

Moderne Methoden der Datenanalyse – Hypothesentests –

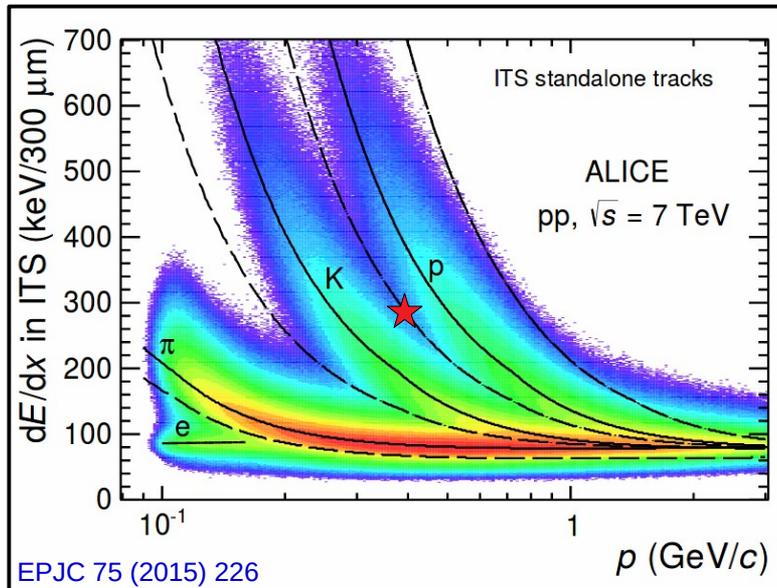
Roger Wolf
18. Juni 2020

Inhalt dieser Vorlesung

- Effizienz und Reinheit
- ROC Kurve
- Poisson-basierter Zweihypothesen Test
- Systematische Unsicherheiten in der Likelihood
- Beispiel: Eigenschaften des Higgs Bosons

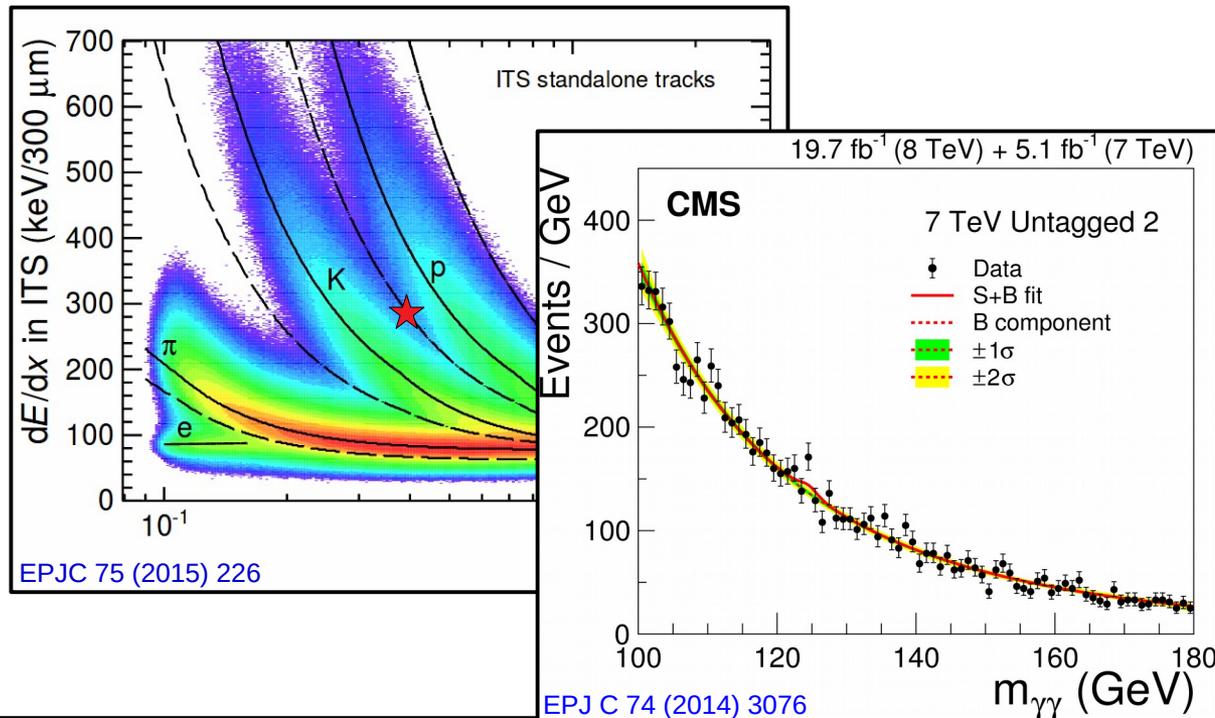
Wichtigste Hypothesentests in der (Teilchen-)physik

- Ist die Energiedeposition eines Teilchens mit vorgegebenem Impuls in Materie kompatibel mit einer Teilchenhypothese eines Elektrons, Pions, Kaons oder Protons?



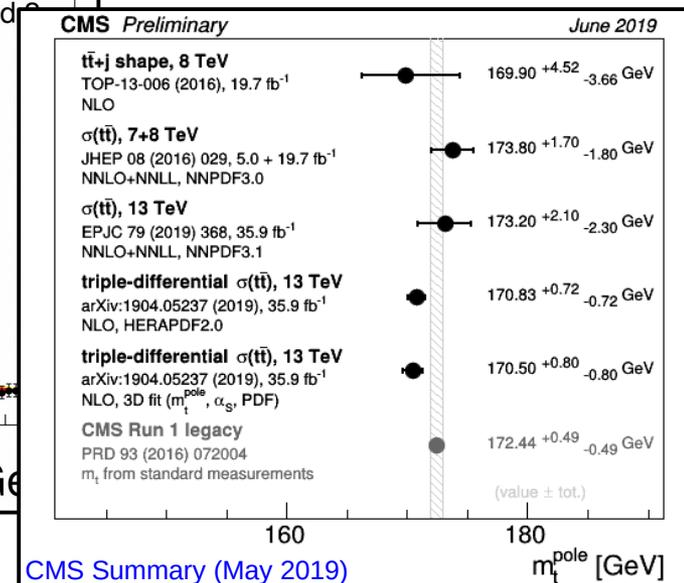
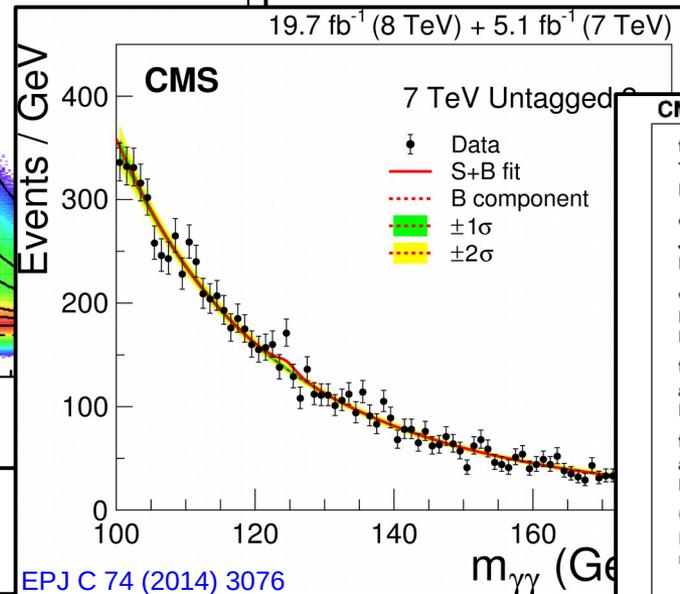
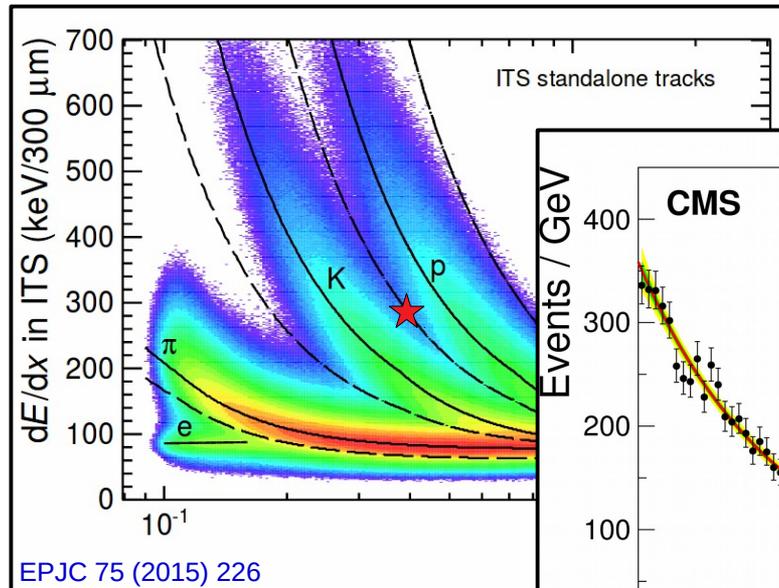
Wichtigste Hypothesentests in der (Teilchen-)physik

- Ist die Energiedeposition eines Teilchens mit vorgegebenem Impuls in Materie kompatibel mit einer Teilchenhypothese eines Elektrons, Pions, Kaons oder Protons?
- Weisen die Daten auf ein neues Teilchen zusätzlich zu einem bekannten Untergrund hin?



Wichtigste Hypothesentests in der (Teilchen-)physik

- Ist die Energiedeposition eines Teilchens mit vorgegebenem Impuls in Materie kompatibel mit einer Teilchenhypothese eines Elektrons, Pions, Kaons oder Protons?
- Weisen die Daten auf ein neues Teilchen zusätzlich zu einem bekannten Untergrund hin?
- Sind zwei oder mehrere Messungen der vermeintlich gleichen physikalischen Größe miteinander kompatibel?



