramètres de nuisance ou de paramètres secondaires

$$O_b(i) = a_0 + a_1 i + b(i); i \in \{1, \dots, N\}$$

Si a_1 est un paramètre de nuisance inconnu, l'application du théorème de Cramer Rao conduit à :

$$I \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \frac{1}{\sigma^2} \begin{bmatrix} N & \sum_{i=1}^N i \\ \sum_{i=1}^N i & \sum_{i=1}^N i^2 \end{bmatrix} = \frac{1}{\sigma^2} \begin{bmatrix} N & \frac{N(N+1)}{2} \\ \frac{N(N+1)}{2} & \frac{N^2(N+1)(2N+1)}{6} \end{bmatrix}$$

$$I^{-1} = \frac{{}^t \text{COM}(I)}{\det(I)} = \sigma^2 \begin{bmatrix} \frac{2(2N+1)}{N(N-1)} & \frac{-6}{N(N-1)} \\ \frac{-6}{N(N-1)} & \frac{12}{N(N-1)(N+1)} \end{bmatrix}$$

$$\sigma_{a_0}^2 \geq \frac{2(2N+1)}{N(N-1)} \underset{N \rightarrow \infty}{\approx} \frac{4\sigma^2}{N}$$

Prévision de performance

Prévision de performances – 1^{ier} cas d'école : estimation en présence de paramètres de nuisance ou de paramètres secondaires

$$O_b(i) = a_0 + a_1 i + b(i); i \in \{1, \dots, N\}$$

a₁ connu $C(a_0) = \frac{\sigma^2}{N}$

a₁ paramètre secondaire $C(a_0^{\text{mv}}) = \frac{\sigma^2}{N} + \frac{(N+1)^2 \sigma_{a_1}^2}{4}$ $C(a_0^{\text{mv}}) = \frac{4N^2 + 5N + 3}{4N^3 + 2N^2} \sigma^2 \underset{N \rightarrow \infty}{\approx} \frac{\sigma^2}{N}$

a₁ paramètre de nuisance $\sigma_{a_0}^2 \geq \frac{2(2N+1)}{N(N-1)} \underset{N \rightarrow \infty}{\approx} \frac{4\sigma^2}{N}$

Prévision de performance

Prévision de performances – 2^{ième} cas d'école : estimation de deux paramètres simultanément en présence de corrélation

$$O_b(i) = a_0 + a_1 + b(i); i \in \{1, \dots, L\}$$

$$O_b(i) = a_0 - a_1 + b(i); i \in \{L+1, \dots, N\}$$

$$\frac{\partial O_b(i)}{\partial a_0} = 1; i \in \{1, \dots, N\}$$

$$\sum_{i=1}^N \left[\frac{\partial O_b(i)}{\partial a_0} \right]^2 = N$$

$$I \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \frac{1}{\sigma^2} \begin{bmatrix} N & 2L-N \\ 2L-N & N \end{bmatrix}$$

$$\frac{\partial O_b(i)}{\partial a_1} = 1; i \in \{1, \dots, L\}$$

$$\sum_{i=1}^N \left[\frac{\partial O_b(i)}{\partial a_1} \right]^2 = N$$

$$\frac{\partial O_b(i)}{\partial a_1} = -1; i \in \{L+1, \dots, N\}$$

$$\sum_{i=1}^N \left[\frac{\partial O_b(i)}{\partial a_1} \right] \left[\frac{\partial O_b(i)}{\partial a_0} \right] = 2L - N$$

$$I^{-1} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \sigma^2 \begin{bmatrix} \frac{N}{-4L^2 + 4LN} & -\frac{2L-N}{-4L^2 + 4LN} \\ -\frac{2L-N}{-4L^2 + 4LN} & \frac{N}{-4L^2 + 4LN} \end{bmatrix}$$

$$\frac{N}{-4L^2 + 4LN}$$

minimale pour $L=N/2$

$$I^{-1} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{bmatrix} \frac{\sigma^2}{N} & 0 \\ 0 & \frac{\sigma^2}{N} \end{bmatrix}$$

Prévision de performance

Prévision de performances – Conclusion

Nous possédons **une expression de la matrice de covariance** (ou une borne inférieure) **d'un estimateur non biaisé de θ_i** avec **la possibilité de prendre en compte la présence de paramètres de nuisance et de paramètres secondaires.**

Cette matrice de covariance est **une quantification des performances statistiques accessibles pour l'estimation de θ_i**

Cette matrice de covariance est **une fonction du vecteur de variables indépendantes et du vecteur des réglages expérimentaux**, la question du choix des valeurs de ces paramètres commandables afin d'optimiser les performances constitue **la conception optimales des expériences.**

Plan de la présentation

I] Introduction

II] Prévisions de Performances

III] Conception Optimale des Expériences

IV] Exemples d'applications passées et à venir

Conception optimale des expériences

Introduction

La phase de prévision de performances a fourni une expression de la matrice de covariance d'un estimateur non biaisé de θ_i .

$$C(\hat{\boldsymbol{\theta}}_i) = G(v_i, \boldsymbol{\theta}_e, \boldsymbol{\theta}_i, \boldsymbol{\theta}_s, \boldsymbol{\theta}_n)$$

La conception optimale des expériences consiste à choisir les paramètres commandables (v_i et θ_e) qui optimisent $C(\hat{\boldsymbol{\theta}}_i)$

Procédure de conception optimale des expériences en 2 pas

pas 1 : choix d'un critère numérique qui résume les caractéristiques d'une matrice de covariance $\text{Qual}(C)$

pas 2 : sélection des paramètres réglables (v_i et θ_e) qui optimisent $\text{Qual}(C)$

