*"It's a capital mistake to theorize before you have all the evidence"*

*It follows that physicists are bad detectives.*

Sherlock Holmes

# Sleuth

Y. Arnoud, B. Knuteson on the behalf of the DØ collaboration

## A Quasi Independent Model for New High $p_T$ Physics at DØ

- New physics ?
- Strategy
- Sleuth algorithm
- Tests
- Conclusions

# New Physics ?

Standard Model has drawbacks

Another model is needed

Supersymmetry is the most promising theory to be looked for …

But where to search ?

General ↑

Any theory of supersymmetry breaking

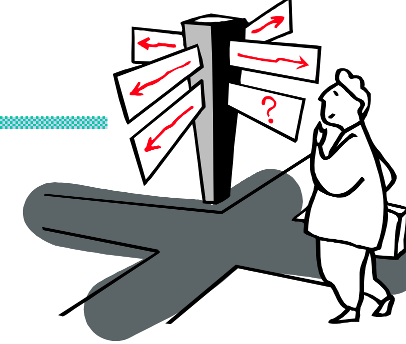
SUGRA

mSUGRA with undefined parameters

mSUGRA with $m_0$=90, $m_{1/2}$=200, $\mu$<0, tan $\beta$ = 20

↓ Constrained

# Usual Approach




Finding SUSY new signals w.r.t. SM processes :

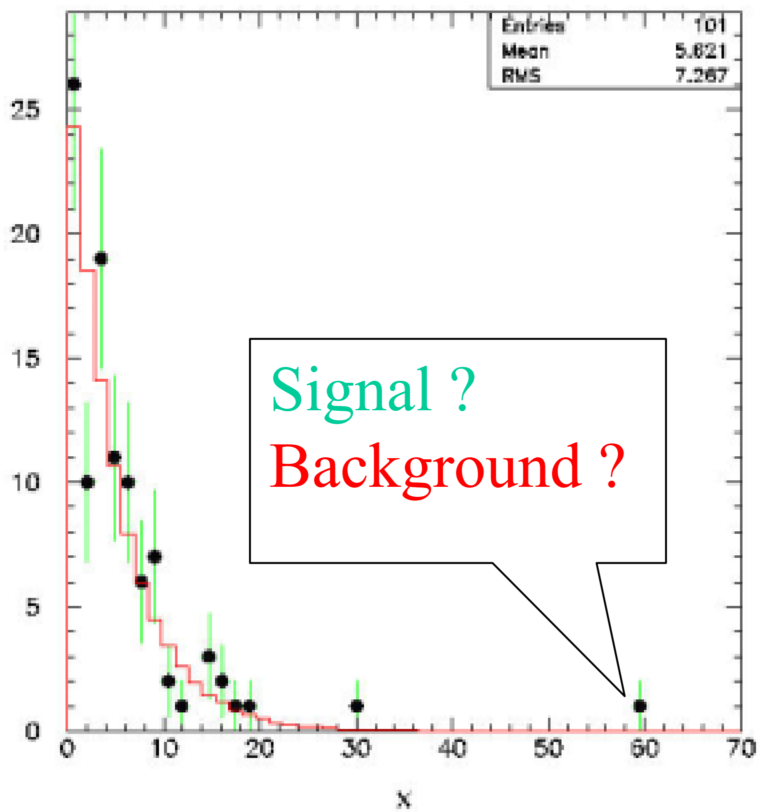
1. Choose the model to be tested and set parameters
2. Study and simulate background processes
3. Optimise cuts (sequential, log likelihood, NN, ...)
4. Compare real data to simulated background processes
5. Cheat a bit (go to step 3. to optimise search or exclusion area...)
6. Discover a new particle ! ... or ...
7. Exclude this set of parameters and goto step 1

