DATA ANALYSE

DETECTOR
PART I

TRIGGER / DAQ
ε

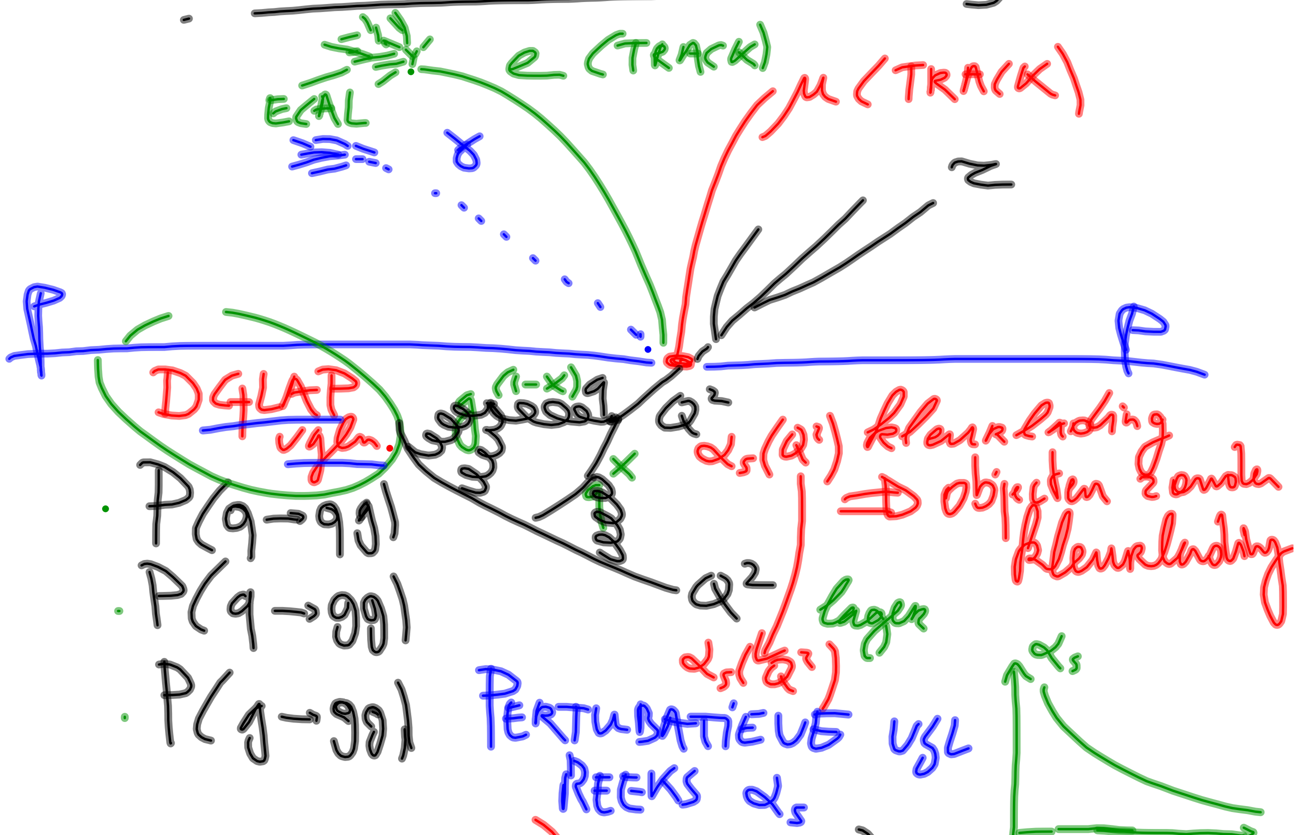
MOTIVATIE
SM / BSM

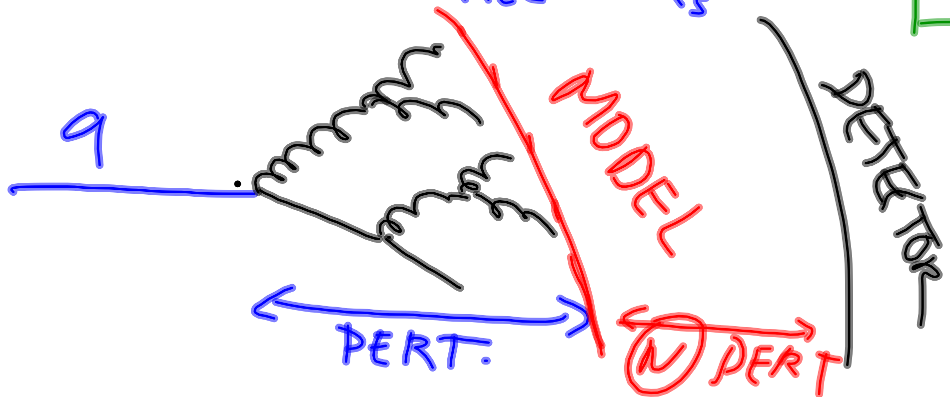
RECONSTRUCTIE
(TRACK)

