



Reconstruction & Analyse de physique

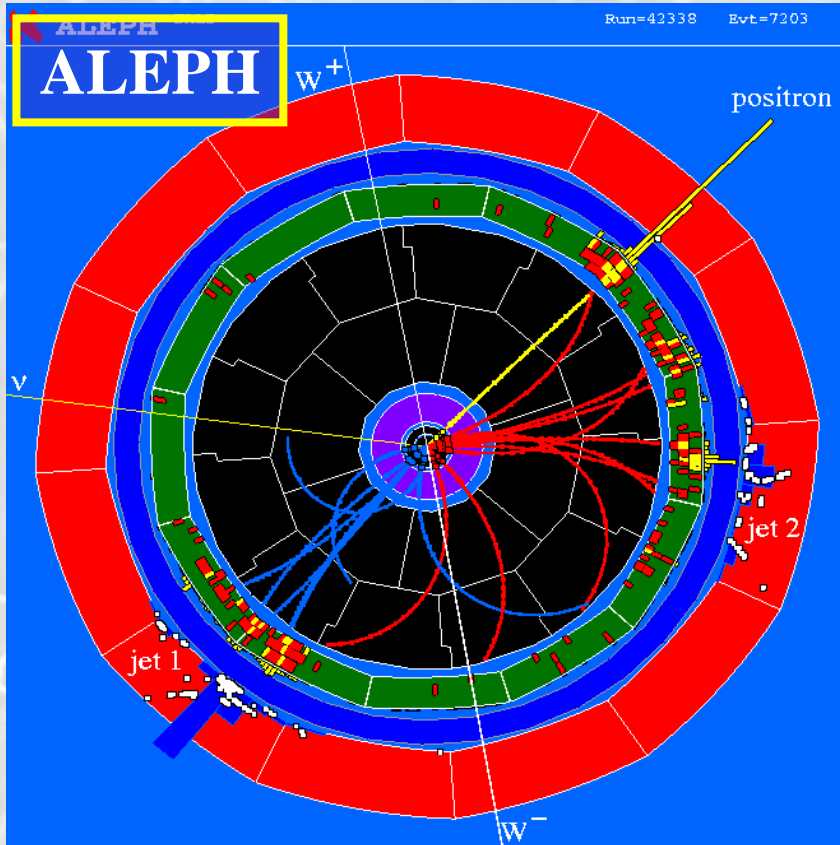
Marc Verderi
LLR-École polytechnique
Du détecteur à la mesure
Roscoff – Juin 2007

INTRODUCTION

Liminaire

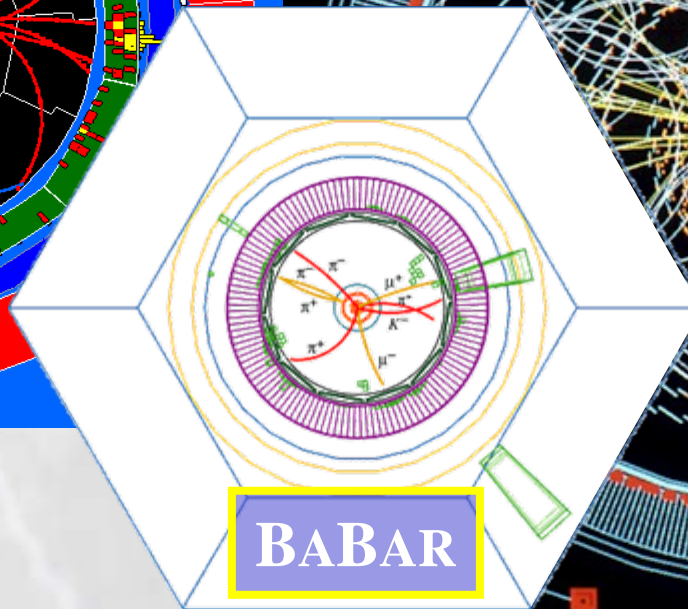
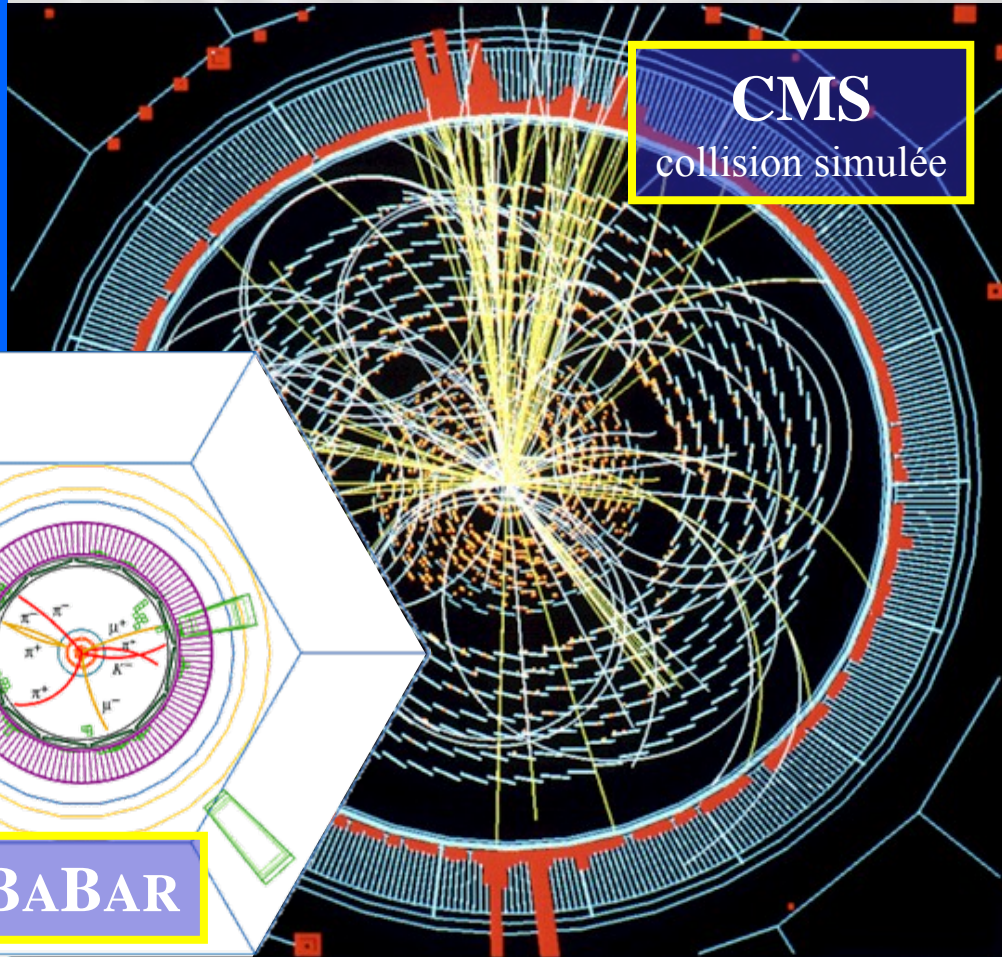
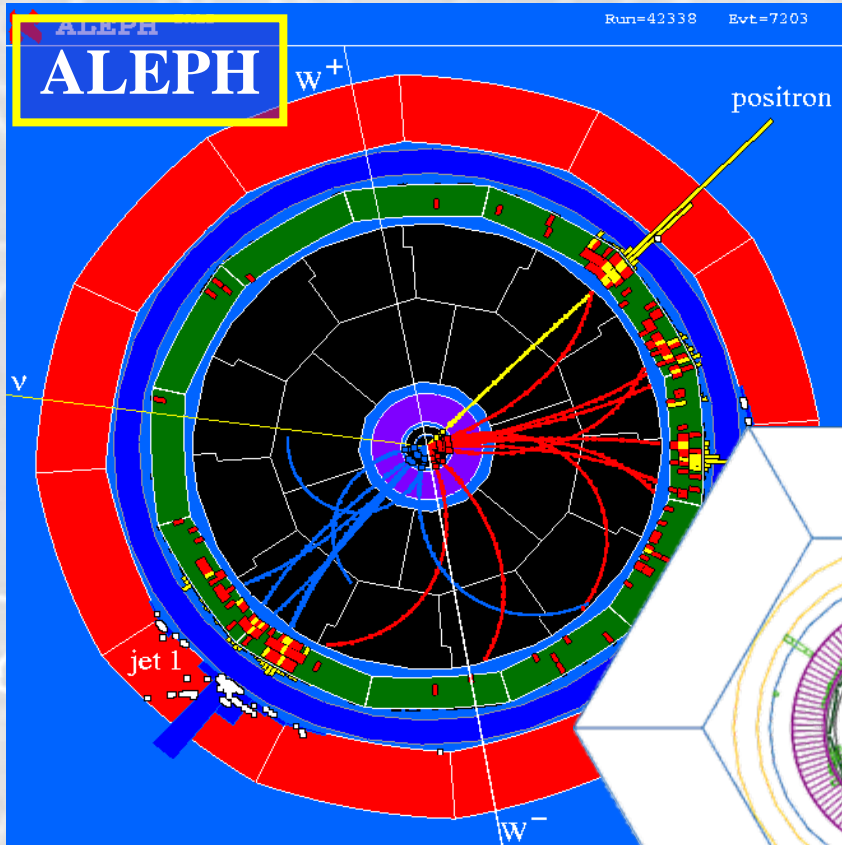
- On suppose un ensemble accélérateur-détecteur(s) en opération et conçu pour répondre à une thématique de physique:
 - Exemples:
 - LEP – {Aleph, Delphi, Opal, L3} = physique du Z^0 , W^\pm
 - PEP-II – BABAR / KEK – BELLE = violation de CP,
 - LHC – {CMS, ATLAS} = recherche du Higgs
- L'accélérateur crée, par collisions, des particules d'intérêt:
 - Exemples:
 - LEP : $e^+e^- \rightarrow Z^0$, $e^+e^- \rightarrow W^+ W^-$
 - PEP-II : $e^+e^- \rightarrow Y(4S) \rightarrow B^0 \text{ anti-}B^0$
 - LHC: gluon gluon \rightarrow Higgs
- Le détecteur permet l'étude de ces particules en collectant et mesurant leur produits de désintégrations:
 - Exemples:
 - $Z^0 \rightarrow$ quark anti-quark, e^+e^- , $\mu^+\mu^-$, $\tau^+\tau^-$
 - $W^\pm \rightarrow$ quark anti-quark', e^\pm neutrino, ...
 - $B^0 \rightarrow J/\Psi K_S^0$ suivi de $J/\Psi \rightarrow e^+e^-$ ou $\mu^+\mu^-$ et $K_S^0 \rightarrow \pi^+\pi^-$, ...
 - Higgs $\rightarrow \gamma\gamma$, ...

Exemples de collisions...



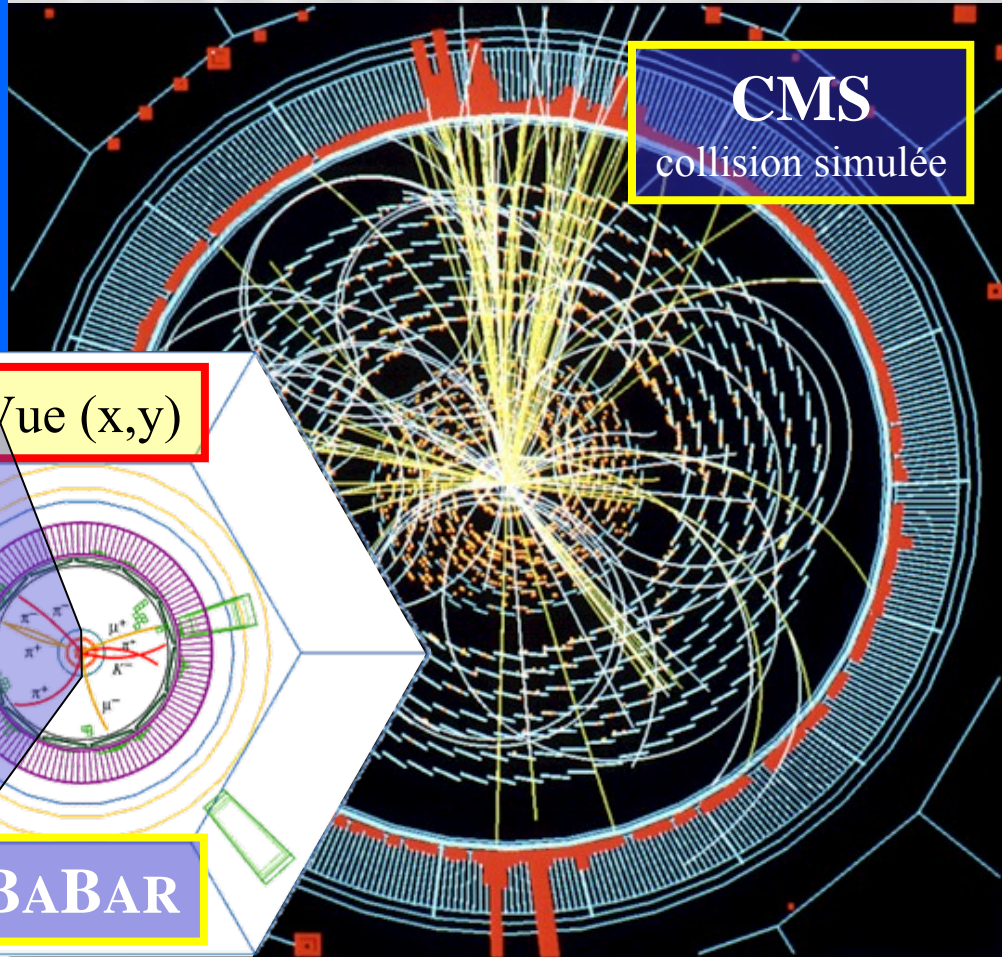
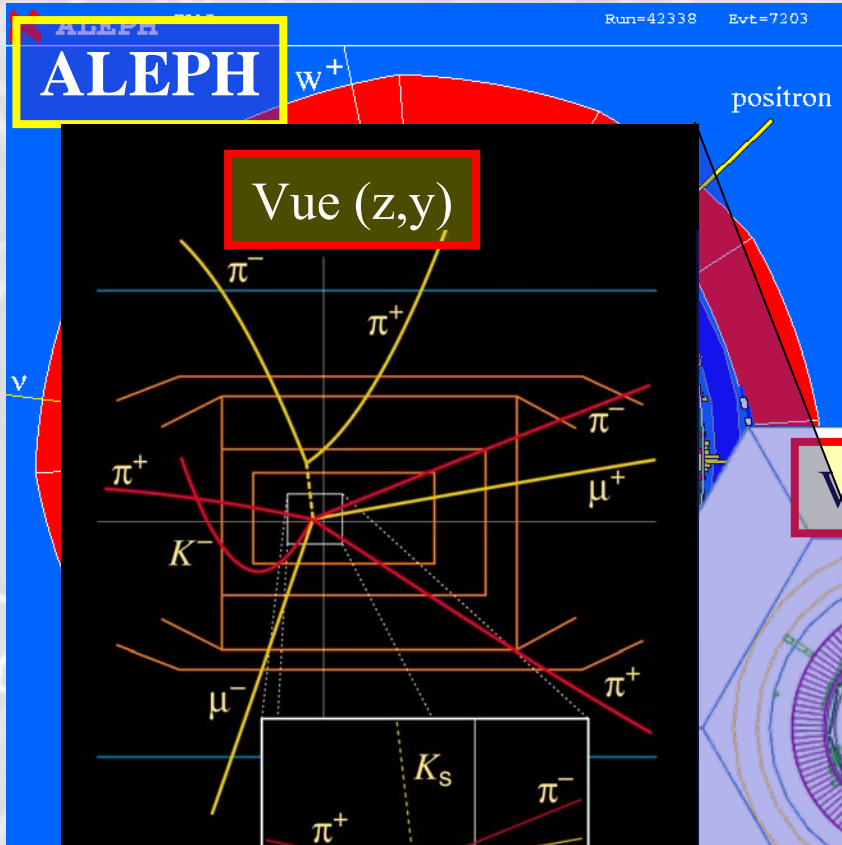
$e^+e^- \rightarrow W^+ W^-$ suivi de $W^- \rightarrow \text{quark anti-quark}'$, $W^+ \rightarrow e^+ \text{ neutrino}$

Exemples de collisions...



$B^0 \rightarrow J/\Psi K_S^0$ suivi de $J/\Psi \rightarrow \mu^+\mu^-$ et $K_S^0 \rightarrow \pi^+\pi^-$

Exemples de collisions...



$B^0 \rightarrow J/\Psi K^0_s$ suivi de $J/\Psi \rightarrow \mu^+\mu^-$ et $K^0_s \rightarrow \pi^+\pi^-$

Les questions abordées

- Une fois les données collectées, par quel(s) processus arrive t-on à une mesure de physique ?
- Quels sont les méthodes et outils utilisés ?

PLAN

- I. LA RECONSTRUCTION DES ÉVÈNEMENTS
- II. LES GRANDES LIGNES D'UNE ANALYSE DE PHYSIQUE
- III. LA SIMULATION
- IV. L'AJUSTEMENT

– PREMIÈRE PARTIE –

LA RECONSTRUCTION DES ÉVÈNEMENTS

Principe

Position voies électroniques

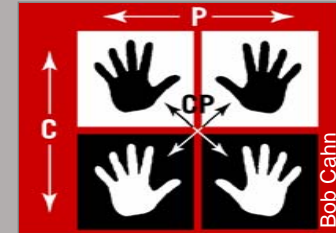
Calibration

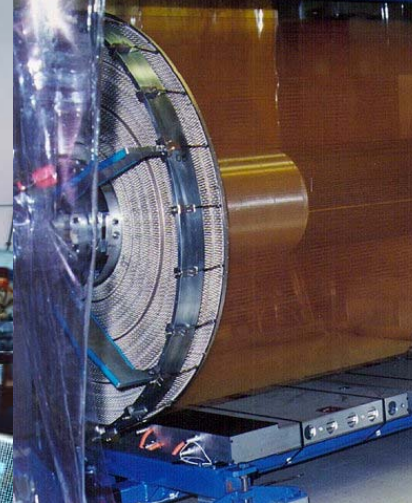
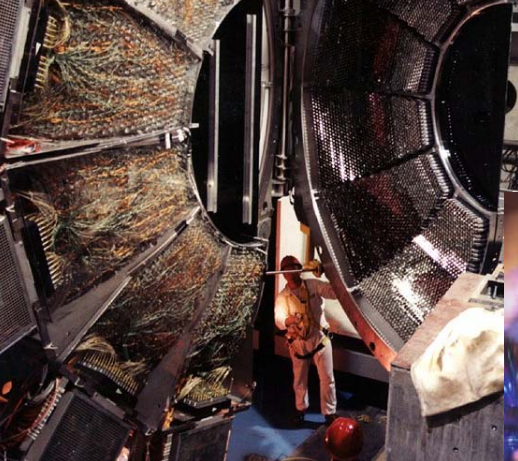
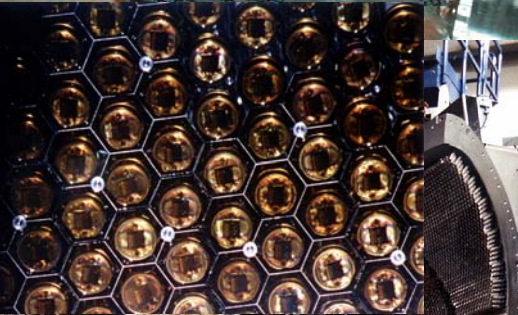
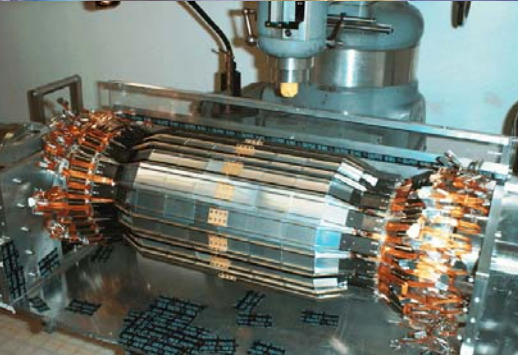
Patterns



Digression: que fait BABAR ?

- Mesure de la violation de CP :
 - C : changement particule \leftrightarrow anti-particule
 - P : transformation $(x,y,z) \leftrightarrow (-x,-y,-z)$
- Violation de CP a(urait) permis... notre existence:
 - Lors du « Big Bang », matière et anti-matière présentes en quantités égales auraient du s'annihiler ne laissant que de la lumière !
 - Violation de CP = (léger/infime) déséquilibre/dissymétrie des lois physiques entre matière et anti-matière.
 - Ce déséquilibre a permis à une infime quantité de matière de survivre, cette matière forme les objets d'aujourd'hui.
- La violation de CP est particulièrement visible dans le système $B^0 - \text{anti-}B^0$

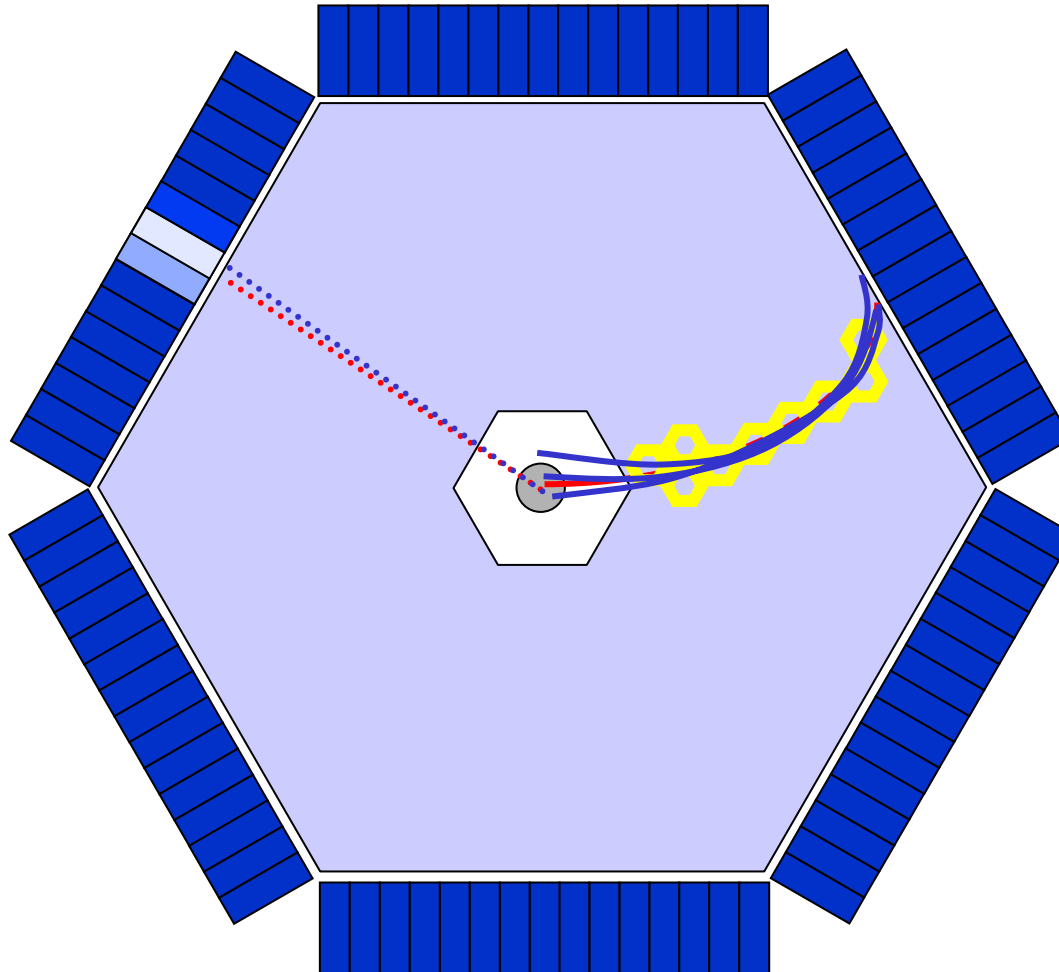




Reconstruction ?

- La reconstruction est l'opération qui permet de passer des données brutes d'un sous-système à des objets qu'il est conçu pour mesurer:
 - Détecteur de vertex:
 - Coups dans cellules \rightarrow vertex
 - Chambre à traces:
 - Coup dans cellules \rightarrow traces chargées
 - Calorimètre
 - Cristaux, cellules, etc... \rightarrow clusters
 - Détecteur Cherenkov:
 - Cellules, photo-multiplicateurs \rightarrow angle cherenkov
 - Etc...

Principe reconstruction



La Répartition des voies électroniques

RUN 0x003F06A

EVT 0x005DC0A

TTT

• La prise en compte des données de calibration, permet de corriger les disparités de voie à voie:

- A un même signal d'entrée, correspondra un même signal de sortie
- On obtiendra une réponse uniforme sur l'ensemble des voies d'un sous-système

• On « met de niveau » les voies d'électroniques

BASE DONNEES
CALO

BASE DONNEES
CHAMBRE
A TRACES

81
50
70
01
90
31
77

Calibration: mais encore...

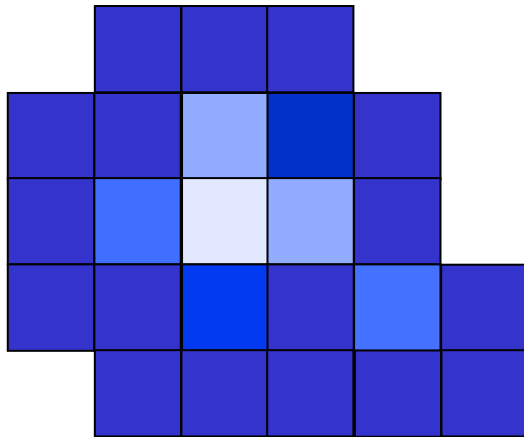
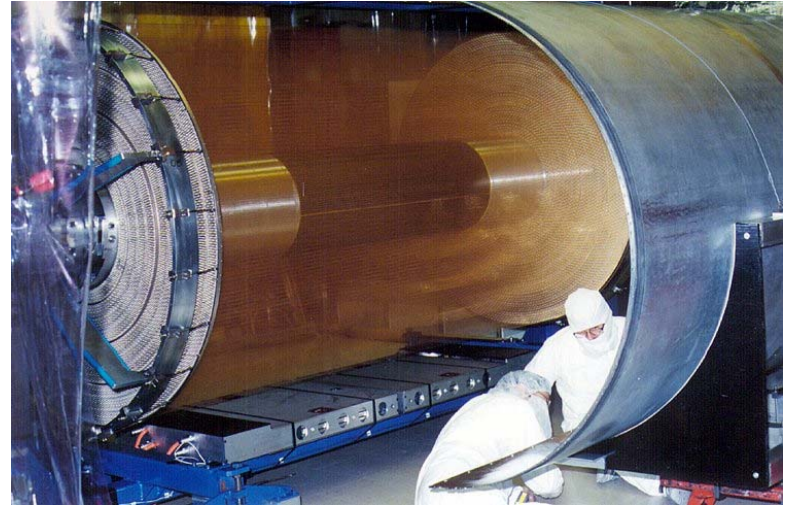
- La calibration ne concerne pas que la réponse des voies d'électroniques
- Certains effets physique doivent être corrigés
- Exemple:
 - Fil idéal:



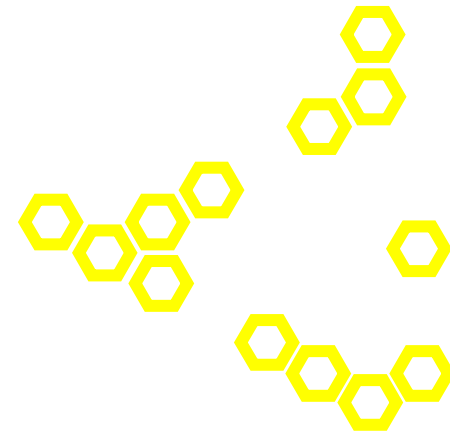
- Fil « réel »:



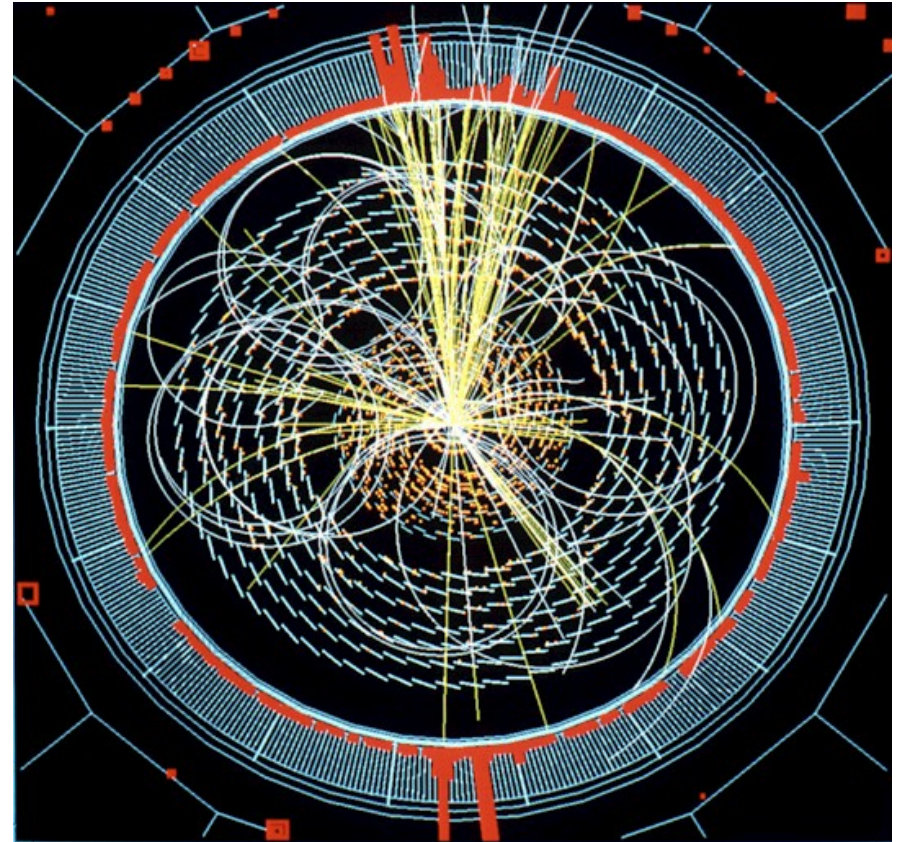
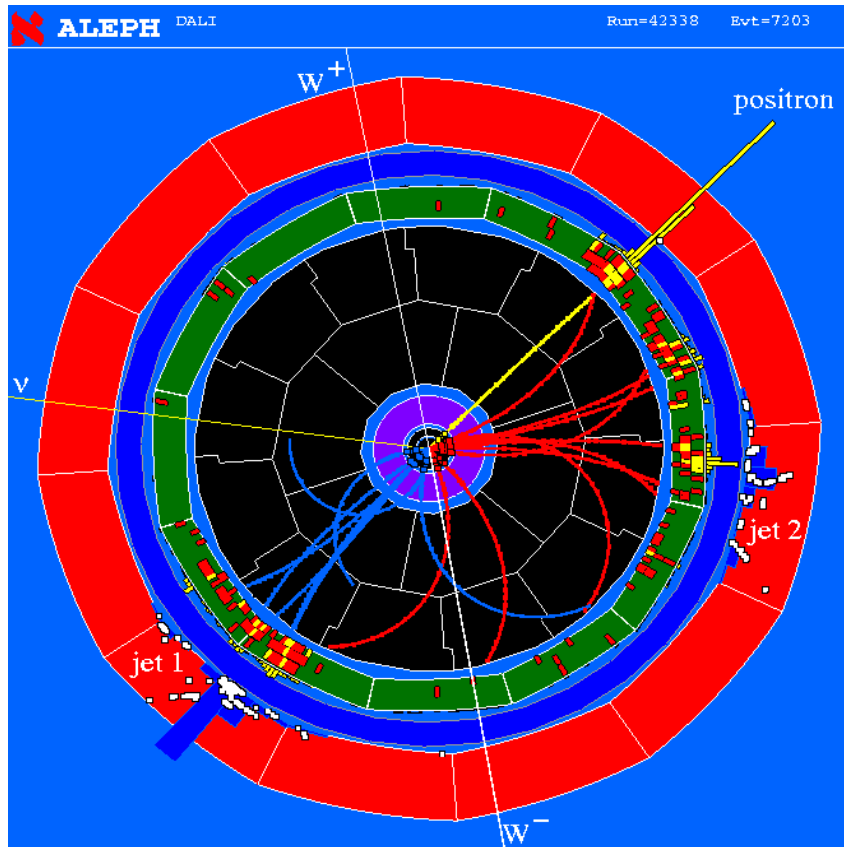
Algorithmie: patterns



- Ex pattern cluster



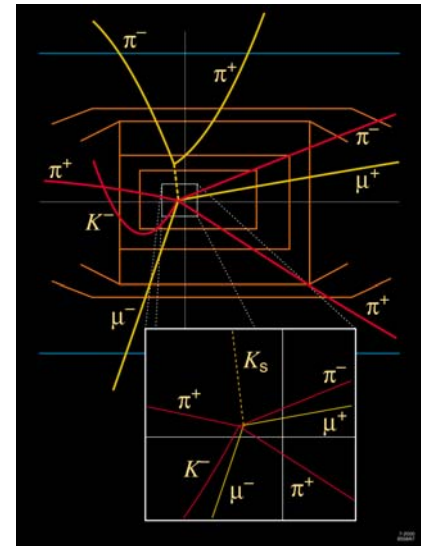
- Ex pattern chambre à traces



- Pas de choix idéal d'algorithme
- La reconstruction n'est donc pas un processus 100% efficace

Et ensuite...