Effizienz und Reinheit

- Wir beschäftigen uns im folgenden mit dem Test: ist ein Teilchen vom Typ-1 oder vom Typ-0?
- In diesem einfachen Fall besteht die folgende Relation zwischen dem Fehler 1. und 2. Art und den häufig verwendeten Begriffen Effizienz und Reinheit:

$$\epsilon = \frac{\text{Richtig erkannte Ereignisse vom Typ-0}}{\text{Alle Ereignisse vom Typ-0}} = 1 - \alpha \quad (\text{Effizienz})$$

$$\rho = \frac{(1 - \alpha) N_0}{(1 - \alpha) N_0 + \beta N_1} \quad (\text{Reinheit})$$

α : Fehler 1. Art

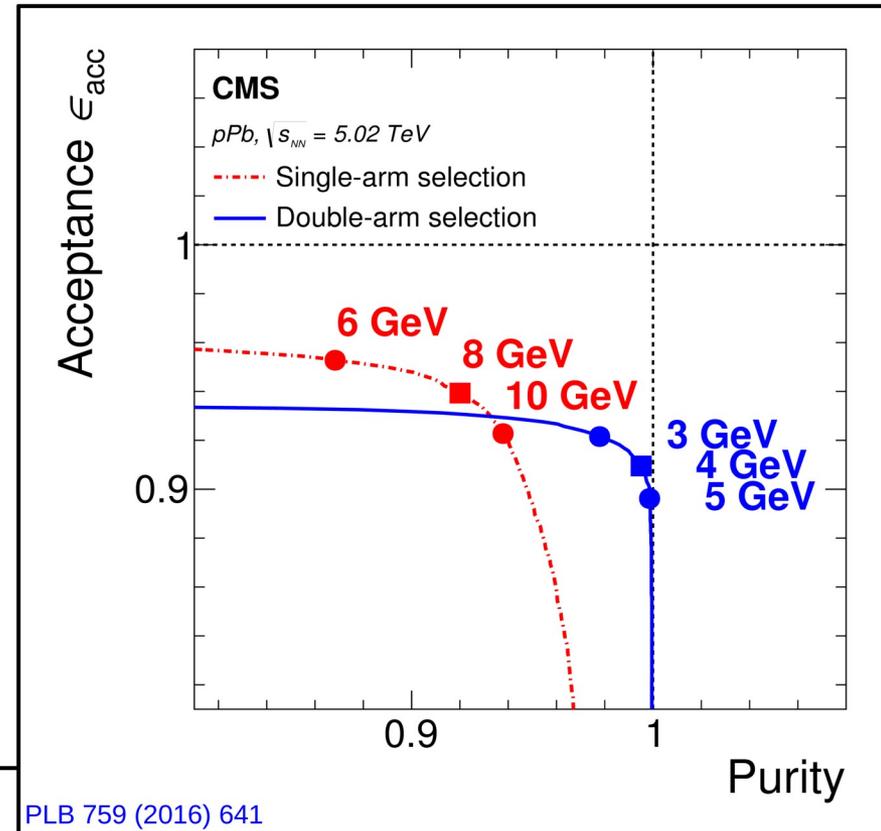
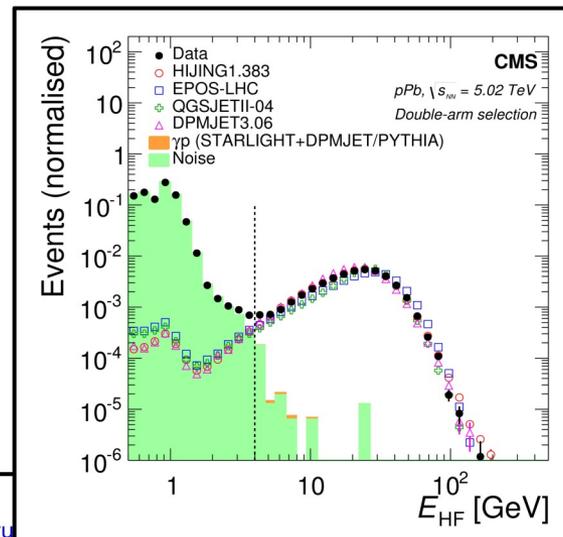
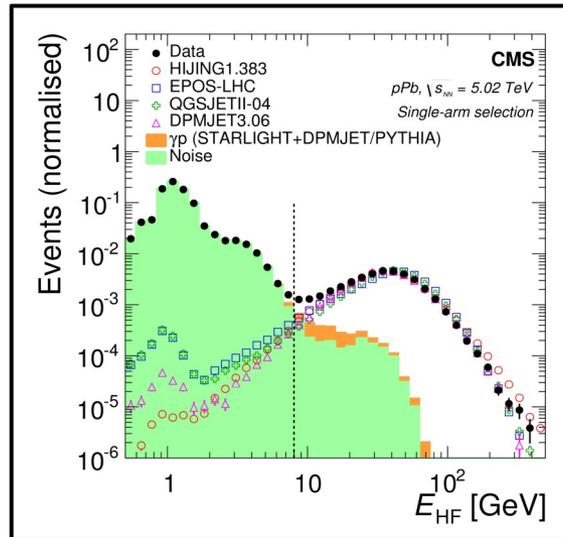
β : Fehler 2. Art

N_0 : Wahre Anzahl der Ereignisse vom Typ-0

N_1 : Wahre Anzahl der Ereignisse vom Typ-1

ROC Kurve

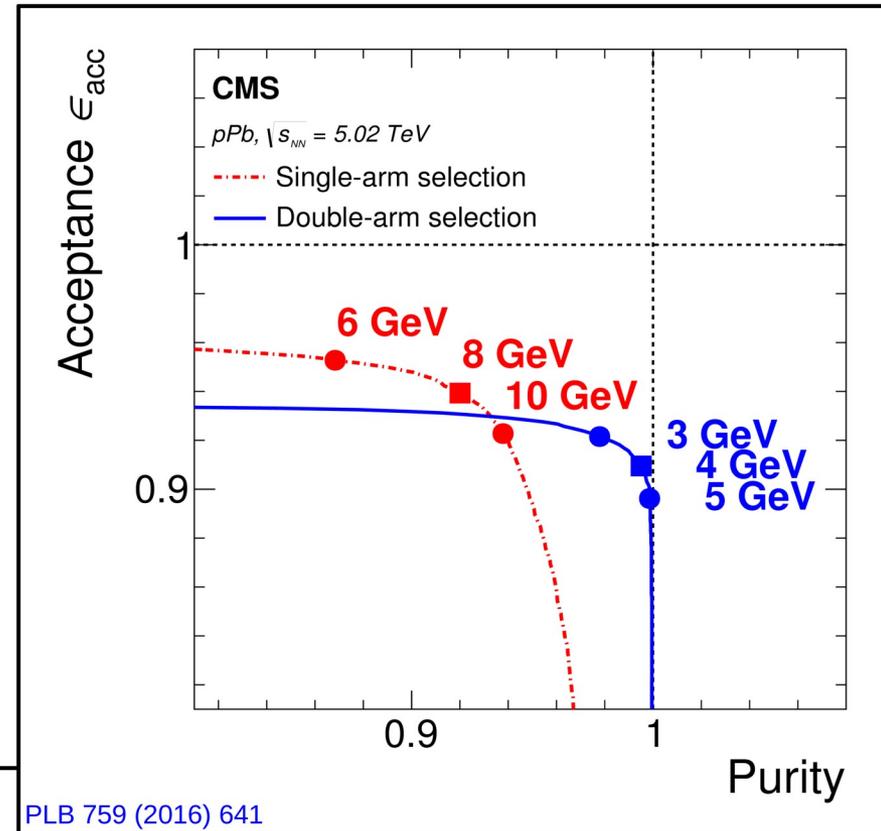
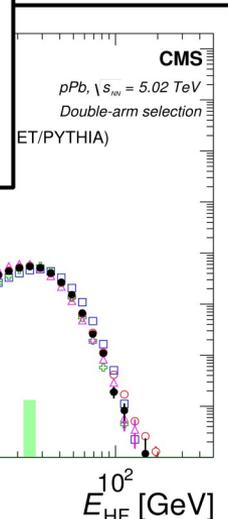
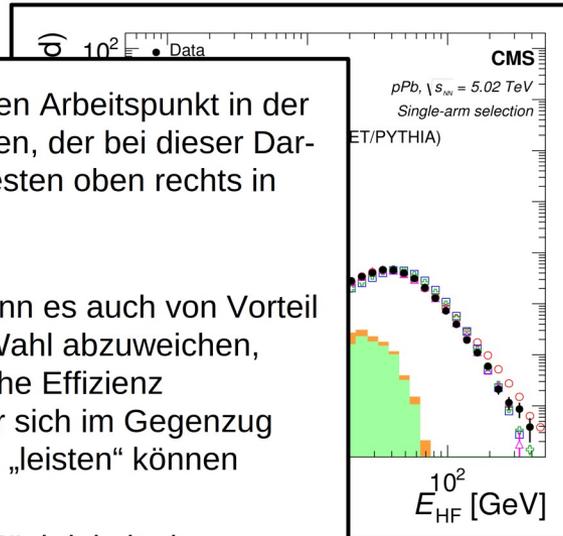
- In solchen Fällen, in denen sich Effizienz und Reinheit definieren lassen wird die Machtfunktion eines Tests auch durch die *Receiver Operating Characteristics* (ROC) Kurve angegeben.



ROC Kurve

- In solchen Fällen, in denen sich Effizienz und Reinheit definieren lassen wird die Machtfunktion eines Tests auch durch die *Receiver Operating Characteristics* (ROC) Kurve angegeben.

- Sie würden i.a. den Arbeitspunkt in der ROC Kurve wählen, der bei dieser Darstellung am weitesten oben rechts in der Kurve liegt.
- In Einzelfällen kann es auch von Vorteil sein von dieser Wahl abzuweichen, wenn Sie z.B. hohe Effizienz bevorzugen, oder sich im Gegenzug niedrige Effizienz „leisten“ können
- Als Maß für die Mächtigkeit eines Algorithmus gibt man in diesen einfachen Fällen oft das Integral unter der ROC Kurve an (*area under curve* AUC).

