Accessing CKM suppressed top decays at the LHC

PIKIMO 13, 11/12/2022

Manuel Szewc, University of Cincinnati

University of CINCINNATI

This talk

Based on a recent preprint arxiv:2209.01222 by D. Faroughy, J.F. Kamenik, M.S. and J. Zupan

A **simple** extension of an existing strategy to measure $|V_{td}|^2+|V_{ts}|^2$ at the LHC

Main idea: Orthogonal *b*- and *q*-taggers define **complementary observables** that **increase the statistical power** of the analysis

This simple extension allows to **measure a non-null $|V_{td}|^2+|V_{ts}|^2$ at 95% CL at the HL-LHC**

The CKM matrix

The CKM matrix encodes the flavor structure of the EW charged currents in the SM. Measuring the CKM → Very powerful precision tests of the SM (with potential sensitivity to BSM effects)

Specially true for its **third row** which is **indirectly** constrained by B-physics measurements. Global CKM fits yield estimates assuming unitarity

A strong **hierarchy** makes **off-diagonal elements** difficult to measure but also more sensitive to any BSM (also due to the large top mass)

$$|V_{tb}^{\mathrm{SM}}| = 999.118^{+0.031}_{-0.036} \times 10^{-3}$$

$$|V_{ts}^{\mathrm{SM}}| = 41.10^{+0.83}_{-0.72} \times 10^{-3}$$

$$|V_{td}^{\mathrm{SM}}| = 8.57^{+0.2}_{-0.18} \times 10^{-3}$$

Direct measurements

$|V_{tx}|$ can be obtained directly from on-shell top-quarks decays.

CMS coll. arxiv:1404.2292, provides the current best measurement using top pair production by measuring

$$\mathcal{R}_b \equiv \frac{\mathcal{B}(t \to bW)}{\sum_{j=d,s,b} \mathcal{B}(t \to jW)} > 0.955 \text{ @ } 95\% \text{ C.L.}$$

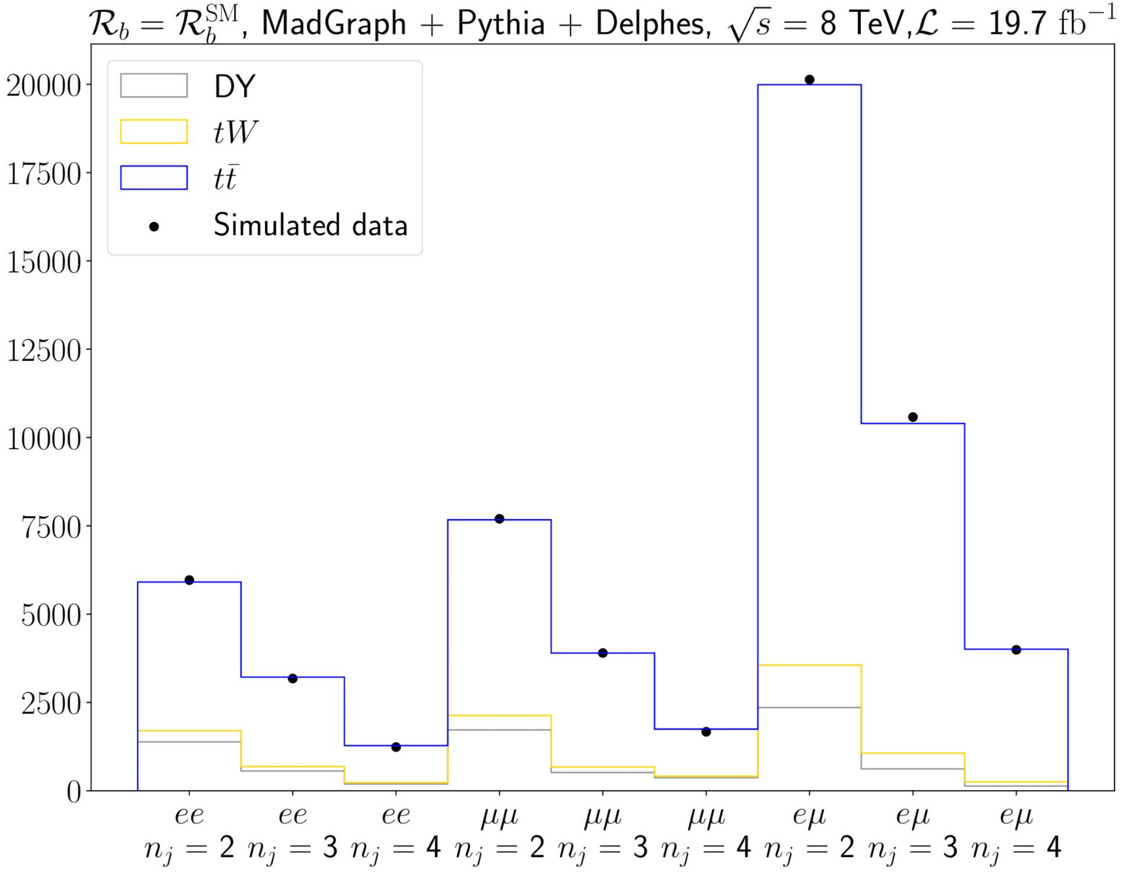
which is translated to

$$\sqrt{|V_{td}|^2 + |V_{ts}|^2} < 0.217 |V_{tb}|$$

The starting point

This is the population for each number of jets n_j and dileptonic channel $\ell\ell'$

We obtained the expected values with Monte Carlo and sample in each category using a Poisson distribution



The CMS analysis

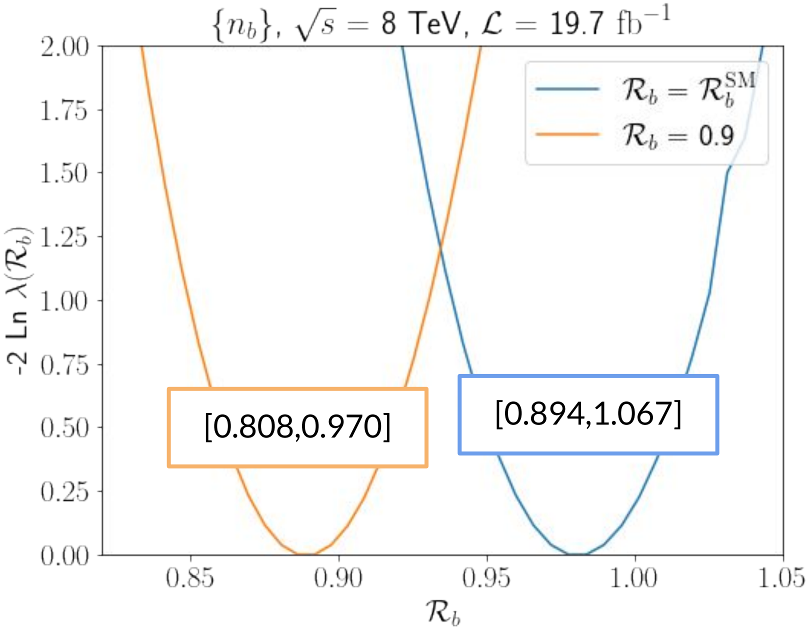
Model the expected number of events for each possible number of b-tags in a $\{n_j, \ell\ell'\}$ category

$$\bar{N}_{\ell\ell'}(n_b|n_j) = P_{\ell\ell'}(n_b|n_j, \mathcal{R}_b, \theta_i)N_{\ell\ell}$$

