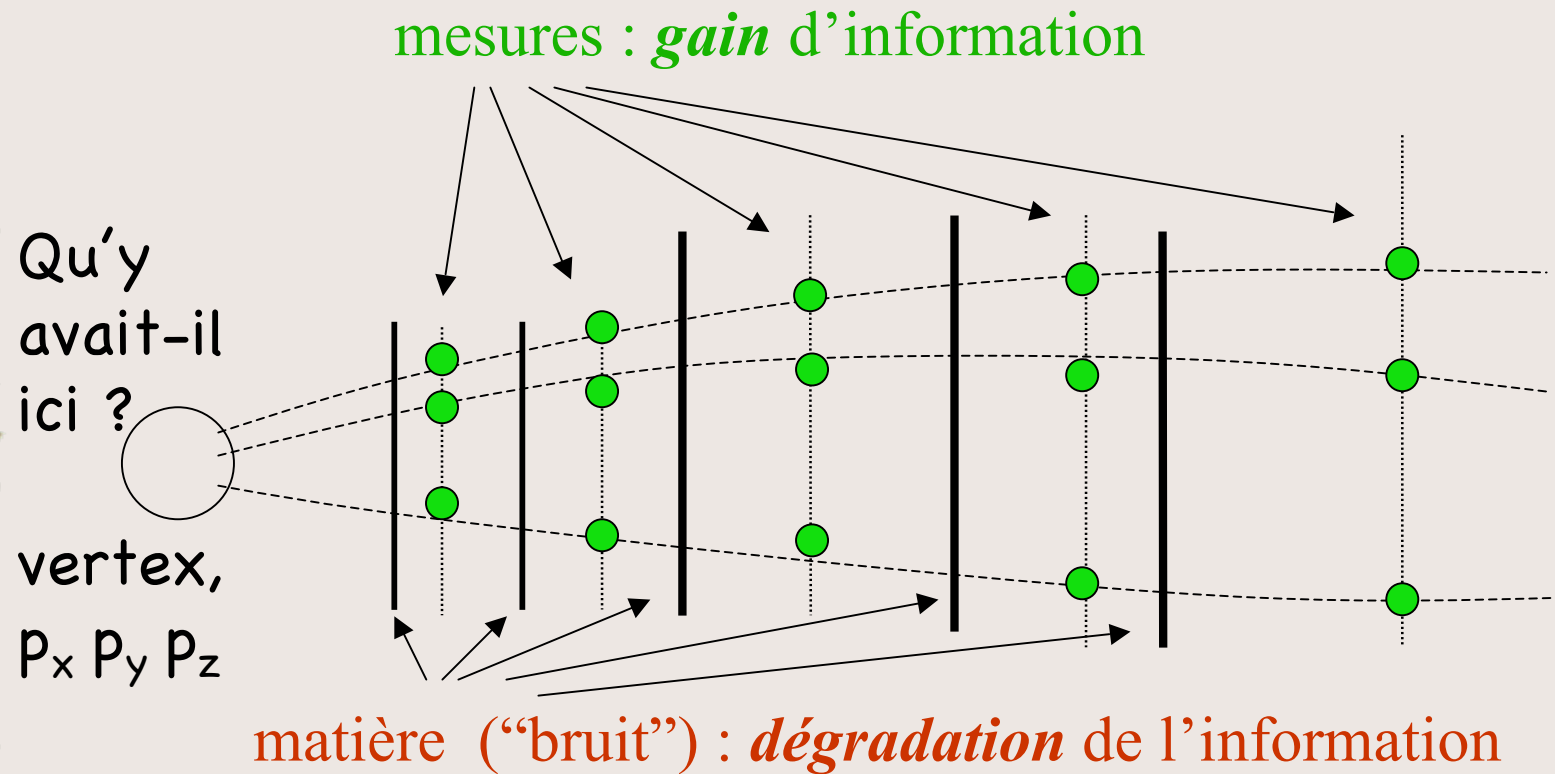


Trajectographie

IN2P3 School of Statistics, Autrans 2012

P. Billoir, LPNHE Paris

Quel est le problème ?



Comment construire le meilleur estimateur des paramètres utiles ?

Les ingrédients

ce qui est supposé connu

- la nature et la précision des mesures
- la nature et l'amplitude des “bruits” dans la matière (diffusion multiple, fluctuation de perte d'énergie)
- l'équation de propagation (forme des trajectoires)

Remarque: les deux derniers points (et parfois le premier) dépendent de *l'état instantané* de la particule (éventuellement de sa *nature*) ; on fera parfois une reconstruction *dans l'hypothèse* e, μ, π , etc

ce qu'on veut reconstruire

- les paramètres “à l'origine” (avant tout bruit, juste après la production)
- éventuellement : les paramètres à la fin (extrapolation chambres à μ) ou en un point intermédiaire
- une “pré-reconstruction” pour lever des ambiguïtés d'attribution
- la position du (ou des) vertex ; les paramètres des particules au vertex *en utilisant la contrainte d'origine commune* (prochain épisode)

L'outil de base pour les trajectoires : le Filtre de Kalman (méthode *progressive*)

Ce qu'on trouve dans les livres... (ici : Wikipedia)

Predict

Predicted (*a priori*) state estimate $\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_k \mathbf{u}_k$

Predicted (*a priori*) estimate covariance $\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_k$

Update

Innovation or measurement residual $\tilde{\mathbf{y}}_k = \mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}$

Innovation (or residual) covariance $\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k$

Optimal Kalman gain $\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1}$

Updated (*a posteriori*) state estimate $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k$

Updated (*a posteriori*) estimate covariance $\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}$

on va essayer de faire l'équivalent de façon intuitive...
et d'aller plus loin !

