

Experience from Searches at the Tevatron

Harrison B. Prosper

Department of Physics, Florida State University, Tallahassee, USA

Abstract

I describe, by way of examples, the experience physicists have gained during two decades of searching for physics, both expected and new, at the Fermilab Tevatron.

1 Introduction

2011 marks the end of the Tevatron program [1] and the rapid rise of the Large Hadron Collider (LHC) [2] as the preeminent accelerator in the world. On 22 April, 2011, the LHC reached a luminosity of $4.67 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$, beating the record set by the Tevatron in 2010. Towards the end of the same month, the LHC produced in one week more than double the integrated luminosity of the datasets that yielded the top quark discovery [3]. The era of the Large Hadron Collider is definitely here; 2011 may be remembered not only as a significant year of transition in high energy physics but perhaps also as the year in which the Standard Model (SM) was finally dethroned.

We have reached this crossroad in large measure because of the achievements of physicists at the Tevatron and other accelerator centres around the world. The goals of the Tevatron program were principally to test the SM and to search for significant deviations from it. Alas, none were found. Rather, numerous predictions of the SM have been confirmed, including

1. the shape of jet transverse momentum spectra,
2. the existence of a 6th quark, the top,
3. the existence of reactions in which top quarks are produced singly,
4. the existence of reactions yielding di-bosons ($WW, ZZ, WZ, W\gamma, Z\gamma$),
5. and properties of B mesons.

These achievements, along with several precision measurements, have established the Standard Model as one of humanity's crowning intellectual achievements [4]. The quantitative agreement between the predictions of the theory and observations is stunning, witness Fig. 1, which shows a comparison of the SM predictions for the jet transverse momentum (p_T) spectra — of jets produced in 1.96 TeV proton antiproton collisions — with the unfolded [5] measurements of the D0 Collaboration. The unfolded results agree with the SM predictions over a dynamic range of 10 orders of magnitude. When searching for new physics, it is not surprising that we take the SM, the null hypothesis, very seriously!

Physicists at the LHC are engaged in an intense search for deviations from the SM, continuing the eclectic approach to searches established at the Tevatron. The Tevatron era is drawing to a close, while that of the LHC is ramping up. Given the theme of this meeting, it is an opportune moment to take stock of the statistical procedures we have used in searches at the Tevatron. This experience may inform what we do at the LHC. One purpose of these proceedings is to encourage closer reflection on what we mean when we say we have found something with "high statistical significance". In this paper, I describe the use of statistical procedures at the Tevatron, in the context of searches, using four case studies: a search for a rare decay of a particle, the search for single top, the search for B_s^0 oscillations, and the search for the Higgs boson.

2 Case Studies

I have chosen to describe four somewhat disparate topics in order to illustrate both the similarities and differences in the statistical approaches that have been pursued at the Tevatron. In reviewing the many

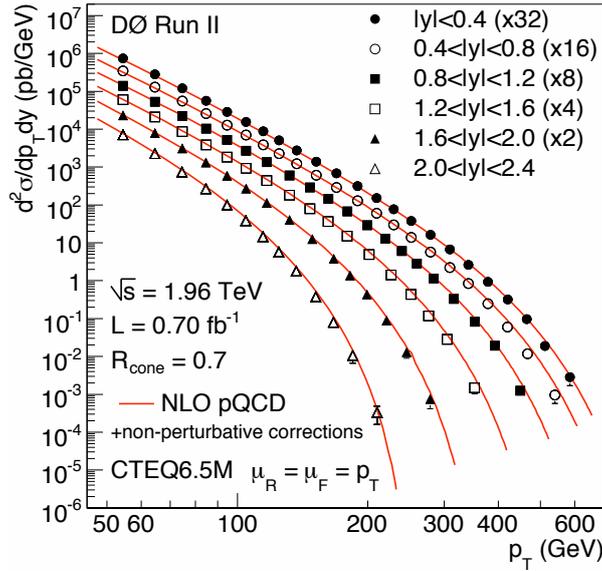


Fig. 1: Comparison of the observed spectrum of jets (points with error bars) with the predictions of the Standard Model (curves) [6].

searches that have been conducted during the period 1991—2011, one notices an interesting sociological evolution. At that start of that period, statistical procedures tended to be described algorithmically with essentially no mention of what statistical procedure was being used nor what quantity was being calculated. Towards the end of that period, however—and one would like to think that this is due in part to the influence of the PHYSTAT series of conferences—words such as *frequentist*, *Bayesian*, *coverage*, *p-value*, *nuisance parameters*, *profile likelihood*, *prior*, etc., began to appear in a few high-profile physics publications. Since these words are now an accepted part of the lexicon of analysis, I shall use them freely in describing the case studies, whether or not such jargon was used in the cited publications.