Set limits using

$$\mathcal{L}(\mathcal{R}_b, \theta_i) = \prod_{\ell\ell'} \prod_{n_j=2}^{4} \prod_{n_b=0}^{n_j} \prod_{n_q=0}^{n_j-n_b} \mathcal{P}(N_{\ell\ell'}|\bar{N}_{\ell\ell'}) \prod_i \rho(\theta_i)$$

We are consistent with the CMS results with our set-up

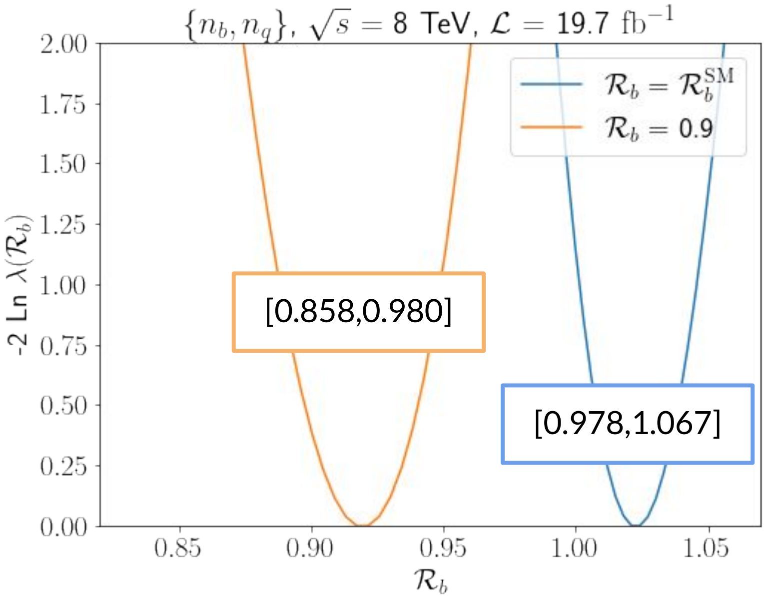


Our extension

We already see an improvement! Smaller CIs and larger discrimination between Rs

Let's incorporate more information

In addition to the number of *b*-tagged jets, we compute the number of *q*-tagged jets, creating non-overlapping $\{n_b, n_q\}$ bins



$$\bar{N}_{\ell\ell'}(n_b|n_j) \rightarrow \bar{N}_{\ell\ell'}(n_b, n_q|n_j)$$

$$\bar{N}_{\ell\ell'}(n_b, n_q|n_j) = P_{\ell\ell'}(n_b, n_q|n_j, \mathcal{R}_b, \theta_i) N_{\ell\ell'}(n_j)$$

Projected sensitivity

We validate the model through consistency checks using MC and project its performance **assuming it is true**

We sample $\{n_b, n_q, n_j, \ell\ell'\}$ counts using the simulated $\{n_j, \ell\ell'\}$ event yields and specific choices for the parameters

We then re-do all our statistical analyses using this pseudo-data

Two large speed-ups: we can incorporate any tagger we want without coding it at the ROOT level and we can explore different choices of R_b and other parameters

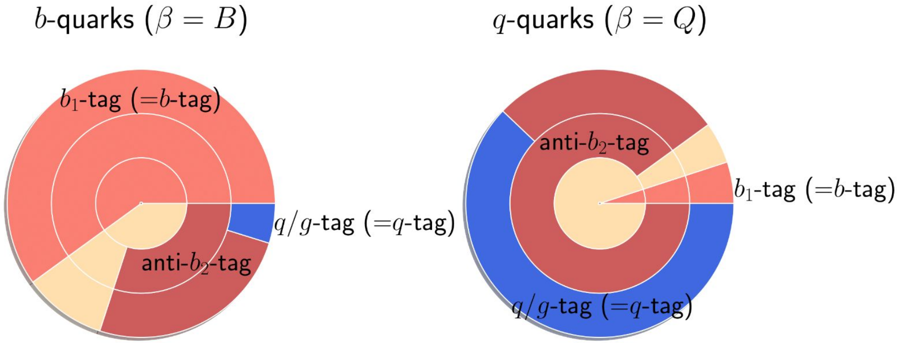
Tagging strategy

We propose to mix state-of-the-art *b*-taggers and quark/gluon taggers to define complementary orthogonal regions.

We choose two Working Points of the former to define a b_1-tagger and an anti-b_2-tagger. This anti-b_2-tagger is combined with the quark/gluon tagger to define a *q*-tagger.

$$\epsilon_\beta^{b_1} \leq \epsilon_\beta^{b_2}$$

$$\epsilon_\beta^{q} = \epsilon_\beta^{\{\mathrm{anti-b_2}\}\cap\{\mathrm{q/g}\}}$$



Reframing measurement as discovery

To compare Working Points in a simple way, we reframe the problem as a signal discovery.

Null hypothesis R_b = 1 vs
Alternative hypothesis R_b = SM

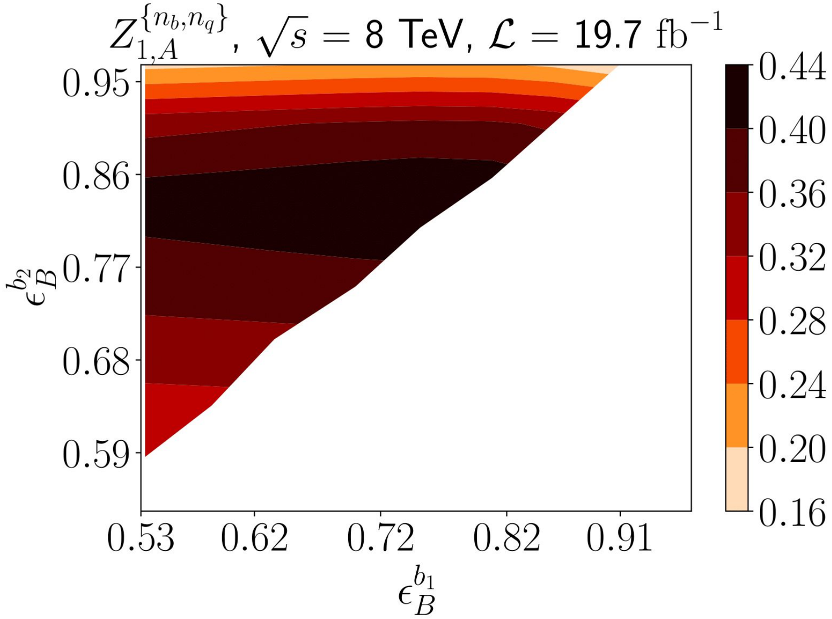
New test statistic + Asimov dataset → **expected discovery significance of the SM hypothesis**

