Prospects for Theory Systematics in Precision Higgs Measurements

Frank Tackmann

Deutsches Elektronen-Synchrotron

HL-LHC workshop
May 11, 2015

Talking About Future Theory Uncertainties ...

Disclaimers:

- By "theory uncertainty/systematics" I mean the uncertainty due to our inability to calculate something
  - I do not mean parametric uncertainties from inputs (PDFs, α_s, etc.). These are also important, but there will be other talks about it
- I will try not to talk (much) about the actual calculations, but also about future possibilities and what is needed to keep the theory under control
  - In all matters of opinion, I will give you mine ...

Talking About Future Theory Uncertainties ...

Disclaimers:

- By "theory uncertainty/systematics" I mean the uncertainty due to our inability to calculate something
  - I do not mean parametric uncertainties from inputs (PDFs, α_s, etc.). These are also important, but there will be other talks about it
- I will try not to talk (much) about the actual calculations, but also about future possibilities and what is needed to keep the theory under control
  - In all matters of opinion, I will give you mine ...

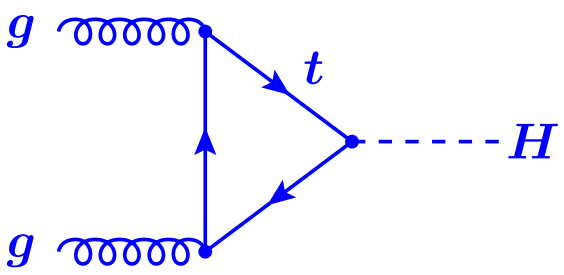
The main message

- It is clear that theory systematics will be increasingly important
- While theory will improve with time, it is essential to utilize increased statistics to help control (avoid, reduce) theory uncertainties

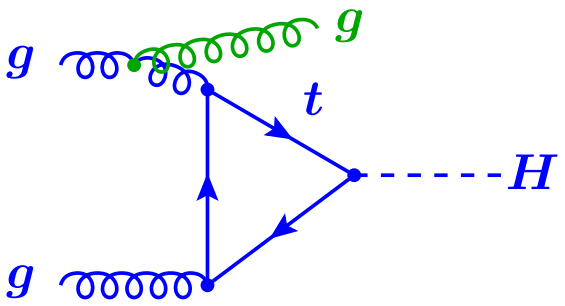
Executive Summary of Higgs Production.