Un problème à un paramètre où est/était la puce ?

abscisse initiale : x_0

à chaque pas de temps (indépendamment)

- on fait une mesure (précision σ)
- elle fait un saut (écart-type τ)

quel est le “meilleur” estimateur de la position x_0 ? x_n ?

intuitivement :

- si $\sigma \ll \tau$: c'est la mesure instantanée ; les autres n'apportent rien
- si $\tau \ll \sigma$: on gagne à moyenner des mesures successives à des temps proches de celui qu'on veut ; mais pas indéfiniment...
- cas général : pas évident...
- l'optimum doit être une combinaison *pondérée* des mesures

Comment évaluer les poids ?

La solution optimale *lourde*

On veut estimer x_0 ; on exprime les mesures :

$$x_0^{\text{mes}} = x_0 + \varepsilon_0$$

$$x_1^{\text{mes}} = x_0 + \eta_1 + \varepsilon_1$$

$$x_2^{\text{mes}} = x_0 + \eta_1 + \eta_2 + \varepsilon_2$$

...

ε_k : erreur sur le point k ; η_k : saut k

D'où la matrice C de variance-covariance des écarts $\Delta x_k = x_k^{\text{mes}} - x_0$:

$$\sigma^2 \quad 0 \quad 0 \quad 0 \quad \dots$$

$$0 \quad \sigma^2 + \tau^2 \quad \tau^2 \quad \tau^2 \quad \dots$$

$$0 \quad \tau^2 \quad \sigma^2 + 2\tau^2 \quad 2\tau^2 \quad \dots$$

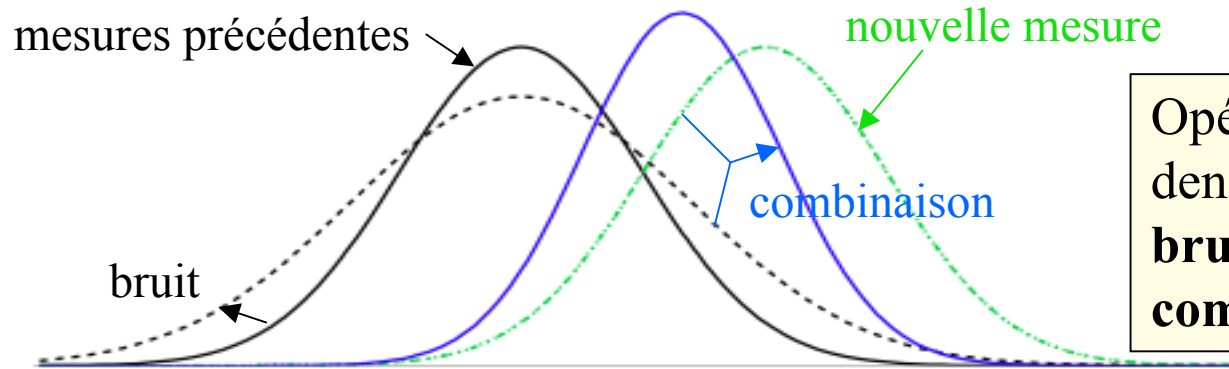
$$0 \quad \tau^2 \quad 2\tau^2 \quad \sigma^2 + 3\tau^2 \quad 3\tau^2 \quad \dots$$

.....

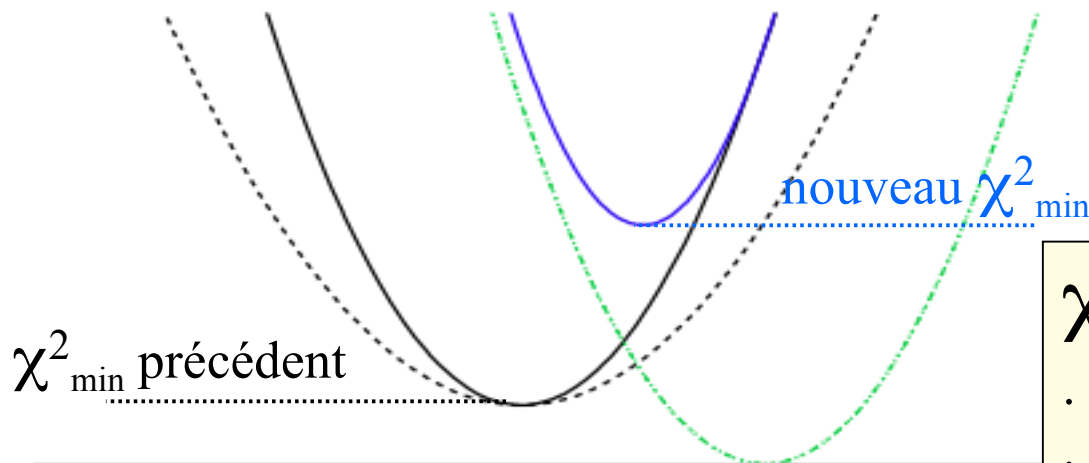
$$\chi^2 = \sum (C^{-1})_{ij} \Delta x_i \Delta x_j \rightarrow x_0^{\text{fit}} = \sum_j (C^{-1})_{ij} x_i^{\text{mes}}$$

avec n mesures, il faut inverser une matrice $(n \times n)$

Une étape d'un estimateur progressif (le fur et la mesure)



Opérations sur la densité de probabilité
bruit : *convolution*
combinaison : *produit*



χ^2 si erreurs gaussiennes :
. forme parabolique
. $\text{Prob}(\chi^2, n_{\text{deg}})$ exploitable

La solution *légère* (progressive)

On veut estimer x_0 : on part de la fin !

x_n^{mes} est un estimateur de x_n , de variance σ^2

$x_{n-1} = x_n - \eta_n \rightarrow x_n^{\text{mes}}$ est un estimateur de x_{n-1} , de variance $\sigma^2 + \tau^2$

x_{n-1}^{mes} est un estimateur de x_{n-1} de variance σ^2 , *indépendant*

→ leur *barycentre* (poids $1/(\sigma^2 + \tau^2)$, $1/\sigma^2$) est l'estimateur optimal $X_{n \rightarrow n-1}$ de x_{n-1}

avec x_n^{mes} et x_{n-1}^{mes} , de variance $\sigma_1^2 = 1/(1/(\sigma^2 + \tau^2) + 1/\sigma^2)$

et aussi un estimateur de x_{n-2} de variance $\sigma_1^2 + \tau^2$, *indépendant* de x_{n-2}^{mes}