- A l'issue de la reconstruction, on dispose d'objets comme:
 - Trace chargées, ou segment de traces (dans détecteur de vertex)
 - Clusters
 - etc...
- On inclue parfois dans la reconstruction la construction d'objets de plus haut niveau, qui combinent les informations de plusieurs sous-systèmes:
 - Trace dans détecteur de vertex + chambre à traces
 - Meilleur pointage au point d'interaction
 - Trace dans chambre à traces + cluster dans calorimètre
 - Identification d'électrons
 - Trace + trace:
 - Reconstructions de particules instables:
 $K_S^0 \rightarrow \pi^+\pi^-$
 - Cluster dans calorimètre électromagnétique sans trace associée
 - Photon
 - Trace + information cherenkov
 - Identification de Kaon
 - Etc...



Résumé

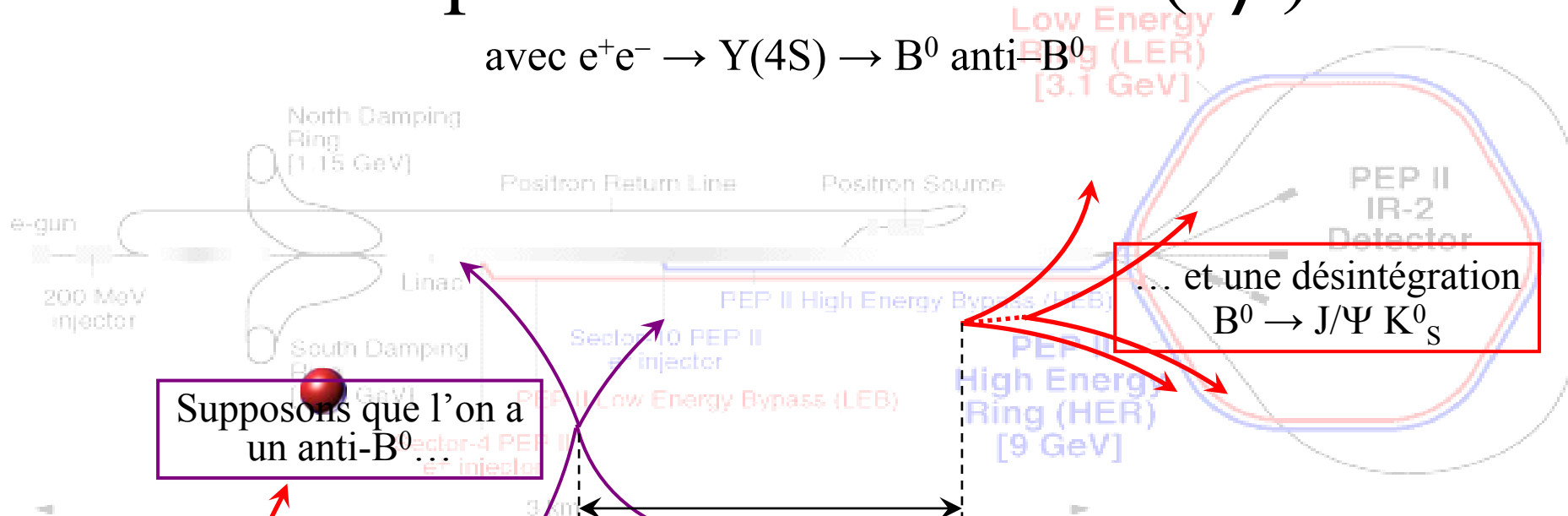
- La reconstruction est l'opération qui, partant des données brutes d'un détecteur, « réalise » sa fonction:
 - Production de vertex, traces, clusters, angles cherenkov...
- Cette reconstruction nécessite la connaissance de:
 - La position des voies d'électroniques
 - De la caractéristique de leur réponses
 - Piédestaux, gains, etc..
 - Ces informations sont consignées dans des bases de données
 - Des données de calibration sont prises régulièrement afin d'assurer le contrôle des voies d'électroniques
 - Ces données de calibration sont injectées dans la reconstruction
- La reconstruction se base sur des algorithmies
 - Elle fait le choix de patterns
 - Qui ont leurs limitations
- Les informations issues de la reconstruction peuvent elles-mêmes être combinées pour
 - Identifier les particules
 - Fabriquer de nouveaux objets/particules:
 - Ex: combinaison de traces chargées + vertexing → obtention de K^0_S
- ... on est à la lisière de l'analyse de physique.

Sur quoi s'appuie une mesure ?

- Le but d'une mesure de physique:
 - Extraire la valeur de paramètre(s) théorique(s)
 - Exemples:
 - Nombre de familles de neutrinos
 - $\sin(2\beta)$
 - Masse du Higgs
- Une mesure de physique exploite certaines désintégrations:
 - Connues pour leur sensibilité à ces paramètres
 - Exemples:
 - Nombre de famille de neutrinos : combine toutes les désintégrations du Z^0
 - Mesure de $\sin(2\beta)$: désintégrations $B^0 \rightarrow J/\Psi K_s^0$
 - Masse du Higgs: toutes désintégrations suffisamment abondantes (et reconnaissables)

Exemple: sensibilité à $\sin(2\beta)$

avec $e^+e^- \rightarrow Y(4S) \rightarrow B^0 \text{ anti-}B^0$



Supposons que l'on a un anti- B^0 ...

Mesure de Δz permet mesure de Δt

Et si $B^0 \rightarrow J/\Psi K_S^0$ survient en premier, on dit $\Delta t < 0$

... et une désintégration $B^0 \rightarrow J/\Psi K_S^0$

Au départ, dans BABAR et BELLE, la paire B anti-B est « boostée » vers l'avant grâce à l'asymétrie des faisceaux e^- (9 GeV), e^+ (3.1 GeV)

Exemple: sensibilité à $\sin(2\beta)$

avec $e^+e^- \rightarrow Y(4S) \rightarrow B^0 \text{ anti-}B^0$

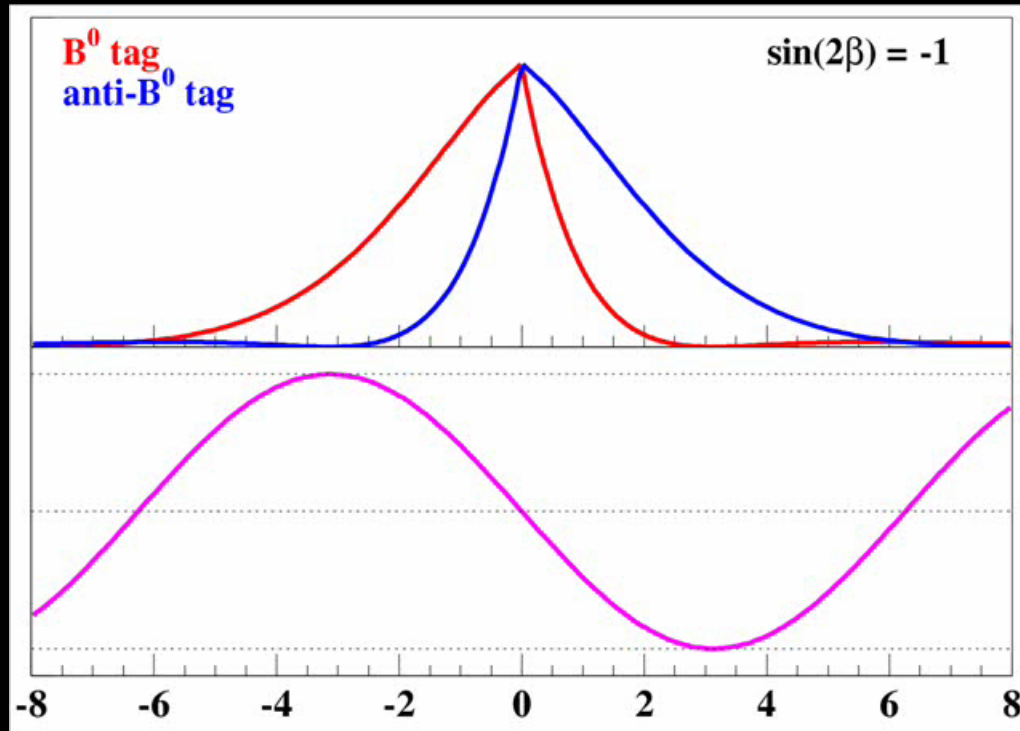
Probabilité de désintégration à Δt
(entre Δt et $\Delta t + \delta\Delta t$)

+1

Asymétrie =

$$\frac{B^0 - \text{anti-}B^0}{B^0 + \text{anti-}B^0}$$

-1



(cliquer sur l'animation pour démarrer)

Δt (ps)

Comment procède la mesure ?

- survol (puis plus de détails) -

- En trois grandes étapes:
 - I. La sélection des désintégrations d'intérêt
 - Désintégration d'intérêt = « signal »
 - On parle de « sélection », tout court
 - II. La modélisation des données sélectionnées
 - Prise en compte des imperfections/effets expérimentaux
 - Ou comment les données retenues représentent une vue « déformée » du modèle théorique à cause des effets expérimentaux
 - Permet ensuite l'extraction ou la mesure elle-même du/des paramètres théoriques recherchés
 - Réalisé par « ajustement » ou « fit »
 - Et fournit une incertitude « statistique » : la fourchette dans laquelle peuvent varier les paramètres du au fait que l'on ne dispose pas d'un nombre infini de désintégrations signal
 - III. L'estimation des incertitudes « systématiques »
 - La modélisation elle-même n'étant pas parfaite, quelles incertitudes cela entraîne-t-il sur les paramètres mesurés ?

Mais tout de même...

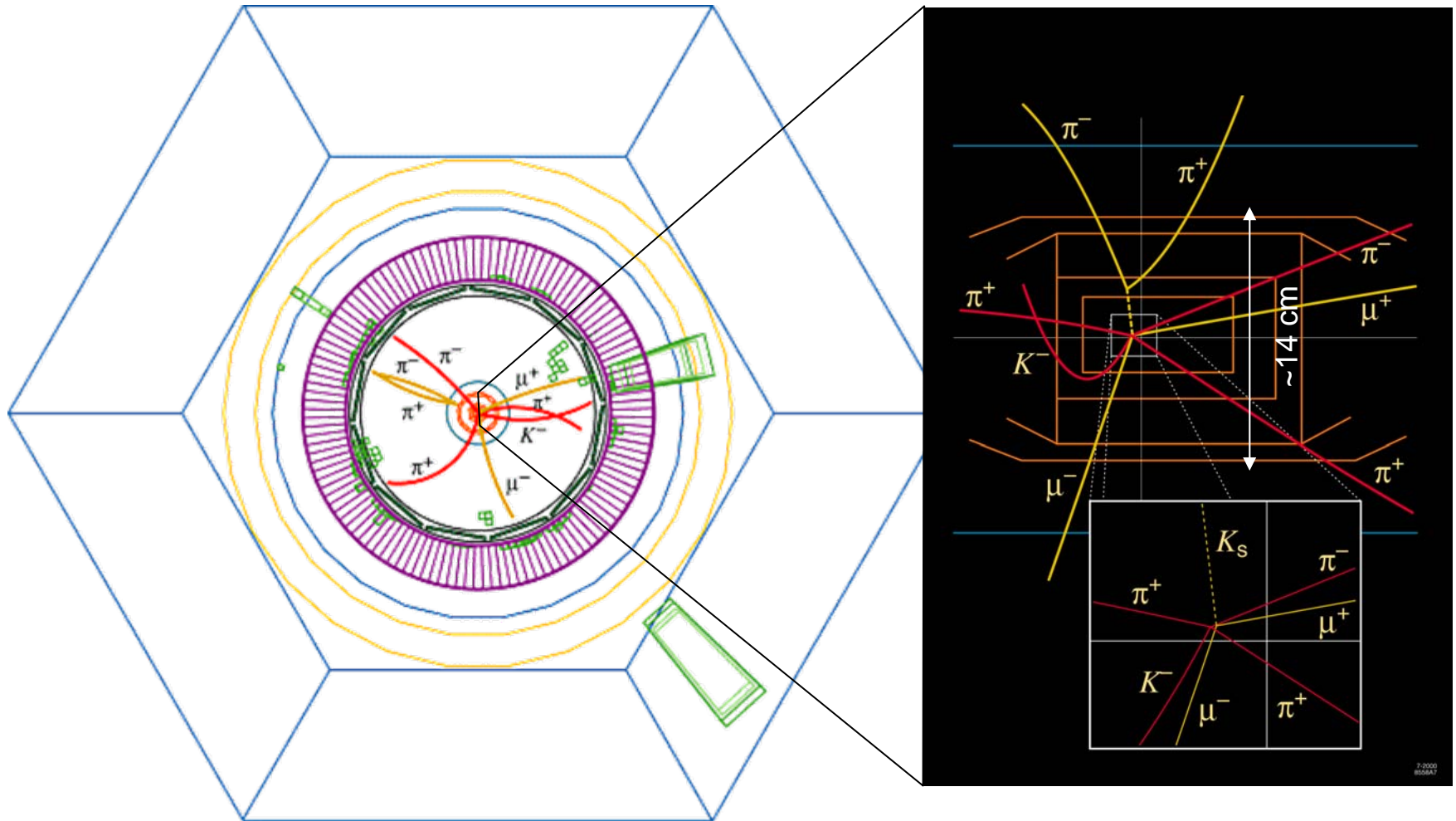
- Si ces grandes étapes se retrouvent dans la plupart des études de physiques, les problèmes à résoudre lors de chacune de ces étapes sont en général ****TRÈS**** différents d'une analyse à une autre !
- Ceci à cause:
 - des différences entre signaux
 - donc de la sélection à mettre en œuvre...
 - et donc des bruits de fonds
 - Et ceci, au sein d'une même expérience !
 - des conditions expérimentales dans lesquels ces signaux sont produits
 - Ex: machines hadroniques versus machines à électrons
 - Recherche/étude du Z^0 , Higgs...
 - Ou aussi production de mésons B:
 - Au LEP, production de B^0 , B^+ mais aussi B_s , etc...
 - A PEP-II/BABAR: production de B^0 et B^+ seulement
 - Etc...
- Les solutions à apporter diffèrent donc entre analyses de physique
- Illustration (non exhaustive) de ces étapes

I. La sélection

- Vise à séparer le signal des autres désintégrations
 - S'appuie sur les/des caractéristiques du signal
 - Qu'est-ce qui le différencie des autres désintégrations ?
- Processus pas 100% efficace
 - Des désintégrations sont perdues
 - A cause de la sélection elle-même
 - Mais au départ aussi du détecteur: n'étant pas complètement « hermétique », certaines particules « échappent » à la détection
 - On perd couramment de 10% à 90% de signal !
- Processus pas 100% sélectif
 - Des désintégrations non désirées sont retenues
 - On parle de « bruit de fond »
 - Les niveaux de bruit de fond sont extrêmement variables: de ~100% de pureté à quelques %, voire moins

Exemple sélection: $B^0 \rightarrow J/\Psi K_S^0$ (0)

$J/\Psi \rightarrow \mu^+\mu^-$ et $K_S^0 \rightarrow \pi^+\pi^-$



Exemple sélection

- Le problème:
 - Rapport de branchement
 - C'est-à-dire une désintégration
 - désintégrations de $B^0 \dots$
- Comment distinguer ces
 - Deux critères:
 - Le J/Ψ :
 - Le J/Ψ (qui est de charge nulle) se désintègre en une paire d'électrons et paire de muons dans chaque cas
 - $J/\Psi \rightarrow e^+e^-$ ou $\mu^+\mu^-$
 - Soit 12 % des désintégrations
 - Les leptons sont plutôt « rares » (la désintégration la plus fréquente est le pion), et les hadrons sont nombreux
 - \Rightarrow Bon critère de distinction
 - Cela veut dire aussi que les désintégrations de J/Ψ !

$J/\psi(1S)$ DECAY MODES	Fraction (Γ_i/Γ)	
hadrons	(87.7 ± 0.5) %	
virtual $\gamma \rightarrow$ hadrons	(17.0 ± 2.0) %	
e^+e^-	(5.93 ± 0.10) %	1548
$\mu^+\mu^-$	(5.88 ± 0.10) %	1545
Decays involving hadronic resonances		
$\rho\pi$	(1.27 ± 0.09) %	1448
$\rho^0\pi^0$	(4.2 ± 0.5) × 10 ⁻³	1448
$a_2(1320)\rho$	(1.09 ± 0.22) %	1123
$\omega\pi^+\pi^+\pi^-\pi^-$	(8.5 ± 3.4) × 10 ⁻³	1392
$\omega\pi^+\pi^-$	(7.2 ± 1.0) × 10 ⁻³	1435
$\omega f_2(1270)$	(4.3 ± 0.6) × 10 ⁻³	1142
$K^*(892)^0\bar{K}_2^*(1430)^0 + c.c.$	(6.7 ± 2.6) × 10 ⁻³	1012
$\omega K^*(892)\bar{K} + c.c.$	(5.3 ± 2.0) × 10 ⁻³	1097
$K^+\bar{K}^*(892)^- + c.c.$	(5.0 ± 0.4) × 10 ⁻³	1373
$K^0\bar{K}^*(892)^0 + c.c.$	(4.2 ± 0.4) × 10 ⁻³	1373
$K_1(1400)^\pm K^\mp$	(3.8 ± 1.4) × 10 ⁻³	1171
$\omega\pi^0\pi^0$	(3.4 ± 0.8) × 10 ⁻³	1436
$b_1(1235)^\pm\pi^\mp$	[gg] (3.0 ± 0.5) × 10 ⁻³	1300
$\omega K^\pm K_S^0\pi^\mp$	[gg] (2.9 ± 0.7) × 10 ⁻³	1210
$b_1(1235)^0\pi^0$	(2.3 ± 0.6) × 10 ⁻³	1407
$\phi K^*(892)\bar{K} + c.c.$	(2.04 ± 0.20) × 10 ⁻³	1496
$\omega K\bar{K}$		1433
$\omega f_0(1710) \rightarrow \pi^+\pi^-\pi^0$	(3.37 ± 0.26) %	1533
$\phi 2(\pi^+\pi^-)\pi^0$	(2.9 ± 0.6) %	1368
$\phi 3(\pi^+\pi^-)\pi^0$	(1.50 ± 0.20) %	1345
$\phi \pi^+\pi^-\pi^0$	(1.20 ± 0.30) %	1407
$\phi \pi^+\pi^-\pi^0 K^+K^-$	(9.0 ± 3.0) × 10 ⁻³	875
$\phi 4(\pi^+\pi^-)\pi^0$	(7.2 ± 2.3) × 10 ⁻³	768
$\phi \pi^+\pi^- K^+K^-$	(1.10 ± 0.29) × 10 ⁻³	938
$\phi (\rho^+\rho^-) \rightarrow (\pi^+\pi^-) + c.c.$	[gg] (1.03 ± 0.13) × 10 ⁻³	697
$\phi \rho^+\rho^- (958)$	(9 ± 4) × 10 ⁻⁴	S=1.7 596
$\phi f_2'(1525)$	(8 ± 4) × 10 ⁻⁴	S=2.7 871
$\phi \pi^+\pi^-$	(8.0 ± 1.2) × 10 ⁻⁴	1365
$\phi K^\pm K_S^0\pi^\mp$	[gg] (7.2 ± 0.9) × 10 ⁻⁴	1114

Extrait PDG
Il y a en 4 pages !

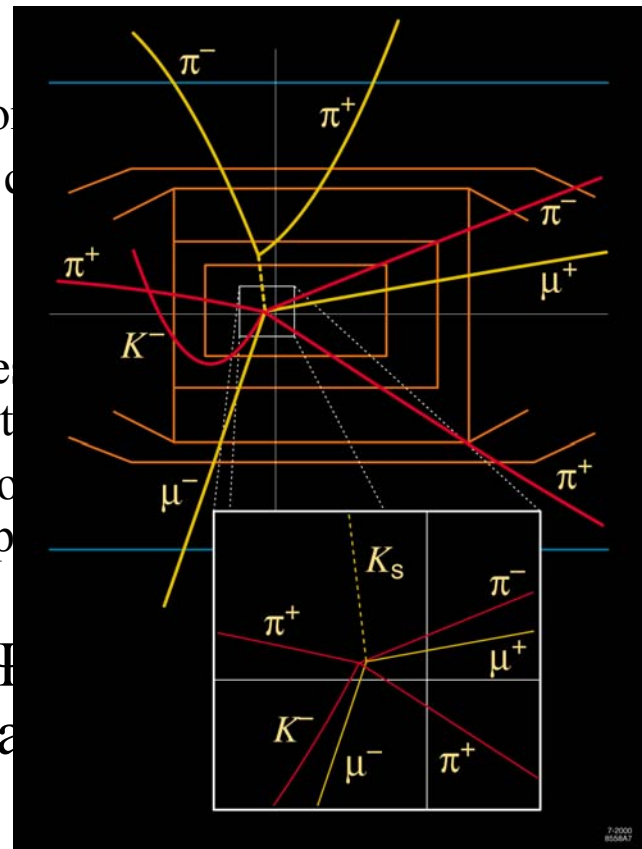
ans

Exemple sélection: $B^0 \rightarrow J/\Psi K_S^0$ (2)

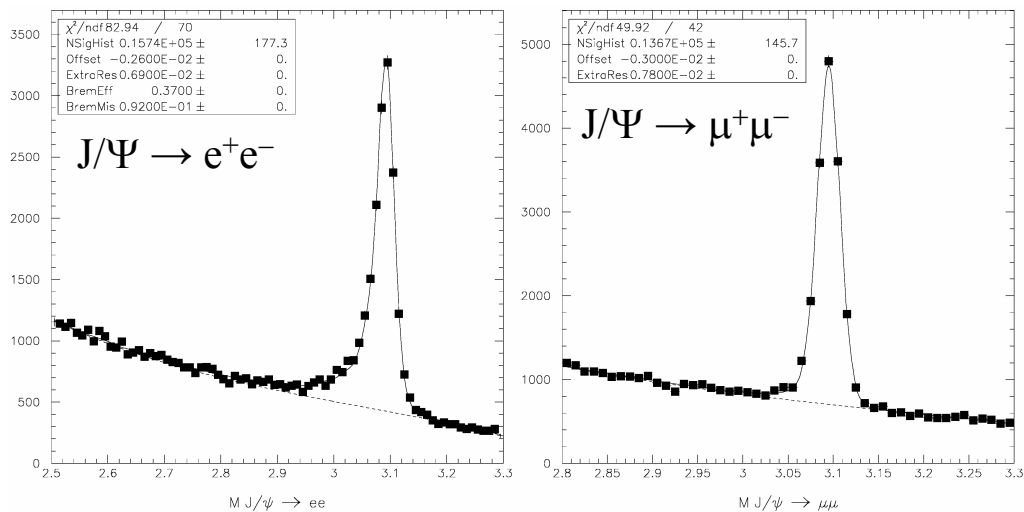
- Le K_S^0 :
 - On utilise sa désintégration en paire de pions
 - » Chargés et neutre (je ne parle que du chargé)
 - » $K_S^0 \rightarrow \pi^+\pi^-$
 - Le K_S^0 « vole » avant de se désintégrer
 - » Son vertex (point de désintégration) est décalé
 - » (d'où proviennent la majorité des particules)
 - » Et donc permet de sélectionner les pics en invariant de masse laissant de côté les nombreux autres pics d'interaction

- On assemble ensuite les (candidats) J/Ψ en combinant les paires de muons qui ont une masse comparée à celle du J/Ψ ($5.279 \text{ GeV}/c^2$)

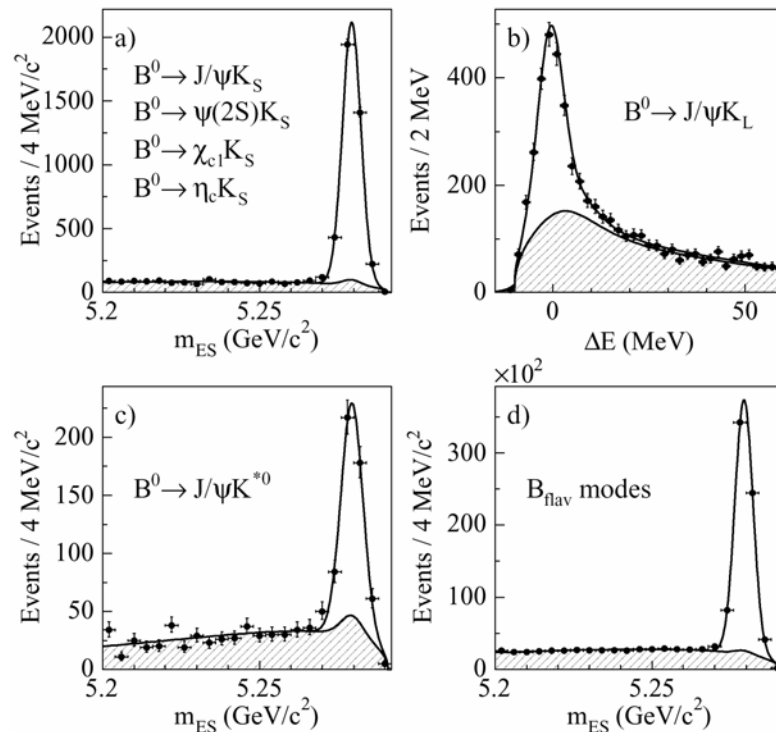
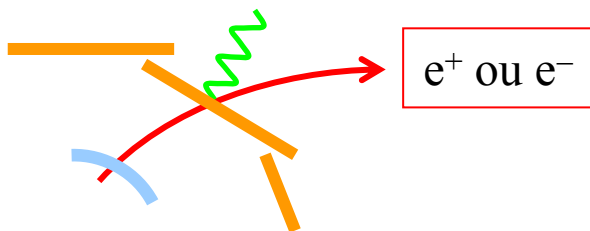
- Assemblage = somme des quadri-vecteurs du J/Ψ et K_S^0
 - En vérité, des contraintes cinématiques sont ajoutées, qui permettent d'affiner la résolution sur la masse etc...



Exemple de candidats sélectionnés



J/Ψ en paire de leptons

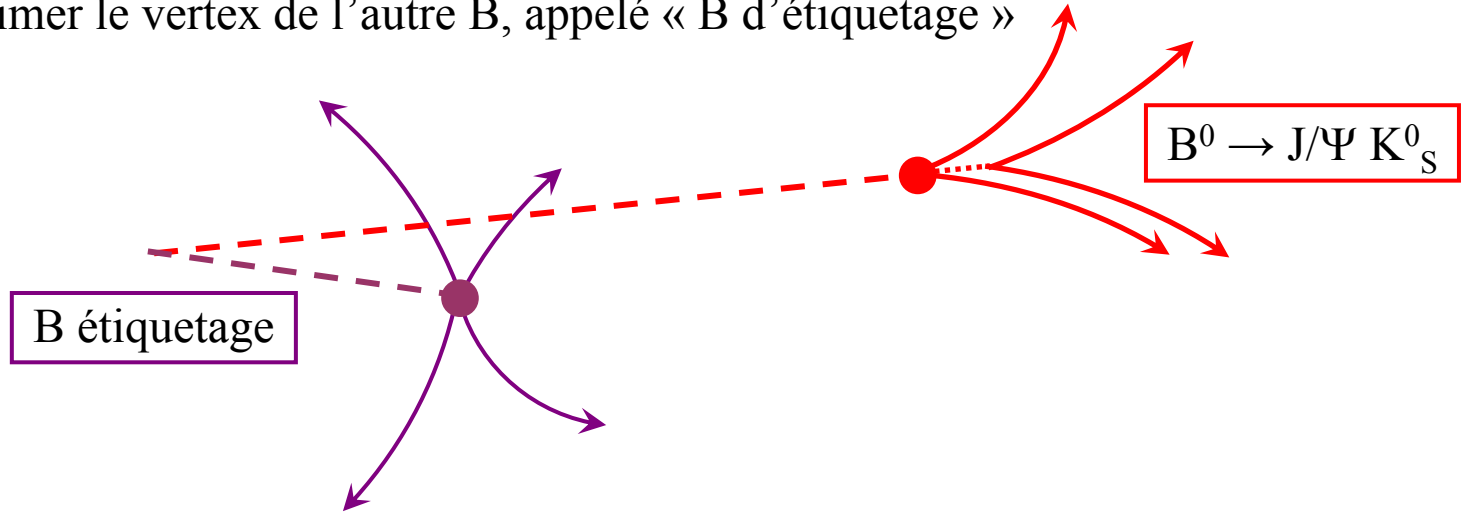


B en divers modes
(sur 227 millions de paires produites !)