Another interesting aspect of the statistical work at the Tevatron, and typical of the field, is that almost all hypotheses tested have been *nested* in that the null hypothesis is a special case of the alternative. The canonical example is the search for a signal s above some background μ . The null hypothesis of no signal, $s = 0$, is nested within the alternative hypotheses that the expected event count is $s + \mu$.

2.0.1 Particle Physics Data

From a statistical viewpoint, high energy physicists perform near-perfect Bernoulli trials, tens of millions of times every second. A trial in the context of high energy physics is a collision between particles—protons against antiprotons at the Tevatron and protons against protons or heavy ions against heavy ions at LHC, while a success is some desired outcome. A success could be say the creation of a Higgs boson one of whose decay products (perhaps a muon) has a momentum that falls within a given momentum bin. Each collision yields about 1MB of data. However, of the tens of millions of collisions that occur per second, it is feasible to record only a few hundred per second. The trick, of course, is to ensure that the ones recorded are potentially the most interesting. The data from each collision, that is, *event*, are compressed by a factor of 10^3 – 10^4 during a process called *event reconstruction*, the goal of which is to infer from the raw data the characteristics of the particles that emanated from the collision point. The cartoon in Fig. 2 illustrates how, ideally, different species of particles are manifested in the particle detectors. It is from the known patterns of particle/detector interactions that the identity of particles can be inferred.

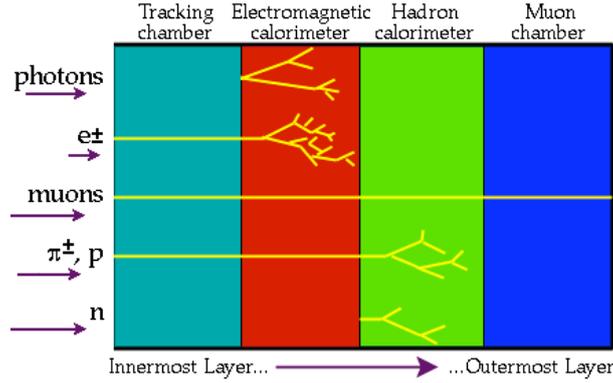


Fig. 2: This cartoon illustrates how, ideally, each species of particle interacts with different components of the detector. In practice, the manner in which particles interact with the detector components is not as clear-cut as this cartoon suggests; ambiguities can arise that lead to particle mis-identification—a jet, for example, could be misconstrued as an electron. (Courtesy CDF Collaboration.)

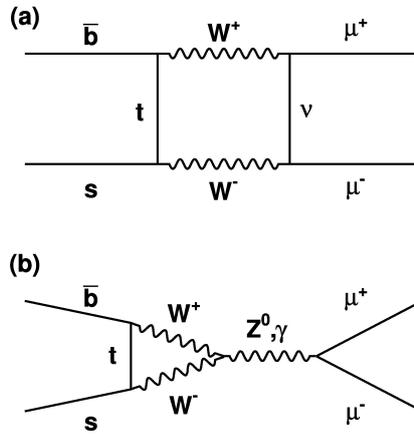


Fig. 3: These diagrams [7] depict a prediction of the SM: the annihilation of a \bar{b} quark and an s quark to a current with zero net charge (a neutral current) that materializes into a pair of oppositely charged muons. Diagram (a) is referred to as a *box* diagram for obvious reasons and diagram (b) is called a *penguin* diagram for reasons that require some imagination.

2.1 Search for a Rare Decay

The search for rare processes, such as the search by the D0 Collaboration described here, is a potentially fruitful way to look for new physics. In many theories of possible new physics, the rates for processes that are rare in the SM are typically predicted to be much higher. Therefore, the observation of a decay rate that differs significantly from the SM prediction would be unambiguous evidence of new physics.