$\sim 2/3$ of Higgs bosons are produced at low p_T

0 jets

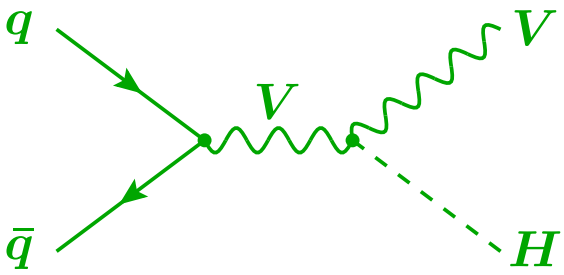


0-1 jets

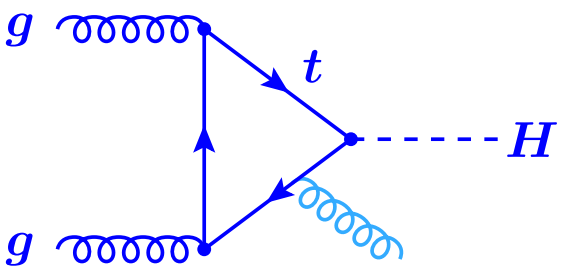


$\sim 1/3$ of Higgs bosons have sizeable p_T

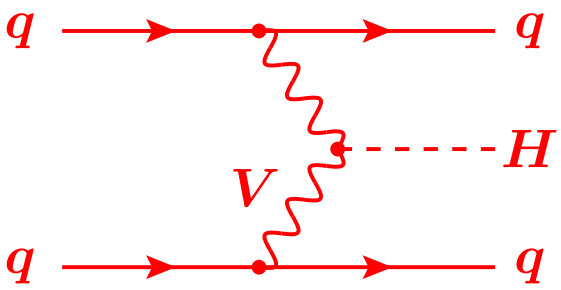
0 jets



1 jet

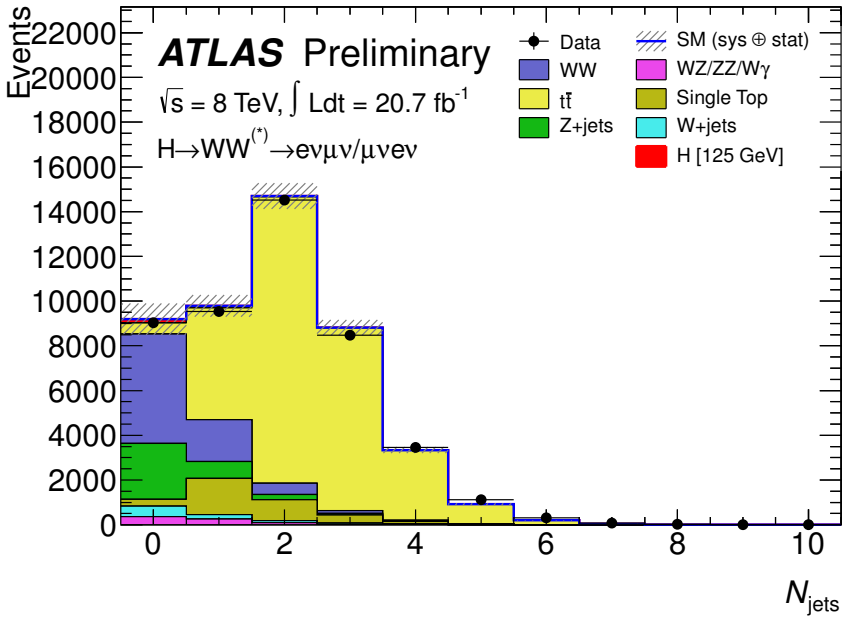


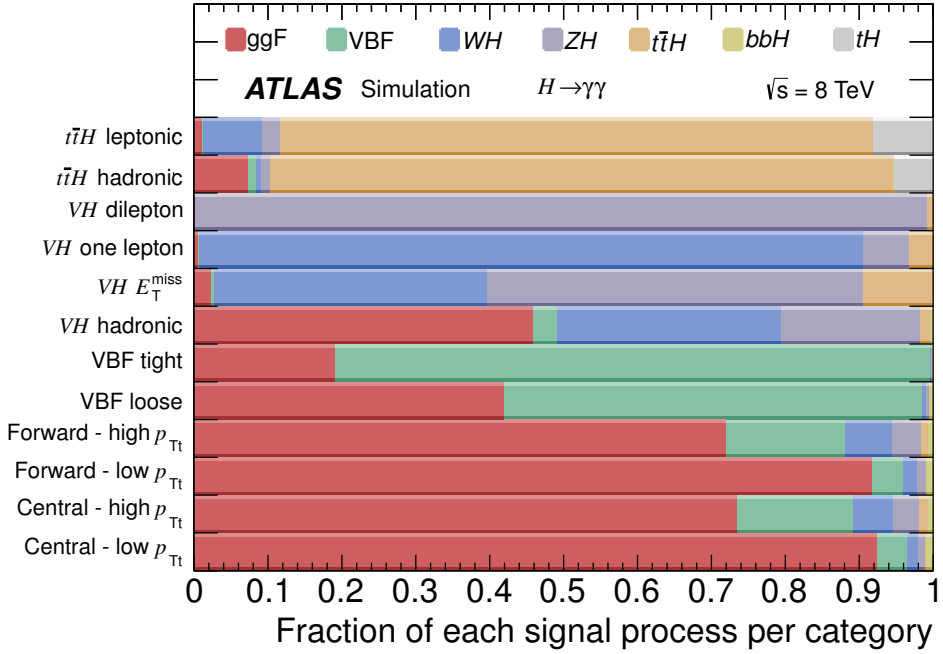
2 jets



- Kinematics and number of jets distinguishes different Higgs processes
- Discriminates against different backgrounds (e.g. in $H \to WW, \tau\tau, b\bar{b}$)

Event Kinematics.





Ultimately, experiments measure the cross sections in different categories = restricted regions of phase space

- Interpretation requires a theory prediction for each exclusive category
- In the end, want to combine categories

⇒ Consistent theory description, treatment of theory uncertainties and their correlations are essential

Event Kinematics.

In CMS [stolen from G. Petrucciani]

	p_T	n(jet)		p_T	n(jet)
$H \to \gamma\gamma$	yes	VBF, VH, ttH	$H \to bb^{(*)}$	no	VBF
$H \to WW$	no	0, 1, VBF, VH	$H \to \mu\mu$	yes	0-1, VBF
$H \to ZZ$	yes	0-1, 2	$H \to Z\gamma$	no	0-1, VBF
$H \to \tau\tau$	yes	0, 1, VBF	$H \to inv^{(*)}$	yes	VBF

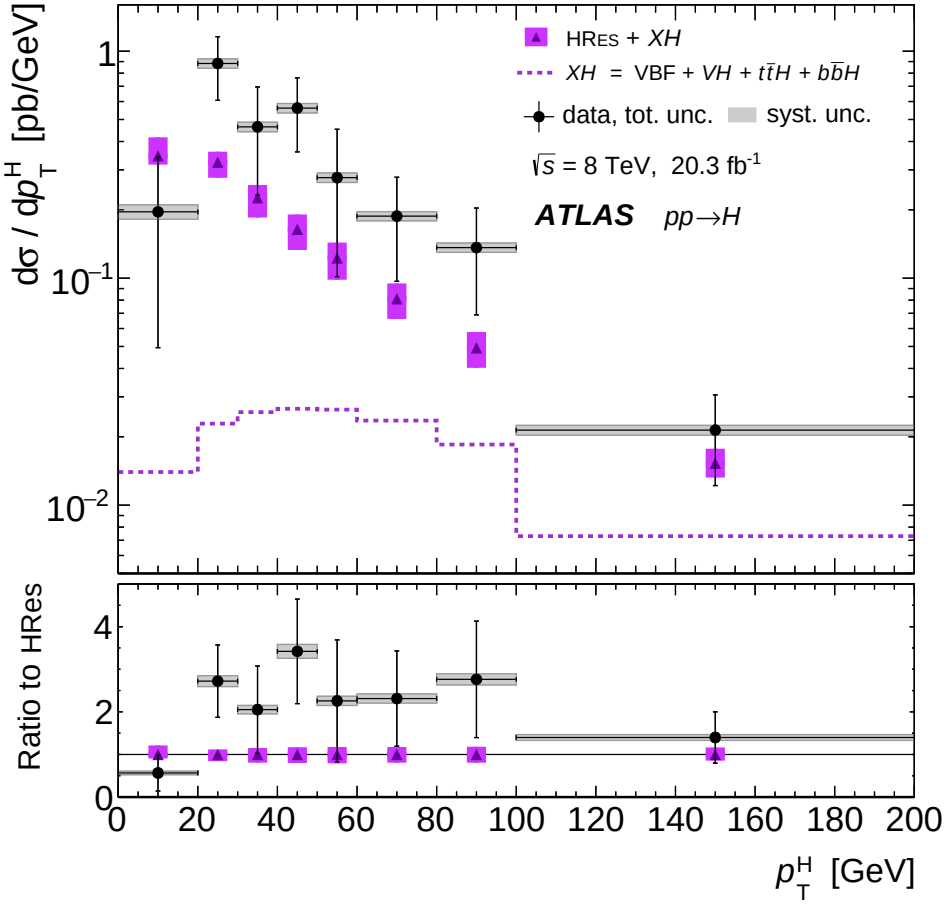
Ultimately, experiments measure the cross sections in different categories = restricted regions of phase space

- Interpretation requires a theory prediction for each exclusive category
- In the end, want to combine categories

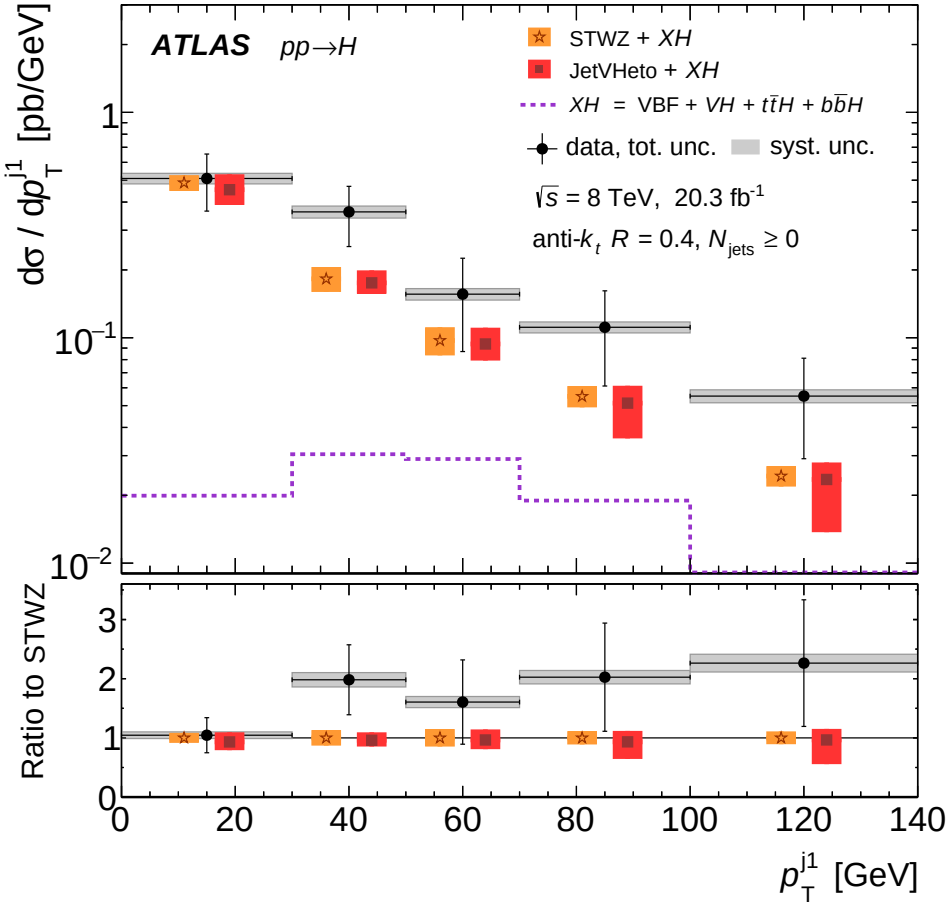
$\Rightarrow$ Consistent theory description, treatment of theory uncertainties and their correlations are essential

Current Status.