- Un outil (quasiment) indispensable pour étudier les critères de sélection est la « simulation »
 - on en reparlera

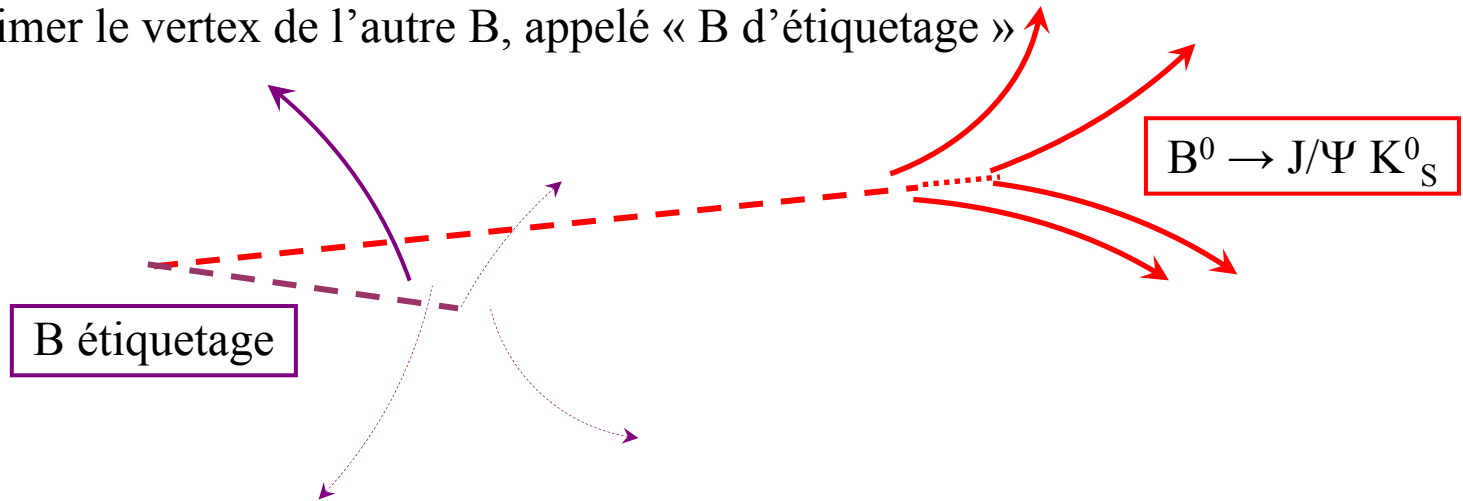
Données collectées

- On peut alors collecter, pour chacun de ces événements sélectionnés, la quantité sur laquelle se base la mesure
- Dans notre exemple avec $\sin(2\beta)$, il s'agit de Δt
- Cela suppose:
 - De mesurer le point de désintégration, ou vertex, du (candidat) $B^0 \rightarrow J/\Psi K_S^0$
 - D'estimer le vertex de l'autre B, appelé « B d'étiquetage »



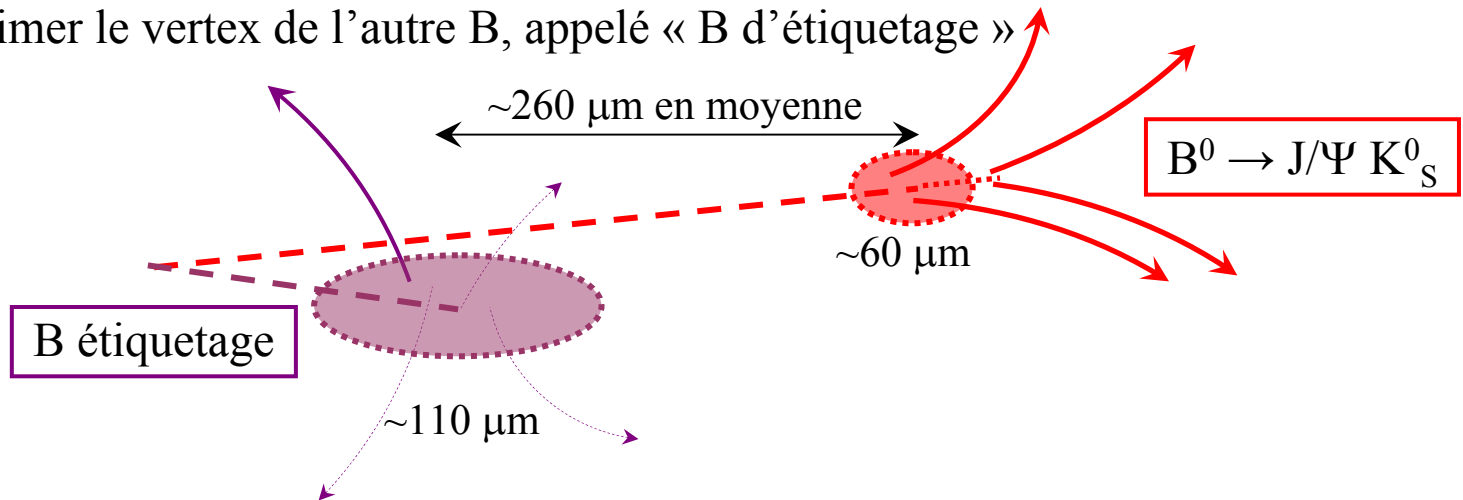
Données collectées

- On peut alors collecter, pour chacun de ces événements sélectionnés, la quantité sur laquelle se base la mesure
- Dans notre exemple avec $\sin(2\beta)$, il s'agit de Δt
- Cela suppose:
 - De mesurer le point de désintégration, ou vertex, du (candidat) $B^0 \rightarrow J/\Psi K_S^0$
 - D'estimer le vertex de l'autre B, appelé « B d'étiquetage »



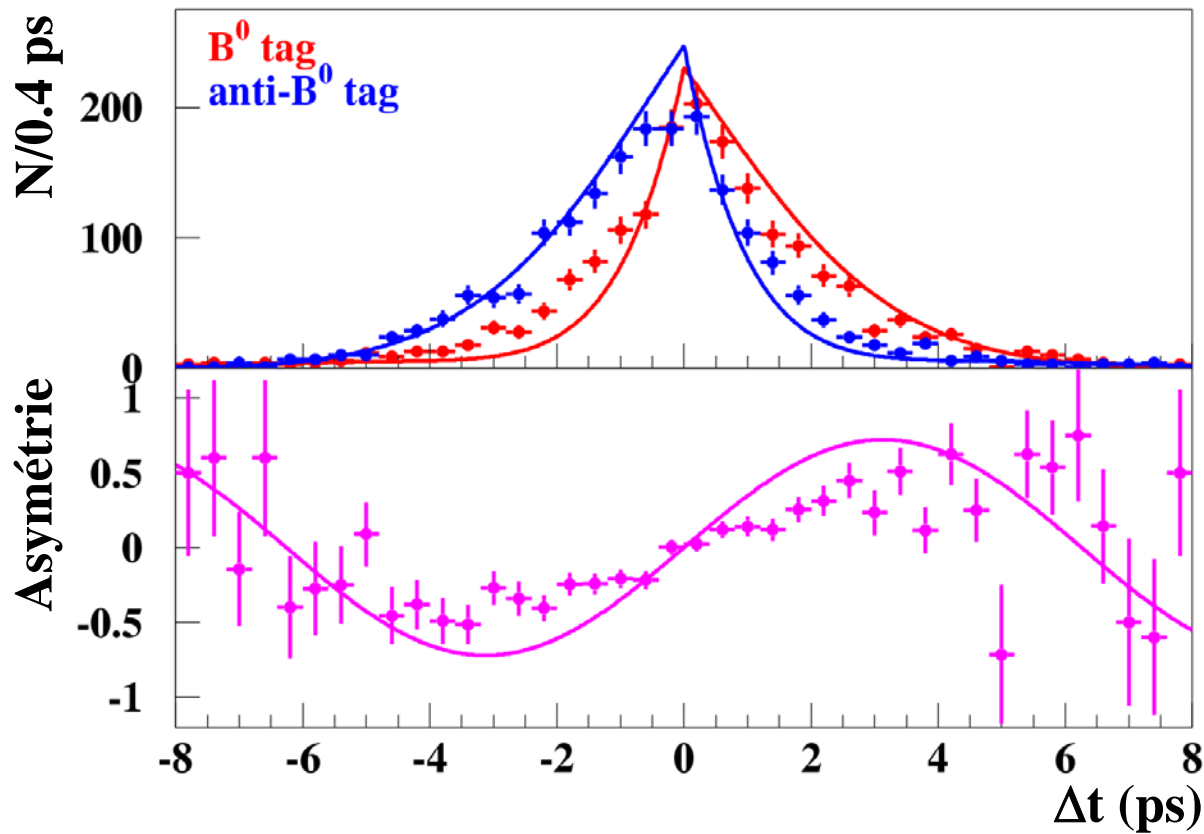
Données collectées

- On peut alors collecter, pour chacun de ces événements sélectionnés, la quantité sur laquelle se base la mesure
- Dans notre exemple avec $\sin(2\beta)$, il s'agit de Δt
- Cela suppose:
 - De mesurer le point de désintégration, ou vertex, du (candidat) $B^0 \rightarrow J/\Psi K_S^0$
 - D'estimer le vertex de l'autre B, appelé « B d'étiquetage »



- Et comme il nous faut distinguer les courbes B^0 et anti- B^0 pour le B d'étiquetage, il faut aussi une méthode qui permette cela
 - Qui, à nouveau, n'est ni 100% efficace : on ne sait parfois pas répondre
 - ni 100% sélective : on se trompe parfois
 - Ce point est critique dans cette analyse, mais trop long à développer ici

Allure des données collectées

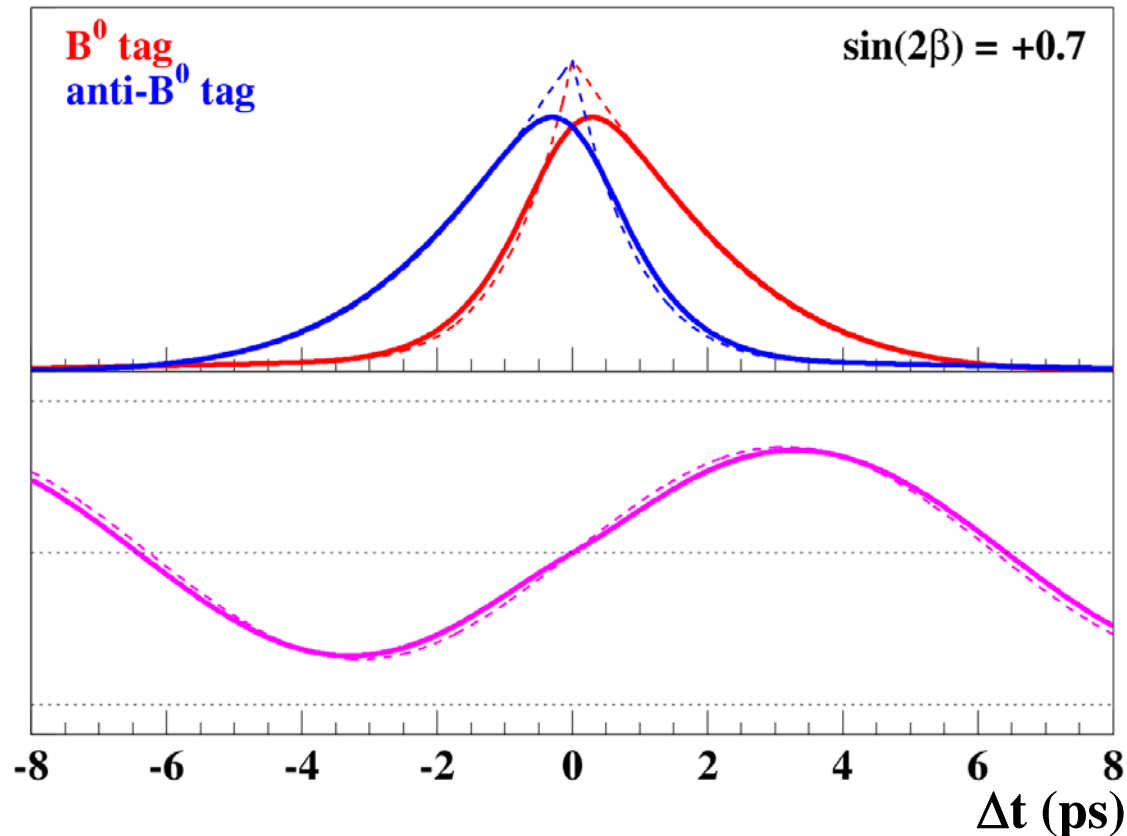


- Et si on superpose le modèle théorique utilisant le résultat publié... 😞
- Normal : on n'a pas tenu compte des effets expérimentaux, l'appareillage n'est pas parfait...

II. La modélisation

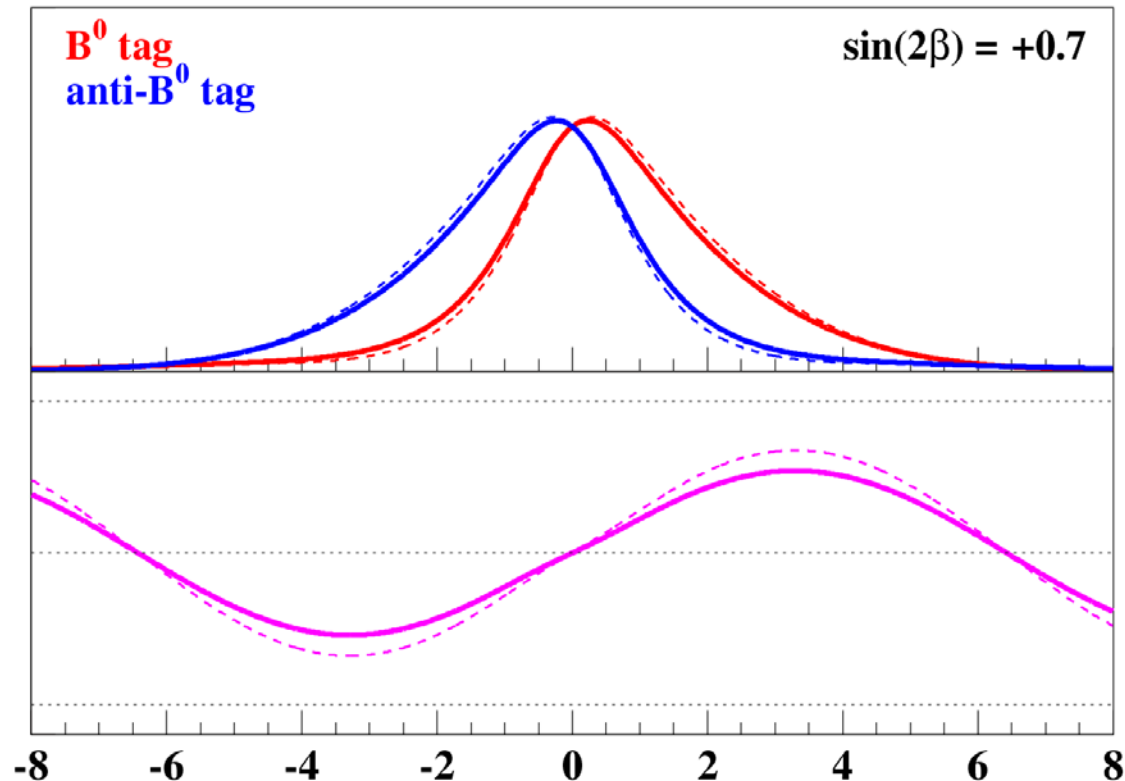
- L'échantillon de données sélectionné/retenu ne peut pas en général être décrit directement par le modèle théorique initial à cause des:
 - Pertes de signal
 - Imperfections de mesures de positions, énergies, etc...
 - Ce sont les effets de résolution
 - Présence de bruits de fond
- Le modèle théorique initial doit donc être adapté au cas expérimental
 - Si l'on ne veut pas fausser la mesure des paramètres
- Là encore, la simulation est utilisée pour :
 - Comprendre et mesurer les effets de résolutions
 - identifier les sources de bruit de fond.

Exemple de « déformations » du modèle théorique par les effets expérimentaux (1)



- Déformation due à la résolution sur Δt (ici 0.5 ps)
 - Qui est due à la résolution sur Δz

Exemple de « déformations » du modèle théorique par les effets expérimentaux (2)

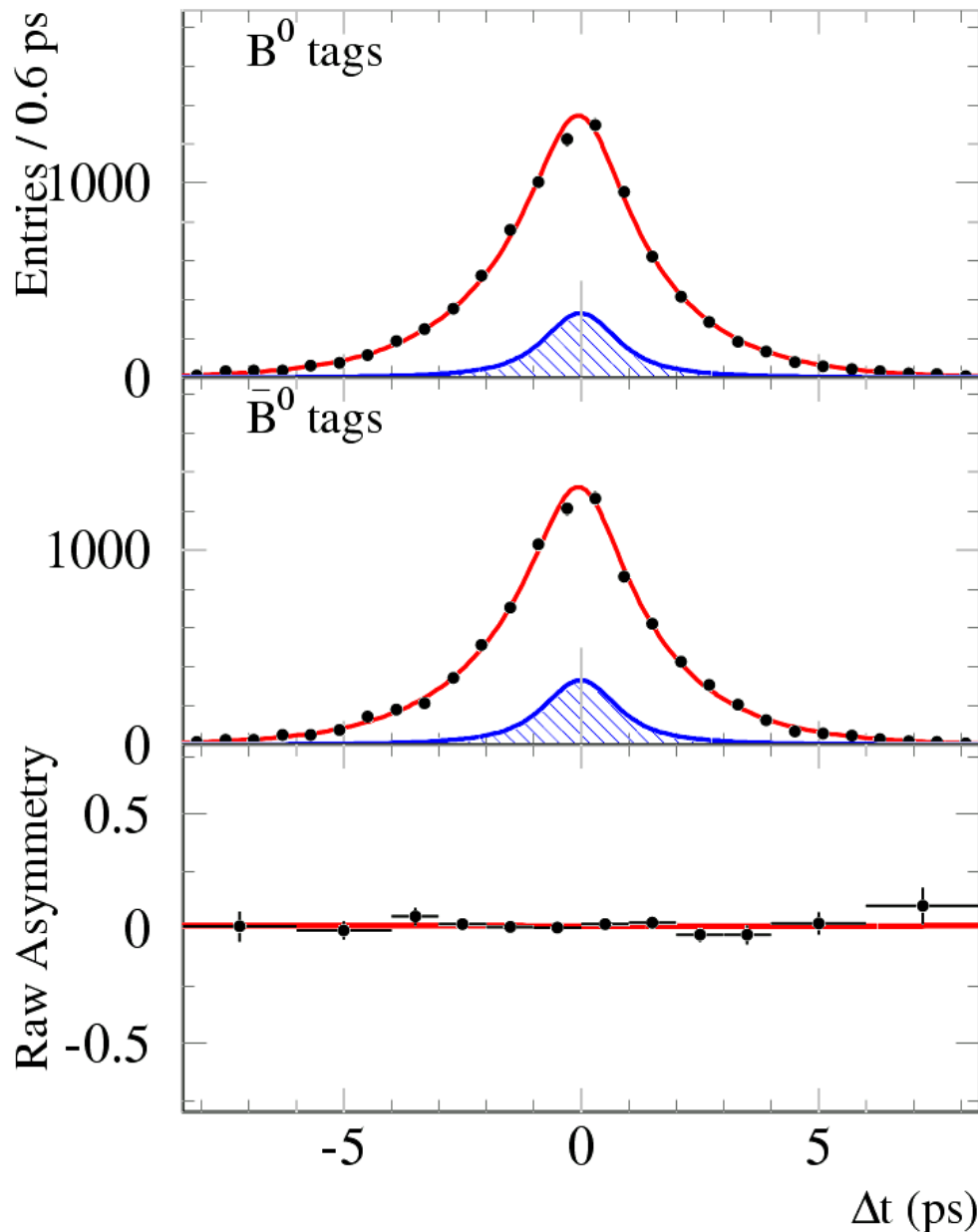


- Déformations due à la résolution sur Δt (ici 0.5 ps) et aux erreurs d'étiquetage (ici 10%)
 - Il en existe d'autres...

Validation(s)

- Avant de se lancer sur un ajustement des paramètres sur les données (c'est-à-dire de faire la mesure), il faut s'assurer que l'on est bien capable... de faire la mesure
 - Ce qu'on appelle la « validation »
- On utilise la simulation/Monté Carlo pour produire des échantillons de données simulées (signal et bruit), que l'on traite comme les données réelles
 - Ces échantillons sont produits avec des valeurs de paramètres connues:
 - Par exemple $\sin(2\beta) = 0.7$
 - Et l'on s'assure que l'ajustement sur le Monté Carlo retourne bien une valeur compatible avec celle utilisée pour la génération.
- On utilise aussi si possible – et autant que possible – des échantillons de données réelles, échantillons de contrôle, pour lesquelles on connaît la valeur du paramètre
 - Par exemple des échantillons pour lesquels $\sin(2\beta) = 0$
 - Permet de « débusquer » d'éventuels effets non pris en compte par la simulation

Exemple échantillon de contrôle

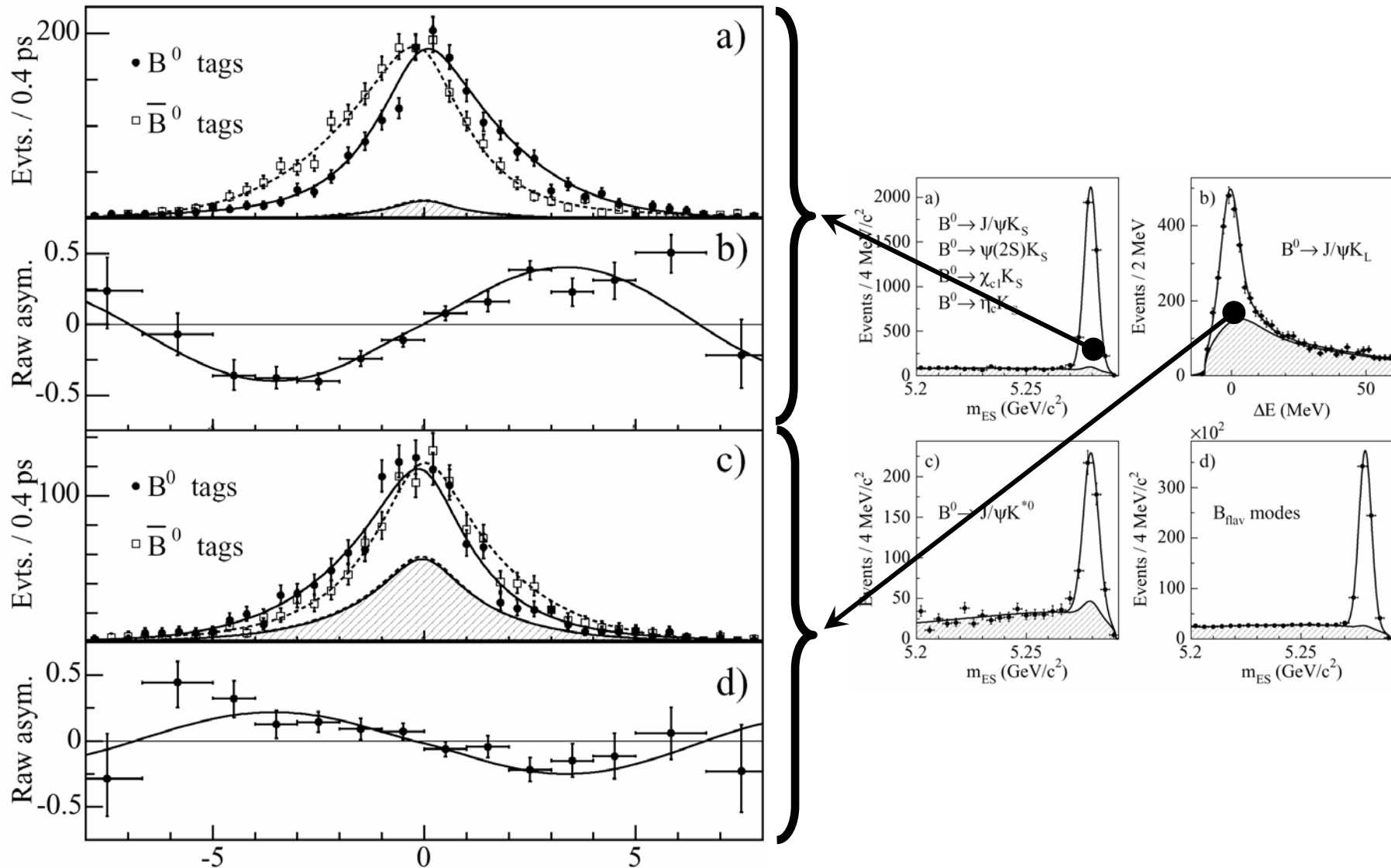


L'ajustement

- Une fois **tous** les effets expérimentaux pris en compte, on peut « ajuster » les paramètres théoriques pour que le modèle « colle au mieux » avec les données
 - On reparlera de comment cet ajustement se fait

Un autre problème pour après...

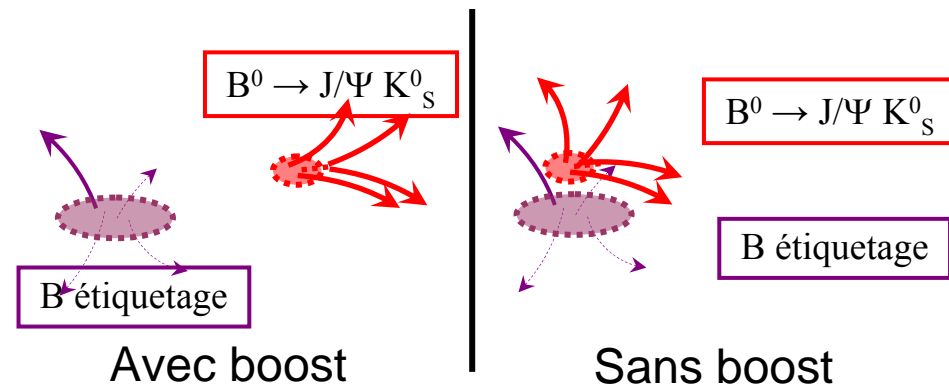
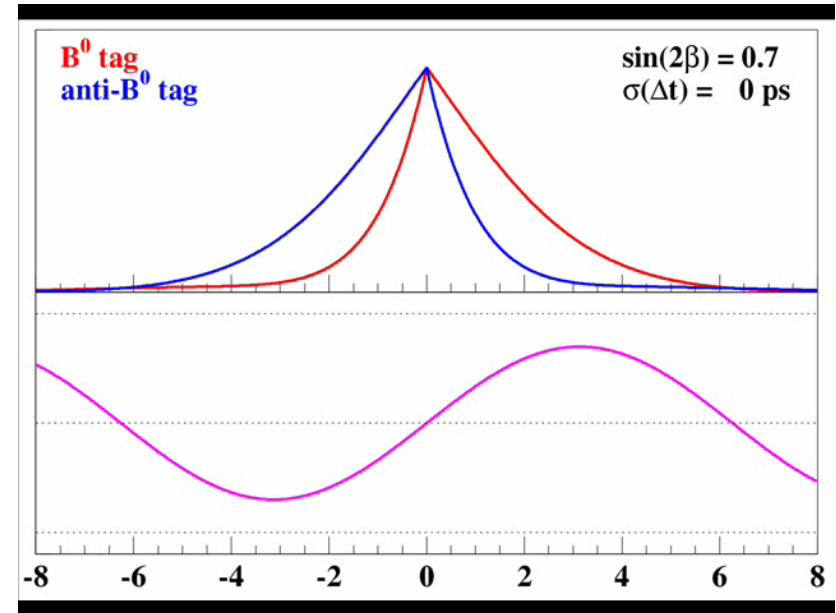
Exemple de résultat d'ajustement



Digression :

retour sur la conception de l'expérience

- On peut illustrer le rôle clef de la conception de l'expérience (accélérateur et détecteur) sur cette mesure :
 - La résolution en Δt , et donc Δz , doit être suffisamment bonne pour permettre la mesure de $\sin(2\beta)$
 - En effet...
- Cette résolution est ici obtenue grâce à un accélérateur asymétrique:
 - Les deux mésons B sont projetés (boostés) vers l'avant
 - (permet d'avoir la différence Δt)
- Et une résolution du détecteur de vertex suffisante
 - Et d'une méthode de reconstruction du vertex du « B d'étiquetage », dont le vertex ne peut être estimé que grossièrement, car incomplètement reconstruit...



III. Les incertitudes systématiques

- Le modèle utilisé pour tenir en compte les effets expérimentaux a lui-même ses imperfections:
 - A-t-on pris en compte tous les effets expérimentaux ?
 - La façon dont ces effets sont pris en compte décrit-elle effectivement ces effets ?
 - Dit autrement: le modèle est-il bien adapté ?
 - Dans quelle mesure les hypothèses faites sont-elles justifiées ?
- Le physicien doit donc faire la critique de son modèle et de là, estimer/quantifier les variations sur les paramètres mesurés
 - ... faire son « travail de physicien »
- A nouveau, derrière cette problématique générale, chaque analyse de physique pose des problèmes spécifiques
 - Il n'y a pas de méthode systématique pour étudier les effets systématiques !

Exemple d'incertitudes dues à des effet systématiques

TABLE III: Sources of systematic error on $\sin 2\beta$ and $|\lambda|$.

Source	$\sigma(\sin 2\beta)$	$\sigma(\lambda)$
CP backgrounds	0.012	0.002
Δt resolution function	0.011	0.003
$J/\psi K_L^0$ backgrounds	0.011	N/A
Mistag fraction differences	0.007	0.001
Beam spot	0.007	0.001
$\Delta m_d, \tau_B, \Delta\Gamma/\Gamma, \lambda $	0.005	0.001
Tag-side interference	0.003	0.012
MC statistics	0.003	0.003
Total systematic error	0.023	0.013

- Incertitudes ajoutées quadratiquement
 - C'est-à-dire qu'on les suppose indépendantes.

Une fois tout ceci fait, on peut publier...

(hep-ex/0408127)

Improved Measurement of CP Asymmetries in $B^0 \rightarrow (c\bar{c})K^{0(*)}$ Decays

We present results on time-dependent CP asymmetries in neutral B decays to several CP eigenstates. The measurements use a data sample of about $227 \times 10^6 \Upsilon(4S) \rightarrow B\bar{B}$ decays collected by the *BABAR* detector at the PEP-II asymmetric-energy B Factory at SLAC. The amplitude of the CP asymmetry, $\sin 2\beta$ in the Standard Model, is derived from decay-time distributions from events in which one neutral B meson is fully reconstructed in a final state containing a charmonium meson and the other B meson is determined to be either a B^0 or \bar{B}^0 from its decay products. We measure $\sin 2\beta = 0.722 \pm 0.040(\text{stat}) \pm 0.023(\text{syst})$ in agreement with the Standard Model expectation.

Résumé...

- Une analyse de physique vise à extraire une valeur de paramètres théoriques
- Ces paramètres sont accessibles *via* l'étude de certaines désintégrations
 - Qui conditionnent aussi la conception de l'expérience
- Les désintégrations signal sont extraites de l'ensemble des données grâce à un processus de « sélection »
 - Qui n'est ni 100% efficace, ni 100% sélectif
- Une modélisation des données extraites est nécessaire pour prendre en compte les diverses imperfections expérimentales afin de pouvoir décrire les données observées
 - Perte d'évènements de signal
 - Résolution(s) du détecteur
 - Présence de bruits de fond
 - Etc...
- Cette modélisation permet alors d'ajuster le(s) paramètre(s) théoriques pour « coller au mieux » aux données observées
- Les imperfections de la modélisation doivent être étudiées afin d'estimer les incertitudes induites sur l'estimation des paramètres théoriques

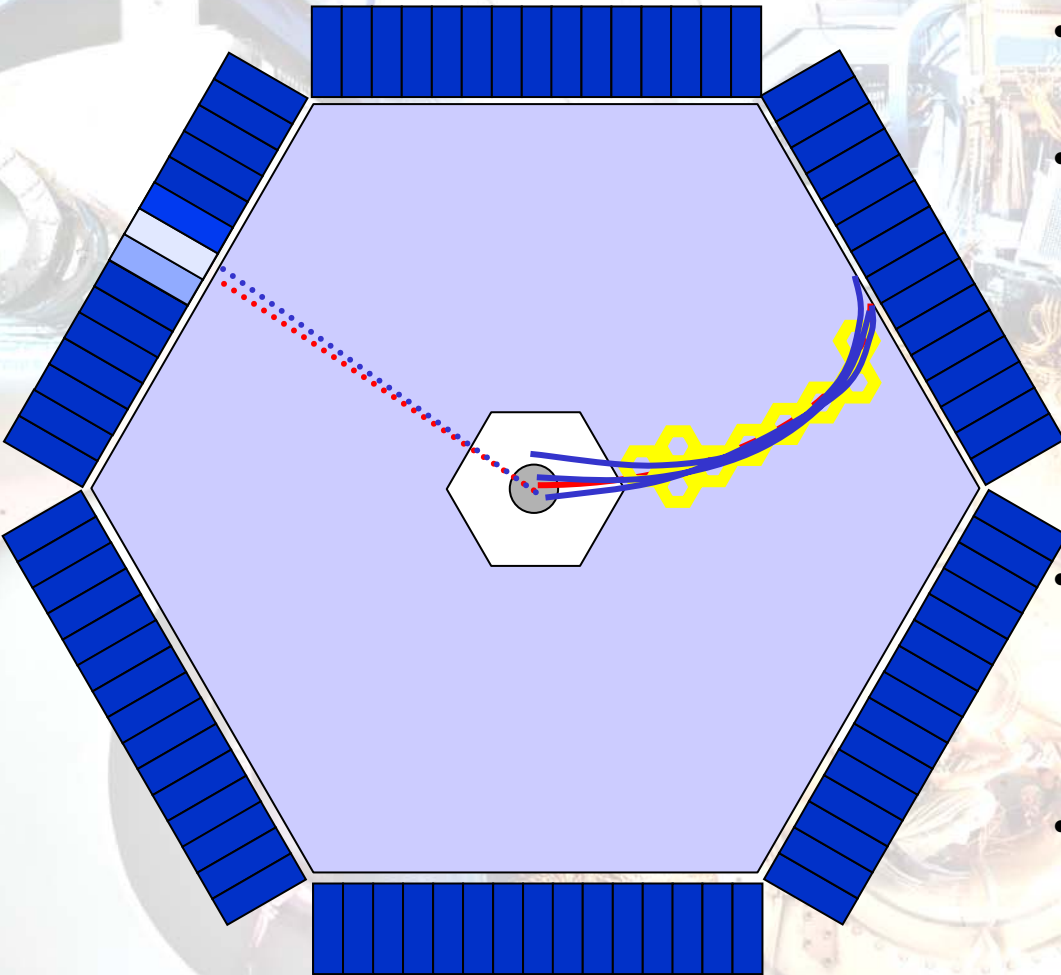
... et perspectives

- Au cours de ce processus d'analyse de physique, nous avons vu l'utilisation et la nécessité d'outils comme
 - La « simulation », utilisée à différents moments
 - Et qui joue un rôle clé
 - D'une méthode permettant de faire « coller au mieux » un modèle sur un échantillon de données
 - Ce qu'on appelle « ajustement » ou « fit »
- Cela fera l'objet des cours suivants.