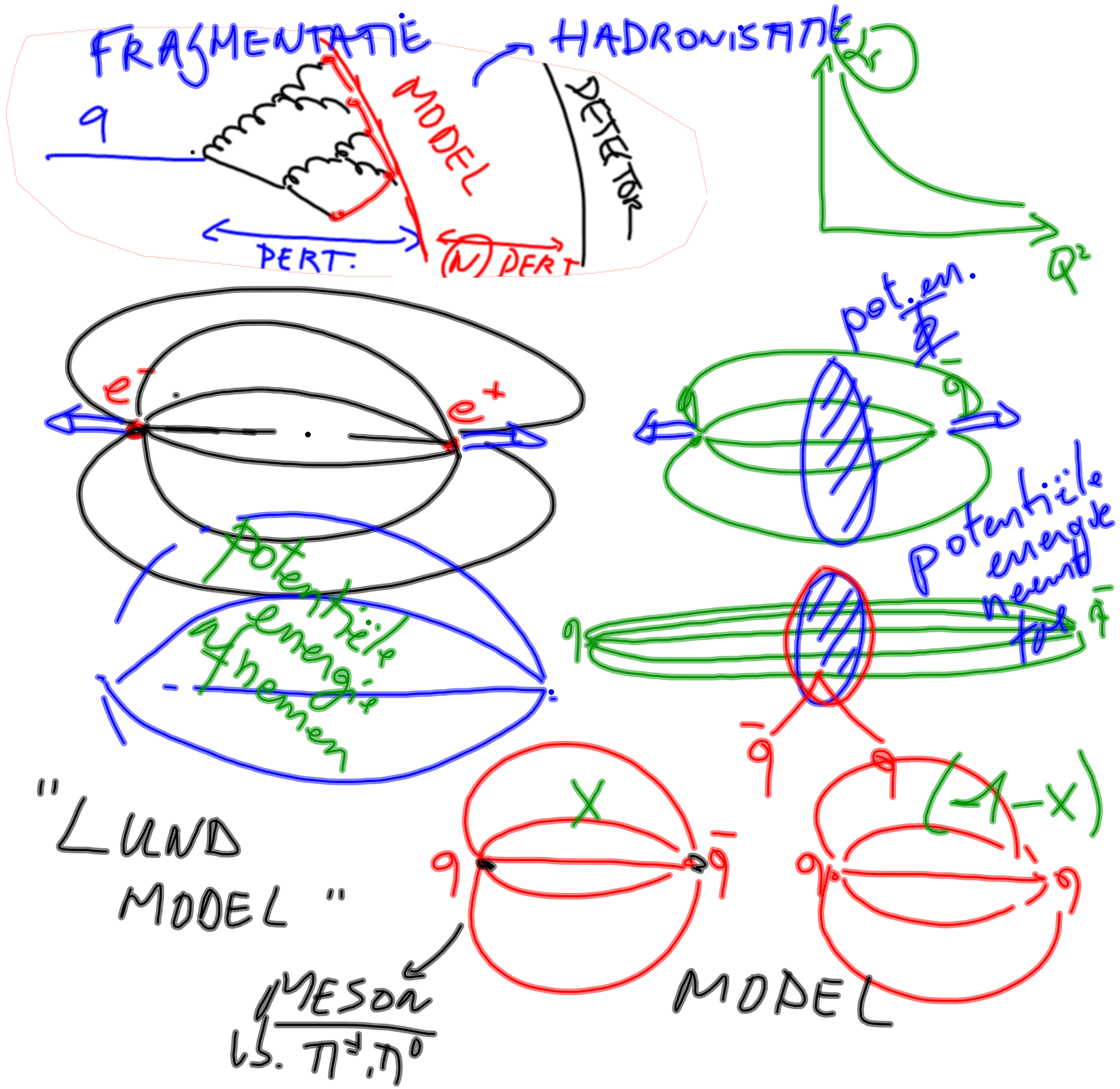
DATA ANALYSE

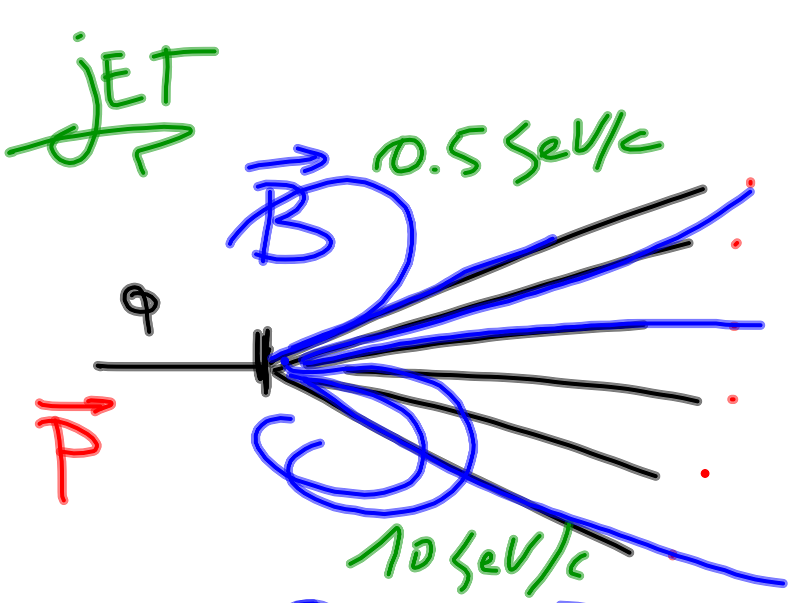
METEN / ONTDEKKEN (Uitsluiten)

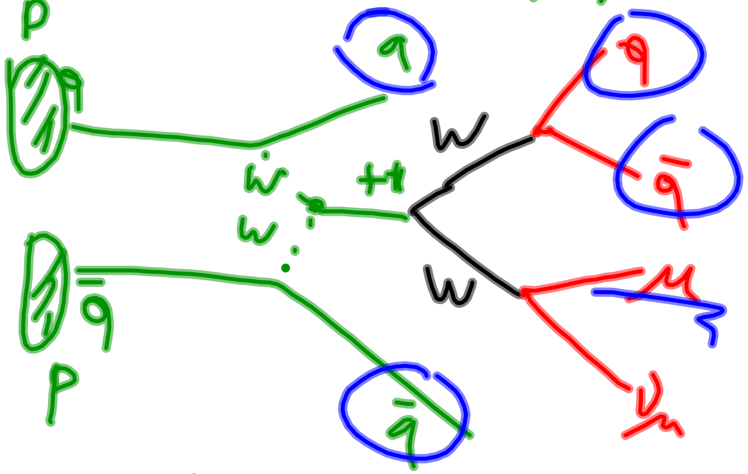
RECONSTRUCTIE V.D. OBJECTEN






FRAGMENTATIE
HADRONISATIE
q
MODEL
DETECTOR
PERT.
(N) PERT
Q²
e⁻
e⁺
pot. en.
potentiële energie
potentiële energie neemt toe
potentiële energie afnemen
q
q̄
"LUND MODEL"
X
(1−x)
MESON
b.v. π±, π⁰
MODEL





$m_W \approx 80\ GeV/c^2$

4-momentum v.d. quarks

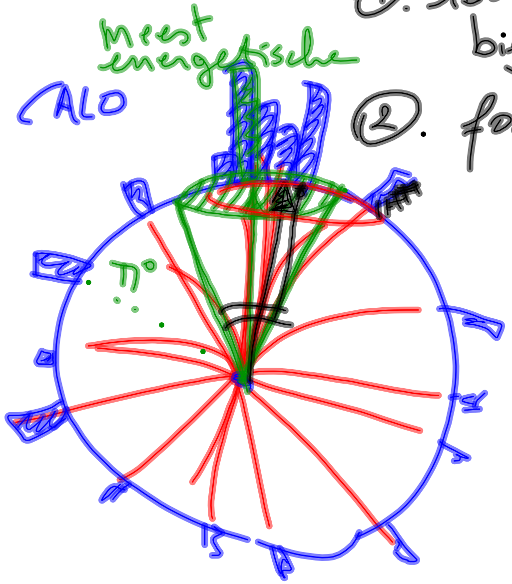
fragmenten $\Longrightarrow$ quark

Reconstrueren

$$\left\{\vec{p}_i\right\} \longrightarrow \vec{p}_{quark}$$

① identificeren welk deeltje bij welke quark

② formule $\vec{p}_{quark}$



$\sum_{i \in \text{cone}} \vec{p}_i \to \boxed{\vec{p}_{jet}}$

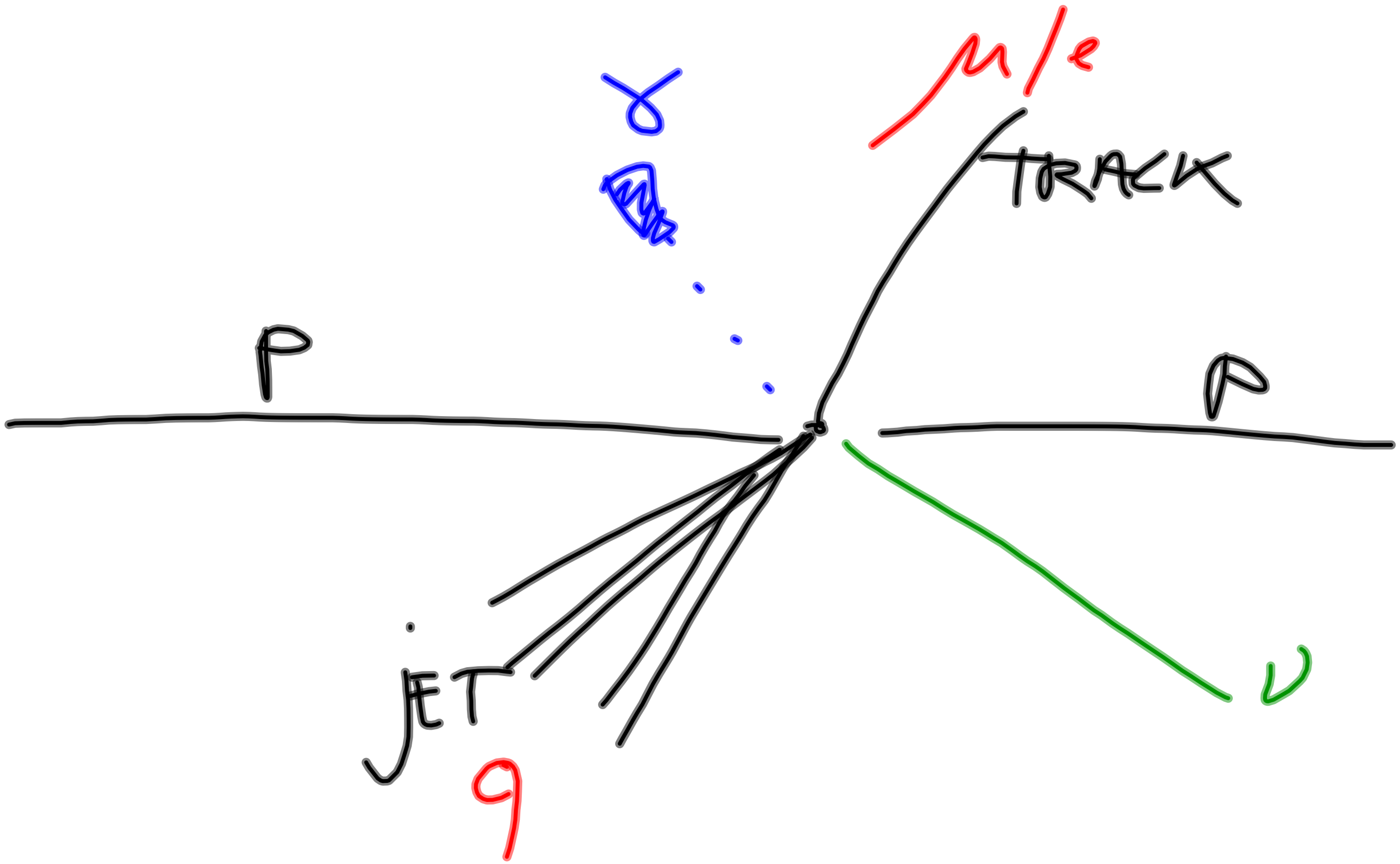
Richting v.d. "jet" verschillend van de richting van

