Belief and Observation: The Top Quark and Other Tales of "Discovery"

T. Ferbel
Universities of Rochester and Maryland

"Seek, and Ye Shall Find" (Matthew 7:7)
(If you search long enough, you're bound to find something!)

Plan for the evening:

- Story of the discovery of the top quark (1995)
- Story of the stillbirth of the b* meson (1964)
- Tragedy of the "splitting" of the A_2 meson (≈1970)
- A tale from "Pathological Science" (Langmuir ≈1930)
- A few words on medical research!!!
- Skip many other quirks: cold fusion, gravity waves, $\beta_\nu>1$, N-rays, leptoquarks at DESY, "OopsLeon," etc, etc, from work of top scientists over past 100 years!
- A word on human frailty (but hope)

References: I. Langmuir, A. Wróblewski, S. Stone, others on web.

History of the top quark

1973: Makoto **Kobayashi** and Toshihide **Maskawa** predict the existence of a third generation of quarks in their attempt to accommodate the observed violation of CP invariance in K^0 decays. (Cabibbo/GIM mixing combine in CKM!)

1974: The November revolution with discovery of the J/ψ and the fourth (GIM) “charm” quark at both BNL and SLAC by Sam **Ting** et al and Burt **Richter** et al, respectively, and, shortly thereafter, the τ lepton by Martin **Perl** et al (also at SLAC), with the τ providing major support for a third generation of fermions.

1975: Haim **Harari**, a great wizard (Israeli theorist) of the era, names the quarks of the third generation "top" and "bottom" to match the "up" and "down" quarks of the first generation, reflecting their "spin up" and "spin down" membership in a new weak-isospin doublet that also restores the numerical quark/lepton symmetry of the next (current) version of the standard model.

1977: Leon **Lederman** et al, discover the fifth *b* quark at Fermilab (bound with the *b* antiquark in Υ states of quarkonium), offering thereby new circumstantial evidence for the existence of a sixth ***t*** quark needed to complete this isospin doublet. Most everyone thought it likely that the mass m_t would be larger than that of m_b, but few expected a factor of ≈35 for the ratio of the masses of these supposed isospin-partner quarks. And few, if any, expected that it would take so many years to finally confirm the existence of the top quark at the Fermilab Tevatron antiproton-proton collider in 1995 (simultaneously by two experiments, led at that time, respectively, by Giorgio **Bellettini** and Bill **Carithers** at CDF and by Paul **Grannis** and Hugh **Montgomery** at DØ).

Initial searches for the top quark started quickly, and negative results reported from e^+e^- colliders at SLAC, DESY and KEK. Following the observations of the *W* and Z bosons at the $p\bar{p}$ collider of the Super Proton Synchrotron (SPS) at CERN in the early 1980s , there were claims of a discovery of the top quark in 1984 at the UA1 experiment (of Carlo **Rubbia** et al) in $W \rightarrow bt \rightarrow bb\ell\nu_\ell$ decays (where ℓ= e or μ). They found 12 events in the lepton+2-jets +$\Delta\mathbf{p}_T$ channel (where $\Delta\mathbf{p}_T$ is the vector imbalance in transverse momentum in each event that can be attributed mainly to the $\mathbf{p}_T$ carried away by the escaping neutrino ν_ℓ). The result was consistent with a top quark mass of m_t=40±1 GeV that originated from $W^+ \rightarrow \bar{b}\, t \rightarrow \bar{b}\, b \ell^+ \nu_\ell$ decays (as well as from the W^- c.c. state). Because 3.5 events were expected from background sources (mainly *W*-boson production with accompanying radiated gluons), at face value, this was a very significant effect corresponding to well over three standard deviations (SD). (Other reports at the time cited 5 events observed with 0.2 expected, which has comparably huge significance.) Although this was very interesting, it was never confirmed by the friendly competition at UA2 of Pierre **Darriulat** et al.

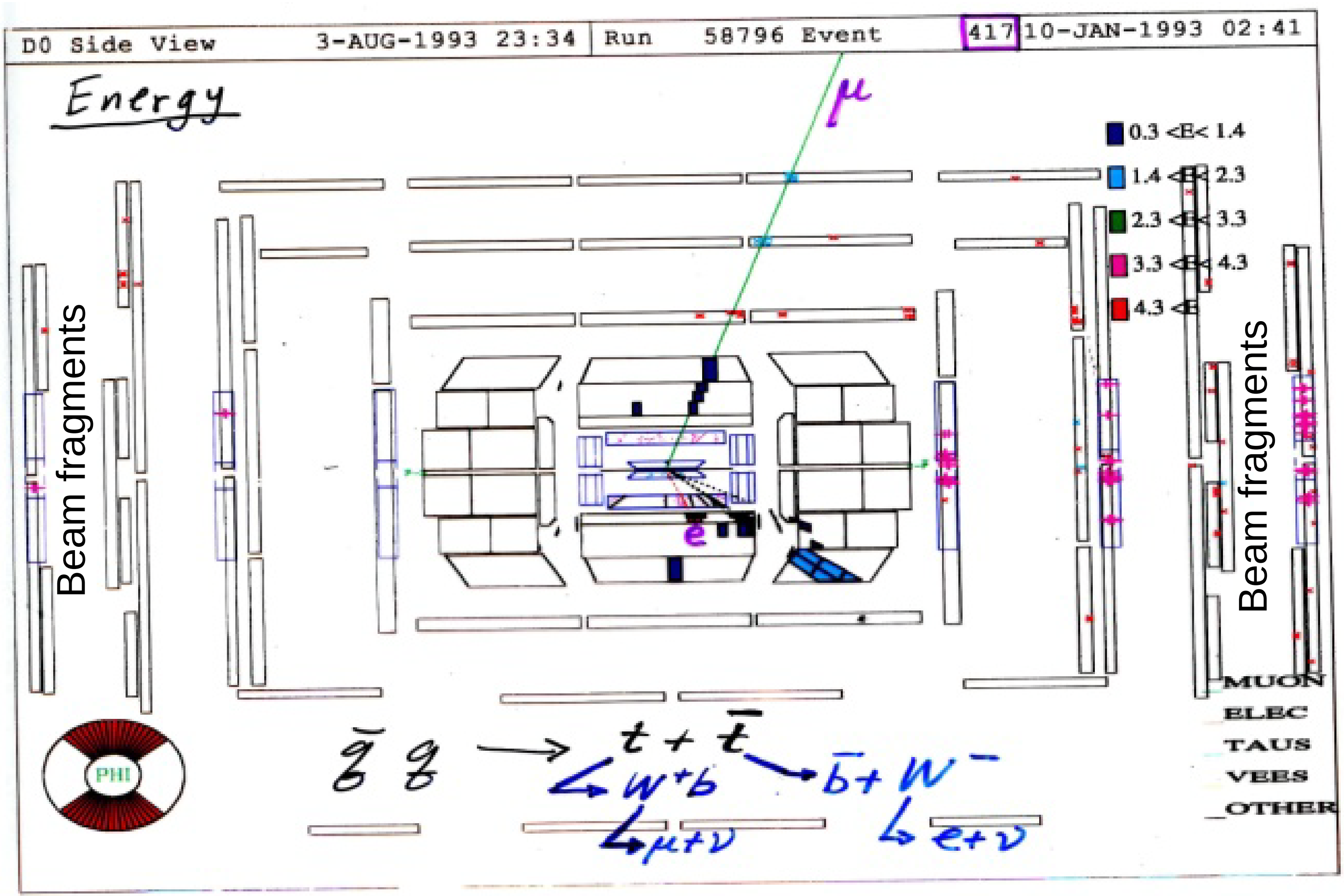
An aside: Enrico **Fermi** was proud of his experimental prowess, and advised his colleagues never to publish any new effect unless it had a significance of more than **3 SD**. Of course, that advice was given in different times, before the explosion in the number of practicing research scientists, and the huge increase in searches for new phenomena that followed the "successes" of science in World War II. As most of us recognize, a competing experiment can impact the significance of any claim of discovery through psychological means as well as through the statistical evaluation of a result (i.e., how to take account of all the other jokers performing the same searches). These are serious concerns, and it used to be rumored that, in trying to maintain his sterling reputation, Fermi multiplied all his uncertainties by a √2. But surely, this “conservatism,” or what some might term foolishness, is not prerequisite for success, and I can’t believe that Fermi would have done that. He might have, however, thought harder about the meaning of the uncertainties!