$\mathrm{d}\sigma/\mathrm{d}p_T^{\mathrm{Higgs}}$



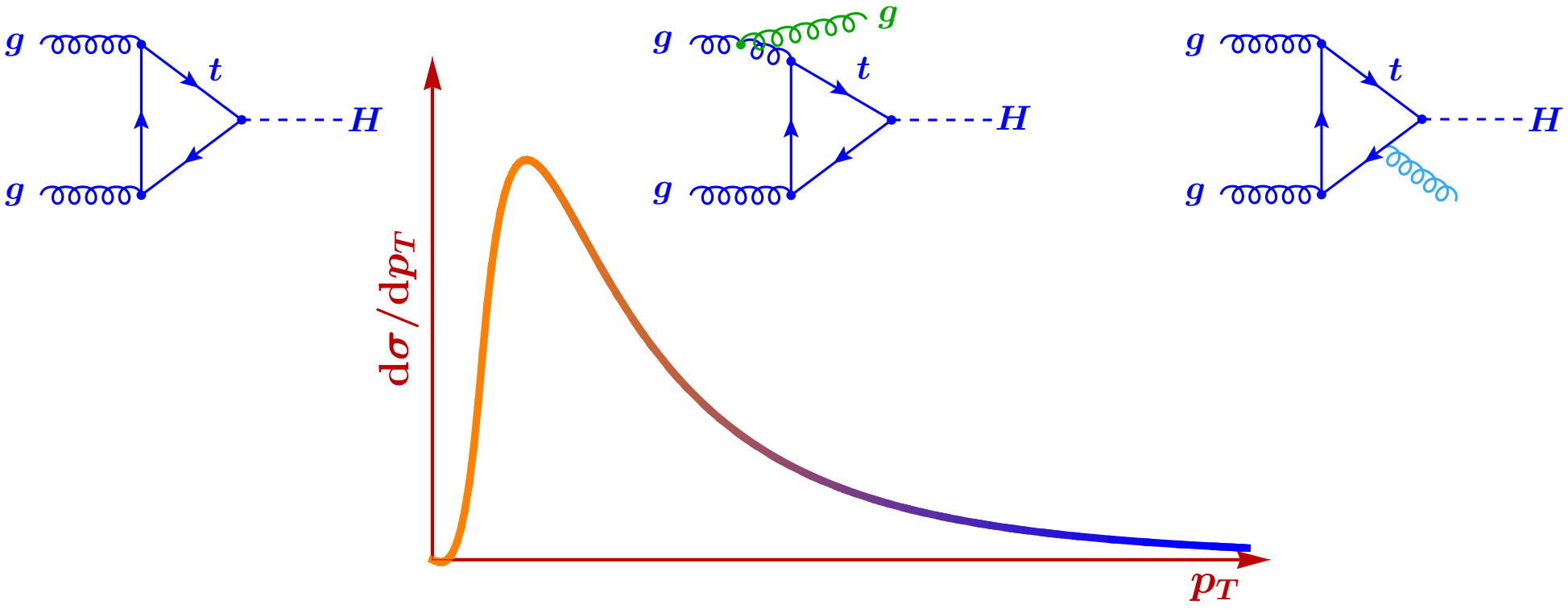
$\mathrm{d}\sigma/\mathrm{d}p_T^{\mathrm{jet}}$



Combined ATLAS $H \to \gamma\gamma$ and $H \to 4\ell$

- Still statistics limited, but theory uncertainties will become relevant with 20 times more statistics, i.e., already at end of Run2

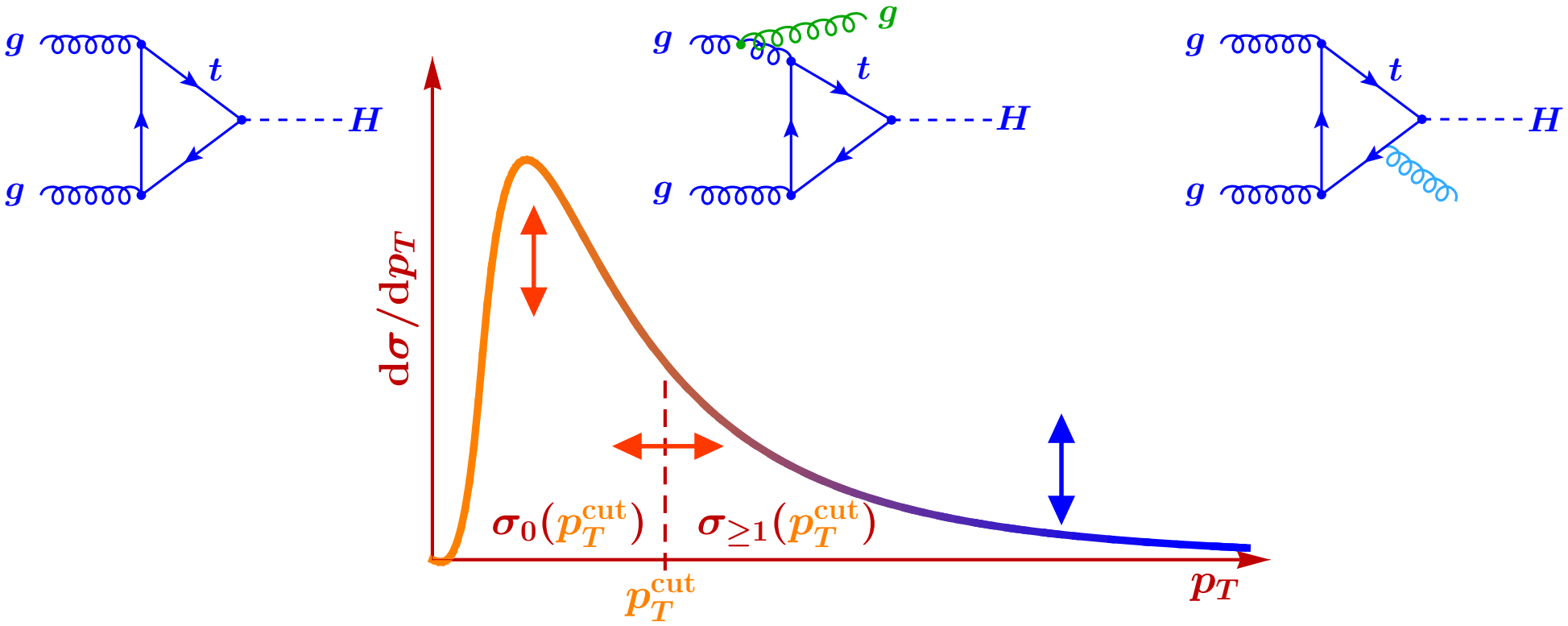
Differential Spectrum.



