Limits and errors on limits

Rajendran Raja
Fermilab

Workshop on Confidence limits
Fermilab, March 27-28, 2000

So, naturalists observe, a flea
Has smaller fleas that on him prey;
And these have smaller still to bite 'em;
And so proceed ad infinitum.

Jonathan Swift

Summary of talk

- Will motivate the concept of errors on limits using a simple Unit Gaussian.
- Will show examples of application using limits-setting programs used in calculating top quark mass limits in published D0 data.
- Will demonstrate the utility of the concept while comparing seemingly contradictory limits.

Error on the mass limit (Caveat:- Not found in Statistical tables)

Being derived from 1st principles

E 417 further analysis

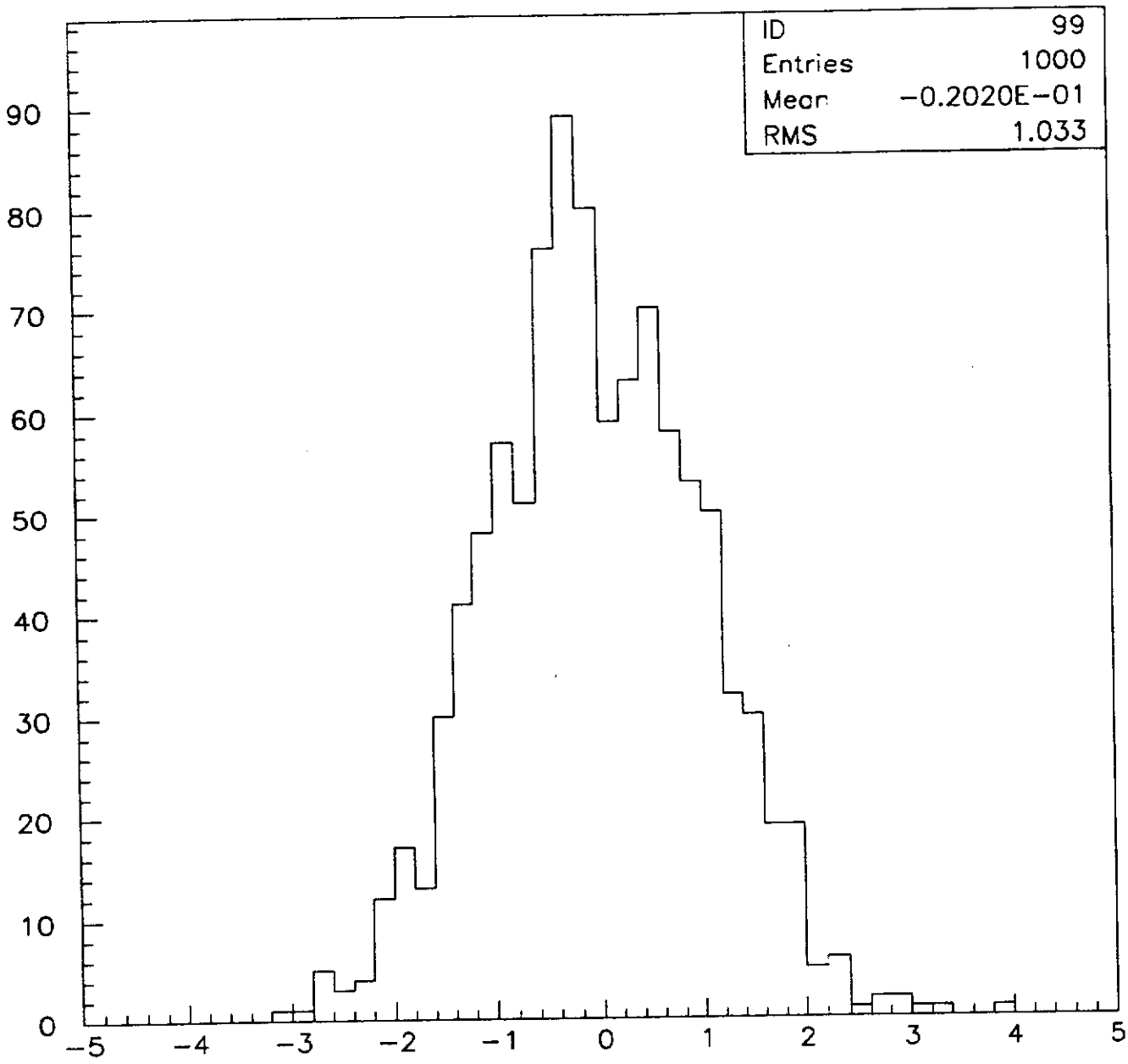
R. Raja
21 Jan 94

- Introduce concept of Error on Errors & limits using Gaussians
- Algorithm to work out Errors on top limiting Cross section using the Taylor Expansion
- Apply it to DØ
- ~~Show that Mass Analysis 1 event max likelihood limits also have errors.~~
- ~~$t\bar{t}$ resonance analysis on E 417~~

UNIT GAUSSIAN

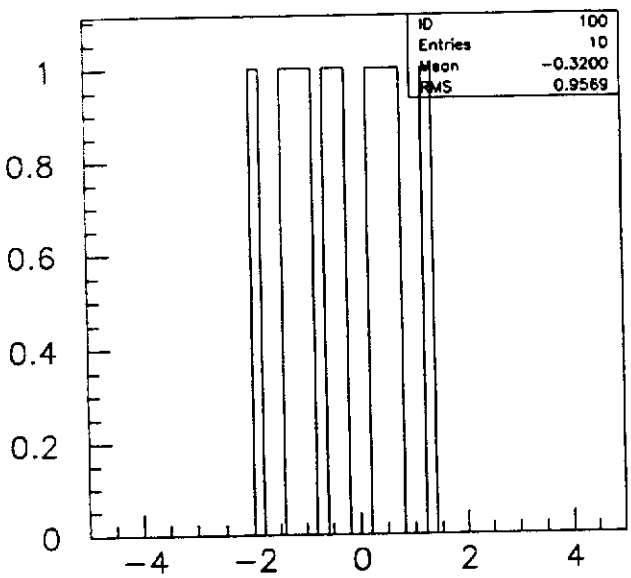
1000 events

$\mu = 0$; $\sigma = 1$

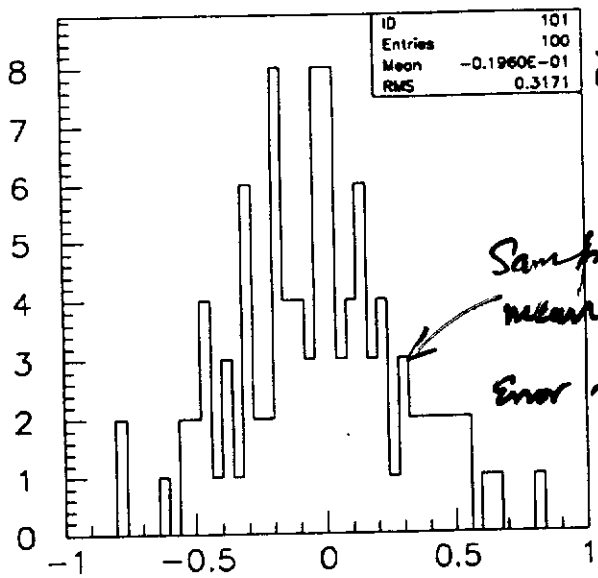


GAUSSIAN DISTRIBUTION

Split into Samples of 10 events
Calculate mean, σ, Error in mean + 97.5% Upperlimit on 10 event Sample.