Now, in fact, the effect at UA1 appears **not** to have been the result of some statistical fluctuation, but rather due to poor modeling of background (from $W \rightarrow \tau\nu_\tau$), as Terry **Wyatt** found, and eventually convinced his powerful colleagues by challenging them with incisive questions about the analysis.

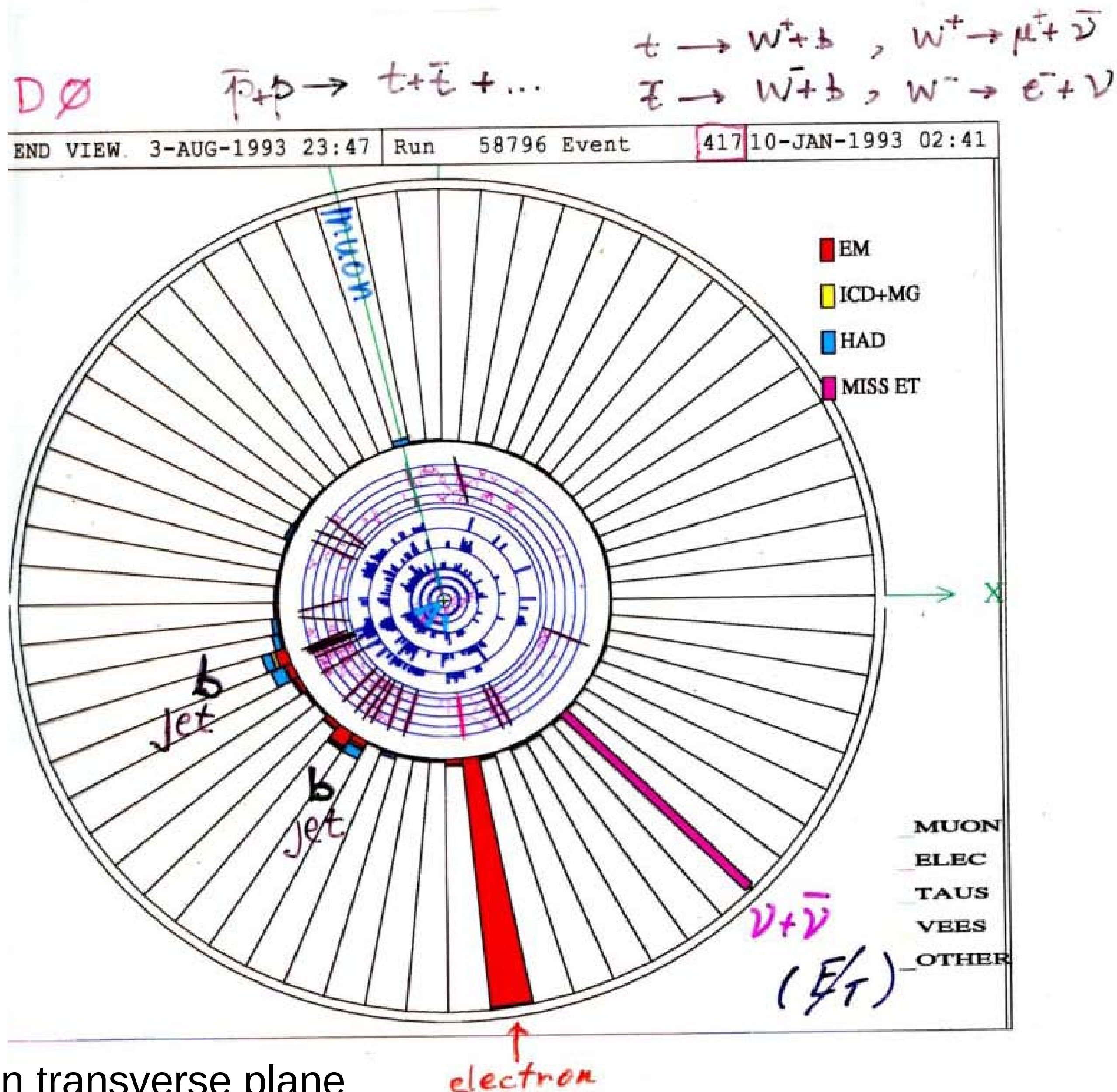
This is when the Fermilab Tevatron entered the scene, and for more than two decades before the start of LHC operations at CERN in 2010, it was the only place in the world with enough energy to produce top quarks. To secure a productive future, Leon **Lederman**, then director of Fermilab, proposed in 1981 to add another general-purpose detector, one that would complement the already approved Collider Detector at Fermilab (CDF) . Leon thought he knew exactly what that other detector should look like (but never told anyone his secret!). Not having been pleased with any of the proposals, he asked Paul **Grannis** to form a fresh collaboration to design a new detector, which is how the DØ experiment came into existence, and by 1991 it joined the more prestigious, and well-operating CDF experiment to find the golden needles expected to be buried in the hay-stack of Tevatron events.

By early of 1993, the two groups saw their first likely top quarks, each with an event that appeared to have the properties for production of $t\bar{t} \rightarrow W^+ b W^- \bar{b} \rightarrow e\nu_e b \mu \nu_\mu b$, namely, an electron, a muon, two jets, and significant $\Delta\mathbf{p}_T$ in each event. More data was collected the following year, and, on April 22, 1994, CDF submitted to PRL "Evidence for" a top quark with m_t of ≈175 GeV/c^2. But DØ did not have any substantive support beyond their suggestive Event #417.

(The famous **DØ Event #417**)



Event was a factor of >10 more likely to be $t\bar{t}$ than next choice (W^+W^-)