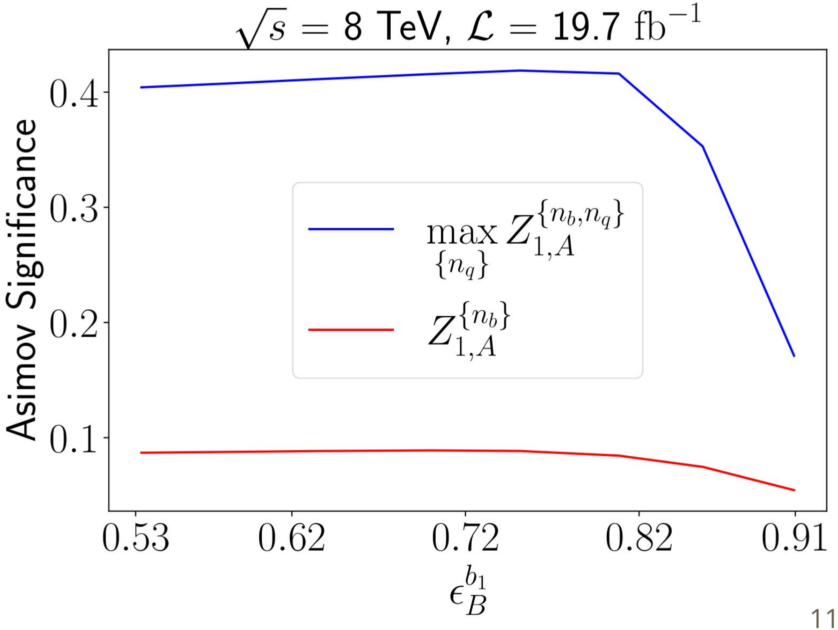
$$\tilde{\lambda}(\mathcal{R}_b) = \begin{cases} \mathcal{L}(\mathcal{R}_b, \hat{\hat{\theta}}_i(\mathcal{R}_b))/\mathcal{L}(\hat{\mathcal{R}}_b, \hat{\theta}_i), & \text{if } \hat{\mathcal{R}}_b \leq 1\,, \\ \mathcal{L}(\mathcal{R}_b, \hat{\hat{\theta}}_i(\mathcal{R}_b))/\mathcal{L}(1, \hat{\hat{\theta}}_i(1)), & \text{if } \hat{\mathcal{R}}_b > 1\,. \end{cases}$$

$$q_1 = \begin{cases} -2 \text{ Ln } \lambda(1) & \text{if } \hat{\mathcal{R}}_b \leq 1\,, \\ 0 & \text{if } \hat{\mathcal{R}}_b > 1\,. \end{cases}$$