Time consuming , could leave some parameter space unsearched

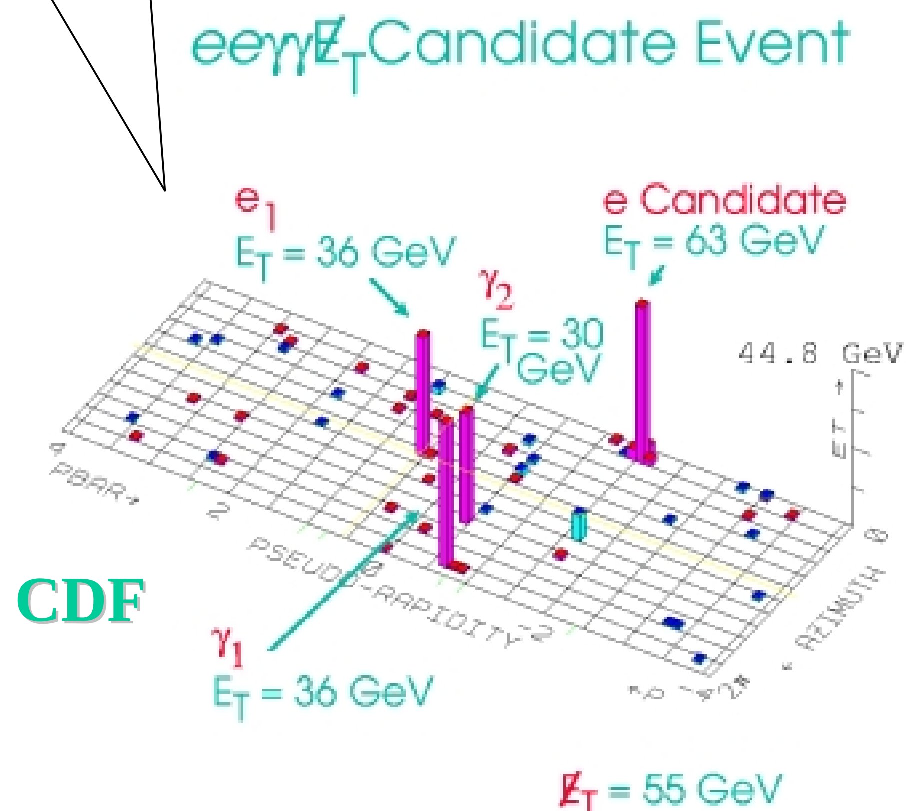
# In case of an abnormal event ...

How to quantify the probability to see such weird event ?

SUSY event ?
SM event ?




Signal ?
Background ?



# What are we looking for ?

## The physics responsible for EWSB

- o Natural scale ~ 250 GeV
- o Expect final states with high transverse momentum
- o Most sensitive variable for our generic search = $p_T$

The DØ collaboration has developed an original algorithm :

- o Looks for any $p_T$ excess
- o As general as possible
- o Provide interestingness of any data event

**SLEUTH !**

# SLEUTH Philosophy

- Variable selection
- Variable transformation
- Definition of regions
- Search for the largest fluctuation w.r.t. known processes
- Study statistical fluctuations of measurements
- Provide reliable numbers on the observed excess

# Final States

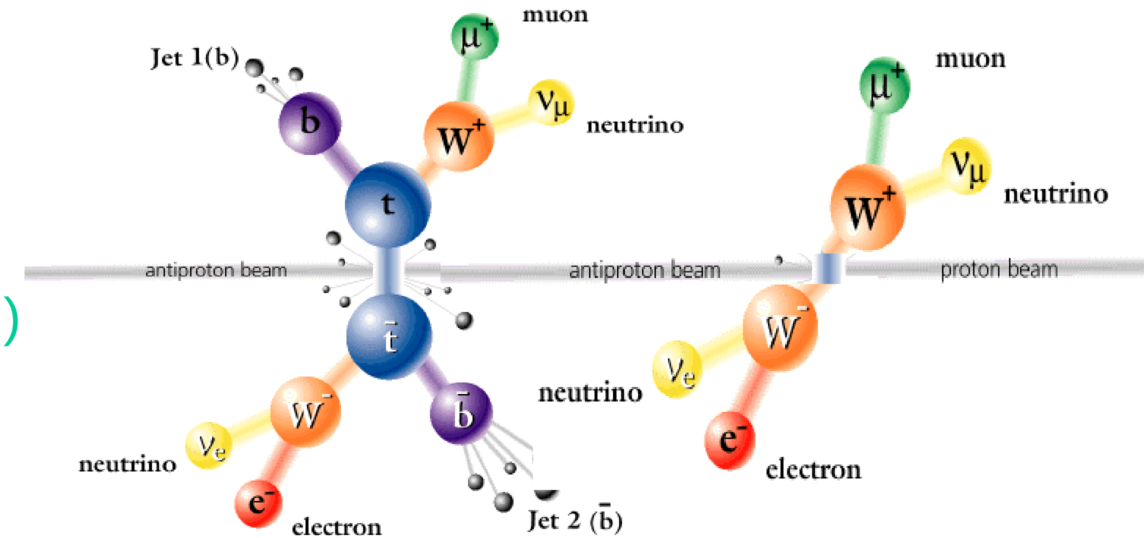
**Inclusive** or **exclusive** final states ?

Analyse inclusive events containing at least 1 e and 1 μ

Eg : p pbar -> e μ X

It could originate from

- t tbar decay (eμjj$\not{E}_t$)
- $W^+ W^-$ decay (eμ$\not{E}_t$)
- ...




different final state distributions, different analysis

Need to have two different sets of selection criteria ...

