Sleuth

A quasi-model-independent new physics search strategy

- Motivation
- Strategy
- Algorithm
- Results

Bruce Knuteson
Berkeley

Motivation

Most searches follow a well-defined set of steps:

- Select a model to be tested
- Find a measurable prediction of the model differing as much as possible from the prediction of the Standard Model
- Check those predictions against the data

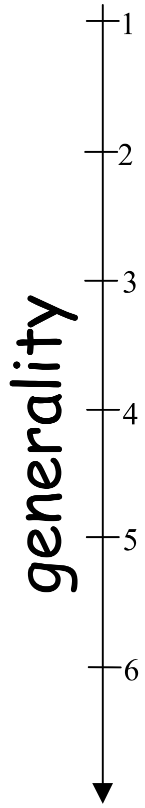
This approach becomes problematic if the number of competing candidate theories is large . . . and it is!

Is it possible to perform some kind of "generic" search?

Sleuth

Motivation

The word "model" can connote varying degrees of generality



1 - A special case of a class of models with definite parameters
    mSUGRA with $M_{1/2}=200$, $M_0=220$, $\tan\beta=2$, $\mu<0$
2 - A special case of a class of models with unspecified parameters
    mSUGRA
3 - A class of models
    SUGRA
4 - A more general class of models
    gravity-mediated supersymmetry
5 - An even more general class of models
    supersymmetry
6 - A set of even more general classes of models
    theories of electroweak symmetry breaking

Most new physics searches have generality $\approx 1\frac{1}{2}$ on this scale

We are shooting for a search strategy with a generality of ≈ 6 . . . .

Motivation

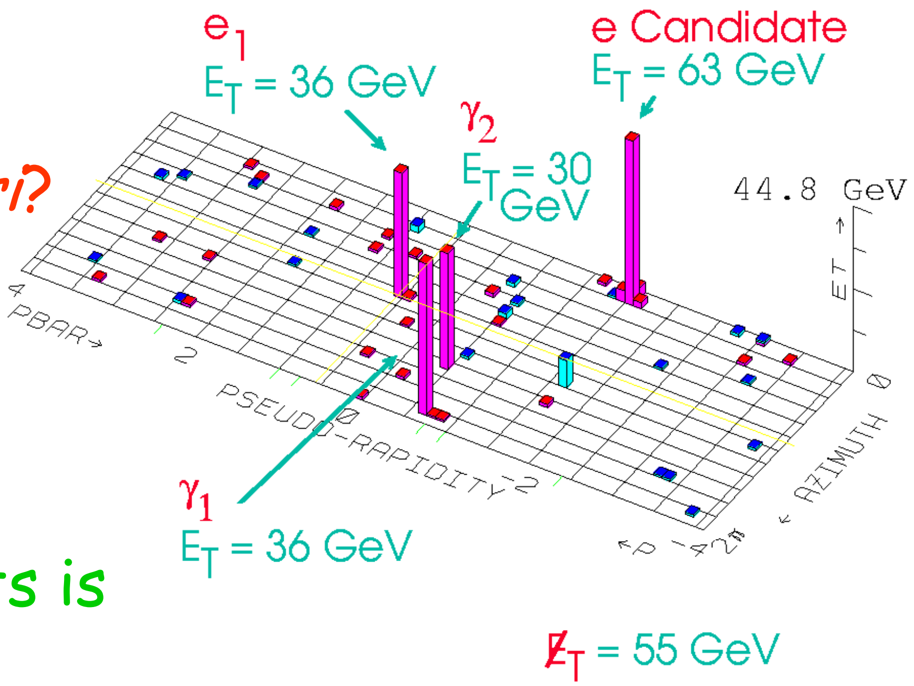
Another, separate issue:

How do we quantify the "interestingness" of a few strange events *a posteriori*?

e.g. Barnett and Hall, PRL 77:3506 (1996)

After all, the probability of seeing exactly those events is zero!

CDF $ee\gamma\gamma\not{E}_T$ Candidate Event



How excited should we be? How can we possibly perform an unbiased analysis after seeing the data?

Sleuth

Motivation

Other advantages of **Sleuth**:

Emphasizes an <u>understanding of the data</u> (rather than what the data have to say about a particular model)

Provides a systematic method for analyzing the entire data set (leaving no stone unturned!)

Allows an approach that keeps attention focused on the most promising channels (rather than optimizing cuts for a signal that does not exist)

Allows for surprises . . .

Sleuth

A quasi-model-independent new physics search strategy

- Motivation
- Strategy →
  - Final states
  - Variables
- Algorithm
- Results

Final states

Initial thought:

Consider inclusive final states, such as e μ X

However:

- The presence of an extra object in an event often qualitatively changes the probable interpretation of the event
- The presence of an extra object in an event generally changes the variables that one would want to use to characterize the event
- Allowing inclusive final states leaves an ambiguity in definition

Therefore:

We consider *exclusive* final states

Final states

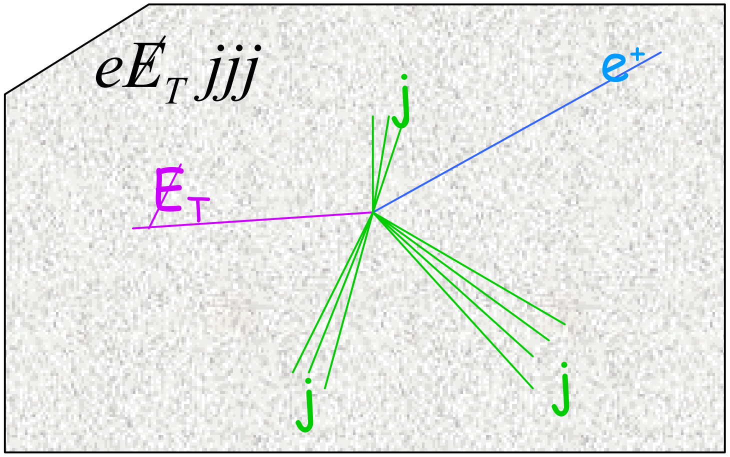
More precisely:

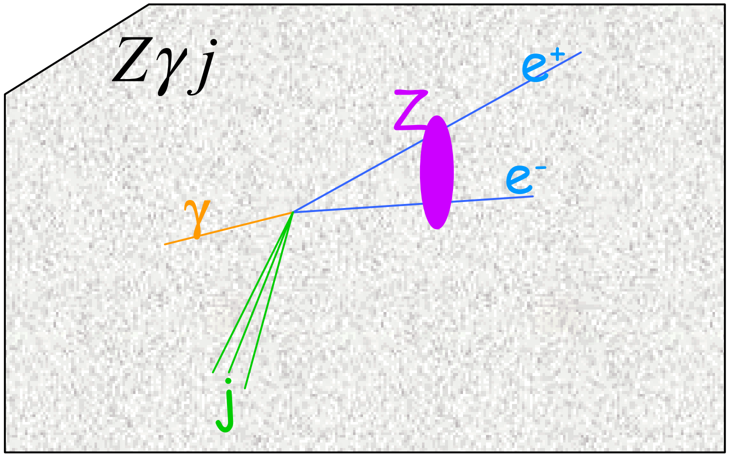
We assume the existence of standard object definitions

These define e, μ, τ, γ, j, b, c, $\not{E}_T$, W, and Z

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

e.g.,





Variables

Initial thought:

Construct a set of variables for each possible final state

However:

- There are a lot of final states!
  - eμX alone comprises several final states
- Our variables need to be robust
  - Otherwise it will be too easy to change them after looking at the data!
- Our variables ought to be
  - well-motivated (sensitive to new physics)
  - simple and few

Therefore:

Instead of choosing a separate set of variables for every conceivable final state, we construct a general rule

$$\mathcal{V} : \text{(final state)} \rightarrow \{\ \text{variables}\ \}$$

Variables

$\mathcal{V}$: (final state) → { variables }

What is it we're looking for?

The physics responsible for EWSB

What do we know about it?

Its natural scale is a few hundred GeV

What characteristics will such events have?

Final state objects with large transverse momentum

What variables do we want to look at?

p_T's

Variables

General:

$\mathcal{V}$: (final state) → { variables }

If the final state contains	Then consider the variable
1 or more lepton	$\sum p_T^{\ell}$
1 or more $\gamma/W/Z$	$\sum p_T^{\gamma/W/Z}$
1 or more jet	$\sum{}' p_T^{j} =$
missing E_T	$\not\!\!E_T$

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

DØ Run I specific:

$$\sum p_T^{\ell} = \sum p_T^{e}$$ (for events containing electrons)

$$\not\!\!E_T = \not\!\!E_T^{cal}$$

Sleuth

A quasi-model-independent new physics search strategy

Motivation

Strategy

Algorithm

Results

Algorithm Overview

For each final state . . .

Input: 1 data file, estimated backgrounds

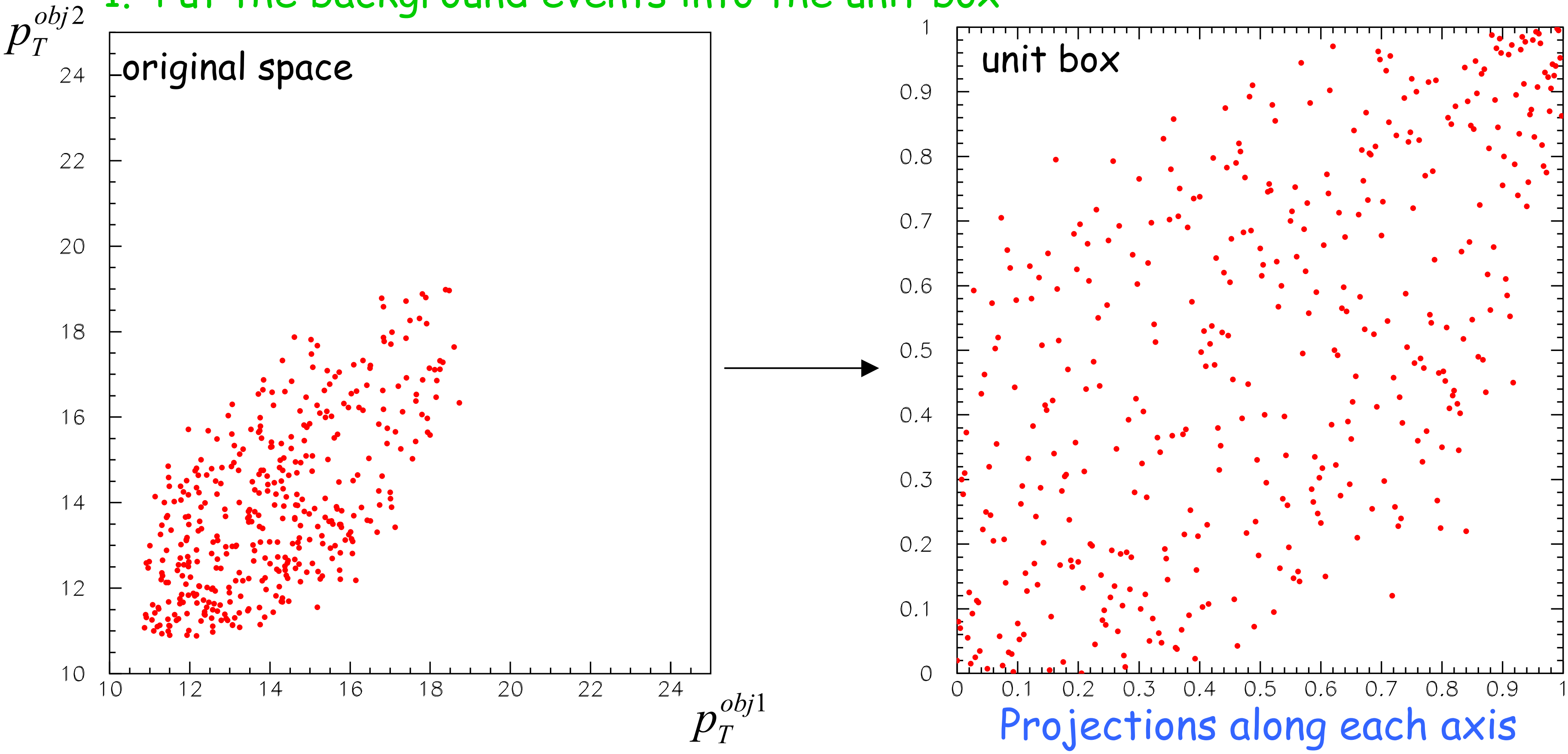
- transform variables into the unit box
- define regions about sets of data points
  - Voronoi diagrams
- define the "interestingness" of an arbitrary region
  - the probability that the background within that region fluctuates up to or beyond the observed number of events
- search the data to find the most interesting region, $\mathcal{R}$
- determine $\mathcal{P}$, the fraction of *hypothetical similar experiments* (hse's) in which you would see something more interesting than $\mathcal{R}$
  - Take account of the fact that we have looked in many different places

Output: $\mathcal{R}$, $\mathcal{P}$

Algorithm — Variable transformation

We begin by applying a variable transformation that makes the background distribution uniform in the "unit box" — $[0,1]^d$

1. Put the background events into the unit box



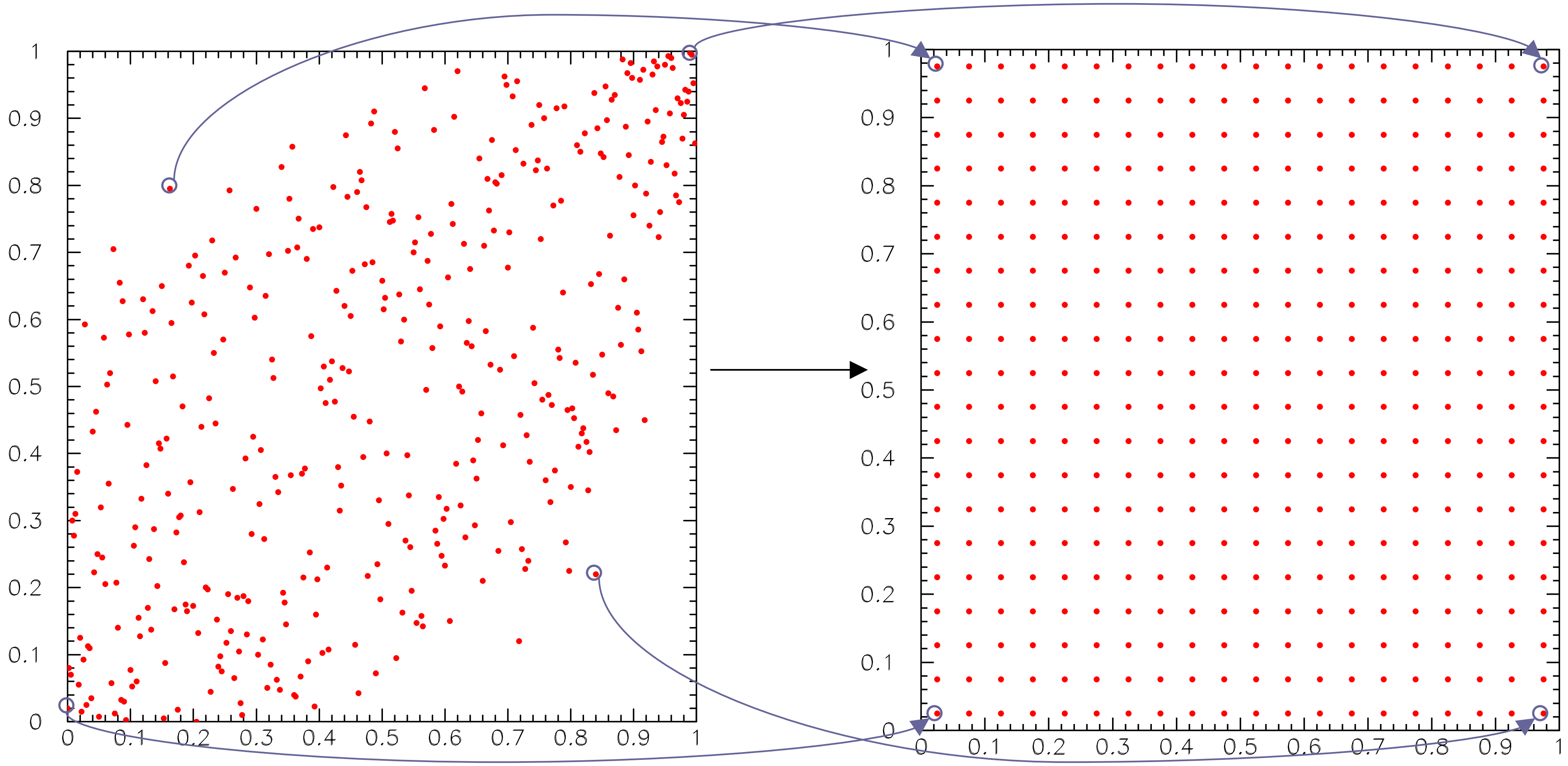
Projections along each axis are now uniform