The goal of the search by D0 [7] was to test the SM prediction,

$$\mathcal{B} = \frac{B_s^0 \rightarrow \mu^+ \mu^-}{B_s^0 \rightarrow \text{anything}} = (3.6 \pm 0.3) \times 10^{-9}. \quad (1)$$

The decay $B_s^0 \rightarrow \mu^+ \mu^-$ is an example of a process in which there is an apparent neutral current (that is, a current with a net charge of zero) between quarks of different flavor, here the \bar{b} and s quarks. This is an example of a so-called flavor changing neutral current (FCNC) interaction, which are rare in the SM. The lowest order Feynman diagrams describing $B_s^0 \rightarrow \mu^+ \mu^-$ are shown in Fig. 3. Table 1 shows the results obtained by D0. These data are described by the 2-count likelihood model

Table 1: D0 results for $B_s^0 \rightarrow \mu^+ \mu^-$: observed event counts, estimated background counts and the scale factors that relate the branching fraction \mathcal{B} to the signals, via $\mathcal{B} = f_i s_i$ with $i = a, b$. The subscripts pertain to the two Tevatron run periods, RunIIa and RunIIb.

Run period	observed count (events)	estimated background count (events)	estimated scale factors ($\times 10^{-9}$)
RunIIa	$n_a = 256$	264 ± 13	4.90 ± 1.00
RunIIb	$n_b = 823$	827 ± 23	1.84 ± 0.36

$$p(n|s, \mu) = \text{Poisson}(n_a|s_a + \mu_a) \text{Poisson}(n_b|s_b + \mu_b), \quad (2)$$

where n , s and μ are the observed counts, expected signal and expected background counts, respectively. The branching fraction \mathcal{B} is related to the expected signals through scale factors f_i , where $\mathcal{B} = f_i s_i$ with $i = a, b$. The likelihood in Eq. (2) therefore contains one parameter of interest, namely the branching fraction \mathcal{B} , and the four nuisance parameters f_a, f_b, μ_a , and μ_b . Information about the nuisance parameters is encoded in an *evidence-based* prior $\pi(f_a, f_b, \mu_a, \mu_b)$, modeled as the product of four normal distributions with the means and standard deviations listed in Table 1, one set for each nuisance parameter. (Given the size of the uncertainties for the scale factors, listed in Table 1, the priors for these parameters were, in fact, truncated Gaussians.)

The likelihood in Eq. (2) was marginalized with respect to the nuisance parameters, f_a, f_b, μ_a , and μ_b to yield the marginal likelihood $p(n|\mathcal{B})$. From this, the limit $\mathcal{B} < 5.1 \times 10^{-8}$ at 95% C.L. was derived using the CL_s method [8]. In the CL_s method one defines the tail probability

$$p_1(\mathcal{B}) = \Pr[t < t_0 | H_1(\mathcal{B})], \quad (3)$$

for some suitable statistic t , for a given (alternative) hypothesis H_1 about the branching fraction \mathcal{B} . One then rejects all values of \mathcal{B} for which $p_1(\mathcal{B}) < \gamma p_1(0)$ and *defines* a $(1 - \gamma)$ C.L. upper limit as the smallest rejected value of \mathcal{B} . The statistic used by D0 is the logarithm of the *Bayes factor* $p(n|\mathcal{B})/p(n|0)$. (It is a Bayes factor rather than a likelihood ratio because the marginal likelihoods entail integrations over priors.)

2.2 Search for Single Top

The goal of this search is to test the SM prediction that the process

$$p + \bar{p} \rightarrow t + X, \quad (4)$$

exists in which the set of particles denoted by X does *not* contain a top quark. (The top quark was discovered [3] through the reaction $p + \bar{p} \rightarrow t\bar{t}$.) The SM predicts how often the reaction in Eq. (4) should occur, which is quantified in terms of the cross section $\sigma(p + \bar{p} \rightarrow t + X) = 3.46 \pm 0.18$ pb (assuming a top quark mass of 170 GeV). At the Tevatron, this cross section corresponds to a production rate of about 1 in 10 billion collisions, which is just under half the rate for the pair production of top quarks. It would seem therefore that the search for single top ought not be that much harder than was the search for top quark pairs ($t\bar{t}$). In fact, owing to the greater similarity between the signal and background events, the search for single top proved to be considerably more challenging. This is illustrated in Fig. 4, which shows the event sample composition before b-tagging (that is, before selecting events with identified b-quark jets), but after the first level of cuts. The signal to background ratio at this stage was a daunting 1 : 260.