→ on peut combiner $X_{n \rightarrow n-1}$ et x_{n-2}^{mes} pour obtenir $X_{n \rightarrow n-2}$

et ainsi de suite jusqu'à x_0^{mes}

chaque étape : dégradation + nouvelle mesure

(le nombre total d'opérations est *proportionnel à n*)

on peut aussi calculer itérativement un χ^2

les bonus (quasi gratuits)

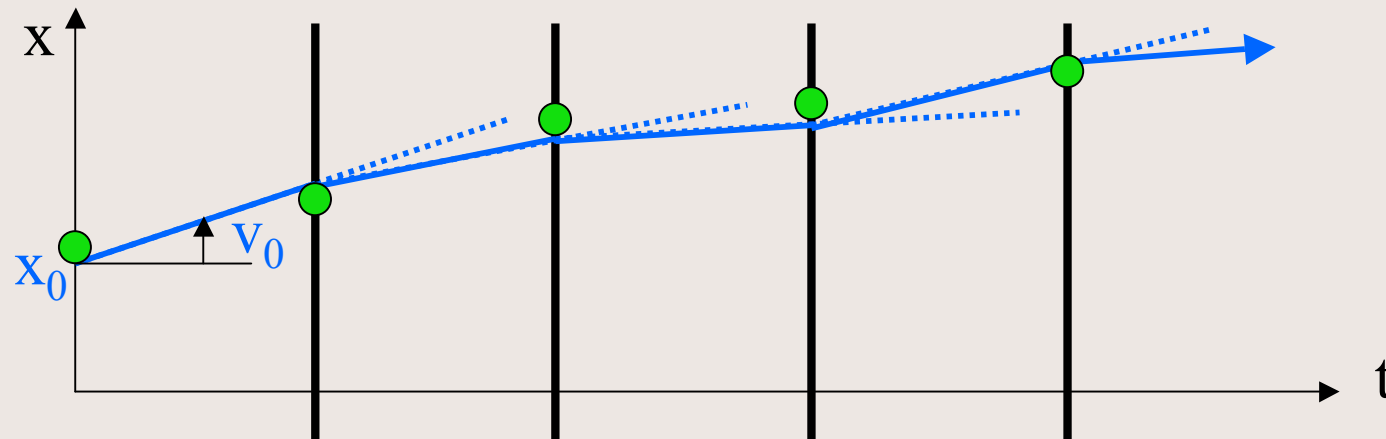
- pour estimer la position finale x_n :
même procédure dans l'autre sens, en partant de x_0^{mes}
- pour estimer x_k (interpolation) : on part des deux bouts jusqu'au point k , et on peut combiner deux estimateurs *indépendants* $X_{n \rightarrow k}$ et $X_{0 \rightarrow k}$, en incorporant x_k^{mes} à un seul des deux.
(équivalent au "smoother" dans le jargon kalmanien)
- on peut aussi avoir un **critère de compatibilité** : on interpole comme ci-dessus, sans incorporer x_k^{mes} : $x_k^{\text{interp}}(\text{sans } x_k^{\text{mes}}) - x_k^{\text{mes}}$ a une variance $V(\text{interp}) + \sigma^2$
- on peut détecter un **saut anormal** en comparant $X_{n \rightarrow k} - X_{0 \rightarrow k}$ à la variance prédite

En bref : avec le "filtre avant" et le "filtre arrière" (en gardant les résultats intermédiaires) on peut avoir tout ce qui précède

Mais: si on modifie quelque chose (ajout ou retrait d'une mesure par exemple), il faut refaire toutes les étapes à partir de ce point. Par exemple: si on travaillait "à la volée" (mesures incorporées en temps réel), le filtre arrière serait lourd ...

problème linéaire à deux paramètres

mouvement à vitesse "bruitée"



conditions initiales: x_0, v_0 (qu'on veut estimer)

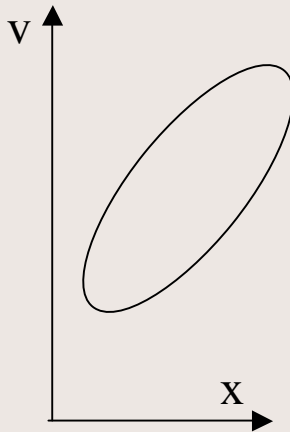
à chaque pas de temps Δt :

- mesure de x (erreur ε_k , variance σ^2)
- variation aléatoire ζ_k de v_k (variance ρ^2)
- déplacement $v_k \cdot \Delta t$

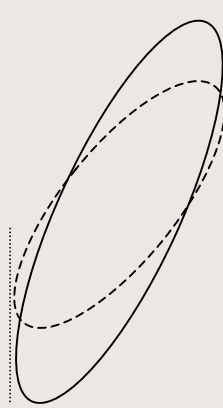
$$x_k^{\text{mes}} = x_0 + (v_0 + \zeta_1) \Delta t + (v_0 + \zeta_1 + \zeta_2) \Delta t + \dots + \varepsilon_k$$

→ *corrélation* $(x_k^{\text{mes}}, x_j^{\text{mes}})$ via les ζ_i

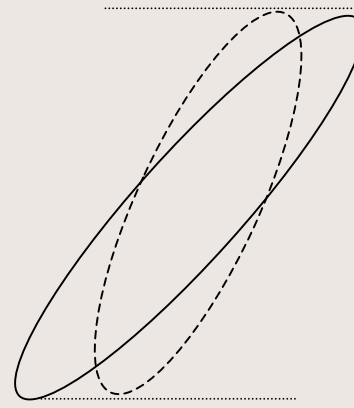
Une étape "fur et mesure" dans le plan (x,v)