Conception optimale des expériences

Définition des critères de qualité

- les critères sont donnés pour une matrice de covariance de taille quelconque
- leur sens est expliqué
- ils sont finalement appliqués à une matrice de covariance 2 par 2

$$C(\boldsymbol{\theta}_i) = \sigma^2 \begin{bmatrix} \frac{1}{a(1-c_{\text{lim}}^2)} & -\frac{c_{\text{lim}}}{\sqrt{ab}(1-c_{\text{lim}}^2)} \\ -\frac{c_{\text{lim}}}{\sqrt{ab}(1-c_{\text{lim}}^2)} & \frac{1}{b(1-c_{\text{lim}}^2)} \end{bmatrix}$$

Conception optimale des expériences

Critères de qualité : 1 Terme diagonal, la trace

Critère 1 : 1 terme diagonal particulier de la matrice de covariance

$$Qual_1(C) = C(l,l) = \sigma^2(\theta_i(l))$$

Pour la matrice 2 par 2

$$Qual_1(C) = C(1,1) = \frac{1}{a(1 - c_{\text{lim}}^2)}$$

Critère 2 : la trace de la matrice de covariance

$$Qual_2(C) = Tr(C) = \sum_{i=1}^P C(i,i)$$

Conception optimale des expériences

Critères de qualité : Le nombre de conditionnement, l'entropie

Critère 3 : le nombre de conditionnement de la matrice de covariance

$$Qual_3(C) = \frac{\lambda_{\max}}{\lambda_{\min}} \quad \lambda_{\max} \text{ et } \lambda_{\min} \text{ valeurs singulières maximales et minimales de } C$$

Pour la matrice 2 par 2 $cond(\mathbf{J}) = \frac{\sqrt{\frac{a^2 + b^2}{ab} + 2c_{\text{lim}}^2} + \sqrt{\left(\frac{a^2 + b^2}{ab} + 2c_{\text{lim}}^2\right)^2 - 4(1 - c_{\text{lim}}^2)^2}}{\sqrt{\frac{a^2 + b^2}{ab} + 2c_{\text{lim}}^2} - \sqrt{\left(\frac{a^2 + b^2}{ab} + 2c_{\text{lim}}^2\right)^2 - 4(1 - c_{\text{lim}}^2)^2}}$

Critère 4 : l'entropie

$$Qual_4(C) = \log(\det(C))$$

Pour la matrice 2 par 2

$$Qual_4(C) = 2 \log(\sigma^2) - \log(1 - c_{\text{lim}}^2) - \log(ab)$$

Conception optimale des expériences

Critères de qualité : 1 coefficient de corrélation particulier, le terme d'information de Kullback

Critère 5 : 1 coefficient de corrélation particulier

$$Qual_5(C) = \left| \frac{C(i, j)}{\sqrt{C(i, i)C(j, j)}} \right|$$

Critère 6 : le terme d'information de Kullback

$$Qual_6(C) = \frac{1}{2} \log \left(\frac{\prod_{i=1}^P C(i, i)}{\det(C)} \right)$$

Pour la matrice 2 par 2

$$Qual_6(C) = -\frac{1}{2} \log(1 - c_{\text{lim}}^2)$$

Conception optimale des expériences

Conclusion sur les critères de qualité

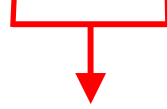
-4 critères (variance particulière, trace, entropie, nombre de conditionnement) **résumé** les caractéristiques d'une matrice de covariance **en prenant en compte les problèmes de quantité d'information et de corrélations entre paramètres**

-2 critères (Coefficient de corrélation particulier, terme d'information de Kullback) **résumé partiellement** les caractéristiques d'une matrice de covariance **en prenant uniquement en compte les problèmes de corrélations entre paramètres**

Conception optimale des expériences

Optimisation des critères de qualité

$$Qual [C(\hat{\theta}_i)] = G(v_i, \theta_e, \theta_i, \theta_s, \theta_n)$$



Commandables par l'expérimentateur

Les paramètres commandables sont supposés être des vecteurs constitués d'un nombre fini N_e ou N_{v_i} de termes et chaque terme peut prendre N valeurs parmi une grille pré-déterminée.

Ainsi, si $E_N = \{1, \dots, N\}$

$$v_i \in (E_N)^{N_{v_i}}$$

$$\theta_e \in (E_N)^{N_{\theta_e}}$$

Conception optimale des expériences

Optimisation des critères de qualité – procédure d'optimisation itérative à 1 pas

- 1] Une itération 'n' a donné un vecteur de variables indépendantes v_i^n
- 2] On construit tous les vecteurs v_i qui diffèrent de v_i^n par un seul terme
- 3] On choisit parmi tous les vecteurs construits celui qui minimise le critère de qualité
- 4] On recommence une nouvelle itération

$$v_i^n = [v_i^n(1) \quad \dots \quad v_i^n(k) \quad \dots \quad v_i^n(N_{vi})]$$

$$\begin{bmatrix} 1 & \dots & v_i^n(k) & \dots & v_i^n(N_{vi}) \\ \dots & \dots & v_i^n(k) & \dots & v_i^n(N_{vi}) \\ N & \dots & v_i^n(k) & \dots & v_i^n(N_{vi}) \end{bmatrix}$$

$$\begin{bmatrix} v_i^n(1) & \dots & 1 & \dots & v_i^n(N_{vi}) \\ v_i^n(1) & \dots & \dots & \dots & v_i^n(N_{vi}) \\ v_i^n(1) & \dots & N & \dots & v_i^n(N_{vi}) \end{bmatrix}$$

Conception optimale des expériences

Optimisation des critères de qualité – procédure d'optimisation itérative à 1 pas