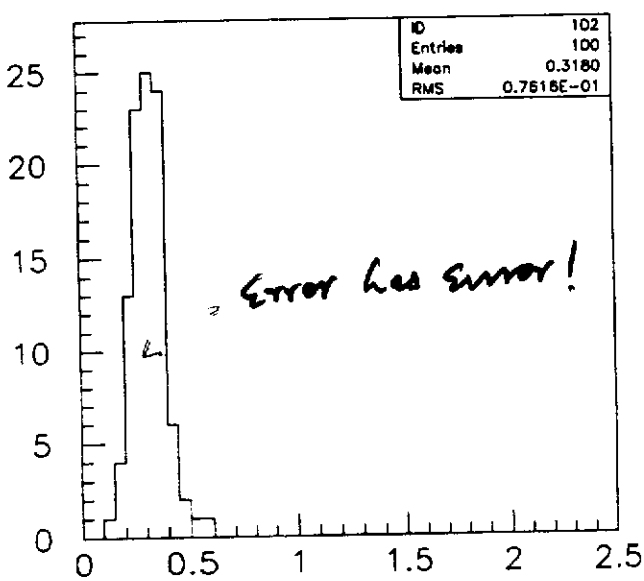


GAUSSIAN SAMPLE

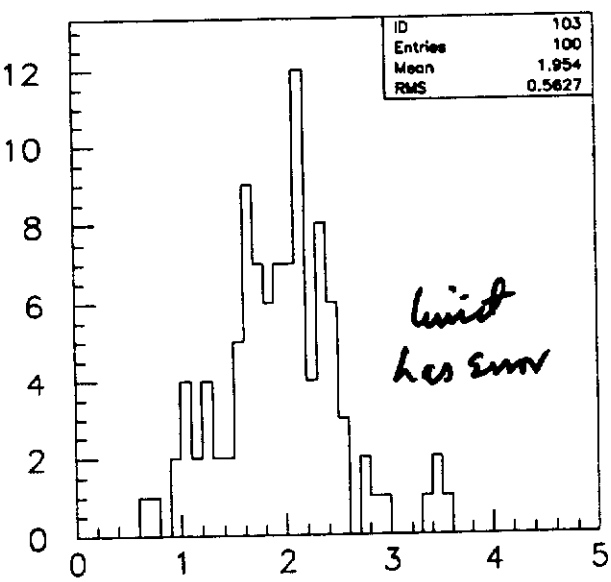


$\frac{1}{\sqrt{10}} \simeq 0.316$

MEAN OF SAMPLE



ERROR IN THE MEAN



95% CL UPPER LIMIT

$$\sigma^2 = E\left[\frac{n}{n-1}\left\{\langle x^2\rangle - \langle x\rangle^2\right\}\right]$$

$$\mu + 1.94\,\sigma = 97.5\%\ \text{UL}$$

Top Quark Search with the D0 1992-1993 data sample- Phys.Rev.D52:4877-4919,1995

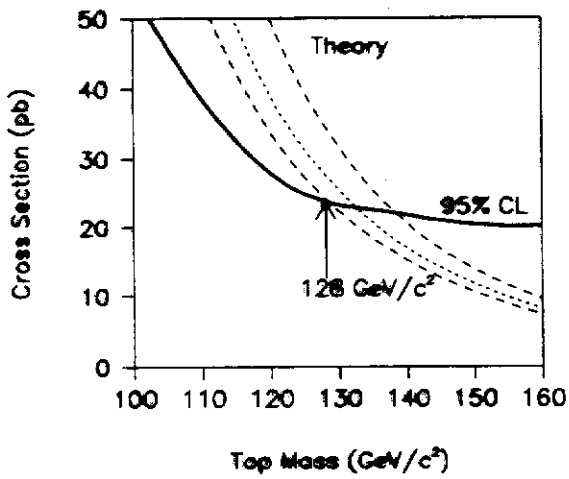


FIG. 44. 95% confidence level on $\sigma_{t\bar{t}}$ as function of top mass from low mass analysis. Also shown are central (dotted line), high and low (dashed lines) theoretical cross section curves [21].

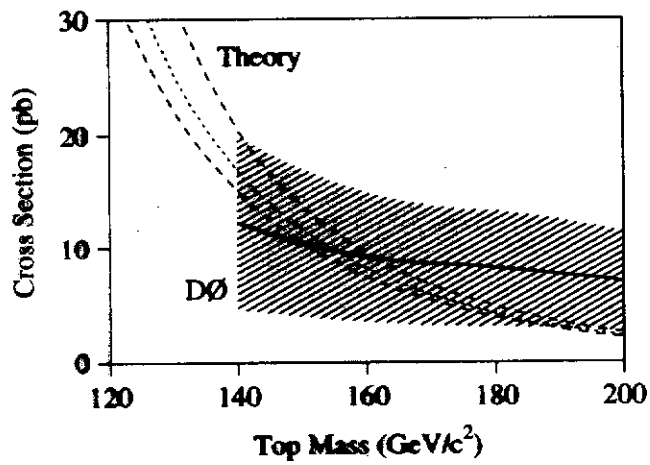


Fig 46: Measured cross section with 1σ errors on it

13.5 pb^{-1}

Should use same algorithm to calculate 45% CL & 50% CL lines.