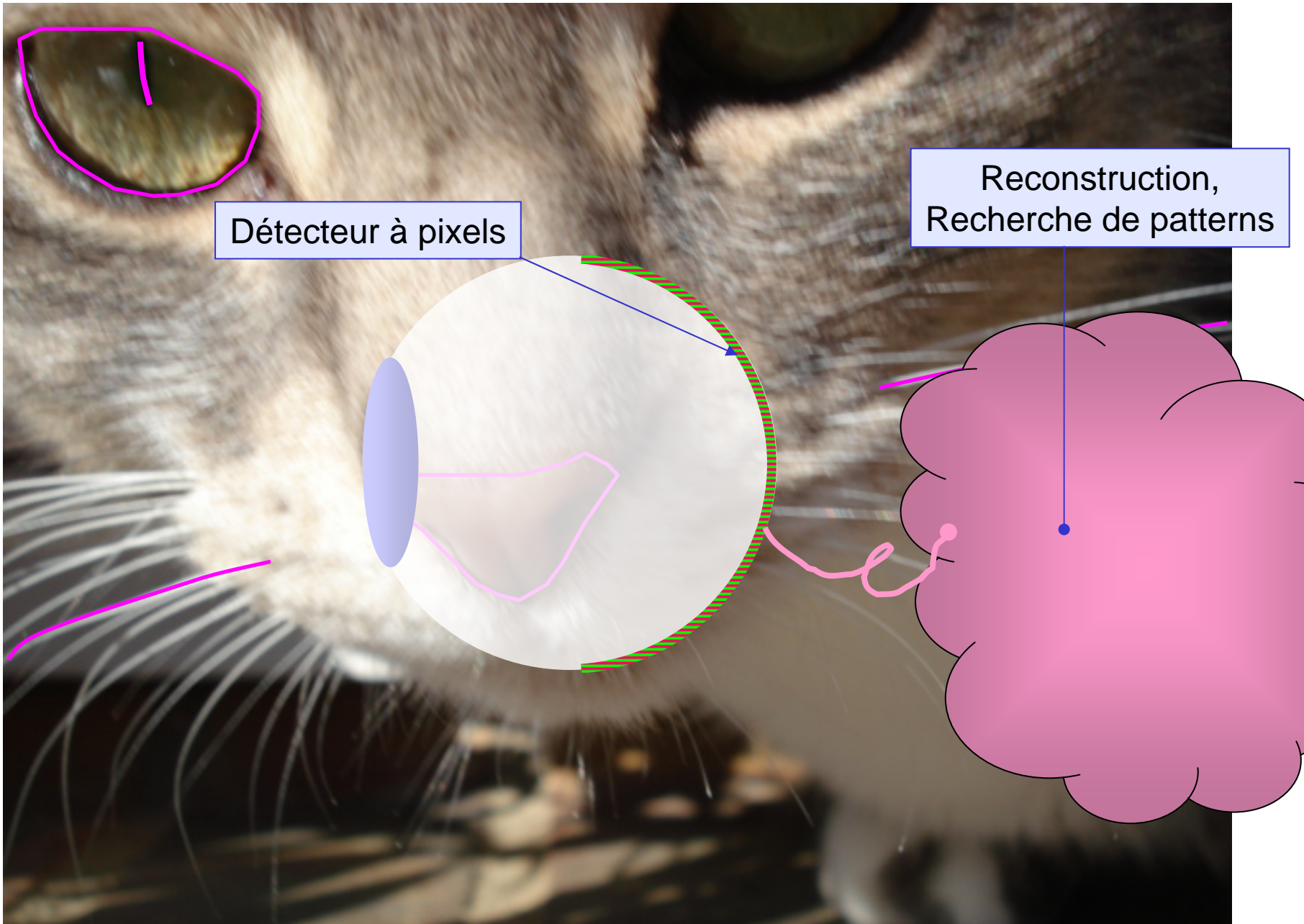
Résumé sessions précédentes

- I. Reconstruction d'évènements
- II. Analyse de physique

Reconstruction d'évènements

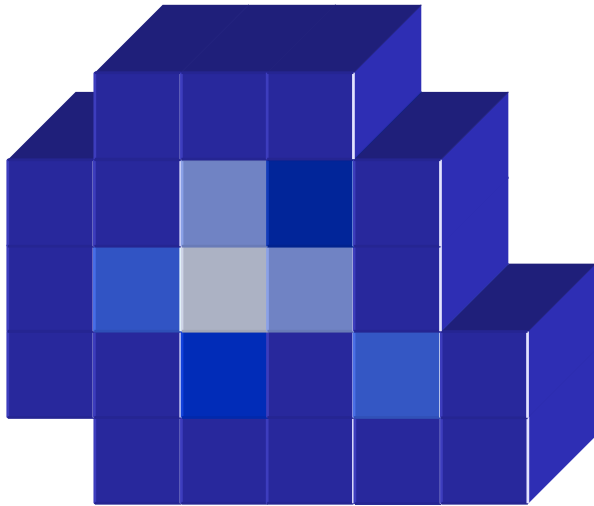


- S'applique à des détecteurs composés de N ($\gg 1$) voies
- Partant des données brutes, la reconstruction réalise la fonction du sous-système:
 - Chambre à trace \rightarrow traces
 - Detector Cherenkov \rightarrow angles Cherenkov
 - Etc.
- Reconstruction = interprétation des données pour reconnaître la/les trace(s) du passage des particules
 - Recherche de « patterns »
- Se nourrie de données:
 - De placement des voies
 - Câblages des voies
 - De la réponse de ces voies
 - Qui l'objet d'un monitoring en ligne

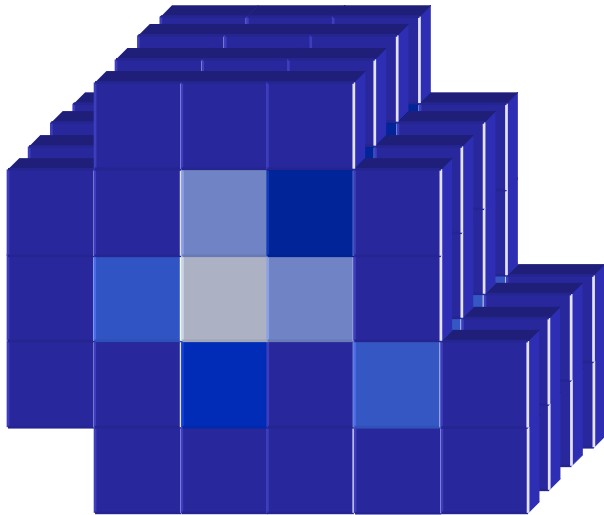


Détecteur à pixels

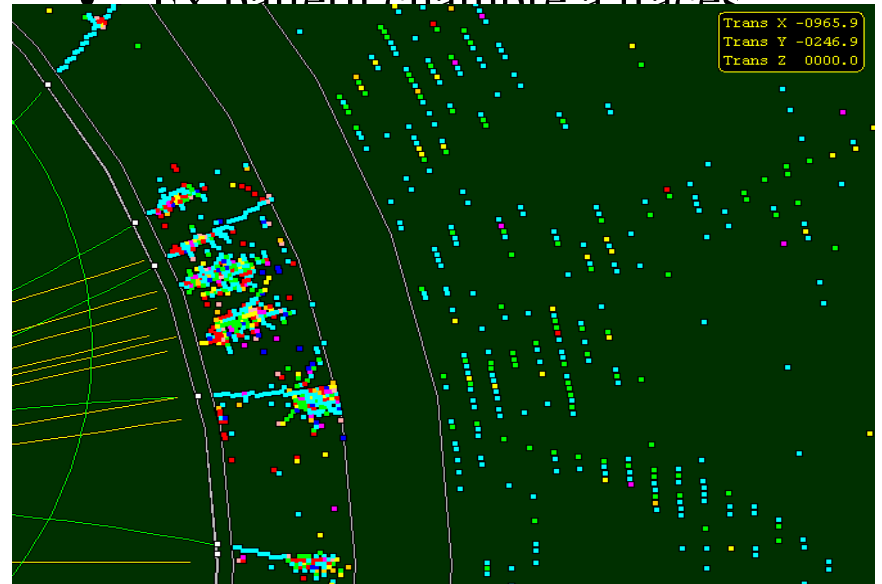
Reconstruction,
Recherche de patterns



- Ex pattern cluster

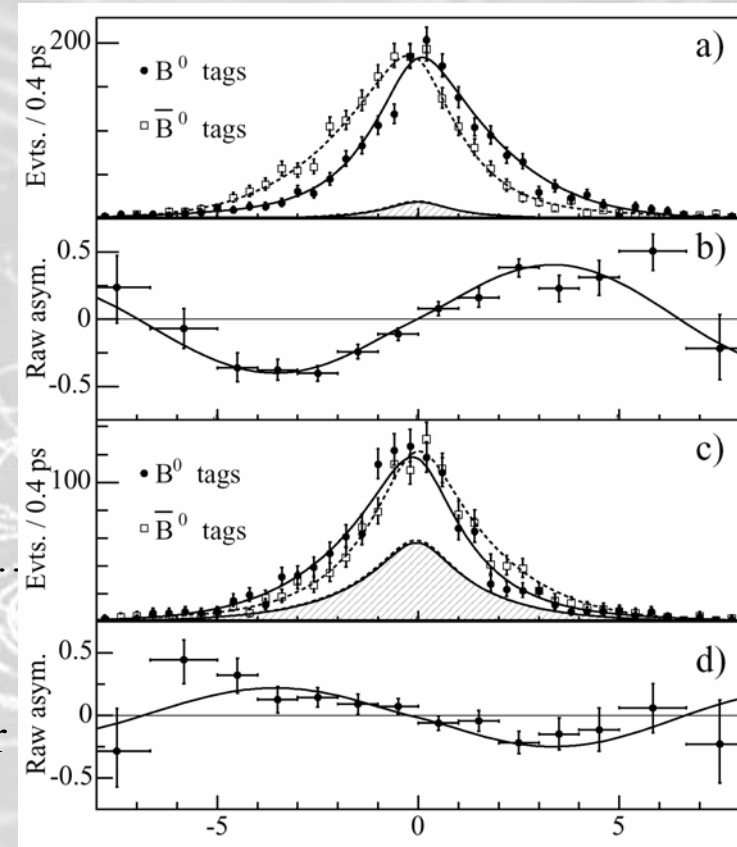


- Ex pattern chambre à traces



Analyse de physique

- Mesure de physique, ou « analyse », exploite certaines désintégrations
 - connues pour leur sensibilité aux paramètres théoriques que l'on veut mesurer
- L'analyse de physique consiste à
 - sélectionner ces désintégrations
 - modéliser les données observées
 - pour prendre en compte les imperfections expérimentales (inefficacités, résolutions, bruits...)
 - et de là réaliser l'ajustement des paramètres théoriques, c'est-à-dire, faire la mesure
 - faire l'autocritique du modèle d'ajustement pour estimer les incertitudes systématiques.
- L'analyse démarre des objets produits par la reconstruction
- Et elle s'appuie sur deux outils généraux
 - la simulation
 - l'ajustement.





– TROISIÈME PARTIE –

LA SIMULATION

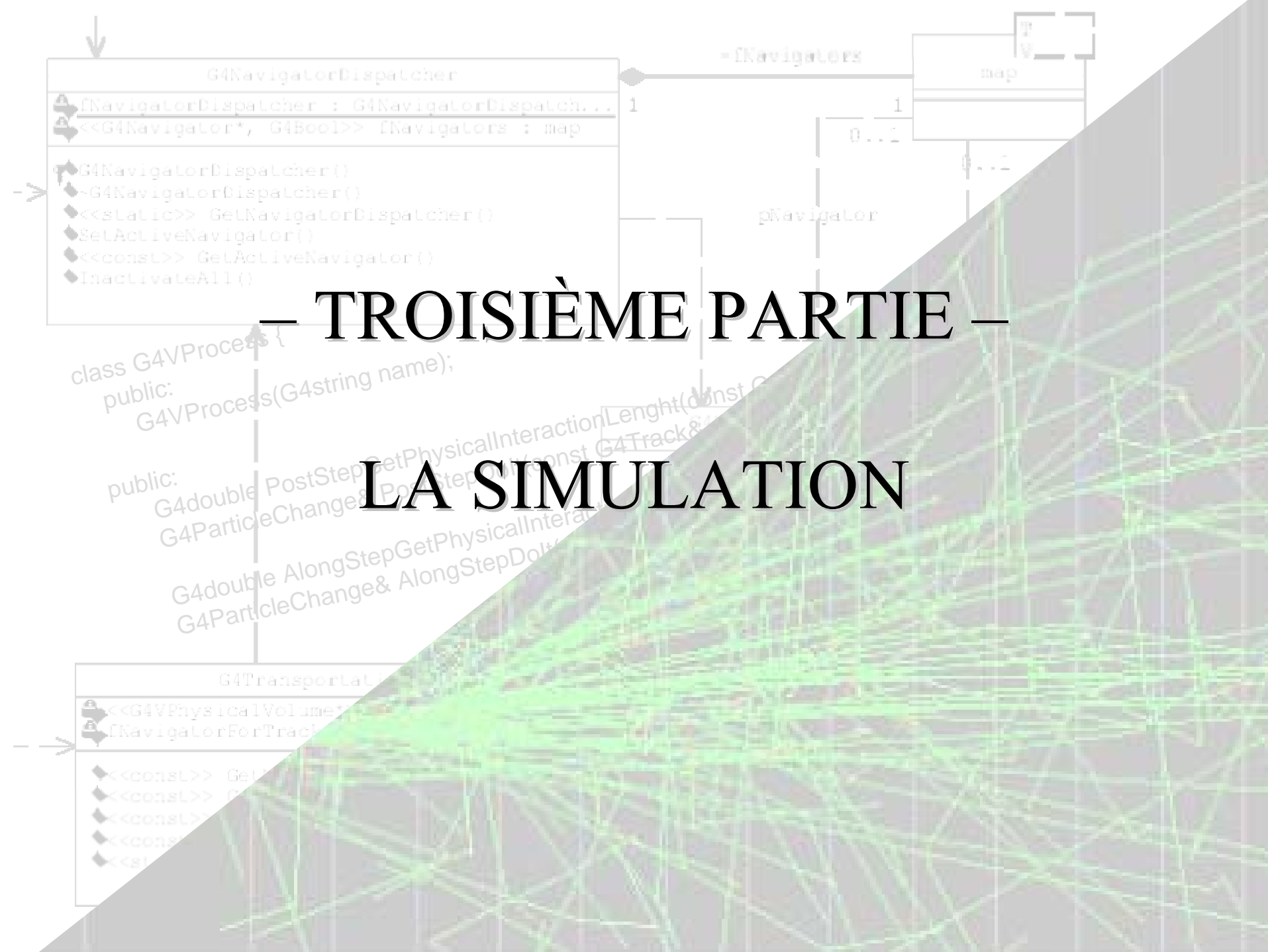
```

class G4VProcess {
public:
  G4VProcess(G4string name);

public:
  G4double PostStepGetPhysicalInteractionLength(const G4Track& track);
  G4ParticleChange* PostStepDoInteraction(const G4Track& track);
  G4double AlongStepGetPhysicalInteractionLength(const G4Track& track);
  G4ParticleChange& AlongStepDoInteraction(const G4Track& track);
};
  
```

```

class G4Transportation : public G4VProcess {
public:
  G4Transportation(G4string name);
  G4Transportation(G4string name, G4int n);
  G4Transportation(G4string name, G4int n, G4int n);
  G4Transportation(G4string name, G4int n, G4int n, G4int n);
  G4Transportation(G4string name, G4int n, G4int n, G4int n, G4int n);
};
  
```



Qu'est-ce que la simulation ?

- La simulation est une réalité virtuelle
 - Captation et modélisation d'une partie de la réalité
 - Exemple: simulateur de vol

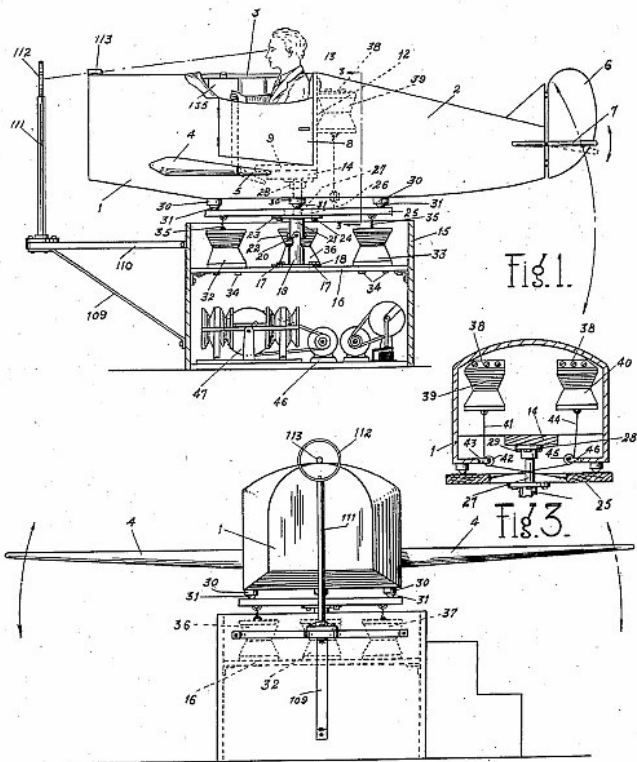
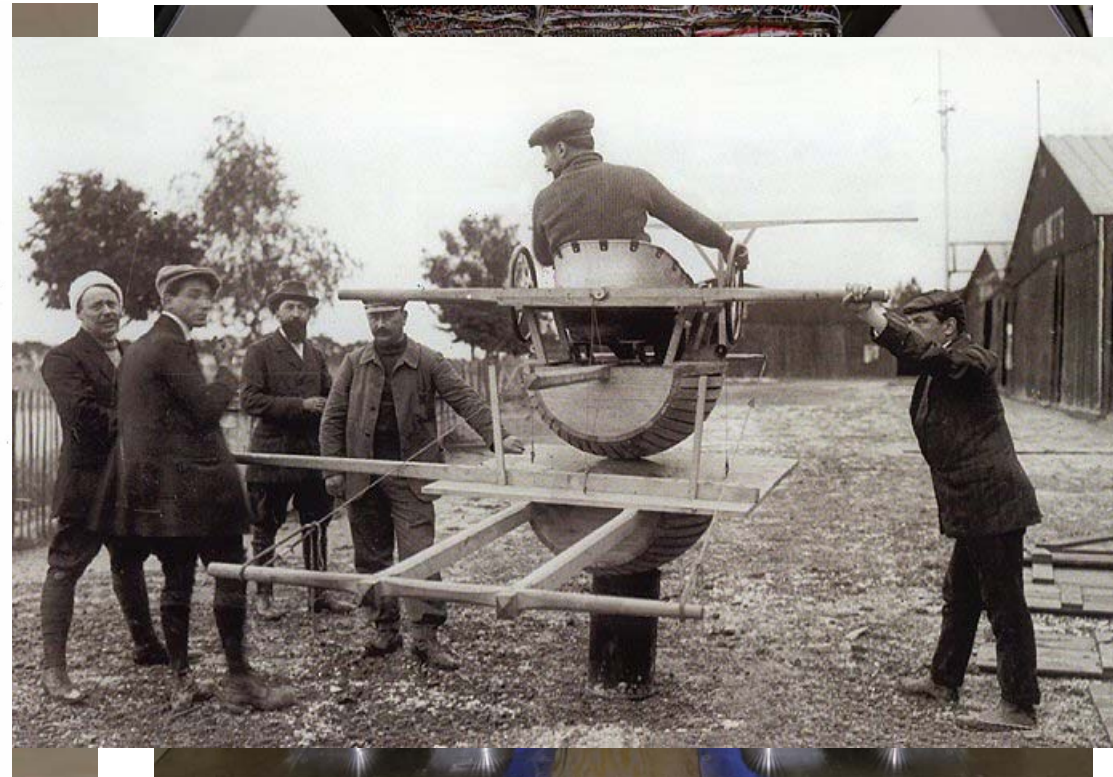


Fig.2.

INVENTOR
EDWIN A. LINK, JR.



NASA RFD simulator
NASA Langley Research Center

3/20/1997

Image # EL-1997-00146

Intérêt de la simulation ?

- Permet de se faire la main pour
 - beaucoup moins cher,
 - beaucoup plus rapidement,
 - (beaucoup moins dangereux,)
 - ...,qu'en vrai.
- Et en contrôlant et sachant tout ce qui se passe
 - Quelle cause a quel(s) effet(s)
 - Pas à pas, inclure davantage de réalisme

L'intérêt de la simulation pour nous ? (1)

- Les détecteurs sont
 - Complexes
 - (chers)
 - Longs à concevoir
 - Longs à mettre en œuvre
- Les évènements à traiter sont
 - Complexes
 - Variés
 - Et ceux qui nous intéressent sont souvent (très) rares
 - Perdus dans un flot inintéressant

L'intérêt de la simulation pour nous ? (2)

-avant la construction de l'expérience-

- On a besoin d'être capable de prédire et d'optimiser les performances générales des détecteurs
 - Résolution en énergie
 - Performance d'identification des particules
 - ...
 - Mais aussi performance des triggers
- On fait des (pré-)études de physique
 - Avec les performances estimées, arrive t'on à réaliser les mesures de physique qui motivent la construction de l'expérience ?
 - Si non, quelles performances doivent être améliorées ?

On utilise ici la simulation comme un outil de CAO

L'intérêt de la simulation pour nous ? (3)

-après la construction de l'expérience-

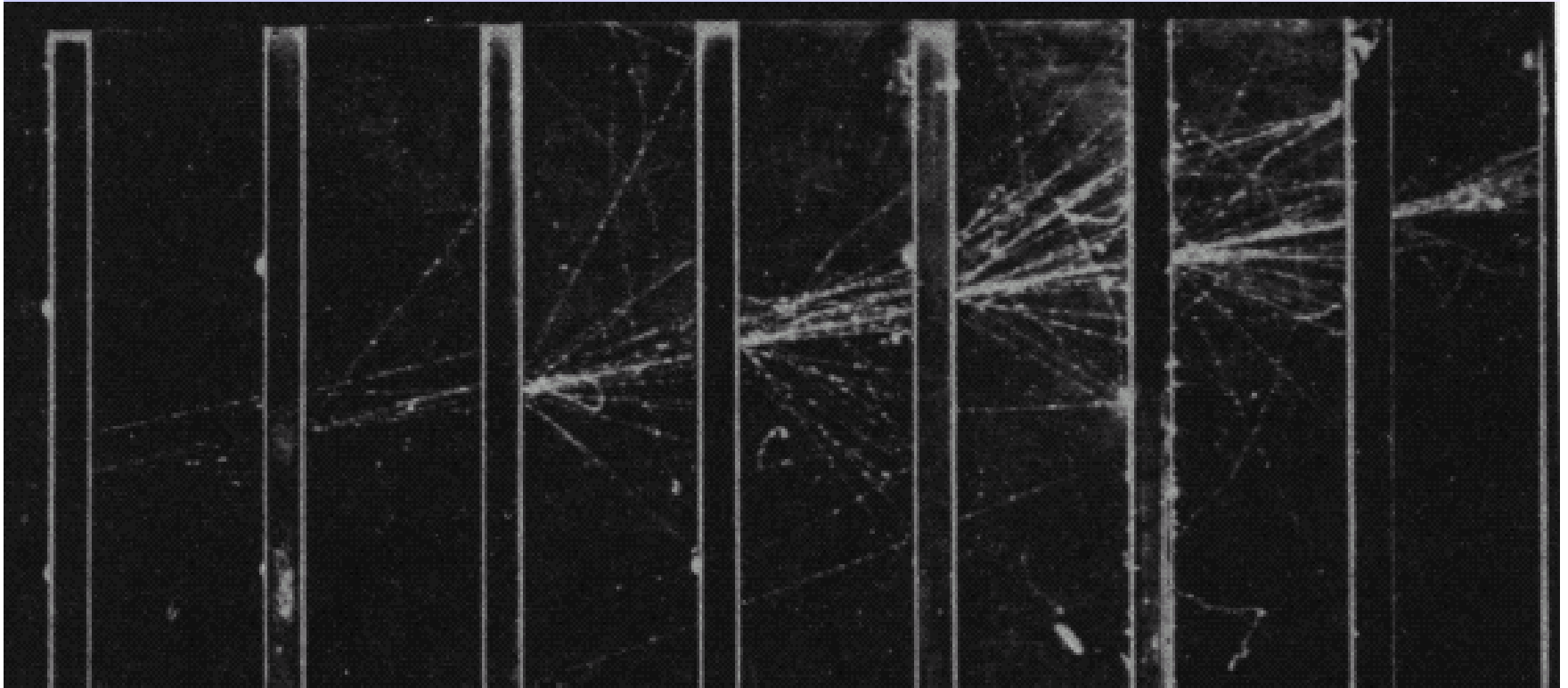
- On a besoin de modéliser les données prises pour
 - Etablir/affiner la stratégie de sélection du signal:
 - Production d'évènements de signal
 - Production d'évènements de bruit
 - Permet de mettre en évidence les différences entre les deux, dans le but de les exploiter dans la sélection.
 - Etudier ce qui pourrait fausser la mesure et en tenir compte:
 - Effets expérimentaux sur le signal
 - Perte d'évènements
 - Effets de résolution (énergie, position, etc...)
 - Composition des bruits passant la sélection:
 - Quels sont les évènements de bruit qui passent la sélection, comment se répartissent ils sur les grandeurs à mesurer, comment en tenir compte pour ne pas fausser la mesure ?

La simulation est ici utilisée comme un outil d'aide à l'analyse de physique

De quelle simulation parle t'on ?

- Celle des interactions de particules dans la matière:
 - Les accélérateurs créent des particules de grande masse ou de haute énergie
 - Qui peuvent traverser ensuite tout ou partie des détecteurs ou se désintégrer, conduisant à d'autres particules de hautes énergies, etc...
- La façon dont ces particules vont se comporter dans le détecteur est ce que l'on a besoin de savoir
 - Avec en particulier les dépôts d'énergie dans les matériaux
 - Ce qui conditionne la réponse des détecteurs
- Ce comportement est régi par les interactions particules-matière
 - Il nous faut donc modéliser ces interactions

On aimerait être capable de reproduire
une image comme celle-ci...



Les étapes d'une simulation

- On crée / calcule le type de désintégration qui nous intéresse
 - Exemples :
 - $e^+e^- \rightarrow Y(4S) \rightarrow B^0 \text{ anti-}B^0$
 - suivi de $B^0 \rightarrow J/\Psi K_S^0$
 - puis de $J/\Psi \rightarrow e^+e^-$ ou $\mu^+\mu^-$ et $K_S^0 \rightarrow \pi^+\pi^-$
 - $\text{gluon gluon} \rightarrow \text{Higgs} \rightarrow \gamma\gamma$
 - En terminant par des particules « standards »
- Ces produits de désintégration sont propagés dans une représentation virtuelle du détecteur
 - En calculant les interactions particules-matières
 - Et les dépôts d'énergie dans la matière
- L'énergie déposée dans les volumes de détection est utilisée pour fabriquer des données semblables aux données réelles
 - A partir de ces dépôts, on va par exemple calculer la collection de lumière dans un cristal, la charge collectée dans un gaz, etc..
 - Et calculer la réponse de la voie d'électronique correspondante

Les étapes d'une simulation

- On crée / calcule le type de désintégration qui nous intéresse

- On parle de « **générateurs** »:
 - Logiciels spécialisés dans le calcul de certaines désintégrations
 - Ils « savent » prendre en compte les paramètres théoriques que l'on veut mesurer
 - Dépend de l'expérience et/ou de la mesure
- Exemples de générateurs: JETSET, EvtGen, TAULO, etc...

- Ces produits de désintégration sont propagés dans une

- C'est la phase dite de « **tracking** »:
 - Elle se base sur des outils généraux, comme Geant4, qui permettent:
 - Et les détecteurs: géométrie, matériaux, etc...

- L'énergie déposée dans les plans de détection est utilisée pour fabriquer des données semblables aux données réelles

- C'est la phase dite de « **digitisation** »:
 - Qui dépend des appareillages spécifiques de l'expérience
 - Et calculer la réponse de la voie d'électronique correspondante

Un logiciel de simulation : **Geant 4**

- Geant4 (« GEometry ANd Tracking ») est un logiciel de simulation des interactions particules-matières, à vocation « généraliste »:
 - Utilisé en physique des hautes énergies, physique des ions lourds, astro-particules
 - Mais aussi dans des domaines de basses énergies : médical (radiothérapie, imagerie), spatial (vols habités, radiation sur électronique), astrophysique...
- Geant4 est le successeur de GEANT3, écrit en FORTRAN
 - GEANT3 étant alors le standard de la physique des hautes énergies

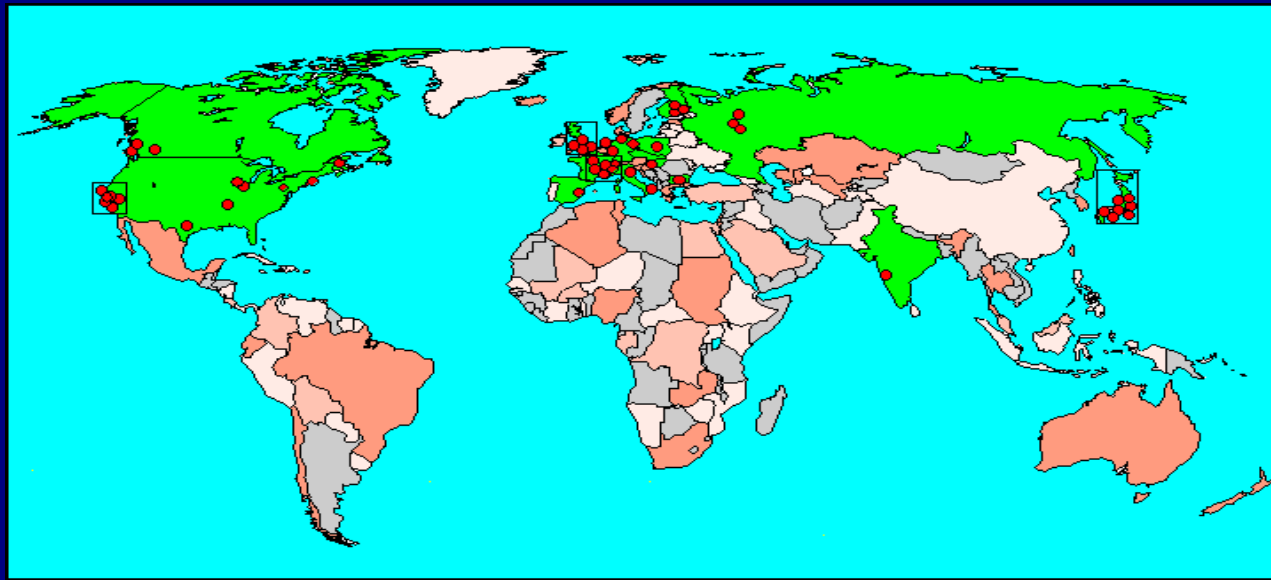
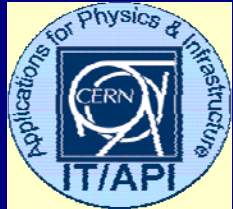
Bref historique de Geant4

- Geant4 a démarré en 1994, sous forme du projet « RD44 »
 - Afin d'étudier la faisabilité et l'intérêt d'un logiciel HEP basé sur des technologies orientées objet.
- Cette étude étant principalement motivée pour palier des limitations de GEANT3
 - Par exemple, l'adjonction d'une nouvelle forme géométrique ou d'un nouveau processus physique impliquait de modifier plusieurs dizaines de routines !
 - Limitations qui peuvent être facilement levées avec une conception orientée objet.
 - Grâce à la notion « d'abstraction »
- Décembre 1998 : première version publique et formation de la collaboration Geant4 afin d'assurer les développements et la maintenance du logiciel.

La Collaboration Geant4

HARP

Stanford
Linear
Accelerator
Center



Univ. Barcelona

PPARC

Lebedev

Helsinki Inst. Ph.

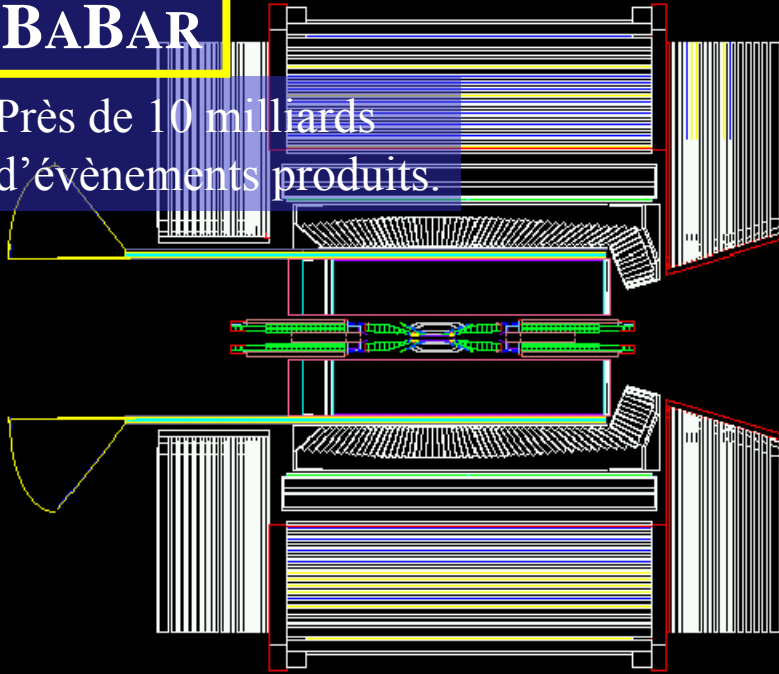


Collaborateurs en individuel
en provenance aussi de
Budker Inst. of Physics
IHEP Protvino
MEPHI Moscow
Pittsburg University

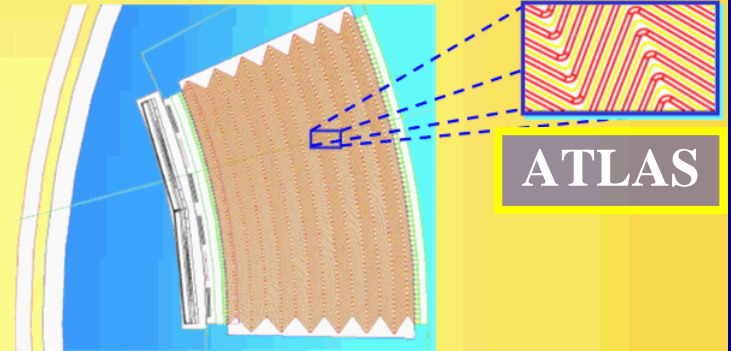
Exemples d'applications (1/5)

BABAR

Près de 10 milliards
d'évènements produits.