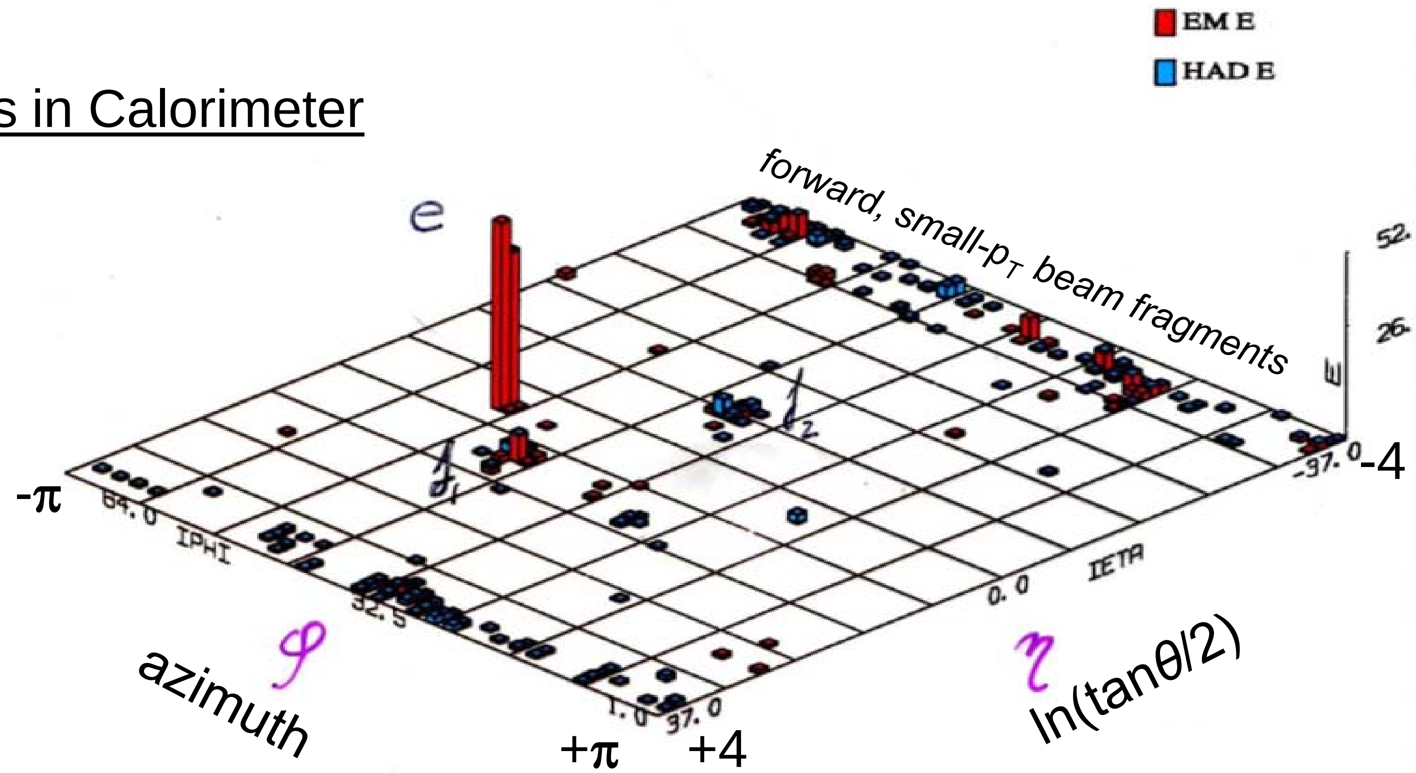
Energies in transverse plane

$t\bar{t}$ (or WW far less likely)

LEGO CAL CAEP	3-AUG-1993 23:46	Run 58796 Event	417	10-JAN-1993 02:41

CALEGO EMIN = 0.5 GeV
CAEP E SUM = 518.4 GeV

<u>Energies in Calorimeter</u>

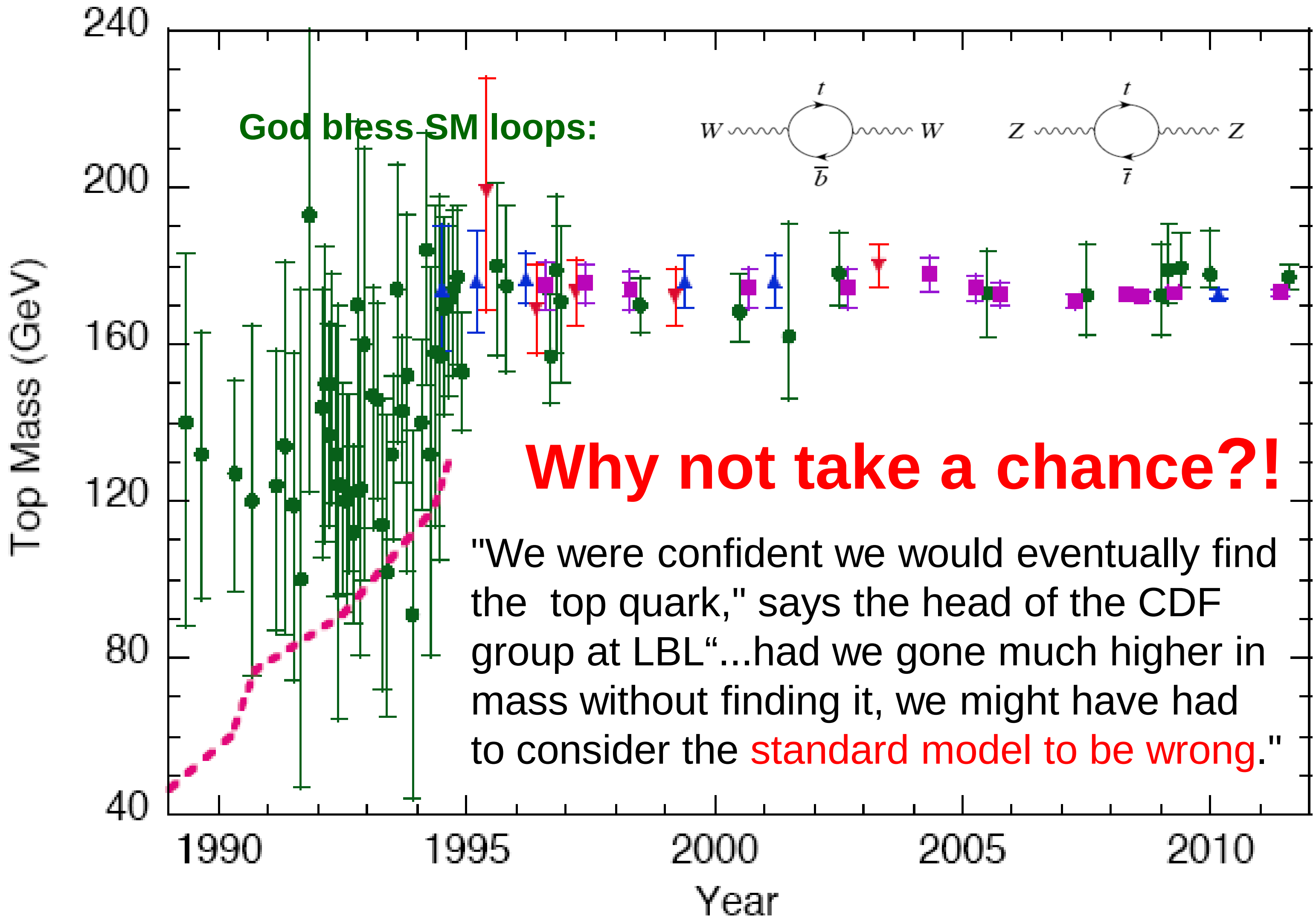


Where is the μ?

What was CDF doing right that DØ was not? The integrated luminosities were similar, the acceptance for top quarks not very different, and yet CDF claimed they expected 5.7 background events, and observed 12: a significance of ≈3 SD! Although not inconsistent with CDF, DØ could not confirm the size of the CDF yield! Clearly, DØ was disappointed. (To be fair, at that time, DØ did not have a silicon detector, which gave greater credence to the early results from CDF, and, despite their famous Event 417, DØ was more inclined to find as large a lower limit on m_t as possible, rather than focus on the harder task of proving that the top quark was truly in the data, as well as on developing multivariate analyses.)

Paul **Tipton**, our former student, who was cool, smart and quite competitive, rejoined Rochester about that time as Asst Prof. He had helped develop CDF's silicon vertex detector at LBL, and was now eager to discover the top quark. I mentioned to him that I overheard his Michigan colleague Dan **Amidei** say that **he** would find the top quark at CDF, to which Paul responded, jokingly, saying, something like "Oh yeh? Let's see who will be first. **I'll find top, even if it's not there!**" Certainly, not an unexpected remark from a self-confident, ambitious young scientist with great moxie! But, as the story developed, it became clear that the CDF cross section for $t\bar{t}$ production was a factor of ≈2 high: as close to inconsistence with the prediction of the standard model (SM) as the significance of the observed effect! But CDF must have thought that it was worth taking a chance on the top's likely existence? (In fact, most of us felt that it was just a matter of time before "discovery!") I recall a "mealy-mouthed" public remark from CDF during all of this: "It is possible that we are seeing a rare statistical fluctuation, but we have a good indication that the top quark **may** be there."