Should work out errors for each!

DØ top-quark production cross section- PRL 79:1203-1208, 1997

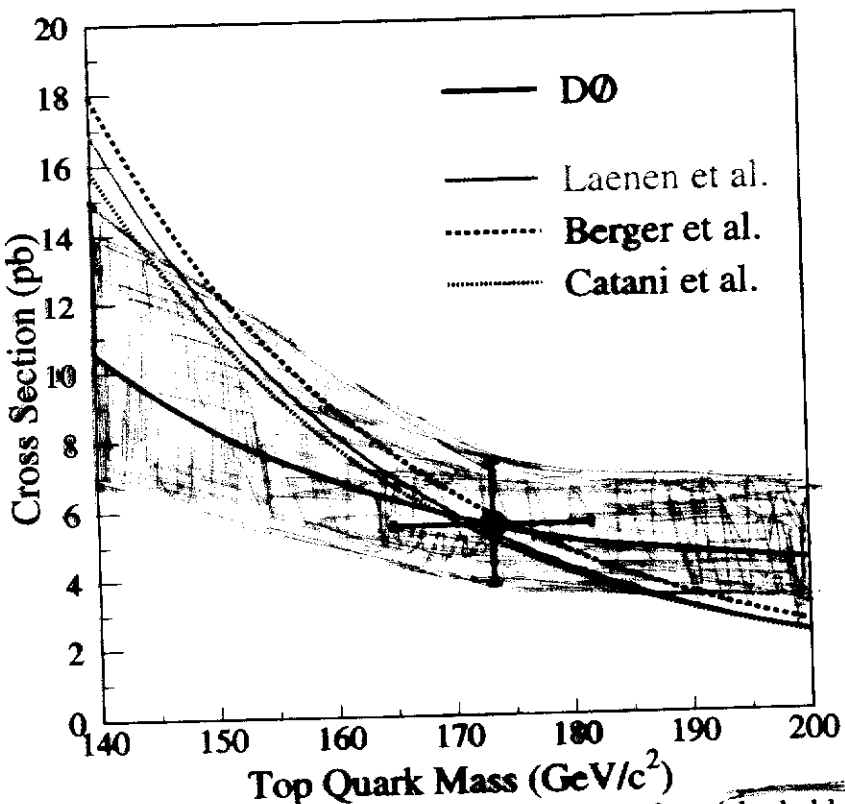
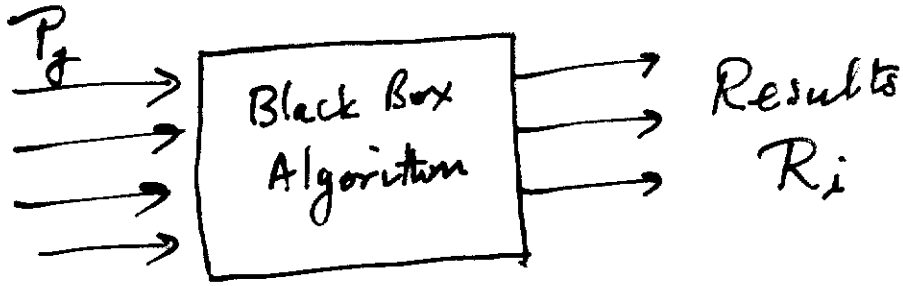


FIG. 3. Measured $t\bar{t}$ production cross section as a function of m_t (shaded band). The point with error bars is the cross section for the measured top quark mass at DØ. Three different theoretical estimates are also shown.

Tyranny of the Taylor Expansion



Parameters

The input parameters $\{P_j\}$ determine the output of results $\{R_i\}$. P_j can fluctuate from calculation to calculation. How does R_i change?

$$\delta R_i = \frac{\partial R_i}{\partial P_\alpha} \delta P_\alpha$$

Sum over α Taylor expansion

$$\delta R_j = \frac{\partial R_j}{\partial P_\beta} \delta P_\beta$$

$$\therefore \langle \delta R_i \delta R_j \rangle = \frac{\partial R_i}{\partial P_\alpha} \frac{\partial R_j}{\partial P_\beta} \langle \delta P_\alpha \delta P_\beta \rangle$$

Average over calculations

$\langle \delta P_\alpha \delta P_\beta \rangle \equiv$ Error matrix of the parameters $\equiv E_{PP}$

$\langle \delta R_i \delta R_j \rangle \equiv$ Error matrix of results. $\equiv E_{RR}$

$$\frac{\partial R_i}{\partial P_\ell} \equiv T = \text{transformation matrix}$$

$$\boxed{E_{RR} = \tilde{T}\, E_{pp}\, T}$$ in matrix notation

i.e Given the error matrix of the input parameters and the transformation matrix T, one can compute the error matrix of the results. No escape!