Algorithm

Variable transformation

2. Map the background events onto a uniform grid

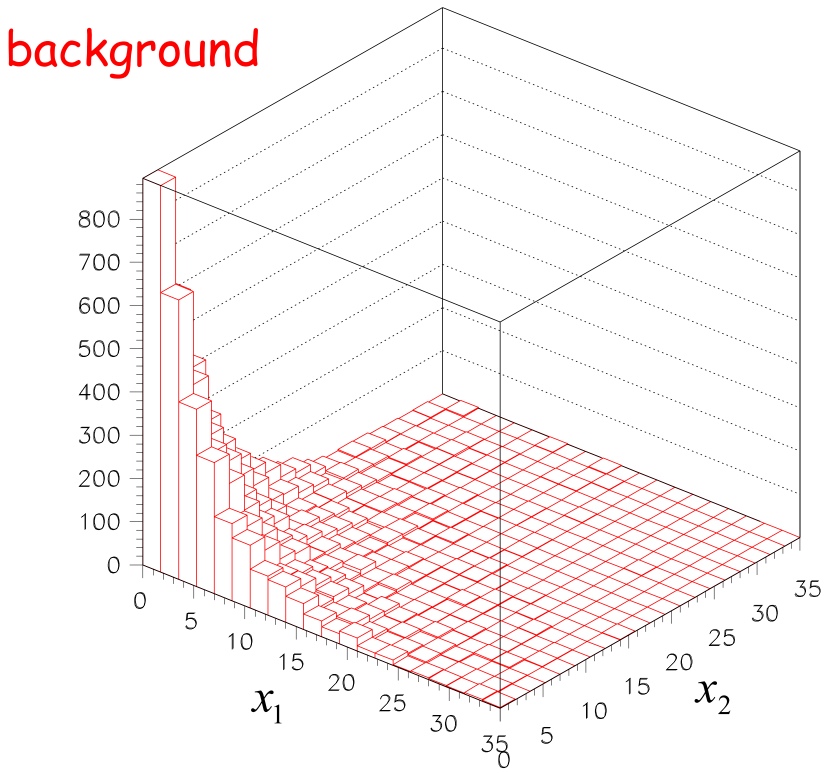
[Iteratively switch pairings to minimize the maximum distance moved]

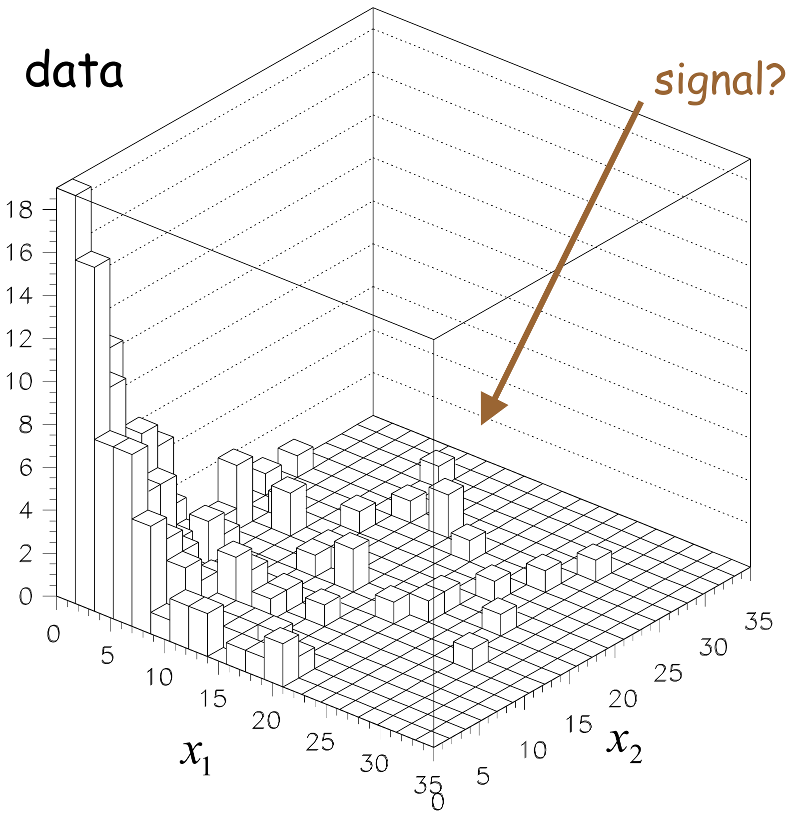


Algorithm

Variable transformation

A quick example of how this might look for data:

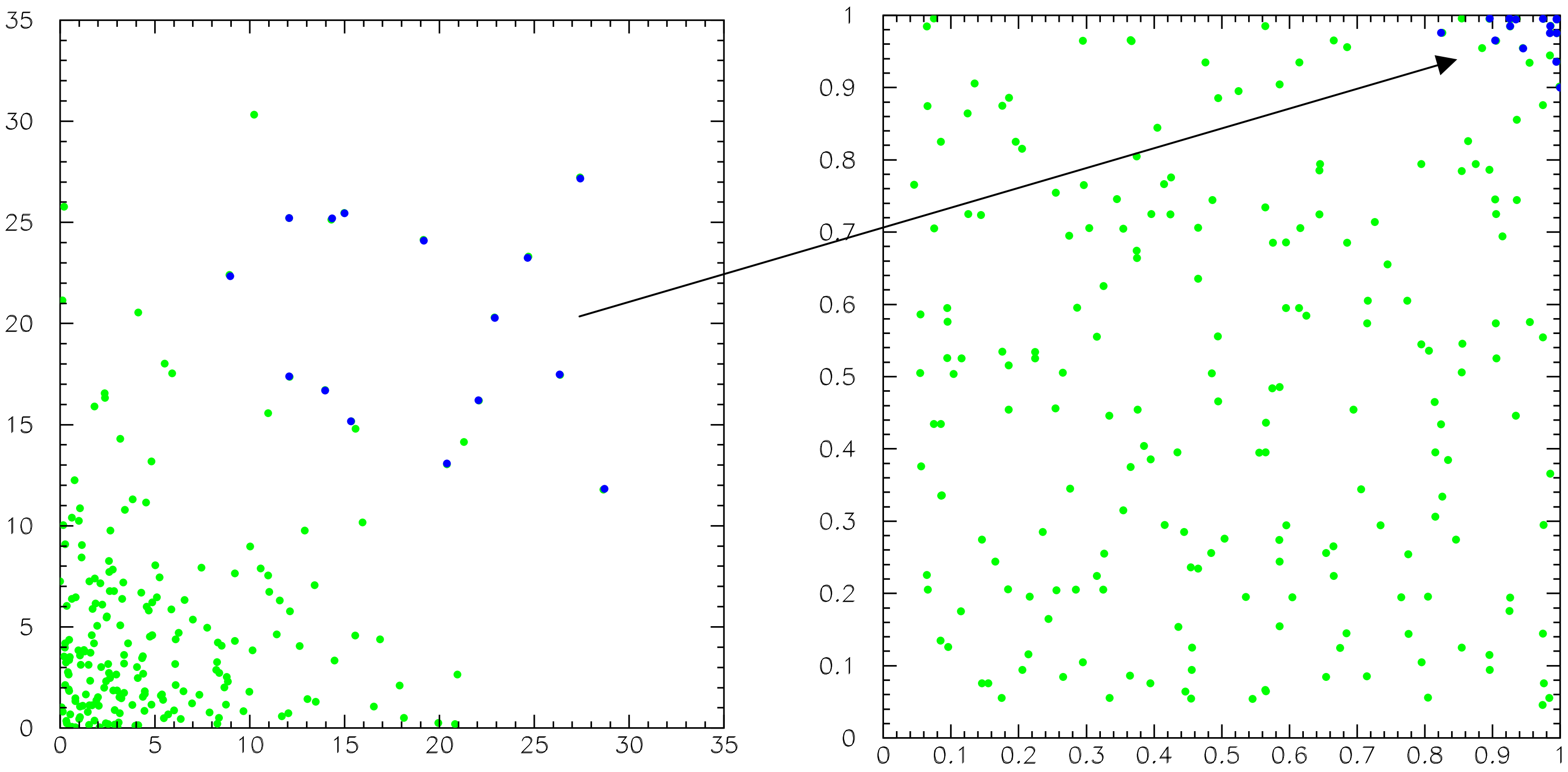




Algorithm

Variable transformation

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



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

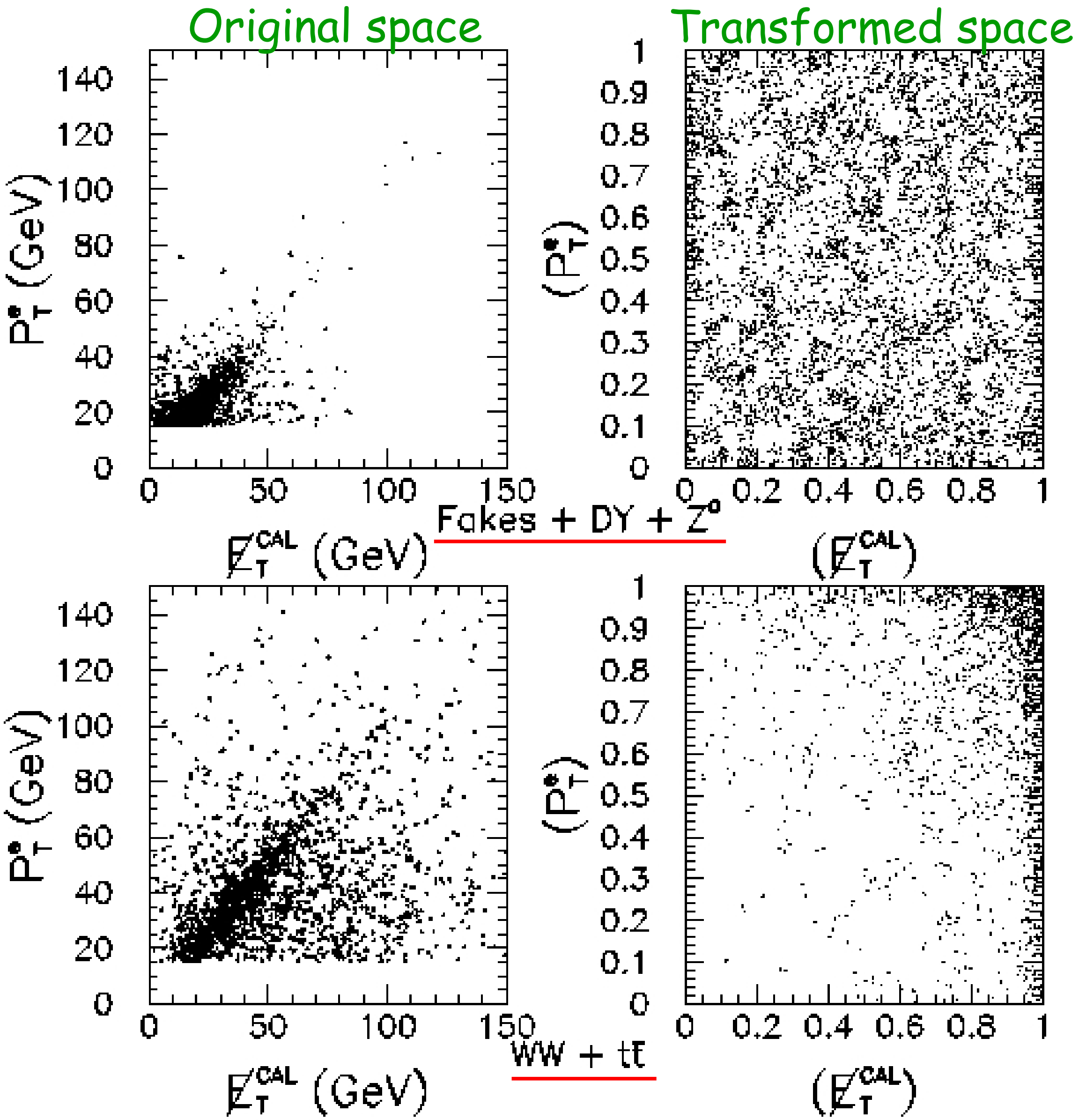
Algorithm

Variable transformation

Backgrounds are Z→ττ and fakes.

(We use (p_T^{ℓ}) $(\not{E}_T)$ to denote transformed variables)

“Signal” is WW and $t\bar{t}$



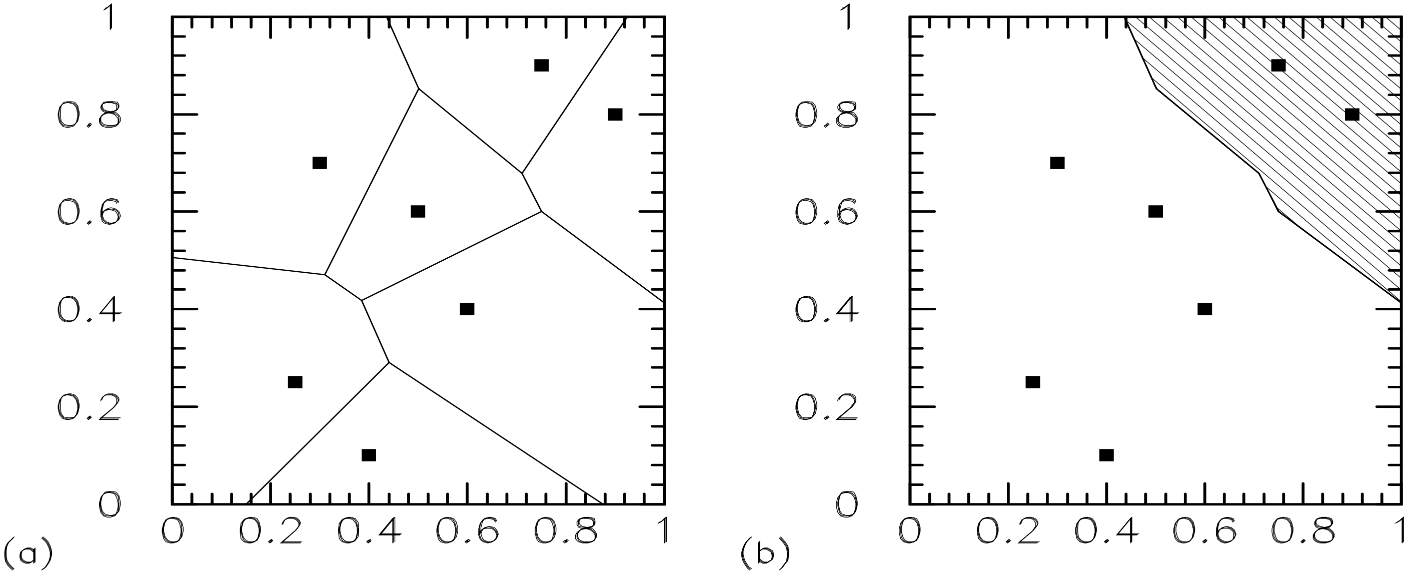
Z→ττ and fakes should together be uniformly distributed. ✓

We expect WW and $t\bar{t}$ to cluster in the upper right hand corner. ✓

Algorithm

Regions

An N-region (about a cluster of N data points) is the set of all values of x closer to a data point in that cluster than to any other data point in the sample.