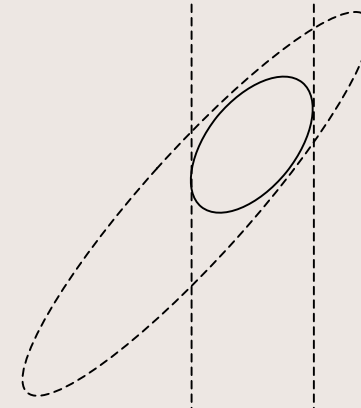
mesures
antérieures
vecteur $p_o(x_o, v_o)$
matrice $C_o(p_o)$



bruit sur v
 $C_b = C_o + B$
 $B = \begin{pmatrix} 0 & 0 \\ 0 & \rho^2 \end{pmatrix}$
*information
dégradée*



propagation
 $p' (x_o + v_o \Delta t, v_o)$
 $C' = D \cdot C_b \cdot D^t$
 $D = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix}$



combinaison avec
nouvelle mesure p_n
 $(W' + W_n) p'' = W' p' + W_n p_n$
 $W' = (C')^{-1}$, $W_n = (C_n)^{-1}$
 $C''(p'') = (W' + W_n)^{-1}$
information gagnée

un (faux) problème technique

comment commencer ?

au début : information insuffisante pour définir p_0 , et avoir C_0, W_0 inversibles

exemple : on mesure x ou une combinaison linéaire de x et $v \rightarrow W$ a une valeur propre nulle (la p.d.f. est une *bande*, et p_0 est indéterminé sur l'axe de la bande)

en pratique, les opérations élémentaires (convolution, propagation, produit) sont toujours possibles :

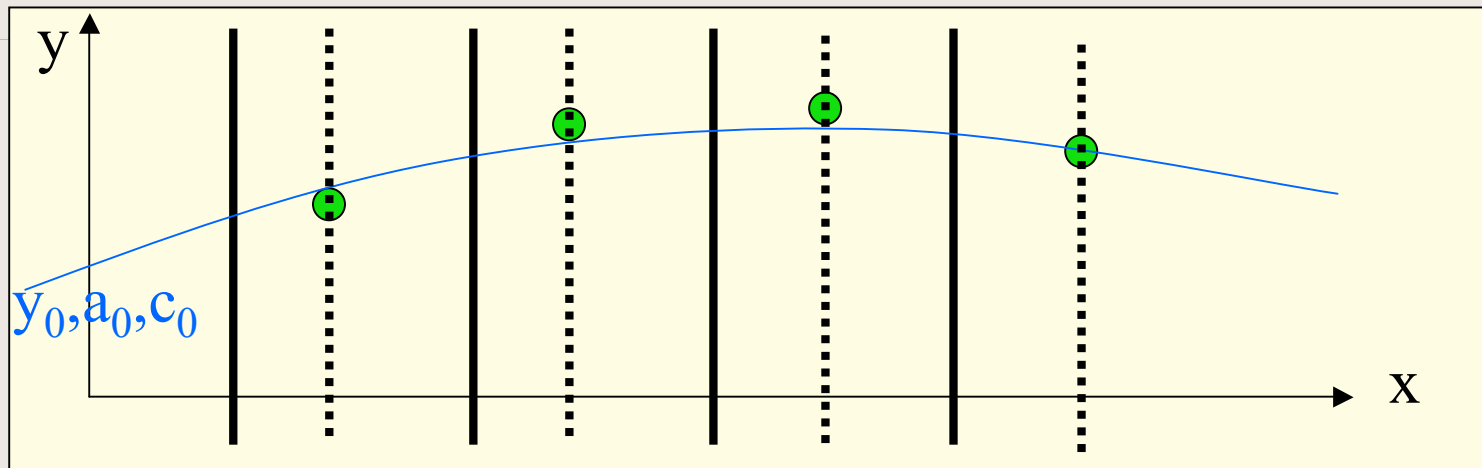
- *convolution* : $(W^{-1}+C)^{-1} = (1+WC)^{-1}.W$
 $1+WC$ est *inversible* dans les cas utiles

- *propagation* : $W' = (D^{-1})^t.W.(D^{-1})$

- *produit* : si W_1 et/ou W_2 est singulière, le système $(W_1+W_2) p = W_1 p_1 + W_2 p_2$ a une solution qui ne dépend pas du choix de p_1 et/ou p_2 sur l'axe de la bande
Cas extrême : deux bandes parallèles : p indéterminé sur une droite, et on peut réinjecter cette information.

on peut travailler avec les matrices de poids dans toutes les étapes

problème à trois paramètres (trajectographe plan simplifié)



modèle parabolique: $y(x) = y_0 + a_0 x + c_0 x^2/2$

($c=1/R$ dans l'approximation de faible pente a)

paramètres physiques : y_0 , p_y/p_x , $1/p$ ($c = 0.3B/p$)

les mesures dépendent linéairement des paramètres

→ même formalisme, avec des matrices (3x3) pour C,W,D

si perte d'énergie non négligeable : Δc dans l'étape "propagation"

bruit : diffusion multiple (sur a) ; fluctuation de $\Delta E \rightarrow$ terme sur c

NB: pour évaluer les termes de bruit, E doit être pré-estimé

approximation linéaire

Dans le monde réel : pas de paramétrisation *linéaire* exacte
une solution possible :

- choisir des paramètres \mathbf{p} commodes (ex: coord. cartésiennes ou cylindriques)
- définir des *lignes/surfaces* de *mesure* (plans, cylindres,...) et de *matière*
(au besoin : le bruit dans un matériau épais peut être “condensé” sur une surface dans une matrice C_b avec corrélation entre position et angle)
- définir une *trajectoire de référence* \mathbf{T}_{ref} proche de la vraie (pré-fit)
- faire la propagation des *écarts* $\delta\mathbf{p}$ de \mathbf{p} par rapport à \mathbf{T}_{ref} dans *l'approximation linéaire* : $D_{S \rightarrow S'} = \partial(\delta\mathbf{p}') / \partial(\delta\mathbf{p}) = \partial\mathbf{p}' / \partial\mathbf{p}$ (matrice jacobienne)
- appliquer le formalisme progressif ; modifier \mathbf{T}_{ref} et itérer si les $\delta\mathbf{p}$ sont trop grands (*on peut aussi changer \mathbf{T}_{ref} en cours de route*)

Le cas général: trajectoire 3D (5 paramètres)

Quels paramètres ?