It was clear from the outset, that only the most sophisticated methods of analysis were likely to yield a successful outcome in a reasonable amount of time. Indeed, the first evidence of the existence of single top reactions [9] and their subsequent definitive observation by CDF [10] and D0 [11] both made

DØ Single Top 2.3 fb⁻¹ Signals and Backgrounds
(All channels combined, before *b*-tagging)

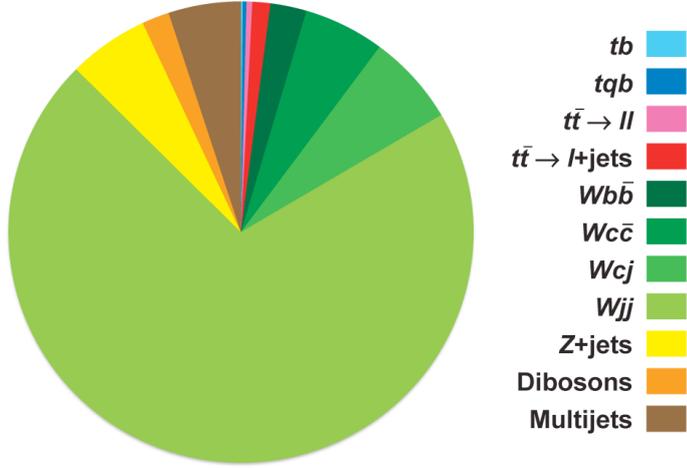


Fig. 4: Predicted composition of the DØ data that were the basis of the single top discovery. CDF predicted a similar composition. The single top signal is the thin wedge at the top of the pie-chart. (Courtesy DØ Collaboration.)

extensive use of multivariate discrimination methods such as boosted decision trees (BDT), Bayesian neural networks (BNN), and *ab initio* semi-analytical calculations of the signal and background probability densities, the so-called Matrix Element (ME) method. This was the first time in high energy physics that a major discovery was based on such methods. The extensive use of Bayesian methods, by DØ, was another first.

After reducing the multivariate data x to a discriminant function $D(x)$, the data were binned into M bins in the variable D (see Fig. 5). The M counts are described by a likelihood function similar in structure to that used in the rare decay search (see Section 2.1),

$$p(n|\sigma, \epsilon, \mu) = \prod_{i=1}^M \text{Poisson}(n_i | \epsilon_i \sigma + \mu_i), \quad (5)$$

where σ , the single top cross section, is the parameter of interest and the $2M$ nuisance parameters ϵ_i and μ_i , respectively, are the expected effective integrated luminosities (integrated luminosity \times signal efficiency \times signal acceptance) and the expected background counts, respectively, while n_i are the observed bin counts. Information about the nuisance parameters was encoded in an evidence-based prior $\pi(\epsilon, \mu)$ modeled as a multivariate normal distribution that took account of the known correlations between the nuisance parameters. The overall prior $\pi(\sigma, \epsilon, \mu)$ was factorized as follows $\pi(\sigma, \epsilon, \mu) = \pi(\epsilon, \mu | \sigma) \pi(\sigma) = \pi(\epsilon, \mu) \pi(\sigma)$ and $\pi(\sigma)$ was taken to be a *flat* prior.

The posterior density resulting from the integration over the nuisance parameters is shown in Fig. 6. The DØ analysts considered Bayes factors, $p(n|\sigma)/p(n|0)$, but chose, in the end, to follow tradition and estimate the significance of the single top observations using a prior-predictive p-value, $p_0 = \text{Pr}[t > t_0 | H_0]$, computed using a null hypothesis (H_0) in which the expected background is marginalized with respect to the background prior. The statistic t (which of course could have been any suitable function of the data) was taken to be the mode of the posterior density, $p(\sigma|n)$. The basic intuition is that larger values of the cross section σ cast greater doubt on the null, that is, on the background-only hypothesis.