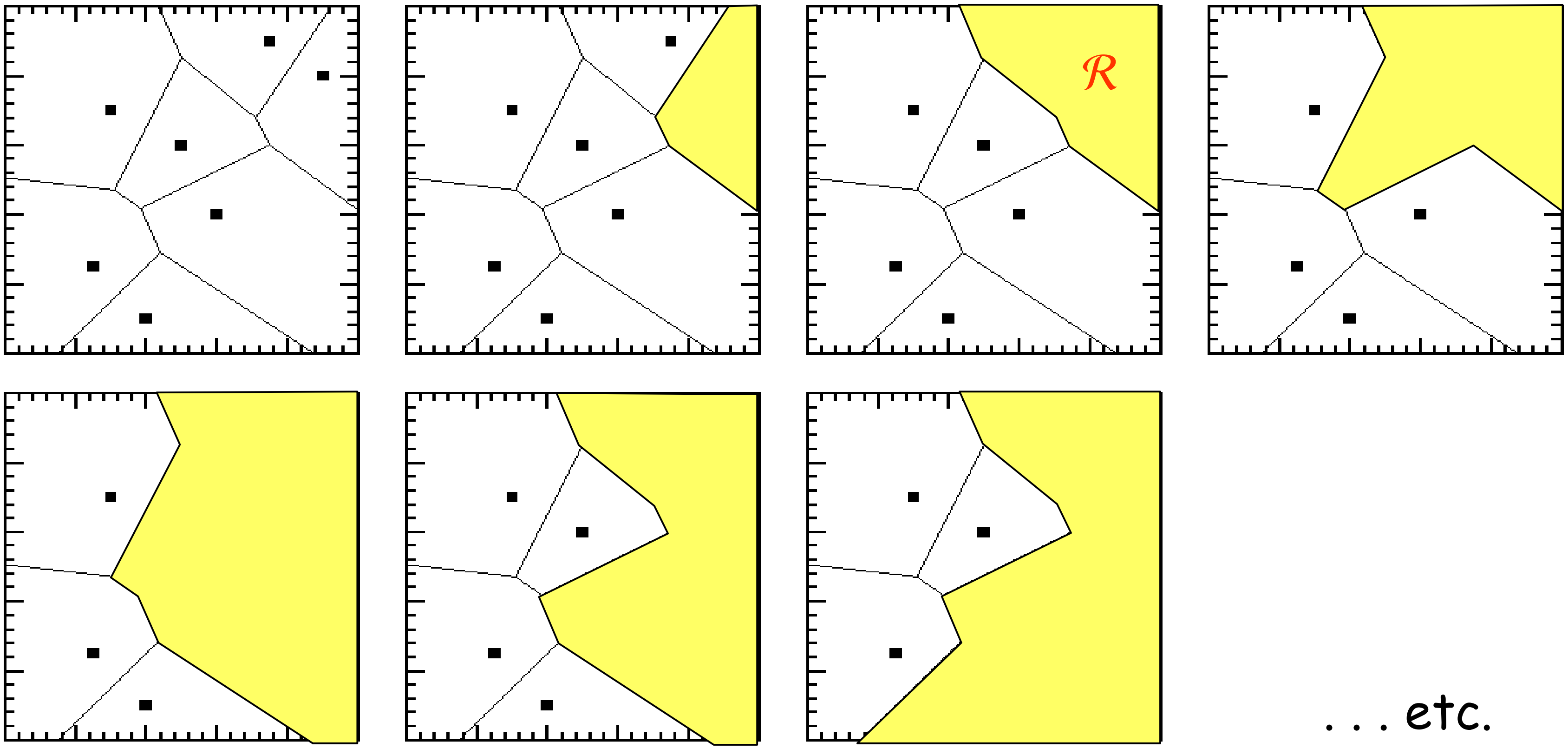
seven 1-regions

one 2-region

Voronoi diagrams

Algorithm Search

Search the space to find the region of greatest excess, $\mathcal{R}$



. . . etc.

Perform many hypothetical similar experiments

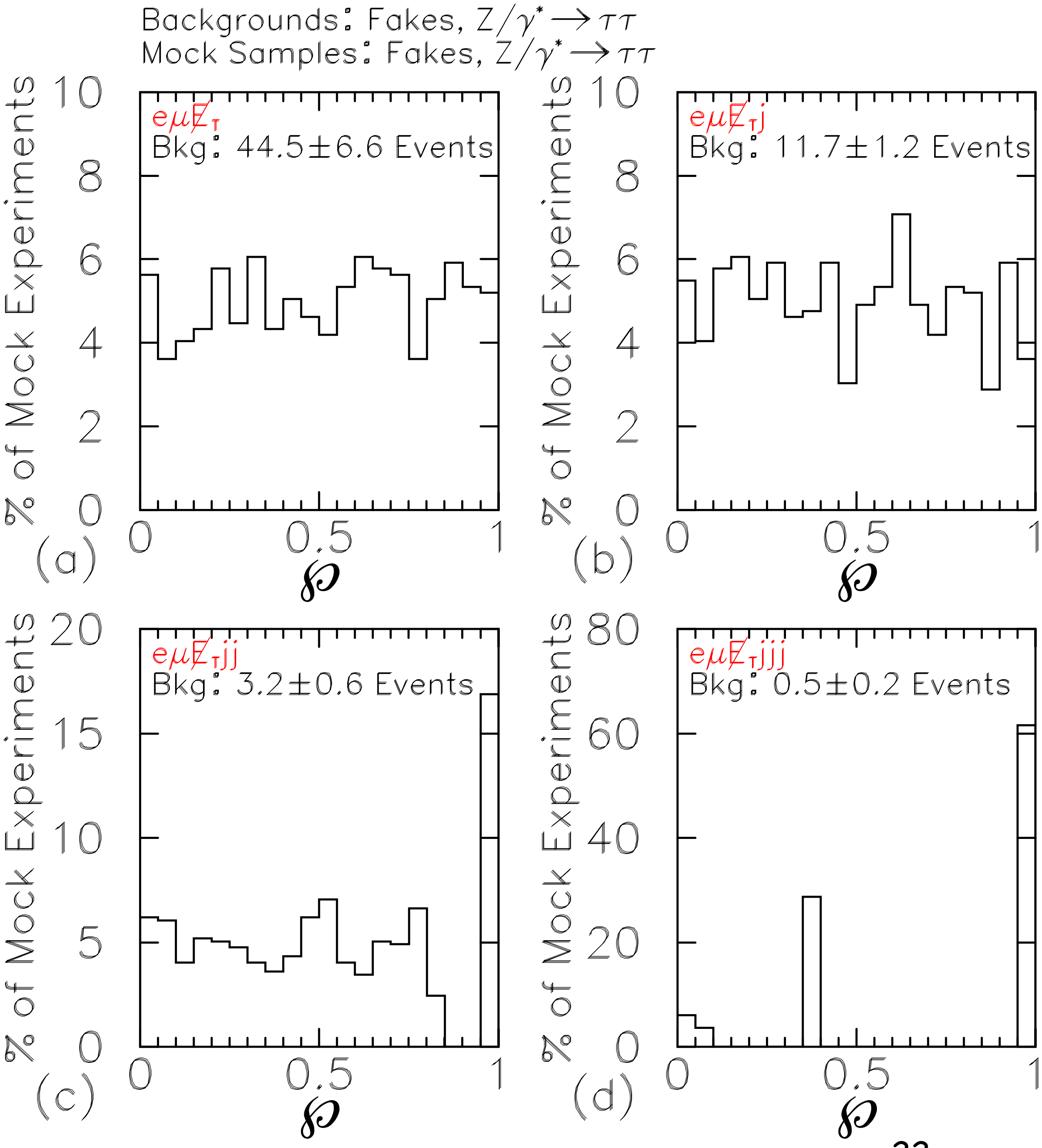
- generate "data samples" from the background distributions
  - Allow numbers of events from each background source to vary according to statistical and systematic errors
- find the most interesting region in each pseudo sample
  - Use same searching algorithm as for the actual data
- compare the most interesting region in each pseudo sample with $\mathcal{R}$
- Determine $\mathcal{P}$, the fraction of *hypothetical similar experiments* in which you see something more interesting than $\mathcal{R}$

Algorithm Example

One entry per mock experiment in these histograms

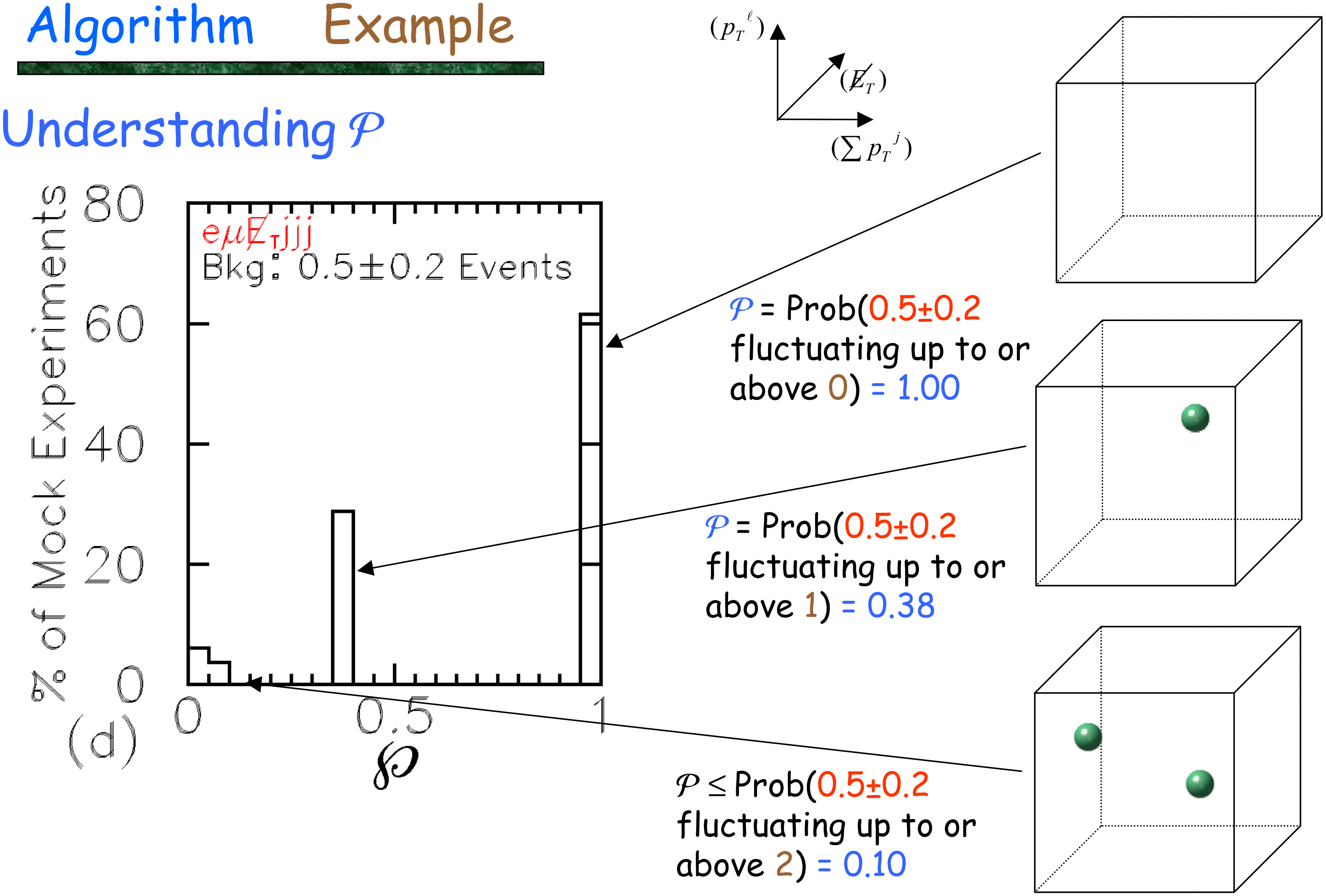
One value of $\mathcal{P}$ is calculated for each mock experiment

$\mathcal{P}$ is the "fraction of hypothetical similar experiments in which you would see something more interesting than what you actually saw" in the mock experiment.



Algorithm Example

Understanding $\mathcal{P}$



$\mathcal{P}$ = Prob(0.5±0.2 fluctuating up to or above 0) = 1.00

$\mathcal{P}$ = Prob(0.5±0.2 fluctuating up to or above 1) = 0.38

$\mathcal{P} \leq$ Prob(0.5±0.2 fluctuating up to or above 2) = 0.10

(depends upon the locations of the points)

$\mathcal{P}$

$\mathcal{P}$ can be written in terms of standard deviations by solving

$$\mathcal{P} = \frac{1}{\sqrt{2\pi}} \int_{\mathcal{P}_{[\sigma]}}^{\infty} e^{-t^2/2}\, dt$$

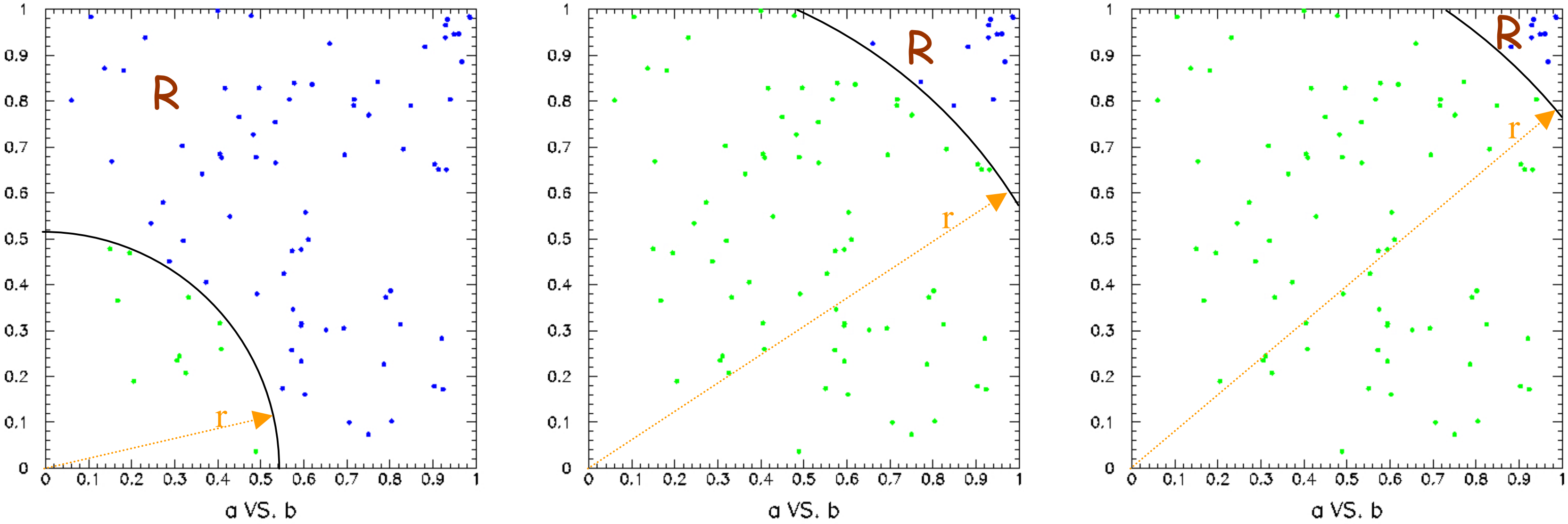
for $\mathcal{P}_{[\sigma]}$

Algorithm   Region criteria

AntiCornerSphere: A region **R** is said to satisfy *AntiCornerSphere* if one can find a number r, such that all data events inside the region are at a distance > r from the origin, and all data events outside the region are at a distance < r from the origin.

Data inside region: blue dots

Data outside region: green dots



We decided for simplicity to impose AntiCornerSphere on the regions used for our initial analysis (eμX)

For remaining final states, we apply more general criteria

Algorithm

Region criteria

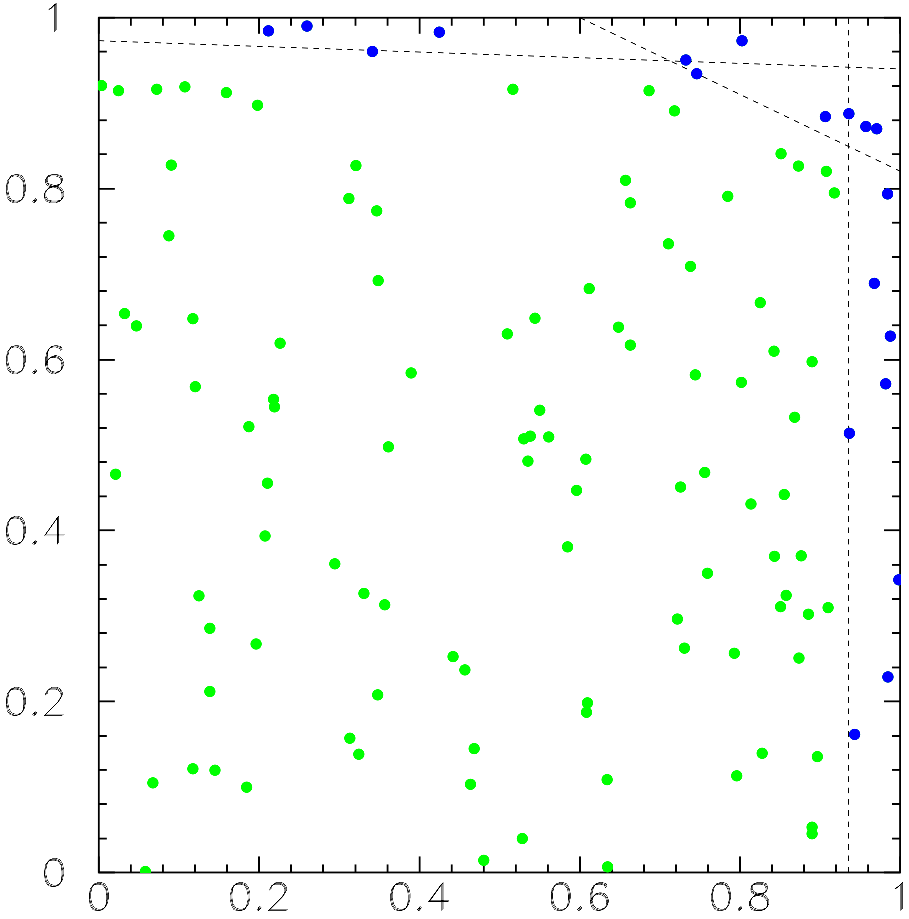
In particular,

Hyperplanes: A region R in a d-dimensional unit box is said to satisfy *Hyperplanes* if, for each data point p inside R, one can draw a (d-1)-dimensional hyperplane through p such that all data points on the side of the hyperplane containing the point 1 (the "upper right-hand corner of the unit box") are inside R.