ça dépend de la géométrie du détecteur (et du champ magnétique)

Exemples:

- pour une cible fixe ou un bouchon (endcap) de collisionneur:
plans de référence perpendiculaires au faisceau (à z fixé)
 - position (x, y)
 - direction (angles θ et ϕ) ou coef. directeurs (u_x, u_y)
 - courbure ($1/R$ ou q/p_t ou q/p) (*avec un signe*)
- pour un “barril” de collisionneur, avec \mathbf{B} suivant z :
cylindres de référence (tube à vide + couches concentriques) :
 - position (angle Φ , z)
 - direction (angles θ et ϕ)
 - courbure ($1/R$ ou q/p_t ou q/p)

mesures indirectes "obliques"

- on mesure une combinaison des paramètres, par exemple $u = \alpha x + \beta y$

$W_u = 1/\sigma_u^2$ ("poids" de la mesure de u)

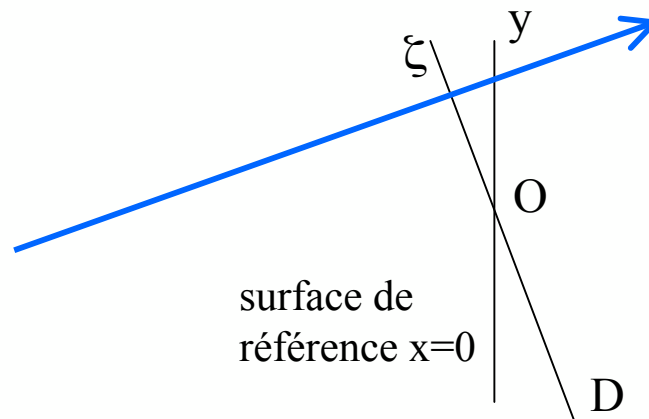
Contribution au χ^2 : "bande" dans le plan (x,y)

$$W_u (u^{\text{mes}} - \alpha x - \beta y)^2 = (x - x^{\text{mes}}, y - y^{\text{mes}})^t W (x - x^{\text{mes}}, y - y^{\text{mes}})$$

$x^{\text{mes}}, y^{\text{mes}}$: n'importe quel point tel que $\alpha x^{\text{mes}} + \beta y^{\text{mes}} = u^{\text{mes}}$

$W = (a, b) \cdot W_u \cdot (a, b)^t$ (matrice singulière)

on mesure dans un détecteur *oblique* par rapport à la surface de mesure



trajectoire de pente a

mesurer ζ (erreur σ) dans D

équivalent à mesurer $y = \zeta (\lambda + \mu a)$ avec

une erreur $|\lambda + \mu a| \cdot \sigma$

λ, μ : constantes géométriques

a : doit être déjà évalué

Formulation générale : $\chi^2 = (\mathbf{p} - \mathbf{p}^{\text{mes}})^t W_p (\mathbf{p} - \mathbf{p}^{\text{mes}})$ avec $W_p = M^t W_m M$

W_m : poids des mesures brutes \mathbf{m} ; M : dépendance de \mathbf{p} par rapport à \mathbf{m}

Mesures "exogènes"

On peut incorporer dans le fit des informations de détecteurs non trajectographiques

exemples:

- E (électron) mesuré dans un calorimètre injecté à l'extrémité de la trace
- ΔE (électron) injecté en un point intermédiaire (plus délicat...)

application à la reconnaissance des traces

on peut coupler la reconnaissance (associations de points à une trajectoire) et le fit progressif

principe de base : on associe de proche en proche et on incorpore les points dans le fit
plus on avance, plus l'extrapolation est précise → plus l'association est sélective

à chaque étape : on peut accepter des ambiguïtés, ou résoudre des ambiguïtés en cours

stratégie générale (point de départ, étapes, itérations):
dépend du détecteur

La vie n'est pas une pure gaussienne tranquille...

Deux sortes de “non-gaussianité”

- “à courte portée” (ex : distrib. uniforme sur (a,b))
se lisse par convolution (limite gaussienne pour les grands nombres)
- “avec longues queues”
peut faire échec à la limite gaussienne

en pratique, pour les particules chargées :

- non-linéarité dans la propagation → distortion de la p.d.f.
- diffusion multiple : faible probabilité d’une diffusion à grand angle (Rutherford)
- pertes d’énergie :
 - . ΔE par ionisation quasi déterministe, avec faibles fluctuations
 - . pertes occasionnelles plus violentes : δ -rays, et surtout bremsstrahlung (**problème majeur pour les électrons**)

approximation gaussienne non satisfaisante ; que faire ?

L'algorithme de Dieu

vecteur \mathbf{p} décrivant l'état de la particule sur une surface (dim. 5)

enchaînement *ad libitum* des opérations élémentaires sur $F(\mathbf{p})$:

- **mesure** (locale) : *multiplication* par $f^{\text{mes}}(m(\mathbf{p}))$
- **bruit** (local) : *convolution* avec $f^{\text{br}}(\mathbf{p})$
- **propagation** : *changement de variable* $F(\mathbf{p}) \rightarrow F^{\text{pr}}(\mathbf{p}^{\text{pr}}(\mathbf{p}))$

Problème : puissance de calcul pour des fonctions de 5 variables !