Consider only **exclusive** final states

# Exclusive Final States

Let us be more precise:

We assume the existence of standard object ID definitions

These define **electrons**, **muons**, **photons**, **taus**, **jets**, **b jets**, **c jets**, $\not{E}_T$, **Z**'s, and **W**'s

All events which contain the same numbers of each of these objects belong to the same final state

# How-To-Build Variables Rule

For each final state, associate one unique set of variables

| If the final state contains | Then consider the variable |
|---|---|
| 1 or more lepton | $p_T^{\ell}$ |
| 1 or more γ/W/Z | $p_T^{\gamma/W/Z}$ |
| 1 or more jet | $p_T^{j}$ |
| missing $E_T$ | $\not{E}_T$ |

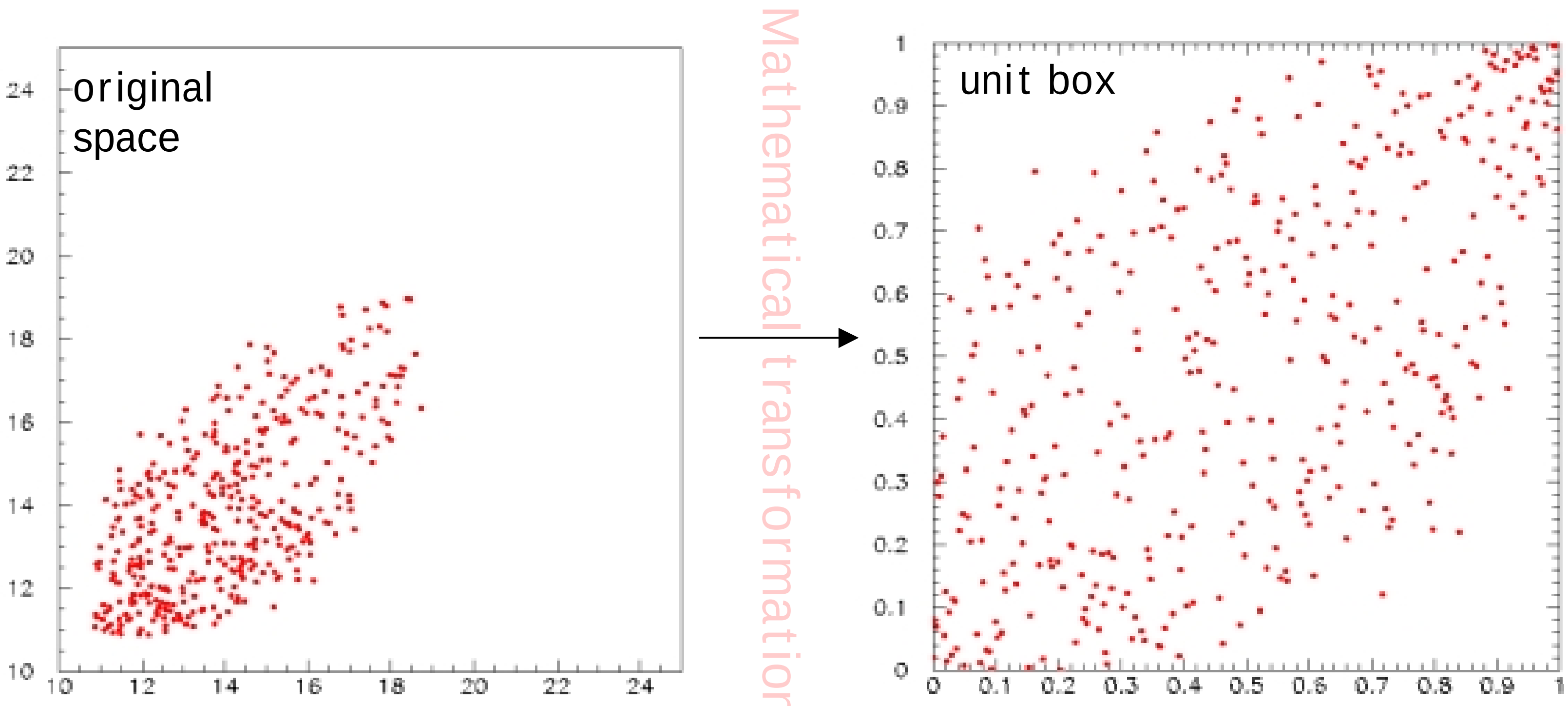
$$p_T^{j} = \begin{cases} p_T^{j_1} & (n_j = 1) \\ \sum_{i=2,n_j} p_T^{j_i} & (n_j \geq 2) \\ \sum_{i=3,n_j} p_T^{j_i} & (alljets)\ \&\ (n_j \geq 3) \end{cases}$$