Soit un vecteur de paramètres commandables initial $\theta_c^0 = (\theta_e^0, v_i^0)$ élément de $[1, N]^{(N_{vi} + N_e)}$, soit N_i le nombre d'itérations souhaitées

for $k = 1:N_i$

1] Génération de l'ensemble des jeux de paramètres commandables

$$\mathcal{E}_k = \left\{ \begin{array}{l} v \text{ tel que } v = \theta_c^{k-1} + (m - \theta_c^{k-1}(l)) \mathbf{i}_l, m \in \{1, \dots, N\}, l \in \{1, \dots, N_e + N_{vi}\}, \\ \mathbf{i}_l \in R^{N_e + N_{vi}}, \mathbf{i}_l = [0 \dots 0 \quad 1 \quad 0 \dots 0] \\ \text{indiciel} \end{array} \right\}$$

2] recherche de l'élément de \mathcal{E}_k qui optimise le critère de qualité

$$\theta_c^k = \arg \min_{v \in \mathcal{E}_k} (\text{Qual}(C))$$

end

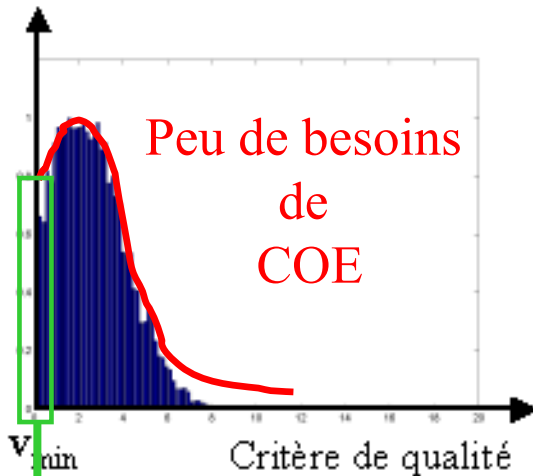
La procédure d'optimisation itérative à un pas n'assure pas la convergence vers un minimum local, en ce sens elle est sous optimale mais elle est cool qu'en même

Conception optimale des expériences

Optimisation des critères de qualité – Analyse de l'exploration exhaustive des jeux de paramètres commandables

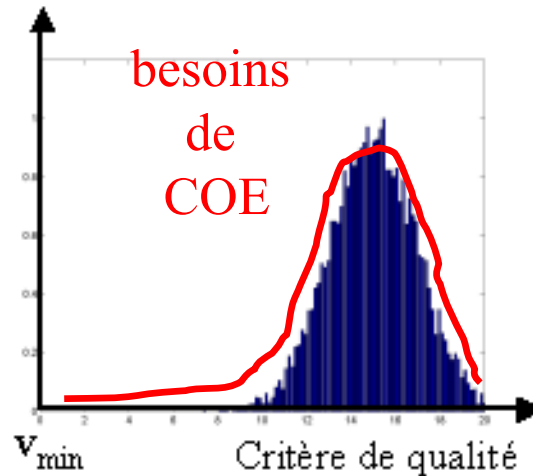
- 1] Création de tous les jeux de paramètres commandables possibles
- 2] Evaluation des critères de qualité de la matrice de covariance pour tous les jeux de paramètres
- 3] Présentation des résultats sous formes d'histogrammes (on compte le nombre de jeux permettant de fournir un critère de qualité dans un intervalle donné) et analyse

Nombre de jeux de réglage expérimentaux



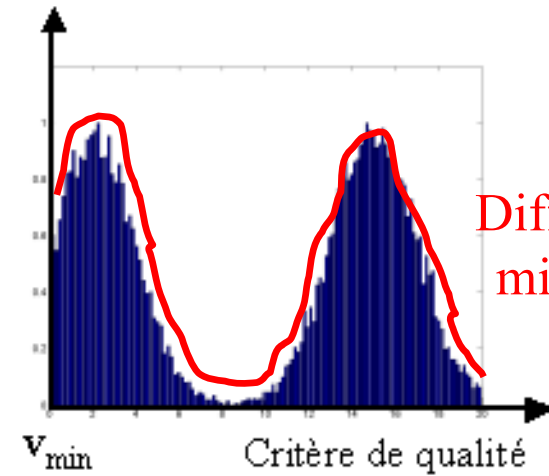
Forme 1

Nombre de jeux de réglage expérimentaux



Forme 2

Nombre de jeux de réglage expérimentaux



Forme 3

Sélection du jeu 'optimal'

Conception optimale des expériences

Cas d'école n°1 : conception optimale en présence de paramètres de nuisance ou paramètres secondaires

Choix de v_i

$$O_b(i) = a_0 + a_1 \times v_i(i) + b(i); i \in \{1, \dots, 20\}; v_i(i) \in \{0, \dots, 19\}$$

a_0 : paramètre d'intérêt

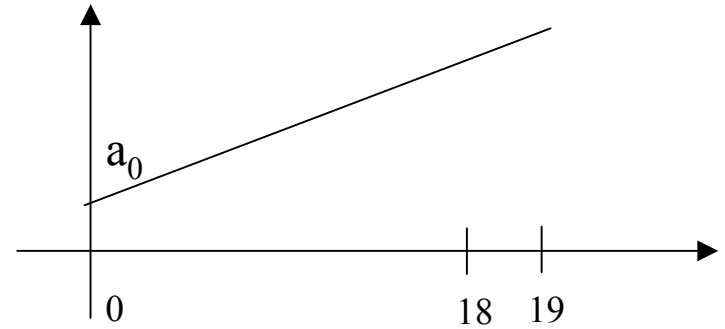
a_1 : paramètre ?