Spectrum describes transition between 0-jet and ≥ 1-jet regions

- Need precise predictions with consistent treatment of theory uncertainties across spectrum (both differential and integrated)

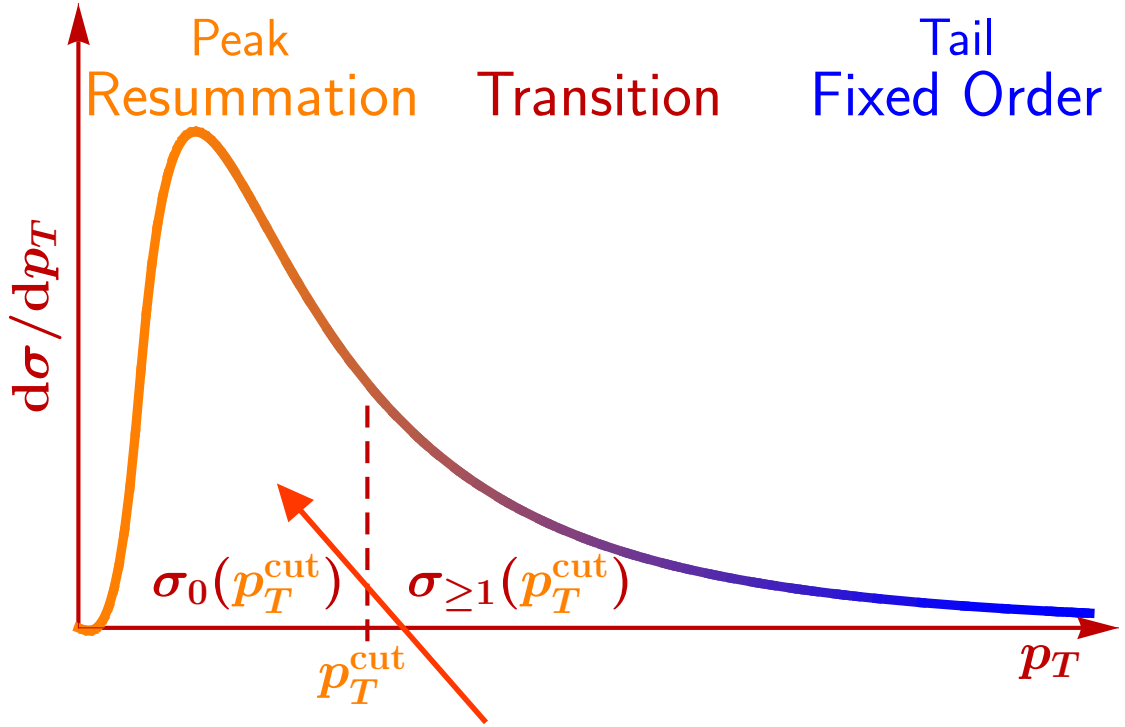
Differential Spectrum.



Spectrum describes transition between 0-jet and $\geq$ 1-jet regions

- Need precise predictions with consistent treatment of theory uncertainties across spectrum (both differential and integrated)
  - *quite nontrivial* because it requires nontrivial correlations (simple factor-2-scale-variation-recipes are not good enough)
  - Understood much better by now, but still open issues to figure out

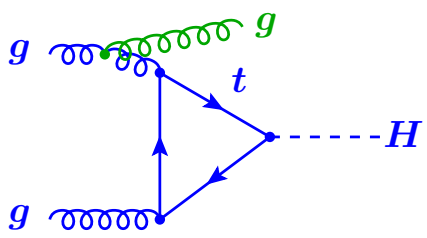
Different Regions Require Different Theory.



where $p_T \equiv p_T^H, p_T^{\rm jet1}$, ...

$p_T \to 0$:
only soft or collinear emissions

$p_T \sim m_H$:
additional hard emissions



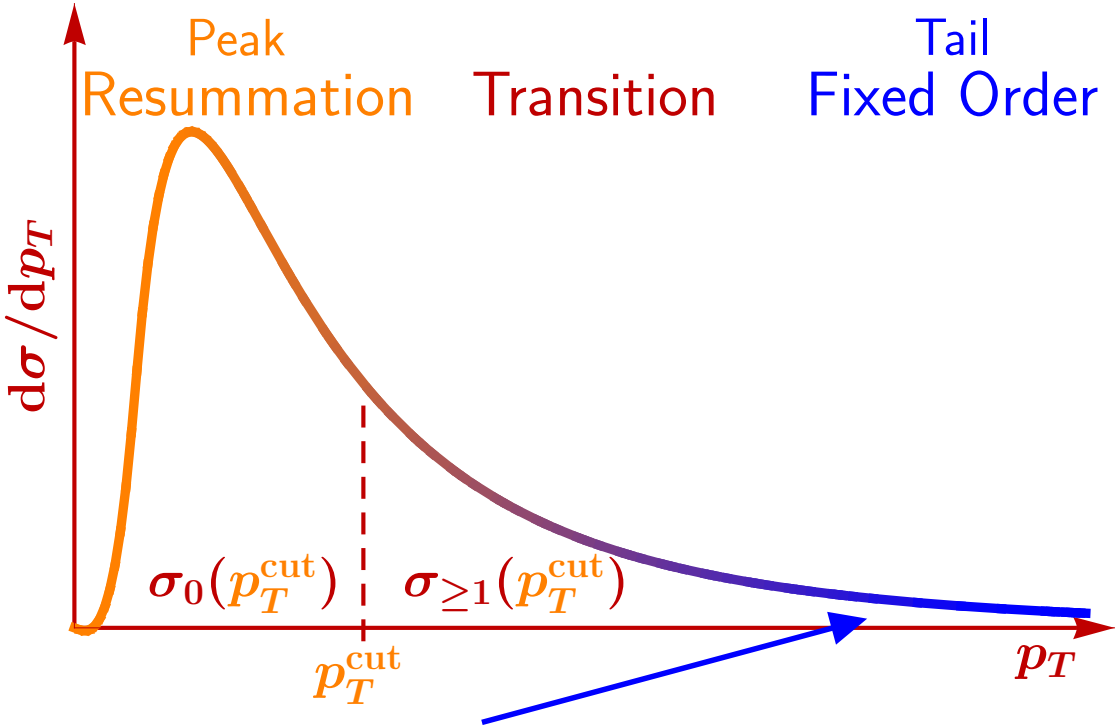
Resummation region (in MC = parton shower regime)