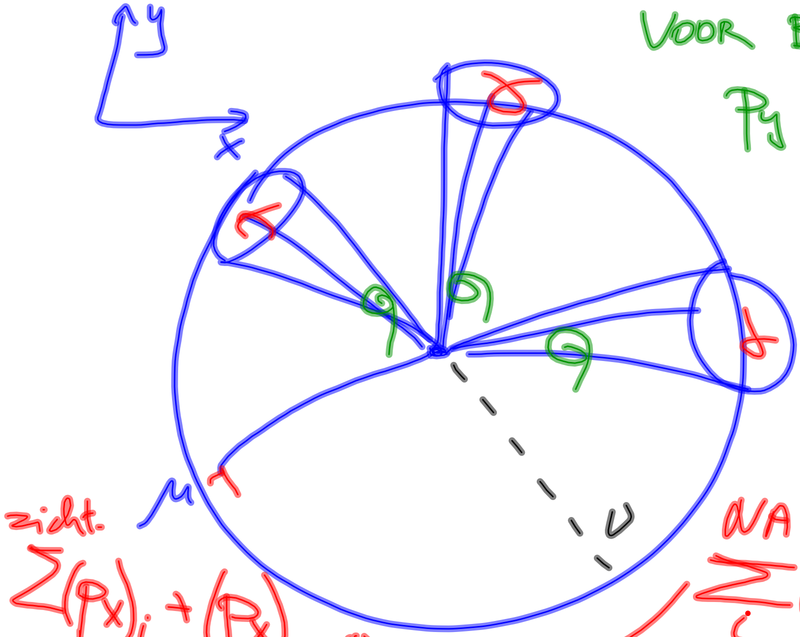
iteratief:
$\Rightarrow$ TOT $(\vec{p}_{jet})_N = (\vec{p}_{jet}$

vb. jet reconstructie algorithme
"cone algorithme"

openingshoek

TAAK: bestuderen die jet algoritmen.
$\longrightarrow$ schrijf er een
~ 5 p. samenva


μ/e
γ
TRACK
P
P
JET
g
ν



VOOR BOTSING:

$p_y = p_x = 0$

NA BOTSING

$\sum_i (p_x)_i = 0$

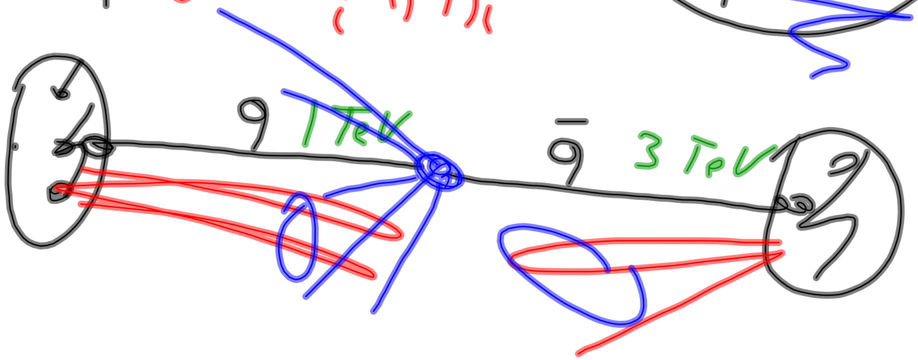
$\sum_i (p_y)_i = 0$

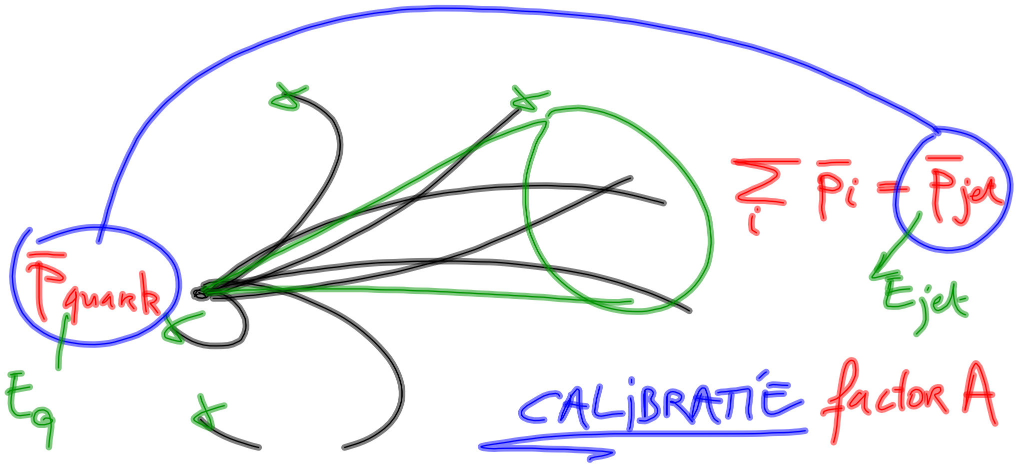
zicht.

$\sum (p_x)_i + (p_x)_{\text{onzicht.}} = 0$

$\Rightarrow \begin{cases} p_x^\nu = -\sum\limits_{i\ \text{zicht.}} (p_x)_i \\ p_y^\nu = -\sum\limits_i (p_y)_i \end{cases}$

$\cancel{p_z^\nu = ?}$





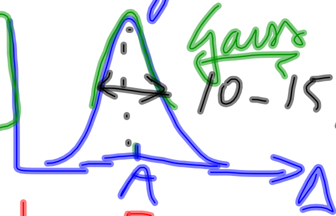
CALIBRATIE factor A

$$\frac{A\cdot E_{jet}}{E_{quark}} \neq 1 \Rightarrow \left\langle \frac{A\cdot E_{jet}}{E_q} \right\rangle = 1$$

SIMULATIE $\longrightarrow$ E_{jet} vs. E_q

$\Delta = E_{jet} / E_q \longrightarrow$ verdeling

$$\cancel{A}\ \frac{1}{A} = \langle E_{jet}/E_q \rangle = E\left[E_{jet}/E_q\right]$$



$$\Longrightarrow E_{jet} \xrightarrow{\text{CALIBR.}} \frac{1}{A}\cdot E_{jet}$$

$A = A(\text{eign. jet})$

p_T, θ

RECO → BOTSING

Signaal / Bck

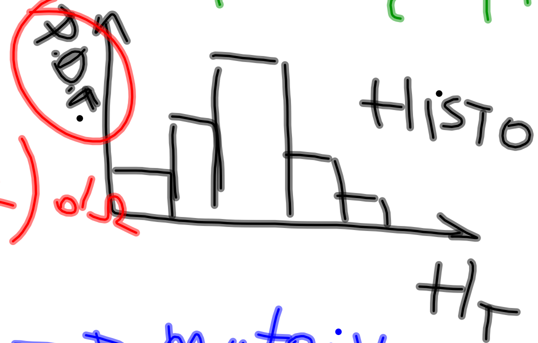
SIMULATIE

THEORIE → signaal $\vec{\Omega}$ observables (eign. v. event)

parameters $\vec{\alpha}$

vb. $H_T \equiv \sum_i (p_T)_i$

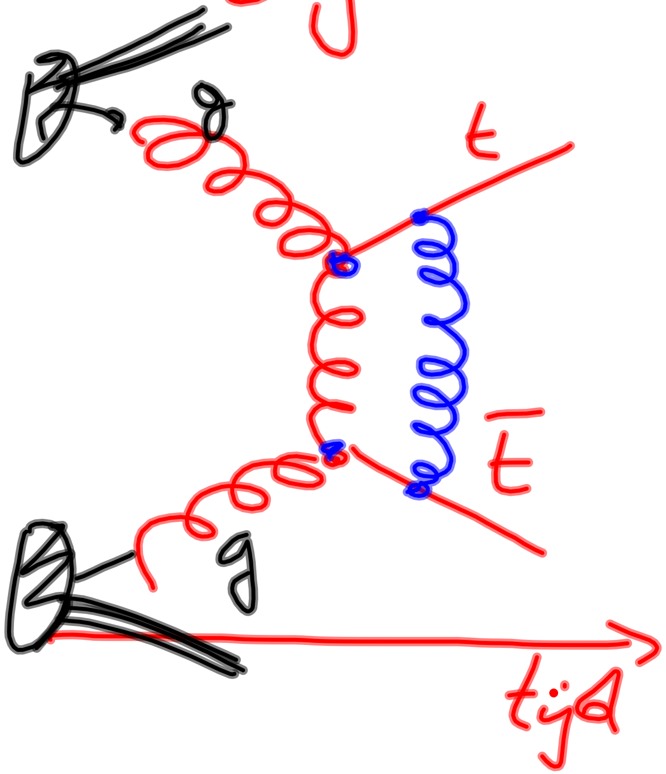
$f(\Omega|\vec{\alpha})$ $1 = \int_\Omega f(\Omega)\, d\Omega$