$v_i = \{0, 1, 2, \dots, 19\}$

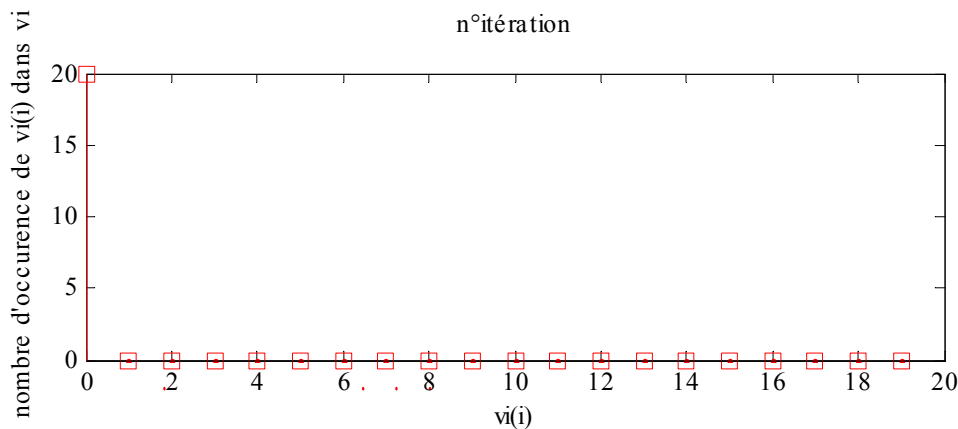
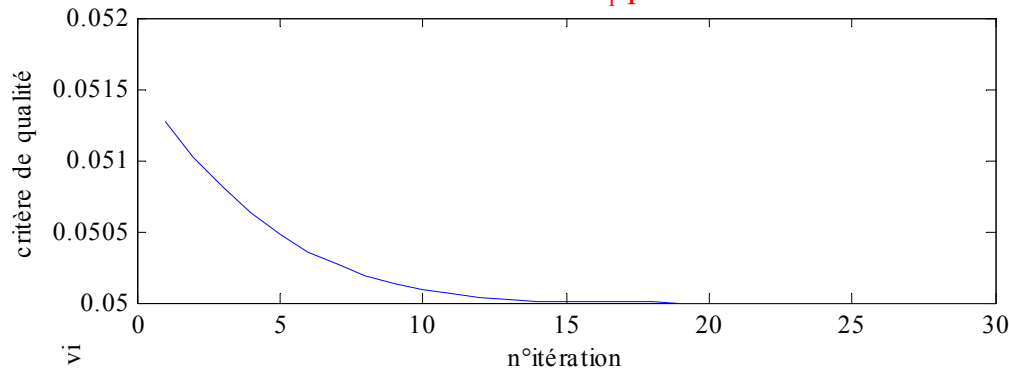
Critère de qualité : trace

Algorithme d'optimisation itérative à 1 pas

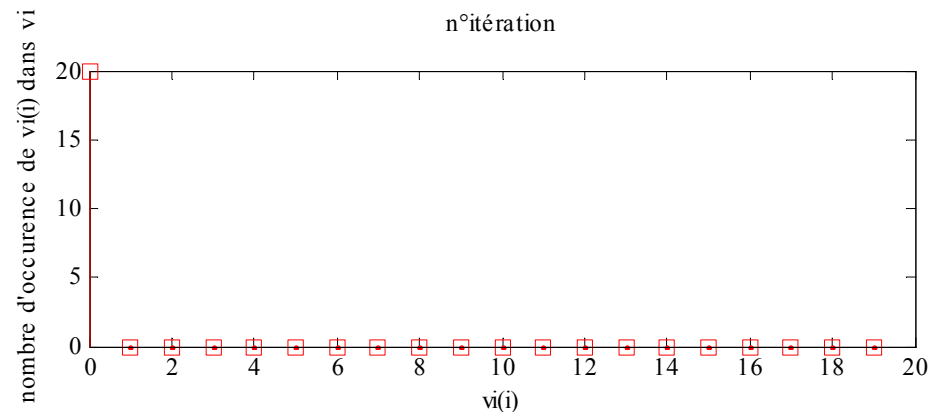
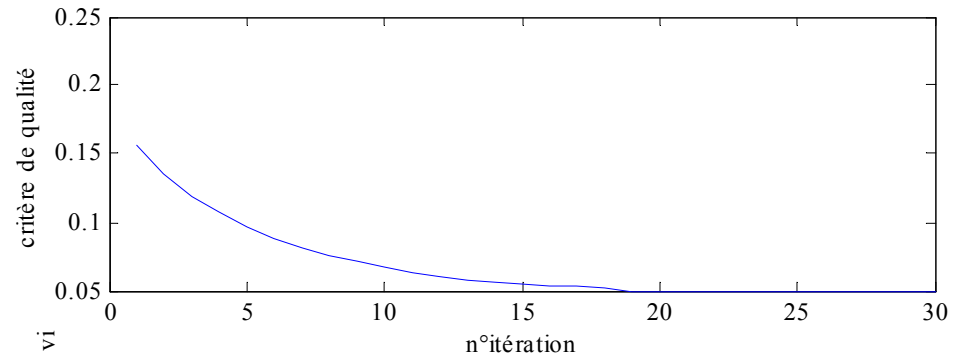
30 itérations



a_1 paramètre secondaire



a_1 paramètre de nuisance



Conception optimale des expériences

Cas d'école n°1 : conception optimale en présence de paramètres de nuisance ou paramètres secondaires

Choix de v_i

a_0 : paramètre d'intérêt

a_1 : paramètre ?