The distribution of t , shown in Fig. 7, was simulated using 67.8 million pseudo datasets, generated with background only, which, for the measured cross section of 3.94 pb, yielded a prior predictive p-value

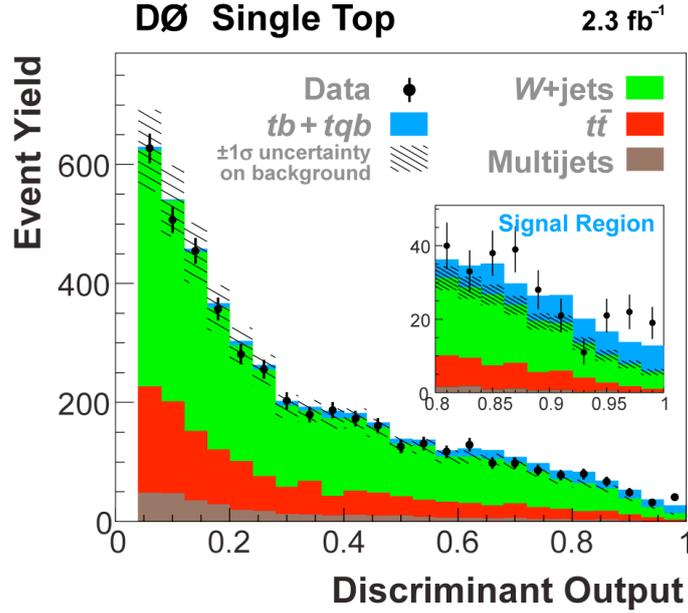


Fig. 5: Distribution of the Bayesian neural network discriminant that combines the discriminants from three separate, but correlated, DØ analyses (BDT, BNN, and ME) [11] into a single overall discriminant $D(x)$. The single top signal is the thin line that becomes slightly more visible as one moves towards the signal region, $D(x) \rightarrow 1$. (Courtesy DØ Collaboration.)

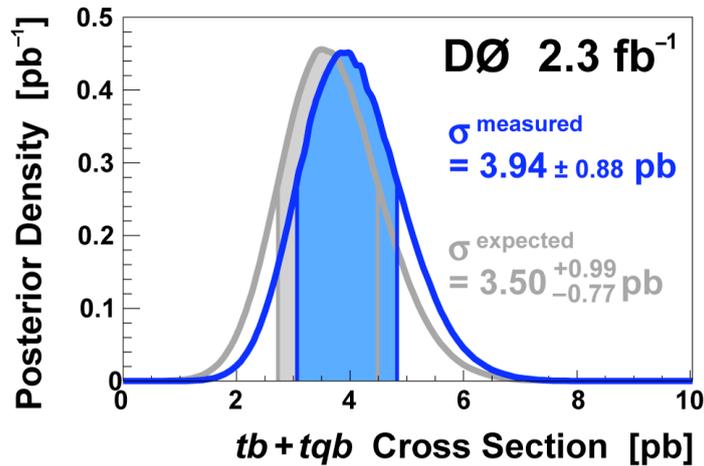


Fig. 6: The posterior density $p(\sigma|n)$ is plotted as a function of the total cross section ($\sigma(p + \bar{p} \rightarrow tb) + \sigma(p + \bar{p} \rightarrow tqb)$) assuming the SM prediction for the ratio of the tb and tqb cross sections. The expected cross section is computed using an *Asimov* dataset (see Cowan, these Proceedings), that is, an artificial dataset in which the “observed” counts are set equal to the sum of the expected background and signal counts, assuming the SM prediction for the signal. (Courtesy DØ Collaboration.)

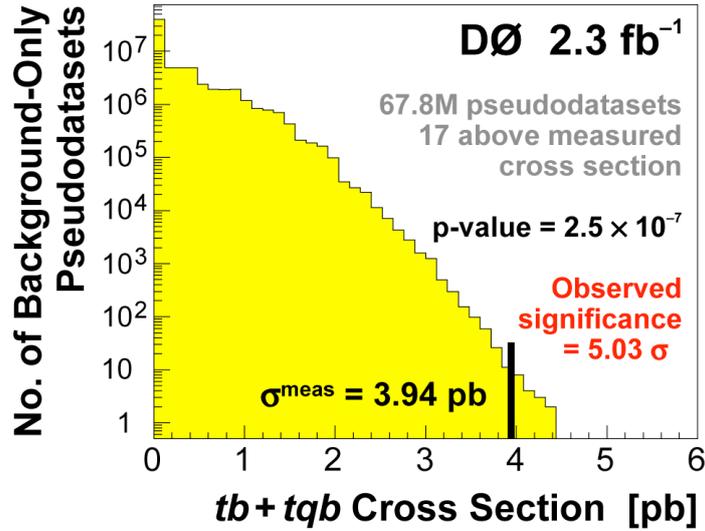


Fig. 7: Distribution of the statistic $t = \text{mode}[p(\sigma|n)]$ over a background-only ensemble of pseudo datasets. (Courtesy D0 Collaboration.)

of 2.5×10^{-7} . Again, by tradition, this was converted to a normal standard deviation scale as an observed significance of 5 standard deviations, the long-accepted threshold in high energy physics for claiming a discovery.