Q.F.T. (Feynman) → matrix element

kans ~ $|\mathcal{M}(\Omega|\vec{\alpha})|^2$

Signal: $(pp \rightarrow t\bar{t})$



$+ \; \ldots \; \underbrace{\left\{ \text{Feynman diagr.} \right\}}$

$$\left| \mathcal{M}(\Omega | \vec{\alpha}) \right|^2$$

THEORIE

vs. H_T

$$P(\Omega) = \int d\Omega' \, R(\Omega, \Omega' | \vec{\alpha}') \cdot | M(\Omega' | \vec{\alpha}) |^2$$

Response functie.

stochastisch

parameters detector

parameters theorie.

observatie

$H_T = 200\,\text{GeV}$

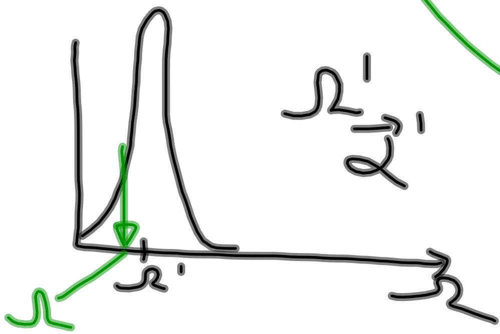
$0 < H_T < \infty$

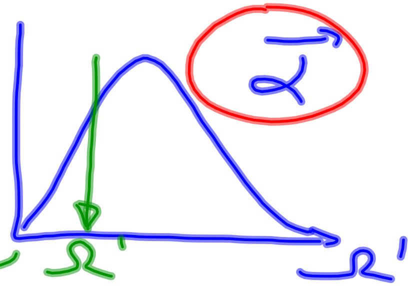
$$E[\Omega] = \int d\Omega \, \Omega \cdot P(\Omega)$$

$E[\Omega](\vec{\alpha}, \vec{\alpha}')$

verw. v. H_T voor signaal gebeurtenis

$$R(\Omega, \Omega' | \vec{\alpha}') \qquad |M(\Omega' | \vec{\alpha})|^2$$





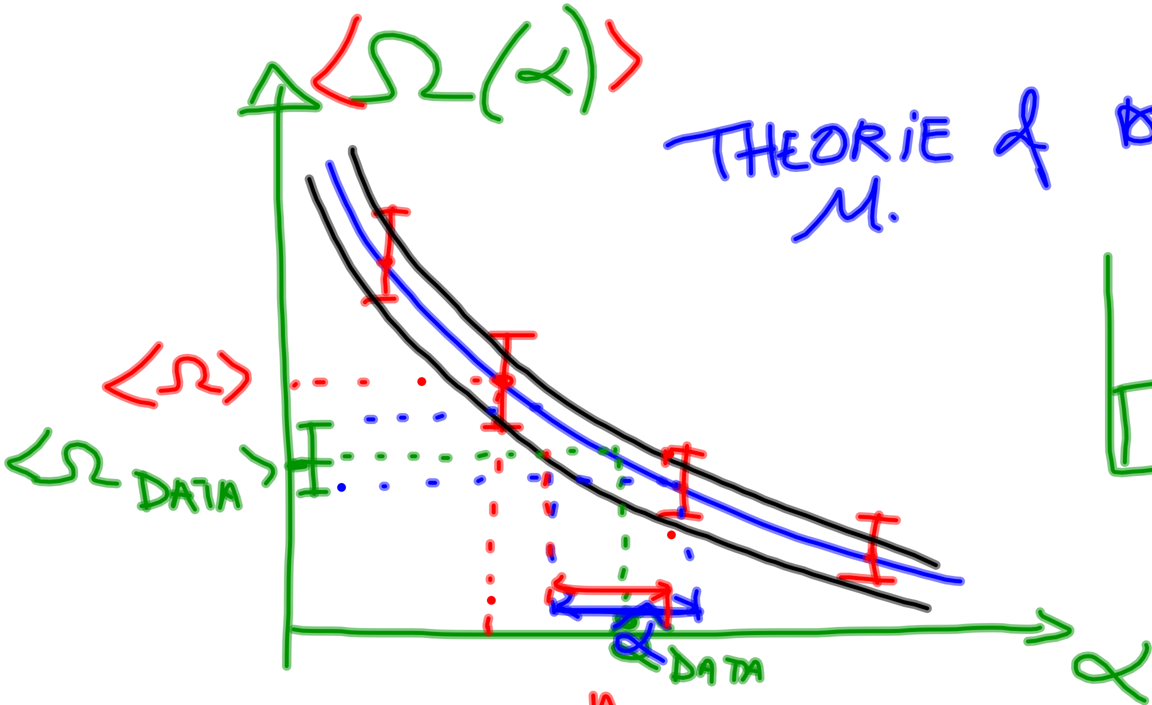
Monte Carlo techn.
(stochastisch)
Random getallen

$$\Longrightarrow \underset{i}{\text{event}} \Longrightarrow \Omega_i$$

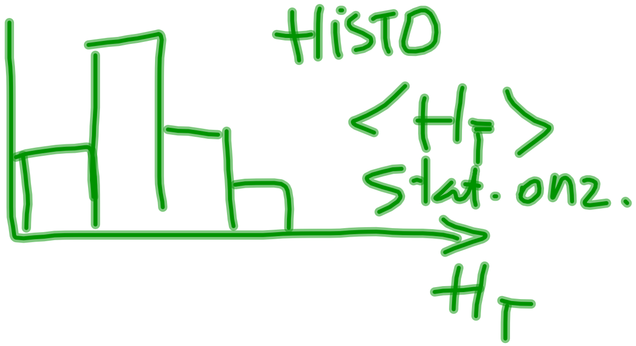
$$\xrightarrow{n} \{\Omega_i\} \quad \vec{\alpha}$$

$$\boxed{\langle \Omega \rangle} = \frac{1}{n} \sum_{i=1}^{n} \Omega_i(\vec{\alpha})$$

n events gesimuleerd $\Longrightarrow \langle \Omega \rangle (\vec{\alpha})$
$\Longrightarrow \langle \Omega \rangle (\alpha)$



THEORIE & DETECTOR
M. R

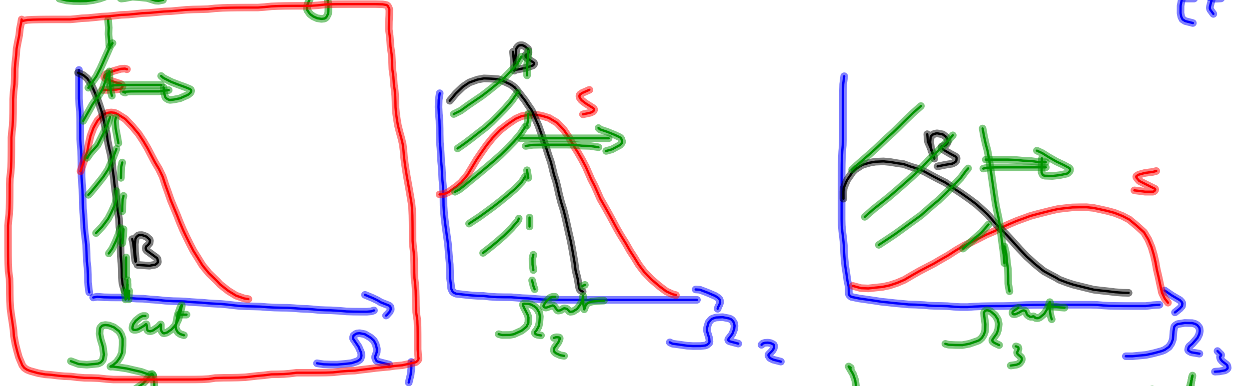


$$\langle \Omega \rangle = \underbrace{\frac{1}{n}\sum_{i=1}^{n}}_{n \text{ event gesim.}} \Omega_i(\alpha) \rightsquigarrow \text{stat. onz.}$$

Signal / Bck

$pp \to t\bar{t}$ {verz.} → complement $pp \to X$ (maar niet $t\bar{t}$)

$\vec{\Omega}$: eign. v. events. (simulatie)