ReasonableSize: We require all regions to contain fewer than 50 data points.

Data points inside region: blue dots

Data points outside region: green dots



Sleuth

A quasi-model-independent new physics search strategy

Motivation

Strategy

Algorithm

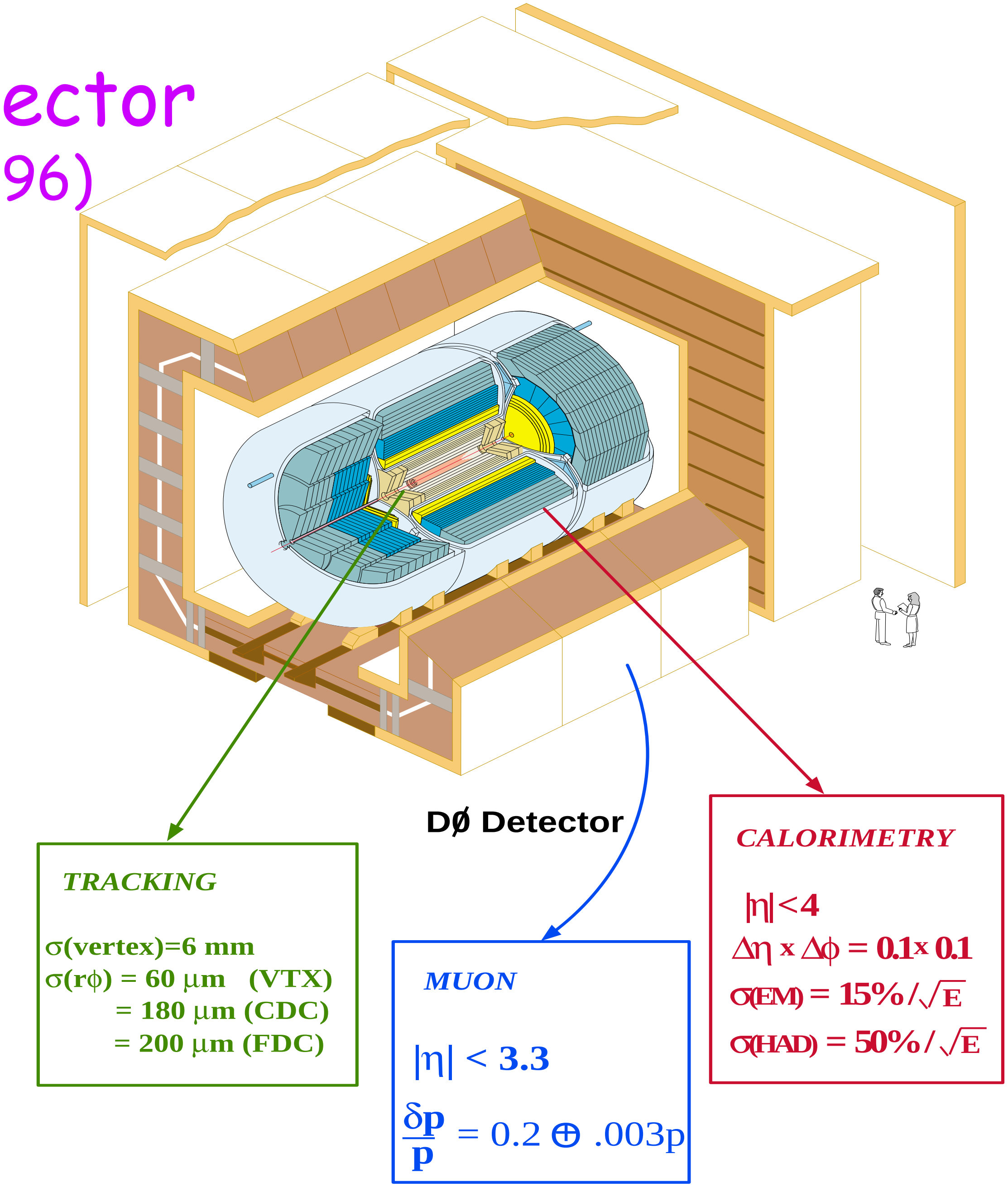
Results →

- $e\mu X$
- W+jets-like
- Z+jets-like
- $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$

Run I detector

(1992-1996)

- Multipurpose detector
  - central tracking
  - muon spectrometer
  - U-LAr sampling calorimeter
- No central magnetic field
- Excellent electromagnetic and hadronic calorimeters



DØ Detector

TRACKING

σ(vertex)=6 mm
$\sigma(r\phi)$ = 60 μm (VTX)
= 180 μm (CDC)
= 200 μm (FDC)

MUON

$|\eta| < \mathbf{3.3}$

$$\frac{\delta p}{p} = 0.2 \oplus .003p$$

CALORIMETRY

$|\eta|<4$
$\Delta\eta \times \Delta\phi = 0.1 \times 0.1$
$\sigma(EM) = 15\%/\sqrt{E}$
$\sigma(HAD) = 50\%/\sqrt{E}$

Particle identification

- Selection criteria (ideally)*:
  - electrons

fiducial	$\lvert\eta\rvert < 1.1$ or $1.5 < \lvert\eta\rvert < 2.5$
identification	based on track match, dE/dx, cluster shape, TRD
high p_T	$p_T > 15$ GeV
isolated	◉

  - photons

fiducial	$\lvert\eta\rvert < 1.1$ or $1.5 < \lvert\eta\rvert < 2.5$
identification	based on cluster shape, track veto
high p_T	$p_T > 15$ GeV
isolated	◉

  - muons

fiducial	$\lvert\eta\rvert < 1.7$
identification	timing, good hits
high p_T	$p_T > 15$ GeV
isolated	⊙

  - taus
    - not identified
  - jets

fiducial	$\lvert\eta\rvert < 2.5$
identification	cone algorithm (R = 0.5)
high p_T	$p_T > 15$ GeV

  - b, c quarks
    - not identified
  - missing transverse energy
    - $\not{E}_T > 15$ GeV
    - "significant"
  - W bosons

$e\not{E}_T$	$30 < m_T^{e\nu} < 110$
$\mu\not{E}_T$	
no second charged lepton	

  - Z bosons

$ee(\gamma)$	$82 < m_{ee(\gamma)} < 100$
$\mu\mu$	$\chi^2(m_{\mu\mu})$

---

*necessary deviations specified later

Systematics

Systematic errors vary among the final states we consider, but roughly:

Systematic uncertainties

jet modeling	20%
trigger / lepton ID eff	10%
cross sections	10%
"faking" probabilities	10%
luminosity	5%

These are handled by the replacement

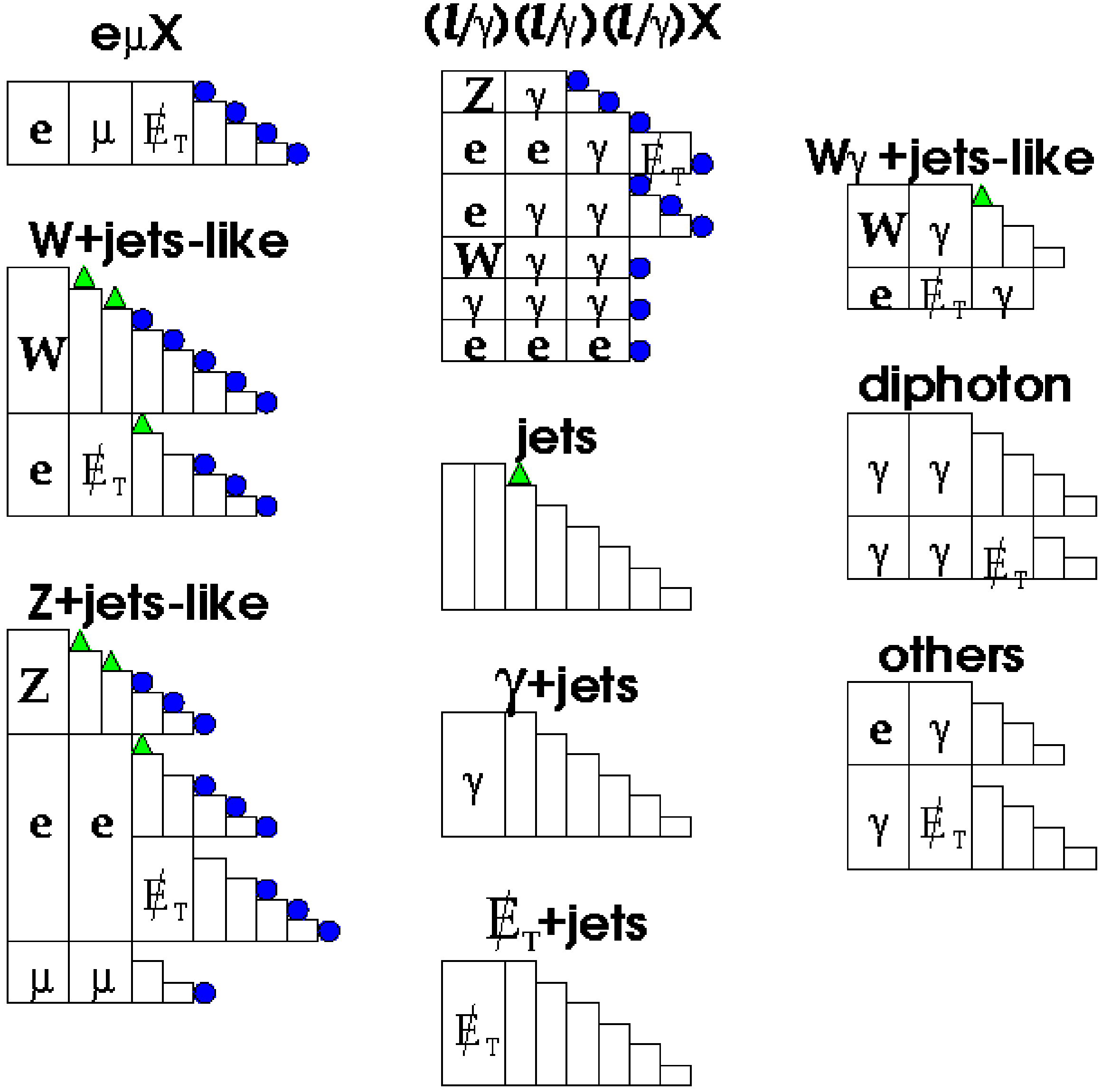
$$p_N{}^R = \sum_{k=N}^{\infty} \frac{e^{-b_R} b_R{}^k}{k!} \rightarrow \quad p_N{}^R = \int db'_R \, \frac{1}{\sqrt{2\pi}\,(\delta\hat{b}_R)} \exp\left(- \frac{(b'_R - \hat{b}_R)^2}{2(\delta\hat{b}_R)^2} \right) \sum_{k=N}^{\infty} \frac{e^{-b'_R} b'_R{}^k}{k!}$$

DØ data

Final states

There were ≈ 80 populated final states at DØ in Run I.

We have applied Sleuth to roughly half of these final states.



- ● analyzed with **Sleuth**
- ▲ analyzed in a spirit similar to **Sleuth**

Sleuth

A quasi-model-independent new physics search strategy

Motivation

Strategy

Algorithm

Results →

- $e\mu X$
- W+jets-like
- Z+jets-like
- $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$

Selection criteria

- Integrated luminosity: 108.3 ± 5.7 pb^{-1}
- Selection criteria:
  - one or more electrons
    - $p_T > 15$ GeV
  - one or more muons
    - $p_T > 15$ GeV
  - zero or more jets
    - $p_T > 15$ GeV
  - no requirement on missing transverse energy
- Leaves 58 events

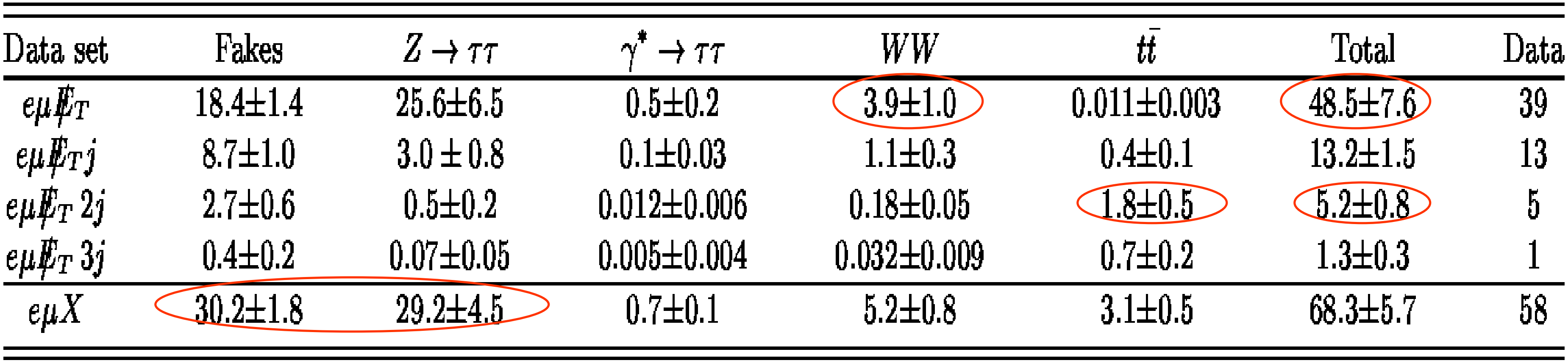
Backgrounds

Dominant backgrounds

- $Z/\gamma^{*} \rightarrow \tau\tau \rightarrow e\mu X$
  - modeled using ISAJET
- "fakes" $\begin{cases} b\bar{b}/c\bar{c} \rightarrow e_{fake}\mu X \\ jW \rightarrow j\mu\nu \rightarrow e_{fake}\mu X \end{cases}$
  - modeled using "bad electron" data
- $WW \rightarrow e\mu X$
  - modeled using Pythia
- $t\bar{t} \rightarrow WWb\bar{b} \rightarrow e\mu X$
  - modeled using Herwig

$e\mu X$ Backgrounds

Events expected



Data set	Fakes	$Z \to \tau\tau$	$\gamma^* \to \tau\tau$	WW	$t\bar{t}$	Total	Data
$e\mu\not{E}_T$	18.4±1.4	25.6±6.5	0.5±0.2	3.9±1.0	0.011±0.003	48.5±7.6	39
$e\mu\not{E}_T j$	8.7±1.0	3.0 ± 0.8	0.1±0.03	1.1±0.3	0.4±0.1	13.2±1.5	13
$e\mu\not{E}_T\, 2j$	2.7±0.6	0.5±0.2	0.012±0.006	0.18±0.05	1.8±0.5	5.2±0.8	5
$e\mu\not{E}_T\, 3j$	0.4±0.2	0.07±0.05	0.005±0.004	0.032±0.009	0.7±0.2	1.3±0.3	1
$e\mu X$	30.2±1.8	29.2±4.5	0.7±0.1	5.2±0.8	3.1±0.5	68.3±5.7	58

- dominant backgrounds are $Z\to\tau\tau$ and “fakes”
- WW contributes ≈ 4 events in $e\mu\not{E}_T$ out of ≈ 49
- $t\bar{t}$ contributes ≈ 2 events in $e\mu\not{E}_T jj$ out of ≈ 5
- good agreement between total numbers expected and observed

Sensitivity check: WW and $t\bar{t}$

We are about to determine how "sensitive" Sleuth is to WW and $t\bar{t}$ in $e\mu X$

To put these signals in context:

DØ's top discovery PRL (1995, 50 pb^{-1}):

- all channels: 17 events with 3.8 ± 0.6 expected — a 4.6σ "effect"
- $e\mu X$ alone: 2 events with 0.12 ± 0.03 expected — a 2.5σ "effect"

DØ's top cross section PRL (1997, 125 pb^{-1}):

- all channels: 39 events with 13.7 ± 2.2 expected
- $e\mu X$ alone: 3 events with 0.21 ± 0.16 expected — a 2.75σ "effect"

Sleuth should never be more sensitive than a dedicated search, so $\approx 2.75\sigma$ is an upper bound on our sensitivity to $t\bar{t}$