Our black box is the Top limit program at a particular mass m_t. The output from the box is the 95% CL upperbound on the cross section. $\boxed{R_i \equiv X}$

Input parameters are as follows:-

Number of channels — fixed

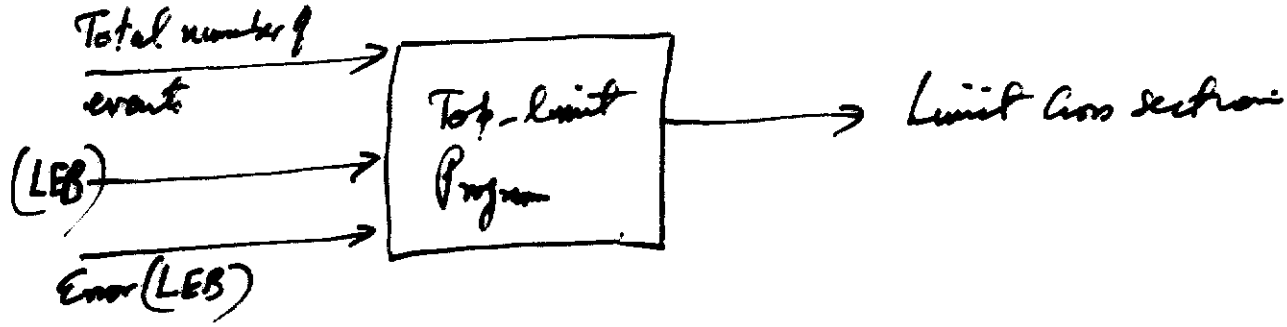
Branching ratios — fixed

Luminosity per channel —
Error in Luminosity /channel —
Efficiency /channel —
Error in efficiency /channel —

} varies from experiment to experiment.

Number of events /channel

Inside the black box of the Top-limit program, there is another black box



$$LEB = \sum_{\text{Channels}} \text{Luminosity} \times \text{Efficiency} \times \text{Branching ratio}$$

Error (LEB) is calculated using these errors.

So actual number of parameters $\equiv 3$

$\therefore E_{pp}$ = Error matrix of Parameters

$$= \begin{pmatrix} \sigma^2_{n_{t\bar{t}}} & 0 & 0 \\ 0 & \sigma^2(LEB) & 0 \\ 0 & 0 & \sigma^2(\text{Error}) \end{pmatrix}$$

Error on Error

$n_{t\bar{t}} = \cancel{7}$

& $\sigma^2_{n_{t\bar{t}}} = \cancel{3}$

$\sigma^2(LEB) = \{\text{Error in } (LEB)\}^2$

Error on (Error in LEB) $\approx 50\%$ ← assumption

Transformation matrix

$$\frac{\partial R}{\partial P_j} \quad \text{determined } \underline{\underline{\text{numerically}}}$$

This enables us to calculate the error matrix (σ^2)
on the limiting x section

$$E^{RR} = \sigma_x^2 = (\cdot \; \cdot \; \cdot) \begin{pmatrix} \cdot & & \\ & \cdot & \\ & & \cdot \end{pmatrix} \begin{pmatrix} - \\ - \\ - \end{pmatrix}$$