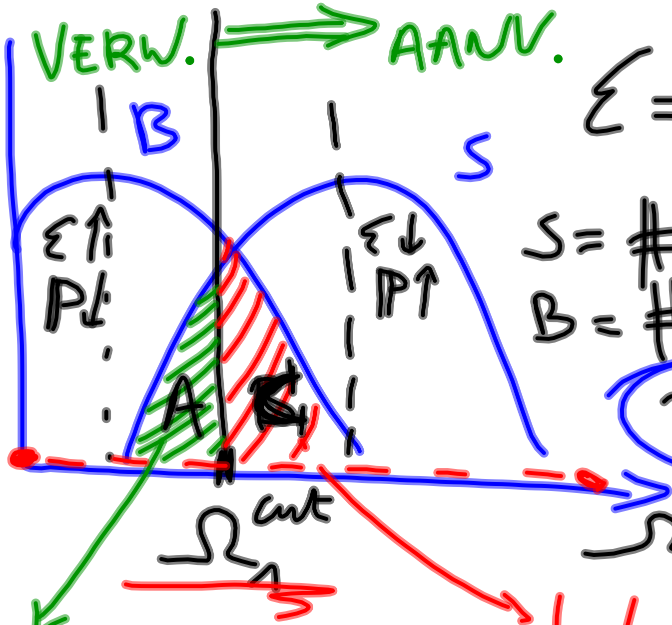
event: $\vec{\Omega} = (\Omega_1, \Omega_2, \Omega_3) \to \begin{cases} \Omega_1 > \Omega_1^{cut} \\ \Omega_2 > \Omega_2^{cut} \\ \Omega_3 > \Omega_3^{cut} \end{cases}$

Hypothese toets

toetsgrootheid Ω_1

aanvaard $\Omega_1 > \Omega_1^{cut}$

verwerp $\Omega_1 < \Omega_1^{cut}$



$$\varepsilon = \mathrm{Eff} = \frac{S_{sel}}{S} = \frac{S-A}{S}$$

$S = \#$ sign.

$B = \#$ bck

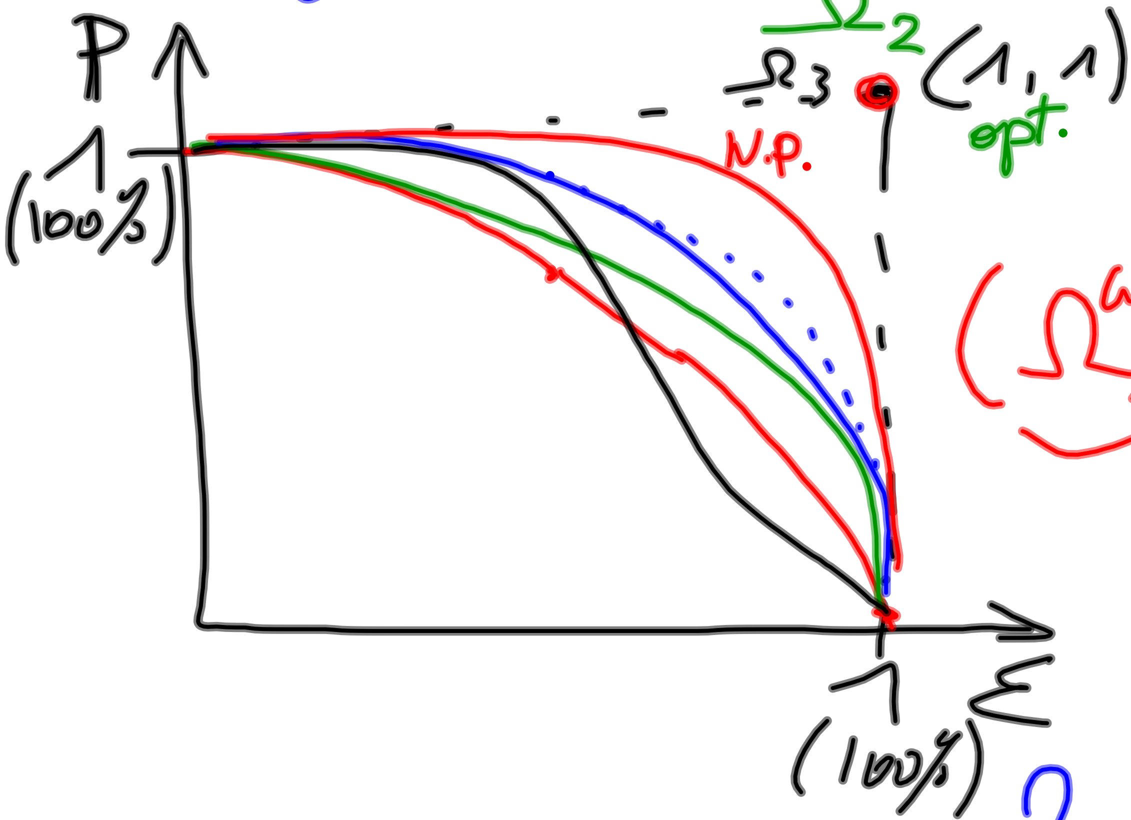
$$\text{purity} = \mathbb{P} = \frac{S_{sel}}{S_{sel}+B_{sel}} = \frac{S-A}{(S-A)+(C)}$$

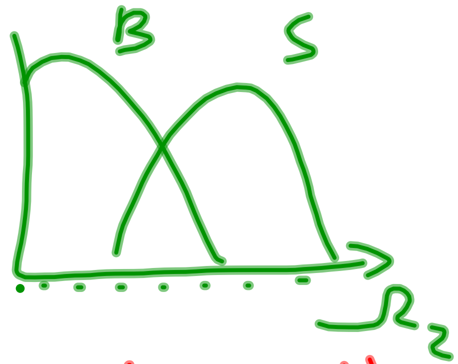
verwerpen maar toch signaal (fout type I)

bck die je toch aanv. (fout type II)

$\Longrightarrow$ zo weinig mogelijk fouten

Ω_1^{cut} vast $\Longrightarrow$ $(\varepsilon, \mathbb{P})$ vast observabele Ω_1





$$(\Omega_1^{cut}, \Omega_2^{cut}, \Omega_3^{cut})$$

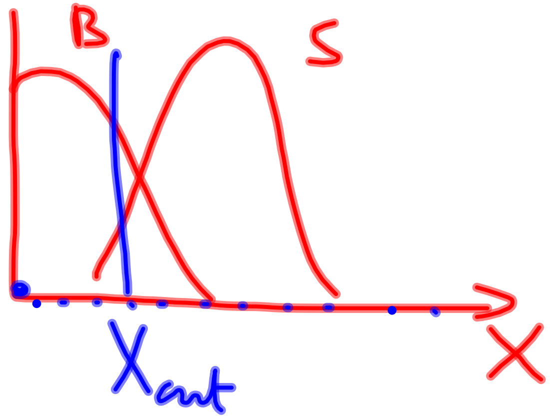
$\Longrightarrow$ combineren ?

$$\vec{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$$

$$\vec{\Omega} = (\Omega_1, \Omega_2, \Omega_3) \quad \text{combineren}$$

$$X = \Omega_1 + (\Omega_2)^3 + \sqrt{\Omega_3} + \Omega_1 \cdot \Omega_2$$

obs.