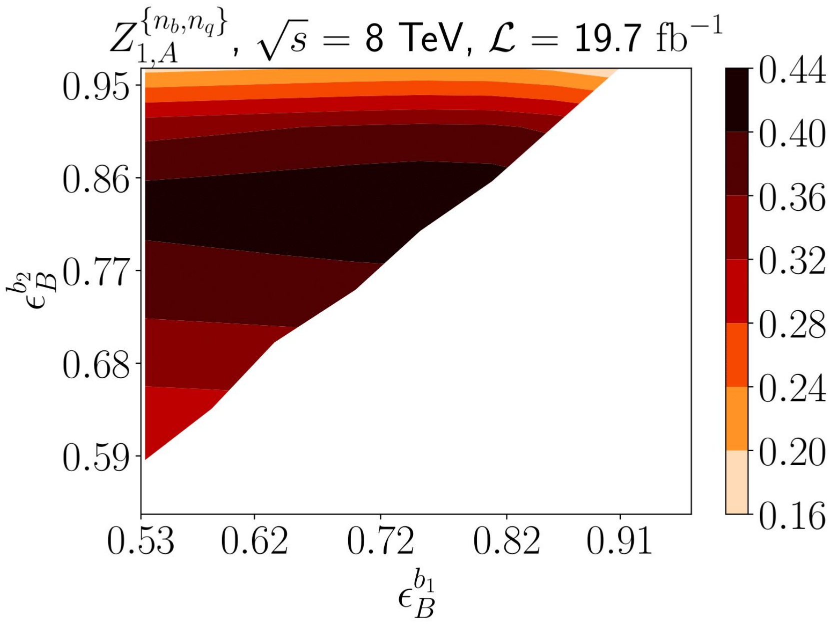
$$Z_{1,A} = \sqrt{q_{1,A}}$$

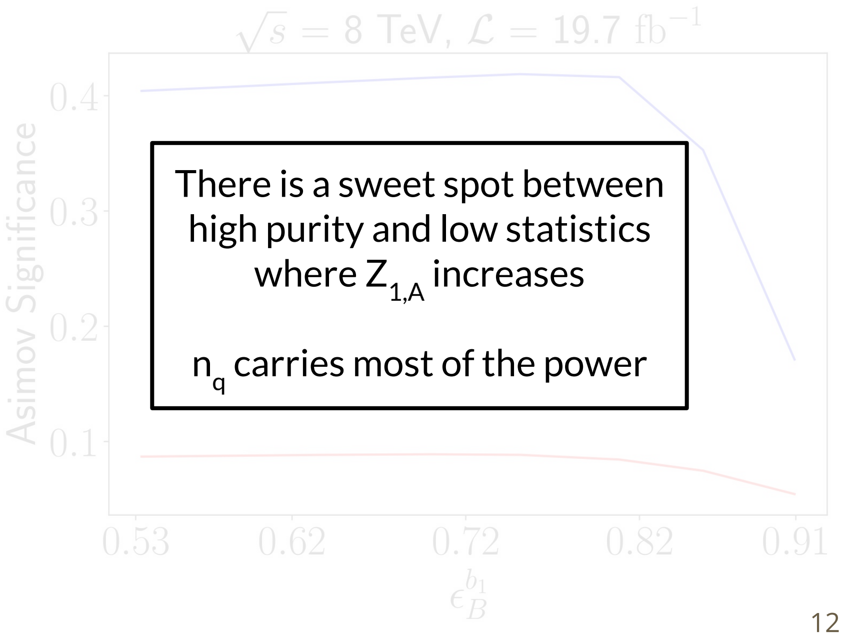
Results: original benchmark





Results: original benchmark

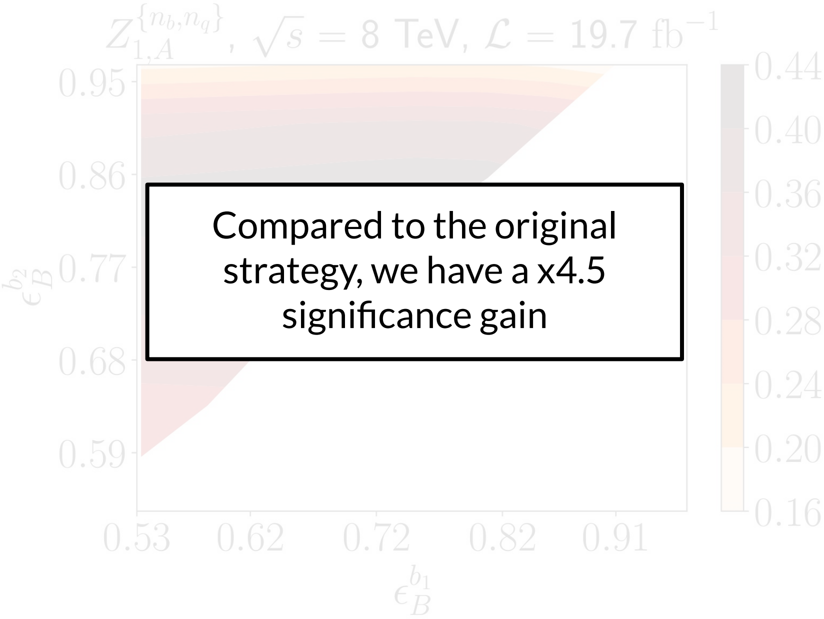




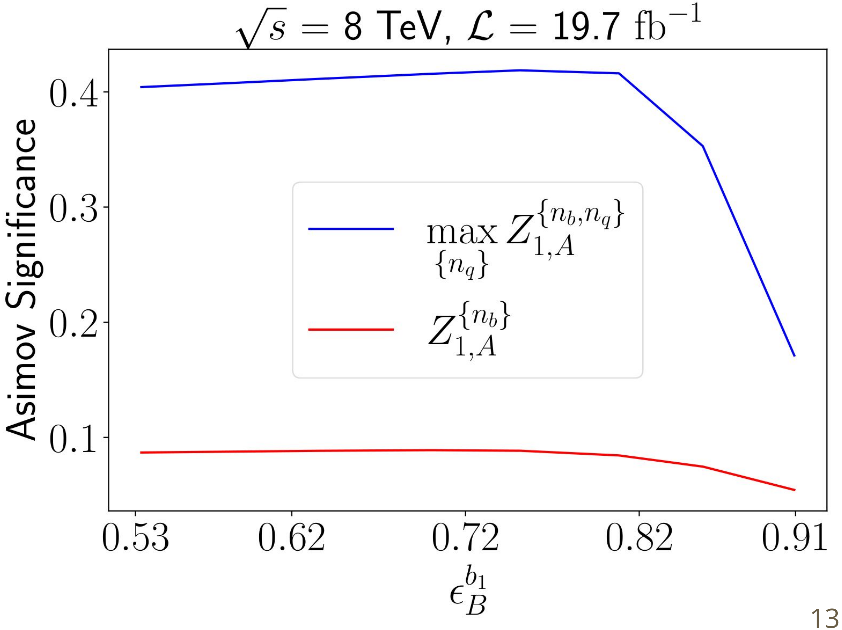
There is a sweet spot between high purity and low statistics where $Z_{1,A}$ increases

n_q carries most of the power

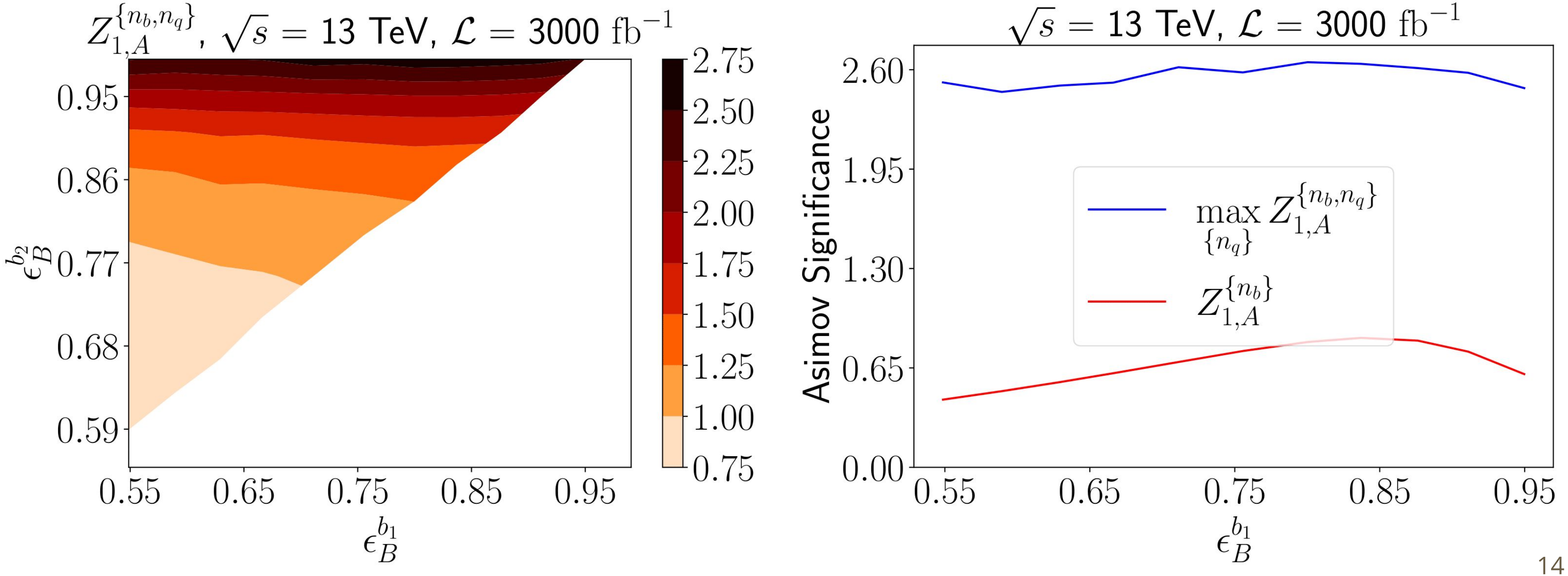
Results: original benchmark



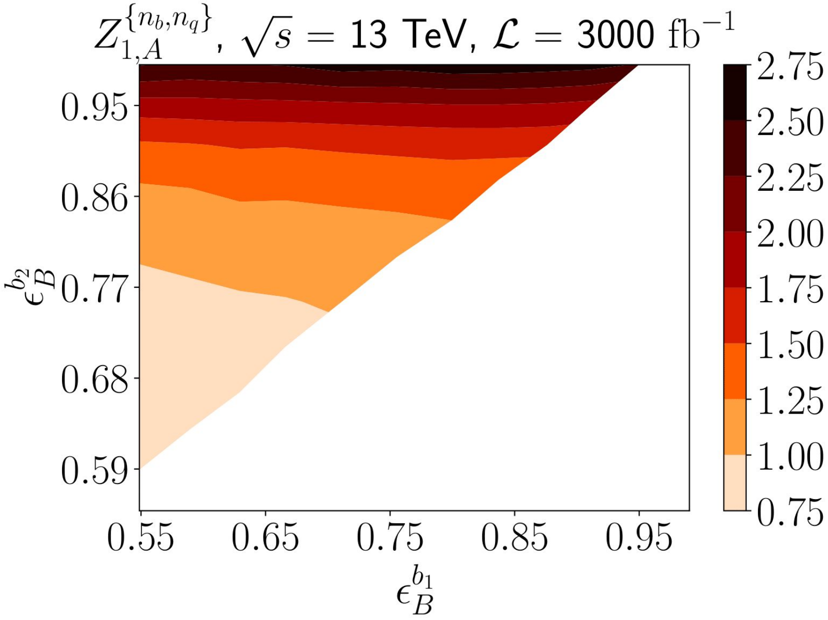
Compared to the original strategy, we have a x4.5 significance gain