There is risk in taking a chance, but, when we feel that the odds for success are high, many of us "will go for it." Some of our major scientists (not to be named in public) have relied at times on their intuition as much as on their data, and, despite having previously made wrong guesses, have managed to survive with their credibility in tact. We are a forgiving and enthusiastic lot, willing to ignore sins, and to focus rather on successes of our colleagues: we remember that they did something important, but not necessarily if it turned out to be right.



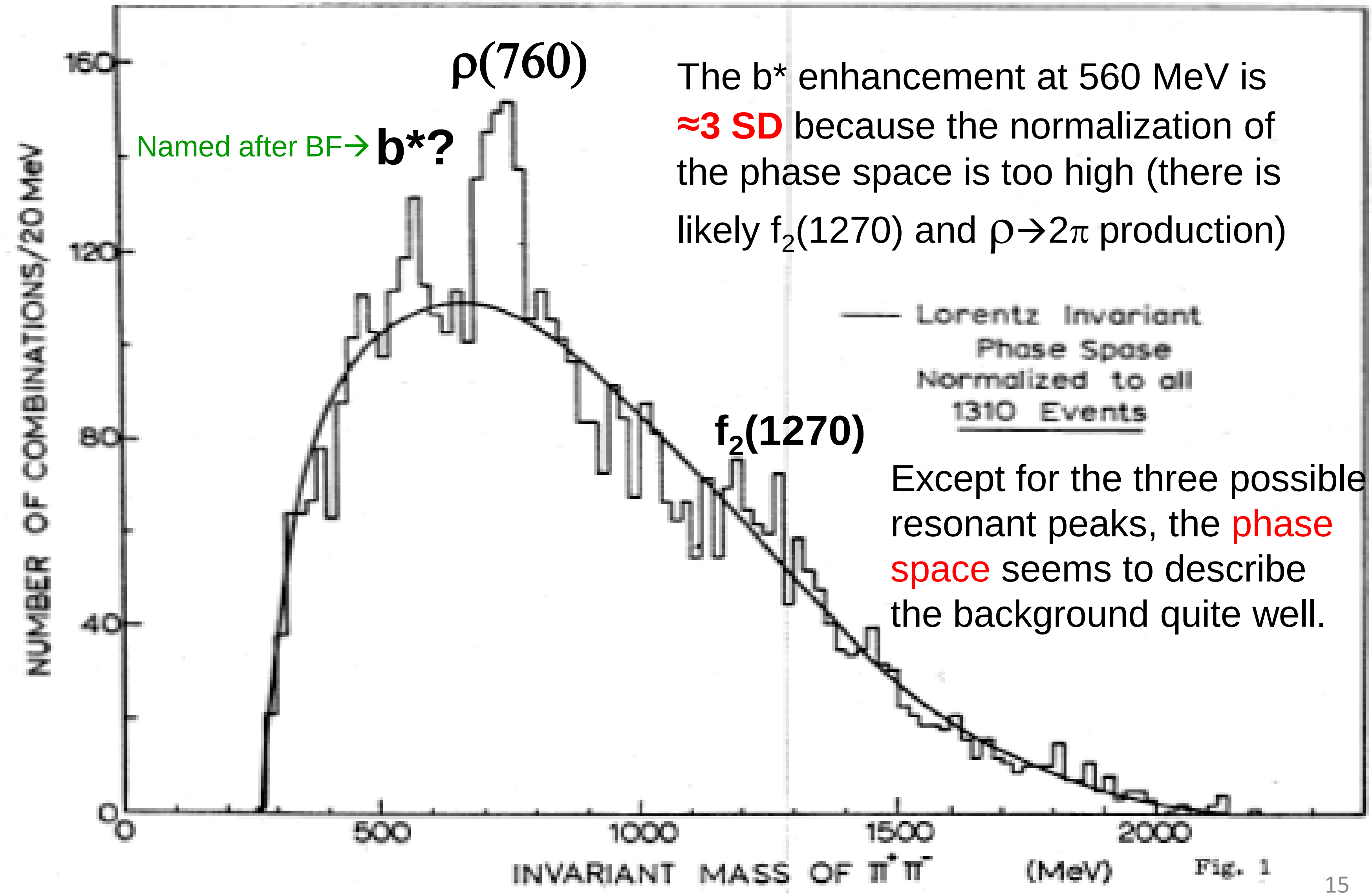
Indirect estimates of mass of top quark from fits to **electroweak** observables (**green**), and **lower bounds** at 95% CL on m_t inferred from direct searches in $\bar{p}p$ collisions at CERN and at the Tevatron (**broken line**), assuming standard $W \to bt$ or $t \to bW$ decays. Results on m_t from CDF (up-triangle) and DØ (down-triangle), and **mean m_t** (**purple square**). Now, **m_t=173±1GeV!** (Adapted from Quigg)

So, on the first round, CDF seemingly outfoxed DØ fair and square. A year later, after an increase in luminosity, both published their definitive results for $t\bar{t}$ production in lepton+jets events ($bWbW \rightarrow bqq'b\ell\nu_\ell$, with ℓ=e or μ), and in dilepton events (such as DØ #417), and each experiment presented an almost ≈5 SD effect! But beware, because CDF and DØ were often too eager:

- The unconfirmed SUSY-like event in $ee\gamma\gamma+\Delta\mathbf{p}_T$ at CDF (≈1996)
- The observation of possible quark substructure from jet production at CDF (≈1996), but not confirmed by DØ, with the excess at large p_T eventually attributed to uncertainties in parton (gluon) distribution functions.
- The $\Omega_b^- \rightarrow J/\psi\ \Omega^-$ observed originally (≈2008) at DØ, but never confirmed in later data (and found at different mass and lower cross section by CDF!)

There were many more humbling inconsistencies. But what is so surprising, especially in such large collaborations, is how results that seem questionable can pass scrutiny! We believe in our own data, and it is difficult to sway colleagues who have become convinced of having made a discovery to wait and view the matter from a different perspective before publishing. Which brings us next to some archaeology of my postdoc years in the mid 1960s.

Search for a meson (b*) produced in $\bar{p} + p \rightarrow \pi^+ + \pi^+ + \pi^- + \pi^- + \pi^0$ fixed-target collisions at 3 – 4 GeV/c (Yale-BNL-CCNY Collaboration)