# Variable transformation - 1

Main idea : flatten background distributions in a unit box

use same transformation scheme for real data

compare space occupancy




Projections along each axis are now uniform

# Mathematical transformation

The height of a sandbox in a d-dimensional unit sandbox is given by the function $b(x)$, where $x$ is a d-component vector. We take the d-dimensional lid of the sandbox and squash the sand flat. The result of this squashing is that a sand grain at position $x$ has moved to a new position $y$, and that the new function $b'(y)$ describing the height of the sand is constant. Given the function $b(x)$, determine the mapping $x \rightarrow y$

$$x_j \rightarrow x'_j = \frac{1}{M} \int_{-\infty}^{x_j} \sum_{i=1}^{M} \frac{1}{\sqrt{2\pi}\sigma_j h} \exp\left( -\frac{(t-\mu_{ij})^2}{2\sigma_j^2 h} \right) dt$$

$\mu_{ij}$ is the value of the $j^{th}$ variable for the $i^{th}$ background event,

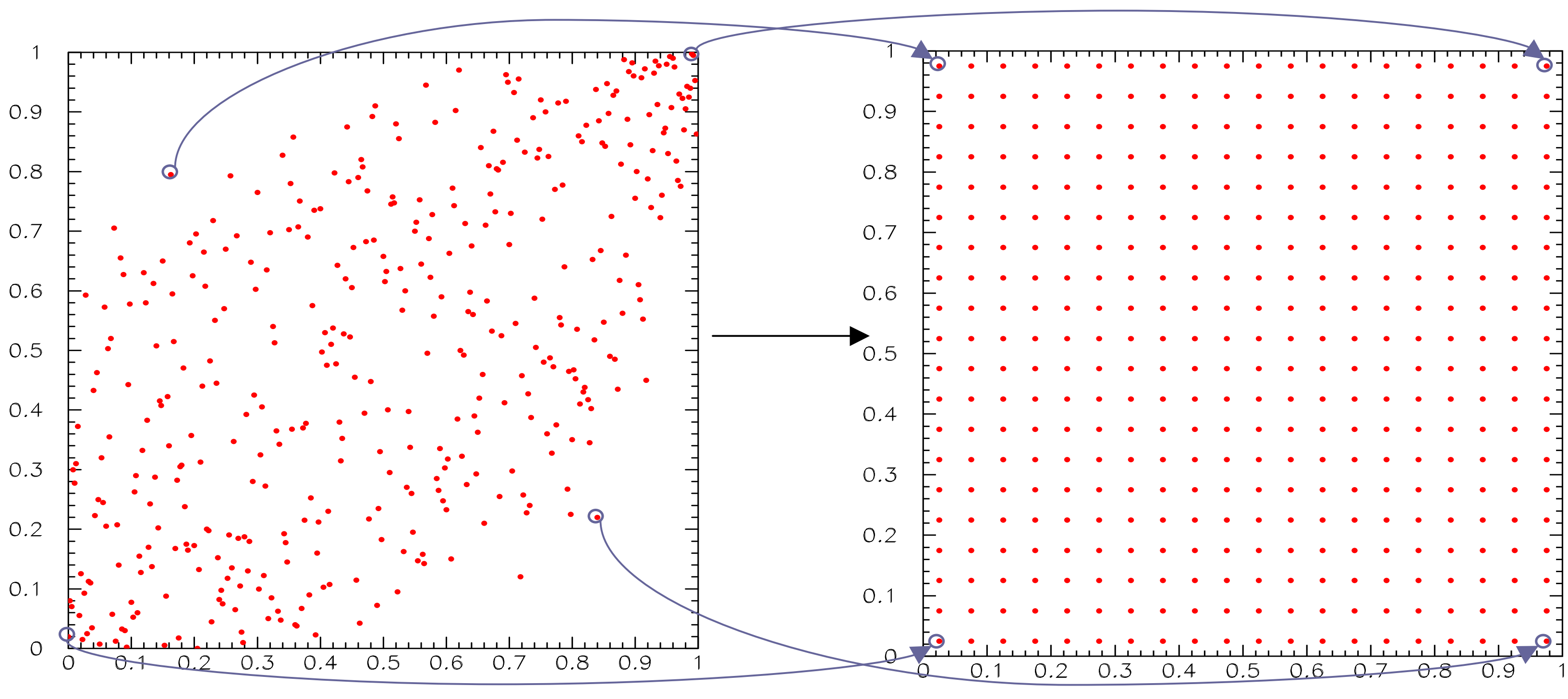
$\sigma_j$ is the standard deviation of the distributi on in the $j^{th}$ variable,

and $h = M^{-\frac{1}{d+4}}$, where $d$ is the dimensionality of the space.