2.3 Search for B_s^0 Oscillations

The goal of this search by the CDF Collaboration [12] was to test the SM prediction that the B_s^0 and \bar{B}_s^0 mesons form an oscillating pair in which each member of the pair changes into its partner at an (extremely) high frequency predicted by the SM. The oscillations are governed by the time-dependent probability densities,

$$\begin{aligned} p_{B_s^0 \rightarrow \bar{B}_s^0}(t|A, \Delta m) &= \frac{1}{2\tau} e^{-t/\tau} [1 - A \cos \Delta m t], \\ p_{\bar{B}_s^0 \rightarrow B_s^0}(t|A, \Delta m) &= \frac{1}{2\tau} e^{-t/\tau} [1 + A \cos \Delta m t], \end{aligned} \quad (6)$$

where A is the amplitude of the oscillations and Δm characterizes its frequency. According to the SM, $A = 1$.

There were two important complications with this search. Firstly, the measured time of decay t of a B meson—inferred from the measured displacement of the B meson decay point from the proton antiproton collision point—was measured with an uncertainty that varied from meson to meson. The uncertainty on t was modeled with a normal distribution with a *heteroscedastic* variance, that is, one that varied from one measurement to the next. Secondly, the oscillation signal was contaminated with background arising from other processes. The probability model for t was therefore taken to be a convolution of a normal density with a mixture model comprising an oscillatory signal plus a background. The likelihood for these data is then just a product of M terms, one for each measured time t_i ,

$$p(t|A, \Delta m) \sim \prod_{i=1}^M N(t_i|t', \sigma_i^2) \otimes [\alpha p(t'|A, \Delta m) + (1 - \alpha) b(t')], \quad (7)$$

where α is the signal fraction and $b(t)$ the background density. Since the oscillation frequency Δm is predicted to be high (about 18 cycles per picosecond), it proved more satisfactory to perform a maximum

likelihood fit for the amplitude A , using $p(t|A, \Delta m)$, for different fixed values of Δm . It was found that at $\Delta m = 17.8$ cycles / ps, $A = 1.21 \pm 0.20$, which is consistent with the SM prediction $A = 1$ and inconsistent with $A = 0$. The amplitude A was then set to unity and Δm was measured to be $17.77 \pm 0.10(\text{stat}) \pm 0.07(\text{syst})$ cycles / ps.

The CDF Collaboration quantified the statistical significance of its results using the p-value $p_0 = \Pr[\Lambda < \Lambda_0 | H_0]$ based on the statistic $\Lambda = \log[p(t|H_0)/p(t|H_1, \Delta m)]$, where $p(t|H_0) \equiv p(t|A = 0)$ and $p(t|H_1, \Delta m) \equiv p(t|A = 1, \Delta m)$ are the densities for the null and alternative hypotheses, respectively. Small values of Λ provide evidence against the null. At $\Delta m = 17.8$ cycles / ps, the value of the test statistic Λ was observed to be -17.26 [12] from which the p-value of 8×10^{-8} was calculated by a Monte Carlo technique. This p-value, being rather smaller than the traditional threshold for discovery, fully justified the title ‘‘Observation of $B_s^0 - \bar{B}_s^0$ Oscillations’’ of the CDF article announcing this result.

2.4 Search for the Higgs

Since the start of the current millennium—and building on the searches at LEP and earlier machines, high energy physicists have been engaged in a relentlessly intensifying search for the Higgs boson. This particle, or something that mimics it, is a critical ingredient of the SM, being a vestige of the mechanism through which mass is introduced into a theory that would otherwise describe an unrealistic world of massless particles. Its fundamental role in the SM is reason enough to sustain the Higgs search effort that began at LEP and the Tevatron and that continues apace at the LHC.