- Spectrum at low $p_T \ll m_H$ and 0-jet cross section at low $p_T^{\rm cut} \ll m_H$
  - Pert. corrections grow large for $p_T \to 0$ due to large logs $\alpha_s^n \ln^m(p_T/m_H)$

$$\sigma_0(p_T^{\rm cut}) = \sigma_B \left[1 - \frac{\alpha_s}{\pi} 2C_A \ln^2 \frac{p_T^{\rm cut}}{m_H} + \mathcal{O}(\alpha_s^2 \ln^4) \right]$$

  - Should be resummed to all orders to obtain reliable precise predictions

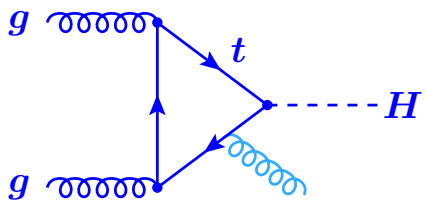
Different Regions Require Different Theory.



where $p_T \equiv p_T^H, p_T^{jet1}, \ldots$

$p_T \to 0$:
only soft or collinear emissions

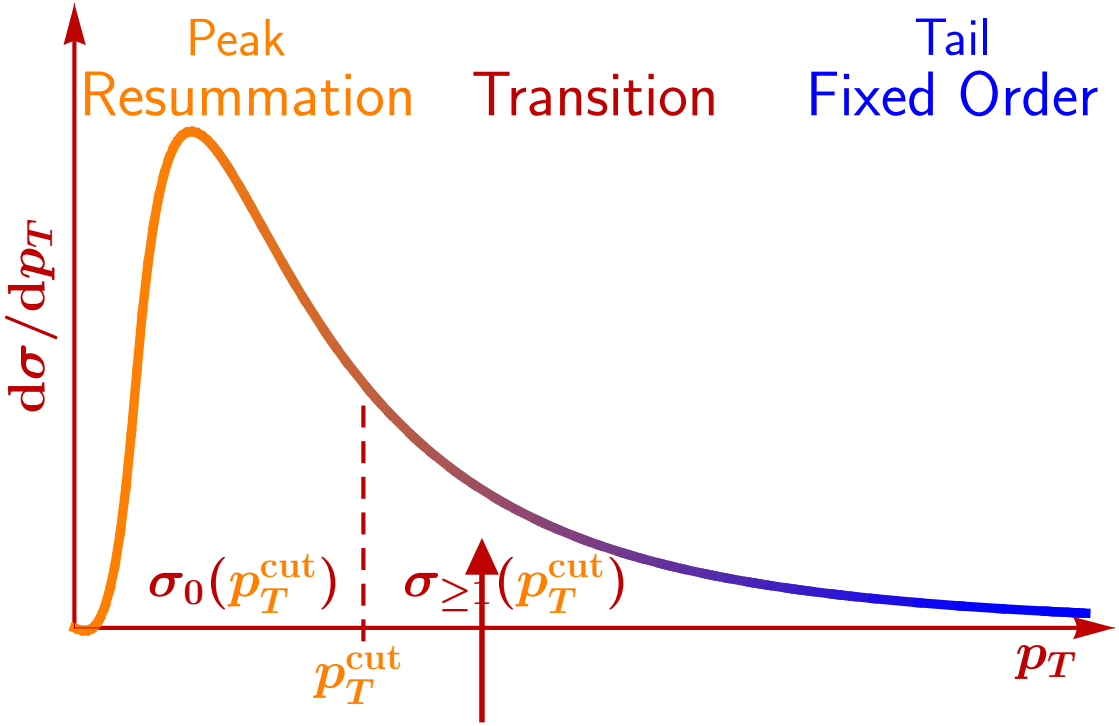
$p_T \sim m_H$:
additional hard emissions



Fixed-order region

- Spectrum at high $p_T \sim m_H$
  - Hard kinematics of inclusive H+1-jet process
  - At some point effective ggH vertex gets resolved
- Integral over $p_T \leq p_T^{cut} \to m_H$
  - H+0-jet cross section $\to$ inclusive total Higgs cross section

Different Regions Require Different Theory.



where $p_T \equiv p_T^H, p_T^{jet1}, \ldots$

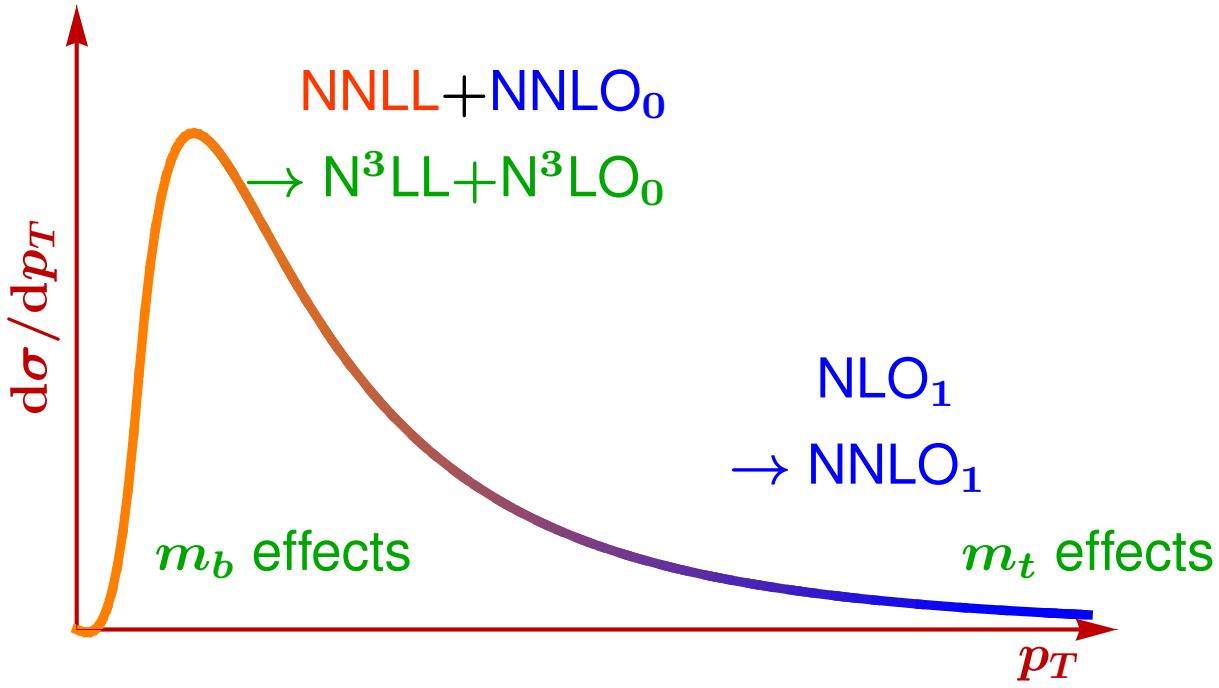
$p_T \to 0$:
only soft or collinear emissions

$p_T \sim m_H$:
additional hard emissions

Transition region

- Experimentally often the most relevant: $p_T^{cut} \sim 25 \ldots 65\,\mathrm{GeV}$
- While theoretically the most subtle
  - Best prediction from properly combining resummation+fixed order
  - Which will then also provide a consistent description with trustable uncertainties over the entire spectrum

Expected Improvements.