$v_i \in \{1, 2, \dots, 20\}$

Critère de qualité : trace

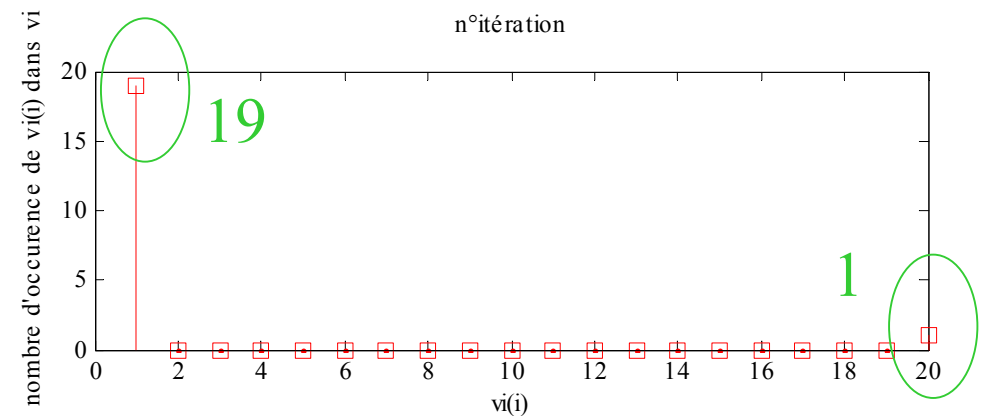
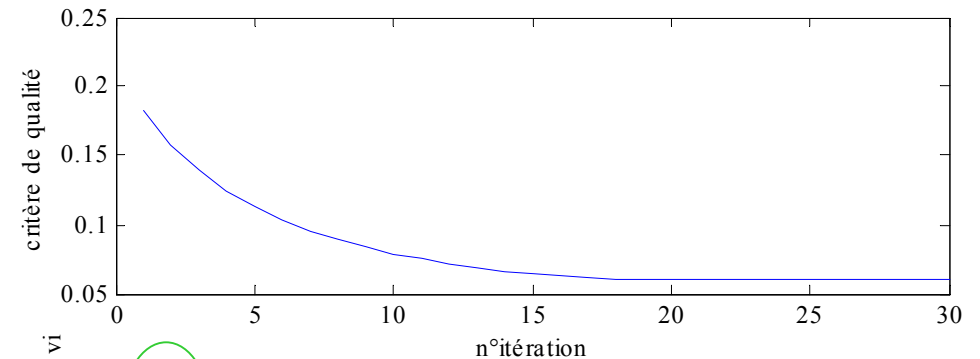
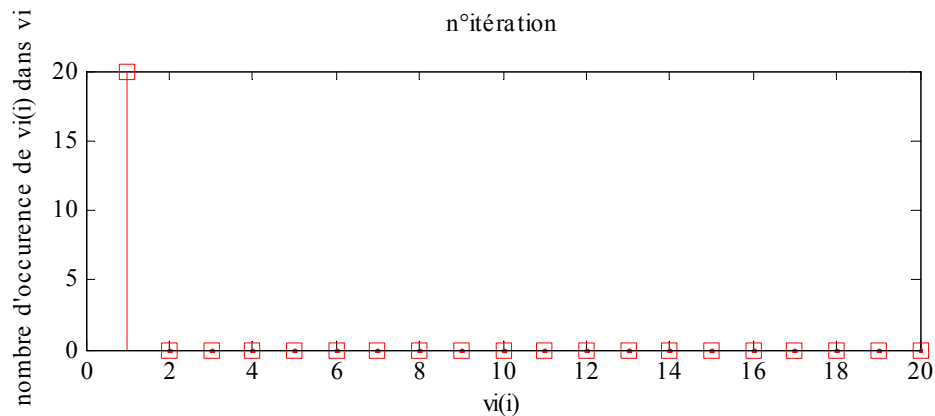
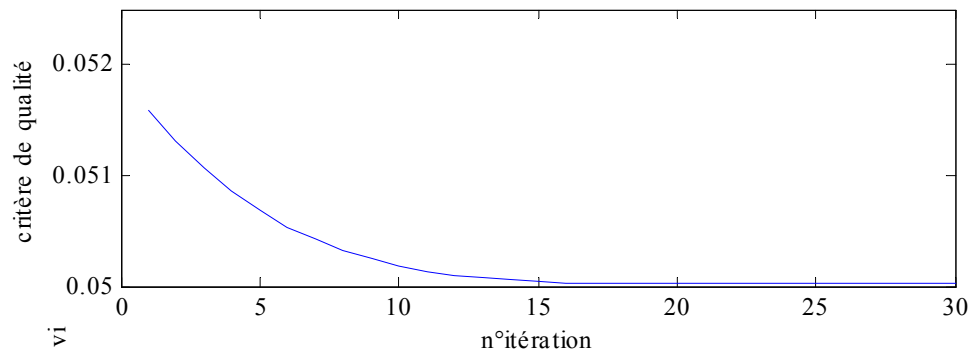
Algorithme d'optimisation itérative à 1 pas

30 itérations

$$O_b(i) = a_0 + a_1 \times v_i(i) + b(i); i \in \{1, \dots, 20\}; v_i(i) \in \{1, \dots, 20\}$$

a_1 paramètre secondaire

a_1 paramètre de nuisance



Conception optimale des expériences

Cas d'école n°2 : conception optimale en présence de corrélation entre 2 paramètres à estimer
Choix de v_i

a_0 : paramètre d'intérêt

a_1 : paramètre d'intérêt

?

$v_{i_0} = \{1, 2, \dots, 20\} \cup \{2, 2, \dots, 21\}$

$N=40$;

$L=20$;