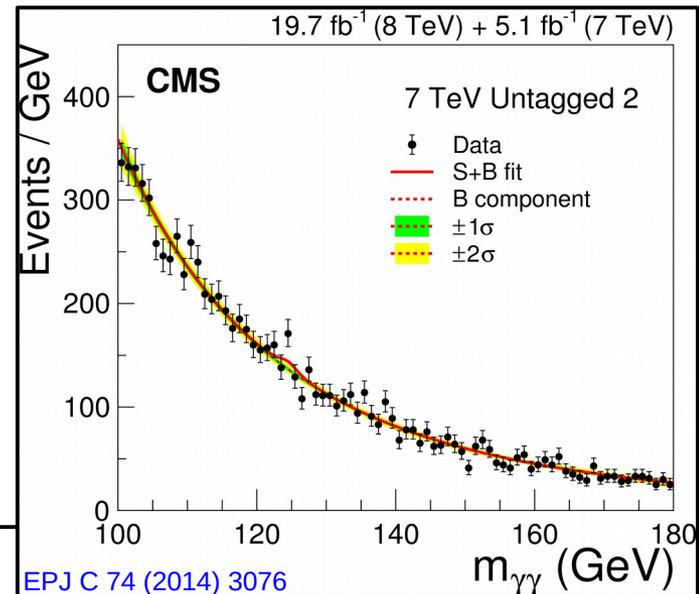


Zweihypothesen Poisson Test

- Ein klassischer und wichtiger Test in der Teilchenphysik, den wir im folgenden tiefer diskutieren werden ist der Zweihypothesen Test basierend auf der Poisson Zählstatistik, z.B.:
- „Weist die Messung auf ein neues Teilchen zusätzlich zu einem bekannten Untergrund hin?“
- In diesem Fall haben wir zwei Hypothesen, z.B.:

H_0 : Etablierte Hypothese (nur Untergrund)

H_1 : Alternative Hypothese (zusätzliches Signal)



Lemma von Neyman-Pearson (einfache Hypothesen)

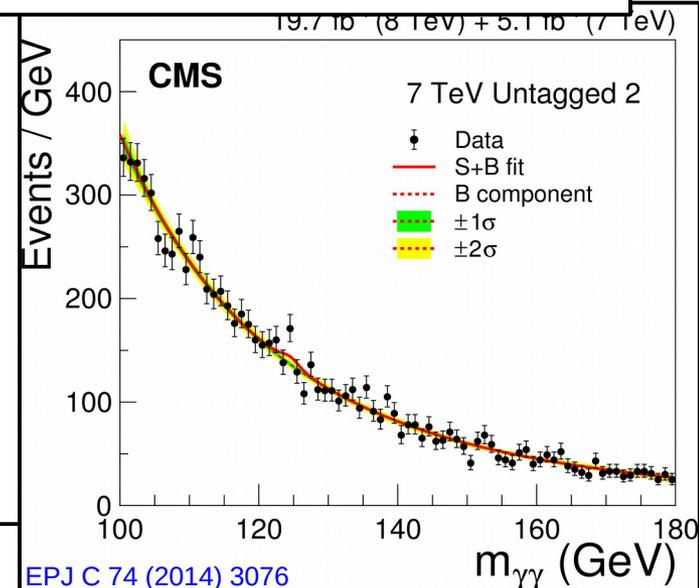
Für den Test zweier einfacher Hypothesen H_0 und H_1 ist für den Likelihood-Quotienten

$$\lambda(\vec{x}) = \frac{\mathcal{L}(\vec{x}|H_0)}{\mathcal{L}(\vec{x}|H_1)}$$

$1 - \beta$ bei vorgegebenem α maximal, d.h. $\lambda(\vec{x})$ ist die mächtigste Teststatistik. Der Likelihood-Quotient $\lambda(\vec{x})$ ist äquivalent zur Differenz der Log-Likelihood:

$$q(\vec{x}) = -2 \ln \left(\frac{\mathcal{L}(\vec{x}|H_0)}{\mathcal{L}(\vec{x}|H_1)} \right) = -2 \left(\ln(\mathcal{L}(\vec{x}|H_0)) - \ln(\mathcal{L}(\vec{x}|H_1)) \right)$$

- Das Lemma von Neyman-Pearson gilt nur für einache Hypothesen, es lässt sich jedoch auf zusammengesetzte Hypothesen verallgemeinern (siehe nächste Folie).



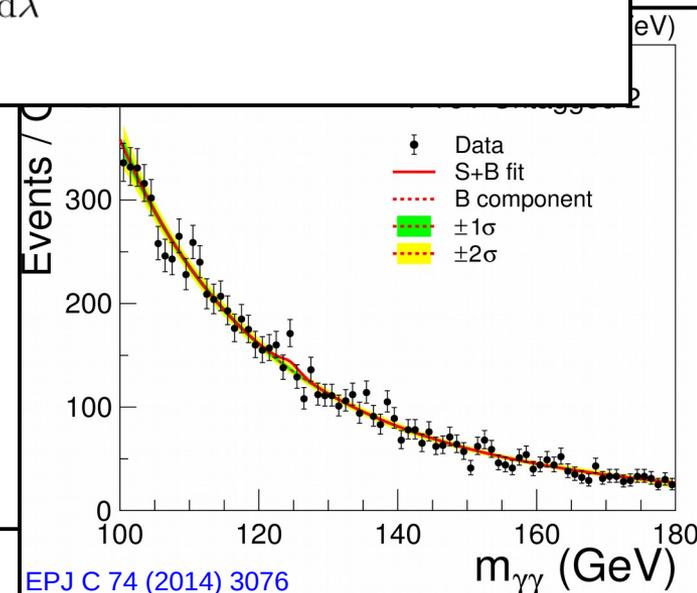
Lemma von Neyman-Pearson (zusammengesetzte Hyp.)

Für eine Verallgemeinerung des Lemmas von Neyman-Pearson bilden Sie den Likelihood-Quotienten aus

$$\lambda(\vec{x}, \{\theta_i\}) = \frac{\max_{\theta_i} (\mathcal{L}(\vec{x}, \{\theta_i\} | H_1))}{\max_{\theta'_i \in \Omega} (\mathcal{L}(\vec{x}, \{\theta'_i\} | H_0))},$$

wobei $\theta'_i \in \Omega$ bedeutet, dass die Parameter $\{\theta'_i\}$ mit H_0 kompatibel sein müssen. Die Hypothese H_0 wird mit der Irrtumswahrscheinlichkeit α verworfen, falls

$$\alpha = P(\lambda > \lambda_\alpha | H_0) = \int_{\lambda_\alpha}^{\infty} g(\lambda | H_0) d\lambda$$



Beispiel CP-Eigenzustand des beobachteten Higgs Bosons

- Goldener Zerfallskanal für solche Messungen: $H \rightarrow ZZ^* \rightarrow 4\ell$

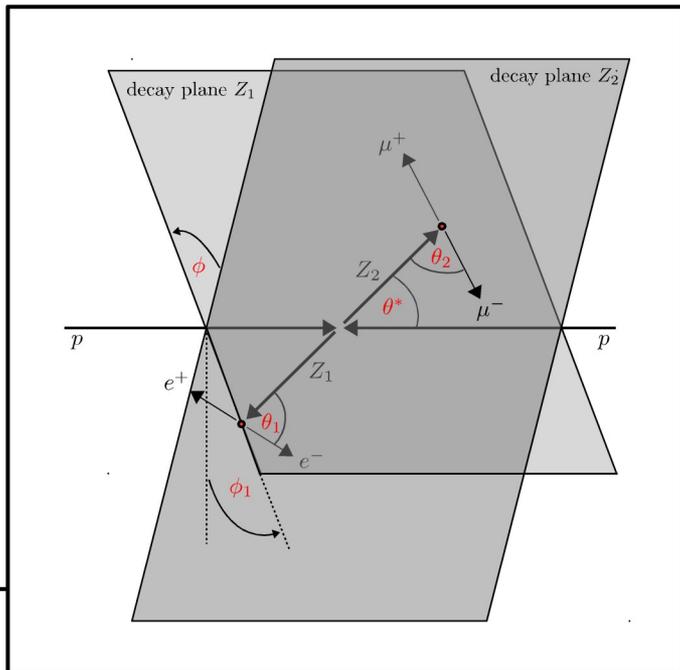
$$P(Y_L^m(\theta, \varphi)) = (-1)^L \cdot Y_L^m(\theta, \varphi)$$

$$P(4\ell) = (-1)^L (-1)^2 (+1)^2 = (-1)^L$$

Paritätsoperator

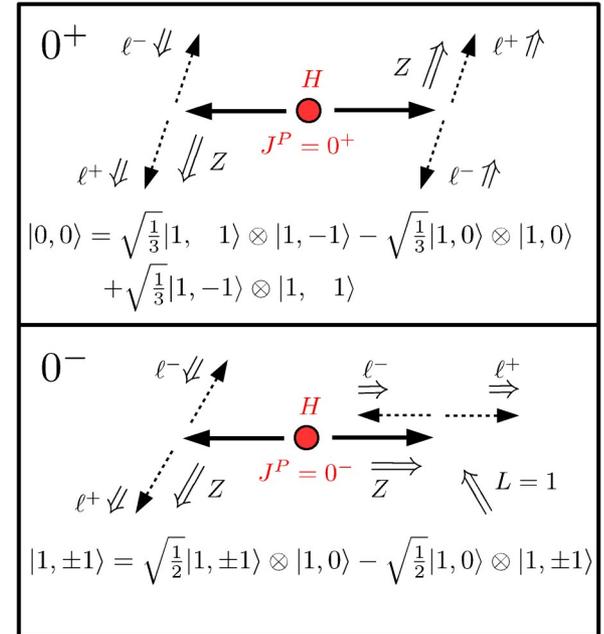
Anteil der Ortswellenfunktion

Anteil der Spins im Endzustand



Characterisierung der Zerfallsebenen des $H \rightarrow ZZ^* \rightarrow 4\ell$ Zerfalls.

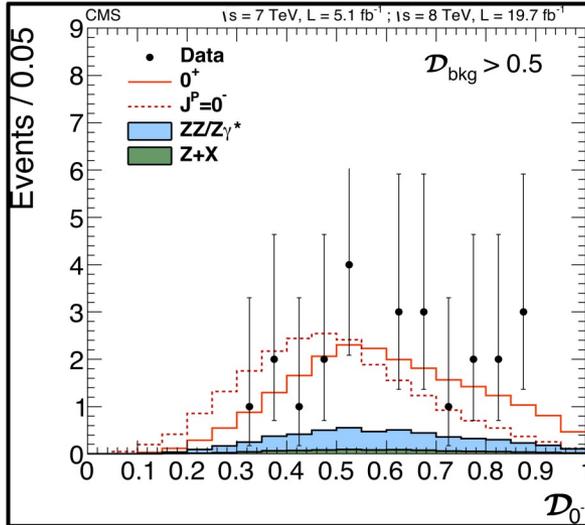
Beispiel, wie Drehimpuls- und Spinkonfiguration Rückschlüsse auf den CP-Eigenzustand des Higgs Bosons zulassen:



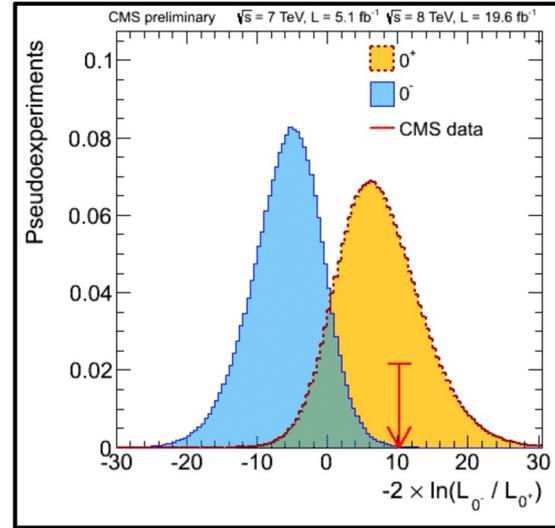
NB: für alle, die sich wundern, warum bei CP Analysen immer nur P diskutiert wird, der Eigenwert von $C(4\ell)$ ist immer +1 (siehe z.B. Kapitel 5 [hier](#) für eine genauere Diskussion).