# Variable transformation - 2

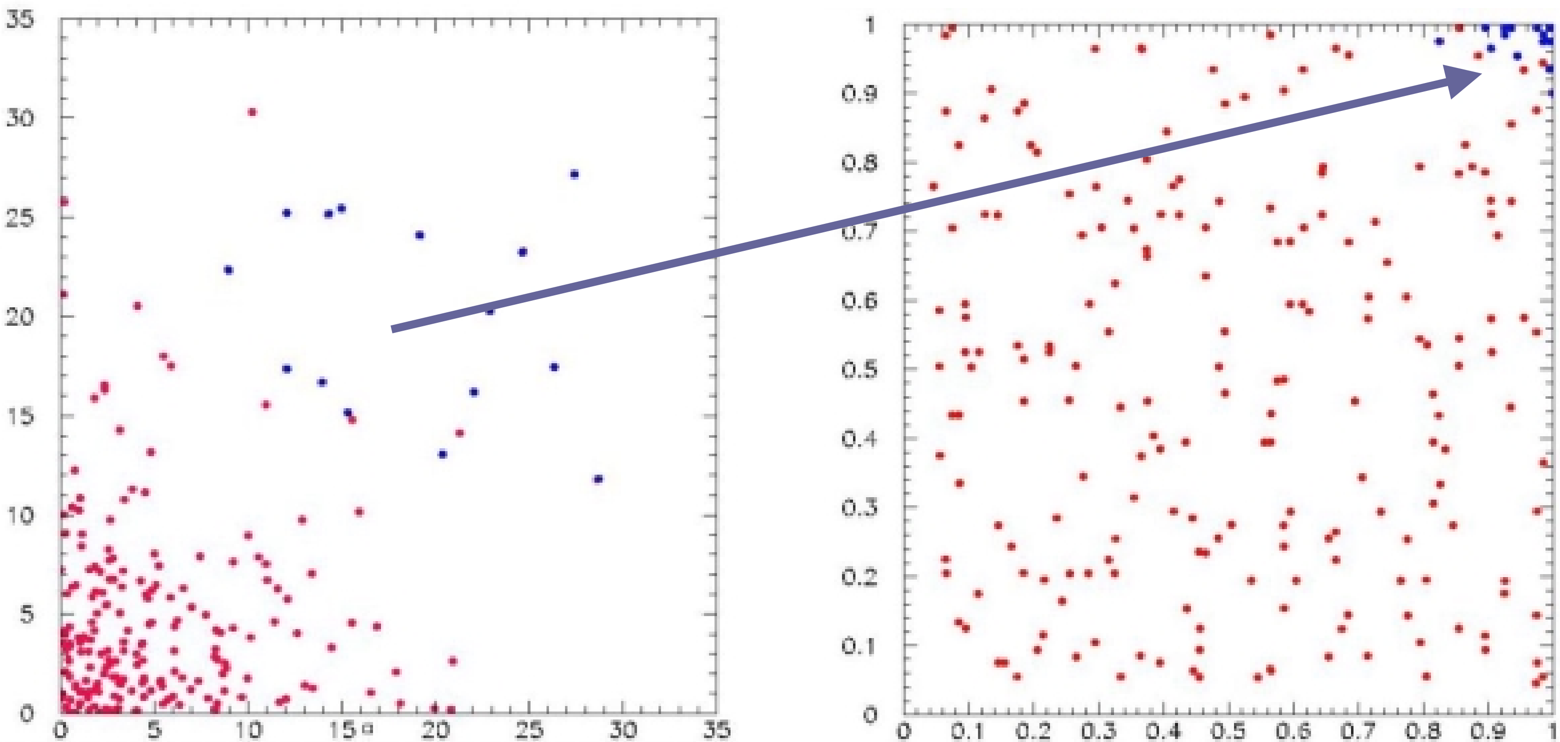
Spread the background events onto a uniform grid

Iteratively switch pairings to minimize the maximum distance moved and keep neighbors



# Dummy example

The transformation maps the signal region into the upper right-hand corner of the unit box




On this simulation, the background data events are uniformly distributed, as desired, and the signal cluster is “obvious”

# N-Regions, look for data excess

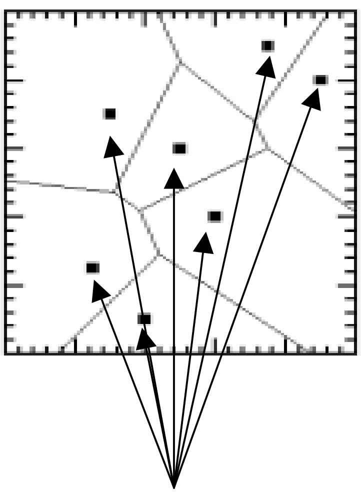
Underlying uniform background by construction