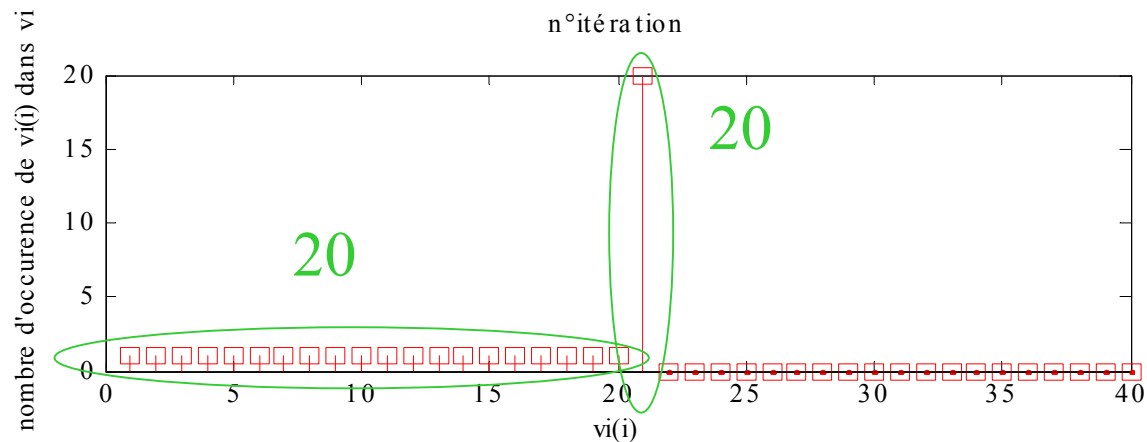
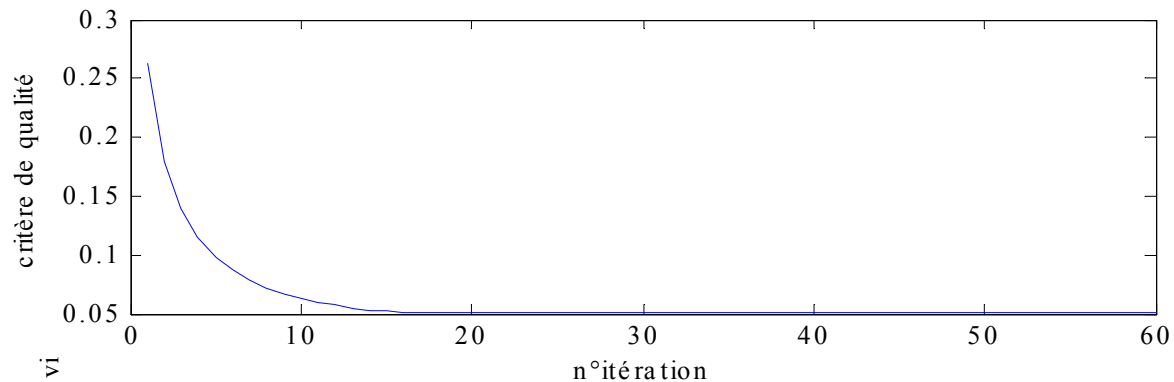
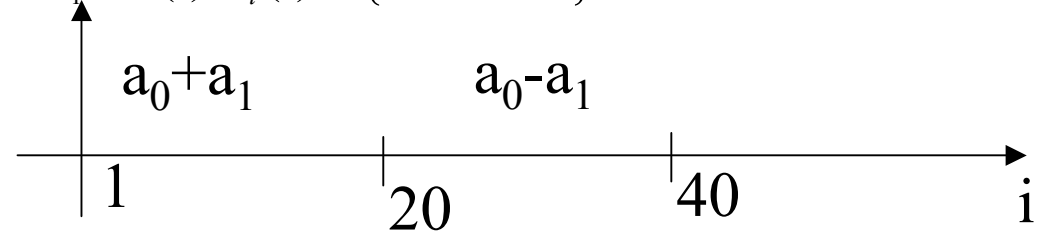
Critère de qualité : trace

Algorithme d'optimisation itérative à 1 pas

60 itérations

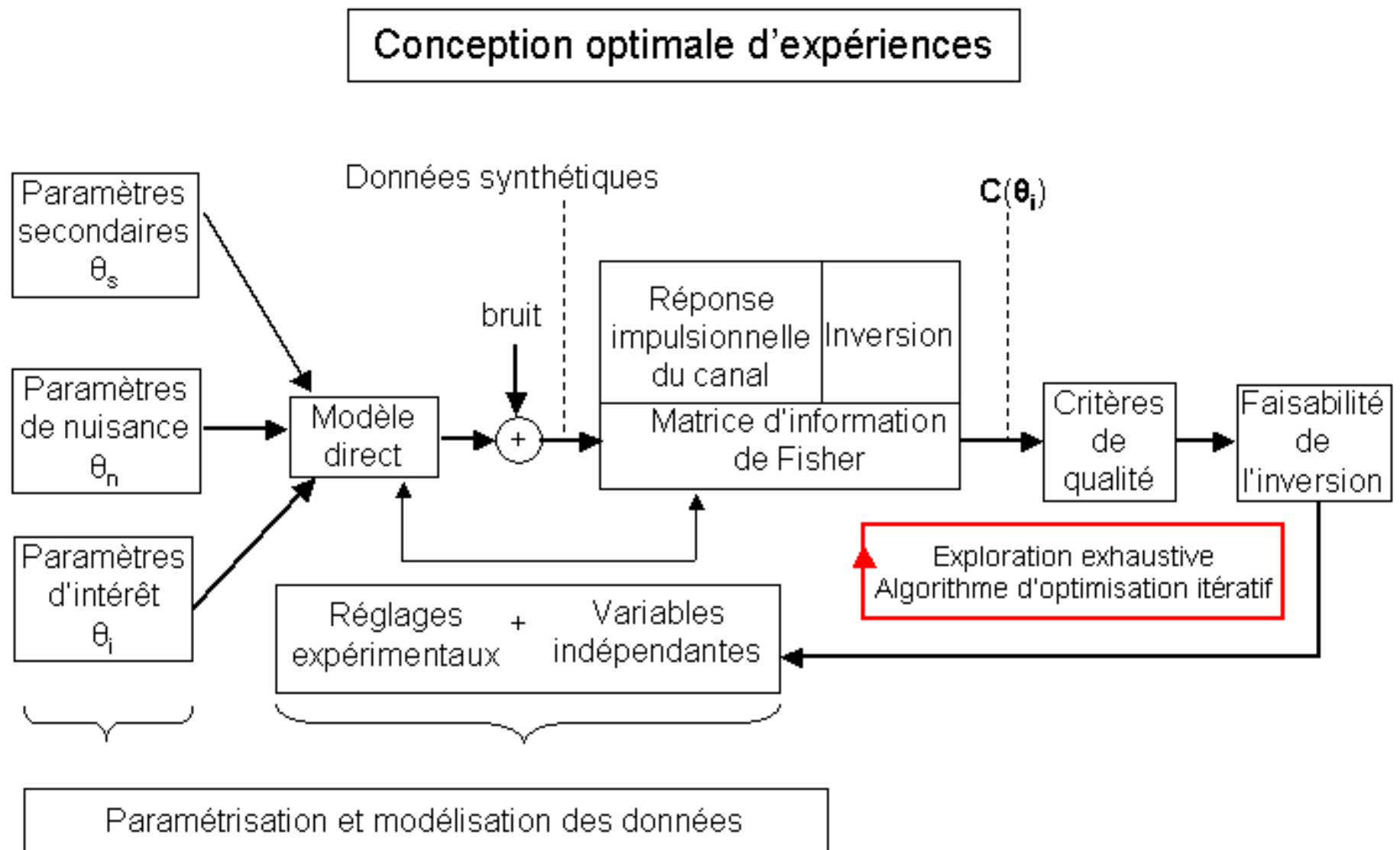
$$O_b(i) = a_0 + a_1 + b(i); v_i(i) \in \{1, \dots, L\}$$