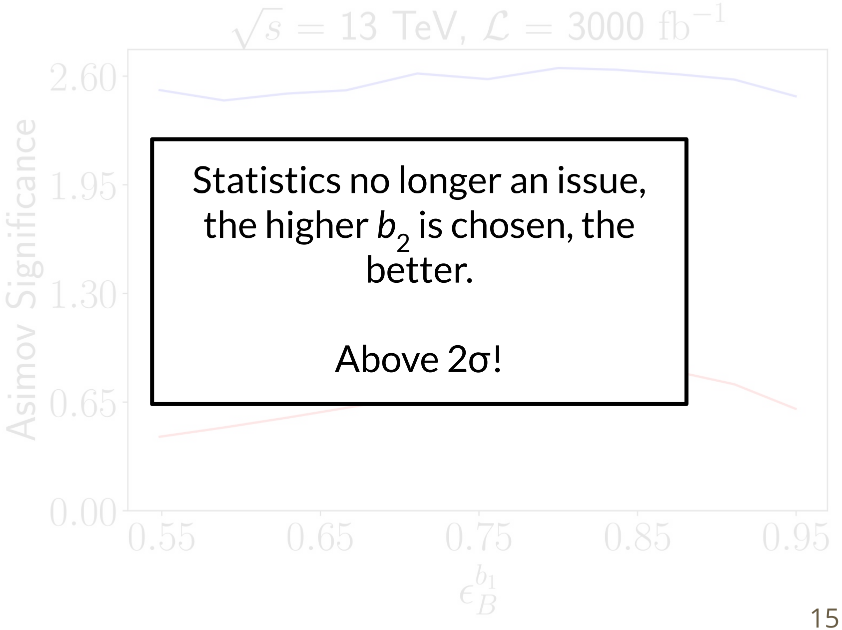


Results: ~HL-LHC projection



Results: ~HL-LHC projection

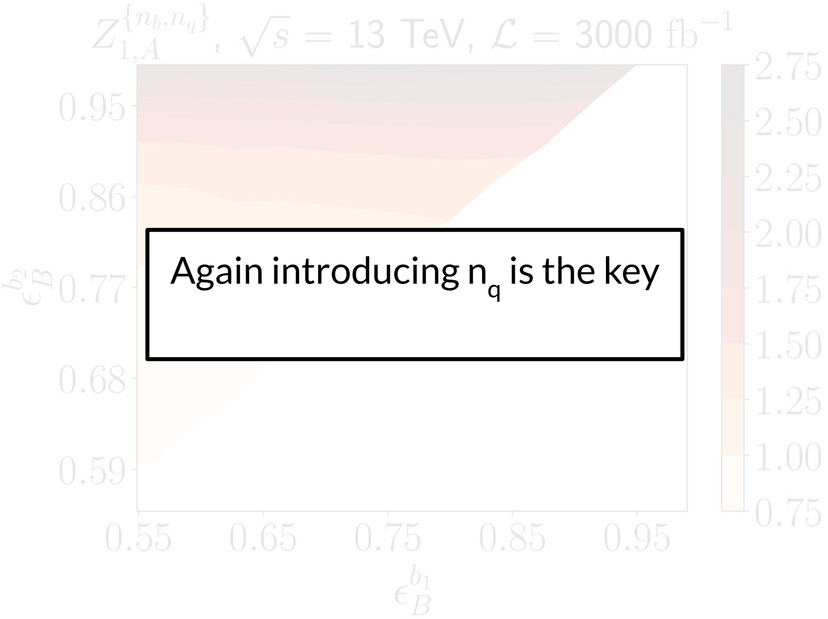




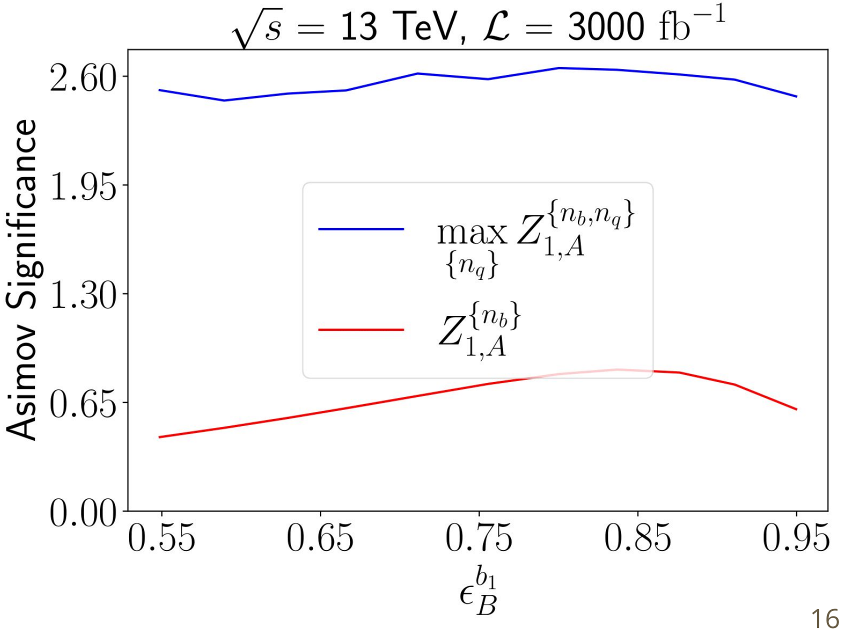
Statistics no longer an issue, the higher b_2 is chosen, the better.

Above 2σ!

Results: ~HL-LHC projection



Again introducing n_q is the key



Outlook

We have extended a previous analysis strategy to incorporate additional information in the form of the number of *q*-tagged jets n_q. We have verified that the probabilistic model captures dileptonic top pair production events appropriately

The proposed strategy allows to measure non-null $|V_{tx}|$ at the 1σ level during Run 2 and at the 2σ level at the HL-LHC

However the probabilistic model is incomplete and additional NPs should be included. It could also be extended to be more physical. For example, the model could incorporate jet kinematics

We have treated tagger efficiency estimation and R_b determination as separate problems. However, they are related and could be treated at the same time


Thank you!

Backup slides

What about other direct measurements?

All other approaches (t-channel single top production, tW associated production, *s*-tagging the top decay products) suffer from low statistics due to the smallness of $|V_{td}|^2+|V_{ts}|^2$ and cannot match the CKM global fits

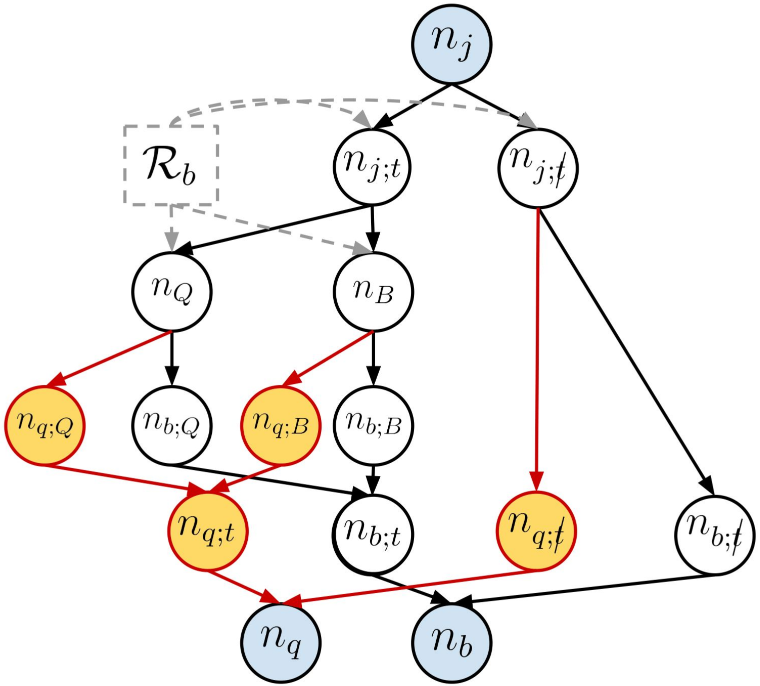
The CMS analysis does not suffer so badly from low statistics, as it considers dileptonic top pair production with no tag requirement on the jets but it lacks discriminating power between SM values and $R_b = 1$