Mais : "On trouve avec le Ciel des accommodements" (*Tartuffe*)

La somme de gaussiennes

approximation de $F(\mathbf{p})$, f^{mes} et f^{br} par une gaussienne ou une *somme de gaussiennes*

$$F(\mathbf{p}) = \sum \alpha_i G_i(\mathbf{p}) \quad \text{avec} \quad G_i(\mathbf{p}) = C_i \exp\left(-(\mathbf{p}-\mathbf{p}_i)^t W_i (\mathbf{p}-\mathbf{p}_i) / 2\right)$$

- ça marche bien dans beaucoup de cas pour f^{mes} et f^{br} (une variable)
- fonction définie et positive partout si tous les α_i le sont
- les opérations (produits, convolutions, propagation linéaire) sont faciles et donnent encore des sommes de gaussiennes

$$\text{produit : } (\mathbf{p}_1, W_1) \times (\mathbf{p}_2, W_2) = ((W_1 + W_2)^{-1} (W_1 \mathbf{p}_1 + W_2 \mathbf{p}_2), W_1 + W_2)$$

$$\text{convolution : } (\mathbf{p}_1, W_1) * (\mathbf{p}_2, W_2) = (\mathbf{p}_1 + \mathbf{p}_2, (W_1^{-1} + W_2^{-1})^{-1})$$

$$\text{propagation : } \delta \mathbf{p}' = D \cdot \delta \mathbf{p} \rightarrow W' = (D^{-1})^t \cdot W \cdot D^{-1}$$

Mais : le nombre de composantes augmente multiplicativement
remèdes proposés:

- suppression des composantes de faible poids
 - fusion de composantes proches en une seule
- *compromis à optimiser au cas par cas, en fonction de l'utilisation finale pour la physique (interprétation des queues de distribution)*

le fit de vertex

A partir d'un ensemble de mesures de n trajectoires (\mathbf{q}_i, W_i) sur des surfaces de référence ($5n$ paramètres) on veut fitter ***simultanément*** $3n+3$ paramètres avec la contrainte de *convergence* de ces trajectoires :

- la position $\mathbf{V}(X,Y,Z)$ d'un vertex d'origine *commun*
 - les impulsions \mathbf{p}_i des particules *en ce point* (3 param. par particule, par ex. $1/p_i, \theta_i, \phi_i$)
- outil : fonction de propagation $\mathbf{q} = \mathbf{F}(\mathbf{V}, \mathbf{p})$

formulation (dans l'approximation du χ^2) :

trouver \mathbf{V} et les \mathbf{p}_i qui minimisent

$$\chi^2 = \sum (\mathbf{q}_i^{\text{mes}} - \mathbf{F}(\mathbf{V}, \mathbf{p}_i))^t W_i (\mathbf{q}_i^{\text{mes}} - \mathbf{F}(\mathbf{V}, \mathbf{p}_i))$$

a priori : minimum dans un espace de dimension $3n+3$

en fait : problème *hiérarchisé*

$$\min_{\mathbf{V}} \left[\sum \min_{\mathbf{p}_i} (\mathbf{q}_i^{\text{mes}} - \mathbf{F}(\mathbf{V}, \mathbf{p}_i))^t W_i (\mathbf{q}_i^{\text{mes}} - \mathbf{F}(\mathbf{V}, \mathbf{p}_i)) \right]$$

les minimisations “interne” et “externe” sont en dimension 3
(l’“emboîtement” reste valable sans l’hypothèse gaussienne))

Cadre plus général : fit "hiérarchique" deux exemples

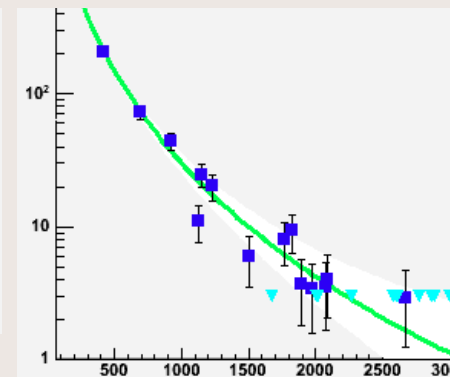
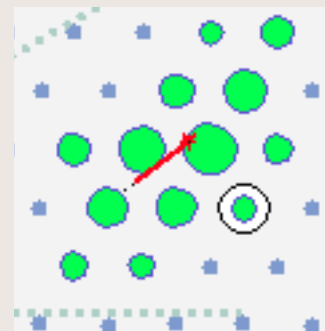
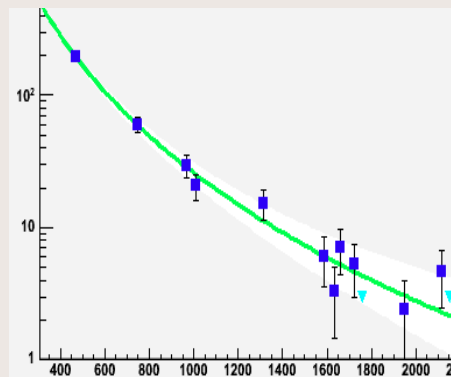
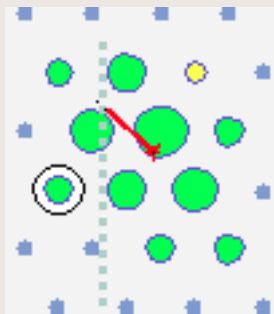
ensemble de signaux de même forme, mais d'amplitudes et de dates variables : $A_i f(t-a_i)$;
chacun est mesuré à n temps $t_k \rightarrow S_{ik}^{\text{mes}} = A_{ik} f(t_k - a_i) + \varepsilon^{\text{mes}}$
la *forme* est définie par p_1, p_2, \dots (paramètres globaux) qu'on veut ajuster
(exemple $f(t) = 0$ pour $t < 0$, $\exp(-p_1 t) - \exp(-p_2 t)$ pour $t > 0$)

un ensemble d'événements du Détecteur de Surface d'AUGER

pour une cuve à la distance r_i de l'axe de la gerbe signal $S_i = A_i f(r_i)$

- paramètres *globaux* (forme) $f(r) = 1/r^p(r+r_1)^q$

- paramètres *particuliers* à chaque événement (gerbe atmosphérique) : position (x_i, y_i) du coeur, et amplitude A_i



linéarisation

au voisinage des valeurs approchées \mathbf{V}_0 (vertex) et \mathbf{p}_{i0} (pour chaque trace) :

$$\mathbf{q}_i = \mathbf{F}_i(\mathbf{V}, \mathbf{p}_i) = \mathbf{q}_{i0} + \mathbf{D}_i \cdot (\mathbf{V} - \mathbf{V}_0) + \mathbf{E}_i \cdot (\mathbf{p}_i - \mathbf{p}_{i0})$$