What can and will happen over the next couple of years

- Adding one more N (just knowing total N^{3}LO is not enough)
- Better understanding of quark mass effects

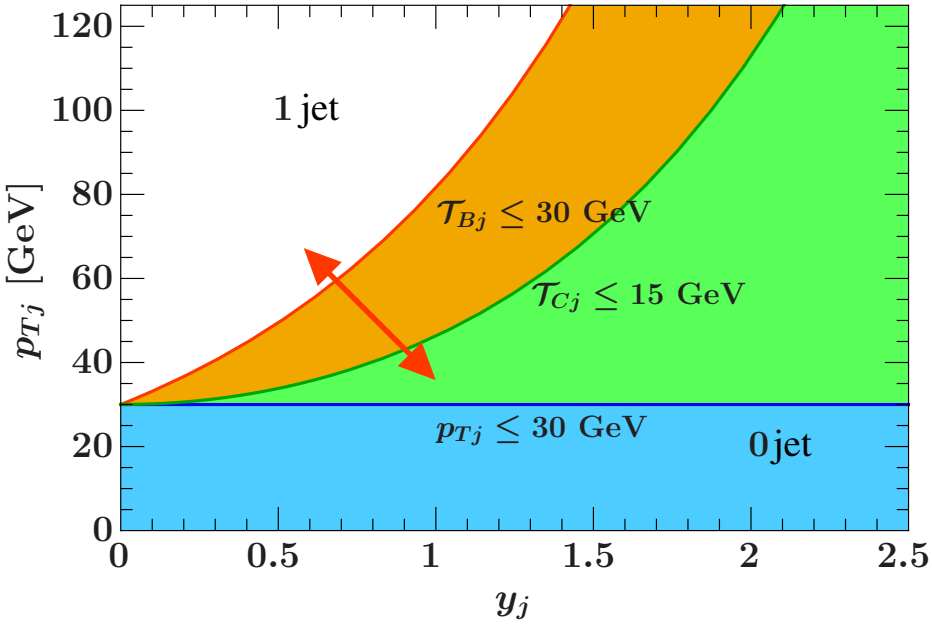
$\Rightarrow$ It is not unrealistic to expect 5-10% accuracy (plus PDF uncertainties)

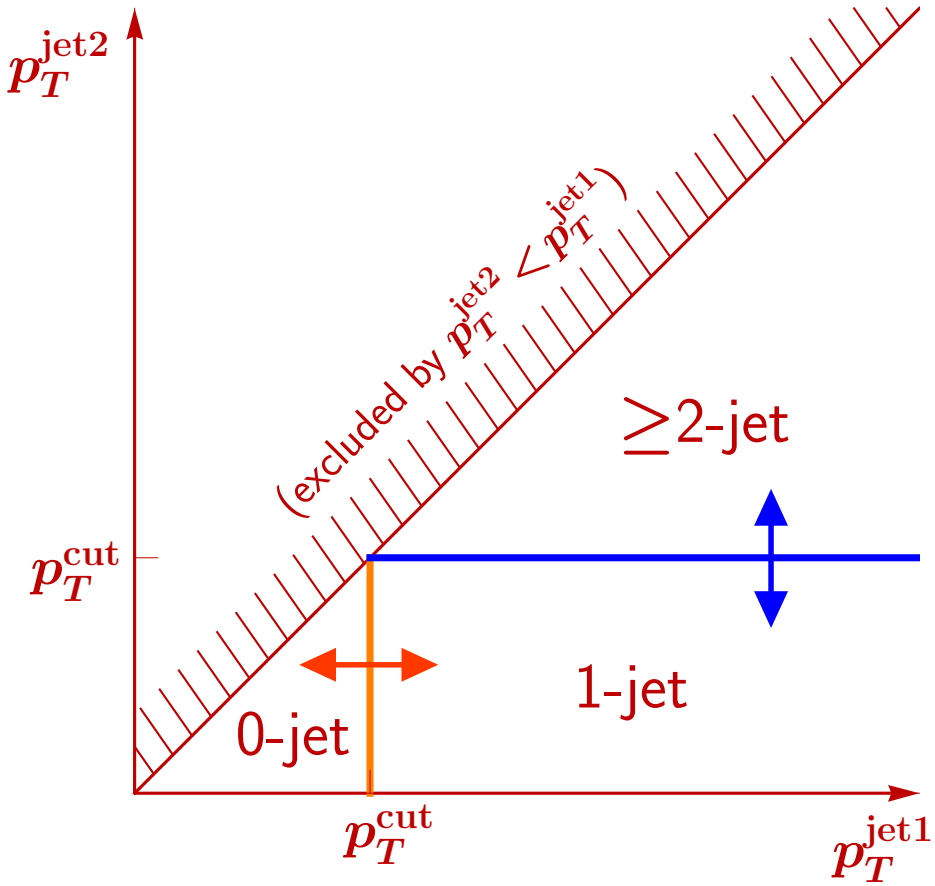
(But it is hard to make accurate predictions, especially about the future.)

However, in many ways this is the simplest possible case ...

In More Dimensions and with More Jets...

... it quickly gets more complicated





Important to better understand distribution and correlations between different variables and across different jet multiplicities

- Resummation still mostly limited to parton shower
- Getting better handle requires multi-differential higher-order resummation
  - Technology is (mostly) developed, should be able to go to NNLL

Hadronization and Underlying Event.

It's only a matter of time (i.e. statistics) that they will become relevant

Hadronization

- Can be understood and controlled with field-theory methods
  - There is plenty of experience from other precision areas (e.g. e^+e^- event shapes, B physics)
  - Application to LHC is certainly possible but has only just started

Underlying Event

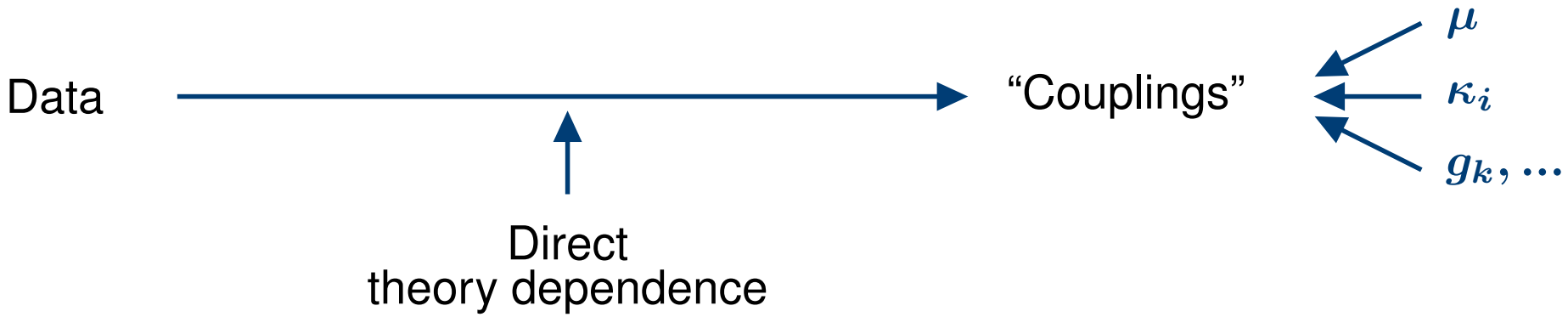
- Really needs to be understood better from first principles
  - But there are reasons to be optimistic
- *Important:* Need to find ways to disentangle UE and pile-up, especially in regions where tracking cannot be used for pileup mitigation

$\Rightarrow$ Precise (and dedicated) measurements would help a lot for both

- but need to probe low jet scales (R dep., low p_T^{jet}, jet mass, N-jettiness, ...)