Explicitly

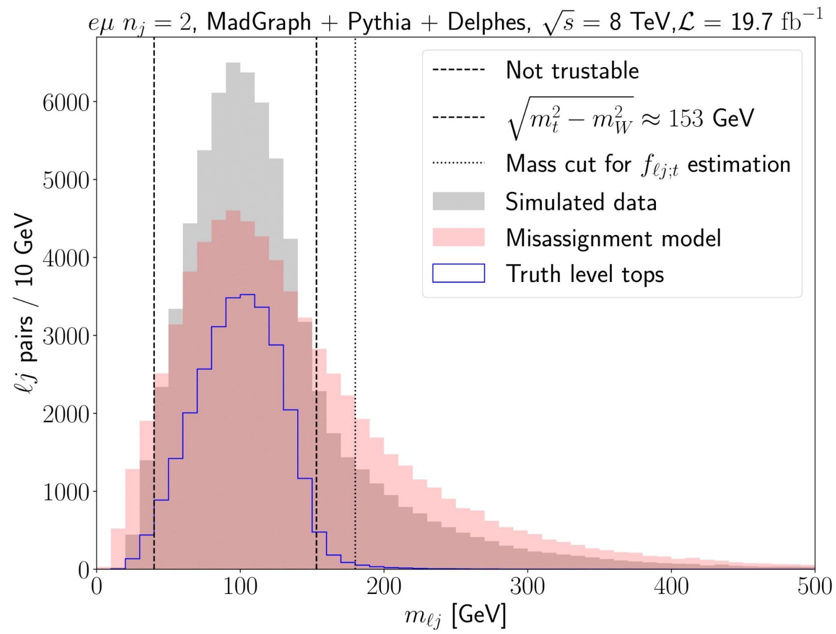
This is the probabilistic model we propose for each ℓℓ' combination

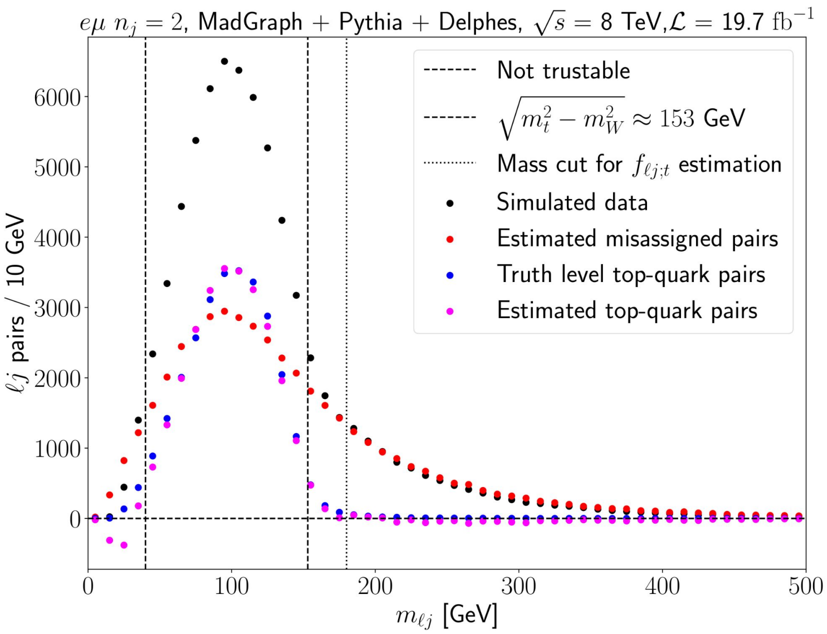
Not shown are some nuisance parameters which need to be fitted from data and determine the probability distributions

Blue is observed, yellow and red are the new additions



Top fraction determination





Model evaluation

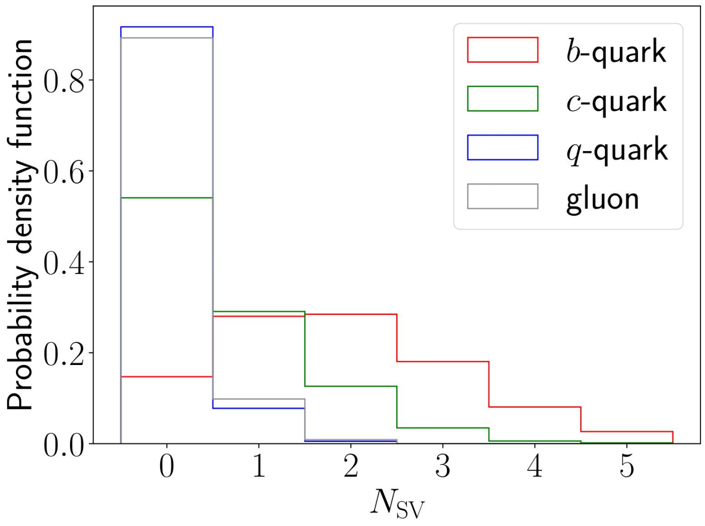
We first want to assess how good the probabilistic model is at capturing the true physics

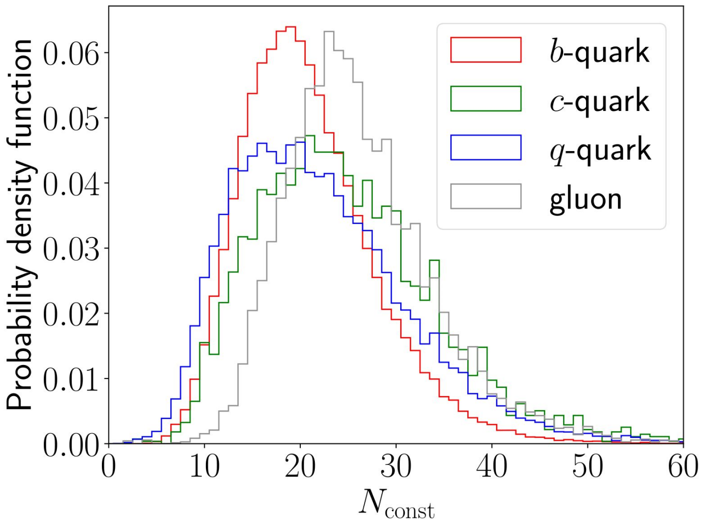
We know it is approximate (all jets have the same efficiencies regardless of kinematics!) but we want to know if it is expressive enough

We build *b*- and *q*-taggers, apply them to Monte Carlo data at the ROOT level and check if the fitted model presents good coverage around the true R_b value

As a bonus, we obtain a first example of how n_q improves the determination of R_b

Homemade *b*- and *q*-taggers





Homemade *b*- and *q*-taggers

High sample purity for n_b, but intermediate for n_q

α-tagger	Cuts	ϵ_Q^α	ϵ_B^α
b-tagger	$N_{\rm SV} \geq 2$	$0.002^{+10\%}_{-10\%}$	$0.49^{+5\%}_{-5\%}$
q-tagger	$N_{\rm SV} = 0$ & $N_{\rm const} < 20$	$0.69^{+20\%}_{-20\%}\%$	$0.16^{+10\%}_{-10\%}$