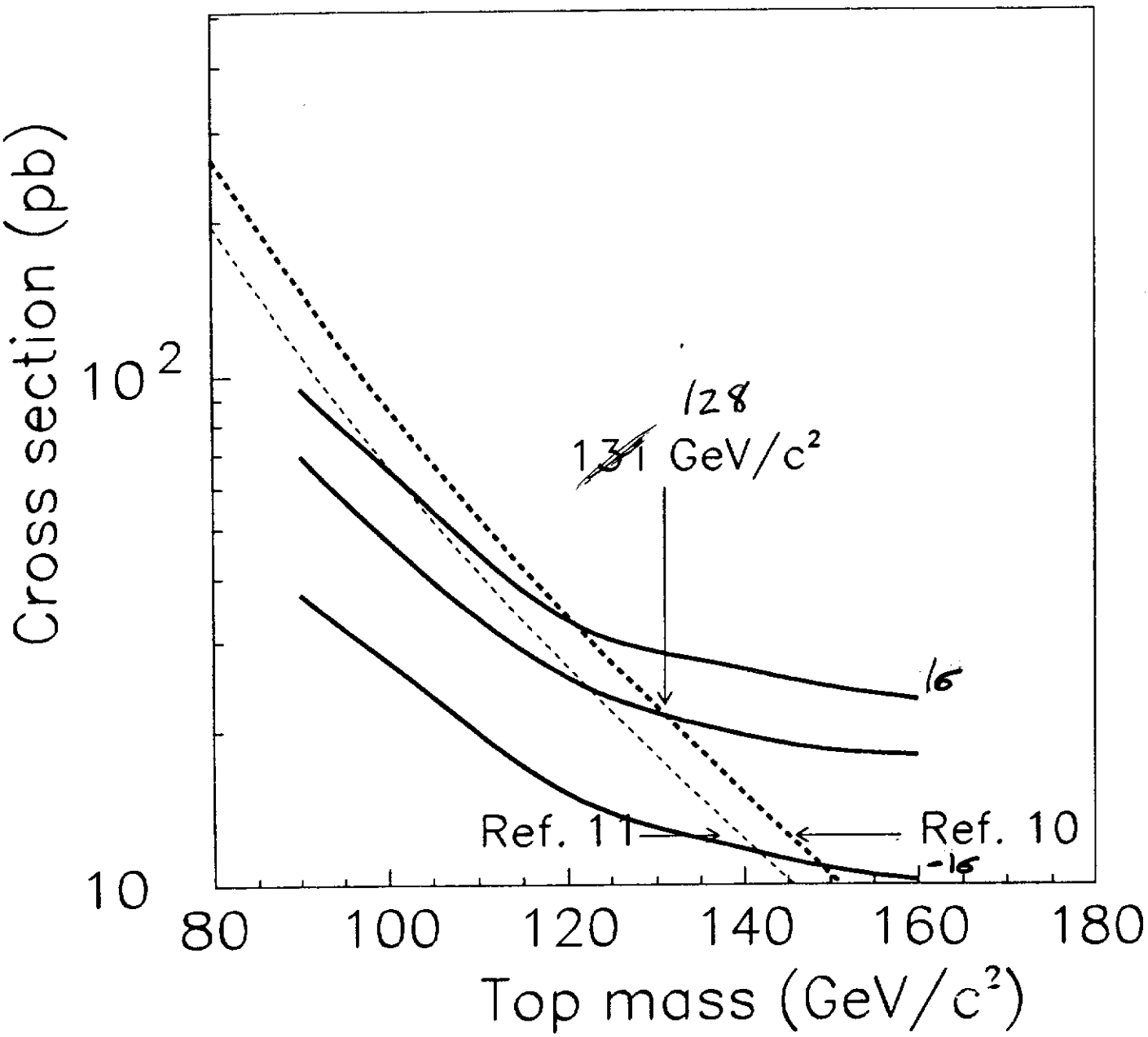
Transformation matrix

$$\frac{\partial R}{\partial \varphi_j} \quad \text{determined } \underline{\underline{\text{numerically}}}$$