ANGULAR DISTRIBUTION OF CHARGED π-MESONS
IN THE CENTER OF MASS SYSTEM
1310 EVENTS

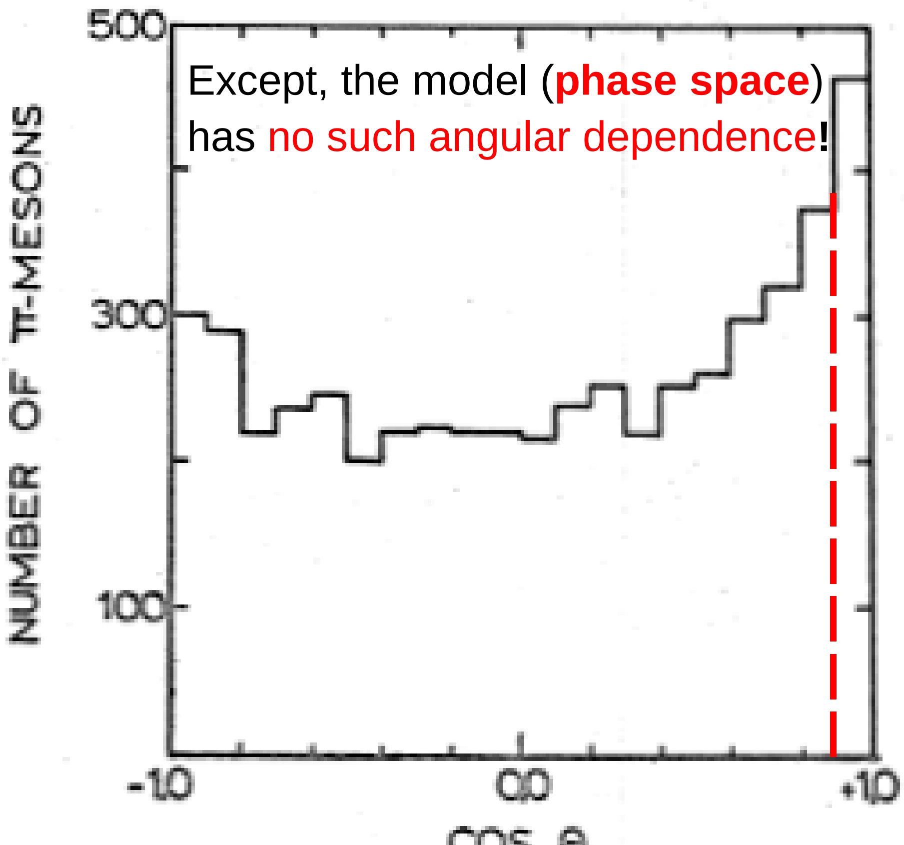


Fig. 2

(cosθ of π^- wrt to $\bar{p}$, and π^+ wrt p)

Yes, this is analogous to the recently misnamed "**charge asymmetry**!" The b* was not seen in e.g., $\pi^- p \rightarrow \pi^+\pi^- n$ reactions, which could be explained if the pions produced at forward angles reflected a different mechanism than in central production. Looking at such forward events (with at least one pion at $\mathbf{cos\theta > 0.9}$) to see if the b* peak is favored in such events, seemed like a sensible idea, and worth checking for any major difference in the b* yield for the two angular production regions.

Now, same data, but with $\cos\theta > 0.9$. **Voilà**! Discovery?

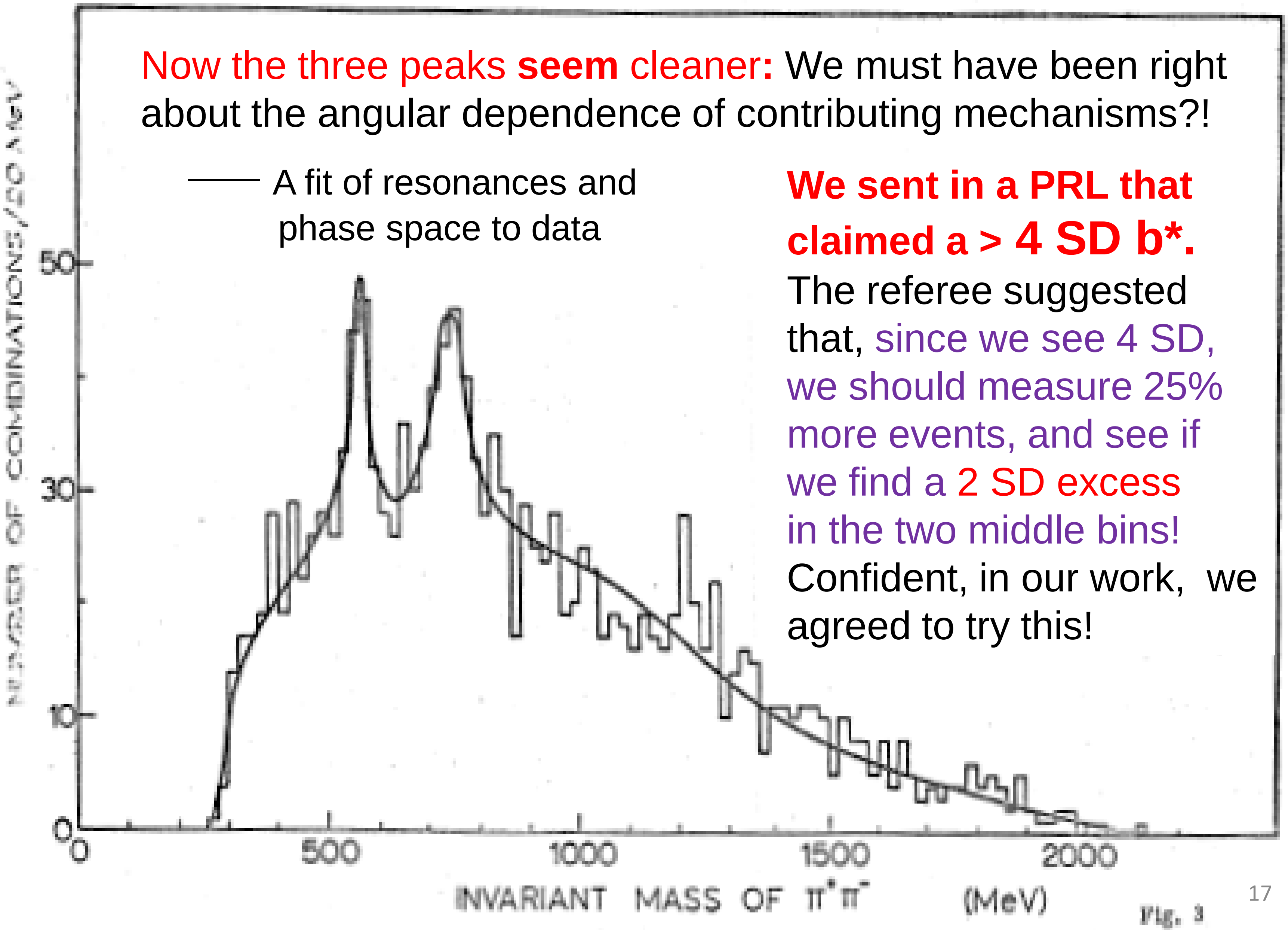
Now the three peaks **seem** cleaner**:** We must have been right about the angular dependence of contributing mechanisms?!