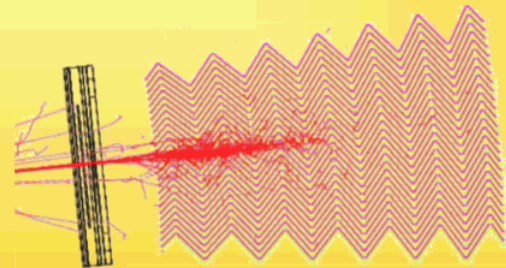


 Electromagnetic Barrel Accordion Calorimeter

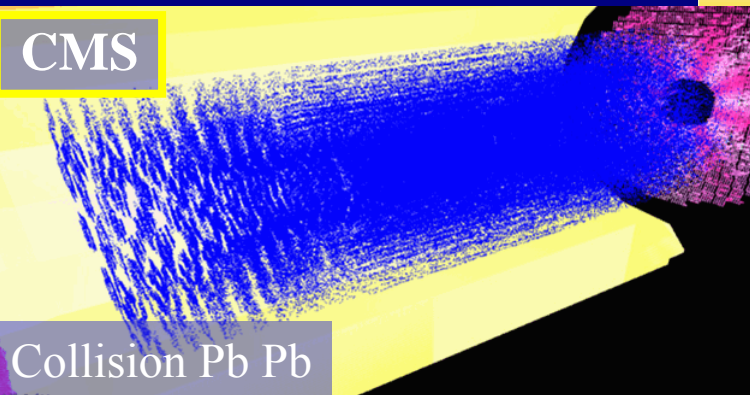


ATLAS

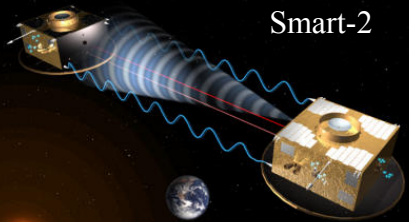
 10 GeV Electron Shower



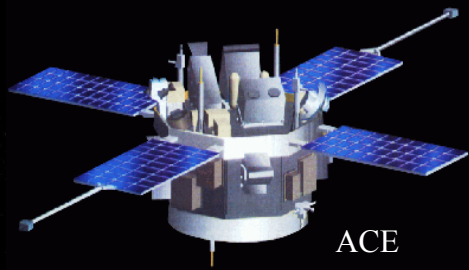
CMS



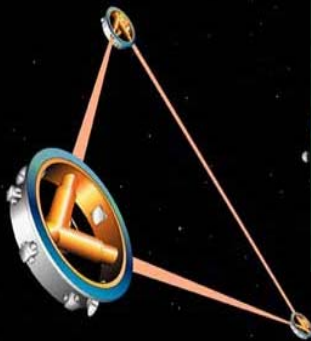
Collision Pb Pb



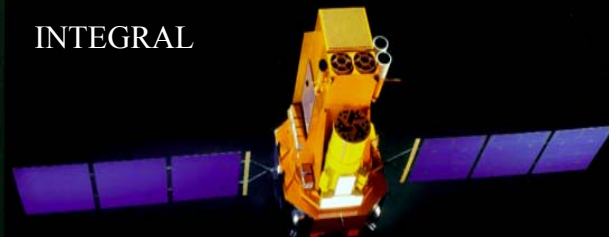
Smart-2



ACE



LISA

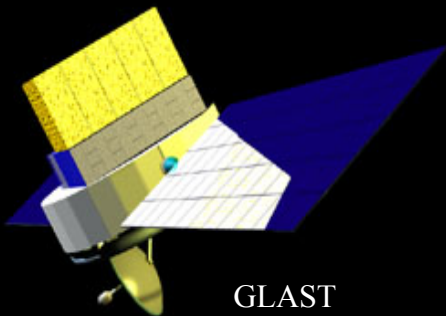


INTEGRAL

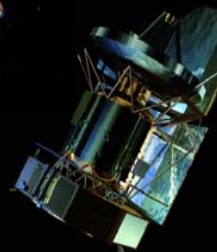
Cassini



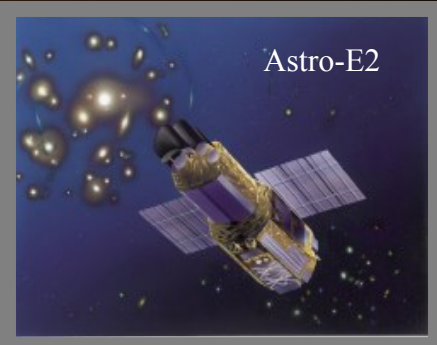
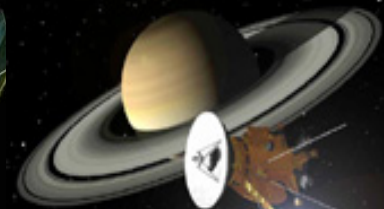
Bepi Colombo



GLAST



Herschel



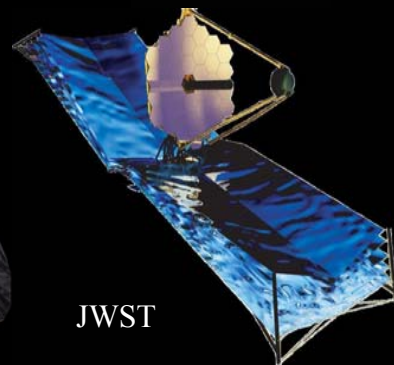
Astro-E2



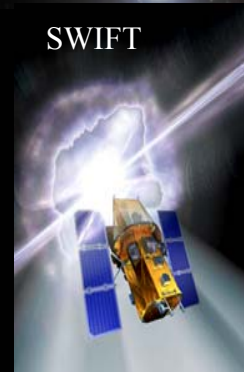
XMM-Newton



GAIA



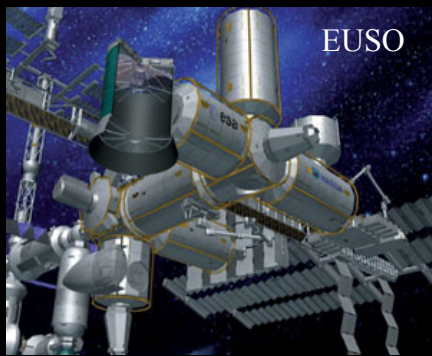
JWST



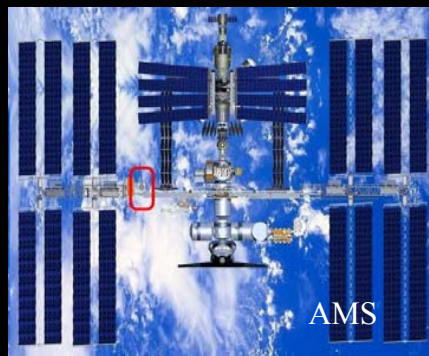
SWIFT



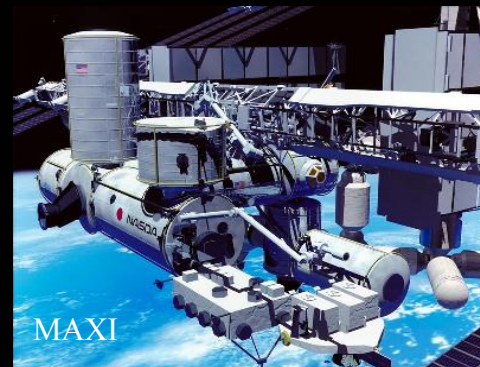
ISS Columbus



EUSO

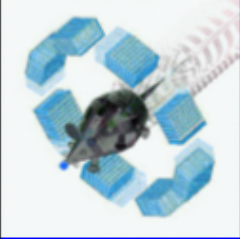


AMS

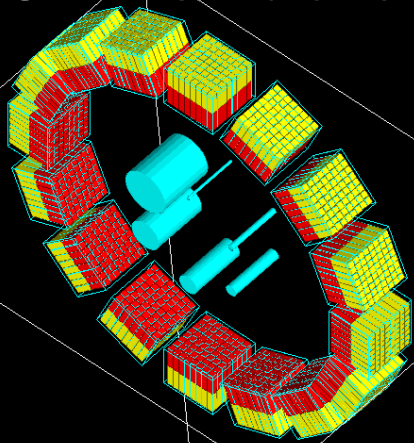


MAXI

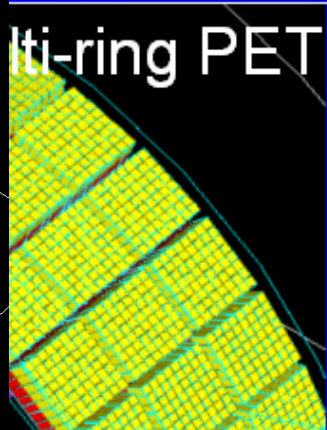
Geometry examples of GATE applications



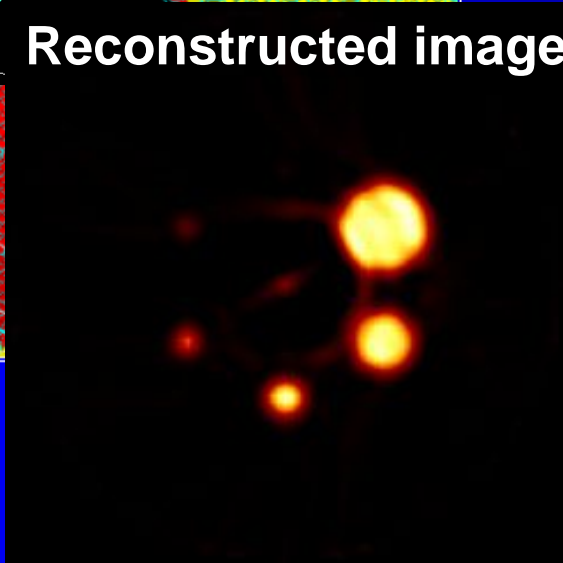
GATE simulation



Multi-ring PET

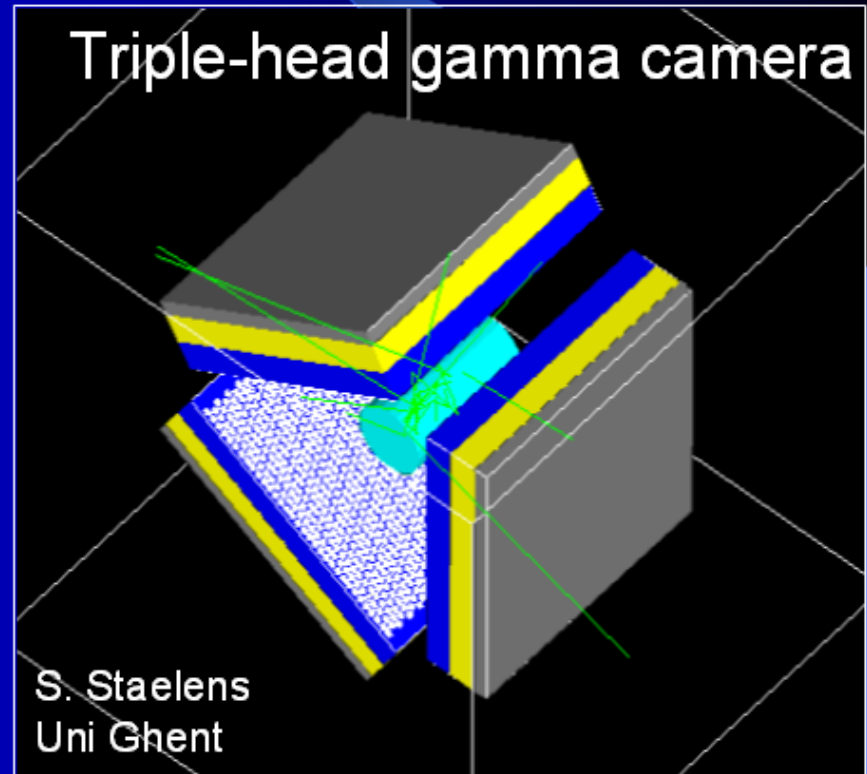


Reconstructed image



D. Strul
IPHE Lausanne

Triple-head gamma camera



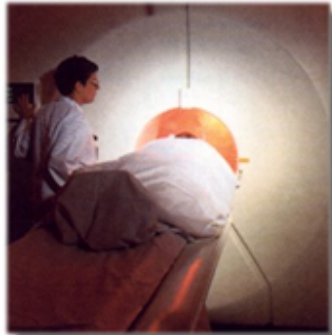
S. Staelens
Uni Ghent



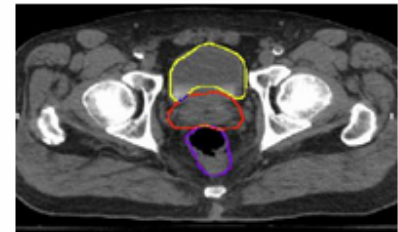
RECHERCHE

Pavillon L'Hôtel-Dieu
Centre hospitalier universitaire de Québec

Geant4 DICOM Interface



Modeling complex structures



file

Reproduce patient's anatomy in a Geant4 application

```
.....DICOM.....UL.....  
.....OB.....UI.....1.2.840.10008.  
5.1.4.1.1.2.....UI.....1.3.12.2.1107.5.  
1.4.40238.5.0.14178572411846.....UI.....  
1.2.840.10008.1.2.1.....UI.....1.3.12.2  
1107.5.1.4.....SH.....SIEMENS_S5VA30A  
.....CS.....ISO_IR 100.....CS* ORIGINAL/P  
RIMARY/AXIAL/CT_SOM5 SPI.....UI.....1.2.  
840.10008.5.1.4.1.1.2.....UI.....1.3.12  
2.1107.5.1.4.40238.5.0.141785724118  
.....DA.....20020109.....  
.....TM.....094310.18600  
987.....3.TM.....0943  
.....CS.....GT.....p.LO.....  
L-DIEU DE QUEBEC  
IS. QUEBEC. QUEBEC  
SH.....HNCT40238 ..
```



UNIVERSITÉ
LAVAL

Authors: L. Archambault, L. Beaulieu, V.-H. Tremblay
(Univ. Laval and l'Hôtel-Dieu, Québec)

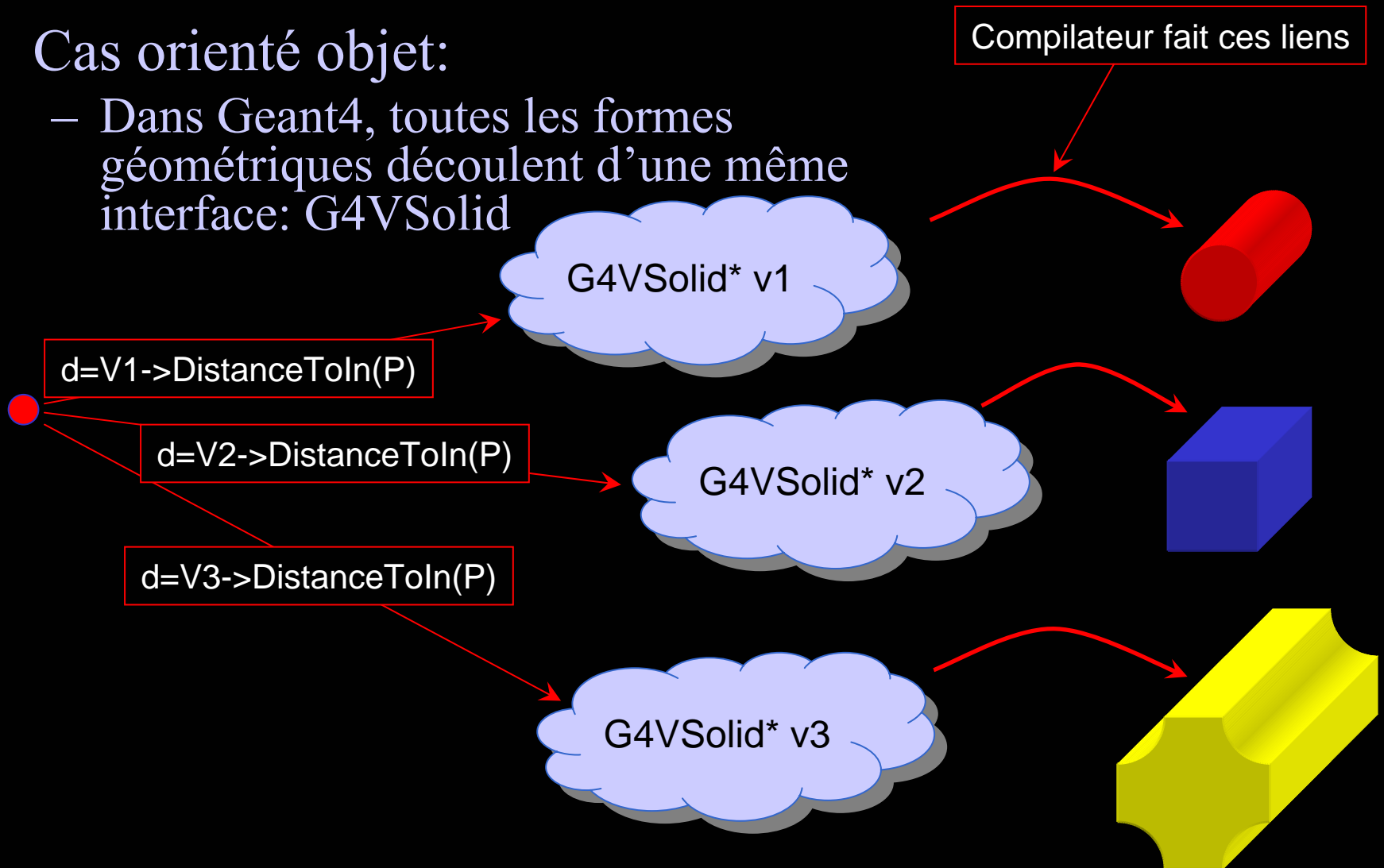
Ce que permet l'abstraction

- L'abstraction permet de concevoir et d'implémenter un algorithme, en résonnant sur des types abstraits:
 - Forme géométrique
 - Processus physique
- On obtient alors un « cœur » de logiciel générique
- L'ajout d'un nouveau processus ou d'une nouvelle forme géométrique ne nécessite pas de modification du cœur.
- L'abstraction est ce qui a permis à Geant4 de s'ouvrir à d'autres applications que les hautes énergies.

Comment cela marche ?

- Cas orienté objet:

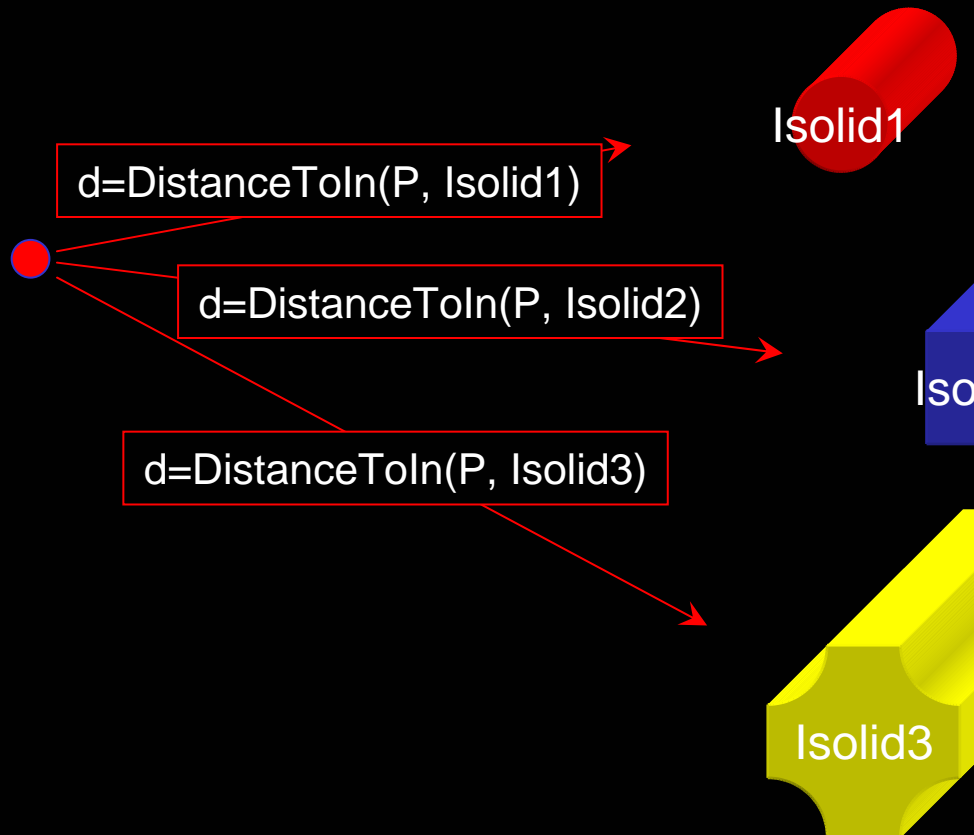
- Dans Geant4, toutes les formes géométriques découlent d'une même interface: G4VSolid



Comment ça joue ?

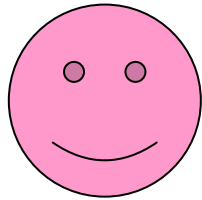
- Cas non-objet:

Développeur doit faire les liens



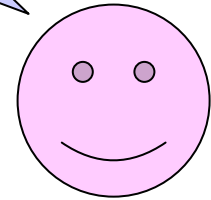
```
real function DistanceToIn(P, Isolid)
real P(3)
real d
integer Isolid

if (Isolid.eq.iBox) then
  d = distanceForBox(P, dir)
else if (Isolid.eq.iCylinder) then
  d = distanceForCylinder(P, dir)
else if (Isolid.eq.iCroixArrondie) then
  d = distanceForCroixArrondie(P, dir)
else if ...
  ...
end
```

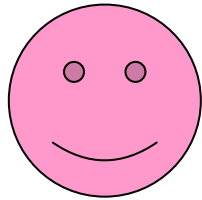


C++

Je peux TOUT faire avec
le FORTRAN !!

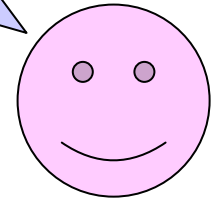


FORTRAN

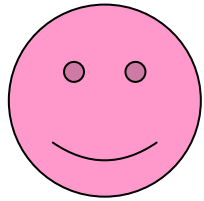


C++

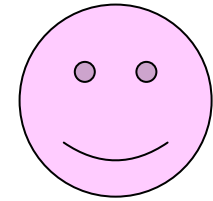
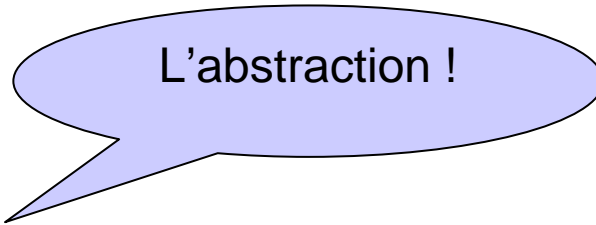
Que peut-on faire en C++
et qu'on ne peut pas faire
en FORTRAN !!?



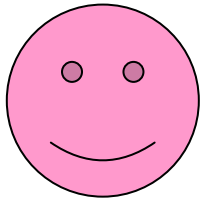
FORTRAN



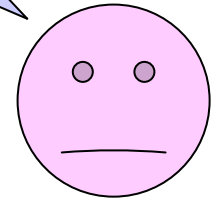
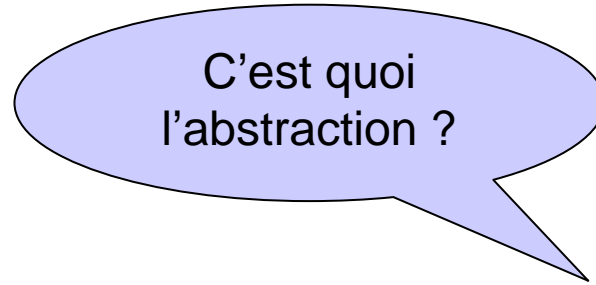
C++



FORTRAN

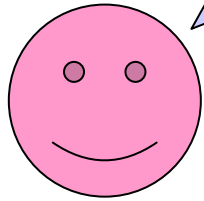


C++



FORTRAN

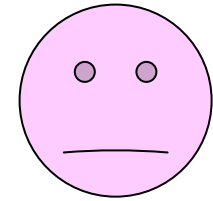
C'est quoi
l'abstraction ?



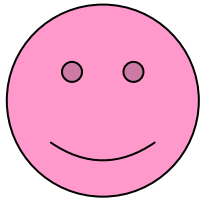
C++



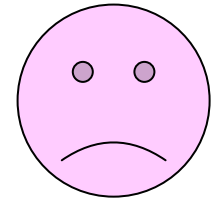
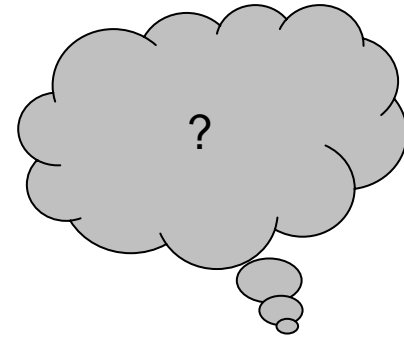
L'abstraction c'est
... blabla... blabla...
Ça permet
...blabla.. blabla...



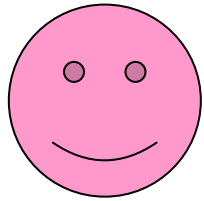
FORTRAN



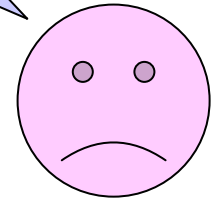
C++



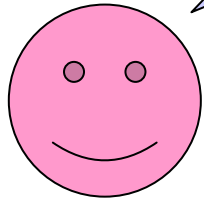
FORTRAN



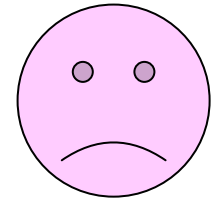
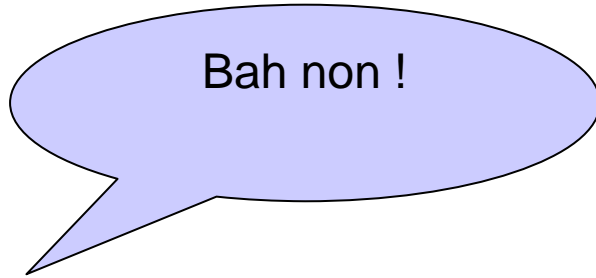
C++



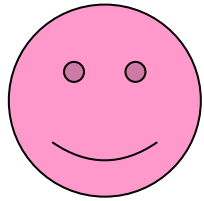
FORTRAN



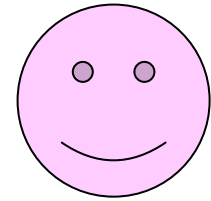
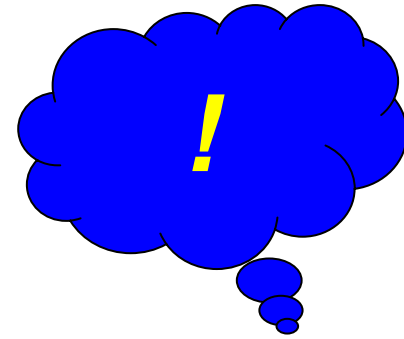
C++



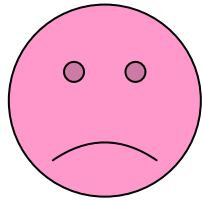
FORTRAN



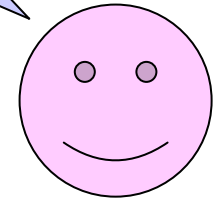
C++



FORTRAN



C++



FORTRAN

Deux classes abstraites clefs dans Geant4

- **G4VSolid:**
 - Interface pour toutes les formes géométriques
 - Boite, cylindre, etc...
 - Un solide doit savoir répondre à
 - Un point (x,y,z) est-il sur sa surface ?
 - À l'intérieur, à l'extérieur
 - À quelle distance de sa surface suivant une direction donnée
 - Etc...
 - Et les classes concrètes (boite, cylindre...) implémentent les réponses à ces questions
- **G4VProcess:**
 - Interface pour tous les processus physiques
 - Permet, au sein d'un même processus, d'implémenter toutes combinaisons d'actions
 - Au repos
 - Désintégration, annihilation...
 - Continue
 - Ionisation...
 - Ponctuelle, en vol
 - Désintégration, rayonnement, δ -ray, interaction particule-noyau...

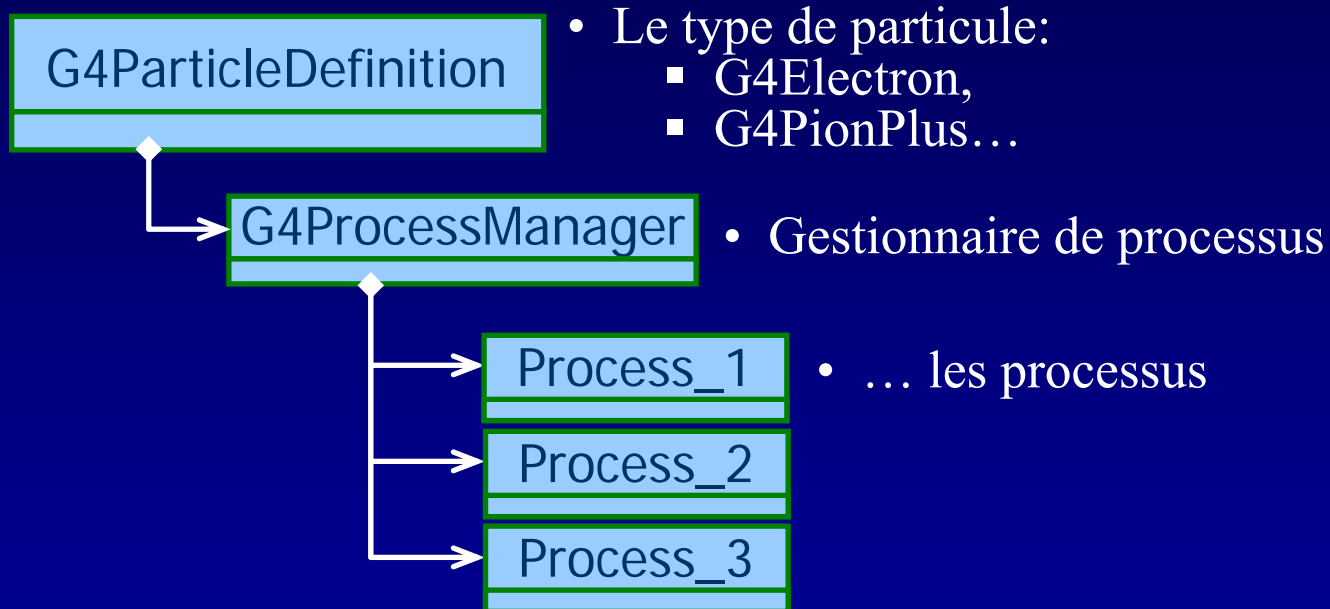
Comment marche un processus ?

- On utilise la technique dite du « Monté Carlo »
- Les processus physiques sont de nature aléatoire
 - Probabilité de conversion d'un photon en paire d'électron
 - Probabilité d'émission d'un photon par un électron passant dans la matière
- Les processus physiques implémentent les « sections efficaces »
 - Qui décrivent les probabilités au-dessus
 - Tenant compte des paramètres physiques (énergie, densité, etc...)
- Un processus physique « tire » aléatoirement le nombre de « longueurs d'interactions », ou la durée, avant qu'il n'agisse sur la particule
 - Par exemple au bout de combien de temps le processus de désintégration va agir
- Puis on met ces processus ensembles, à agir sur la particule, avec un arbitre: le « stepping » (on en reparlera)
- Si un processus intervient, il détermine alors le sort de la particule:
 - Sa perte d'énergie
 - Son éventuelle disparition
 - L'éventuelle création de particules secondaires
- Tout cela, sous le contrôle du stepping

Particules et processus

- Les particules sont définies par leurs propriétés statiques telles que masse, charge, spin, etc...
- *Elles ne sont, en elles-mêmes, sensibles à aucune physique !*
- Elle acquièrent cette sensibilité physique *via* les processus physiques qu'on leur associe
 - Et pour un électron, on associera l'ionisation, le bremsstrahlung, le processus de transport
 - Pour un positron, on associera en plus l'annihilation

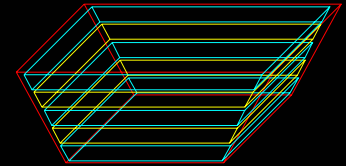
Sensibilité à la physique



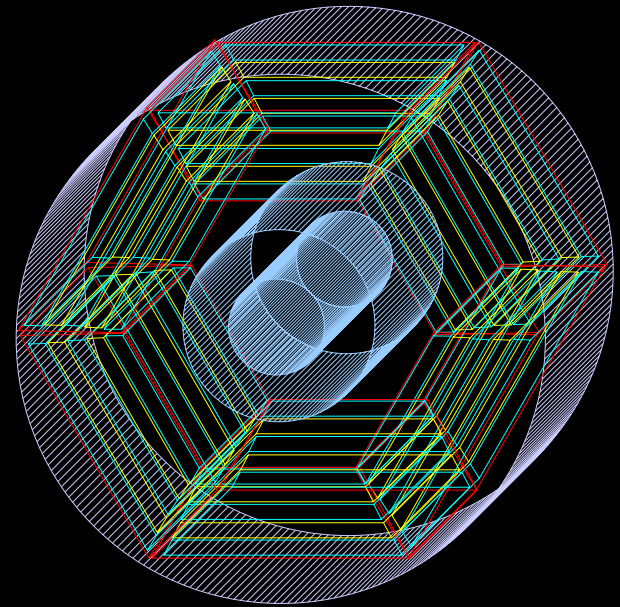
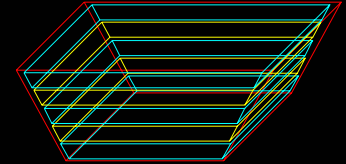
Volumes / géométrie

- Volumes décrits par:
 - Forme
 - Matériel
 - Dont les propriétés sont utilisées dans le calcul des interactions
 - Sous-structure
 - Ensemble de ces trois caractéristiques constitue un « **volume logique** »
 - Placement par rapport au volume « mère »
 - Le volume devient un « **volume physique** »
- Il existe un volume particulier, le « world » volume, qui définit le contenant de toute la géométrie
 - Et toute particule sortant de ce volume est supprimée de la suite de la simulation.
- Un volume peut enfin être déclaré « sensible »
 - C'est un volume correspondant à un volume de lecture d'une voie d'électronique
 - Et dans lequel le ou les dépôts d'énergie seront donc utilisés pour calculer la réponse de cette voie.