(\mathbf{E} et \mathbf{D} : matrices (5×3))

avec $\Delta \mathbf{q}_i = \mathbf{q}_i^{\text{mes}} - \mathbf{q}_{i0}$, on fitte $\delta \mathbf{V} = \mathbf{V} - \mathbf{V}_0$ et les $\delta \mathbf{p}_i = \mathbf{p}_i - \mathbf{p}_{i0}$ pour minimiser

$$\chi^2 = \sum (\Delta \mathbf{q}_i - \mathbf{D}_i \delta \mathbf{V} - \mathbf{E}_i \delta \mathbf{p}_i)^t \mathbf{W}_i (\Delta \mathbf{q}_i - \mathbf{D}_i \delta \mathbf{V} - \mathbf{E}_i \delta \mathbf{p}_i)$$

un bloc de 3 équations sur l'ensemble des paramètres:

$$\mathbf{A} \delta \mathbf{V} + \sum \mathbf{B}_i \delta \mathbf{p}_i = \mathbf{T} \quad (1)$$

avec $\mathbf{A} = \sum \mathbf{D}_i^t \mathbf{W}_i \mathbf{D}_i$, $\mathbf{B}_i = \mathbf{D}_i^t \mathbf{W}_i \mathbf{E}_i$, $\mathbf{T} = \sum \mathbf{D}_i^t \mathbf{W}_i \Delta \mathbf{q}_i$

n blocs de 3 équations sur \mathbf{V} et *un* \mathbf{p}_i :

$$\mathbf{B}_i^t \delta \mathbf{V} + \mathbf{C}_i \delta \mathbf{p}_i = \mathbf{U}_i \quad (2)$$

avec $\mathbf{C}_i = \mathbf{E}_i^t \mathbf{W}_i \mathbf{E}_i$, $\mathbf{U} = \sum \mathbf{E}_i^t \mathbf{W}_i \Delta \mathbf{q}_i$

résolution "par blocs"

à partir des équations (2) on exprime les $\delta \mathbf{p}_i$ en fonction de $\delta \mathbf{V}$

$$\delta \mathbf{p}_i = \mathbf{C}_i^{-1} (\mathbf{U}_i - \mathbf{B}_i^t \delta \mathbf{V}) \quad (3)$$

et on injecte ces expressions dans (1) pour obtenir une équation en $\delta \mathbf{V}$

$$(\mathbf{A} - \sum \mathbf{B}_i \mathbf{C}_i^{-1} \mathbf{B}_i^t) \delta \mathbf{V} = \mathbf{T} - \sum \mathbf{B}_i \mathbf{C}_i^{-1} \mathbf{U}_i \quad (4)$$

(4) donne $\delta \mathbf{V}$ puis chacune des éq. (3) donne $\delta \mathbf{p}_i$

en prime, la matrice $(3n+3) \times (3n+3)$ de covariance ...

$$\text{cov}(\mathbf{V}, \mathbf{V}) = (\mathbf{A} - \sum \mathbf{B}_i \mathbf{C}_i^{-1} \mathbf{B}_i^t)^{-1}$$

$$\text{cov}(\mathbf{V}, \mathbf{p}_i) = - \text{cov}(\mathbf{V}, \mathbf{V}) \mathbf{B}_i \mathbf{C}_i^{-1}$$

$$\text{cov}(\mathbf{p}_i, \mathbf{p}_j) = \delta_{ij} \mathbf{C}_i^{-1} + \mathbf{C}_i^{-1} \mathbf{B}_i^t \text{cov}(\mathbf{V}, \mathbf{V}) \mathbf{B}_j \mathbf{C}_j^{-1}$$

flexibilité

(ajout ou suppression d'une particule)

on veut ajouter une trajectoire (mesure \mathbf{q}_{n+1} , \mathbf{W}_{n+1})

on ajoute un triplet de paramètres $\delta \mathbf{p}_{n+1}$

dans (1) on ajoute $\mathbf{D}_{n+1}^t \mathbf{W}_{n+1} \mathbf{D}_{n+1}$ à \mathbf{A} , et un terme $\mathbf{B}_{n+1} = \mathbf{D}_{n+1}^t \mathbf{W}_{n+1} \mathbf{E}_{n+1}$

dans (2) on ajoute un bloc d'équations $\mathbf{B}_{n+1}^t \delta \mathbf{V} + \mathbf{C}_{n+1} \delta \mathbf{p}_{n+1} = \mathbf{U}_{n+1}$

si on prend comme **valeurs de départ** (\mathbf{V}_0 , \mathbf{p}_{i0} pour $i=1 \dots n$) le résultat du fit **précédent** (avec n particules) :

$$(\mathbf{A} + \mathbf{A}_{n+1}) \delta \mathbf{V} + \sum \mathbf{B}_i \delta \mathbf{p}_i = \mathbf{T}_{n+1}$$

$$\mathbf{B}_i^t \delta \mathbf{V} + \mathbf{C}_i \delta \mathbf{p}_i = \mathbf{0} \text{ pour } i=1 \dots n$$

$$\mathbf{B}_{n+1}^t \delta \mathbf{V} + \mathbf{C}_{n+1} \delta \mathbf{p}_{n+1} = \mathbf{U}_{n+1}$$

résolution :