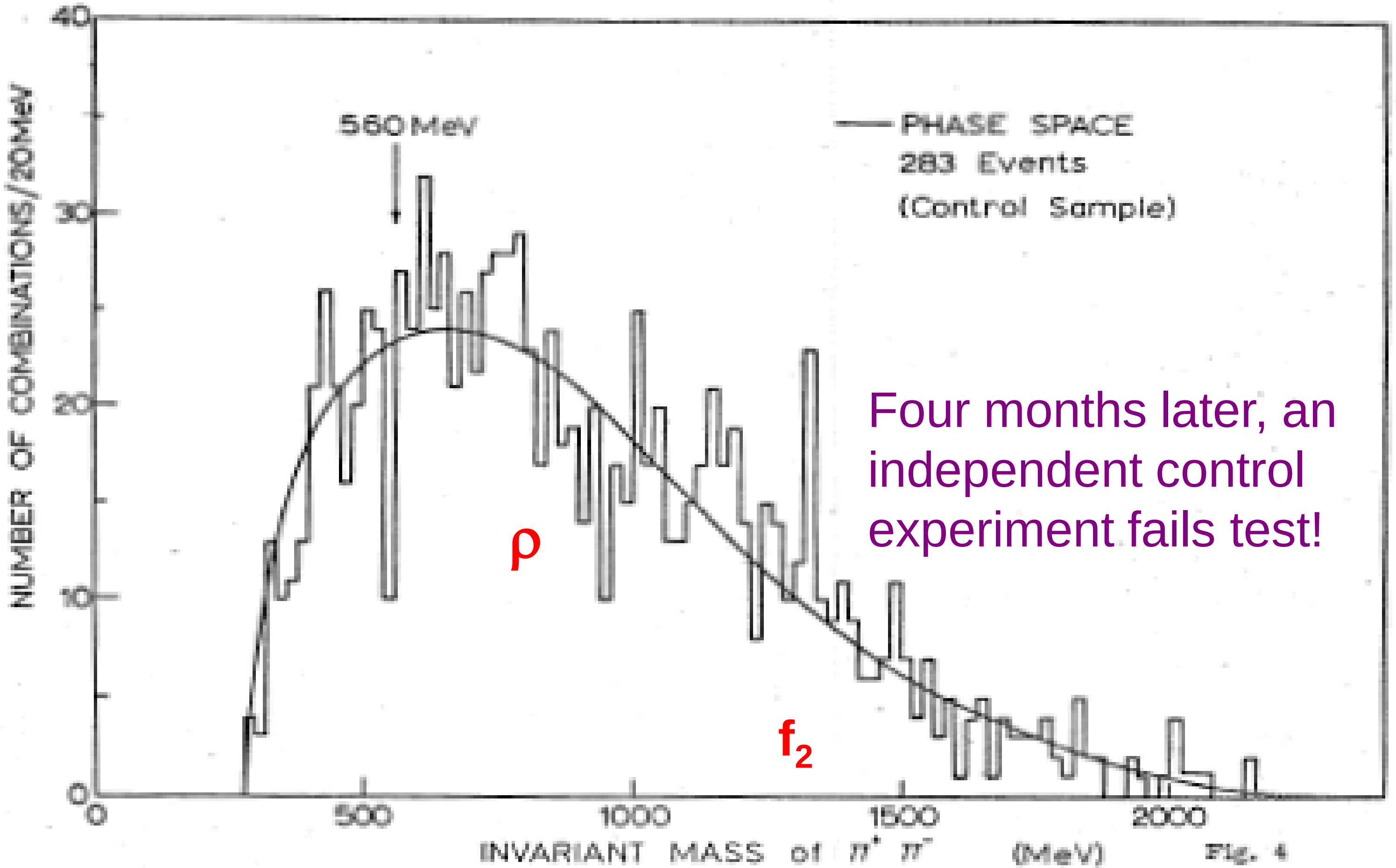
—— A fit of resonances and phase space to data

We sent in a PRL that claimed a > 4 SD b*.

The referee suggested that, since we see 4 SD, we should measure 25% more events, and see if we find a 2 SD excess in the two middle bins!

Confident, in our work, we agreed to try this!





Oh horrors! Found a 2 SD deficit! This meant that our beautiful (incalculable) effect was likely an unlikely statistical fluctuation! We were all very disappointed (especially Barbara), but we felt relieved , somewhat like martyrs, turning defeat into a victorious lesson in statistics! We didn't publish, but not everyone gives up that easily!

Splitting the A_2 Meson (late 1960s)

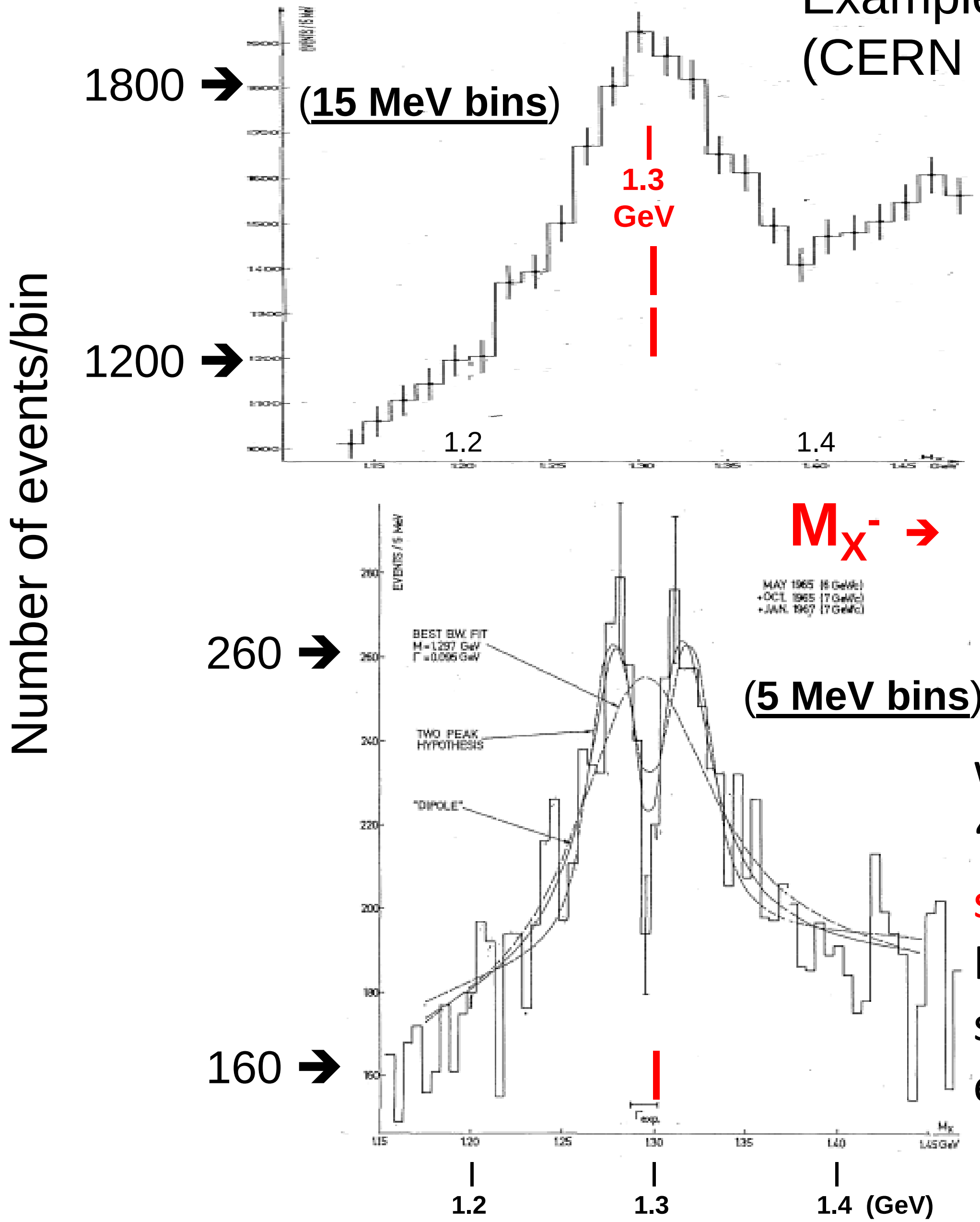
Two CERN experiments reported a **precision study** of then recently observed $J^{PG}=2^{+-}$ a_2 (1320) meson, finding that the **mass spectrum did not follow a Breit-Wigner** form but rather a **double-pole-like structure**. This meant that a particle with a width of ≈100 MeV, did **not undergo exponential decay with time!**)
These remarkable results were found using "missing-mass" spectrometers, with the mass measured from the recoiling proton in inclusive $\pi^- p \to p + X^-$:

$$M_{X^-}{}^2 = (E_{\pi^-} + M_p - E_p^{rec})^2 - (\mathbf{p}_\pi - \mathbf{p}^{rec})^2$$