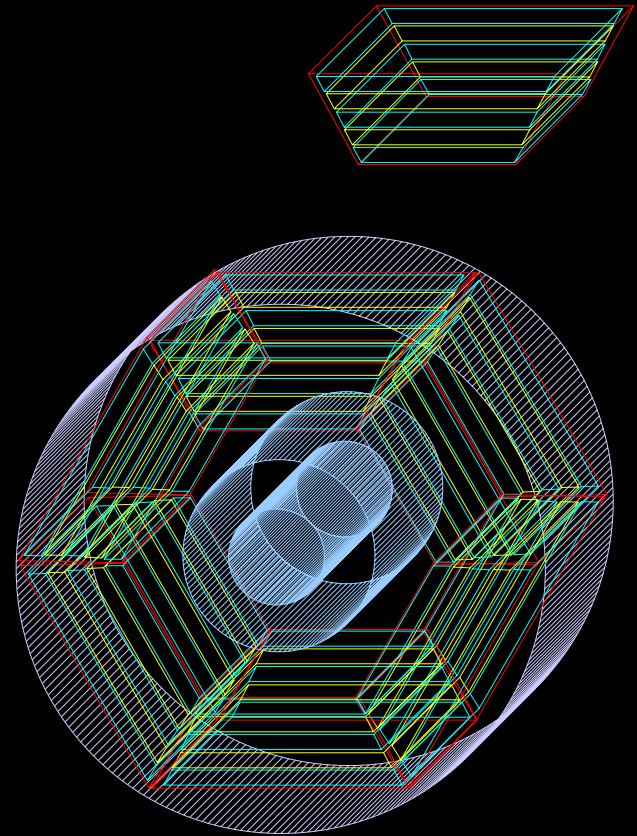
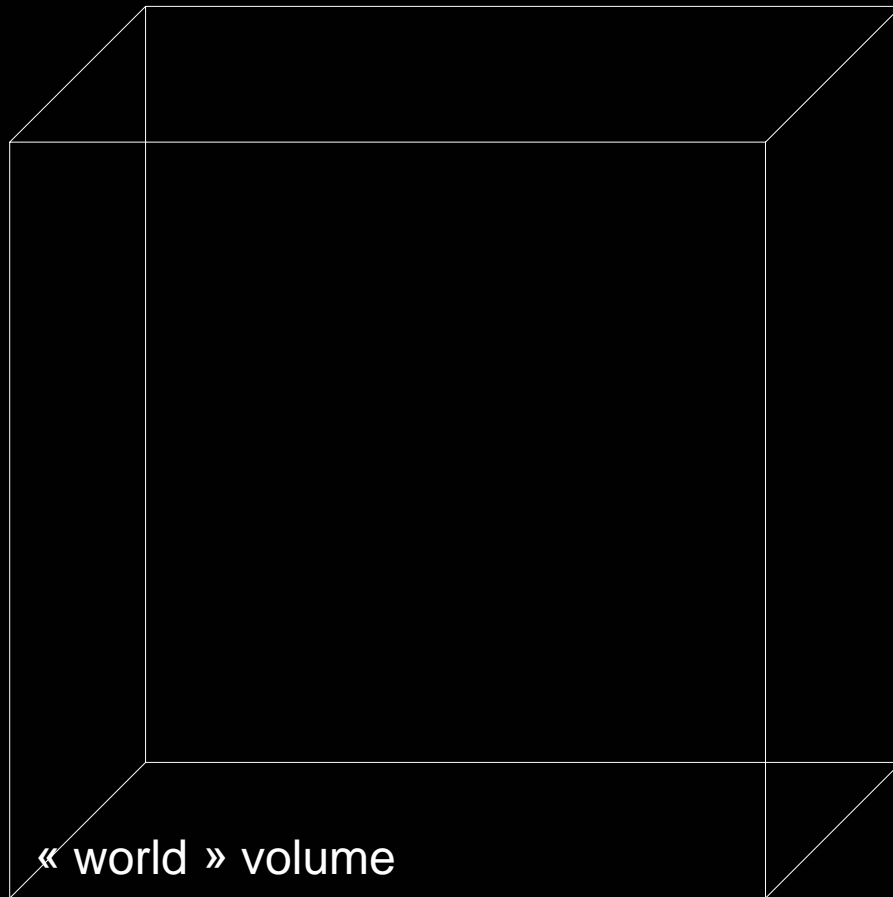
Construction d'une géométrie



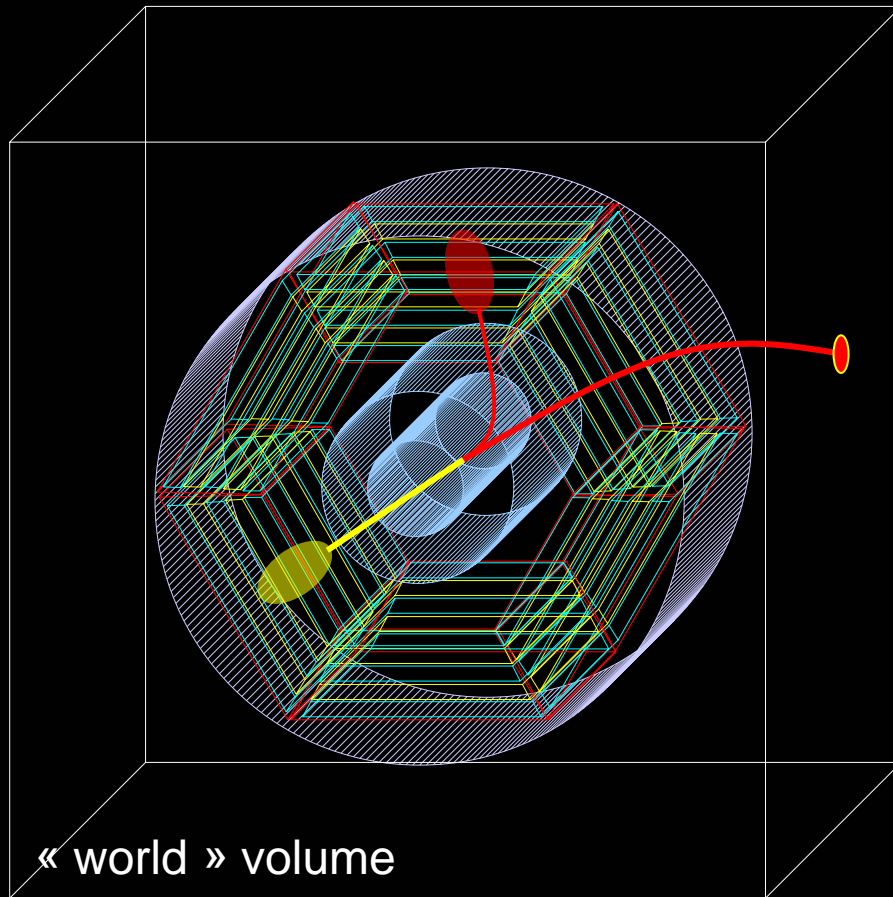
Construction d'une géométrie



Construction d'une géométrie



Construction d'une géométrie

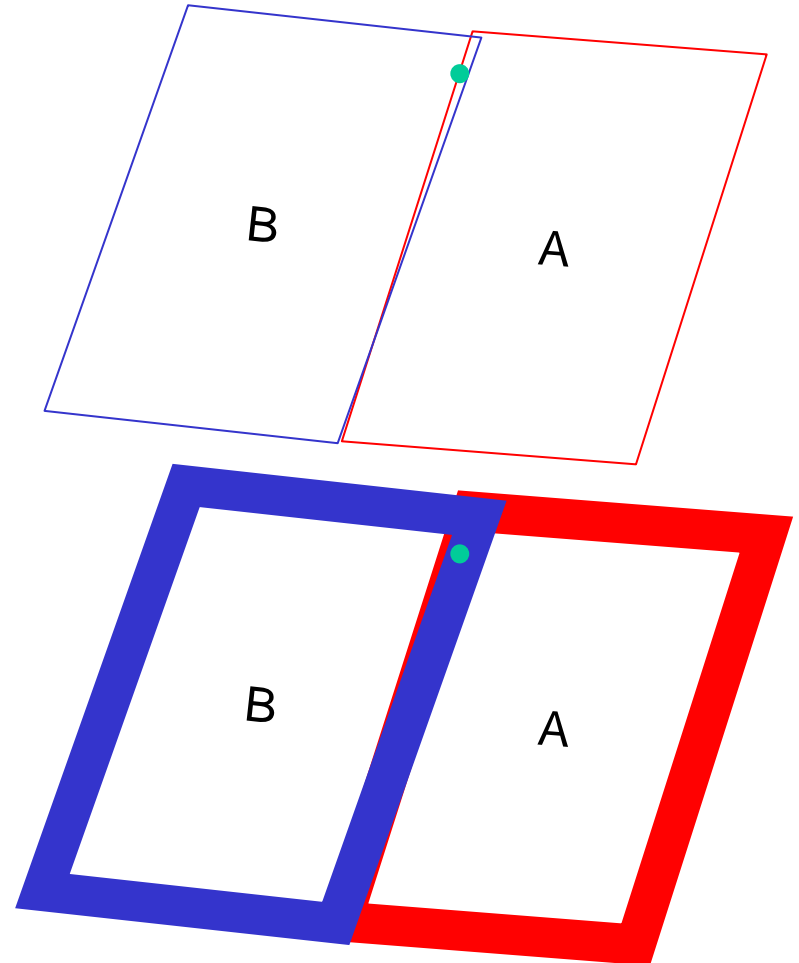


Concrètement...

```
//  
// World  
//  
solidWorld = new G4Box("World", //its name  
                      WorldSizeX/2,WorldSizeYZ/2,WorldSizeYZ/2); //its size  
  
logicWorld = new G4LogicalVolume(solidWorld, //its solid  
                                defaultMaterial, //its material  
                                "World"); //its name  
  
physiWorld = new G4PVPlacement(0, //no rotation  
                              G4ThreeVector(), //at (0,0,0)  
                              logicWorld, //its logical volume  
                              "World", //its name  
                              0, //its mother volume  
                              false, //no boolean operation  
                              0); //copy number  
  
//  
// Calorimeter  
//  
solidCalor=0; logicCalor=0; physiCalor=0;  
solidLayer=0; logicLayer=0; physiLayer=0;  
  
if (CalorThickness > 0.)  
{ solidCalor = new G4Box("Calorimeter", //its name  
                        CalorThickness/2,CalorSizeYZ/2,CalorSizeYZ/2); //size  
  
  logicCalor = new G4LogicalVolume(solidCalor, //its solid  
                                  defaultMaterial, //its material  
                                  "Calorimeter"); //its name  
  
  physiCalor = new G4PVPlacement(0, //no rotation  
                                G4ThreeVector(), //at (0,0,0)  
                                logicCalor, //its logical volume  
                                "Calorimeter", //its name  
                                logicWorld, //its mother volume  
                                false, //no boolean operation  
                                0); //copy number
```

Volumes et tolérance numérique

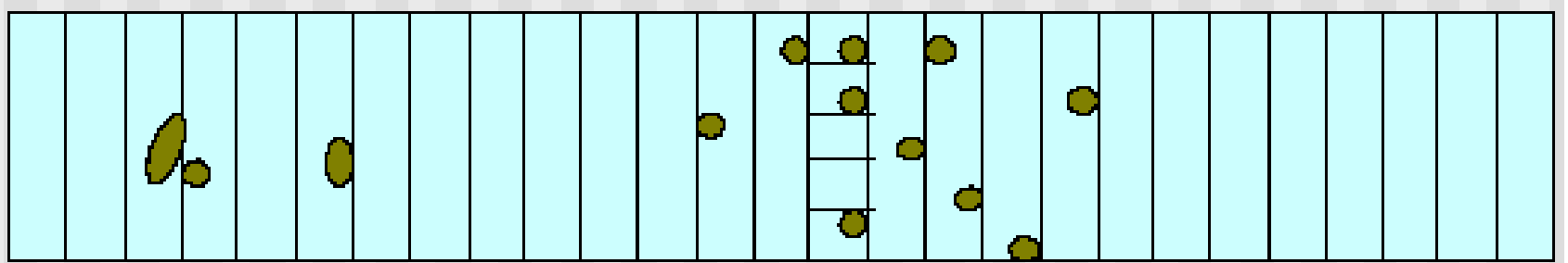
- Une complication est due aux arrondis numériques, qui peuvent créer des confusions
 - Dans quel volume est ce point ?
- Pour contrôler ce problème, Geant4 introduit une « épaisseur » aux surfaces
 - Cette épaisseur étant supérieure aux arrondis numériques
 - Permet de dire qu'un point est sur la surface
 - Et de prendre des précautions



Volumes et navigation

- La navigation consiste à savoir répondre à la question:
 - Soit un point $P(x,y,z)$, dans quel volume est ce point P ?
 - Afin notamment de connaître le matériau courant
 - Il faut avoir conscience que la géométrie « n'existe pas »
 - On donne en réalité une suite de définitions de volumes et de relations entre ces volumes, basées sur des conventions
 - Le « navigateur » est plus proprement dit un « interpréteur » de géométrie
- Les performances de la navigation deviennent critiques dès lors que l'on considère de grands nombres de volumes
 - Exemple : ATLAS ~ 10 millions de volumes !
- La navigation de Geant4 repose sur la technique de « voxélisation »
 - La géométrie est découpée en une grille 3D, chacune des cellules de cette grille étant occupée par un petit nombre de volumes

À quoi ressemblent ces voxels ?

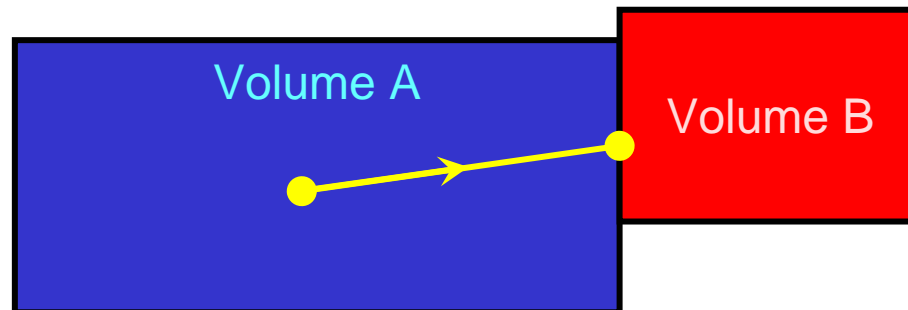


Le cœur de la simulation: le stepping

- Geant4 fait progresser les particules par « pas »
 - « pas » = « step » en anglais
 - ...cette progression s'appelle donc le « stepping »
- C'est lors de cette progression qu'intervient le calcul des processus physiques, en se basant
 - sur les propriétés de la particule en cours (particule primaire)
 - du milieu courant
 - et qui aboutit :
 - au calcul de l'énergie déposée
 - et à la création d'éventuelles particules secondaires
 - et qui modifie donc l'état de la primaire
 - Énergie, direction, etc...
- L'ensemble de ces opérations de stepping constitue le « tracking »

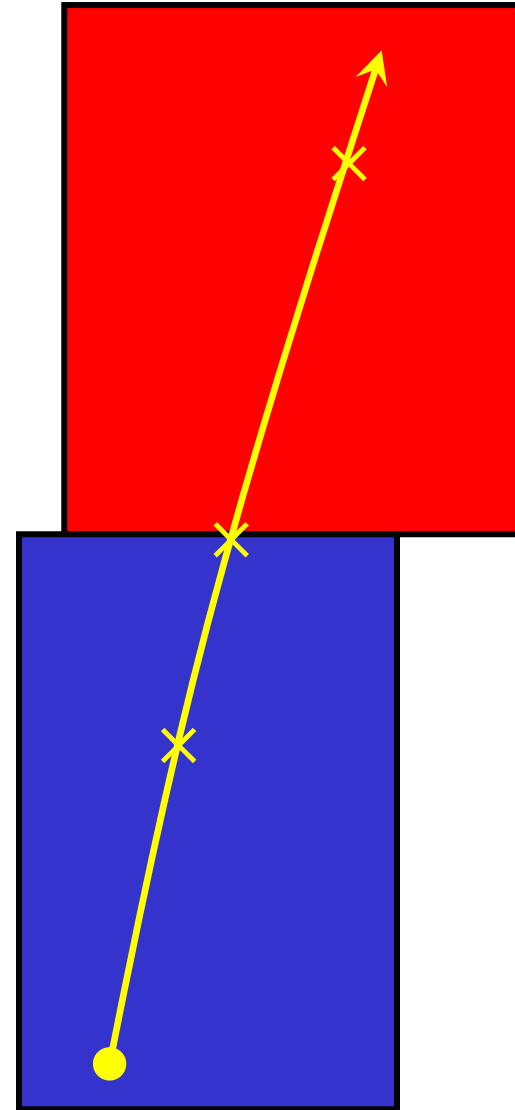
Un step se déroule dans un seul et même volume

- Ceci afin de garantir un seul et même matériau tout au long du step:
 - Le même que celui utilisé pour faire les calculs physiques au point de départ du step
- C'est le processus de transport qui assure cette fonction
 - C'est aussi lui qui déplace la particule, en mettant à jour ses coordonnées
- Si une frontière de volume doit être franchie, le processus de transport demande à arrêter le step sur la surface du volume.



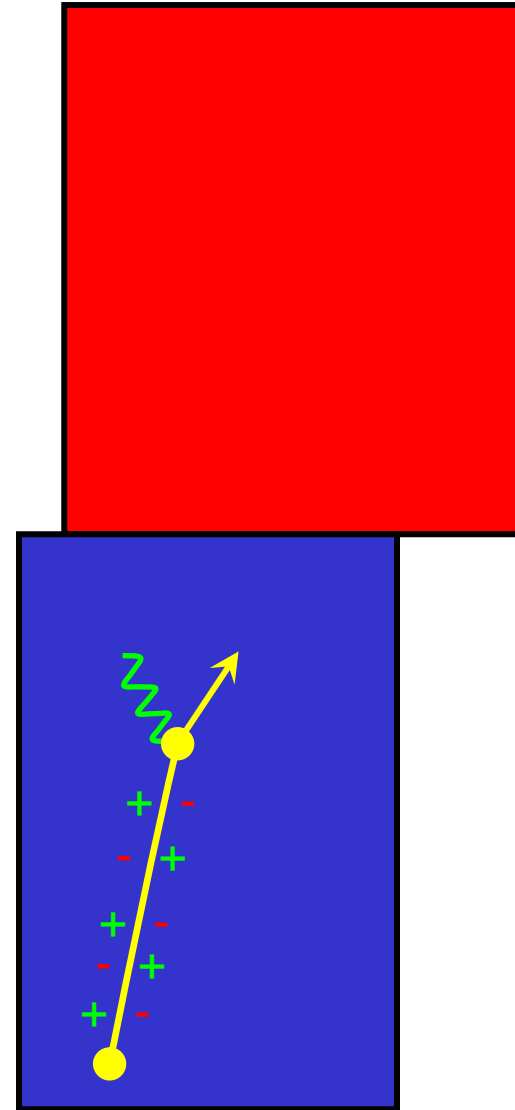
Histoire d'un step

- Détermination de la longueur du step:
 - Pour ce faire, le stepping demande à chaque processus quelle longueur de vol il autorise pour la particule
 - Par exemple, s'il s'agit d'un positron, on peut avoir:
 - le processus « annihilation » répond 22.1 cm
 - Le bremstrahlung répond 7.3 cm
 - Le processus de transport (géométrie) répond 12.1 cm
 - Dans ce cas la longueur de step sera 7.3 cm



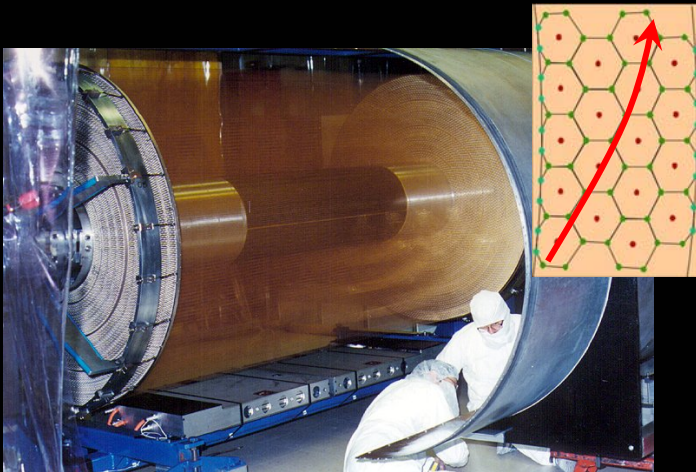
Histoire d'un step

- Détermination de la longueur du step:
 - Pour ce faire, le stepping demande à chaque processus quelle longueur de vol il autorise pour la particule
 - Par exemple, s'il s'agit d'un positron, on peut avoir:
 - le processus « annihilation » répond 22.1 cm
 - Le bremstrahlung répond 7.3 cm
 - Le processus de transport (géométrie) répond 12.1 cm
 - Dans ce cas la longueur de step sera 7.3 cm
- Application des processus physiques:
 - En deux temps:
 - D'abord les contributions « continues »
 - comme l'ionisation
 - et calcul de l'énergie déposée dans la matière
 - Puis les processus ponctuels:
 - Ici, le bremstrahlung: émission d'un photon
 - Dans ce cas, pas de dépôt supplémentaire d'énergie
- Enfin, si le volume dans lequel s'est déroulé le step est sensible : appel au code utilisateur pour traiter le dépôt d'énergie



Que fait-on des dépôts dans les volumes « sensibles » ?

- Ces dépôts sont le point de départ pour simuler la réponse du détecteur
- Chambres à traces :
 - On s'intéresse aux particules individuellement
 - Chaque dépôt est converti en une réponse de voie d'électronique
- Calorimètres:
 - Il y a développement de gerbe
 - Cascade de secondaires dans un cristal par exemple
 - C'est la somme des dépôts qui conduit à une réponse de voie d'électronique



- Ce sont de grandes lignes, chaque expérience a ensuite ses variantes
 - Par ex: considérant les dépôts segmentant un cristal le long de son axe.
- Ces réponses de voie d'électronique sont alors traitées par la reconstruction.

Histoire d'un évènement

- Au début de l'évènement, le générateur fournit les particules « primaires » à tracer
 - Photons issus de la désintégration du Higgs,
 - Particules filles de la désintégration d'un méson B
 - Etc...
- Chacune de ces particules est tracée jusqu'à
 - Disparaître par interaction
 - Ou sortir du « world » volume
 - Lors de son tracking, cette particule crée des particules secondaires.
- Ces particules secondaires sont à leur tour tracées, et ainsi de suite, jusqu'à disparition complète de toutes les particules.
- Reste alors les dépôts d'énergie dans les volumes « sensibles »
 - Et toutes informations que l'utilisateur a bien voulu conserver
 - Comme les relations mère \rightarrow {filles}

Les cuts

- Les « cuts » sont un sujet récurrent pour qui fait une simulation.
 - Ils sont la manifestation d'une limitation de la modélisation.
- Qu'est-ce que ces « cuts » ? (en soit le mot « cut » est un peu malheureux)
 - C'est qu'on ne peut pas descendre à des particules d'énergie arbitrairement basse.
 - Cela en ferait trop !
 - On est obligé de décider d'une limite: ce sont les cuts qui définissent cela
 - Dans Geant4, ces cuts sont un seuil de production en énergie
 - Ils sont utilisés pour empêcher des particules de trop basse énergie de naître.
 - (Dans GEANT3 les cuts sont production et tracking)
- Et pour certains processus, on n'a pas le choix
 - Ils présentent des divergences infra-rouges:
 - Ils génèrent des nombres infinis / très élevés de particules de plus en plus basse énergie.
 - Ces processus sont le bremstrahlung et l'ionisation/production de δ -rays
- Mais ces secondaires auraient emporté une énergie ? Qu'en fait-on ?
 - L'énergie correspondante est comptabilisée comme contribution continue.
- Il s'agit de trouver alors un compromis entre vitesse et précision.
 - Cela dépend complètement de l'expérience, du sous-détecteur, etc...

Mise en œuvre de la simulation

- Cela consiste, pour un utilisateur a
 - Décrire sa géométrie de détecteur
 - Volumes, matériaux, volumes « sensibles »
 - Décrire la physique mise en jeu:
 - Sélectionner les processus physiques d'intérêt pour sa simulation
 - Domaine en énergie
 - Adapter les « cuts » : seuil de production
 - Sélectionner les particules présentes
 - Raccorder son générateur d'évènements à la simulation

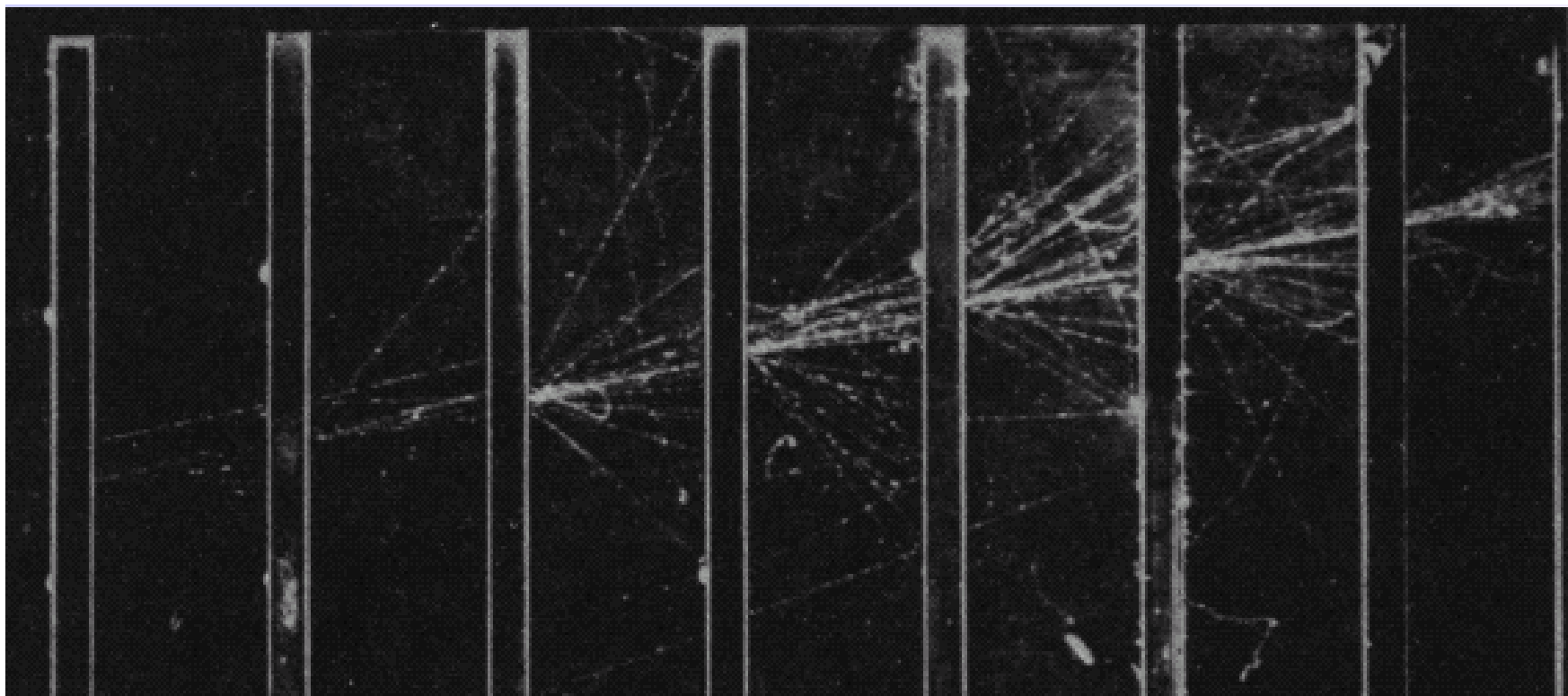
Démos ?

VISUALISATION

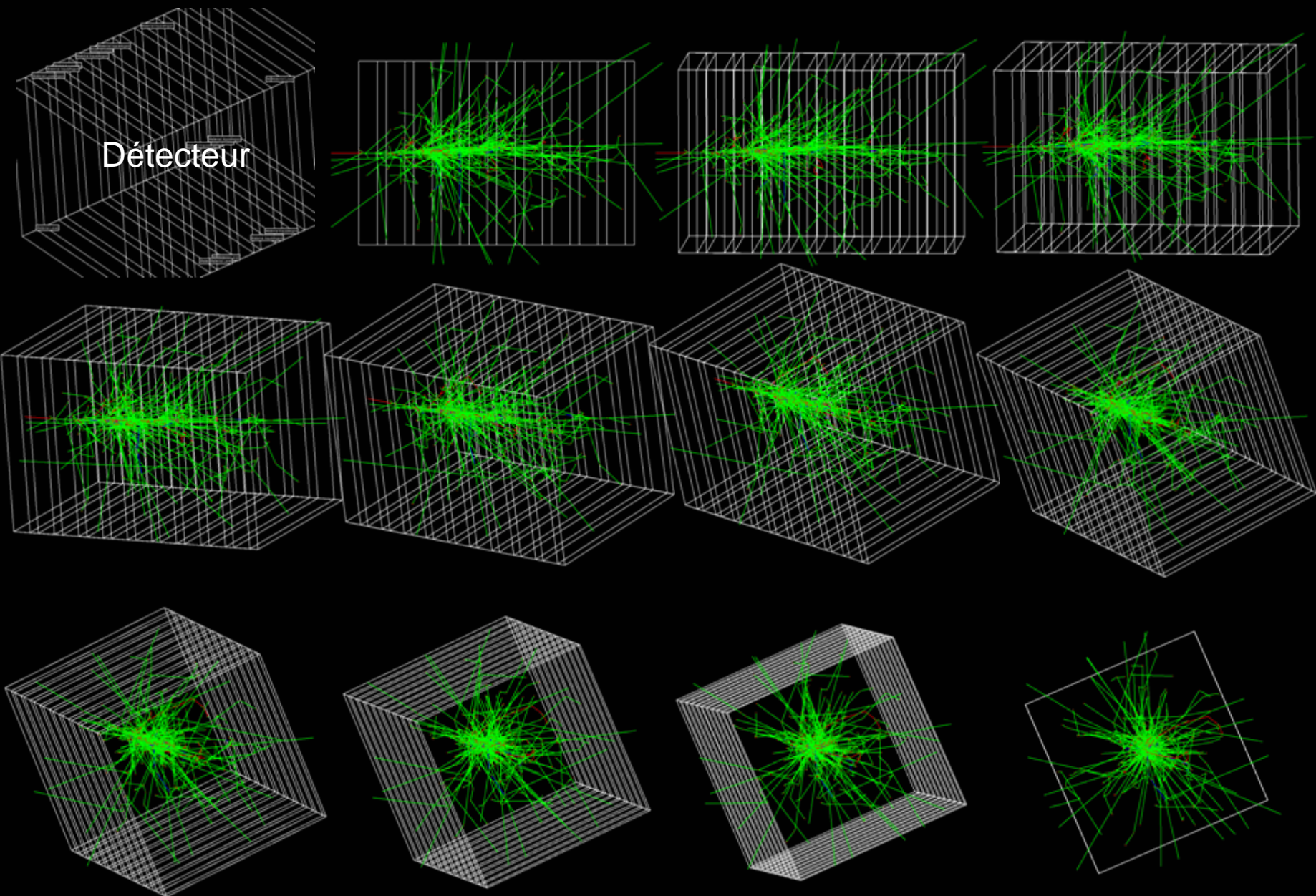
Électron incident de 1 GeV dans un calorimètre simple à échantillonnage.