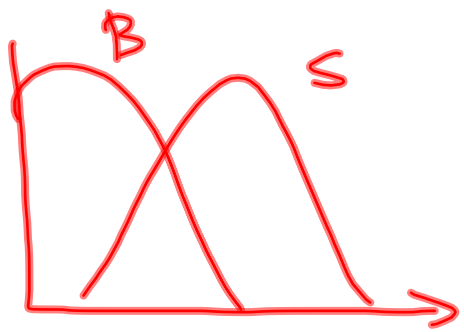
? optimale combinatie

Neyman - Pearson hypothese toets

$\{\Omega_i\}$: eign. $\Rightarrow \vec{\Omega}$

$$\lambda_j = \frac{PDF_S(\vec{\Omega}_j)}{PDF_B(\vec{\Omega}_j)}$$

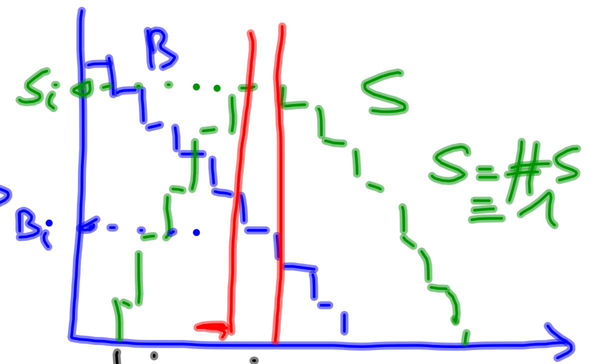
j : event



P.D.F. $\quad \Omega_1$

$$\int d\Omega_1 \, f(\Omega_1) = 1$$

bins. $\Longrightarrow$



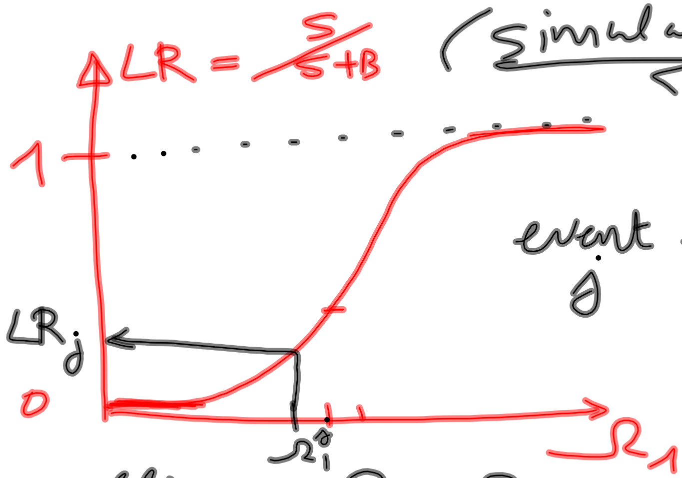
bin i $\quad \Omega_1$

$$P_i(B) = B_i/B = B_i \qquad P_i(S) = S_i/S = S_i$$

$\Longrightarrow$ event j in bin i wat is de kans dat event j signaal is

Likelihood Ratio $\longrightarrow$ $\boxed{LR_i = \dfrac{S_i}{S_i + B_i}} \sim 2/3$

$$\frac{S_i}{S_i + B_i} = \frac{1}{\frac{S_i + B_i}{S_i}} = \frac{1}{1 + \left(\frac{B_i}{S_i}\right)}$$



(simulatie)

event $\vec{\Omega} \rightarrow \Omega_1$
j

Zelfde voor Ω_2, Ω_3

$\Rightarrow \boxed{LR(\Omega_1)_j}, LR(\Omega_2)_j, LR(\Omega_3)_j$

voor elk event j

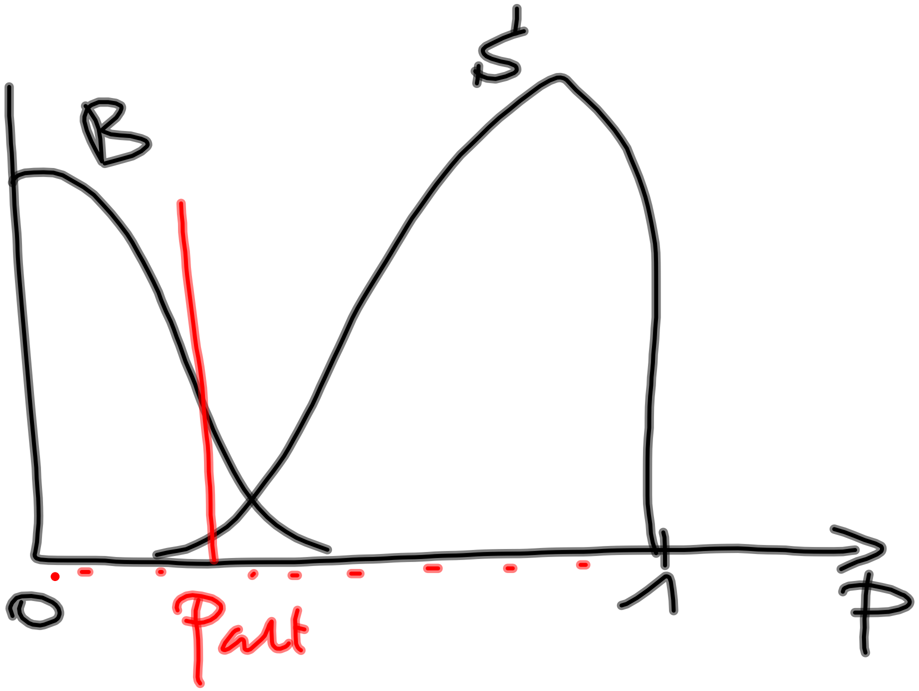
kans op S'

$$P(\vec{\Omega}) = P(\Omega_1) * P(\Omega_2) * P(\Omega_3)$$

PDF($\vec{\Omega}$) om sign te zijn

Neyman-Pearson $\Rightarrow$ optimale toetsgrootheid

! indien Ω_1 Ω_2 Ω_3 niet gecorreleerd zijn !

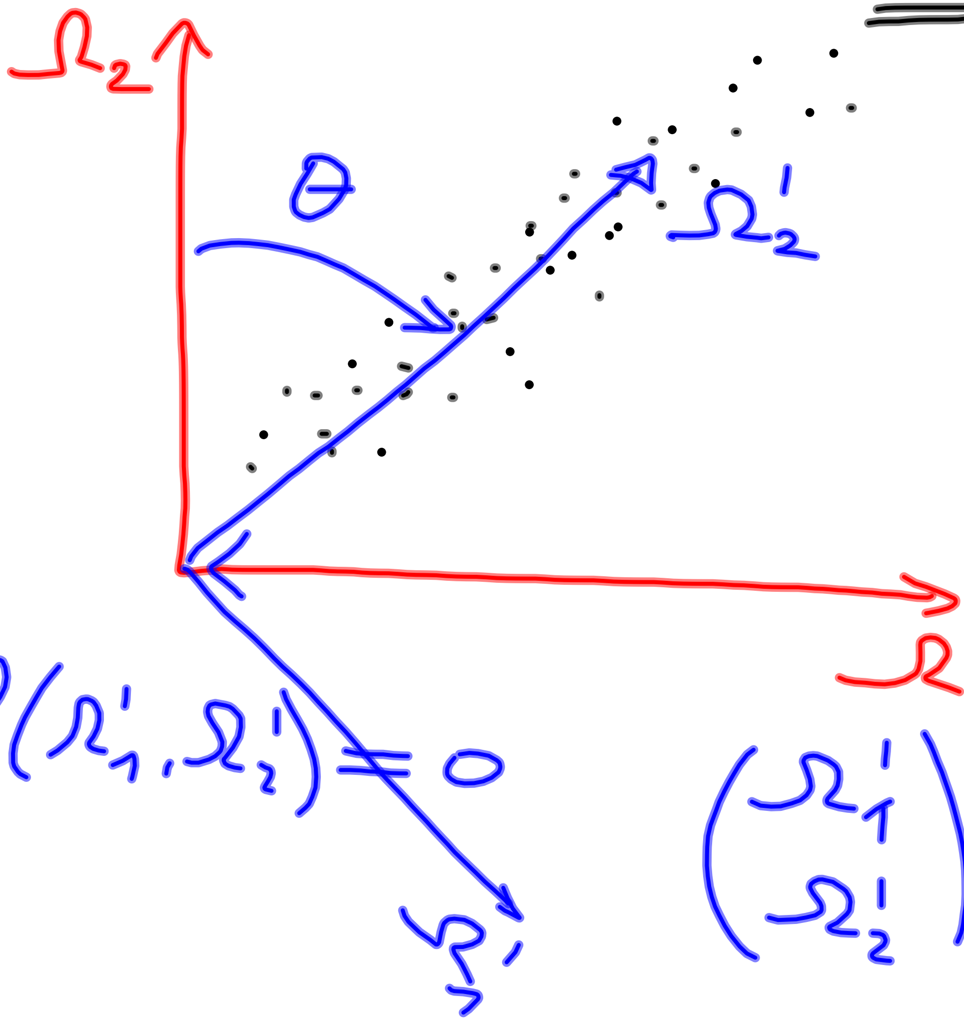

B
S
0
Part
1
P

signal $\}i\}$

$\Longrightarrow$ LR methode

Ω_1' & Ω_2'