Beispiel CP-Eigenzustand des beobachteten Higgs Bosons

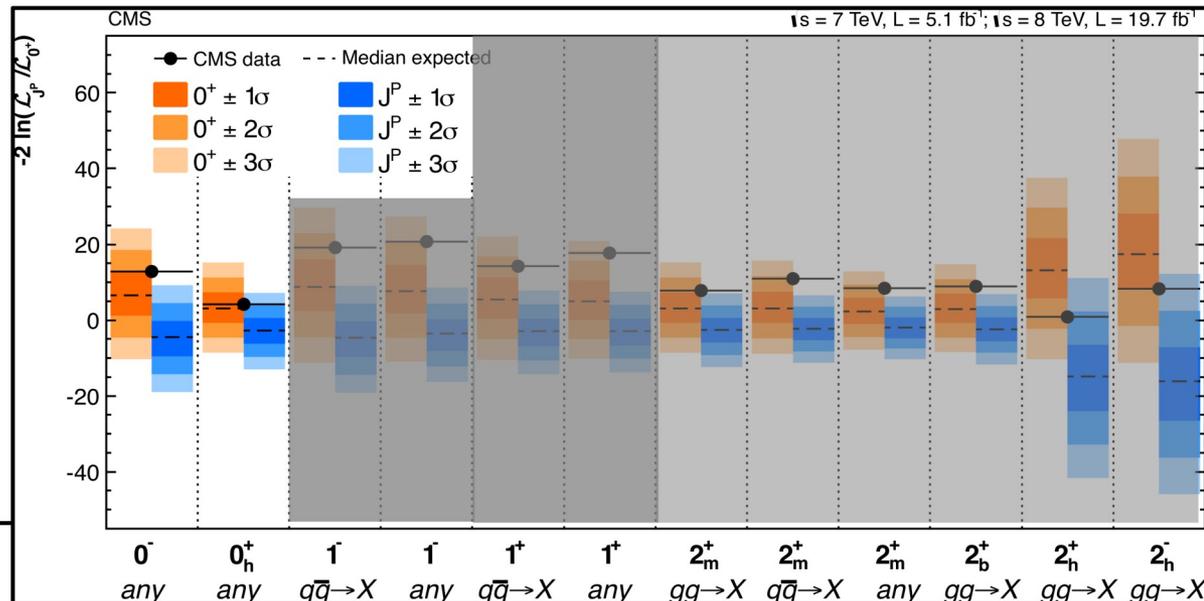
- Goldener Zerfallskanal für solche Messungen: $H \rightarrow ZZ^* \rightarrow 4\ell$



PRD 89 (2014) 092007



Test reiner Spin-Paritäts Hypothesen basierend auf $\mathcal{O}(50)$ Ereignissen:



Theorem von Wilks

Für Stichproben $\{x_i\}$, $i = 0, 1, \dots, n$ einer Zufallsvariablen x der Länge $n \rightarrow \infty$ gilt: Wird die Grundgesamtheit der Zufallsvariablen x durch eine Wahrscheinlichkeitsdichte $p(x, \{\theta_j\})$ beschrieben, und legt die Hypothese H_0 $r = m_1 - m_0$ der Parameter $\{\theta_j\}$ fest, so folgt die Testfunktion

$$q(\vec{x}, \vec{\theta}) = -2 \ln \left(\lambda(\vec{x}, \vec{\theta}) \right)$$

einer χ^2 -Verteilung mit r Freiheitsgraden. Dabei entspricht m_i der Anzahl der freien Parameter der Hypothese H_i .

Annals Math. Statist. 9 (1938) 1

- **Beispiel:** Zählexperiment mit einem Signal (s) zusätzlich zu bekanntem Untergrund (b)

$$q(\vec{x}, \theta) = -2 \ln \left(\frac{\mathcal{L}(\vec{x}|\theta s + b)}{\mathcal{L}(\vec{x}|b)} \right) \propto \chi^2(x', 1) \longleftarrow \chi^2\text{-Verteilung mit einem Freiheitsgrad}$$

- D.h. wenn Sie sich sicher sind, dass Ihre Stichprobe den Anforderungen von Wilks' Theorem entspricht können Sie p-Werte und Irrtumswahrscheinlichkeiten analytisch und ohne die Auswertung von Pseudo-Experimenten bestimmen.

Beispiel-1: Gaußverteilte Stichprobe

- **Hypothesentest:** Hat eine Gaußverteilte Stichprobe $\{x_i\}$ der Länge n den Mittelwert μ_0 ?
- Likelihood:

$$\mathcal{L}(\{x_i\}, \mu, \sigma) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right)$$

- Likelihood-Quotient:

$$\lambda(\{x_i\}, \mu_0, \sigma) = \exp \left(-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n (x_i - \mu_0)^2 - \sum_{i=1}^n (x_i - \bar{x})^2 \right) \right)$$

Maximiert \mathcal{L} im gesamten Parameterraum

$$q(\{x_i\}, \mu_0, \sigma) = -2 \ln(\lambda) = \frac{n}{\sigma^2} (\mu_0 - \bar{x})^2$$

μ_0 fest vorgegeben für H_0

$$\begin{aligned} \sum_{i=1}^n (x_i - \mu_0)^2 - \sum_{i=1}^n (x_i - \bar{x})^2 &= \left(\sum_{i,j=1}^n x_i x_j - 2 \underbrace{\sum_{i=1}^n x_i \mu_0 + n \mu_0^2}_{\equiv 2n\bar{x}\mu_0} \right) - \left(\sum_{i,j=1}^n x_i x_j - 2 \underbrace{\sum_{i=1}^n x_i \bar{x} + n \bar{x}^2}_{\equiv 2n\bar{x}^2} \right) \\ &= n\mu_0^2 + 2n\bar{x}\mu_0 - n\bar{x}^2 = n(\mu_0 - \bar{x})^2 \end{aligned}$$

Beispiel-1: Gaußverteilte Stichprobe

- **Hypothesentest:** Hat eine Gaußverteilte Stichprobe $\{x_i\}$ der Länge n den Mittelwert μ_0 ?

- Likelihood:

$$\mathcal{L}(\{x_i\}, \mu, \sigma) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right)$$

- Likelihood-Quotient:

$$\lambda(\{x_i\}, \mu_0, \sigma) = \exp \left(-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n (x_i - \mu_0)^2 - \sum_{i=1}^n (x_i - \bar{x})^2 \right) \right)$$

Maximiert \mathcal{L} im gesamten Parameterraum

$$q(\{x_i\}, \mu_0, \sigma) = -2 \ln(\lambda) = \frac{n}{\sigma^2} (\mu_0 - \bar{x})^2$$

μ_0 fest vorgegeben für H_0

- Die Stichprobenverteilung für \bar{x} ist bekannt: $f(\bar{x}) = N(\bar{x}, \mu_0, \sigma/\sqrt{n})$.
- Wir ermitteln die Stichprobenverteilung $g(q)$ durch Variablentransformation bei vorgegebener Stichprobenverteilung $f(\bar{x}) = N(\bar{x}, \mu_0, \sigma/\sqrt{n})$.

Beispiel-1: Gaußverteilte Stichprobe

- **Hypothesentest:** Hat eine Gaußverteilte Stichprobe $\{x_i\}$ der Länge n den Mittelwert μ_0 ?
- **NB:** Der Wertebereich von $g(q)$ wird zweimal überdeckt, einmal für $\bar{x} > \mu_0$ und einmal für $\bar{x} < \mu_0$. Es gilt also:

$$g(q) = 2 \left| \frac{d\bar{x}}{dq} \right| f(\bar{x}) \quad ; \quad f(\bar{x}) = N(\bar{x}, \mu_0, \sigma/\sqrt{n}) = \sqrt{\frac{n}{2\pi\sigma^2}} \exp\left(-\frac{n}{2\sigma^2} (\mu_0 - \bar{x})^2\right)$$

$$q(\bar{x}) = \frac{n}{\sigma^2} (\mu_0 - \bar{x})^2 \quad ; \quad \bar{x} = \mu_0 - \sqrt{\frac{\sigma^2}{n}} q^{\frac{1}{2}} \quad ; \quad \left| \frac{d\bar{x}}{dq} \right| = \frac{1}{2} \sqrt{\frac{\sigma^2}{n}} q^{-\frac{1}{2}}$$

$$g(q) = 2 \frac{1}{2} \sqrt{\frac{\sigma^2}{n}} q^{-\frac{1}{2}} \cdot \sqrt{\frac{n}{2\pi\sigma^2}} \exp\left(-\frac{n}{2\sigma^2} (\bar{x} - \mu_0)^2\right) \Big|_{\bar{x}=\sqrt{\frac{\sigma^2}{n}} q^{\frac{1}{2}} + \mu_0} = \frac{1}{\sqrt{2\pi}} q^{-\frac{1}{2}} \exp\left(-\frac{q}{2}\right)$$

- In diesem Beispiel folgt $g(q)$ der χ^2 -Verteilung für $r = 1$ Freiheitsgrade für beliebe Werte von n

$$\chi^2(x, r = 1) = \frac{1}{2^{r/2} \Gamma(r/2)} x^{r/2-1} e^{-x/2} \Big|_{n=1} = \frac{1}{\sqrt{2\pi}} q^{-\frac{1}{2}} \exp\left(-\frac{q}{2}\right)$$

Beispiel-2: Zählexperiment mit Poisson-Statistik

- **Hypothesentest:** Kleines Signal s auf bekanntem Untergrund $b \gg s$.

$$\mathcal{P}_0(n, b) = \frac{b^n}{n!} e^{-b} \quad ; \quad \mathcal{P}_1(n, s + b) = \frac{(s + b)^n}{n!} e^{-(s+b)} \quad ; \quad n = s + b$$

$$\lambda = \frac{P_0(n, b)}{P_1(n, s + b)}$$

$$q = -2 \ln \left(\left(\frac{b}{s + b} \right)^n e^{-b+s+b} \right) = -2 \left(n \ln \left(\frac{b}{s + b} \right) + s \right) = 2 \left((s + b) \ln \left(\frac{s}{b} + 1 \right) - s \right)$$

mit: $s \ll b$

$$\ln(1 + x) = \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu+1}}{\nu} x^{\nu} = \frac{s}{b} - \frac{1}{2} \frac{s^2}{b^2} + \mathcal{O} \left((s/b)^3 \right) \quad \text{für } x \in (-1; 1]$$

$$q \approx 2 \left((s + b) \left(\frac{s}{b} - \frac{1}{2} \frac{s^2}{b^2} \right) - s \right) = 2 \left(\frac{s^2}{b} + s - \frac{1}{2} \frac{s^3}{b^2} - \frac{1}{2} \frac{s^2}{b} - s \right) \approx \frac{s^2}{b}$$

Beispiel-2: Zählexperiment mit Poisson-Statistik

- Dem Theorem von Willks folgend gilt:

$$q = \chi^2(x', r = 1) \Big|_{n_{\text{obs}}} = \frac{(n_{\text{obs}} - \mu_0)^2}{\sigma_0^2} = \frac{s_{\text{obs}}^2}{\sigma_0^2} \quad \text{NB: Vergleichen Sie mir [Lecture 05](#) Folie 18.}$$

mit $\mu_0 = b$, $n_{\text{obs}} = (s_{\text{obs}} + b) \equiv y$ und

$$z = \sqrt{\frac{(y - \mu_0)^2}{\sigma_0^2}} = \boxed{\sqrt{q} \approx \frac{s_{\text{obs}}}{\sqrt{b}}} \quad (\text{Signifikanz der Beobachtung})^1)$$

- Man erkennt an dieser Rechnung das die altbekannte Formel für die Berechnung der Signifikanz bei einem einfachen Zählexperiment

$$z = \frac{s_{\text{obs}}}{\sqrt{b}}$$

nur näherungsweise gilt. NB: Die Näherung besteht in der Annahme $\sigma_0 = \sqrt{b}$.