But, of course, the Higgs boson may not exist. From a certain point of view, it would be a spectacularly exciting outcome were it to be shown convincingly that no such particle exists with a mass less than about 1 TeV. On the other hand, finding it rather than not finding it is pretty exciting too! If a low-mass neutral Higgs boson exists, we would be in a position akin to that during the search for the top quark. During that search, we ‘‘knew’’ everything about the top quark since all of its characteristics, with the exception of its mass, were predicted in detail from the SM. Moreover, the mass of the top quark was inferred from radiative corrections to precision measurements. Likewise for the Higgs searches: if a SM Higgs boson exists, we know a lot about it [13]. Indeed, the searches for the Higgs boson rely extensively on detailed predictions from the SM. When the Tevatron data from CDF and D0 are analyzed in the context of the SM, one obtains the results shown in the left plot of Fig. 8. In this figure is plotted the 95% credible level (C.L.) upper limit R^{up} , given by

$$0.95 = \int_0^{R^{up}} p(R|n, m_H) dR, \quad (8)$$

as a function of the Higgs mass hypothesis, where the posterior density is given by

$$p(R|n, m_H) \propto \prod_{i=1}^{N_C} \prod_{j=1}^{N_{bi}} \text{Poisson}(n_{ij} | R s_{ij} + \mu_{ij}) \pi(R, s_{ij}, \mu_{ij}, m_H), \quad (9)$$

and $R \equiv \sigma/\sigma_{SM}$, with σ_{SM} the predicted SM cross section for the creation of Higgs bosons of a given mass. The index i ranges from 1 to N_C final state channels, while the index j is over the N_{bi} data bins in the i^{th} channel. The quantities s_{ij} are the predicted signals, for a given Higgs boson mass, assuming the validity of the Standard Model. The prior incorporates the uncertainty in these predictions. Systematic uncertainties can be incorporated by representing the prior $\pi(R, s, \mu, m_H)$ as an integration,

$$\pi(R, s, \mu, m_H) = \int \pi(R, s, \mu, m_H | \theta) \pi(\theta) d\theta, \quad (10)$$

with respect to (hyper) parameters θ that characterize the systematic effects.

In the right plot in Fig. 8 is displayed a summary of the LEP results: the negative log-likelihood as a function of the Higgs mass.

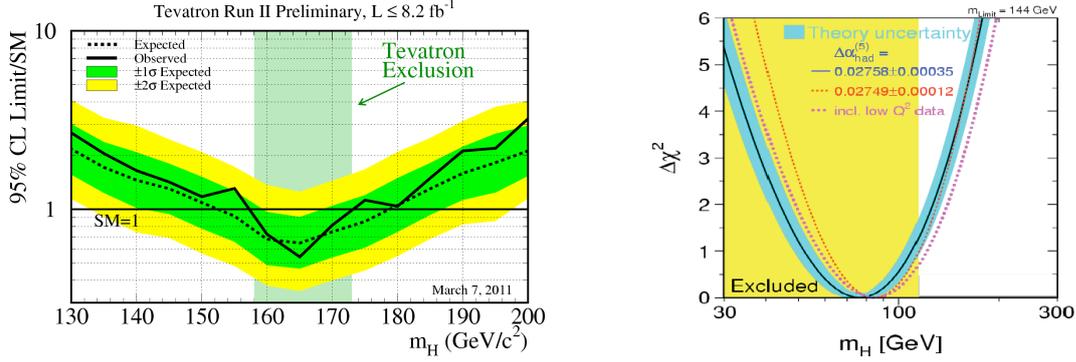


Fig. 8: (left) The plot summarizes the conclusions of the Tevatron New Phenomena and Higgs Working Group (TEVNPHWG) regarding the SM Higgs boson. The dark line is the observed upper limit from a Bayesian calculation, while the bands indicate by how much the limits may be expected to fluctuate in a large number of simulated repetitions of the (combined) CDF and D0 experiments. Note: the bands depend on the ensemble used to calculate them. Therefore, a different ensemble would yield different bands. It is useful to keep this mind and, it is hoped, keep in check the temptation to over-interpret the bands, useful though they are. (right) The plot summarizes the conclusions of the LEP experiments. The negative log-likelihood curve, which favours a low-mass Higgs, is the result of a fit to the LEP precision measurements. The vertical band on the left of the plot ends at 114 GeV, the lower limit set by the LEP experiments as a result of direct searches for the SM Higgs boson. (Courtesy TEVNPHWG.)

So what, if anything, can we say about the Higgs hypothesis, given the assumption that its description in the SM is correct? An answer to this question would require a coherent integration of the information presented in Fig. 8. In principle, the LEP curve provides the probability density $p(m_H|H_1)$ (assuming a flat prior in the Higgs mass—though a reference prior would be better), while the Tevatron results provide $p(s|m_H, H_1)$, where now s denotes the expected Higgs signal and H_1 denotes the Higgs hypothesis and the numerous assumptions on which the LEP and Tevatron results depend¹.