Unfortunately, these experiments were not managed properly, and a series of misjudgments by senior scientists led to results that eventually came to be ignored by the community. However, in the prime of that "A_2-splitting mania" of ≈1968, the **intoxication of the discovery silenced all doubters**: At the 1968 conference on Meson Spectroscopy, Don Miller from Berkeley was supposed to have responded for the "opposition," however, the reception and accolades of the audience heaped on the proponents discouraged Miller from presenting his results. This was a pity, as it compounded the flight from reason. Several

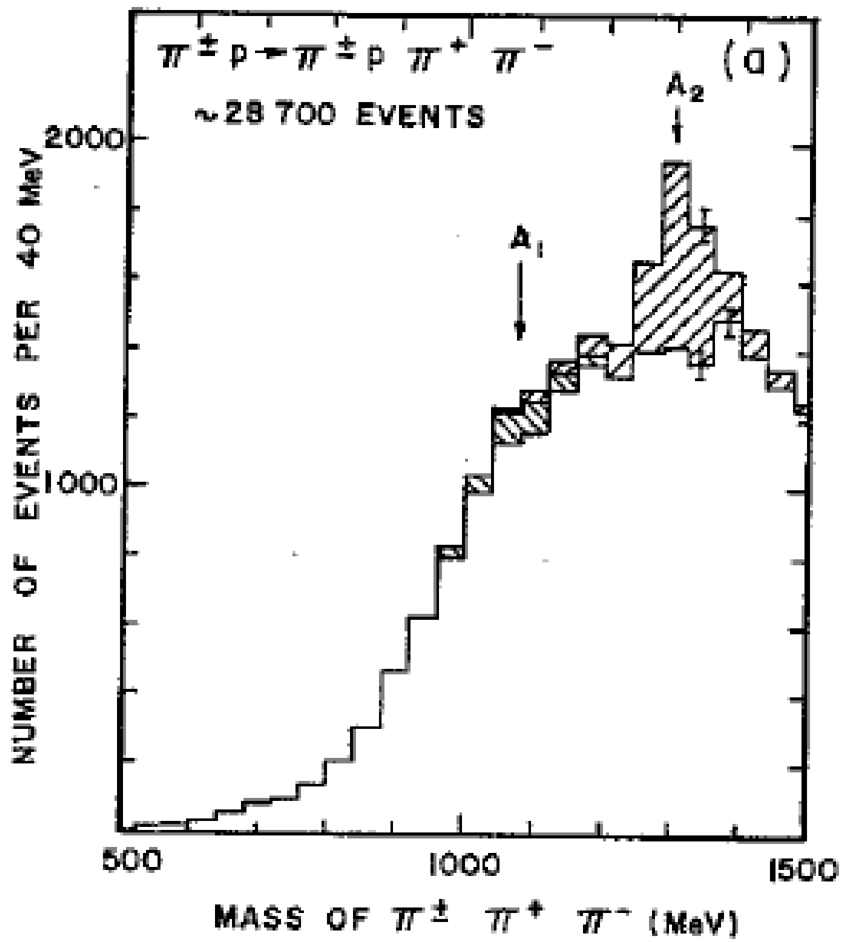
other groups (bubble chamber data analyses) provided **rapid support** for the splitting of the A_2 peak, and a few eminent theorists tried to make sense of this new phenomenon. But it soon became clear that some procedures of the **analysis at CERN** were **flawed**. It seems to me, that most egregious was the discard of data for runs that did not show a split peak. There was surely **no intent to defraud**, but, after checking a few configurations of the spectrometer, and becoming perhaps too quickly convinced that there was a true effect, the group **decided to use the presence of a central dip to select** events, so that, when there was no splitting, the experimenters checked to see whether there was any problem that could affect the resolution of that specific run. And, naturally, found some reason to remove the data of supposed "poorer"quality! It was felt that to see the tiny **mass splitting required data of best precision**! Also, to assure lack of bias(!), the accepted data runs also had their mass spectra shifted to the mean value of the A_2 peak to preserve the "correct" calibration! Experiments at Brookhaven Lab (led by Northeastern University et al) and SLAC/LBL eventually found no splitting, which brought an end to this rather sad story. This "protective" attitude towards data, bears resemblance to adventures of the past chronicled by Irving **Langmuir**, one of whose many stories we will turn to shortly, after a final look at several A2 mass spectra.

Examples of “observed” A_2 spectra (CERN missing mass spectrometers)



All data from $\pi^- + p \rightarrow p + X^-$ (of average resolution) show a peak at 1.3 GeV.

While 40% of the data with “better” resolution (≈5 MeV) show a deep dip at 1.3 GeV. It’s hard to prove or disprove such sensitive kinds of effects, especially with small statistics!



The figure shows the **$\pi^\pm\pi^+\pi^-$ mass** spectrum from a **compilation** of exclusive events with different mass resolutions, for all available bubble chamber data of the time. The peak should be flat-topped if resolution matters for the effect, but result is nice and sharp. This was published two years before the 1968 Meeting to check whether the A_2 had a legitimate resonant distribution. Although Miller knew about this, he still chose not to present his measurements at the meeting. I took this to mean that the conferees had lost their mental balance, and immediately started offering good odds against the A_2 splitting, and was thereby rewarded with some Kleingeld for my timely "chuzpah!"

Irving Langmuir on Pathological Science: Davis and Barnes Effect (1930)

(Langmuir, a Nobel Laureate in Chemistry, spent most of his career at General Electric, and became interested in these issues late in his life.)

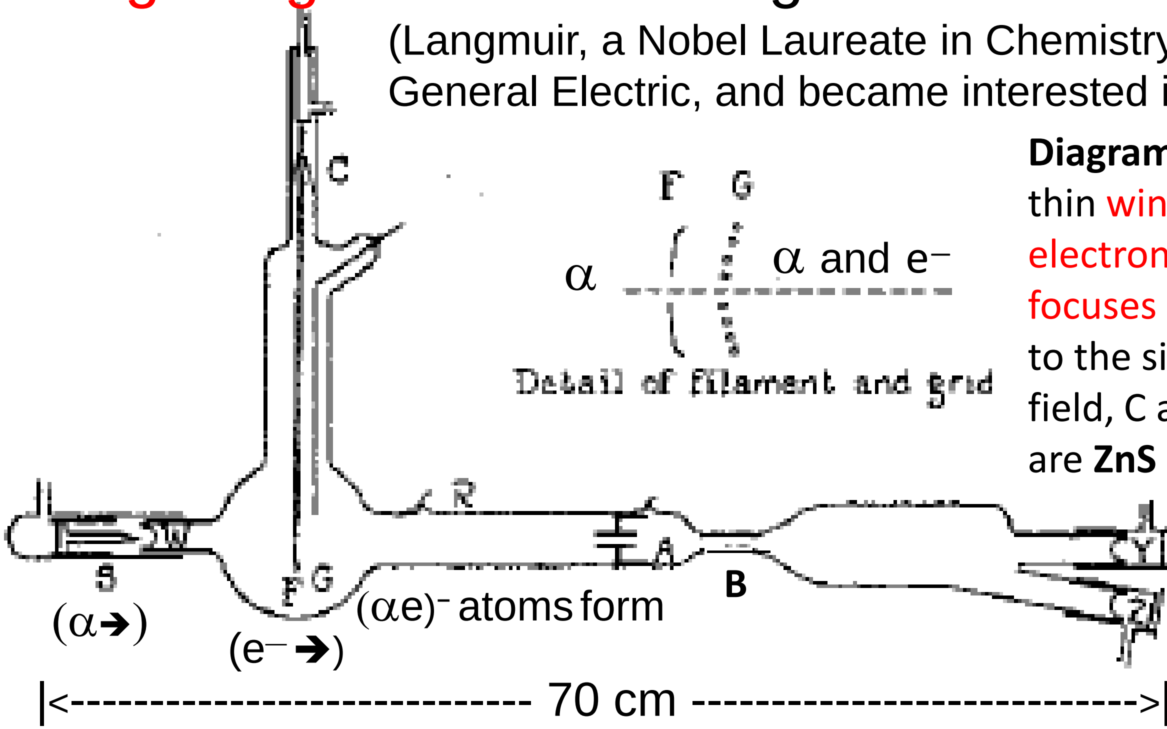


Diagram of apparatus: S is an α source, W a thin window, F is a filament that produces electrons, G an accelerating voltage grid that focuses the e⁻ to a second anode A, R is a lead to the silvered surface, **B** is magnetic bending field, C are water-cooled Cu seals, and Y and Z are **ZnS** scintillation screens (viewed by eye!)