¹⁾ zur Erinnerung: die Signifikanz ist die Abweichung der Beobachtung von der Erwartung in Vielfachen der Standardabweichung.

Systematische Unsicherheiten in der Likelihood

- Systematische Unsicherheiten können mit Hilfe von Störparametern (*nuisance parameters*) in die Likelihood eingebracht werden. Wir diskutieren das Vorgehen am Fall eines Signals auf bekanntem Untergrund von [Lecture 05](#) (Folie 9):
- Erinnerung:** ΔNLL ist äquivalent zu einem Likelihood-Quotienten

$$\Delta\text{NLL} = -2 \ln \left(\frac{\mathcal{L}(\vec{x}, \theta | \mu s + b)}{\mathcal{L}(\vec{x}, \theta | b)} \right)$$

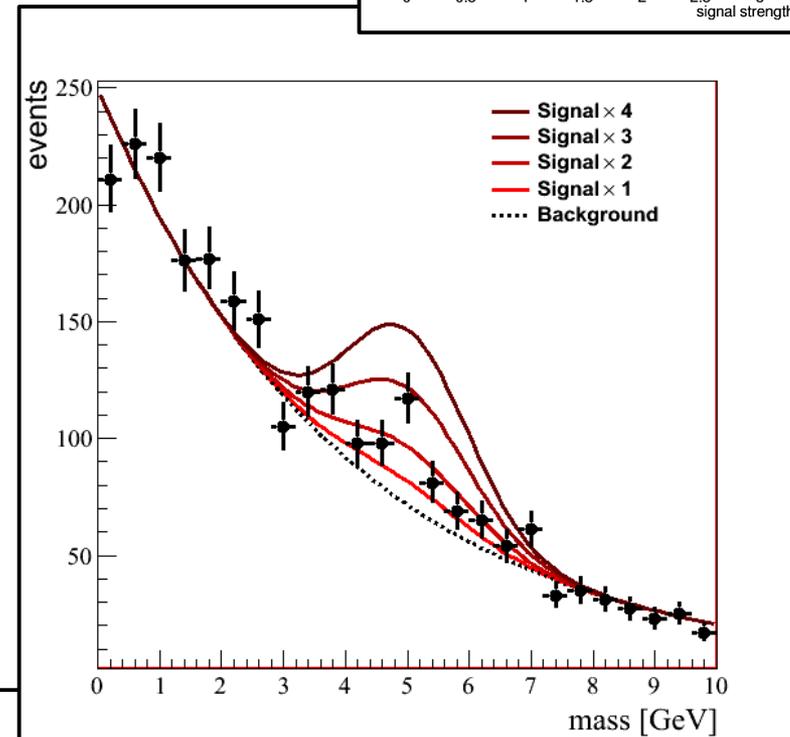
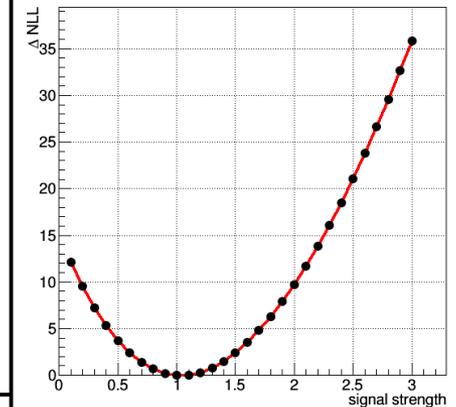
$$\mathcal{L}(\{x_k\}, \{\theta_j\}) = \prod_{k \leq 25} \mathcal{P}(\{x_k\}, \nu_k(\{\theta_j\}))$$

Produkt Poissondichte
für 25 bins.

$$\nu_k(\{\theta_j\}) = \underbrace{\theta_0 e^{-\theta_1 x_k}}_{\text{Untergrund}} + \underbrace{\theta_2 e^{-(\theta_3 - x_k)^2}}_{\text{Signal}}$$

Ein lauffähiges RooT macro finden Sie [hier](#).

Likelihood Scan



Systematische Unsicherheiten in der Likelihood

- Wir führen einen Störparameter für die Normierung der Untergrundverteilung ein. Hierfür nehmen wir an, dass die Untergrund-Normierung θ_0 in unserem Modell nicht exakt bekannt ist, sondern nur mit einer rel. Unsicherheit von $\sigma(\theta_0)/\theta_0 = \pm 5\%$.

$$\nu_k(\{\theta_j\}) = \underbrace{\tilde{\theta}_0 f(\theta_0/\tilde{\theta}_0, 1, \sigma(\tilde{\theta}_0)/\tilde{\theta}_0)}_{\equiv \rho(\theta_0|\tilde{\theta}_0)} e^{-\theta_1 x_k} + \theta_2 e^{-(\theta_3 - x_k)^2}$$

Wahrscheinlichkeitsdichte für θ_0 bei gegebener Abschätzung $\tilde{\theta}_0$.

Mögliche wahre Werte für θ_0 (auszuintegrieren).

Beste Abschätzung des wahren Wertes in Likelihood

$$\mathcal{L}(\{x_k\}, \{\theta_j\}) = \prod_{k \leq 25} \underbrace{\mathcal{P}(\{x_k\}, \nu_k(\{\theta_j\}))}_{\text{Produkt Poissondichte für 25 bins.}}$$

$$\nu_k(\{\theta_j\}) = \underbrace{\theta_0 e^{-\theta_1 x_k}}_{\text{Untergrund}} + \underbrace{\theta_2 e^{-(\theta_3 - x_k)^2}}_{\text{Signal}}$$

Ein lauffähiges RooT macro finden Sie [hier](#).

Systematische Unsicherheiten in der Likelihood

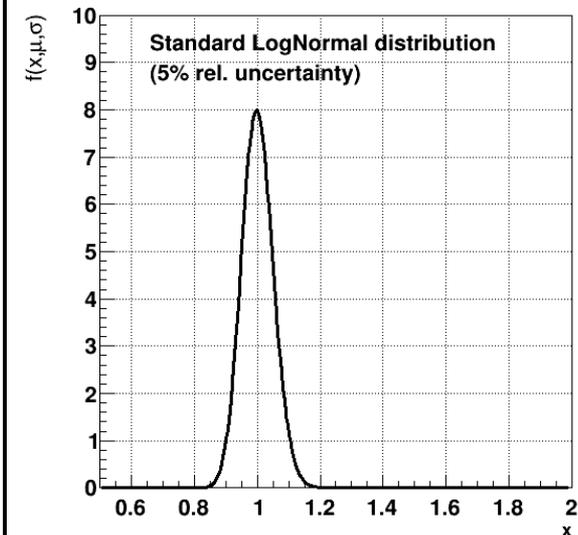
- Wir führen einen Störparameter für die Normierung der Untergrundverteilung ein. Hierfür nehmen wir an, dass die Untergrund-Normierung θ_0 in unserem Modell nicht exakt bekannt ist, sondern nur mit einer rel. Unsicherheit von $\sigma(\theta_0)/\theta_0 = \pm 5\%$.

$$\nu_k(\{\theta_j\}) = \underbrace{\tilde{\theta}_0 f(\theta_0/\tilde{\theta}_0, 1, \sigma(\tilde{\theta}_0)/\tilde{\theta}_0)}_{\equiv \rho(\theta_0|\tilde{\theta}_0)} e^{-\theta_1 x_k} + \theta_2 e^{-(\theta_3 - x_k)^2}$$

Wahrscheinlichkeitsdichte für θ_0 bei gegebener Abschätzung $\tilde{\theta}_0$.

Mögliche wahre Werte für θ_0 (auszuintegrieren).

Gewählte Wahrscheinlichkeitsdichte
 $f(x, \mu = 1, \sigma = 0.05 \mu)$ für Normierung.



Beste Abschätzung des wahren Wertes in Likelihood

NB: Eine Diskussion der Log-Normalverteilung finden Sie im [backup](#) zu dieser Vorlesung.

Systematische Unsicherheiten in der Likelihood

- Wir führen einen Störparameter für die Normierung der Untergrundverteilung ein. Hierfür nehmen wir an, dass die Untergrund-Normierung θ_0 in unserem Modell nicht exakt bekannt ist, sondern nur mit einer rel. Unsicherheit von $\sigma(\theta_0)/\theta_0 = \pm 5\%$.

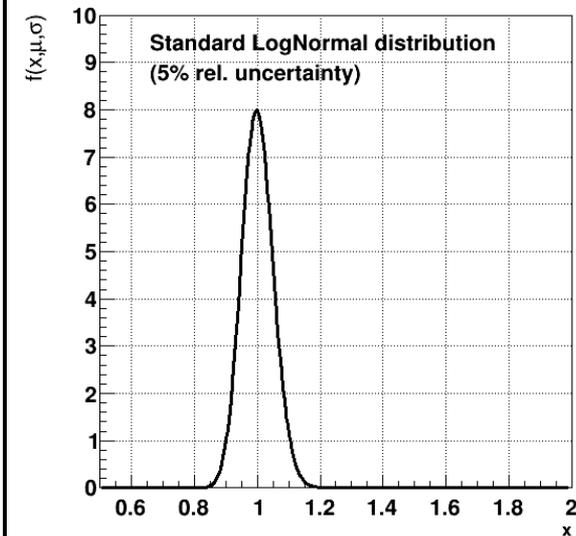
$$\nu_k(\{\theta_j\}) = \underbrace{\tilde{\theta}_0 f(\theta_0/\tilde{\theta}_0, 1, \sigma(\tilde{\theta}_0)/\tilde{\theta}_0)}_{\equiv \rho(\theta_0|\tilde{\theta}_0)} e^{-\theta_1 x_k} + \theta_2 e^{-(\theta_3 - x_k)^2}$$

Wahrscheinlichkeitsdichte für θ_0 bei gegebener Abschätzung $\tilde{\theta}_0$.