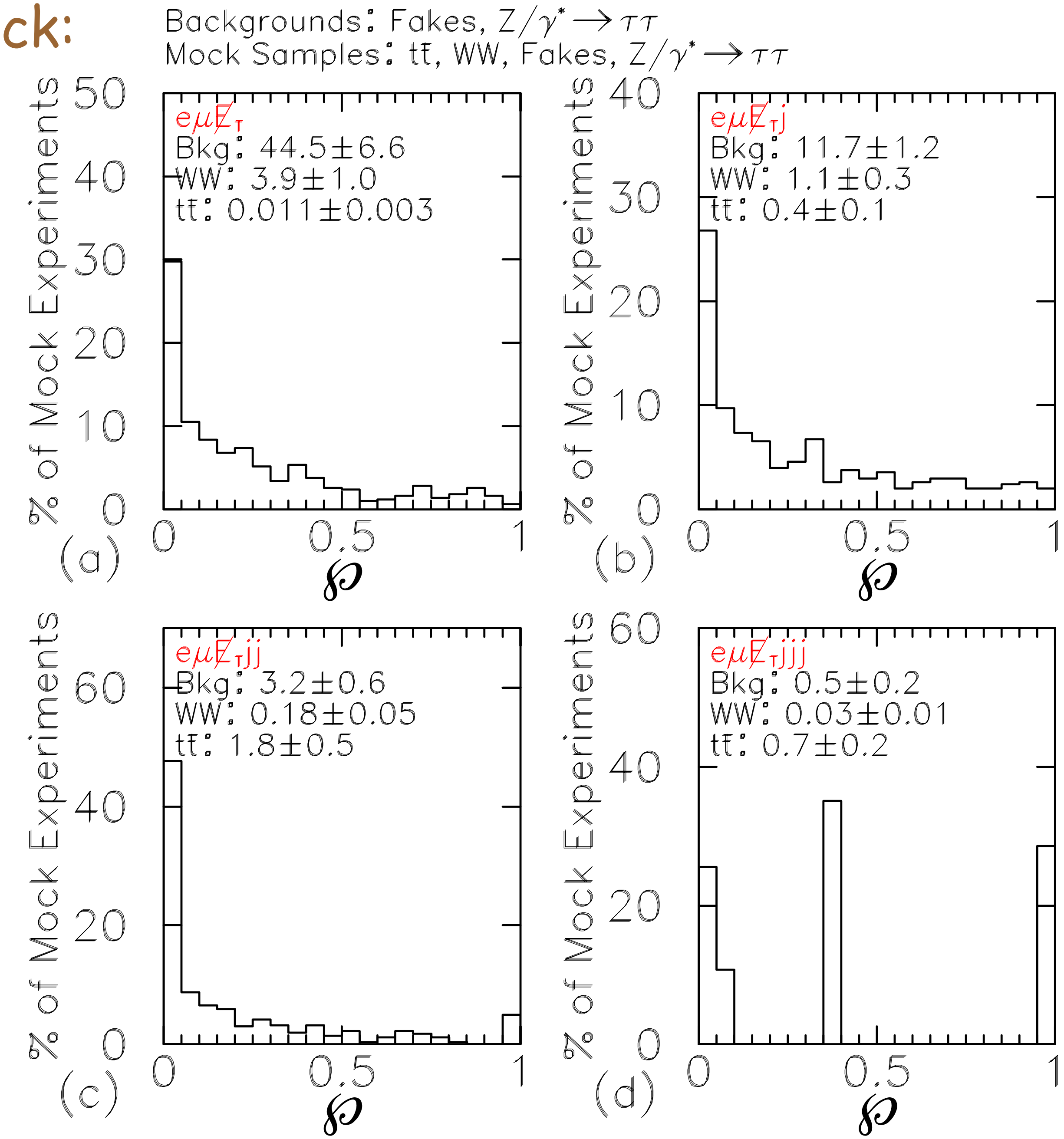
(We've given ourselves a difficult test)

eμX Sensitivity check: WW and $t\bar{t}$

Add WW and $t\bar{t}$ events to the mock experiments, but not to the background estimate

The numbers of WW and $t\bar{t}$ events are allowed to fluctuate according to statistical and systematic errors

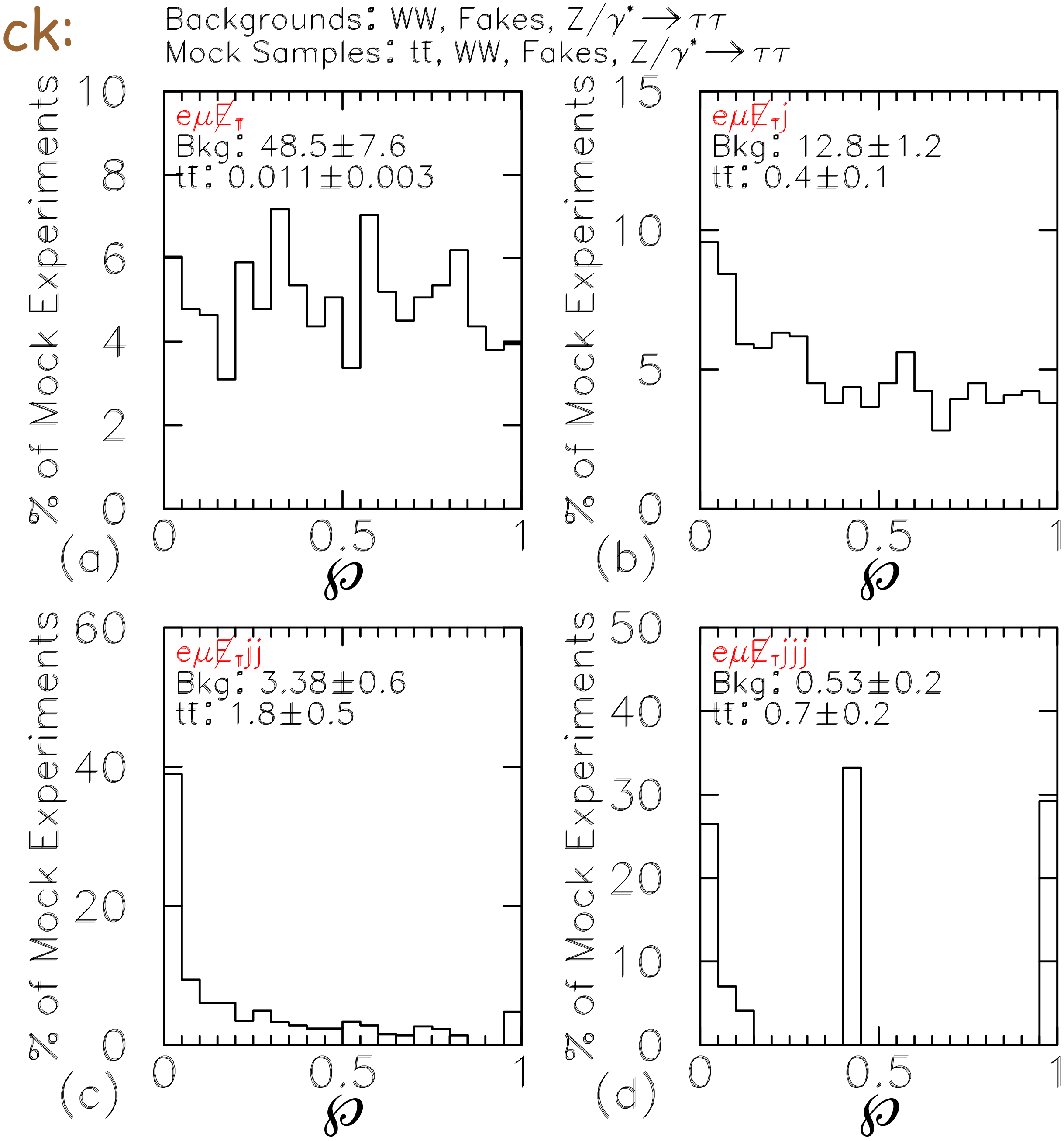
We see that $\mathcal{P}$ is often small, due to WW in (a) and (b), and $t\bar{t}$ in (c) and (d)



eμX Sensitivity check: $t\bar{t}$

Continue to add WW and $t\bar{t}$ events to the mock experiments, but now add WW to the background estimate

We see that $\mathcal{P}$ is often small, due to a bit of $t\bar{t}$ in (b), and more in (c) and (d)



eμX

DØ data

Let the backgrounds include

1)

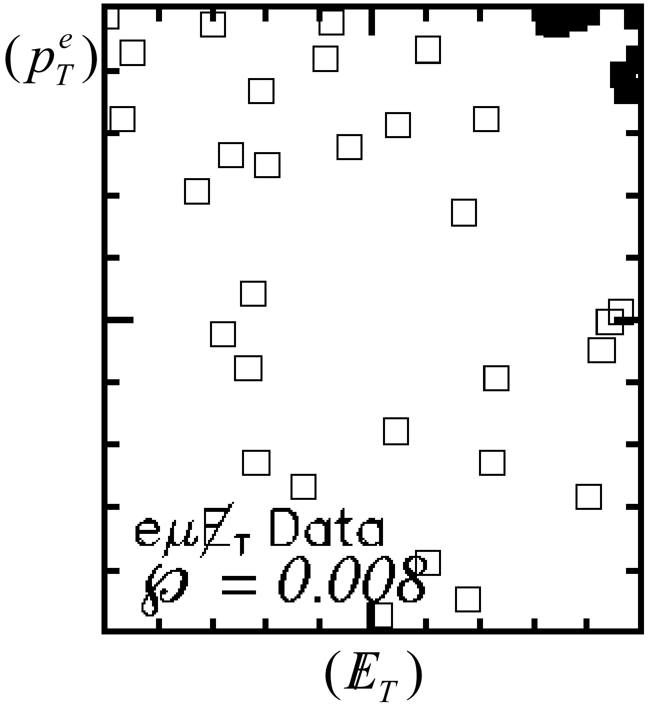
- fakes
- Z→ττ
- WW
- t$\bar{t}$

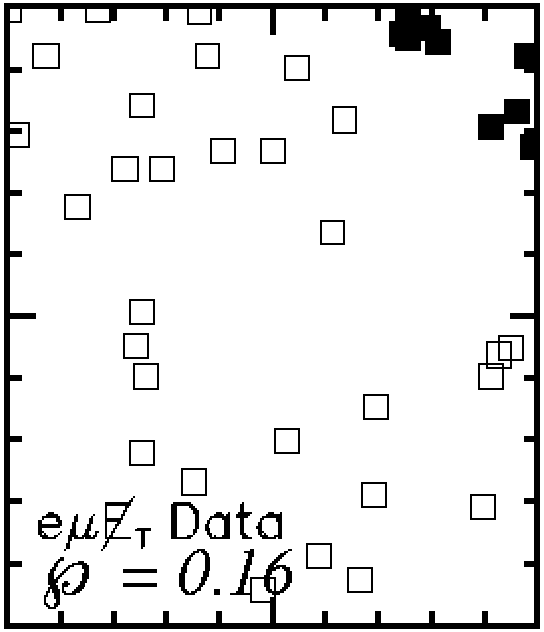
2)

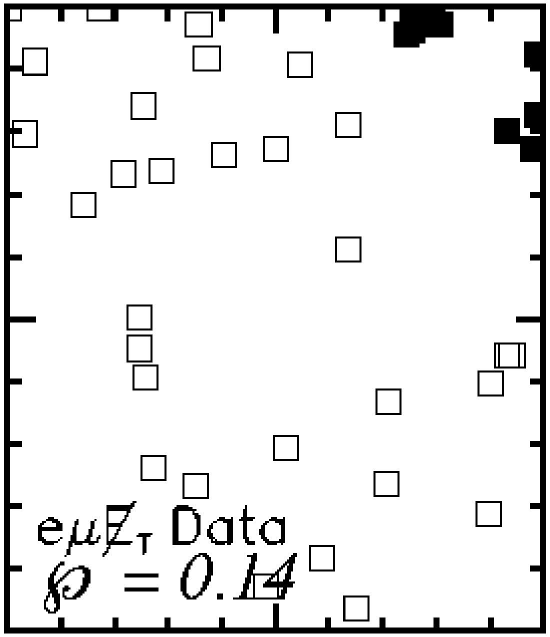
- fakes
- Z→ττ
- WW
- t$\bar{t}$

3)

- fakes
- Z→ττ
- WW
- t$\bar{t}$







eμX

DØ data

Let the backgrounds include

1)
- fakes
- Z→ττ
- WW
- $t\bar{t}$

DØ data

Data Set	$\mathcal{P}$
$e\mu\not\!\!E_T$	2.4σ
$e\mu\not\!\!E_Tj$	0.4σ
$e\mu\not\!\!E_Tjj$	2.3σ
$e\mu\not\!\!E_Tjjj$	0.3σ
Combined	1.9σ

Excesses corresponding (presumably) to WW and $t\bar{t}$

2)
- fakes
- Z→ττ
- WW
- $t\bar{t}$

DØ data

Data Set	$\mathcal{P}$
$e\mu\not\!\!E_T$	1.1σ
$e\mu\not\!\!E_Tj$	0.1σ
$e\mu\not\!\!E_Tjj$	1.9σ
$e\mu\not\!\!E_Tjjj$	0.2σ
Combined	1.2σ

Excess corresponding (presumably) to $t\bar{t}$

3)
- fakes
- Z→ττ
- WW
- $t\bar{t}$

DØ data

Data Set	$\mathcal{P}$
$e\mu\not\!\!E_T$	1.1σ
$e\mu\not\!\!E_Tj$	0.1σ
$e\mu\not\!\!E_Tjj$	0.5σ
$e\mu\not\!\!E_Tjjj$	-0.5σ
Combined	-0.6σ

No evidence for new physics

$\tilde{\mathcal{P}}$

How do we combine the results of several final states?

Introduce $\tilde{\mathcal{P}}$, the fraction of *hypothetical similar experimental runs* in which you would see something more interesting than what you actually observe

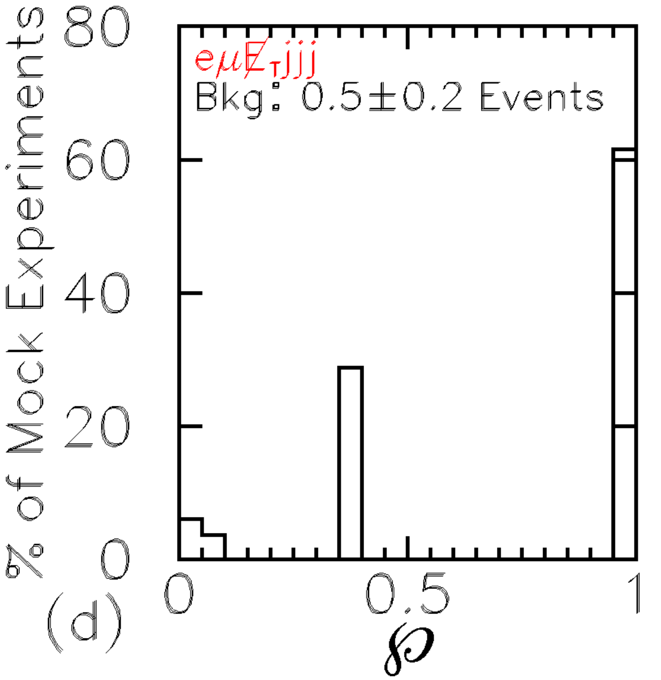
Replace the naive $\tilde{\mathcal{P}} = 1 - (1 - \mathcal{P}_{min})^{N_{fs}}$

by $\tilde{\mathcal{P}} = 1 - \prod_{i=1}^{N_{fs}} (1 - s_i)$

($\mathcal{P}$ is 1 - "the probability that nothing is more interesting")

where

$s_i =$ "the integral of this histogram from 0 up to $\mathcal{P}_{min}$"



Sleuth

A quasi-model-independent new physics search strategy

Motivation

Strategy

Algorithm

Results →

- $e\mu X$
- W+jets-like
- Z+jets-like
- $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$

W+jets-like

Selection criteria

- Electron channel
  - Integrated luminosity: 115 ± 6 pb^{-1}
  - Selection criteria:
    - one electron
      - $p_T > 20$ GeV
    - zero photons, muons
    - two or more jets
      - $p_T > 20$ GeV
    - missing transverse energy
      - $\not{E}_T > 30$ GeV
      - $m_T^{e\nu} > 30$ GeV
      - $\Delta\phi_{j\nu} > 0.25$
      - $p_T^W > 40$ GeV
  - Leaves 470 events
- Muon channel
  - Integrated luminosity: 94 ± 5 pb^{-1}
  - Selection criteria:
    - one muon
      - $p_T > 25$ GeV
      - $|\eta| < 0.95$
    - zero electrons, photons
    - two or more jets
      - $p_T > 15$ GeV
      - $|\eta| < 1.5$ for leading jet
    - missing transverse energy
      - $\not{E}_T > 30$ GeV
      - $|\Delta\phi_{\mu\nu} - \pi| > 0.1$
      - $p_T^W > 40$ GeV
  - Leaves 69 events