Map data points in the unit box with the same transformation

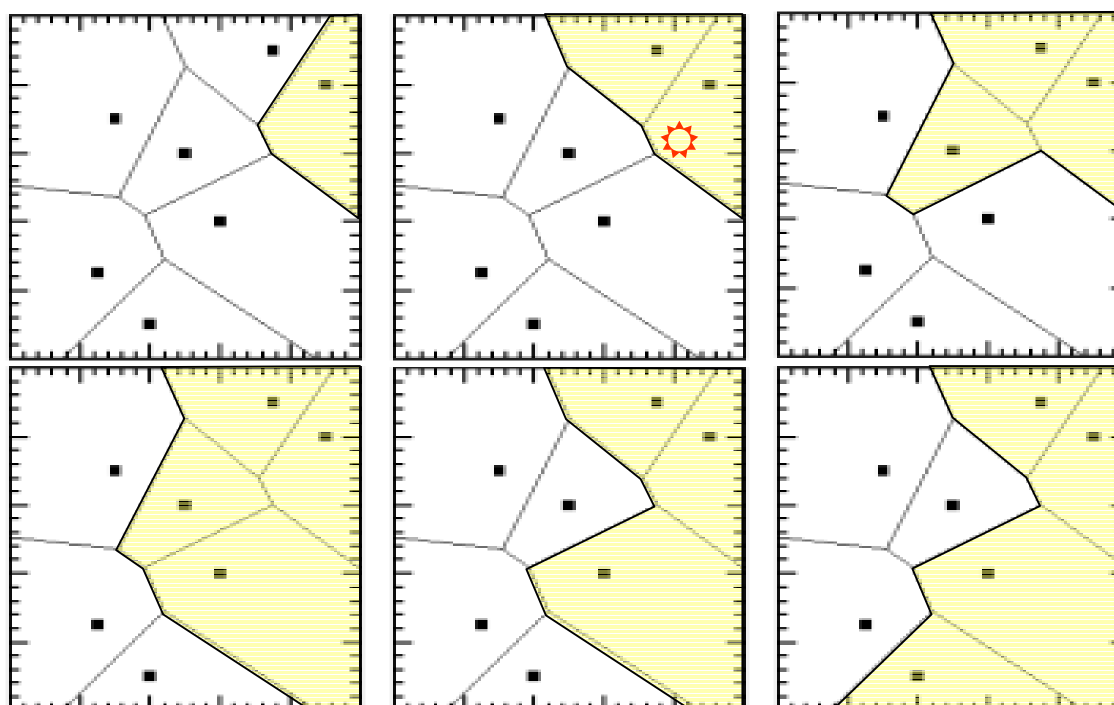
Group data points one by one (1-Region)

two by two (2-Region), ...

Search the N-region of greatest data excess ☼, i.e the N-region in which the probability $p_N^R$ for the observed background to fluctuate above observed events is the smallest



Data points mapped into the unit box



Voronoi diagrams

# Probability Calculation

The algorithm:

Input: 1 data file, 1 background file   Output : region of excess, probability

1. Define an *N-region* about any *connected* cluster of N data points → $R$
2. Estimate the background expected within that N-region R → $b_R$
3. Calculate the probability that $b_R$ fluctuates up to or above N → $p_N^R$
4. Determine the N-region R for which $p_N^R$ is minimum. This region is tagged as potentially interesting !

$$P(\text{data}) = \min_R (p_N^R)$$

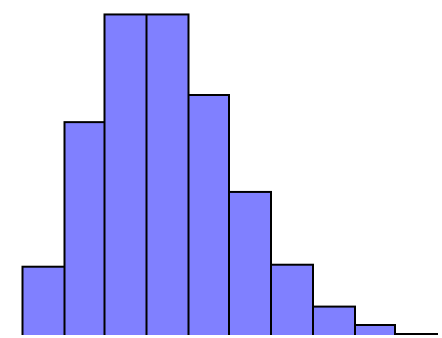
# Statistical Fluctuations

Lowest probability = P(data)

How likely is it to find such a low probability with Hypothetical Similar Experiments (hse) varying a bit the backgrounds ?

- If most *hse* give lower probabilities, our excess is not significant
- If only a small fraction of *hse* give a lower probability, our excess is significant

Replace the background $b_R$ with random numbers generated from the background distribution, and search for the lowest probability P(hse)



Determine the fraction $P$ of *hse* in which the $p_N$ (hse) is smaller than the observed $p_N$ (data)

$P$ will provide the appropriate measurement of the degree of interest

# Testing Sleuth ...

Sleuth algorithm was applied to Run I DØ data containing 1 e and 1 μ :