$\Longrightarrow$ optimaal



lin. corr. tussen

Ω_1 & Ω_2

$\rho(\Omega_1', \Omega_2') \neq 0$

$$\begin{pmatrix} \Omega_1' \\ \Omega_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \Omega_1 \\ \Omega_2 \end{pmatrix}$$

methode

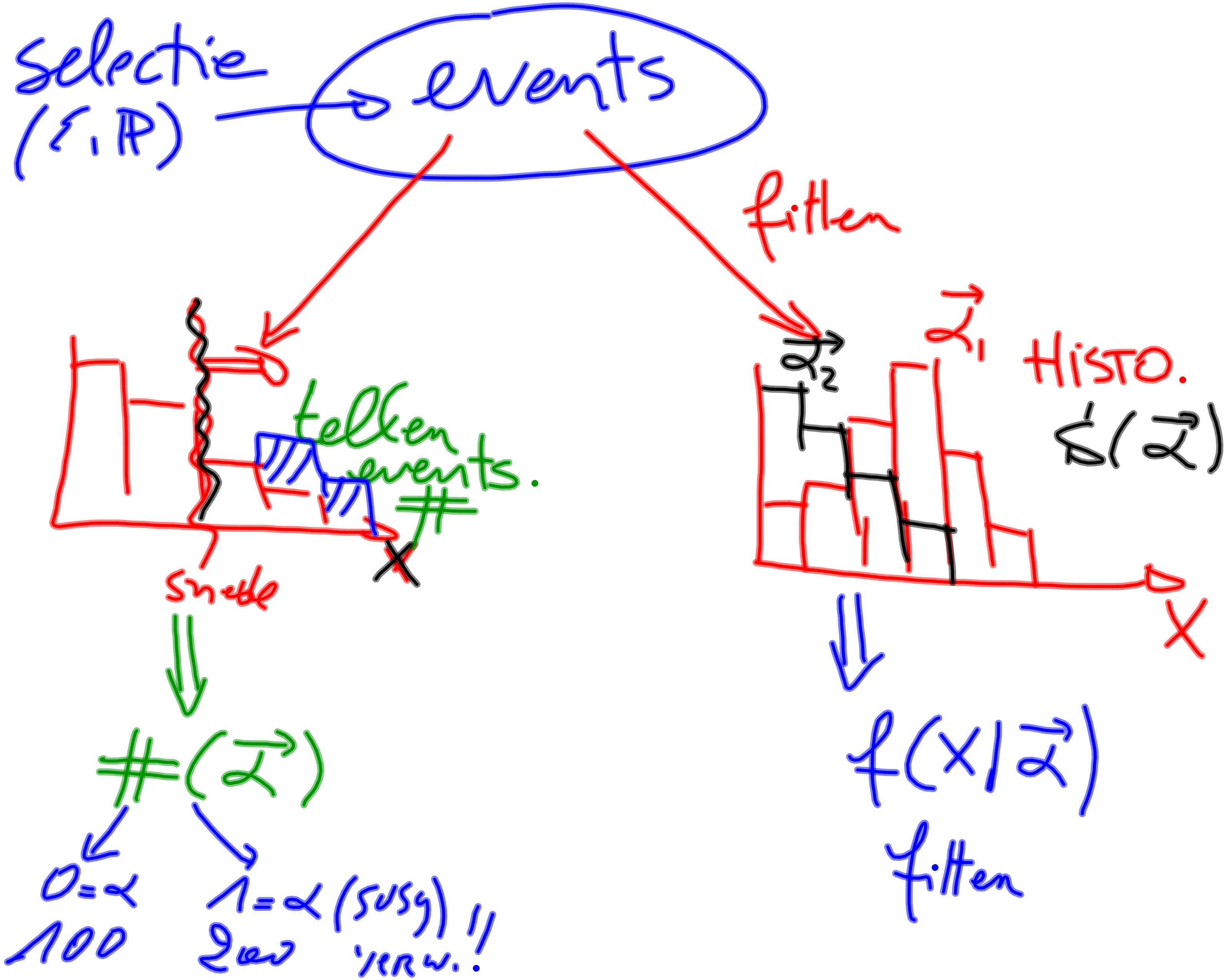
$$S \longleftrightarrow B$$

LR

$$\longrightarrow (\mathcal{E}, \mathbb{P}) \longleftarrow$$

$\rightsquigarrow$ Multi-Variate Analyse tools.
(M.V.A)

vb. Neurale Netwerken
Fisher discr.
Support Vector Machines


selectie
(E, P)
events
fitten
tellen events
#
snede
$\vec{\alpha}_1$
$\vec{\alpha}_2$
HISTO.
$s(\vec{\alpha})$
X
$\#(\vec{\alpha})$
$0 = \alpha$
100
$1 = \alpha$ (SUSY)
$f(X|\vec{\alpha})$
fitten

TELLEN (geïntegreerde lumi $\mathcal{L}$)

$$\mathcal{L} = \int_{t_0}^{t_1} L\, dt$$

$\# S$ (vb. Higgs) $\Delta t = t_1 - t_0$

$\# B$ (vb. SM. zonder Higgs)

$\sqrt{Var[S]} = \sqrt{S}$

Simulatie $\left\{ \begin{array}{l} E[\#S] \equiv S \sim \text{Poisson} \\ E[\#B] \equiv B \sim \text{Poisson} \end{array} \right.$

$\sqrt{Var[B]} = \sqrt{B}$

significantie van signaal

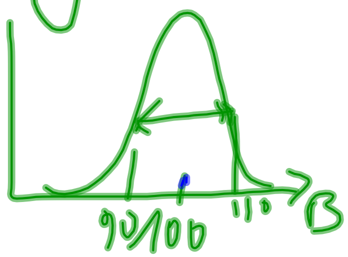
$B = 100 \longrightarrow$ onz. $\sqrt{100} = 10$

$E[B] = 100 \pm 10 \longrightarrow$ Gaussian

$E[S] = 2$

$= 20$

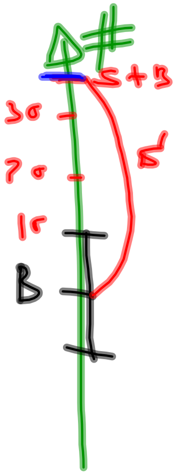
$DATA = S + B$



$E[\#S] = S$

$E[\#B] = B$

significantie : $\frac{S}{\sqrt{Var[S] + Var[B]}}$



$$\boxed{Sign = \frac{S}{\sqrt{S+B}}}$$

$B = 100$

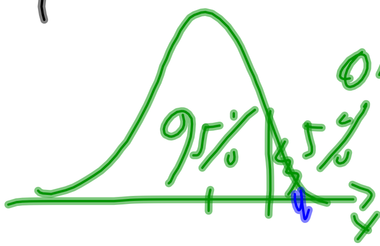
$S = 44$ $\Rightarrow$ $sign = \frac{44}{\sqrt{144}} = \frac{44}{12}$ $\sim 3.6\,\sigma$

$\Rightarrow$ REGEL : 5σ $\Leftarrow$ ONTDEKKEN

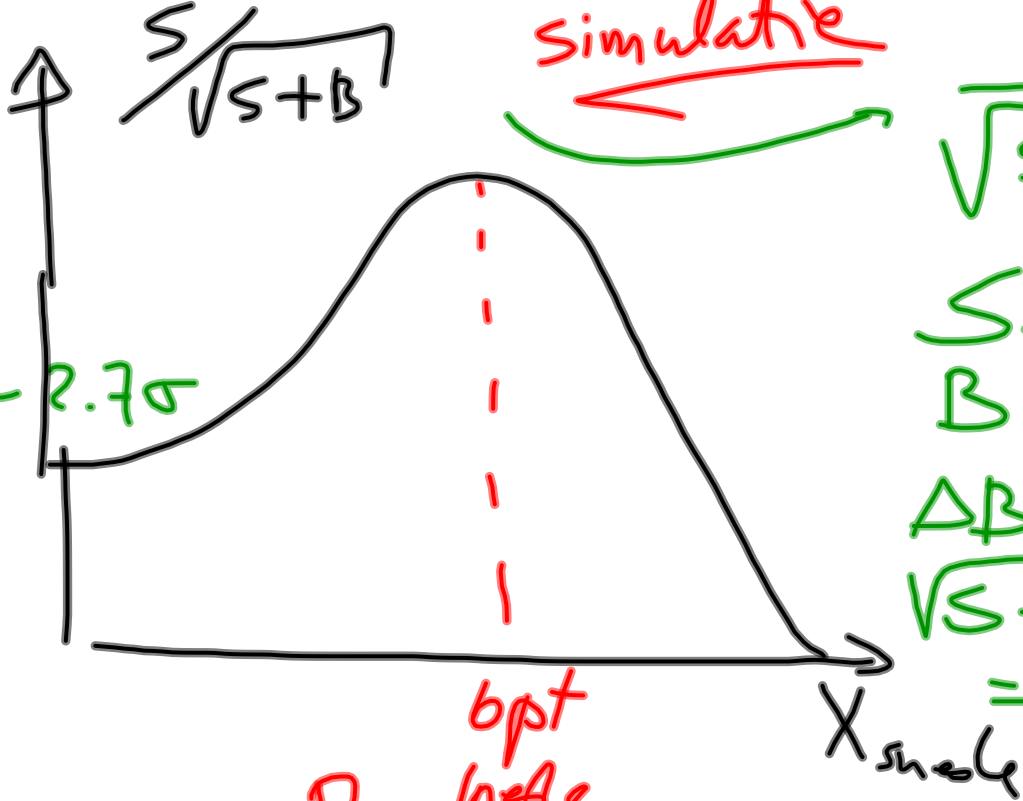
Stel 2σ $\Rightarrow$ 2.5% foute ond.