W+jets-like Backgrounds

Dominant backgrounds

- Electron channel
  - (W→eν)+jets
    - model with Vecbos + Herwig
  - QCD “fakes”
    - model with data
  - $t\bar{t}$
    - model with Herwig
- Muon channel
  - (W→μν)+jets
    - model with Vecbos + Herwig
  - (Z→μμ)+jets
    - model with Vecbos + Herwig
  - WW, $t\bar{t}$
    - model with Pythia, Herwig

All Monte Carlo events are run through DØGEANT

W+jets-like Backgrounds

Events expected

DØ preliminary

Final State	W+jets	QCD fakes	$t\bar{t}$	Total	Data
$e\not\!\!E_T\,2j$	6.7 ± 1.4	3.3 ± 0.9	1.7 ± 0.6	11.6 ± 1.7	7
$e\not\!\!E_T\,3j$	1.0 ± 0.4	0.48 ± 0.22	1.0 ± 0.4	2.5 ± 0.6	5
$e\not\!\!E_T\,4j$	0.15 ± 0.11	0.38 ± 0.19	0.26 ± 0.09	0.80 ± 0.24	2
$W(\to e\not\!\!E_T)\,2j$	333.9 ± 50.5	12.0 ± 2.6	4.0 ± 1.4	349.9 ± 50.6	387
$W(\to e\not\!\!E_T)\,3j$	57.0 ± 9.0	3.4 ± 0.9	6.0 ± 2.1	66.3 ± 9.3	56
$W(\to e\not\!\!E_T)\,4j$	5.9 ± 1.3	1.1 ± 0.4	3.9 ± 1.4	10.9 ± 1.9	11
$W(\to e\not\!\!E_T)\,5j$	0.8 ± 0.3	0.19 ± 0.12	0.73 ± 0.26	1.8 ± 0.4	1
$W(\to e\not\!\!E_T)\,6j$	0.12 ± 0.06	0.030 ± 0.015	0.10 ± 0.04	0.25 ± 0.07	1

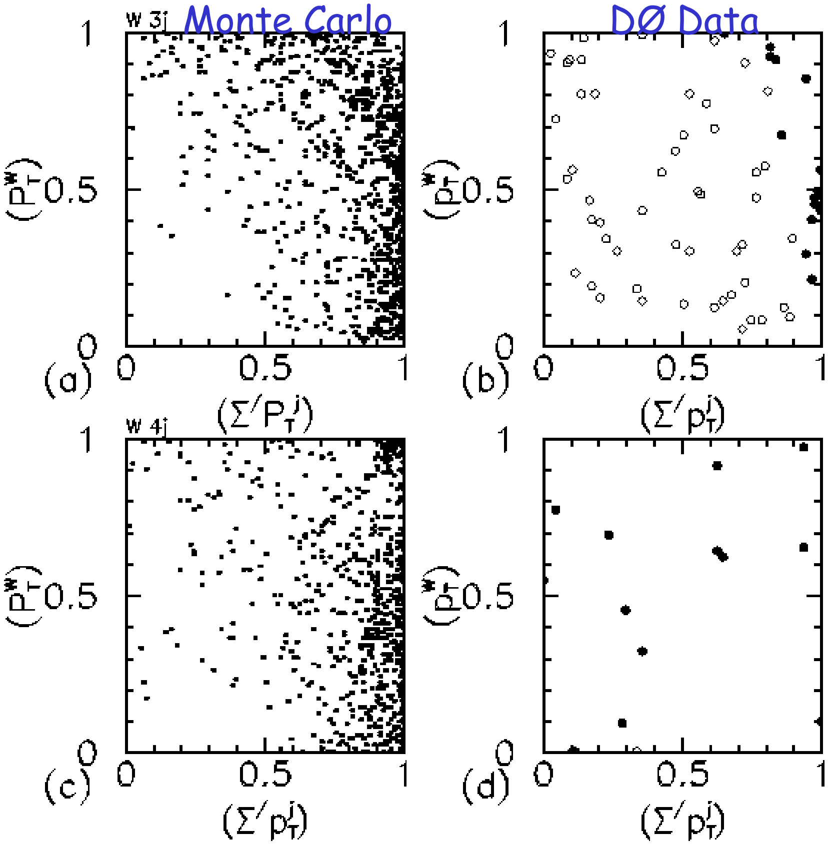
Final State	W+jets	Z+jets	WW	$t\bar{t}$	Total	Data
$W(\to \mu\not\!\!E_T)\,2j$	47.7 ± 14.7	1.6 ± 0.4	0.5 ± 0.3	0.42 ± 0.14	50.2 ± 14.7	54
$W(\to \mu\not\!\!E_T)\,3j$	9.5 ± 3.4	0.27 ± 0.08	0.41 ± 0.26	0.58 ± 0.20	10.8 ± 3.4	11
$W(\to \mu\not\!\!E_T)\,4j$	2.8 ± 1.3	0.022 ± 0.011	–	0.61 ± 0.21	3.5 ± 1.3	4

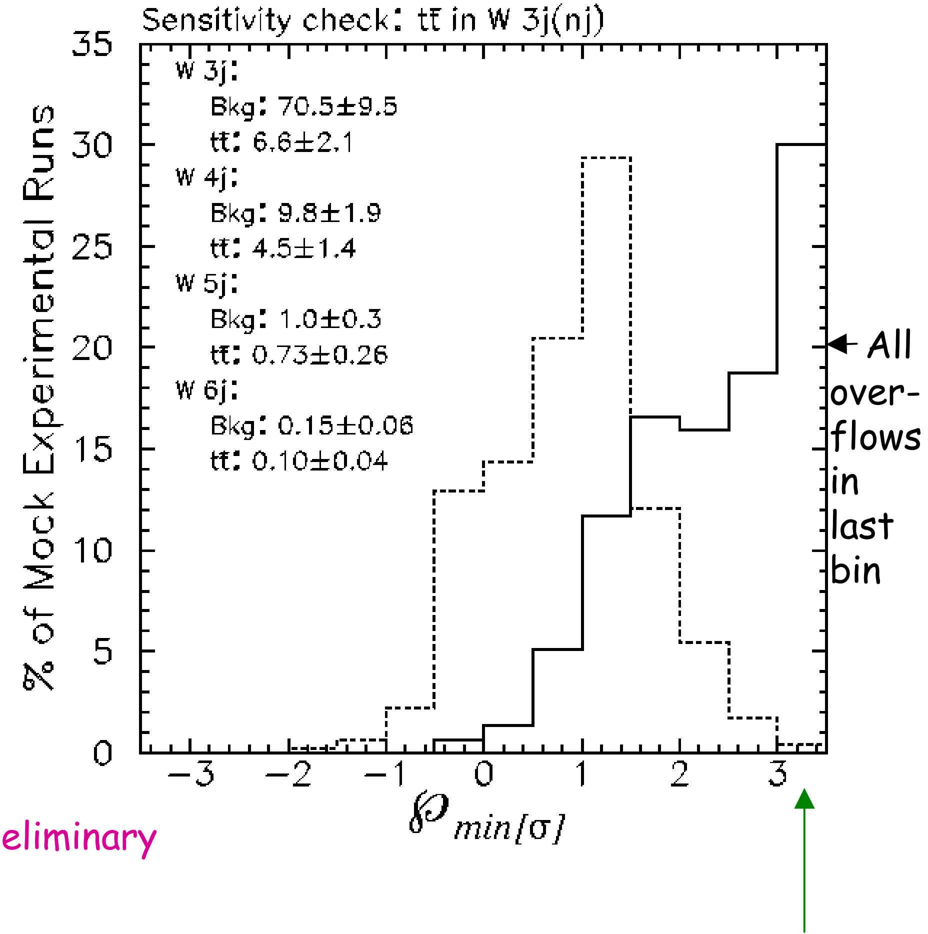
- dominant backgrounds are W+jets
- $t\bar{t}$ contributes ≈ 7 events in W 3j out of ≈ 77
- good agreement between total numbers expected and observed

W+jets-like

Sensitivity check: $t\bar{t}$

Could Sleuth have found $t\bar{t}$ in the lepton+jets channel?





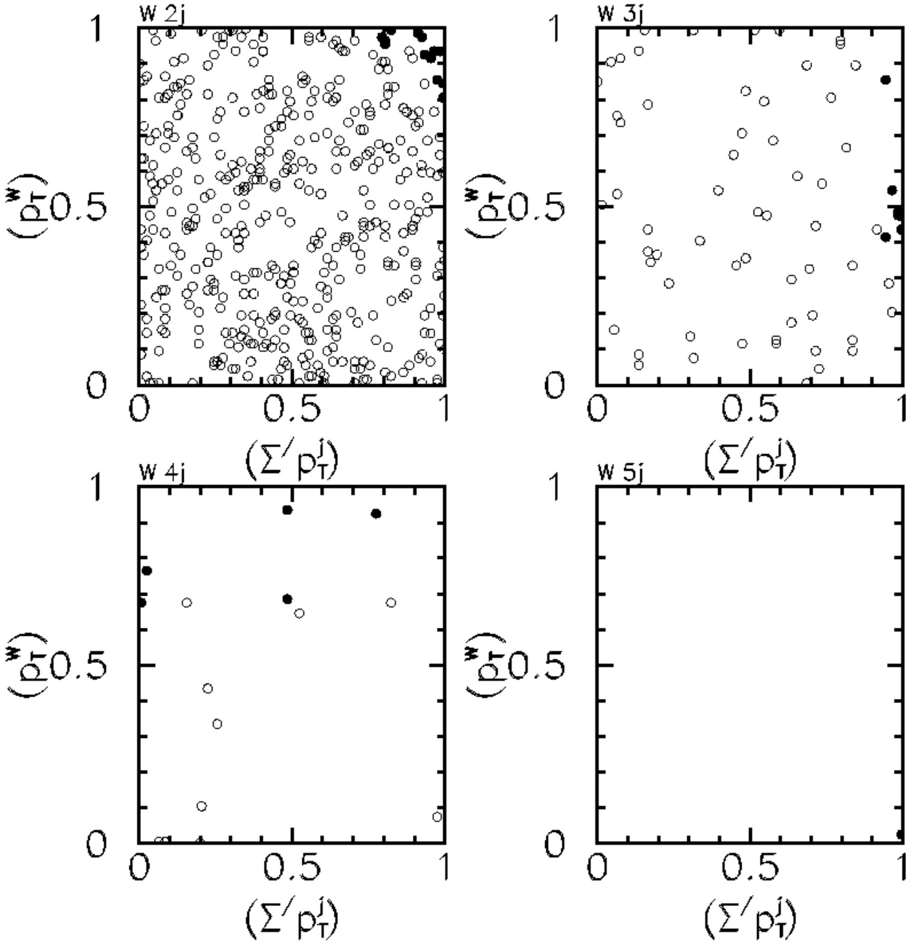
DØ preliminary

Sleuth finds $\mathcal{P}_{min} > 3\sigma$ in 30% of an ensemble of mock experimental runs

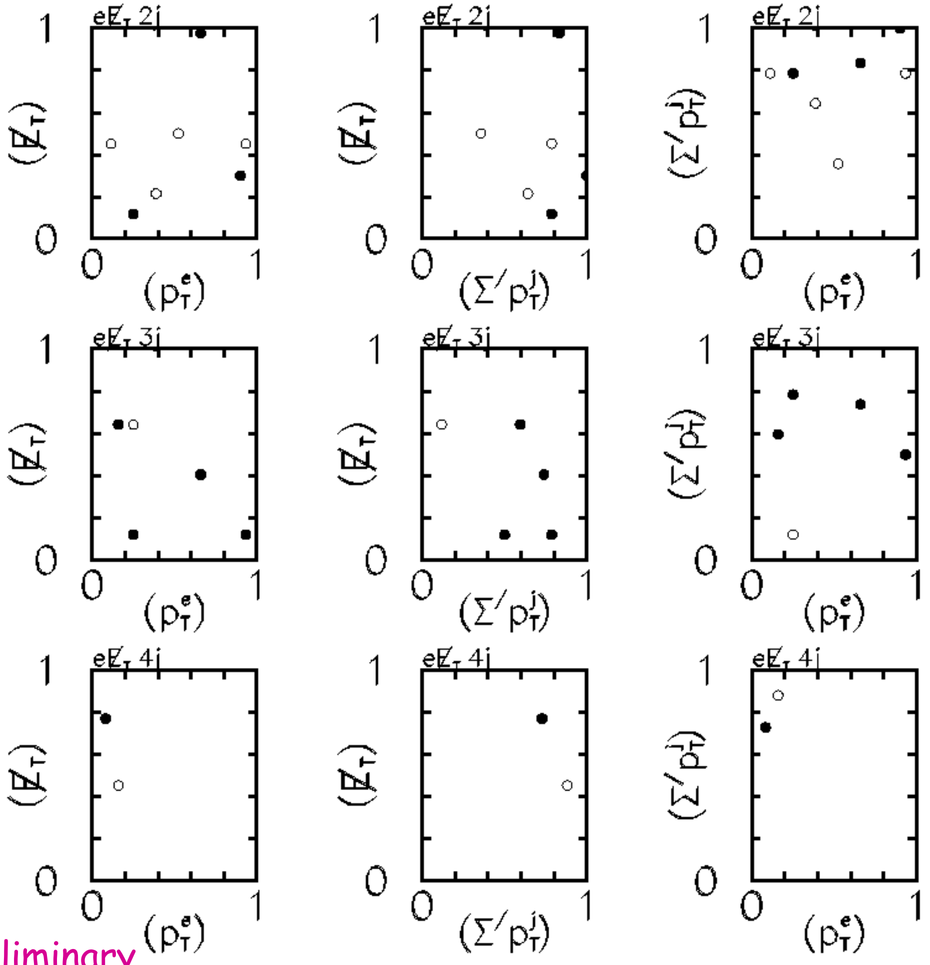
W+jets-like

DØ data

Wjj(nj)



$e\not\!\!E_T jj(nj)$



DØ preliminary

W+jets-like

Results

DØ preliminary

Data set	$\mathcal{P}$
$e\not\!\!E_T\ 2j$	0.76
$e\not\!\!E_T\ 3j$	0.17
$e\not\!\!E_T\ 4j$	0.13
$W\ 2j$	0.29
$W\ 3j$	0.23
$W\ 4j$	0.53
$W\ 5j$	0.81
$W\ 6j$	0.22

No hints of new high p_T physics observed

Sleuth

A quasi-model-independent new physics search strategy

Motivation

Strategy

Algorithm

Results →

- $e\mu X$
- W+jets-like
- Z+jets-like
- $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$

Z+jets-like

Selection criteria

- Electron channel
  - Integrated luminosity: 123 ± 6 pb^{-1}
  - Selection criteria:
    - two electrons
      - $p_T > 20$ GeV
    - zero photons, muons
    - two or more jets
      - $p_T > 20$ GeV
  - Leaves 137 events
- Muon channel
  - Integrated luminosity: 94 ± 5 pb^{-1}
  - Selection criteria:
    - two muons
      - $p_T > 20$ GeV
      - $|\eta| < 1.0, 1.7$
    - zero electrons, photons
    - two or more jets
      - $p_T > 20$ GeV
  - Leaves 6 events