Using same base cuts as for eμ top analysis

same e/μ identification

e/μ transverse momentum > 15 GeV

Leaves 58 events over 108.3 ± 5.7 $pb^{-1}$

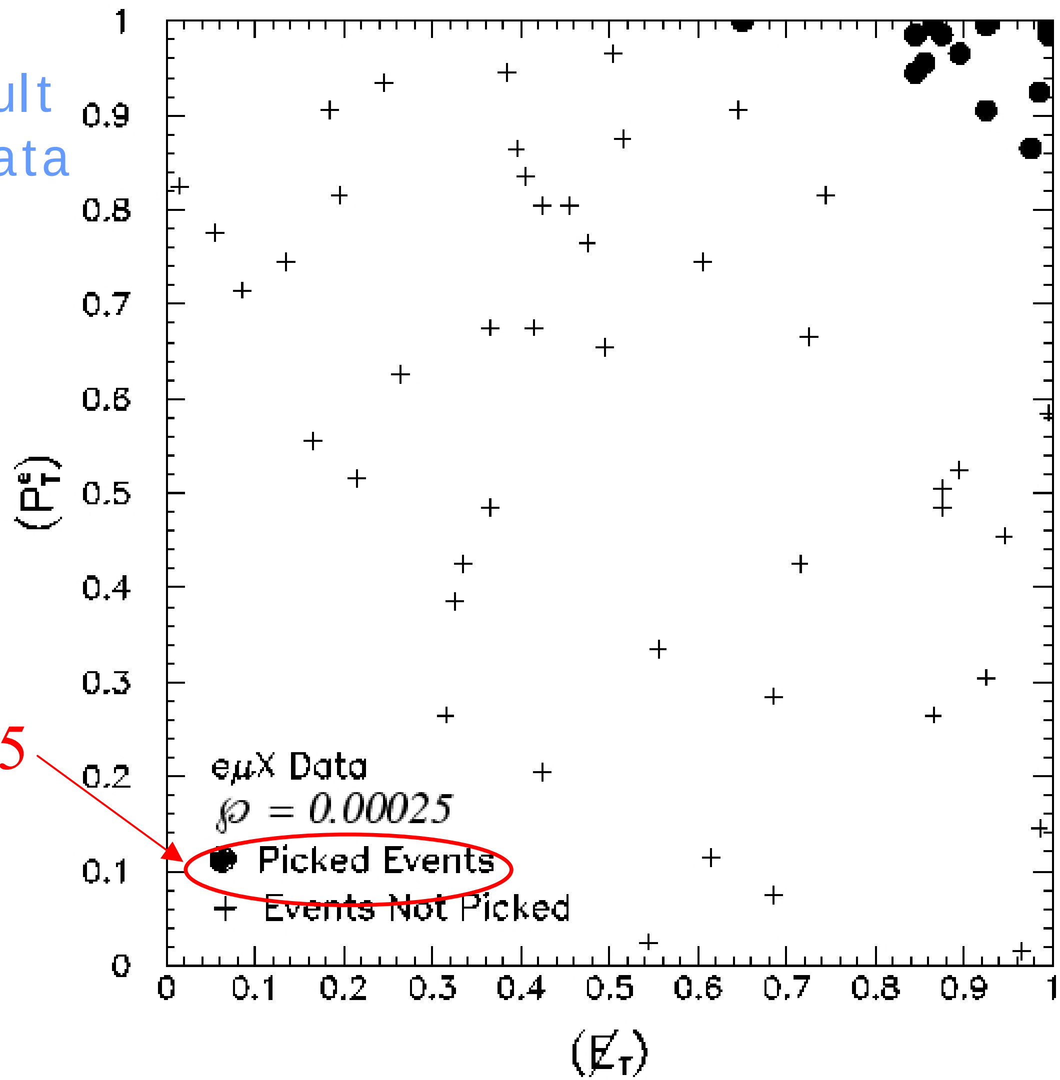
| Final States | Backgrounds |
|---|---|
| $e\mu \not{E}_T$<br>$e\mu \not{E}_T\, j$<br>$e\mu \not{E}_T\, j\, j$<br>$e\mu \not{E}_T\, j\, j\, j$ | $Z/\gamma^* \to \tau\tau \to e\mu X$<br>"fakes": $b\bar{b}/c\bar{c} \to e_{fake}\mu X$; $jW \to j\mu\nu \to e_{fake}\mu X$<br>$WW \to e\mu X$<br>$t\bar{t} \to WWb\bar{b} \to e\mu X$ |

Can we find evidence for WW and t tbar production ?

# …to look for WW and $t\bar{t}$ production

Here is the result on DØ's $e\mu X$ data

$P = 0.00025$ (3.5σ) !