Given the densities, $p(m_H|H_1)$ and $p(s|m_H, H_1)$, it would be natural from a Bayesian viewpoint to compute the expected signal density,

$$p(s|H_1) = \int_{100}^{200} p(s|m_H, H_1) p(m_H|H_1) dm_H, \quad (11)$$

by marginalizing over the Higgs boson mass, m_H . Then, physicists at the LHC, if so inclined, could use $p(s|H_1)$ as an evidence-based prior $\pi(s)$ in their LHC Higgs searches. How might it be used? It could be used, for example, in conjunction with LHC likelihoods to test the Higgs hypothesis using a Bayes factor (see Berger, these Proceedings),

$$B_{10} = \int_0^\infty p(\text{LHC-data}|s + \mu) \pi(\mu) \pi(s) d\mu ds / \int_0^\infty p(\text{LHC-data}|\mu) \pi(\mu) d\mu, \quad (12)$$

which can be mapped to a scale akin to n -sigma using the transformation $Z = \sqrt{2 \log B_{10}}$.

3 Conclusions

Numerous discoveries have been made at the Tevatron in spite of our eclectic (and sometimes baroque) approach to interpreting results in a statistical manner. If there is one theme throughout, it is that we

¹In practice, because these curves are summaries, they do not provide enough information to actually carry out the integration of this information. Here is a compelling case for making the full probability model, plus the observations, available using, for example, RooStats workspaces.

remain ferociously fond of exact frequentist coverage. Hence the Herculean efforts to achieve “exact” 5 standard deviation results. Moreover, p-values remain the principle measure of “surprise”: if a p-value is small enough, we judge that something surprising and presumably exciting has happened. Rarely has the notion of *power* been explicitly addressed in Tevatron analyses, though simple measures of experimental sensitivity have become routine, such as the notion of expected limits. The Poisson model remains ubiquitous as does the use of the normal distribution as a model for systematic uncertainties. However, there is a growing realization that we can, and should, do a better job of designing probability models using more appropriate functions, such as gamma or log-normal densities, for modeling systematic uncertainties. The RooFit/RooStats system now makes this possible (see Schott, these proceedings).

Bayesian methods have made significant inroads, witness for example the discovery of single top by $D\bar{0}$, which was Bayesian through and through, until the very end when a p-value was used to quantify the significance of the observations.

Physicists are still prone to statistical invention, even when perfectly satisfactory alternatives exist. But the good news is that we can be taught!

Acknowledgements

I wish to thank Luc Demortier and the organizers of this very interesting conference, in particular, Louis Lyons and Albert de Roeck.

References

- [1] The Tevatron, Fermilab, <http://www.fnal.gov/pub/science/accelerator>.
- [2] The Large Hadron Collider, CERN, <http://public.web.cern.ch/public/en/lhc/lhc-en.html>.
- [3] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995); D0 Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).
- [4] See for example, J. Lefrancois, “The standard model: 30 years of glory” in Proceedings: Techniques And Concepts In High-Energy Physics, Eds. H.B. Prosper and M. Danilov, (Dordrecht, The Netherlands, Kluwer, 2001, 411p., Nato Science Series C: Mathematical and Physical Sciences, Vol. 566).
- [5] See the contributions on unfolding in these Proceedings.
- [6] D0 Collaboration, V.M. Abazov, *et al.*, Phys. Rev. Lett. **101**, 062001 (2008).
- [7] D0 Collaboration, V.M. Abazov, *et al.*, Phys. Rev. Lett. **B693**, 539 (2010); aXiv:1006.3469 [hep-ex].
- [8] Luc Demortier, these Proceedings.
- [9] D0 Collaboration, V.M. Abazov, *et al.*, Phys. Rev. Lett. **98**, 18102 (2007); Phys. Rev. D**68**, 012005 (2008).
- [10] CDF Collaboration, T. Aaltonen *et al.*, Phys. Rev. Lett. **103**, 092002 (2009).
- [11] D0 Collaboration, V.M. Abazov *et al.*, Phys. Rev. Lett. **103**, 092001 (2009).
- [12] CDF Collaboration, A. Abulencia *et al.*, Phys. Rev. Lett. **97**, 242003 (2006).
- [13] TEVNPHWG, <http://tevnpnphwg.fnal.gov>.