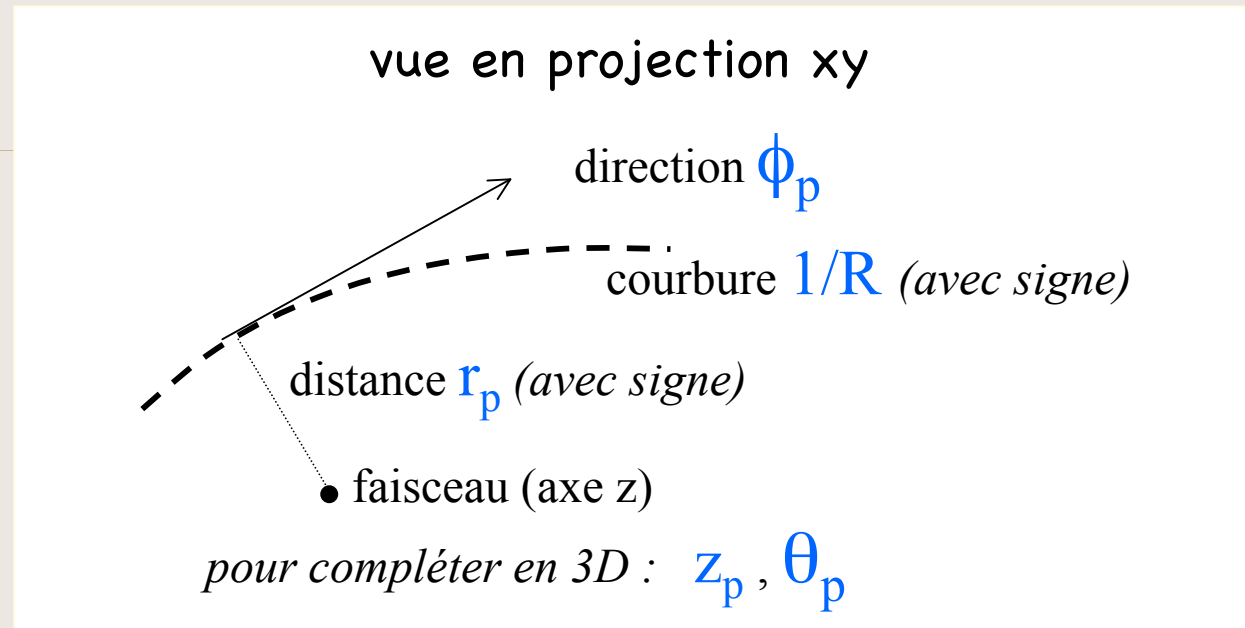
$$(\mathbf{A} - \sum \mathbf{B}_i \mathbf{C}_i^{-1} \mathbf{B}_i^t + \mathbf{A}_{n+1} - \mathbf{B}_{n+1} \mathbf{C}_{n+1}^{-1} \mathbf{B}_{n+1}^t) \delta \mathbf{V} = \mathbf{T}_{n+1} - \mathbf{B}_{n+1} \mathbf{C}_{n+1}^{-1} \mathbf{U}_{n+1}$$

(on ne calcule que les termes en **rouge** : procédure rapide \rightarrow on peut essayer beaucoup de combinaisons)

suppression d'une trajectoire = addition avec un poids (-W)

Remarque : le faisceau peut être considéré comme une trajectoire incluse dans le vertex principal (mesure très précise de x, y , mais z indéterminé)

Les paramètres de "périgée"



Avantages :

- propagation depuis une surface de référence sans singularité pour la matrice D
(grâce au signe géométrique donné à r : + si la trajectoire passe à gauche de l'axe z)
- propagation à courte distance vers le vertex : approximation linéaire toujours valide
- les paramètres au périgée ont un sens physique

une bonne stratégie : extrapoler toutes les traces à leur périgée (une fois pour toutes) puis faire itérativement le fit du (ou des) vertex

vertex avec contraintes

V^0 (désintégration neutre $\rightarrow + -$) : on peut imposer la convergence des trajectoires **et** une hypothèse de masse ; éventuellement la contrainte que \mathbf{p} pointe vers le vertex principal

$\gamma \rightarrow e^+ e^-$: on impose la convergence **et** le parallélisme des trajectoires ; éventuellement la contrainte que \mathbf{p} pointe vers le vertex principal

multiplicateurs de Lagrange: outil généraliste

$\min_{\mathbf{p}} (F(\mathbf{p}))$ avec la contrainte $C(\mathbf{p}) = 0 \Leftrightarrow \min_{\mathbf{p}, \lambda} (F(\mathbf{p}) + \lambda C(\mathbf{p}))$

commode dans l'approximation suivante (couramment valable) :

- la log-vraisemblance est quadratique autour du minimum
- la contrainte est linéarisable autour du minimum

dans ce cas : système linéaire $\rightarrow \mathbf{p}$ en fonction de λ , puis élimination linéaire de λ

généralisation évidente à plusieurs contraintes

$\min_{\mathbf{p}, \lambda_1, \lambda_2, \dots} (F(\mathbf{p}) + \lambda_1 C_1(\mathbf{p}) + \lambda_2 C_2(\mathbf{p}) + \dots)$

En résumé...

- on a défini des opérations élémentaires sur les paramètres de trajectoire (ajout d'une mesure, ajout d'un bruit, propagation)
- si on les enchaîne dans l'ordre, on travaille à chaque étape sur des quantités indépendantes
- dans l'approximation linéaire (presque toujours applicable dans les cas utiles), les opérations sont des manipulations simples de 5-vecteurs et de matrices (5x5)
- dans le cas gaussien (ou quasi-gaussien) on peut obtenir des tests de qualité globaux ou intermédiaires sous forme de $\text{Prob}(\chi^2)$
- on peut tenir compte d'effets non gaussiens si nécessaire
- on peut coupler l'ajustement des trajectoires au "pattern recognition" (grande variété de stratégies possibles)
- de même, dans l'ajustement de vertex, on peut procéder par "essais et erreurs" (ajouter ou retirer une trace est peu coûteux), ajouter des contraintes