Miracles were observed at Columbia University! At 590 V on the accelerating grid, electrons were expected to move with a mean velocity of the α particles. The two objects were found to combine in a 5 cm section to form bound αe− “atomic states” that continued down the tube to be counted on the ZnS screens at Y or Z by eye through a microscope. The voltage was read to extraordinary accuracy. Despite that the scale had only 1000 divisions from 0 to 1000 V, the capture of the electrons occurred only at very specific values (set to accuracy <0.01 V), and each corresponding to the energy levels calculated for the Balmer series of the helium atom. The apparatus was checked without electrons energized via the filament, and showed similar findings. Naturally, these results raised enormous interest in the scientific community!

It was clear that there were **new physical principles** involved in these studies.

First of all, among the many problems, the kinematics did not make sense:

- The energy of capture into orbit must be a factor of two larger than energy in orbit, and so energies observed for capture can't be same as the binding energies in the Balmer series of the Bohr atom (and this needs γ emission!)
- The observed 80% capture rates (independent of electron current) for each level were impossible to explain.
- The precision claimed for the apparatus was impossibly good, yet results were self consistent and repeatable (by the proponents).
- Intensities of the α source were too small, and spread in velocity of the α particles too great to match the narrow range of capture energies.
- Effect was observed even with the cathode filament cold (no electrons)!

The experimenters had ready answers for all these deficiencies, e.g., electrons don't have to be there, as they are waves, so even cold electrons contribute!

All of this was sufficiently interesting to get Sommerfeld to propose a model.

Eventually, Langmuir figured out that Barnes and Davis, again, without intent to defraud, blinded by their convictions, were hallucinating/making up data!

Issues in Medical and Epidemiological Research

Research in these areas is fraught with influence from communities with financial interests and unscientific, but highly focused, political goals! It also seems that an appreciation of subtle statistical issues is often lacking in these investigations. There are a huge number of independent studies, which can enhance chance of getting individual false results, and add to the difficulty of calculating correct probabilities. We see periodic reports in the media that appear to be very naive, some even violating the laws of physics! Nevertheless, **there are heroes** in these areas who try to alert their colleagues. One astute individual who has written on this subject is John P. A. **Ioannidis**. In his comment "**An Epidemic of False Claims"** (*Sci Am*, June 2011), he states: "False positives and exaggerated results in peer-reviewed scientific studies have reached epidemic proportions in recent years in economics, the social sciences, and even the natural sciences, and is particularly egregious in biomedicine. Studies claiming some drug or treatment as beneficial have turned out not to be true: just look at conflicting findings about beta-carotene, vitamin treatments, Vioxx and Avandia. Even when effects are genuine, their magnitude is often smaller than originally claimed.

"Is there something wrong with the scientific method?"

Jonah Lehrer, December 13, 2010 (New Yorker) ← **familiar name?!**

"…all sorts of **well-established**, multiply confirmed **findings** have started to **look increasingly uncertain**. It's as if our facts were losing their truth: **claims** that have been **enshrined in textbooks are suddenly not provable**. This phenomenon doesn't yet have an official name, but it's occurring across a **wide range of fields**, from psychology to ecology. In the field of medicine, the phenomenon seems extremely widespread. There is a forthcoming analysis demonstrating that **the efficacy of antidepressants has gone down as much as threefold** in recent decades. Just because an idea is true doesn't mean it can be proved. And just because an idea can be proved it doesn't mean it's true. When the experiments are done, **we still have to choose what to believe**."

This is pretty scary talk – for all of science!

Nevertheless, Lehrer understands the problem. Quoting Ioannidis, he says that the **main problem is that too many researchers engage in what he calls "significance chasing,"** or finding ways to interpret the data so that they pass the statistical test of significance—the ninety-five-per-cent boundary invented by Tsar of the frequentists Ronald **Fisher** (and preached by Fermi). "Scientists are so eager to pass this magical mark that they start playing around with the numbers, to find anything that seems worthy." Recently, **Ioannidis** has become increasingly blunt about the **pervasiveness of the problem**. One of his most cited papers has the provocative title: "**Why Most Published Research Findings Are False.**" It's clearly all **Fermi's and Fisher's fault**! (TF confesses that, for shock value, he added the above slights of frequentists!)

Many reports abound that make absolutely no scientific sense!

The entire history of power lines causing cancer shows a depressing lack of scientific understanding (photoelectric effect) by scientists and the media, e.g.,

According to a Scientific Report in *Nature:* "A Prospective Study of In-utero Exposure to Magnetic Fields and the Risk of Childhood Obesity." The study reveals the shocking truth: Mom was overdosing on her cell phone during her pregnancy. Women participating in the study carried a meter during pregnancy that measured magnetic field levels. Their 733 children were followed for 13 years. According to the report, **prenatal exposure to high magnetic field level was found to be associated with increased risk of being obese**. The article concluded: "Maternal exposure to a high magnetic field during pregnancy may be a **new and previously unknown factor contributing to the world-wide epidemic of childhood obesity**." (From **Bob Park's "What's New"** column.)

This is a huge problem that HEP has handled better than most fields!

Comments on pathological science

According to sage Irving **Langmuir**, pathological “small science” can often be characterized through some of the following properties:

- The maximum observed **effect originates from a barely detectable cause**, and the magnitude of the effect is **often independent of size of the cause**!
- The effect is often **close to the limit of experimental sensitivity**, with many measurements needed because of its **small statistical significance**.
- There are **claims for need of great accuracy**.
- **Fantastic theories are generated** to accommodate the result.
- **Criticism is met by ad hoc excuses** thought up on the spur of the moment.
- Ratio of **supporters to critics rises to near 50%** and then falls slowly to nil.

No question but that the wizard from G.E. understood human nature, and some of his symptoms have certainly plagued the false positives in our own field. The more regimented, large collaborations will hopefully keep this problem in check, but it will not disappear because even physicists are just too human.

Summary and a comment about mistakes

Despite the ephemeral nature of many "discoveries," it is crucial to seek, without which we cannot find! When you seek, no matter how experienced you are, you should also keep your wits about you, and understand what you are doing, and how statistics can fool you when you do not stay alert. But also, as was pointed out in a recent article by David Kaiser and Angela Creager in *Sci Am, June 2012,* blunders can often have major impact through stimulating search for a claimed observation not even believed to be correct (not even wrong). Because further **work in inspired areas can lead to even greater discoveries**. For example, the discovery of CP violation in kaon decay owes many thanks to Bob Adair's now forgotten result on K^0_S regeneration, and LIGO and other gravity-wave detectors should be grateful to the enthusiasm engendered in this field many years ago through the work of Joe Weber! **So, if ye seek deep, there's more chance that ye shall find something truly grand! So, I end, after all, on a positive note!**