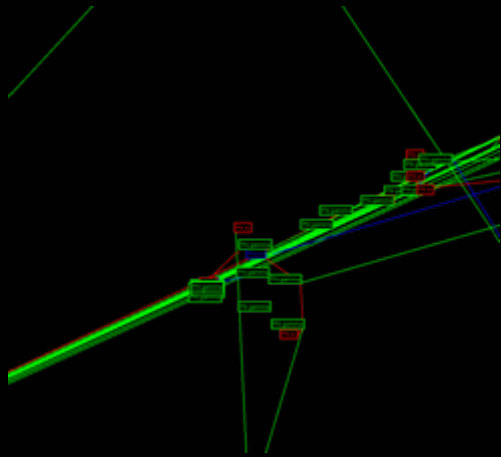
Images obtenues avec « wired »:

<http://www.slac.stanford.edu/~wiredces>

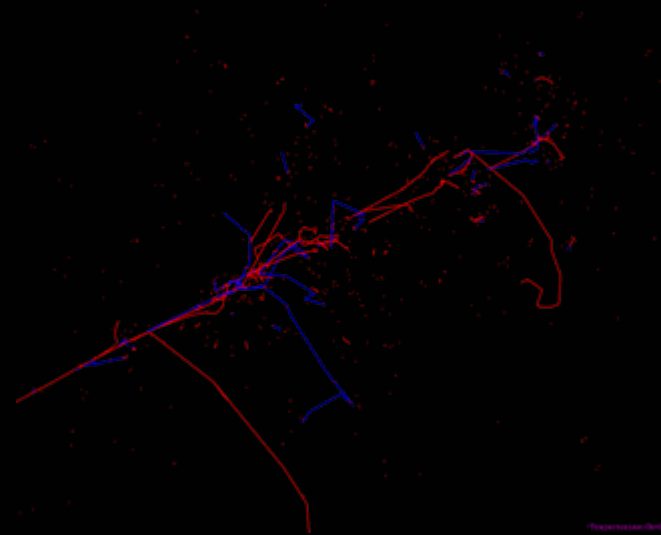
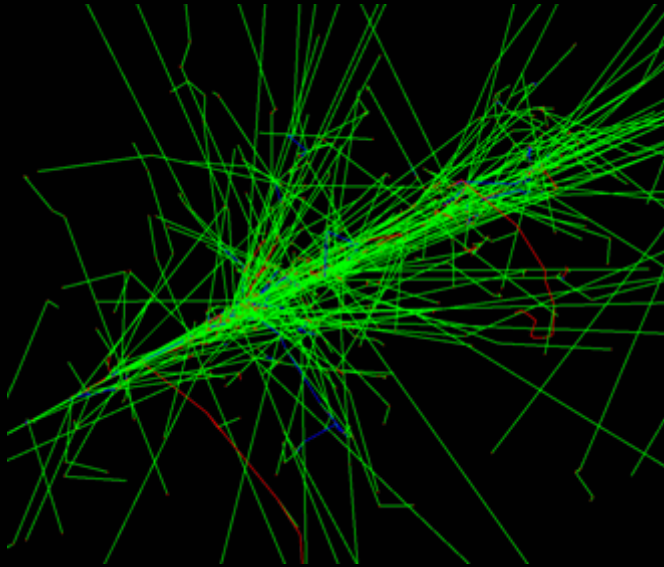


Décteur





Détail début de gerbe, avec nom des particules



La même vue de la gerbe avec les photons, en vert, sur la gauche, et sans montrer les photons, à droite. La présence d'électrons loin du cœur de la gerbe illustre le transport de la gerbe par les photons. Image facile à réaliser par simulation !

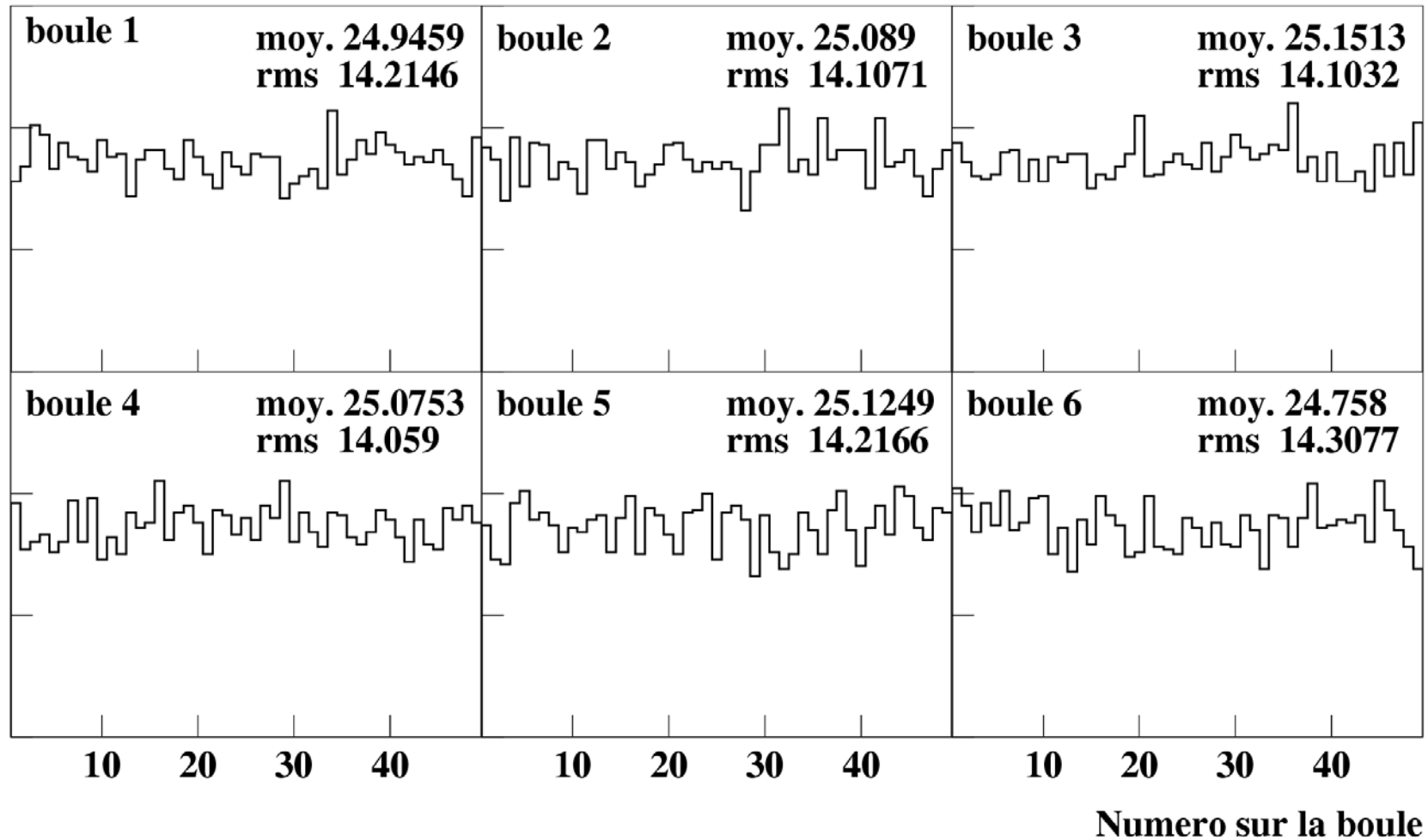
– QUATRIÈME PARTIE –

L'AJUSTEMENT

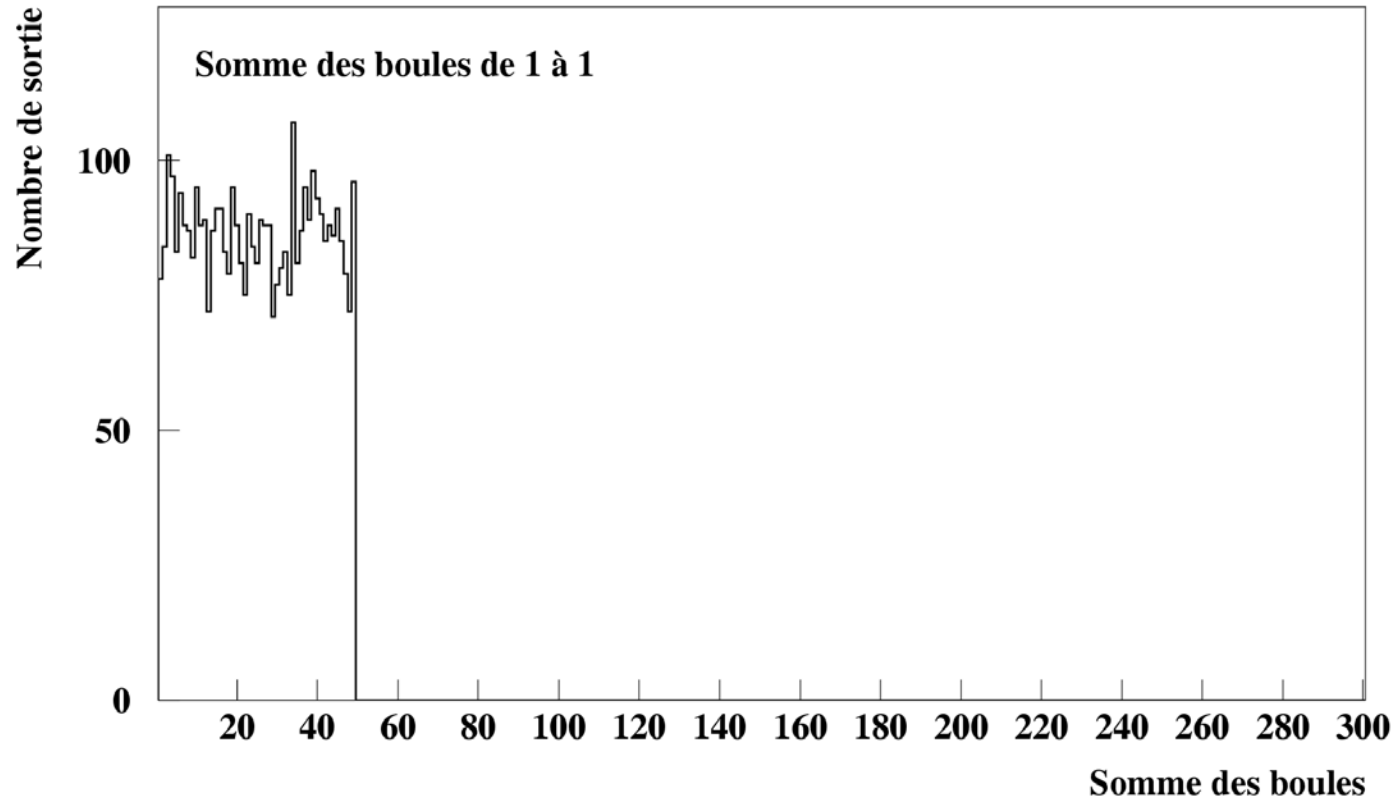
Introduction

- Nous avons vu, dans une analyse de physique, le besoin pour faire « coller au mieux » un modèle sur un échantillon de données, en ajustant les paramètres de ce modèle.
 - Dans notre cas, il s’agissait d’ajuster (notamment) $\sin(2\beta)$.
- Ce problème d’ajustement est un problème général, celui des « statistiques »
 - Il y a les « probabilités » : partant d’une distribution (théorique) connue, et contrôlée par des paramètres, connus eux aussi, à quoi ressemble un échantillon de données ?
 - Et les « statistiques » : partant d’un échantillon de données observées, et supposant une distribution, comment estimer les paramètres contrôlant la distribution ?
 - Et dans quelle fourchette peuvent varier ces paramètres ?

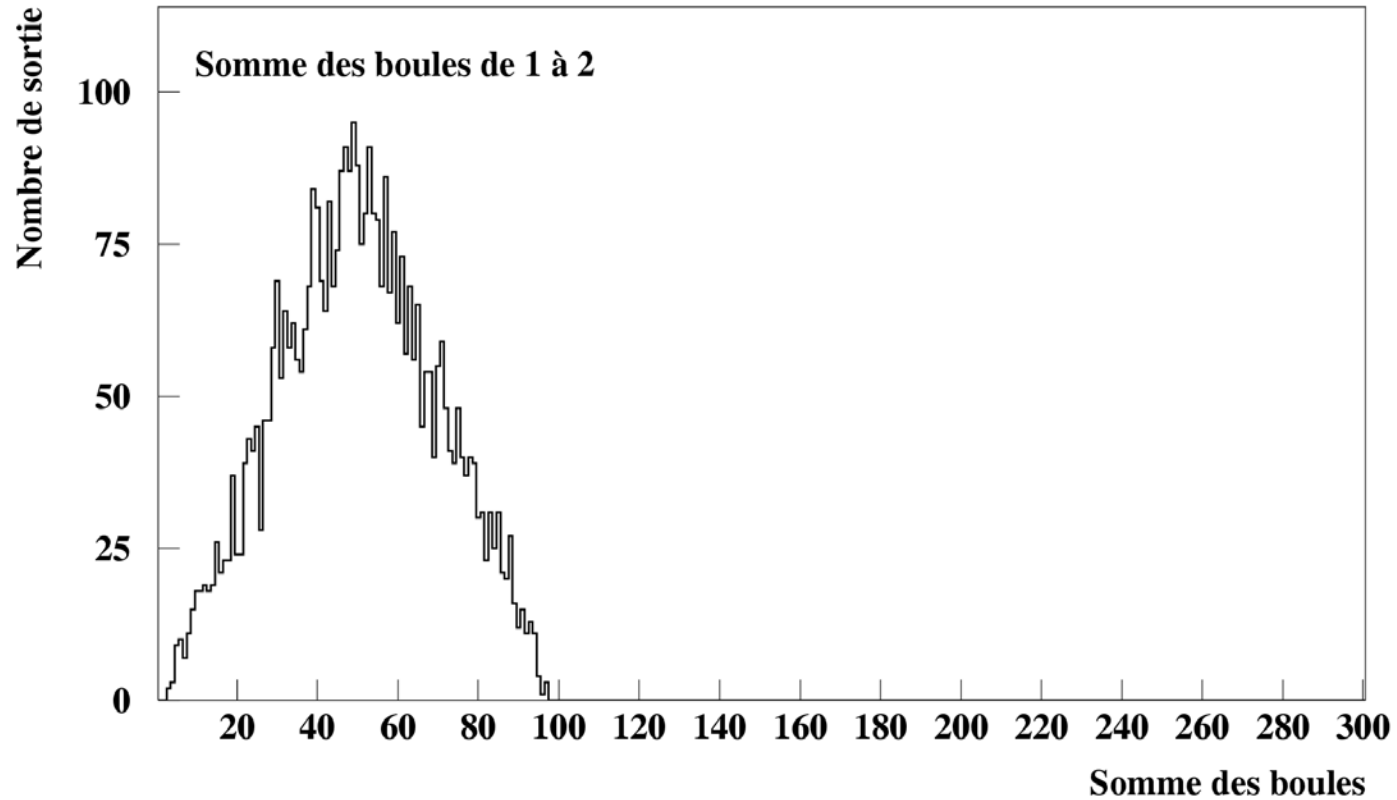
Les tirages du LOTO !



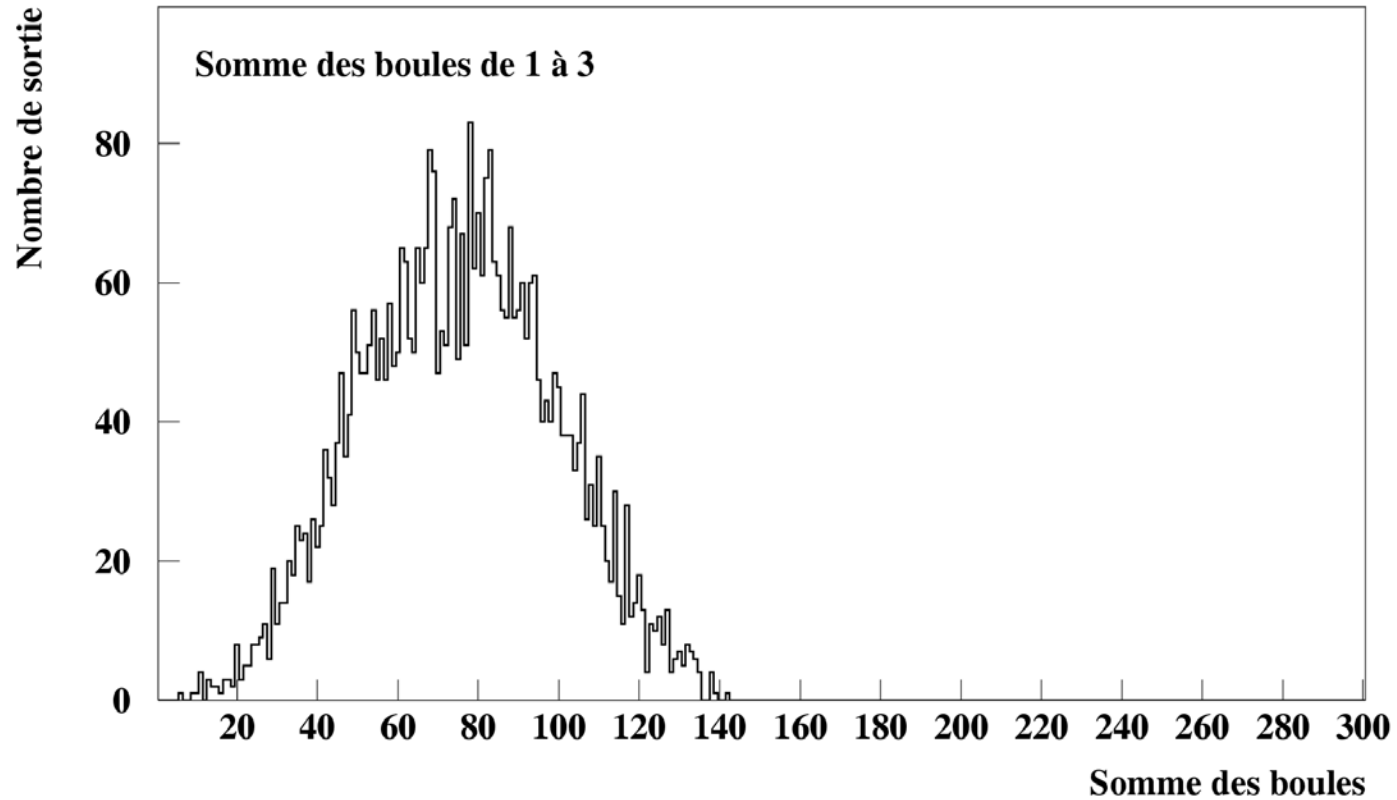
Somme des numéros



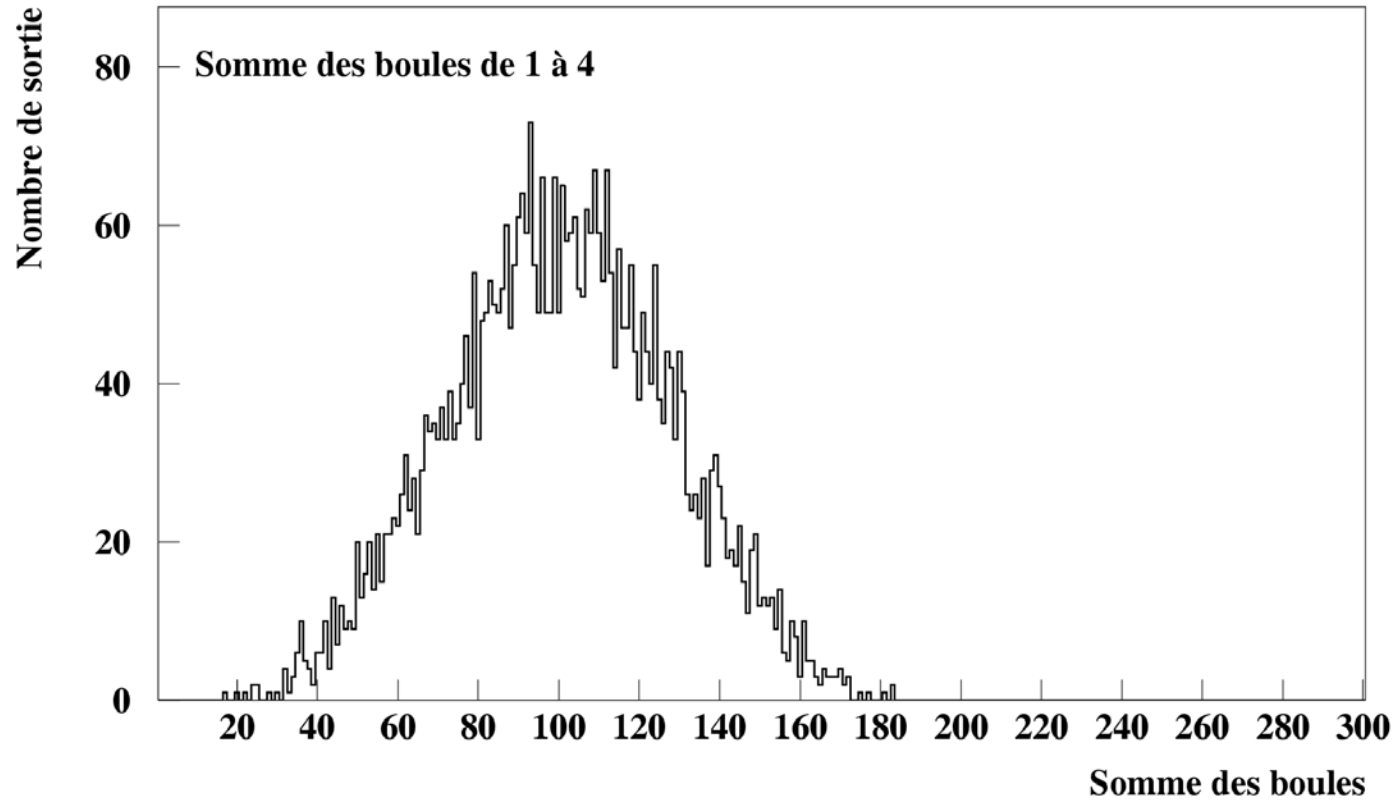
Somme des numéros



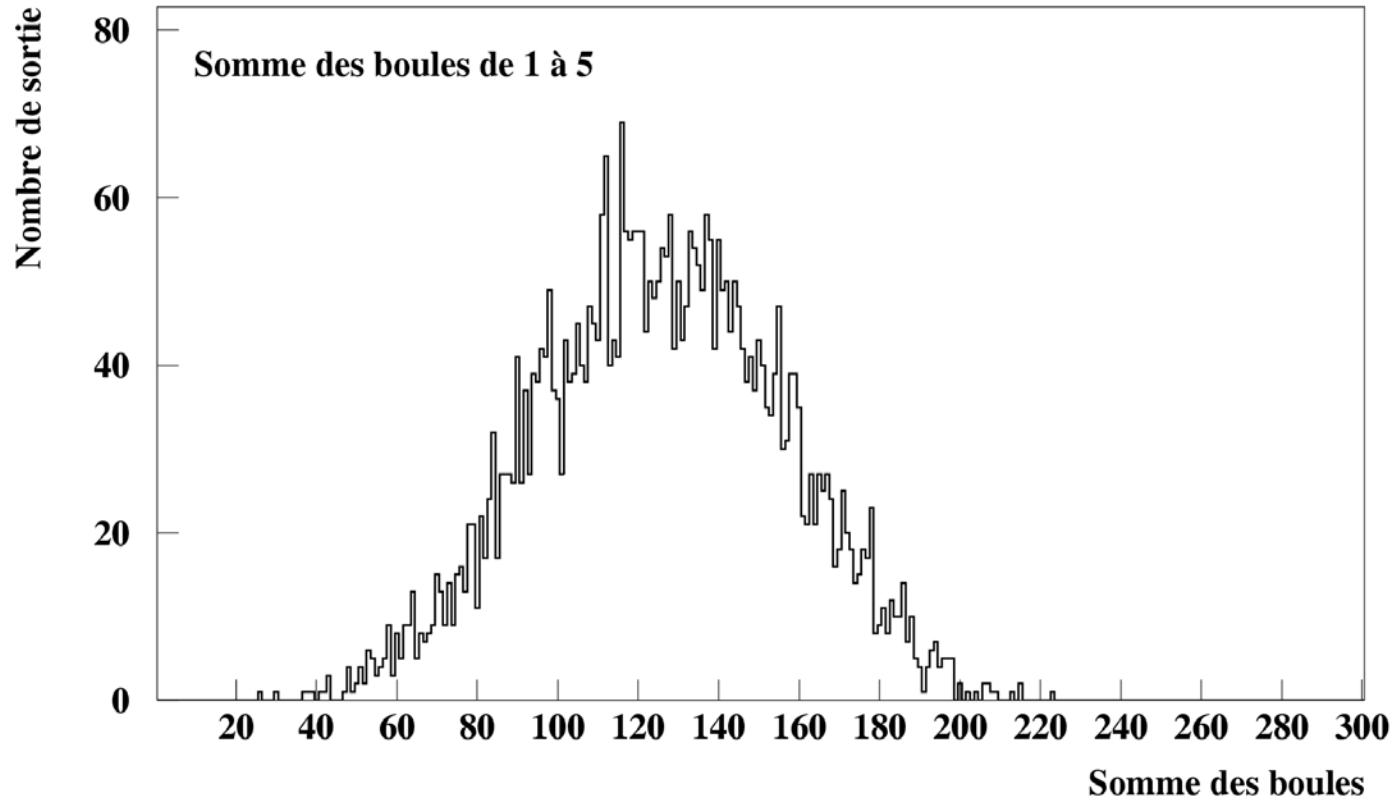
Somme des numéros



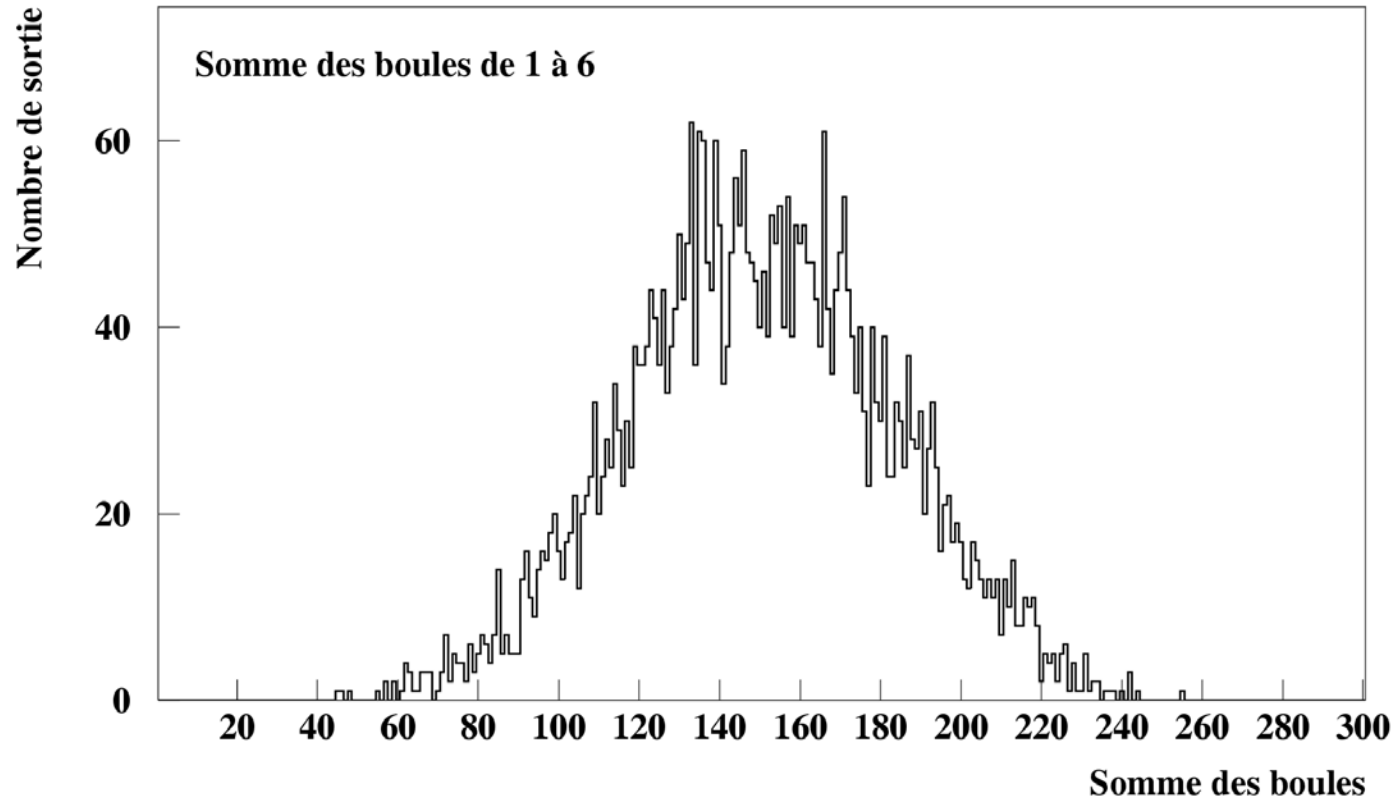
Somme des numéros



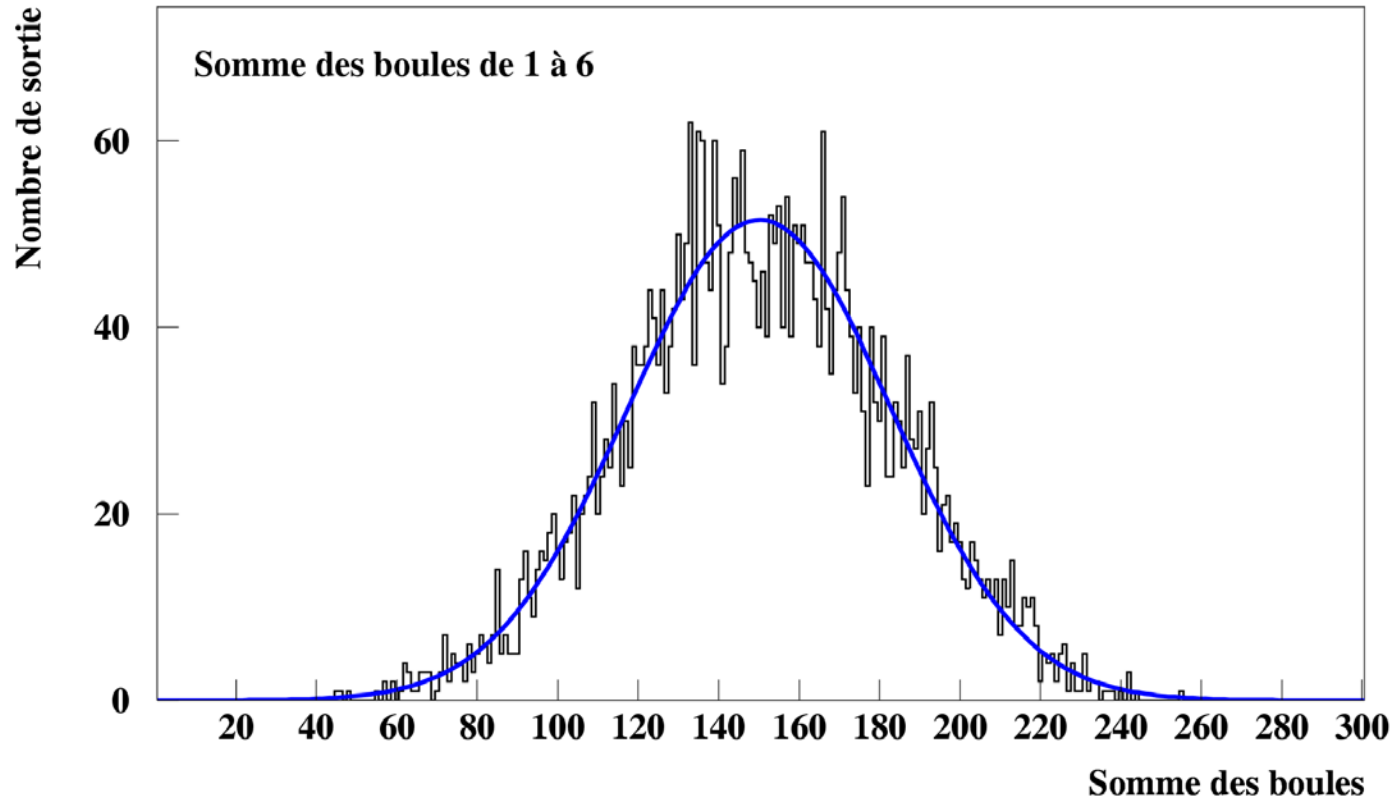
Somme des numéros



Somme des numéros



Somme des numéros



- Ajustement par une « gaussienne »

« central limit theorem »

- C'est pour cela que les physiciens aiment les gaussiennes:
 - Résultat d'une mesure (énergie, position...) peut-être vu comme provenant d'une suite de contributions, chacune apportant sa fluctuation
 - Fluctuation du nombre de traces déposant de l'énergie dans un cristal, fluctuation des dépôts d'énergie de ces traces, fluctuation du bruit électronique, etc...
 - Cela n'a rien de surprenant si en bout de chaîne on se retrouve avec des quantités qui ont des distributions gaussiennes
- C'est le « central limit theorem »:
 - « Si une variable aléatoire X est formée de la somme de N variables aléatoires x_i indépendantes, de moyenne μ_i et de variance σ_i , alors X est distribuée suivant une gaussienne de valeur moyenne $\mu = \sum_{i=1,N} \mu_i$ et de variance $\sigma^2 = \sum_{i=1,N} \sigma_i^2$, dans la limite $N \rightarrow \infty$. »
- Avec quelques bémols dans notre cas:
 - Les x_i doivent être continues (mais j'ai pris des gros bins)
 - On a $N=6$
 - Mais résultat déjà pas si mal !