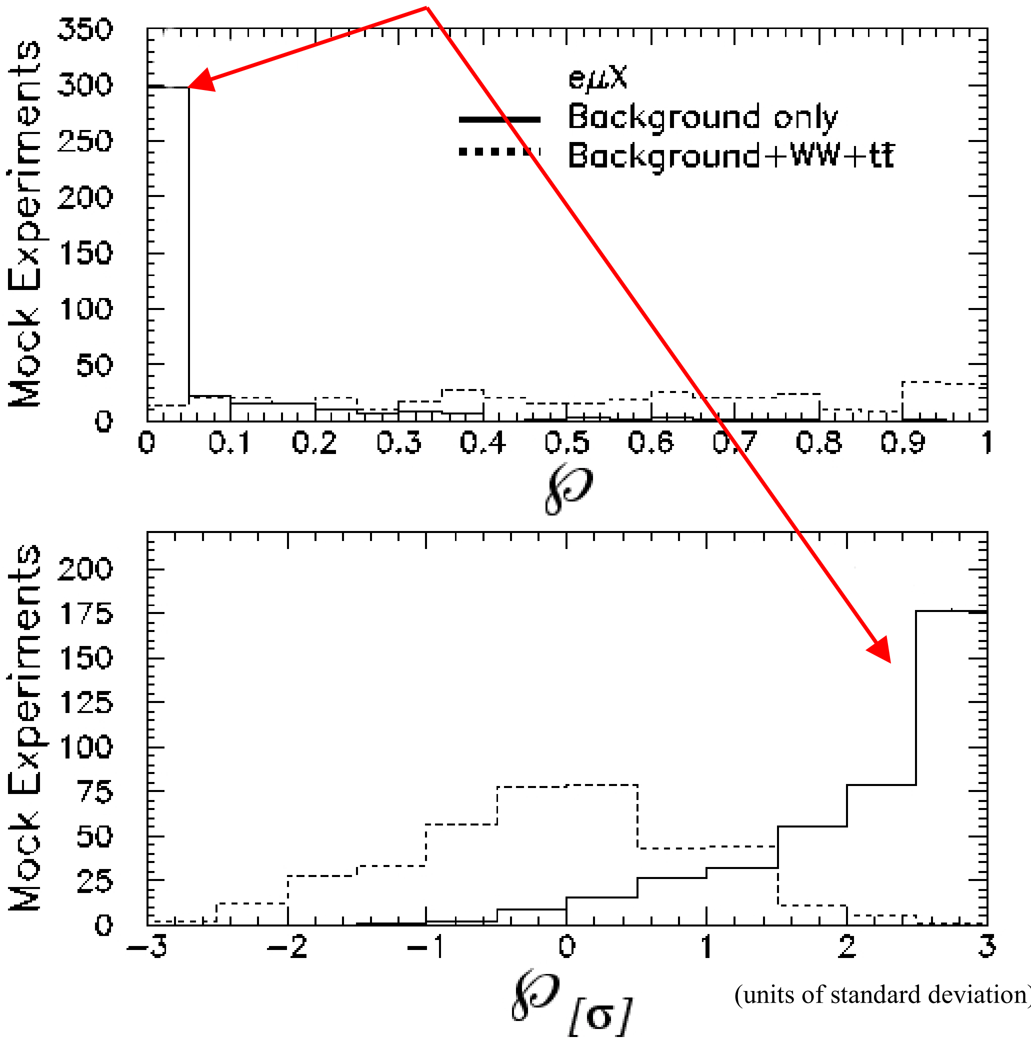

We do indeed find WW and tt (data excess in this region !)

Remember : underlying flat background, all crosses and points shown are data points

# P(hse) distributions

We can indeed pull out WW and $t\bar{t}$ :

*P* peaked at 0 = indication for new physics !




Comparaison with
the official DØ t tbar analysis :

| Sleuth (bkg + WW) | DØ eμX analysis |
|---|---|
| **1.9 σ** effect | **2.75 σ** effect |

# Sleuth beyond Standard Model

Take into account all SM backgrounds and run Sleuth on all available DØ data

## Summary of results

| Data Set | $\mathcal{P}$ |
|---|---|
| $e\mu\not{E}_T$ | 0.12 |
| $e\mu\not{E}_T\, j$ | 0.68 |
| $e\mu\not{E}_T\, jj$ | 0.50 |
| $e\mu\not{E}_T\, jjj$ | 0.63 |
| $\mathcal{P}$ | 0.40 |

Disappointingly good agreement with the Standard Model

No evidence for new high $p_T$ physics in eμX

# Conclusions

- **Sleuth** is a quasi-model-independent search strategy for new high $p_T$ physics developed by B. Knuteson
  - Defines final states and variables
  - Systematically searches for and quantifies regions of excess
- **Sleuth** allows an *a posteriori* analysis of interesting events
- **Sleuth** appears sensitive to new physics . . .
- **. . .** but finds no evidence of new physics in DØ data
- **Sleuth** has the potential for being an extremely useful tool
  - **Looking forward to Run II!**

hep-ex/0006011 PRD
hep-ex/0011067 PRD
hep-ex/0011071 PRL