$$O_b(i) = a_0 - a_1 + b(i); v_i(i) \in \{L+1, \dots, N\}$$



Conception optimale des expériences

Conclusion : bientôt ouf



Plan de la présentation

I] Introduction

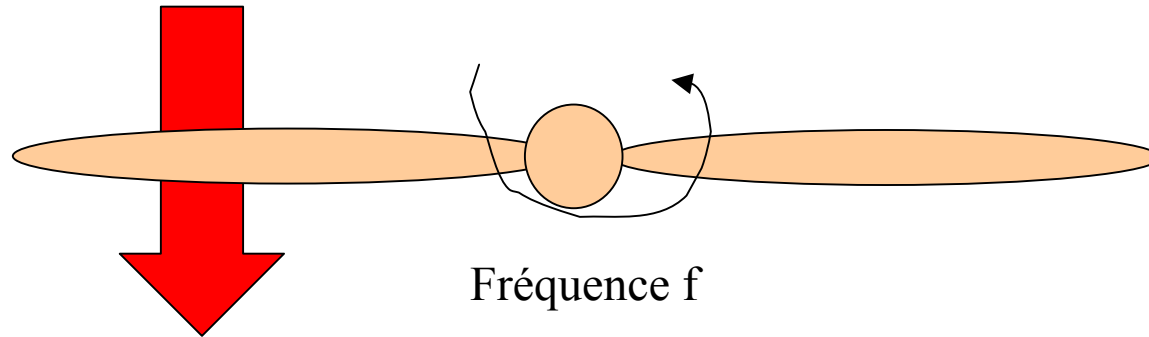
II] Prévisions de Performances

III] Conception Optimale des Expériences

IV] Exemples d'applications passées et à venir

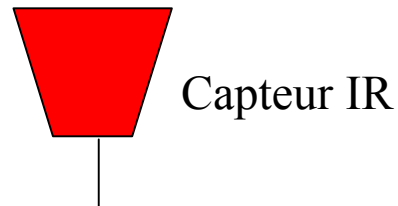
Exemples d'applications passées et à venir

Métrologie des propriétés thermiques de collages spatiaux



métal	Conductivité, diffusivité
colle	Conductivité, diffusivité
métal	Conductivité, diffusivité