Defining a test statistic

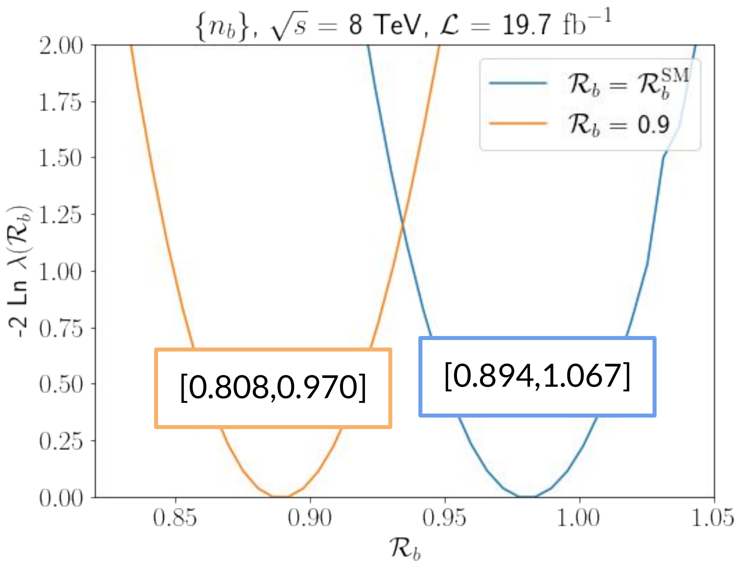
To measure a parameter in presence of systematic uncertainties, we define the Profile Likelihood Ratio

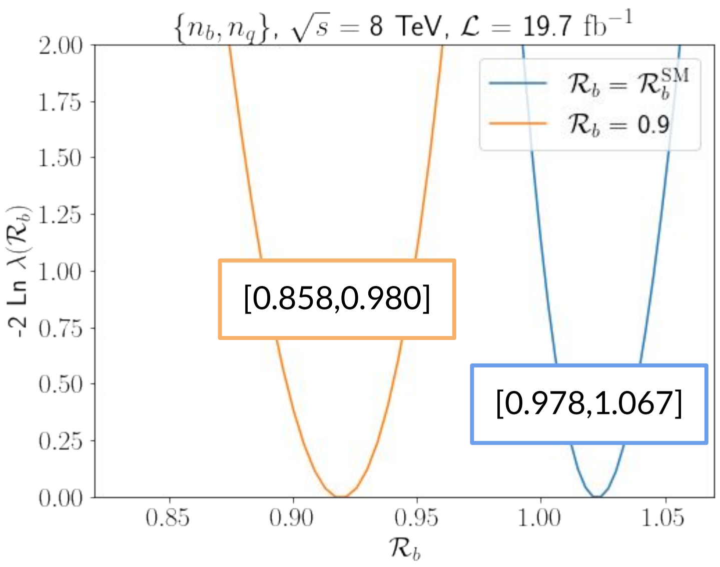
$$\lambda(\mathcal{R}_b) = \frac{\mathcal{L}(\mathcal{R}_b, \hat{\hat{\theta}}_i(\mathcal{R}_b))}{\mathcal{L}(\hat{\mathcal{R}}_b, \hat{\theta}_i)}$$

and the associated test statistic, which helps us compute the 95% Confidence Intervals

$$-2\ \mathrm{Ln}\ \lambda(\mathcal{R}_b)$$

It already improves the results!





Results

The model is able to capture the essential features of the data

It is able to differentiate between different values of R_b

The addition of the number of *q*-tagged jets considerably improves the fit by tightening the Confidence Intervals even with large systematic uncertainties

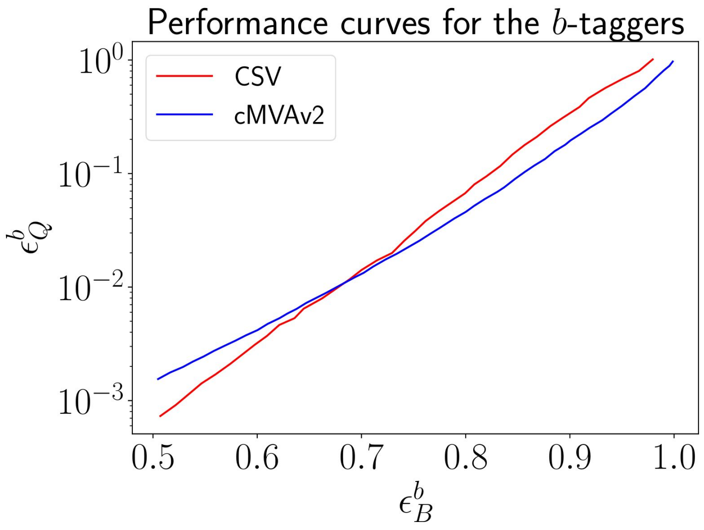
Results

Observables	$\mathcal{R}_b = \mathcal{R}_b^{\rm SM}$	$\mathcal{R}_b = 0.9$
$\{n_b\}$	[0.894, 1.067]	[0.808, 0.970]
$\{n_b, n_q\}$	[0.978, 1.067]	[0.858, 0.980]

Incorporating state-of-the-art *b*-taggers

We digitalize two state-of-the-art *b*-taggers, one for 8 TeV and another for 13 TeV

We explore different Working Points to assess the difference trade-offs between sample purity and sample statistics



Arbitrary *q*-taggers

We chose the following definition.

$$\epsilon^{q}_{B} = 0.16(1 - \epsilon^{b_2}_{B})$$

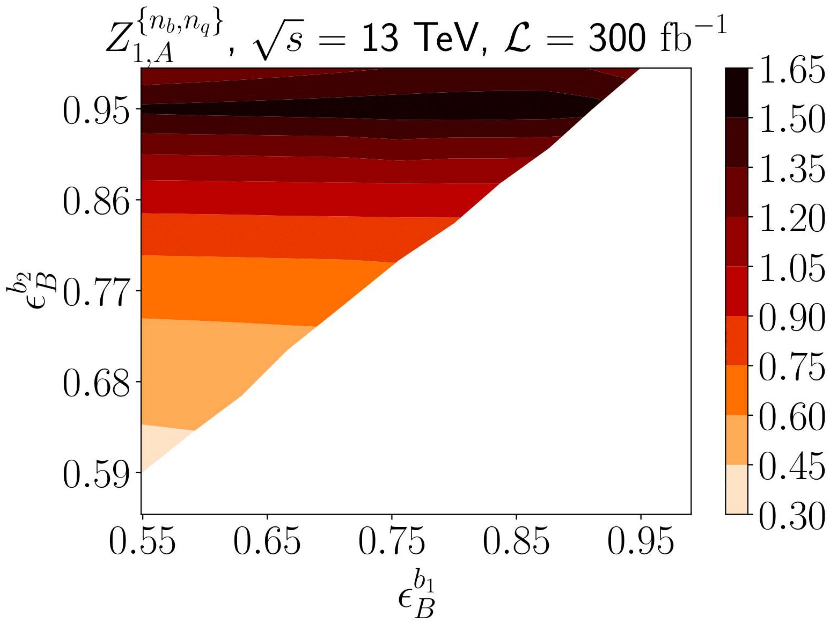
$$\epsilon^{q}_{Q} = 0.69(1 - \epsilon^{b_2}_{Q})$$