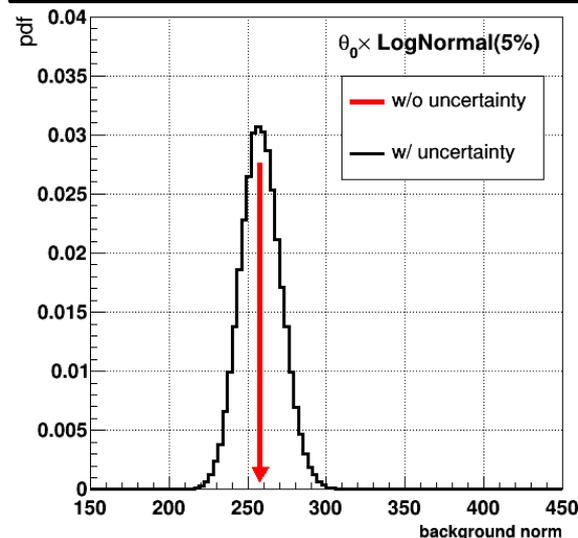
Mögliche wahre Werte für θ_0 (auszuintegrieren).

Beste Abschätzung des wahren Wertes in Likelihood

Gewählte Wahrscheinlichkeitsdichte $f(x, \mu = 1, \sigma = 0.05 \mu)$ für Normierung.



Gewählte Wahrscheinlichkeitsdichte $\rho(\theta_0|\tilde{\theta}_0)$ für θ_0 .



- Für $\rho(\theta_0|\tilde{\theta}_0)$ ist θ_0 der (wahre) Parameter in der Likelihood und $\tilde{\theta}_0$ der in der Likelihood angenommene Schätzwert.
- Bei der Bestimmung der Testverteilung wird der Parameter θ_0 ausintegriert (marginalisiert).
- Dies erfolgt i.d.R. mit Hilfe der MC Methode.

Systematische Unsicherheiten in der Likelihood

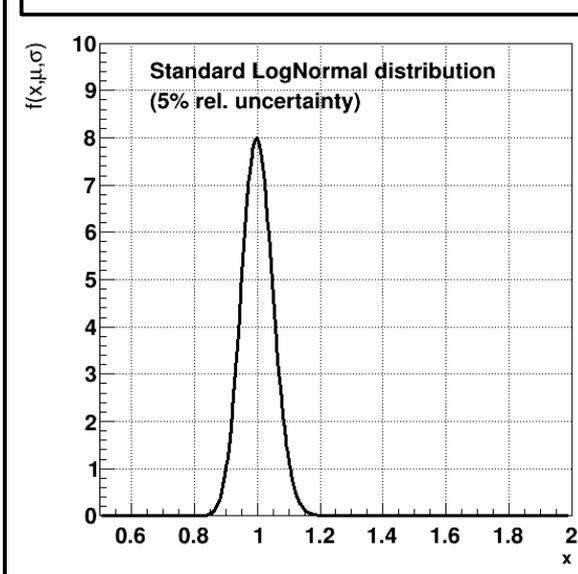
- Wir führen einen Störparameter für die Normierung der Untergrundverteilung ein. Hierfür nehmen wir an, dass die Untergrund-Normierung θ_0 in unserem Modell nicht exakt bekannt ist, sondern nur mit einer rel. Unsicherheit von $\sigma(\theta_0)/\theta_0 = \pm 5\%$.

$$\nu_k(\{\theta_j\}) = \underbrace{\tilde{\theta}_0 f(\theta_0/\tilde{\theta}_0, 1, \sigma(\tilde{\theta}_0)/\tilde{\theta}_0)}_{\equiv \rho(\theta_0|\tilde{\theta}_0)} e^{-\theta_1 x_k} + \theta_2 e^{-(\theta_3 - x_k)^2}$$

Wahrscheinlichkeitsdichte für θ_0 bei gegebener Abschätzung $\tilde{\theta}_0$.

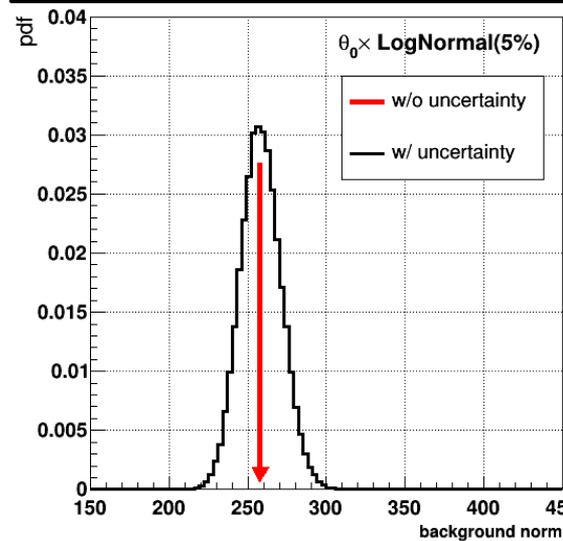
Mögliche wahre Werte für θ_0 (auszuintegrieren).

Gewählte Wahrscheinlichkeitsdichte $f(x, \mu = 1, \sigma = 0.05 \mu)$ für Normierung.

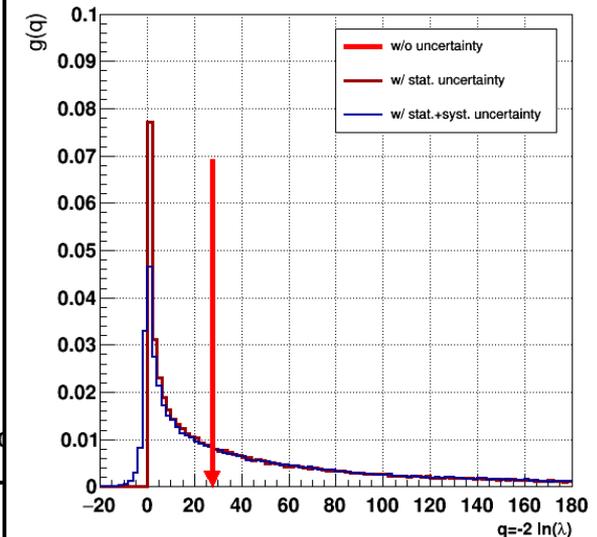


Beste Abschätzung des wahren Wertes in Likelihood

Gewählte Wahrscheinlichkeitsdichte $\rho(\theta_0|\tilde{\theta}_0)$ für θ_0 .



Die Testverteilung $g(q)$ weitet sich durch systematische Unsicherheiten erwartungsgemäß aus.



Ein lauffähiges Root macro finden Sie [hier](#).

Die LEP Teststatistik

- Eine exakt solche Teststatistik wurde (um 2000) zur Suche nach dem SM Higgs Boson bei LEP verwendet:

$$\mathcal{L}(\vec{x} | b(\{\theta_j\})) = \prod_k \mathcal{P}(\{x_k\} | b_k(\{\theta_j\}))$$

$$\mathcal{L}(\vec{x} | \mu s(\{\theta_j\}) + b(\{\theta_j\})) = \prod_k \mathcal{P}(\{x_k\} | \mu s_k(\{\theta_j\}) + b_k(\{\theta_j\}))$$

$$q_\mu = -2 \ln \left(\frac{\mathcal{L}(\vec{x} | \mu s(\{\theta_j\}) + b(\{\theta_j\}))}{\mathcal{L}(\vec{x} | b(\{\theta_j\}))} \right), \quad 0 \leq \mu$$

- Der Likelihood-Quotient besteht aus H_1 : "SM" über H_0 : "Es gibt kein Higgs Boson im SM". Alle Unsicherheiten werden durch Störparameter in die Likelihood eingeführt und wie auf den vorigen Folien beschrieben marginalisiert.

Die LEP Teststatistik

- Eine exakt solche Teststatistik wurde (um 2000) zur Suche nach dem SM Higgs Boson bei LEP verwendet:

$$\mathcal{L}(\vec{x} | b(\{\theta_j\})) = \prod_k \mathcal{P}(\{x_k\} | b_k(\{\theta_j\}))$$

$$\mathcal{L}(\vec{x} | \mu s(\{\theta_j\}) + b(\{\theta_j\})) = \prod_k \mathcal{P}(\{x_k\} | \mu s_k(\{\theta_j\}) + b_k(\{\theta_j\}))$$

$$q_\mu = -2 \ln \left(\frac{\mathcal{L}(\vec{x} | \mu s(\{\theta_j\}) + b(\{\theta_j\}))}{\mathcal{L}(\vec{x} | b(\{\theta_j\}))} \right), \quad 0 \leq \mu$$

- Der Likelihood-Quotient besteht aus H_1 : "SM" über H_0 : "Es gibt kein Higgs Boson im SM". Alle Unsicherheiten werden durch Störparameter in die Likelihood eingeführt und wie auf den vorigen Folien beschrieben marginalisiert.
- **NB**: Dieser Ansatz ist in seinem Grundsatz bayesianischer Natur. Die Wahrscheinlichkeitsdichten $\rho(\theta_j | \tilde{\theta}_j)$ für die Störparameter θ_j entsprechen Wahrscheinlichkeitsdichten für den „wahren“ Wert des Parameters. In der frequentistischen Interpretation gibt es nur den wahren Wert, für den es keine Wahrscheinlichkeitsdichte gibt. Dieser Ansatz wird daher auch als frequentistisch-bayesianischer Hybridansatz bezeichnet.

Die Tevatron Teststatistik

- Bei der Suche nach dem SM Higgs Boson am Tevatron (um 2005) wurde die LEP Teststatistik angesichts großer systematischer Unsicherheiten modifiziert:

$$\mathcal{L}(\vec{x} | b(\{\theta_j\})) = \prod_k \mathcal{P}(\{x_k\} | b_k(\{\theta_j\})) \cdot \prod_j \rho(\theta_j | \tilde{\theta}_j)$$

$$\mathcal{L}(\vec{x} | \mu s(\{\theta_j\}) + b(\{\theta_j\})) = \prod_k \mathcal{P}(\{x_k\} | \mu s(\{\theta_j\}) + b_k(\{\theta_j\})) \cdot \prod_j \rho(\theta_j | \tilde{\theta}_j)$$

$$q_\mu = -2 \ln \left(\frac{\mathcal{L}(\vec{x} | \mu s(\{\hat{\theta}_j\}) + b(\{\hat{\theta}_j\}))}{\mathcal{L}(\vec{x} | b(\{\hat{\theta}_j\}))} \right), \quad 0 \leq \mu$$

- Der Hybridansatz zur Marginalisierung der Störparameter wurde beibehalten aber die Poisson-Likelihoods in Zähler und Nenner einfach um die Wahrscheinlichkeitsdichten der Störparameter erweitert.
- Die Störparameter der Likelihoods für Zähler und Nenner wurden dann vor Evaluation der Teststatistik mit Hilfe der ML Methode abgeschätzt. Diese Abschätzung erlaubt es die Unsicherheiten auf die Störparameter (post-fit) zu reduzieren und sogar Störparameter als freie Parameter abzuschätzen, die dann ihre Unsicherheit erst aus der Abschätzung erhalten.