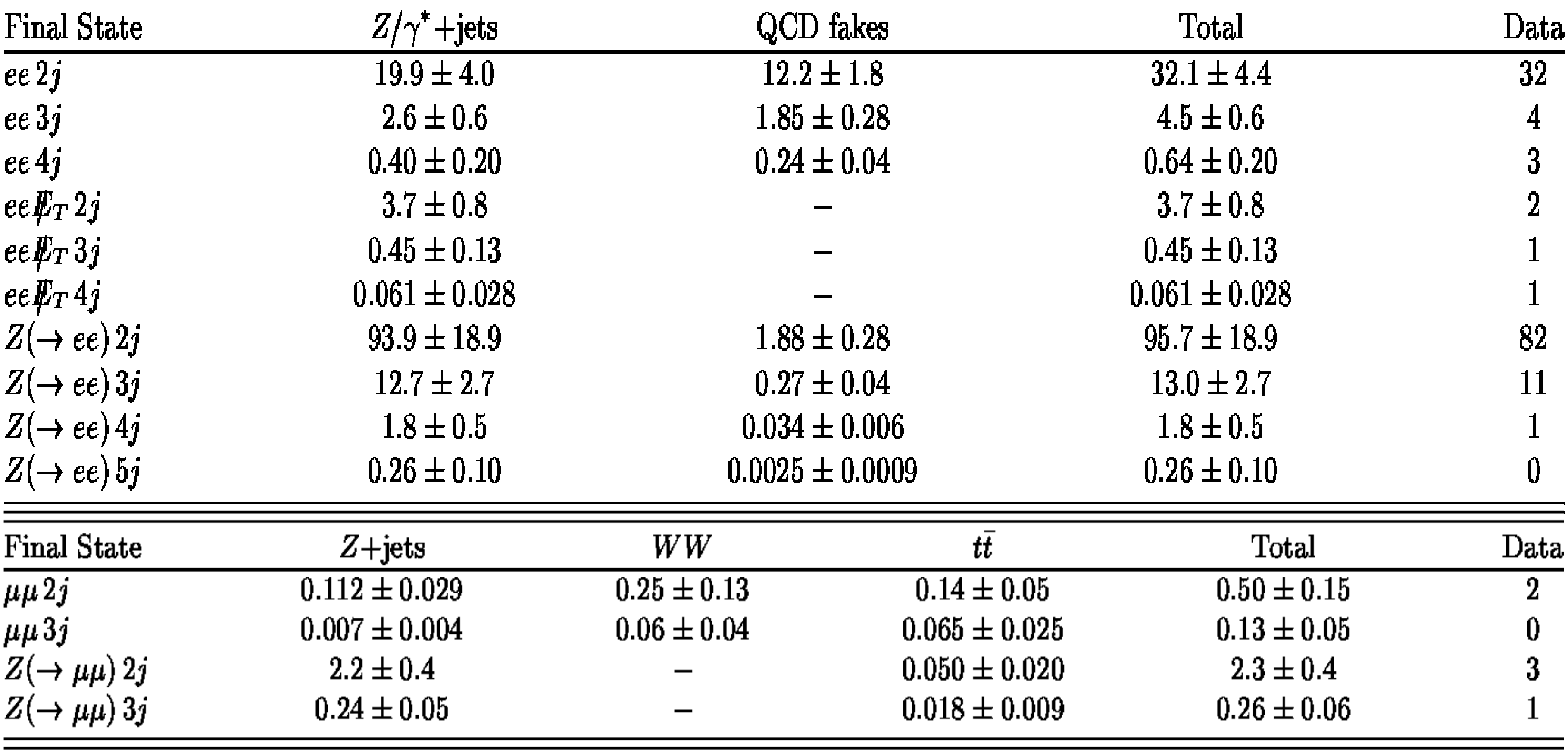
Dominant backgrounds

- Electron channel
  - $(Z/\gamma^*\to ee)$+jets
    - model with Isajet
    - normalization fixed to $Z+\geq 2$jets data in Z boson region
  - QCD "fakes"
    - model with data
- Muon channel
  - $(Z/\gamma^*\to \mu\mu)$+jets
    - model with Vecbos + Herwig
  - $(WW \to \mu\mu\nu\nu)$+jets
    - model with Pythia
  - $(t\bar{t}\to \mu\mu\nu\nu jj)$+jets
    - model with Herwig

All Monte Carlo events are run through DØGEANT

Z+jets-like Backgrounds

Events expected


DØ preliminary

Final State	Z/γ^*+jets	QCD fakes	Total	Data
$ee\,2j$	19.9 ± 4.0	12.2 ± 1.8	32.1 ± 4.4	32
$ee\,3j$	2.6 ± 0.6	1.85 ± 0.28	4.5 ± 0.6	4
$ee\,4j$	0.40 ± 0.20	0.24 ± 0.04	0.64 ± 0.20	3
$ee\not\!\!E_T\,2j$	3.7 ± 0.8	–	3.7 ± 0.8	2
$ee\not\!\!E_T\,3j$	0.45 ± 0.13	–	0.45 ± 0.13	1
$ee\not\!\!E_T\,4j$	0.061 ± 0.028	–	0.061 ± 0.028	1
$Z(\to ee)\,2j$	93.9 ± 18.9	1.88 ± 0.28	95.7 ± 18.9	82
$Z(\to ee)\,3j$	12.7 ± 2.7	0.27 ± 0.04	13.0 ± 2.7	11
$Z(\to ee)\,4j$	1.8 ± 0.5	0.034 ± 0.006	1.8 ± 0.5	1
$Z(\to ee)\,5j$	0.26 ± 0.10	0.0025 ± 0.0009	0.26 ± 0.10	0

Final State	Z+jets	WW	$t\bar{t}$	Total	Data
$\mu\mu\,2j$	0.112 ± 0.029	0.25 ± 0.13	0.14 ± 0.05	0.50 ± 0.15	2
$\mu\mu\,3j$	0.007 ± 0.004	0.06 ± 0.04	0.065 ± 0.025	0.13 ± 0.05	0
$Z(\to \mu\mu)\,2j$	2.2 ± 0.4	–	0.050 ± 0.020	2.3 ± 0.4	3
$Z(\to \mu\mu)\,3j$	0.24 ± 0.05	–	0.018 ± 0.009	0.26 ± 0.06	1

- dominant background is Z+jets
- “fakes” become significant outside the Z window
- good agreement between total numbers expected and observed

Z+jets-like

Sensitivity check: Leptoquarks

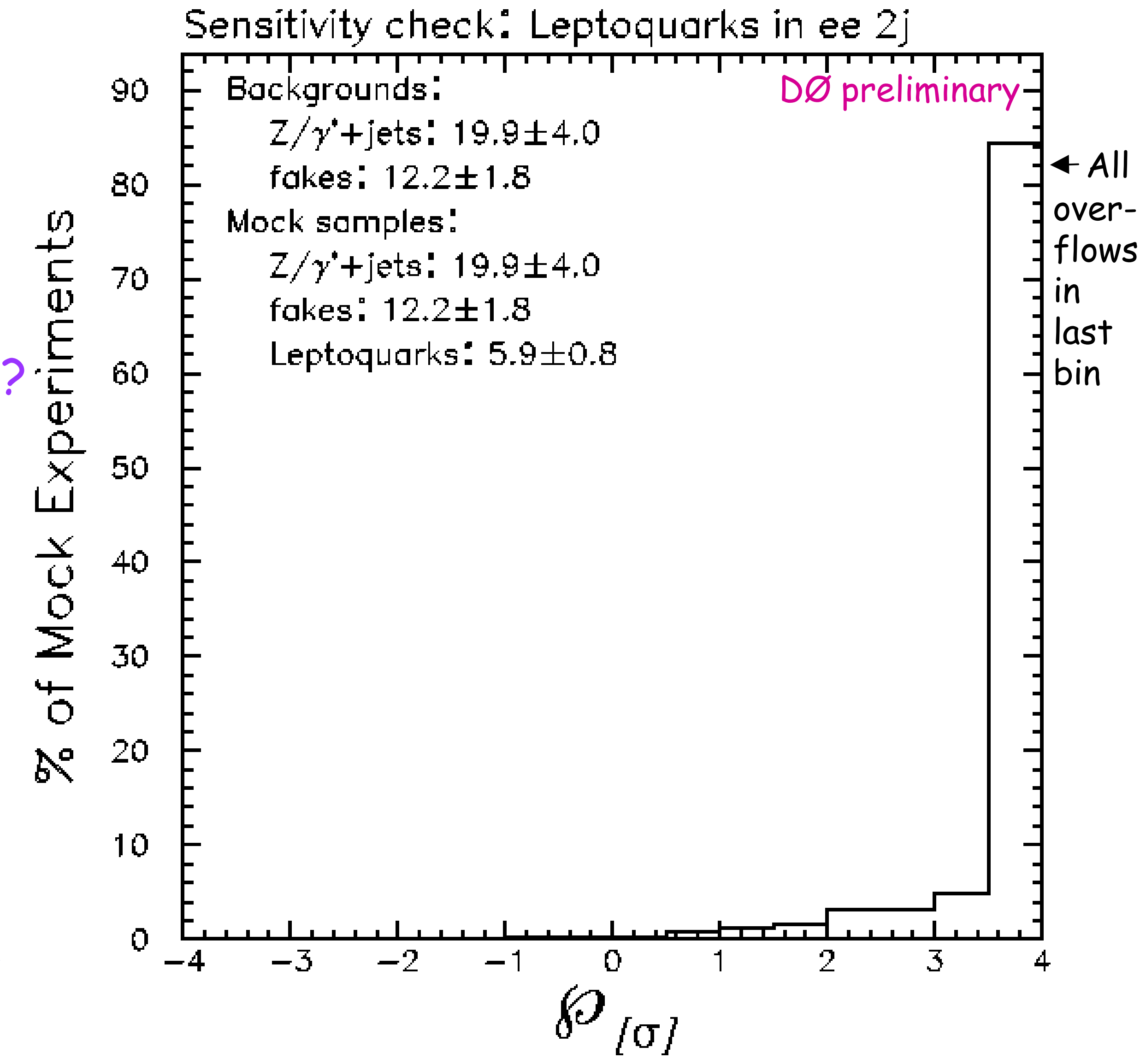
We can also run mock experiments with hypothetical signals

What if our data contained leptoquarks?

(Assume scalar, $\beta = 1$, $m_{LQ} = 170$ GeV)

Sleuth finds $\mathcal{P} > 3.5\sigma$ in > 80% of the mock experiments

(Remember that Sleuth "knows" nothing about leptoquarks!)

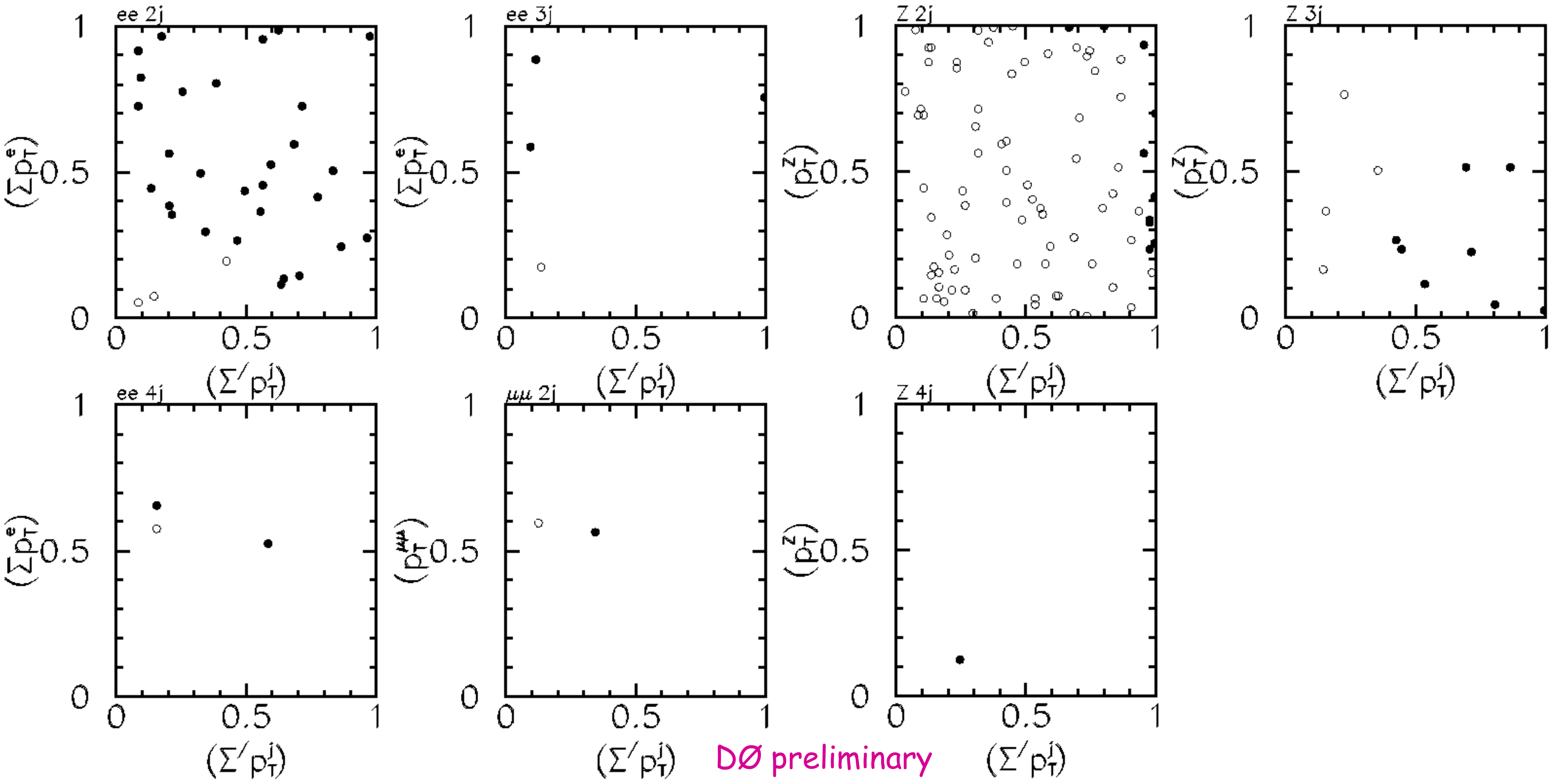


Z+jets-like

DØ data

(ee/μμ)jj(nj)

Zjj(nj)



DØ preliminary

Z+jets-like

DØ data

Results

DØ preliminary

Data set	$\mathcal{P}$
$ee\,2j$	0.72
$ee\,3j$	0.61
$ee\,4j$	0.04
$ee\not{E}_T\,2j$	0.68
$ee\not{E}_T\,3j$	0.36
$ee\not{E}_T\,4j$	0.06
$\mu\mu\,2j$	0.08
$\mu\mu\,3j$	1.00
$Z\,2j$	0.52
$Z\,3j$	0.71
$Z\,4j$	0.83
$Z\,5j$	1.00

No hints of new high p_T physics observed

Sleuth

A quasi-model-independent new physics search strategy

Motivation

Strategy

Algorithm

Results →

- eμX
- W+jets-like
- Z+jets-like
- $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$

$(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$

Selection criteria

- Integrated luminosity: 120 ± 8 pb^{-1}
- Selection criteria:
  - Three or more of
    - electrons
      - $p_T > 15$ GeV
    - photons
      - $p_T > 15$ GeV
    - muons
      - $p_T > 15$ GeV
  - zero or more jets
      - $p_T > 15$ GeV
  - no requirement on missing transverse energy
- Includes e.g. eee, eeγ, e$\gamma\gamma$, $\mu\mu\gamma\gamma$, ee$\mu\not{E}_T$j, etc.
- Leaves 21 events

Backgrounds

- Dominant backgrounds
  - $Z\gamma$
    - model with a LO matrix element Monte Carlo
      - Ulrich Baur
  - WZ(→leptons)
    - model with Pythia

- Lesser backgrounds
  - Zj
    - model with Pythia
  - $W\gamma\gamma$
    - model with a LO matrix element Monte Carlo
      - Ulrich Baur

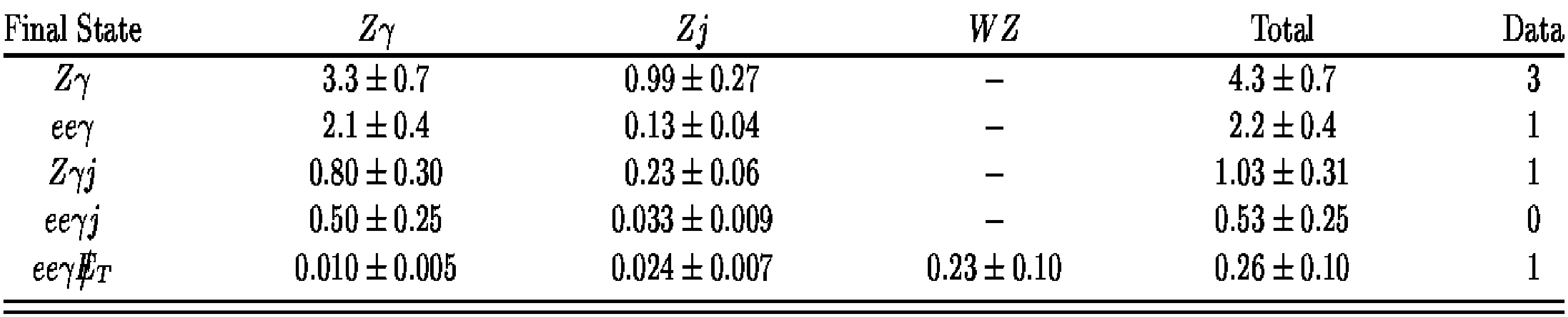
(Mis)Identification matrix

	e	γ
e	0.61 ± 0.04	0.28 ± 0.03
γ	0.16 ± 0.016	0.73 ± 0.012
j	0.00035 ± 0.000035	0.00125 ± 0.00013