Nombre de

Somme des boules

20 40 60 80 100 120 140 160 180 200 220 240 260 280 300

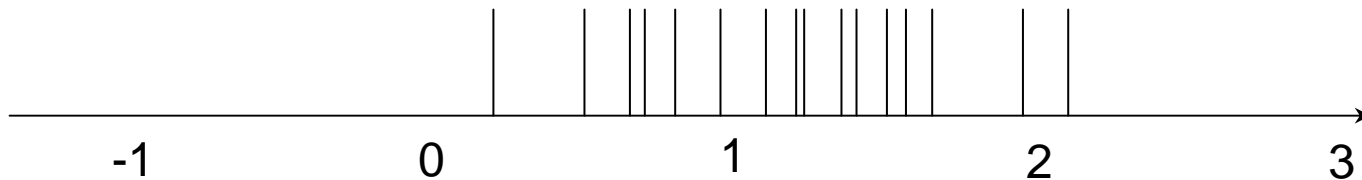
Somme des boules

PROBLEMATIQUE DES STATISTIQUES

- Oui mais voilà:
 - Ce que nous avons vu là concerne le problème des probabilités:
 - On connaît des distributions de départ
 - « Densité de probabilité », dans le cas de variables continues
 - » « Probability density function » ou *pdf*, en Anglais
 - Uniforme = plate, pour le tirages des numéros,
 - Gaussienne pour la somme des numéros
 - Et on observe des échantillons de données suivant ces distributions.
- Dans une analyse de physique, c'est le « contraire »:
 - On observe des données et on aimerait bien remonter à la distribution sous-jacente.
 - En général on a un modèle pour ces données, mais les paramètres (par exemple $\sin(2\beta)$) contrôlant la densité de probabilité sont inconnus.
- Dit autrement, on veut « estimer » ces paramètres.
 - La notion « d'estimateur » a un sens précis en statistique !
- Des estimateurs, ils en existent de nombreux (et on a le droit d'en inventer !)
 - Tous ne sont pas aussi performants
 - La méthode d'estimation la plus populaire est le « maximum de vraisemblance » car elle fournit en général l'estimation la plus précise (possible).

Le problème

- On observe:



- Que valent μ et σ ?

La méthode du maximum de vraisemblance

- Supposons que l'on collecte les N valeurs $\{x_i\} = \{\dots\}$
- Et que la distribution sous-jacente est une gaussienne

$$g(x; \mu_{\text{vraie}}, \sigma_{\text{vraie}})$$

- de moyenne μ_{vraie} et d'écart type σ_{vraie} inconnus
 - et que l'on cherche donc à estimer.
- On se dit:

C'est un abus de langage, la vraie probabilité est $g(x; \mu, \sigma) \cdot dx$, c'est-à-dire la probabilité d'observer une valeur entre x et $x+dx$. Mais c'est sans importance pour la suite.

- La probabilité d'observer (par ex.) 0.967, en fonction de paramètres μ et σ , est

$$g(0.967; \mu, \sigma)$$

- Si on considère les N observations $\{x_i\}$, la probabilité d'observer ces valeurs est le produit des probabilités:

$$L(\{x_i\}; \mu, \sigma) = g(x_1; \mu, \sigma) \cdot g(x_2; \mu, \sigma) \cdot \dots \cdot g(x_N; \mu, \sigma) = \prod_{i=1, N} g(x_i; \mu, \sigma)$$

- Ce produit est appelé la « vraisemblance » (« likelihood » en anglais).
- Et l'on se dit que les valeurs μ_{max} et σ_{max} maximisant cette vraisemblance devraient approcher les valeurs vraies μ_{vraie} et σ_{vraie} .

Log(vraisemblance)

- Techniquement, le produit

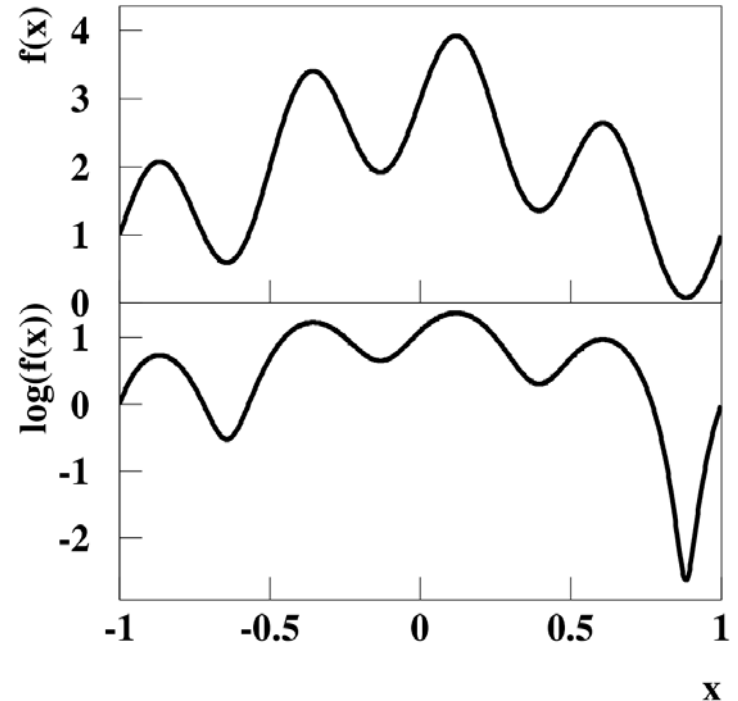
$$L(\{x_i\}; \mu, \sigma) = \prod_{i=1, N} g(x_i; \mu, \sigma)$$

aboutit à de très petits nombres, ce qui peut poser des problèmes numériques.

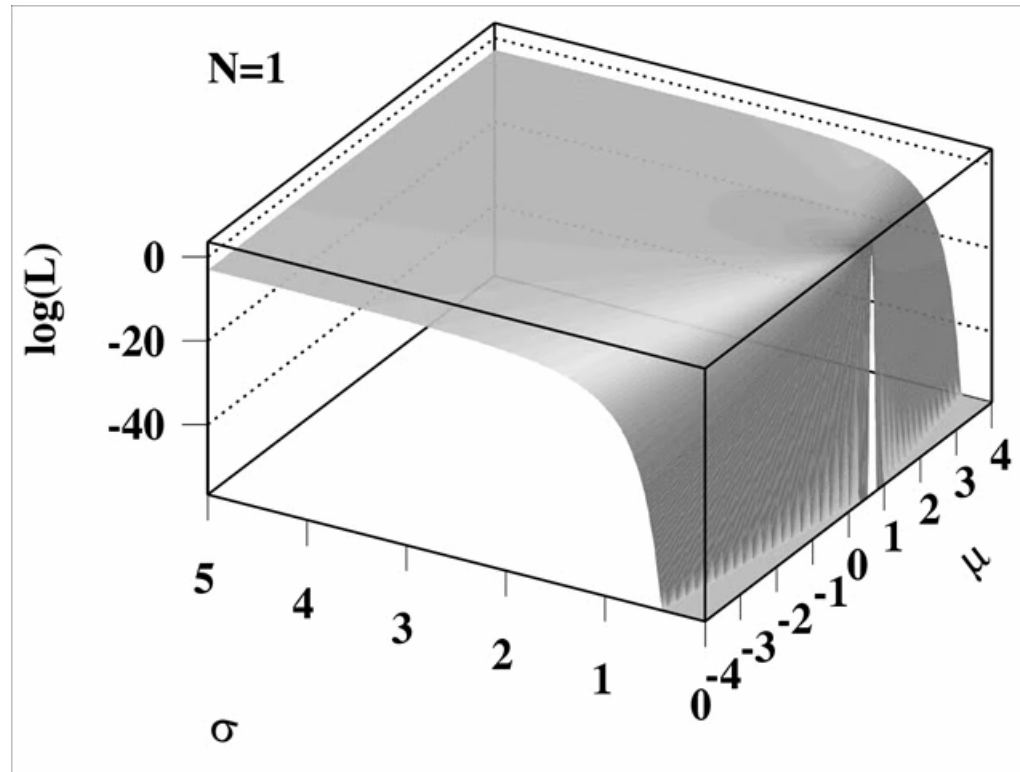
- On préfère utiliser le logarithme de cette quantité, on passe alors à la somme:

$$\log(L(\{x_i\}; \mu, \sigma)) = \sum_{i=1, N} \log(g(x_i; \mu, \sigma))$$

- (Mais ce n'est pas qu'une question de problème technique)
- S'agissant de la recherche d'un maximum, cela ne change pas le problème.
- Illustration de comment cette somme évolue avec le nombre d'évènements.
 - Dans la suite, des données $\{x_i\}$ ont été générées suivant une gaussienne
 - de valeur moyenne 1 et d'écart type 2
 - et l'on trace le Log(vraisemblance) pour différents N (de 1 à 1024).



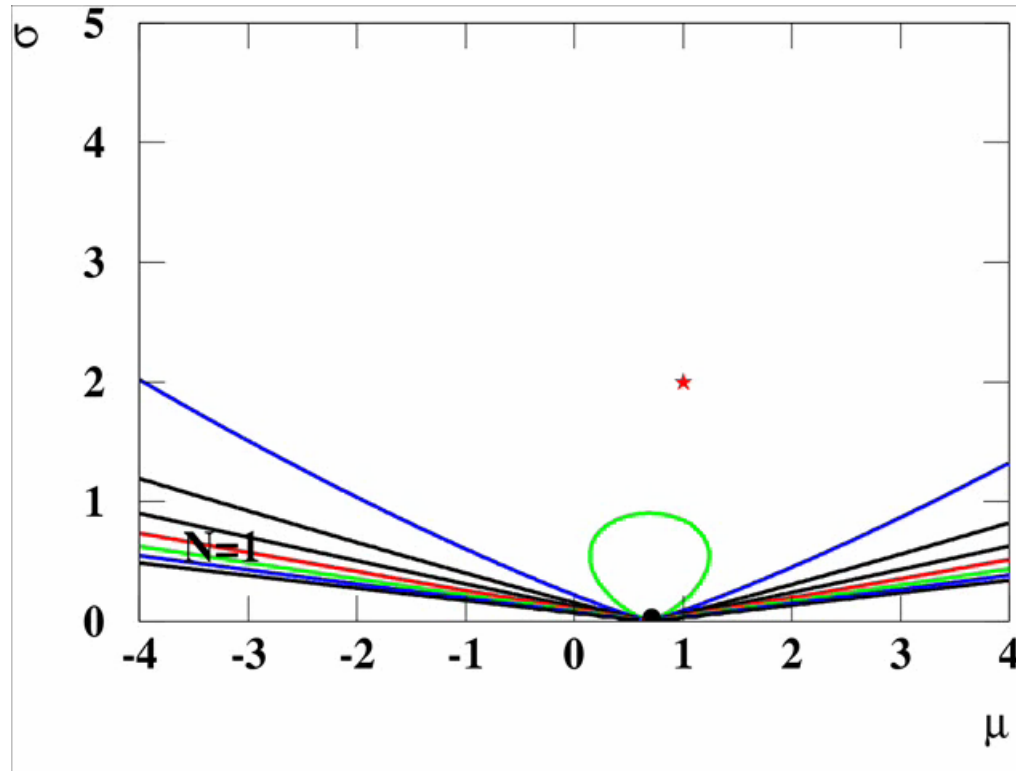
Log(vraisemblance)



(cliquer sur l'animation pour démarrer)

- Plot de $\log(\text{vraisemblance})$:
 - Hauteur de la courbe montrée: de $\{\max \text{ à } \max - 11^2/2\}$

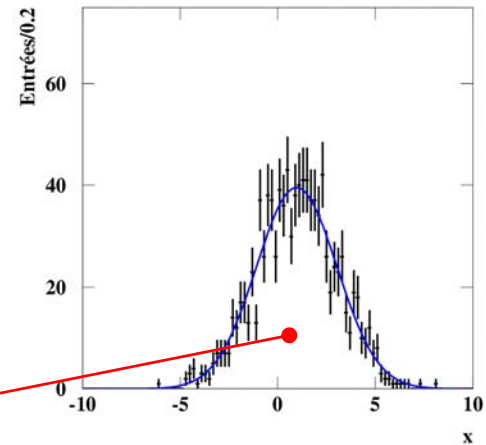
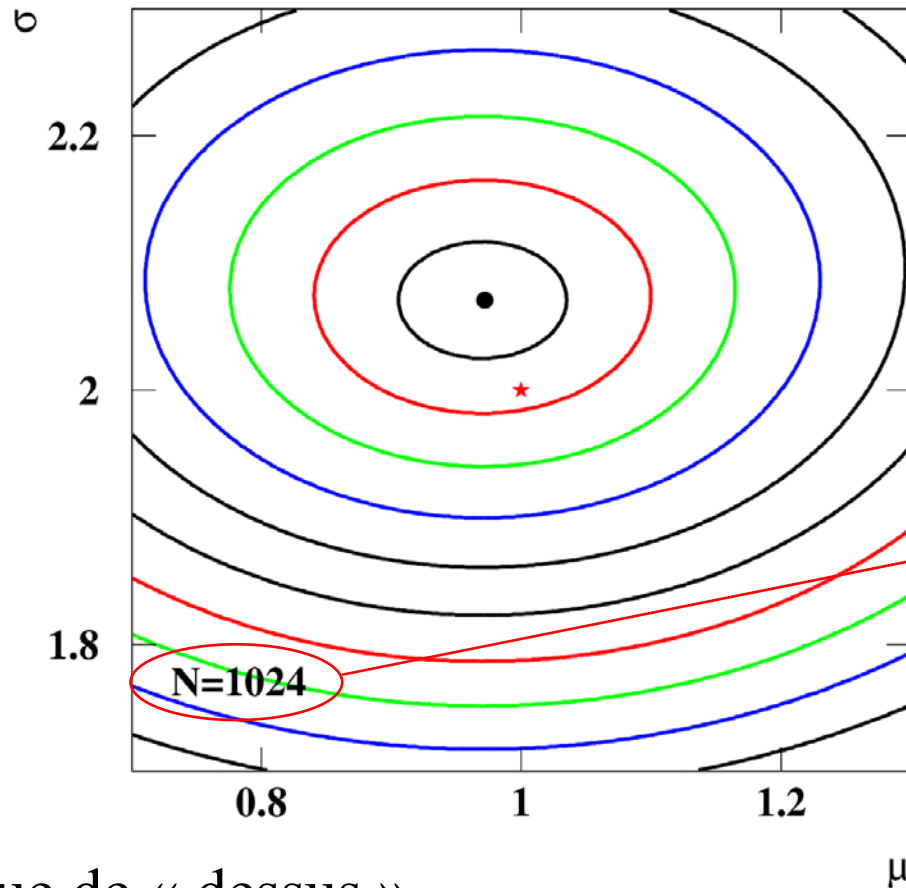
Log(vraisemblance)



(cliquer sur l'animation pour démarrer)

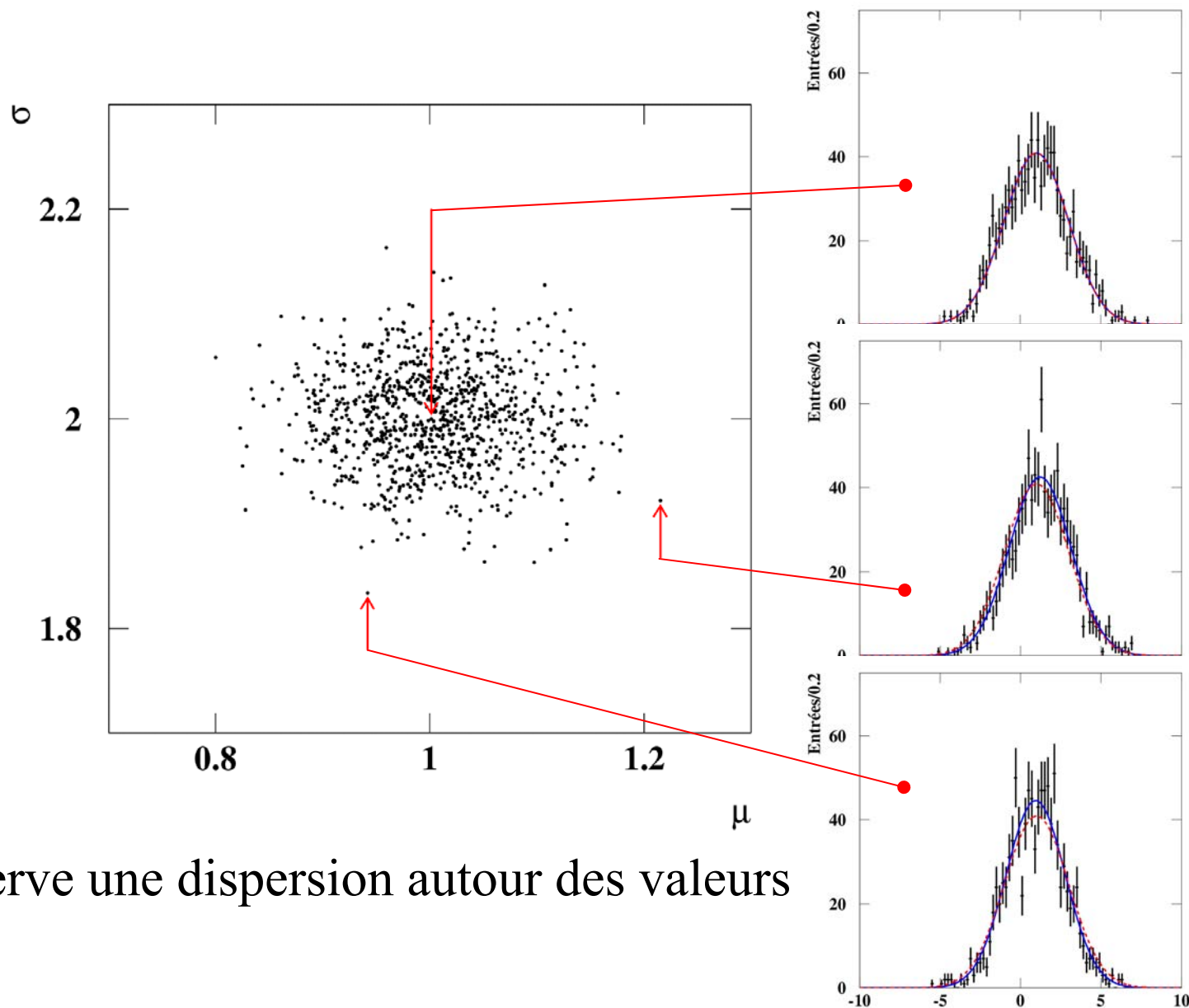
- Idem, vue de « dessus »
 - Les contours correspondent à des variations de $-1^2/2$, $-2^2/2$, $-3^2/2$... $-10^2/2$ par rapport au maximum

Log(vraisemblance)



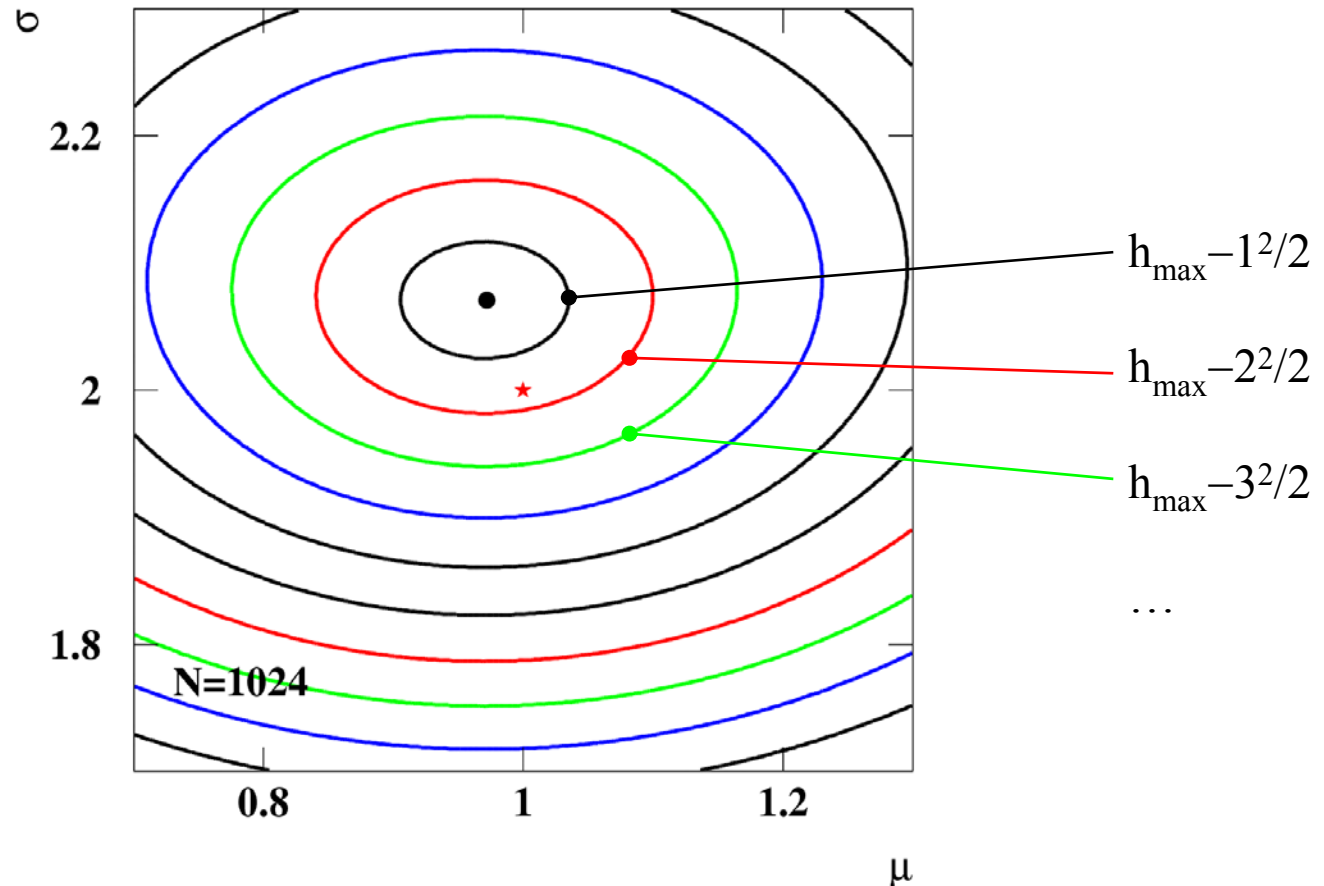
- Idem, vue de « dessus »
 - Les contours correspondent à des variations de $-1^2, -2^2, -3^2 \dots -10^2$ par rapport au maximum

Et répétant 1000 fois la même « expérience »...



- On observe une dispersion autour des valeurs vraies.

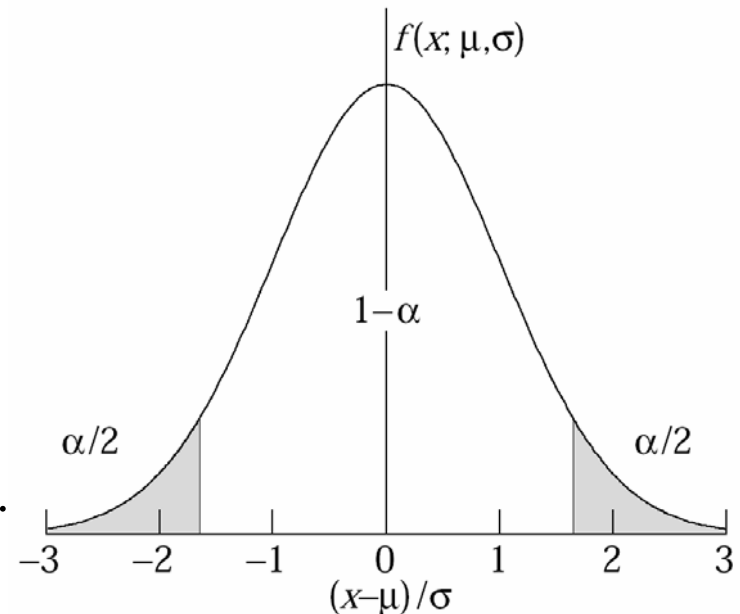
Dispersion & incertitudes statistiques



- Valeur « vraie », i.e. générée, se trouve à environ $(1.8)^2/2$ plus bas que le maximum.
- Quid des autres expériences ?

Incertitude statistique

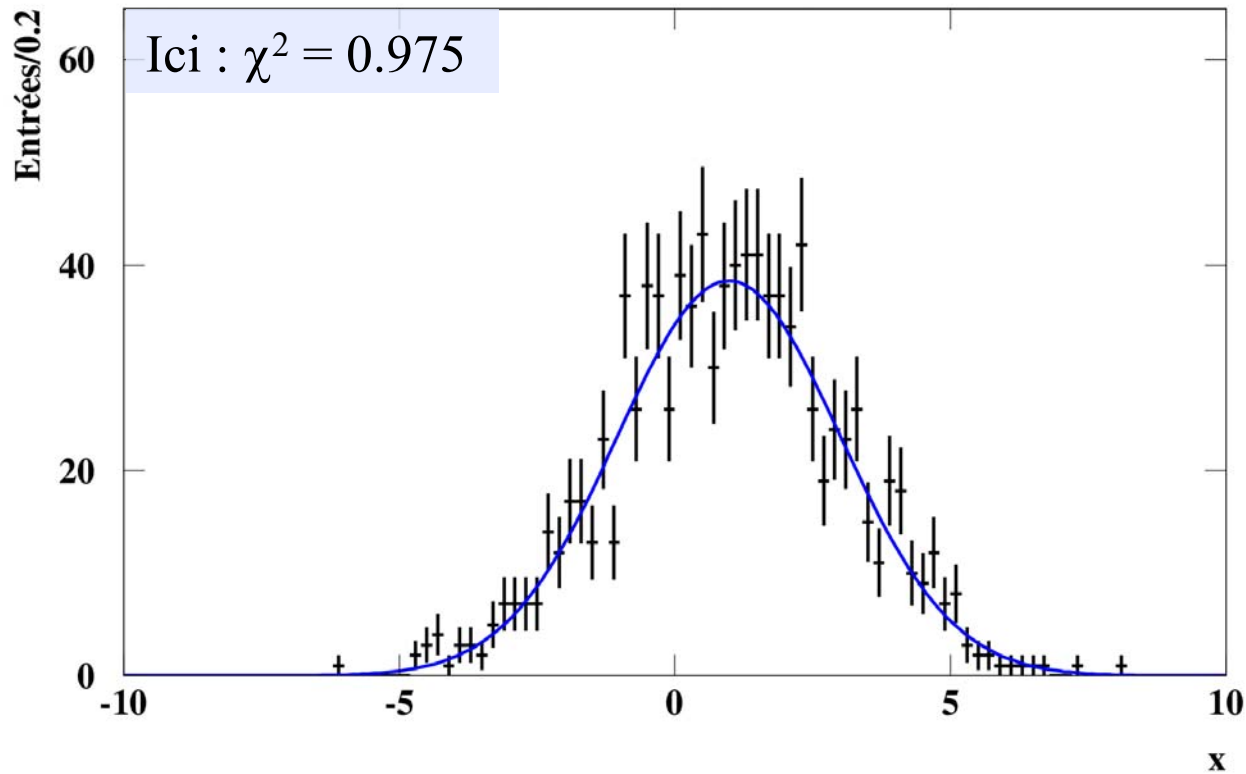
- Dans la limite des hautes statistiques ($N \rightarrow \infty$) on montre alors que les maxima se répartissent suivant une gaussienne
 - Dont la moyenne correspond aux valeurs vraies
 - Dont le(s) écart(s) type(s) peuvent être estimés par la largeur du $\log(\text{Vraisemblance})$ à $\max - 1/2$.
- C'est dans cette interprétation gaussienne des incertitudes que l'on rapporte le résultat.
 - Et si l'on s'écarte de cette situation, on le décrit
- On donne:
 - La valeur centrale de l'ajustement
 - Le(s) sigma(s) de cette gaussienne
- Cela permet de définir la probabilité pour que tel ou tel intervalle contienne / ne contienne pas le résultat vrai.



α (%)	δ	α (%)	δ
31.73	1σ	20	1.28σ
4.55	2σ	10	1.64σ
0.27	3σ	5	1.96σ
6.3×10^{-3}	4σ	1	2.58σ
5.7×10^{-5}	5σ	0.1	3.29σ
2.0×10^{-7}	6σ	0.01	3.89σ

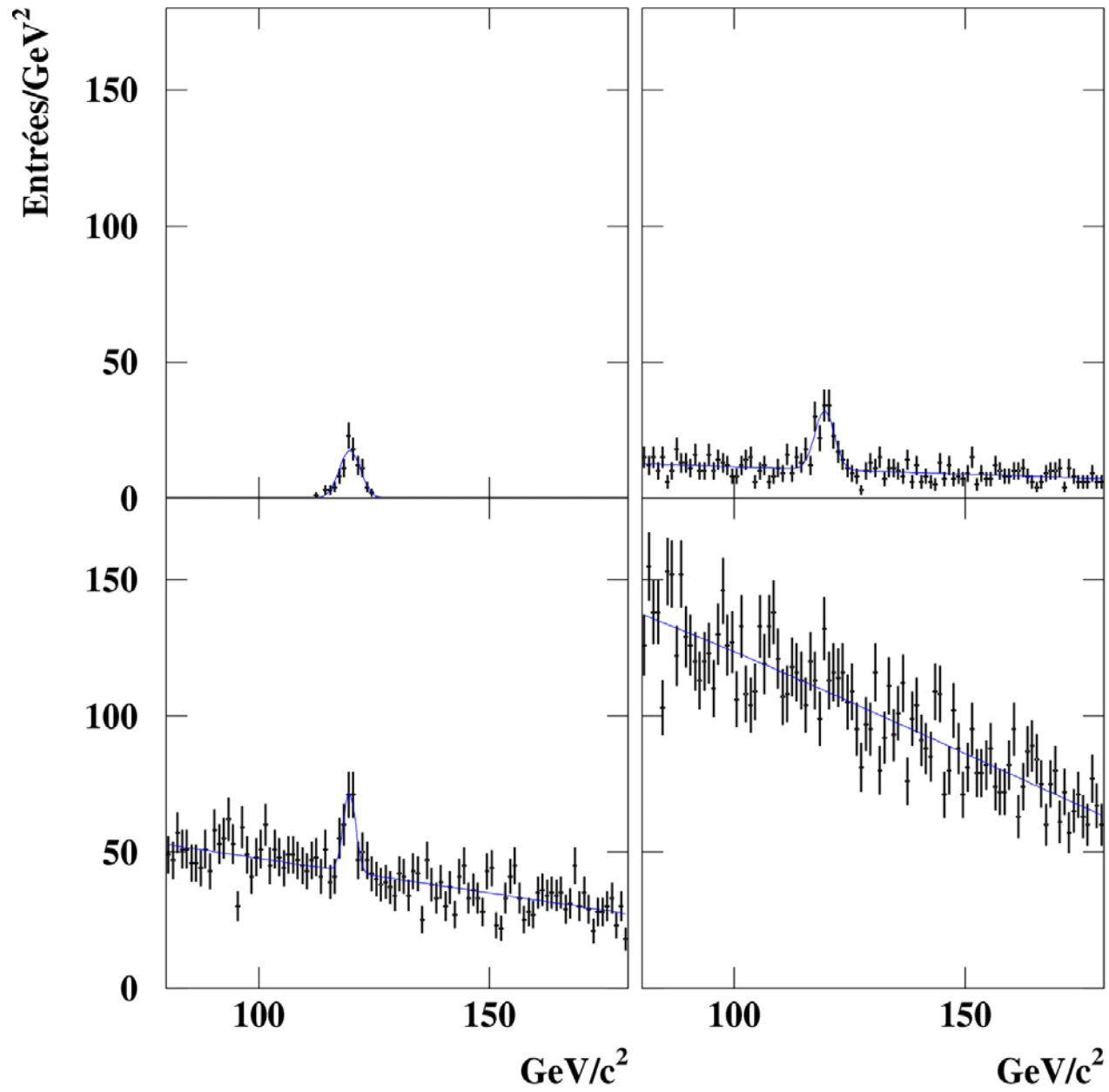
Une autre méthode: le χ^2

$$\chi^2(\text{parametres}) = \sum_{i=1}^N \left(\frac{y_i - \text{modele}(x_i; \text{parametres})}{\sigma_i} \right)^2$$



Recherche de signal

- Situation où l'on cherche à mettre en évidence l'existence d'une nouvelle particule
 - Higgs
- Typiquement, après sélection, on cherche un « pic » de masse
 - Sur un bruit de fond plus ou moins important
- Lorsque le résultat est publié, si la recherche est positive, on donne une « P-value »:
 - Probabilité qu'une fluctuation du bruit de fond soit la cause du signal observé



Et en pratique ?

- En pratique on utilise (très souvent) le programme MINUIT
 - Du bon vieux FORTRAN !
- MINUIT est un programme de minimisation qui travaille dans un nombre arbitraire de dimensions
 - On maximise le $\log(\text{vraisemblance})$ en minimisant $-\log(\text{vraisemblance})\dots$
 - On minimise les χ^2
- Et MINUIT sait déterminer les « iso-courbes ».
 - Ce qui permet de déterminer les incertitudes

Quelques références

- Geant4:
 - WEB:
 - <http://wwwasd.web.cern.ch/wwwasd/geant4/geant4.html>
 - Documentation
 - Nombreux liens vers exposés et publication sur sujets précis (rubrique « Results & Publications »)
 - Publications generales:
 - Nuclear Instruments and Methods in Physics Research A 506 (2003) 250-303, and IEEE Transactions on Nuclear Science 53 No. 1 (2006) 270-278.
- Statistiques:
 - Nombreux ouvrages, mon préféré (fait par physicien des particules):
 - Glen Cowan, “ Statistical Data Analysis ”, Oxford University Press
 - Voir aussi page WEB de l’auteur:
 - <http://www.pp.rhul.ac.uk/~cowan/sda/>
 - Chapitre PDG sur les stats