Résistance thermique de l'interface

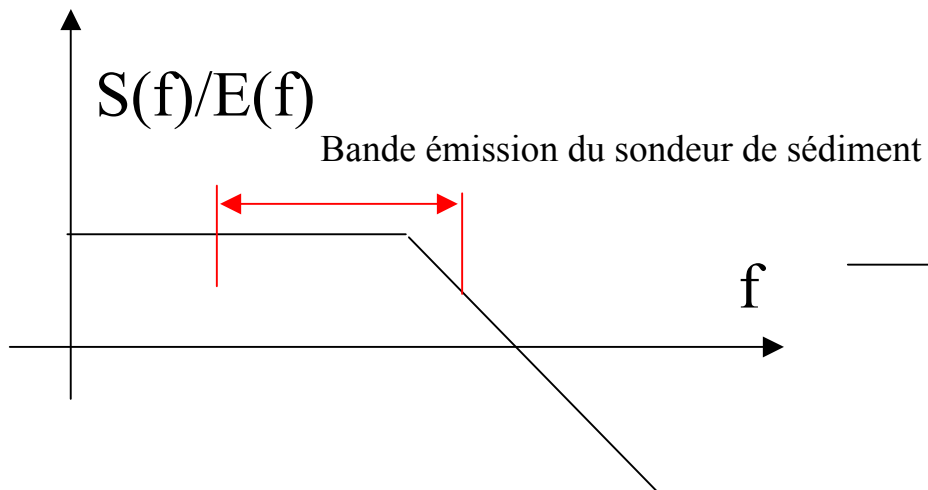
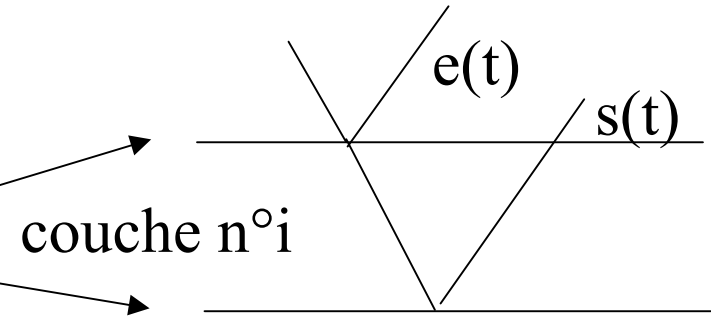
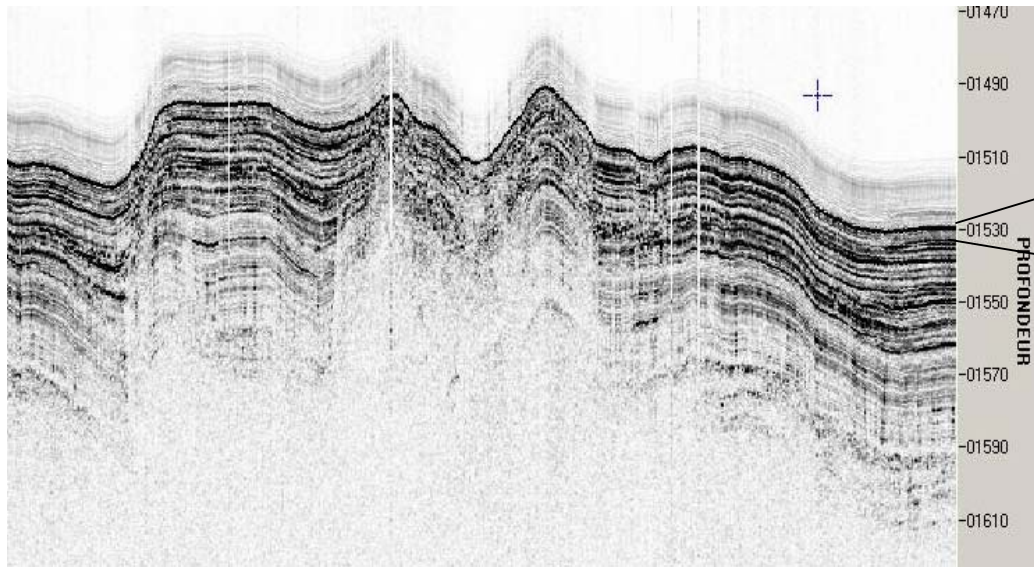


Prédiction de performance => détection et suivi de l'usure du collage par mesure des propriétés thermiques de l'interface

Conception optimale d'expériences => choix de fréquences découplant les paramètres 'métal' et 'collage'

Exemple d'applications passées et à venir

Mesure du coefficient d'atténuation des couches sédimentaires à partir d'un sondeur de sédiments



Prévision de performances

Exemple d'applications passées et à venir

Identification des paramètres de la première couche d'un fond marin par inversion de l'index de rétro-diffusion

Quantifier l'influence de la fréquence sur les performances (3 gammes disponibles 12 KHz, 80 KHz, 450 KHz)

T_2 Choix de l'excitation

T_5 Étage d'inversion

- Critère d'accord T_{5-1}

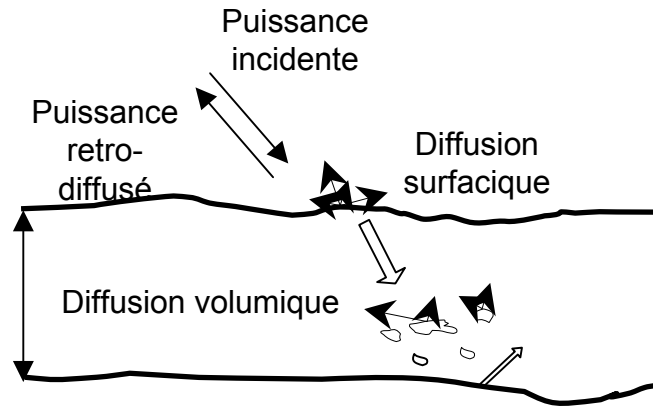
- Optimisation du critère T_{5-2}

T_3

Émission T_{3-1}
Propagation T_{3-2}
Mesures T_{3-3}

T_4

Étage de pré-traitement

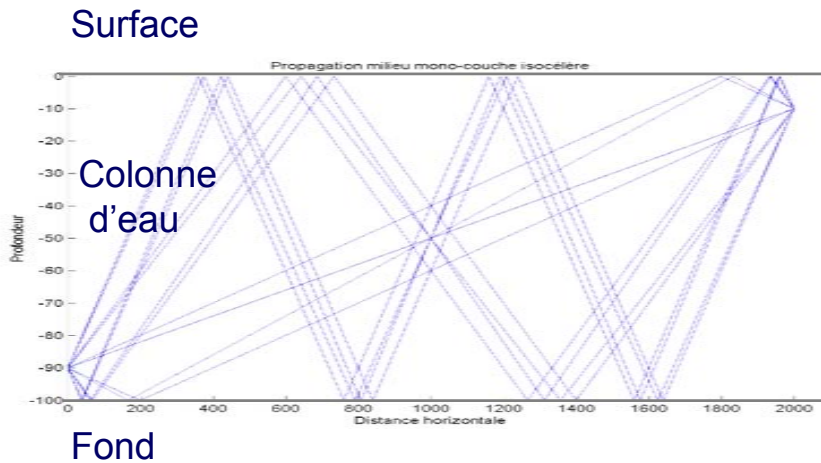


6 paramètres d'intérêt

Quantifier les conséquences de la non-maîtrise de la formation de voies + diagrammes émission, réception
Sondeur multi-faisceaux

Exemple d'applications passées et à venir

Tomographie acoustique océanique passive



Actif idéal

Signal connu, positions connues

Actif Réelle

Signal presque connu, positions presque connues

Actif discret

Signal presque connu, positions presque connues, faible RSB

Passif assisté

Signal inconnu, positions presque connues, faible RSB

Passif autonome

Signal inconnu, positions inconnues, faible RSB