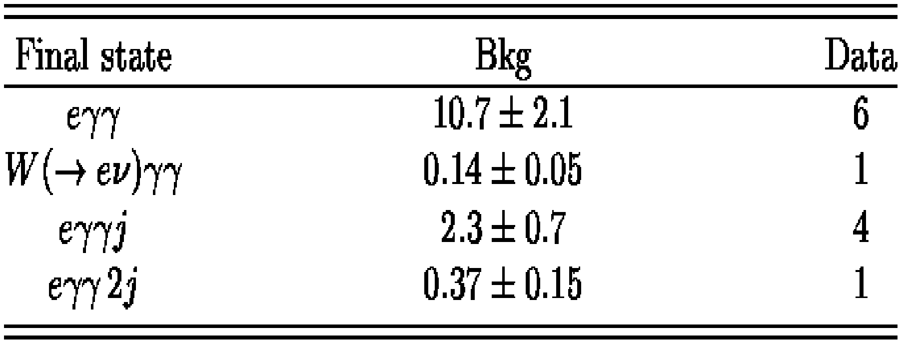
Monte Carlo events are run through a fast smearing routine

Events expected

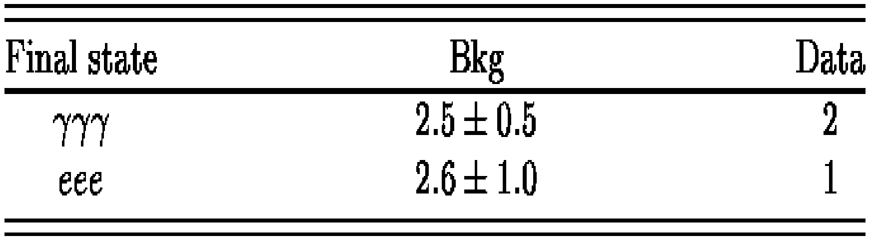
DØ preliminary



Final State	$Z\gamma$	Zj	WZ	Total	Data
$Z\gamma$	3.3 ± 0.7	0.99 ± 0.27	–	4.3 ± 0.7	3
$ee\gamma$	2.1 ± 0.4	0.13 ± 0.04	–	2.2 ± 0.4	1
$Z\gamma j$	0.80 ± 0.30	0.23 ± 0.06	–	1.03 ± 0.31	1
$ee\gamma j$	0.50 ± 0.25	0.033 ± 0.009	–	0.53 ± 0.25	0
$ee\gamma \not{E}_T$	0.010 ± 0.005	0.024 ± 0.007	0.23 ± 0.10	0.26 ± 0.10	1



Final state	Bkg	Data
$e\gamma\gamma$	10.7 ± 2.1	6
$W(\to e\nu)\gamma\gamma$	0.14 ± 0.05	1
$e\gamma\gamma j$	2.3 ± 0.7	4
$e\gamma\gamma 2j$	0.37 ± 0.15	1



Final state	Bkg	Data
$\gamma\gamma\gamma$	2.5 ± 0.5	2
eee	2.6 ± 1.0	1

- dominant background is $Z\gamma$ (possibly with a misidentified EM object)
- WZ is contributes to final states with $\not{E}_T$
- good agreement between total numbers expected and observed

Sleuth

A quasi-model-independent new physics search strategy

Motivation

Strategy

Algorithm

Results

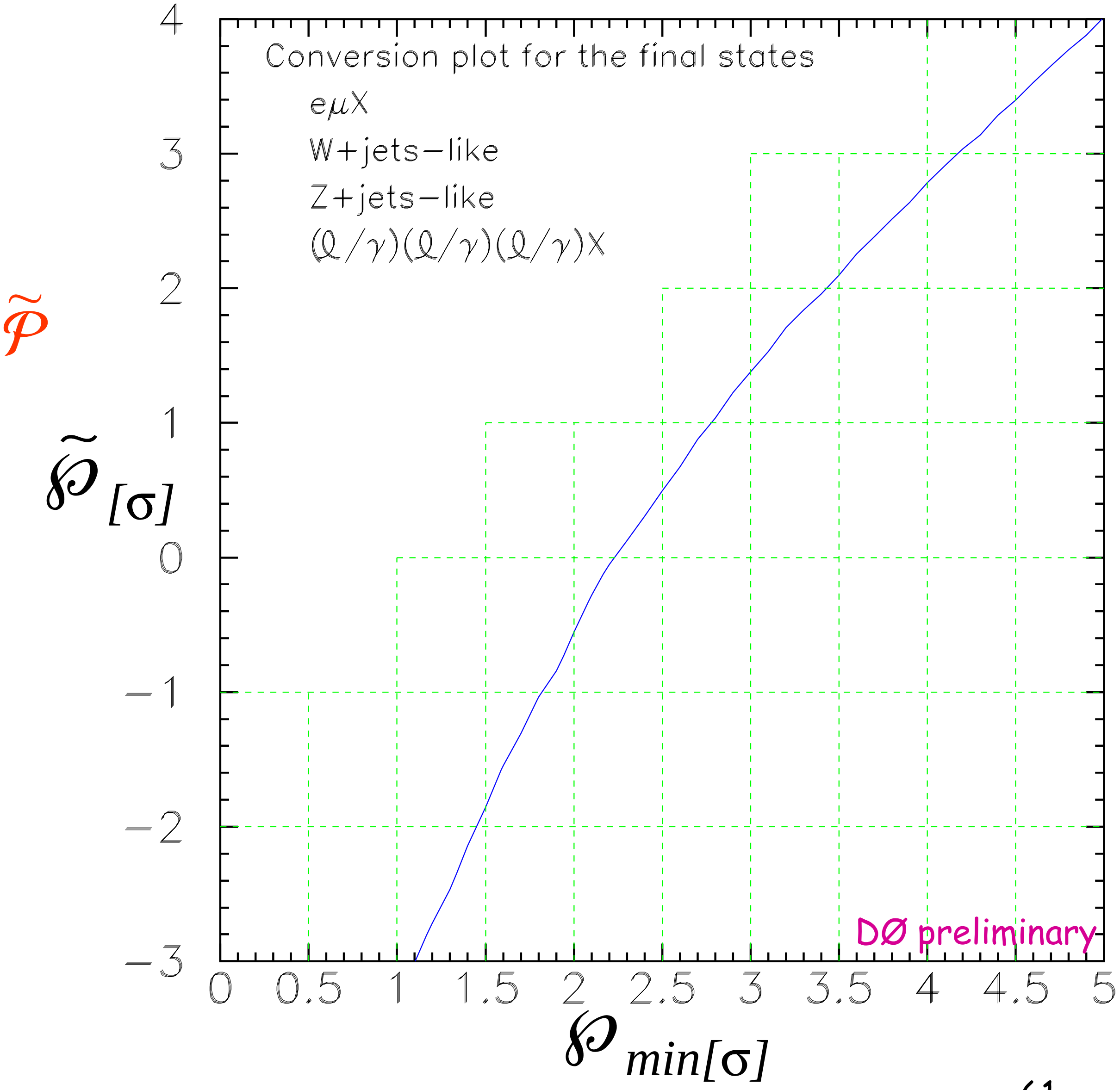
Summary

Results

Combining many final states

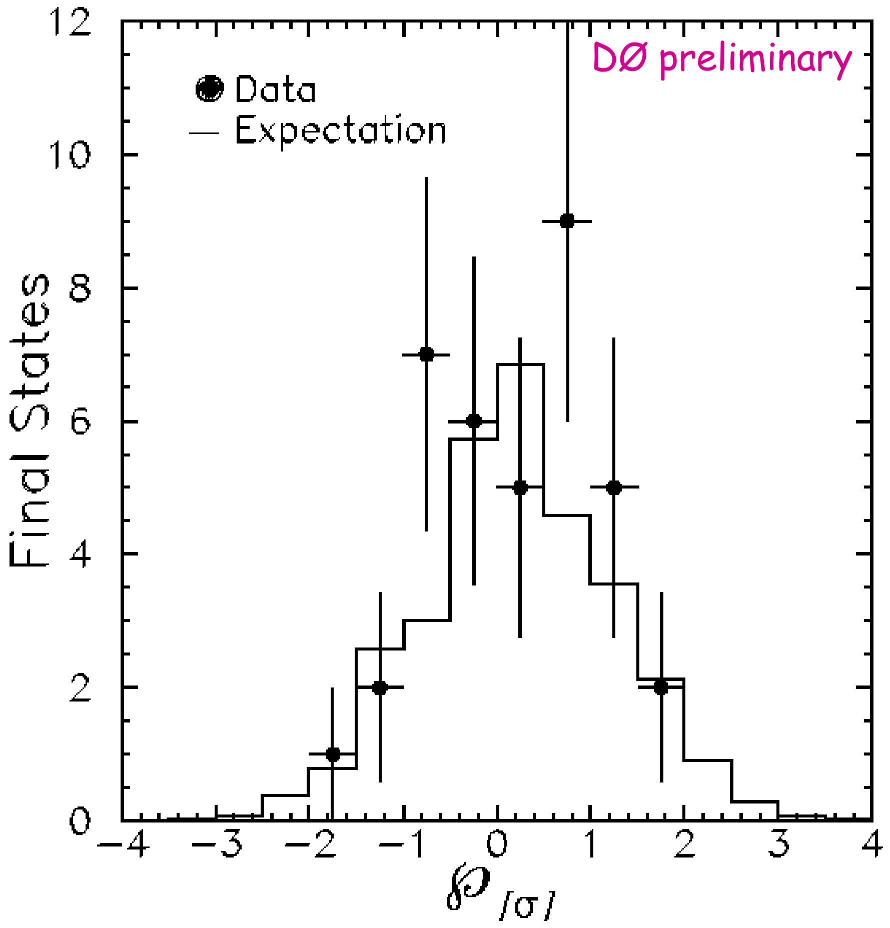
We can account for the fact that we have looked at many different final states by computing $\tilde{\mathcal{P}}$

The correspondence between $\tilde{\mathcal{P}}$ and the minimum $\mathcal{P}$ found for the final states that we have considered is shown here



Results

DØ data



Results agree well with expectation

No evidence of new physics is observed

Data set	DØ preliminary	$\mathcal{P}$
	$e\mu X$	
$e\mu\not{E}_T$		0.14 $(+1.08\sigma)$
$e\mu\not{E}_T j$		0.45 $(+0.13\sigma)$
$e\mu\not{E}_T\,2j$		0.31 $(+0.50\sigma)$
$e\mu\not{E}_T\,3j$		0.71 (-0.55σ)
	W+jets-like	
$W\,2j$		0.29 $(+0.55\sigma)$
$W\,3j$		0.23 $(+0.74\sigma)$
$W\,4j$		0.53 (-0.08σ)
$W\,5j$		0.81 (-0.88σ)
$W\,6j$		0.22 $(+0.77\sigma)$
$e\not{E}_T\,2j$		0.76 (-0.71σ)
$e\not{E}_T\,3j$		0.17 $(+0.95\sigma)$
$e\not{E}_T\,4j$		0.13 $(+1.13\sigma)$
	Z+jets-like	
$Z\,2j$		0.52 (-0.05σ)
$Z\,3j$		0.71 (-0.55σ)
$Z\,4j$		0.83 (-0.95σ)
$ee\,2j$		0.72 (-0.58σ)
$ee\,3j$		0.61 (-0.28σ)
$ee\,4j$		→ 0.04 $(+1.75\sigma)$
$ee\not{E}_T\,2j$		0.68 (-0.47σ)
$ee\not{E}_T\,3j$		0.36 $(+0.36\sigma)$
$ee\not{E}_T\,4j$		→ 0.06 $(+1.55\sigma)$
$\mu\mu\,2j$		0.08 $(+1.41\sigma)$
	$(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$	
eee		0.89 (-1.23σ)
$Z\gamma$		0.84 (-0.99σ)
$Z\gamma j$		0.63 (-0.33σ)
$ee\gamma$		0.88 (-1.17σ)
$ee\gamma\not{E}_T$		0.23 $(+0.74\sigma)$
$e\gamma\gamma$		0.66 (-0.41σ)
$e\gamma\gamma j$		0.21 $(+0.81\sigma)$
$e\gamma\gamma\,2j$		0.30 $(+0.52\sigma)$
$W\gamma\gamma$		0.18 $(+0.92\sigma)$
$\gamma\gamma\gamma$		0.41 $(+0.23\sigma)$
$\tilde{\mathcal{P}}$		→ 0.89 (-1.23σ)

Results

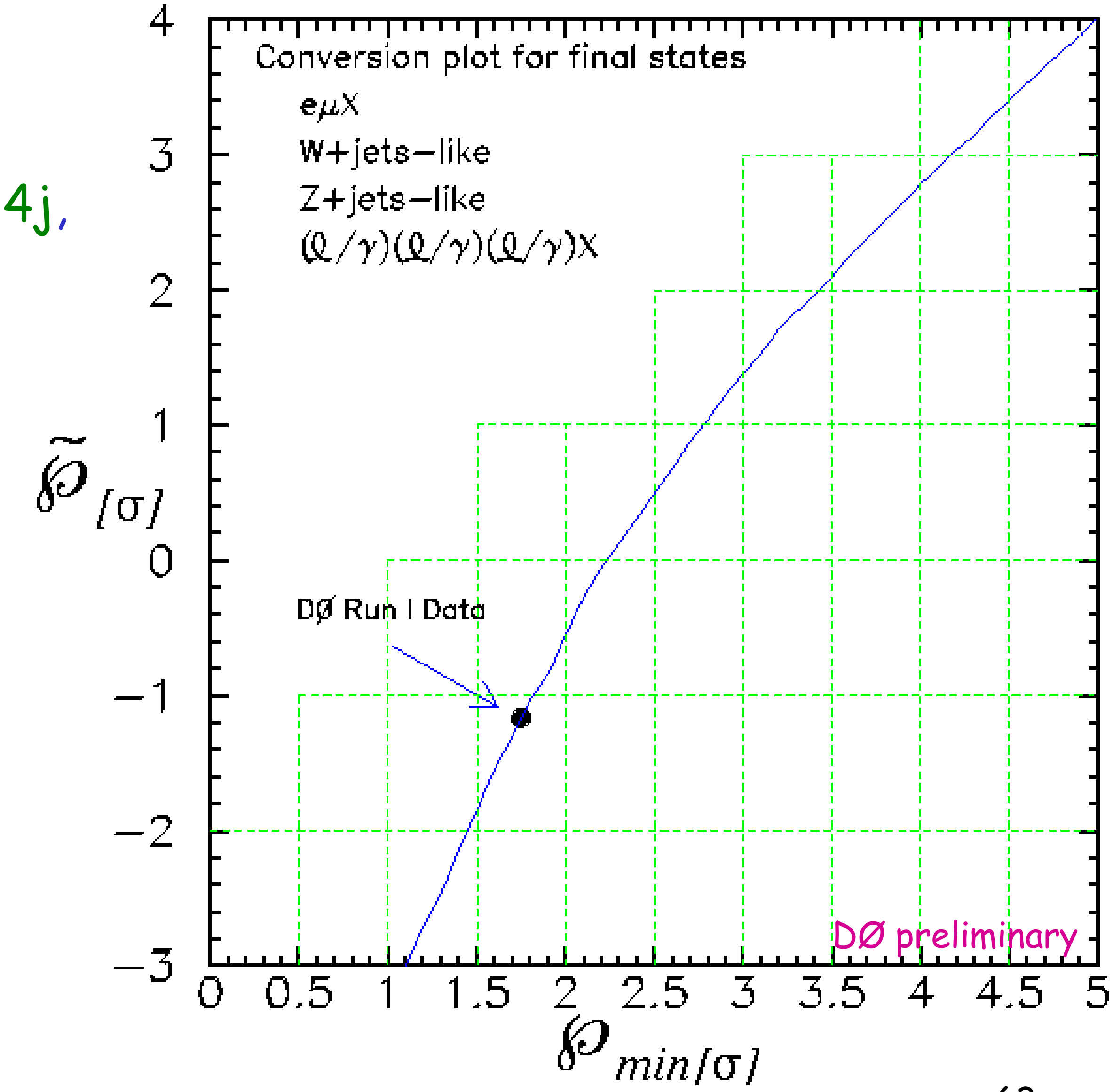
We find

$\mathcal{P}_{min} = 0.04$ ($+1.7\sigma$)

from the final state ee 4j,

corresponding to

$\tilde{\mathcal{P}} = 0.89$ (-1.2σ)



Conclusions

- **Sleuth** is a quasi-model-independent search strategy for new high p_T physics
  - Defines final states and variables
  - Systematically searches for and quantifies regions of excess
- Allows an *a posteriori* analysis of interesting events
- **Sleuth** appears sensitive to new physics
- But finds no evidence of new physics in DØ data
- Should be a useful data-driven search engine in Run II

hep-ex/0006011