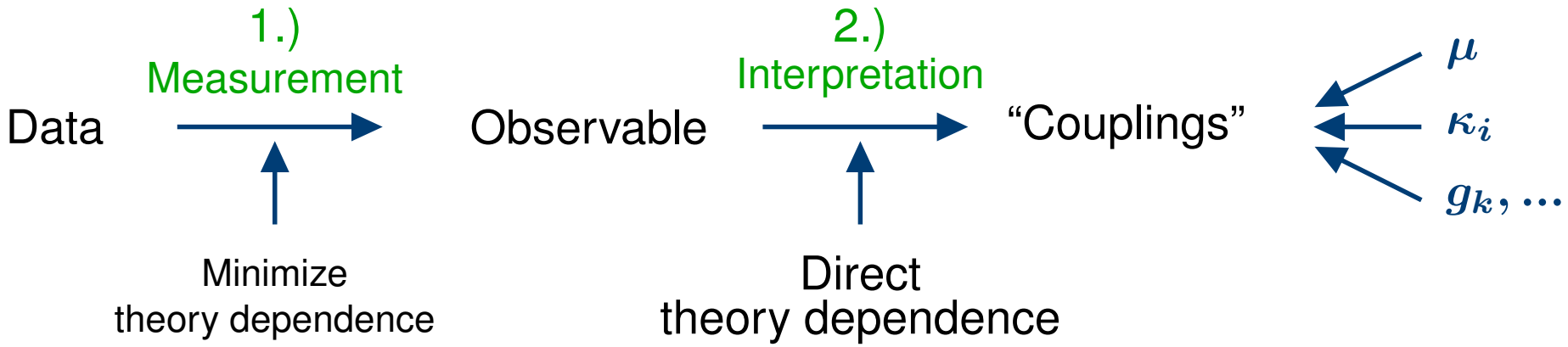
Managing Theory Dependence.

Discovery mode:



Managing Theory Dependence.

Precision mode:

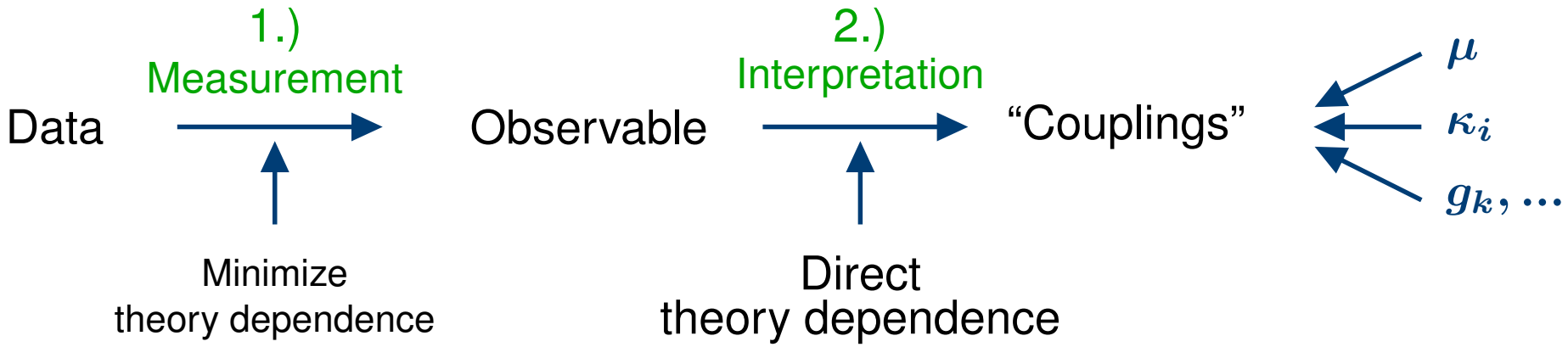


Defining a sensible and practical "Observable" interface

- Many options: Differential/fiducial cross sections, pseudo-observables, ...
- Goal should be to minimize the *theory dependence* in Measurement in two respects
  - Underlying physics model (SM, Higgs EFT, specific NP models, ...)
  - Theory systematics in signal (and also background) distributions $\rightarrow$ I will mostly refer to this
- Far from trivial and easier said than done ...

Managing Theory Dependence.

Precision mode:



Defining a sensible and practical "Observable" interface

- Ultimately theory uncertainties cannot be avoided, instead one can
  - Decouple them: Move them from 1.) to 2.)
  - Can choose the observables used in step 2.) according to desired theory cleanliness

  (So the point is not that theorists should have all the fun with 2.), but to factorize the theory dependence)

$\Rightarrow$ More data helps a lot here because it provides more flexibility

Theory Dependence in the Measurements.

When designing and optimizing measurements

- Minimizing any direct theory-dependence is more important than trying to minimize the eventual σ(measurement) $\oplus$ σ(interpretation)

Measurements will always have some (hopefully) subleading theory dependence through signal MC

- In a perfect world, in 1.) signal MC should only be needed to correct measurements for detector effects
- Prospects look good that perturbative aspects of MC models will be brought under better control
  - Going to NNLO
  - Incorporating improved resummations
  - Reliable perturbative uncertainties

$\Rightarrow$ Very important also to be able to evaluate theory (in)dependence of measurements

Example: Jet-binning in $H \to WW$.

There is no point in trying to find "the one optimal" jet-veto cut, rather optimize for minimal dependence on signal and background modelling

With sufficient data you can

1.) Measure the cross section
   - As a function of the jet-veto cut (or directly the spectrum)
   - Using alternative jet-veto variables
   - Using different jet radii

2.) Fit everything to theory
   - Requires predictions with consistent description of whole spectrum
   - Fit will pick the region with best σ(measurement) $\oplus$ σ(theory)
   - At the same time test/constrain the theory

$\Rightarrow$ This is not easy (nobody said it would be), but probably the best (and maybe the only) way to precision

Combining Different Channels.