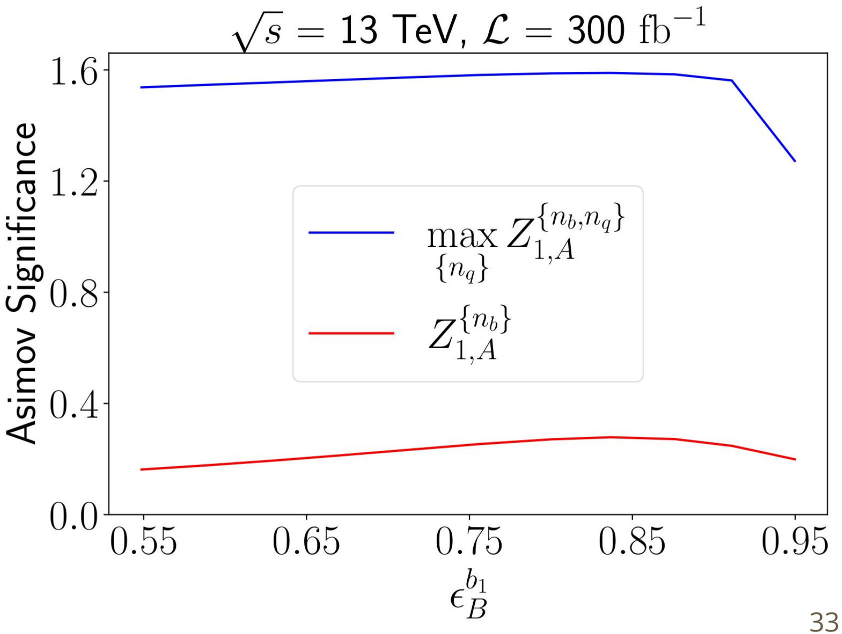
$$\epsilon^{\alpha}_{j;\not{t}} = 0.85\epsilon^{\alpha}_{Q}$$

Systematic uncertainties

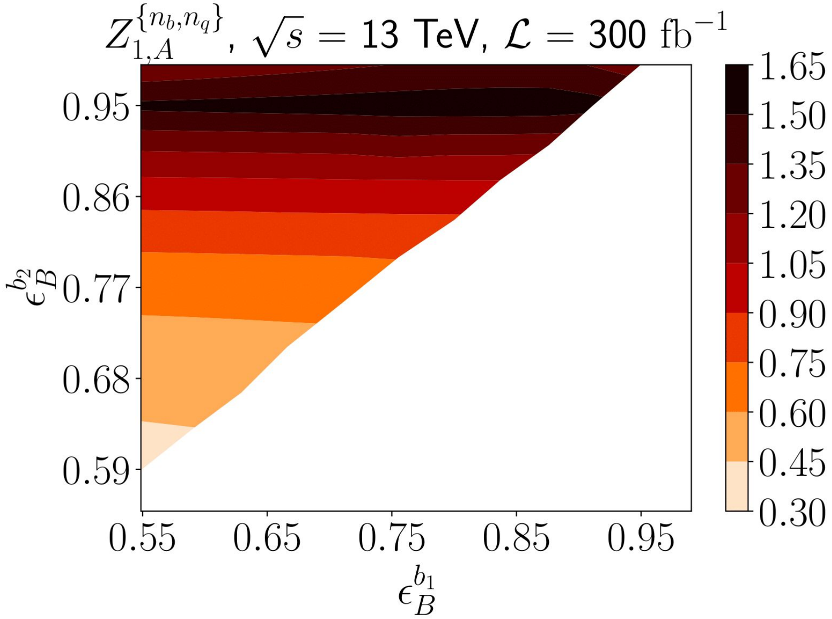
Nuisance Param.	Uncertainty
$\epsilon_B^{b_1}$	2%
$\epsilon_Q^{b_1}$	11%
$\epsilon_{B,Q}^{q}$	5%

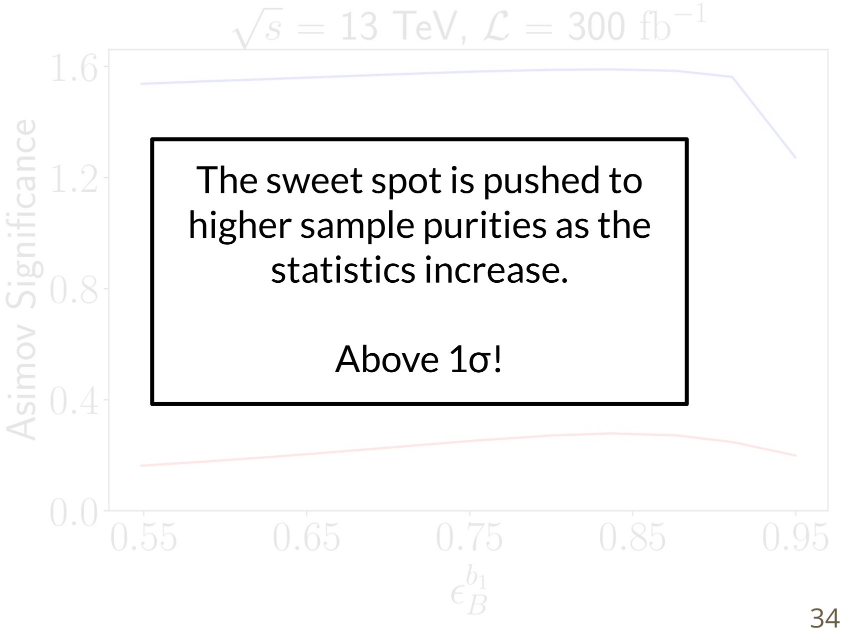
Results: Run 2 projection





Results: Run 2 projection

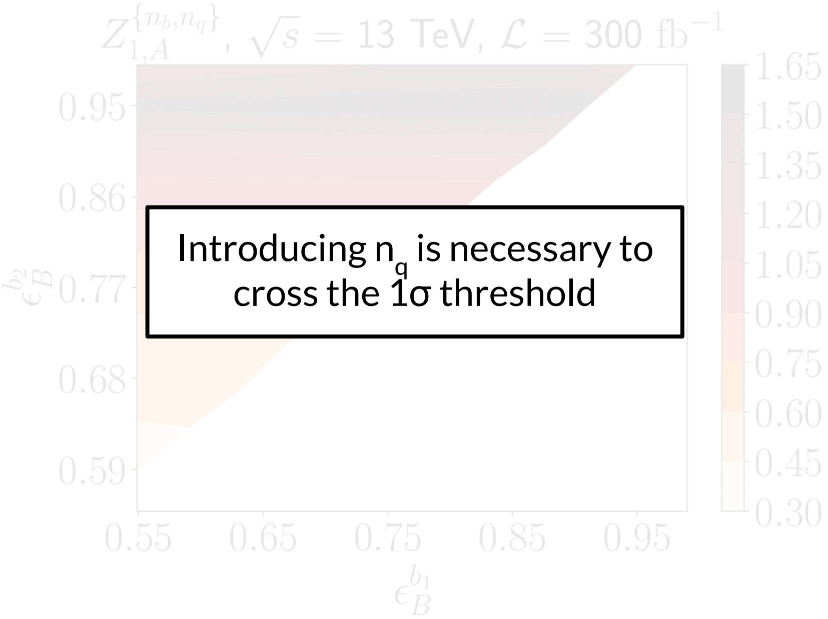




The sweet spot is pushed to higher sample purities as the statistics increase.

Above 1σ!

Results: Run 2 projection



Introducing n_q is necessary to cross the 1σ threshold