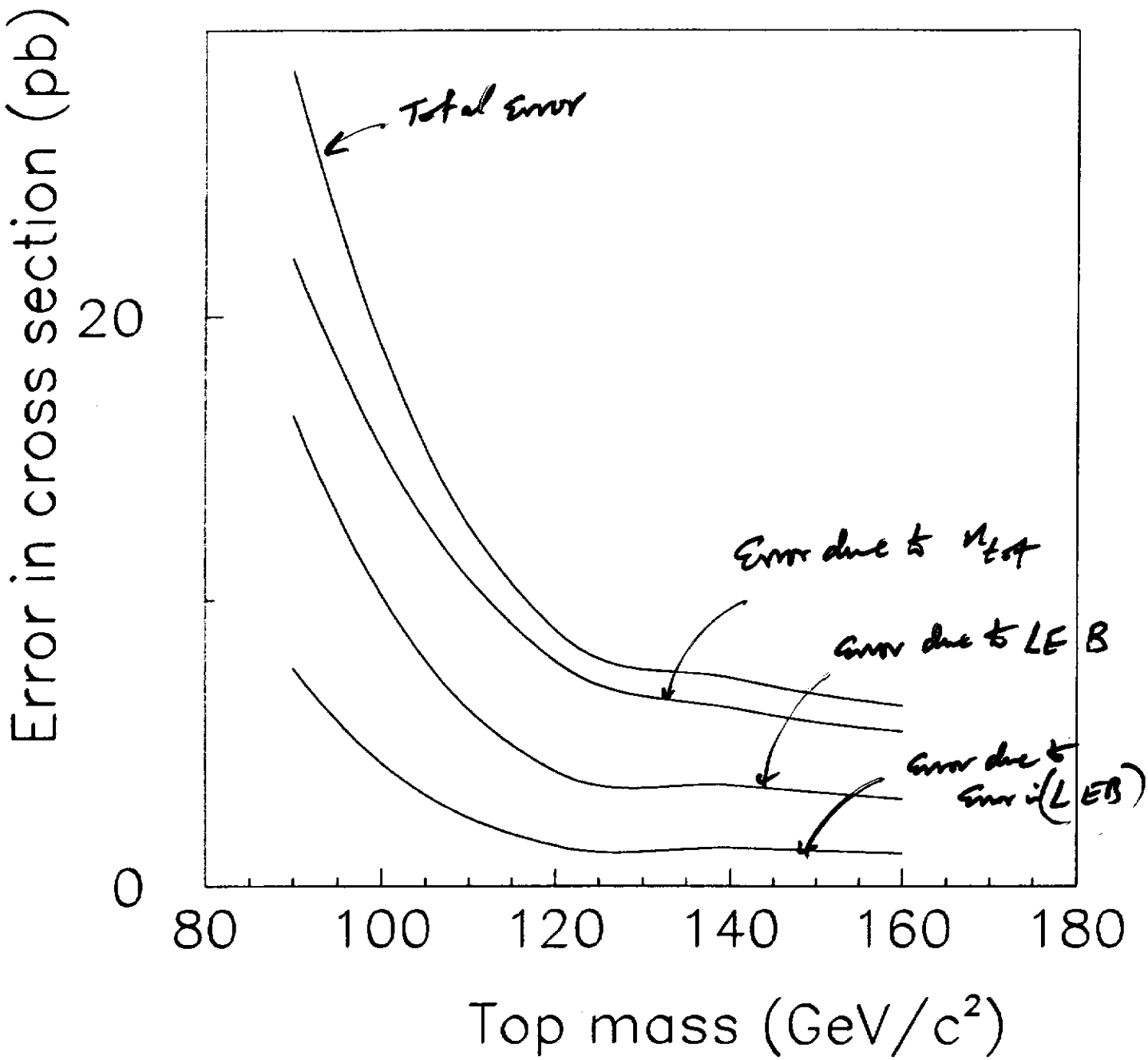
This enables us to calculate the error matrix (σ^2)
on the limiting x section

$$E^{RR} = \sigma_x^2 = \begin{pmatrix} \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} - \\ - \\ - \end{pmatrix}$$

The estimate of the 95% CL upper [illegible] of the [illegible] mass distribution is 128 GeV/c² and the error in this estimate is +17 −11 GeV/c²



$t\bar{t}$ xsection > ~~131~~ 128 GeV/c² $^{+17}_{-11}$
at 95% CL


Error in cross section (pb)
20
0
80
100
120
140
160
180
Top mass (GeV/c²)
Total error
Error due to N_{tot}
Error due to LEB
Error due to Error in (LEB)

Conclusions

- The concept of 'error bands' on limits should be of utility for limits calculated using small statistics.
- It should be seen to be on a footing similar to quoting mean values (50% CL) and their errors.
- It is a mathematically rigorous concept. We have shown how to calculate errors on limits using a completely general black-box limits program.

So, physicists observe, an error
Has smaller errors that on it prey;
And these have smaller still to modulate 'em;
And so proceed ad infinitum.

Apologies to Jonathan Swift