5%

$\Longrightarrow$



one-sided



$\frac{44}{16} \sim 2.7\sigma$

simulatie

$$\frac{S}{\sqrt{S+B+(\Delta B)^2}}$$

$S = 44$

$B = 100$

$\Delta B = 10$

$$\sqrt{S+B+(\Delta B)^2} = \sqrt{100+44+100} = \sqrt{244} \approx 16$$

Stat. onz.! snede.

$$\sqrt{Var[S]+Var[B]+(\Delta B)^2}$$

$B \rightarrow B+\Delta B \Rightarrow$ syst. onz. ΔB Syst. onz.

$\vec{\alpha}'_1 \leftrightarrow \vec{\alpha}'_2$

FITTEN (BINNED HISTO)

Histo. THEORIE

$$y \longrightarrow f(x|\vec{\alpha})$$

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - f(x_i|\vec{\alpha})}{\sigma_i} \right)^2 \longrightarrow \left\{ \left. \frac{\partial \chi^2(\vec{\alpha})}{\partial \alpha_j} \right|_{\vec{\alpha}=\hat{\vec{\alpha}}} = 0 \right\}$$

$$\Downarrow$$

$$\hat{\vec{\alpha}}$$

meting

$\sigma_i = \sqrt{y_i}$

gemakkelijk

wel bias

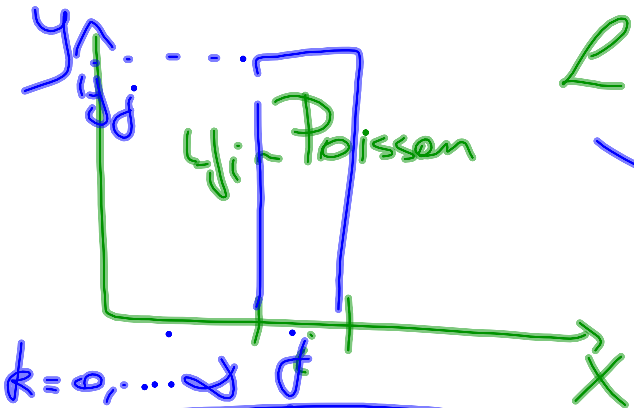
theorie

$\sigma_i = \sqrt{f(x_i|\vec{\alpha})}$

niet-lineaire vgh.

$\vec{\alpha}$ juist $\longrightarrow$ geen bias op $\vec{\alpha}$

BINNED HISTO



niet in bins

alle events

$$\mathcal{L}(\vec{x}\,|\,\vec{\alpha}) = \prod_{i=1}^{n} \underbrace{P(x_i\,|\,\vec{\alpha})}_{\text{niet Poisson}}$$

bins

$$\boxed{P_y(k) = \frac{\lambda^k}{k!} e^{-\lambda}} \qquad \mathcal{L}(\vec{y}\,|\,\vec{\alpha}) = \prod_{\text{bins } j} \underbrace{P(y_j\,|\,\vec{\alpha})}_{\text{Poisson.}}$$

$\lambda = E[y]$

$\vec{\alpha}$ $\qquad \lambda = \lambda(\vec{\alpha})$

max lik

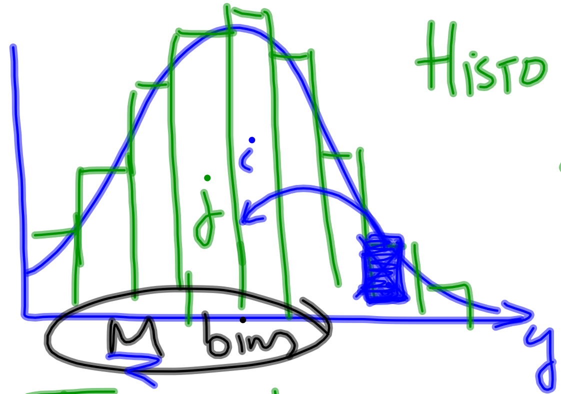
$$\hat{\alpha} = \ldots \Longleftarrow \left.\frac{\partial \mathcal{L}(\vec{y}\,|\,\alpha)}{\partial \alpha}\right|_{\alpha = \hat{\alpha}} = 0$$

$$X = \int \tilde{R} \otimes Th$$

CMS/ATLAS

"UNFOLDING"

variabele $y \longrightarrow \underbrace{f(y)}_{\text{PDF}}$



Histo

bin j : $p_j = \int_{\text{bin } j} dy\, f(y)$

TRUE Histo (theorie) $\longrightarrow \mu_j = \mu_{TOT} \cdot p_j$

$\mu_{TOT} = \sum_j p_j$

Realiteit $\xrightarrow{\text{smearing}}$

$$f_{meas}(y) = \int \underbrace{R(x|y)}_{\text{DETECTOR EFFECTEN}} f_{True}(y)\, dy$$

binned

$$\nu_i = \sum_{j=1}^{M} \underbrace{R_{ij}}_{\text{migratie } j \to i} \cdot \mu_j$$

$i \in \{1 \dots N\}$

E[·]

DETECTOR EFFECTEN

$$\nu_i = \sum_{j=1}^{M} R_{ij} \cdot \mu_j$$

$i \in \{1 \ldots N\}$

migratie $j \to i$

constant

DATA : $\vec{n} = (n_1 \ldots \boxed{n_i} \ldots n_N)$

N bins

stochastisch.

$\nu_i = E[n_i]$

met achtergrond

$$\nu_i = \sum_{j=1}^{M} R_{ij} \mu_j + \beta_j$$

verw. v.d. bck (simulatie)

$$\Longrightarrow \quad E[\vec{n}] = \underbrace{\vec{\nu} = \tilde{R} \cdot \vec{\mu} + \vec{\beta}}$$

$$\vec{\mu} = \tilde{R}^{-1}(\vec{\nu} - \vec{\beta})$$

$$E[\vec{n}] = \vec{\nu} = \underbrace{\tilde{R} \cdot \vec{\mu} + \vec{\beta}}$$

$$\vec{\mu} = \tilde{R}^{-1}(\vec{\nu} - \vec{\beta}) \longrightarrow \hat{\vec{\mu}} = \tilde{R}^{-1}(\vec{n} - \vec{\beta})$$

schatten

$n_i \sim$ Poisson

$$P(n_i|\nu_i) = \frac{\nu_i^{n_i}}{n_i!} \cdot e^{-\nu_i}$$

$$\mathcal{L} = \prod_{i=1}^{N} P(n_i|\nu_i) \longrightarrow \log \mathcal{L} = \sum_{i=1}^{N} \left(n_i \log \nu_i - \nu_i\right.$$

termen met ν_i afhankelijkheid

$$\left.\frac{\partial \log \mathcal{L}}{\partial \nu_i}\right|_{\nu_i = \hat{\nu}_i} = 0$$

$$\Longrightarrow n_i \cdot \frac{1}{\hat{\nu}_i} - 1 = 0 \Longrightarrow \hat{\nu}_i = n_i$$

$$\hat{\vec{\mu}} = \tilde{R}^{-1}(\vec{n} - \vec{\beta}) \longrightarrow E[\hat{\vec{\mu}}] = E\left[\tilde{R}^{-1}(\vec{n} - \vec{\beta})\right]$$

$$= \tilde{R}^{-1}\, E[\vec{n} - \vec{\beta}]$$

$$= \tilde{R}^{-1}\big(\underbrace{E[\vec{n}]}_{\nu_i} - \vec{\beta}\big)$$

$$= \vec{\mu}$$

$$\mathrm{Var}[y] = \lambda = E[y]$$

geen bias

$$\mathrm{cov}(\hat{\mu}_i, \hat{\mu}_j) = \boxed{\sum_{k,l=1}^{N}} (\tilde{R}^{-1})_{ik} (\tilde{R}^{-1})_{jl} \cdot \underbrace{\mathrm{cov}(n_k, n_l)}_{\text{onafh.}\ \delta_{kl} \cdot \nu_k}$$

groot

$$\hat{\mu}_i = c_i (n_i - \beta_i)$$

$$c_i = \frac{\mu_i^{MC}}{\nu_i^{MC}}$$

μ_i^{MC} → theorie

ν_i^{MC} → geobs

c_i: simulatie

bias: $E[\hat{\mu}_i] - \mu_i = \frac{\mu_i^{MC}}{\nu_i^{MC}} (\nu_i - \beta_i) - \mu_i \neq 0$

onz. $cov(\hat{\mu}_i, \hat{\mu}_j) = \underbrace{c_i^2}_{\approx 1} cov(n_i, n_j)$

Neyman methode voor betrouwbaarheids intervallen