Anmerkungen...

... zu den bisher diskutierten Teststatistiken:

- Durch die Marginalisierung der Störparameter ist das Theorem von Wilks nicht mehr anwendbar. Dies führt dazu, dass in beiden Fällen die Testverteilung von q_μ aufwändig mit Hilfe von MC Methoden bestimmt werden muss.
- Die Behandlung der Störparameter in der Tevatron Teststatistik ist zu der Behandlung der Störparameter in der LEP Teststatistik mathematisch nicht äquivalent! Beide erfassen jedoch den Umstand der Anwesenheit von Messunsicherheiten, was zu einer Verbreiterung der Testverteilung führt.
- Mir persönlich erscheint die LEP Methode intuitiver und auch „richtiger“, die Tevatron Methode praktikabler und daher einfacher zu realisieren.¹ Im Zweifel entscheidet die Eigenschaft der Abdeckung (*coverage*) über Anwendbarkeit und Erfolg eines Modells. Wir werden diese Eigenschaft in der nächsten Vorlesung diskutieren.

¹ Ich finde, dass es z.B. ein ganz gutes Thema einer Bachelorarbeit sein könnte, die Unterschiede beider Methoden im Detail zu verstehen.

Die LHC Teststatistik

- Mit der Suche nach dem SM Higgs Boson am LHC (ab 2011) ist man dazu übergegangen eine *Profile-Likelihood* Teststatistik zu verwenden:

$$\mathcal{L}(\vec{x} | \mu s(\{\theta_j\}) + b(\{\theta_j\})) = \prod_k \mathcal{P}(\{x_k\} | \mu s(\{\theta_j\}) + b_k(\{\theta_j\})) \cdot \prod_j \rho(\tilde{\theta}_j | \theta_j)$$

$$q_\mu = -2 \ln \left(\frac{\mathcal{L}(\vec{x} | \mu s(\{\hat{\theta}_{\mu,j}\}) + b(\{\hat{\theta}_{\mu,j}\}))}{\mathcal{L}(\vec{x} | \hat{\mu} s(\{\hat{\theta}_j\}) + b(\{\hat{\theta}_j\}))} \right), \quad 0 \leq \hat{\mu} \leq \mu$$

- Dabei werden im Nenner alle freien Parameter (d.h. sowohl $\{\theta_j\}$ als auch μ) global abgeschätzt. Im Zähler werden die $\{\theta_j\}$ bei vorgegebenem μ abgeschätzt. Zähler und Nenner sind so kontinuierlich ineinander überführbar.
- Diese Teststatistik hat den Vorteil, dass Wilks' Theorem hier wieder anwenbar ist, d.h. die Testverteilung läßt sich asymptotisch ohne Zuhilfenahme von MC Methoden analytisch angeben. Sie finden mehr Details hierzu unter [EPJ C 71 \(2011\) 1554](#).

Die LHC Teststatistik

- Zudem werden die Wahrscheinlichkeitsdichten der Störparameter $\rho(\theta_j|\tilde{\theta}_j)$ unter Verwendung des Satzes von Bayes (implizit) als Posteriorwahrscheinlichkeiten für den Ausgang virtueller oder realer Messungen bei vorgegebener Wahrheit $\rho(\tilde{\theta}_j|\theta_j)$ neu interpretiert.

$$\rho(\tilde{\theta}_j|\theta_j) \propto \rho(\theta_j|\tilde{\theta}_j) \pi_{\theta_j}(\theta_j)$$

- Bei flachem prior $\pi_{\theta_j}(\theta_j)$ impliziert dies z.B. bei der Normal- und Log-Normalverteilung lediglich die „Vertauschung“ der Variablen θ_j und $\tilde{\theta}_j$.

$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\theta_j - \tilde{\theta}_j)^2\right) \longrightarrow \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\tilde{\theta}_j - \theta_j)^2\right)$$

- Durch diese konzeptionell unterschiedliche Interpretation wird die bayesianische Wahrscheinlichkeitsdichte $\rho(\theta_j|\tilde{\theta}_j)$ für den wahren Wert bei vorgegebenem Schätzwert zur einer rein frequentistisch interpretierbaren Wahrscheinlichkeitsdichte für den Ausgang einer (virtuellen/angenommenen) Messung bei vorgegebenem, wahren Wert $\rho(\tilde{\theta}_j|\theta_j)$.
- Auf der technischen Seite spart man sich die Marginalisierung der Störparameter.
- Für eine detailliertere Lektüre siehe [CMS NOTE-2011/005](#).

Testverteilung des Profile-Likelihood Quotienten

- Im Grenzwert großer Stichproben folgt der Profile-Likelihood Quotient einer nicht-zentralen χ^2 -Verteilung mit einem Freiheitsgrad:

$$\chi^2(q_\mu, \Lambda, r = 1) = \frac{1}{2\sqrt{q_\mu}} \frac{1}{\sqrt{2\pi}} \left(\exp\left(-\frac{1}{2} \left(\sqrt{q_\mu} + \sqrt{\Lambda}\right)^2\right) + \exp\left(-\frac{1}{2} \left(\sqrt{q_\mu} - \sqrt{\Lambda}\right)^2\right) \right)$$

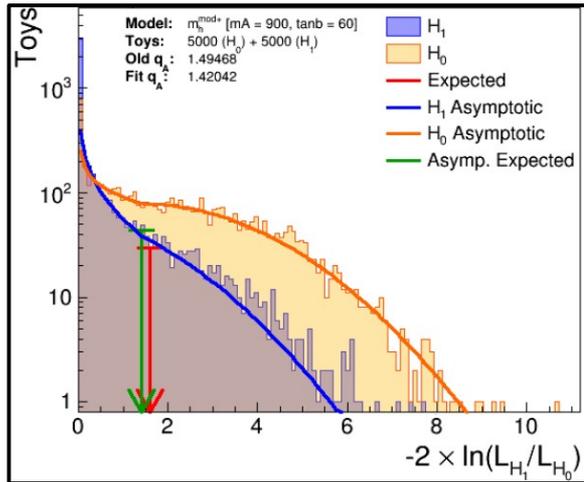
mit:

$$\Lambda = \frac{(\mu - \mu')^2}{\sigma^2}$$

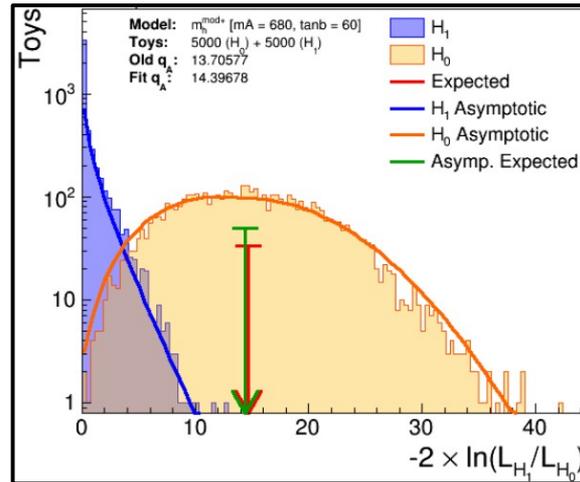
- Dabei ist μ der Parameter, der von q_μ getestet wird und μ' die Variable.
- Für $\mu = \mu'$ geht die nicht-zentrale χ^2 -Verteilung in eine normale χ^2 -Verteilung über.

Testverteilung des Profile-Likelihood Quotienten

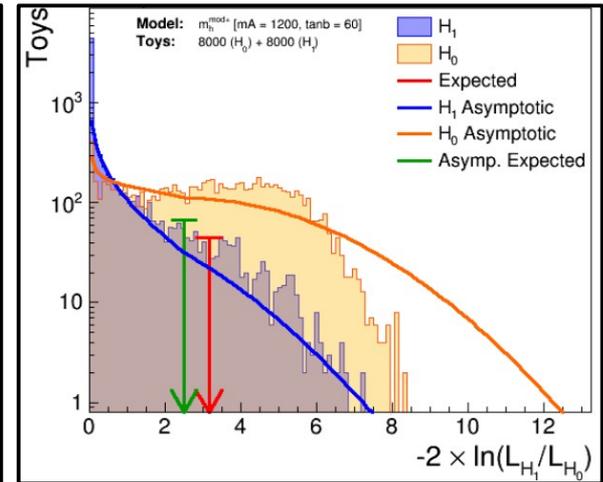
- Im folgenden sind einige Beispiele für verschiedene H_0 - und H_1 -Hypothesen gezeigt. Für mehr Details siehe [Masterarbeit Dennis Roy](#).



Asymptotische Formeln folgen
Verteilung aus der MC Methode.



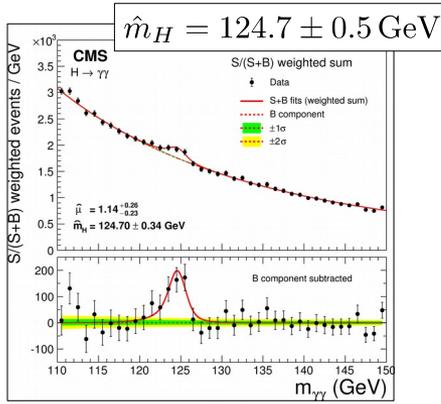
Asymptotische Formeln folgen
Verteilung aus der MC Methode.



Asymptotische Formeln folgen
Verteilung aus MC Methode nicht.

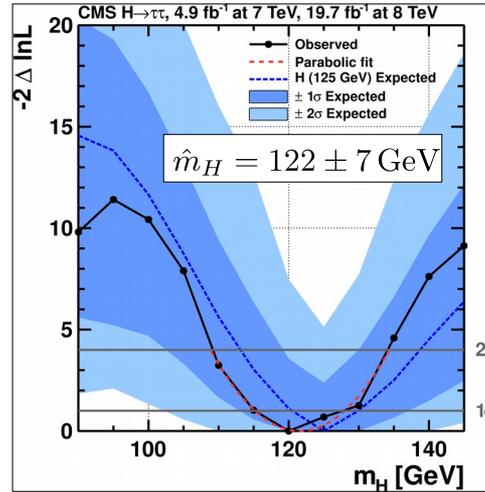
Kompatibilität: Higgs@CMS

EPJ C 74 (2014) 3076

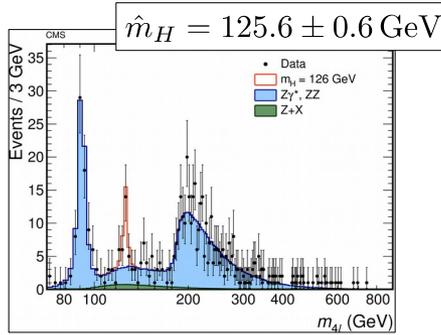


Gezeigt sind hier die Run-1 Ergebnisse von CMS.

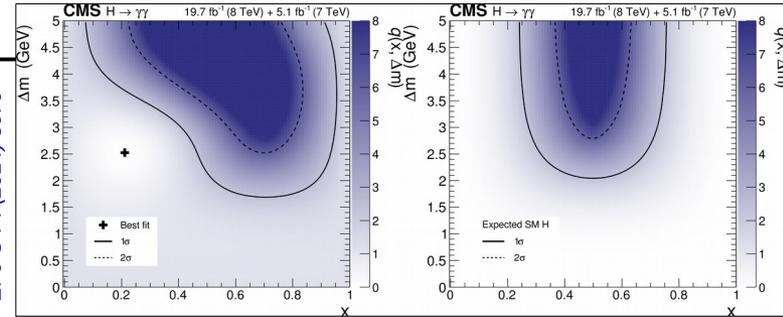
JHEP 05 (2014) 104



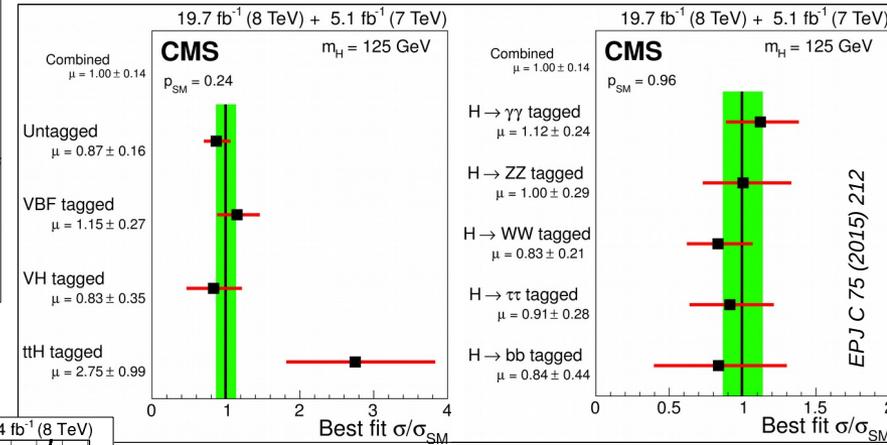
PRD 89 (2014) 092007



EPJ C 74 (2014) 3076

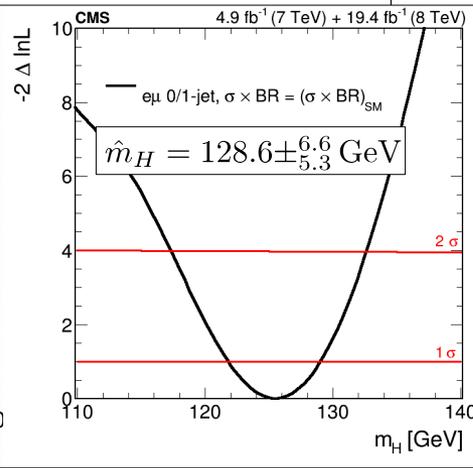
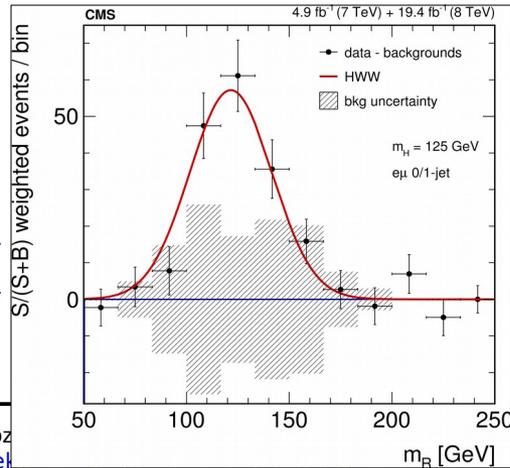


Coupling across production modes or decay channels:



EPJ C 75 (2015) 212

JHEP 01 (2014) 096



EPJ C 75 (2015) 212

- Event categories : 227
 - Nuisance parameters: $\mathcal{O}(2500)$
 - 16 MB binary file of stat. model (~145 MB in human readable form).
- $\mu = \sigma/\sigma_{SM} = 1.00 \pm 0.14$
- $p\text{-value} = 84\%$

Kopplungsstruktur

- Event categories : 574
 - Nuisance parameters: 4268
- $\mu = \sigma/\sigma_{SM} = 1.09 \pm 0.11$

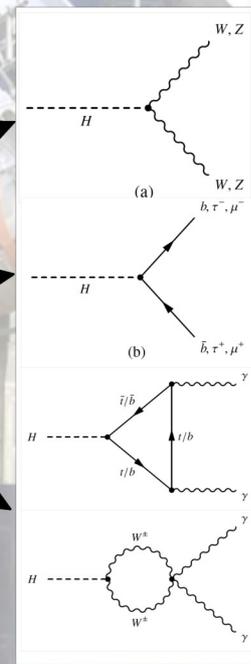
• ATLAS+CMS LHC Run-1 combination:

Considered production modes:

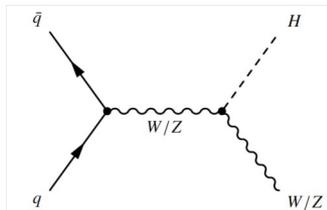
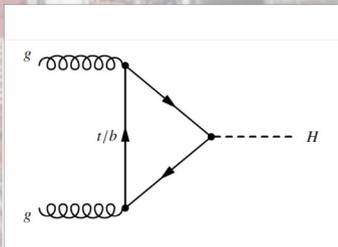
Production process	Cross section [pb]		Order of calculation
	$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$	
ggF	15.0 ± 1.6	19.2 ± 2.0	NNLO(QCD)+NLO(EW)
VBF	1.22 ± 0.03	1.58 ± 0.04	NLO(QCD+EW)+~NNLO(QCD)
WH	0.577 ± 0.016	0.703 ± 0.018	NNLO(QCD)+NLO(EW)
ZH	0.334 ± 0.013	0.414 ± 0.016	NNLO(QCD)+NLO(EW)
[ggZH]	0.023 ± 0.007	0.032 ± 0.010	NLO(QCD)
bbH	0.156 ± 0.021	0.203 ± 0.028	5FS NNLO(QCD) + 4FS NLO(QCD)
ttH	0.086 ± 0.009	0.129 ± 0.014	NLO(QCD)
tH	0.012 ± 0.001	0.018 ± 0.001	NLO(QCD)
Total	17.4 ± 1.6	22.3 ± 2.0	

Considered decay channels:

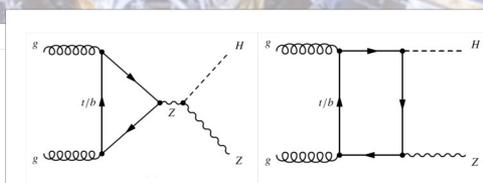
Decay channel	Branching ratio [%]
$H \rightarrow bb$	57.5 ± 1.9
$H \rightarrow WW$	21.6 ± 0.9
$H \rightarrow gg$	8.56 ± 0.86
$H \rightarrow \tau\tau$	6.30 ± 0.36
$H \rightarrow cc$	2.90 ± 0.35
$H \rightarrow ZZ$	2.67 ± 0.11
$H \rightarrow \gamma\gamma$	0.228 ± 0.011
$H \rightarrow Z\gamma$	0.155 ± 0.014
$H \rightarrow \mu\mu$	0.022 ± 0.001



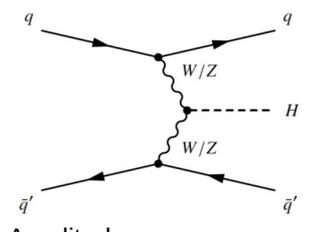
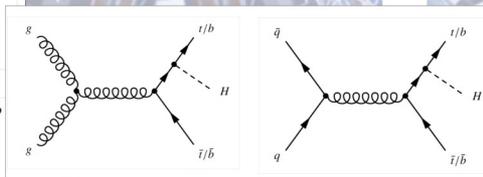
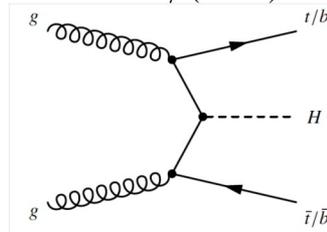
Main production modes:



$gg \rightarrow ZH$ (10% to $ZHbb$)

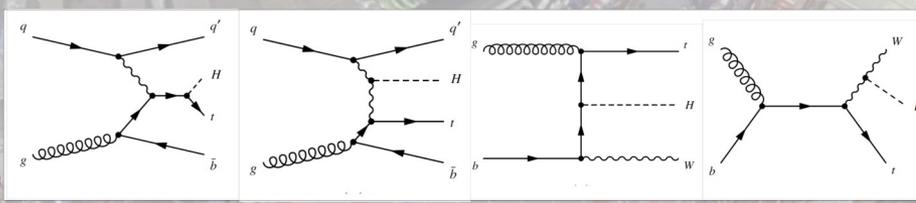


$t\bar{t}H / (b\bar{b}H)$



Amplitudes

$tqH + tHW$



Zusammenfassung

- Effizienz und Reinheit
- ROC Kurve
- Poisson-basierter Zweihypothesen Test:
 - Lemma von Neyman-Pearson
 - Theorem von Wilks
- Systematische Unsicherheiten in der Likelihood
 - LEP Teststatistik
 - Tevatron Teststatistik
 - LHC Teststatistik
- Beispiel: Eigenschaften des Higgs Bosons

Backup

Erinnerung: Log-Normalverteilung

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{x} e^{-(\ln(x)-\mu)^2/2\sigma^2}$$

$$E[x] = e^{\mu+1/2\sigma^2} \quad (\text{Erwartungswert})$$

$$\text{var}[x] = e^{2\mu+\sigma^2} \cdot (e^{\sigma^2} - 1) \quad (\text{Varianz})$$

- Variablensubstitution:

$$x' = \ln(x) \quad dx' = d \ln(x) = \frac{dx}{x}$$

führt die Log-Normalverteilung in Normalverteilung über:

$$f(x, \mu, \sigma) dx = \varphi(x', \mu, \sigma) dx' \Big|_{x'=\ln(x)}$$

- **NB:** So wie die Summe einer großen Zahl von Zufallsvariablen unabhängig von ihrer Verteilungsfunktion einer Normalverteilung zustrebt, strebt das Produkt aus einer großen Zahl von Zufallsvariablen unabhängig von ihrer Verteilungsfunktion einer Log-Normalverteilung zu.