1.) Measure

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_T}(H \to \gamma\gamma)\,, \qquad \frac{\mathrm{d}\sigma}{\mathrm{d}p_T}(H \to ZZ)\,, \qquad \frac{\mathrm{d}\sigma}{\mathrm{d}p_T}(H \to WW)$$

- Could consider ratios if this improves the measurement step, i.e., if it helps to reduce some dependence on signal modelling

2.) For combined interpretation

- Measuring ratios is not necessary here for theory uncertainties to drop out, the combined fit automatically takes this into account
- What is relevant is to
  - Measure similar kinematics in $H \to \gamma\gamma/ZZ$ as in $H \to WW$
  - Properly correlate theory uncertainties between different channels

Alternatives: Rapidity-Weighted Jet Bins.

Generalize p_T^{jet} by defining

$$\mathcal{T}_{fj} = p_{Tj}\, f(y_j)$$

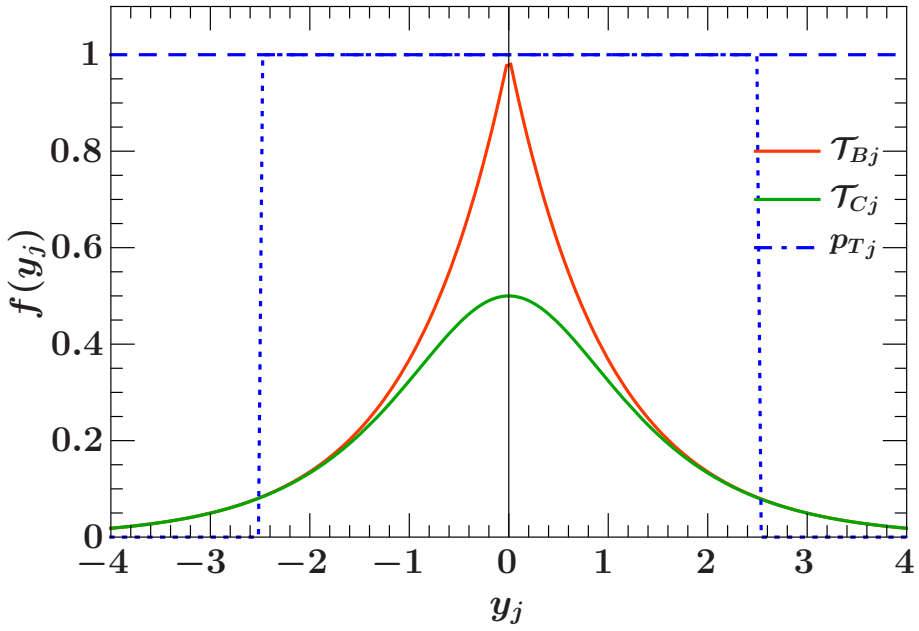
$$\mathcal{T}_f^{\text{jet}} = \max_{j \in J(R)} \mathcal{T}_{fj}$$

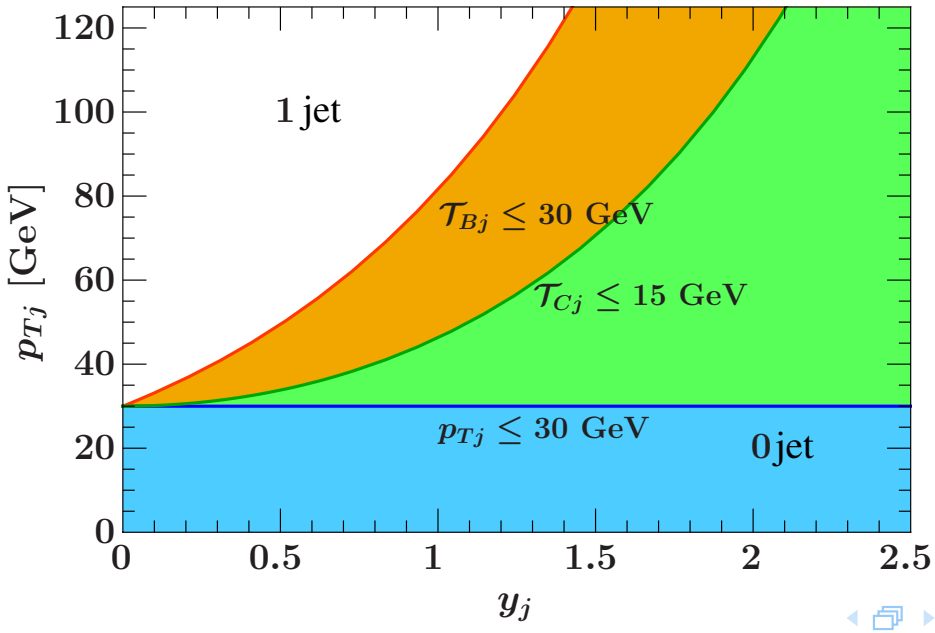
- Can choose different rapidity weighting functions such that $\mathcal{T}_f^{\text{jet}}$
  - can be resummed
  - is insensitive to forward rapidities
- Count jets according to $\mathcal{T}_{fj}$

  0 jets: $\sigma_0(\mathcal{T}_f^{\text{jet}} < \mathcal{T}^{\text{cut}})$

  $\geq$ 1 jets: $\sigma_{\geq 1}(\mathcal{T}_f^{\text{jet}} > \mathcal{T}^{\text{cut}})$

$\Rightarrow$ Provides different slicing through jet phase space

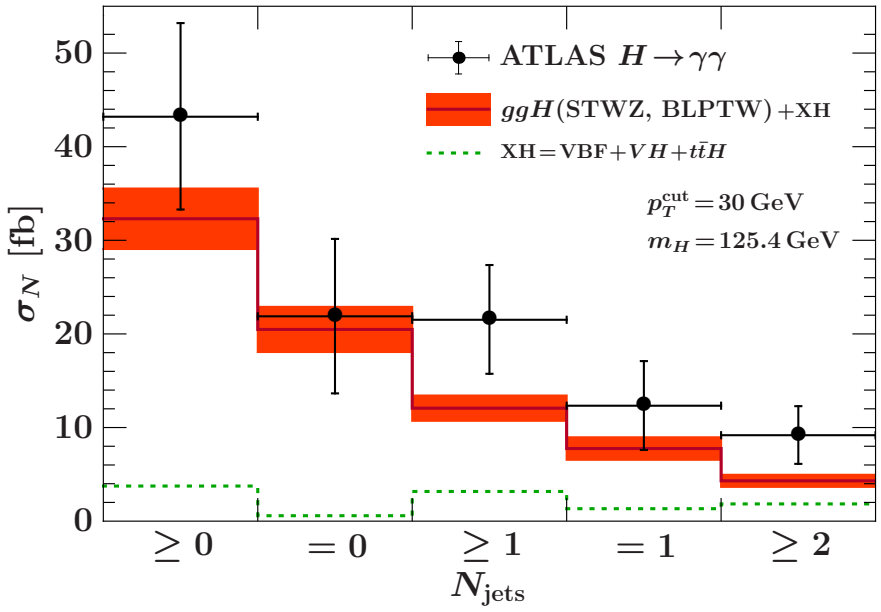




Example: Resolving Jet Bins.

(only using $H \to \gamma\gamma$ here)

Pretend uncertainties are much smaller and we want to resolve the "excess"



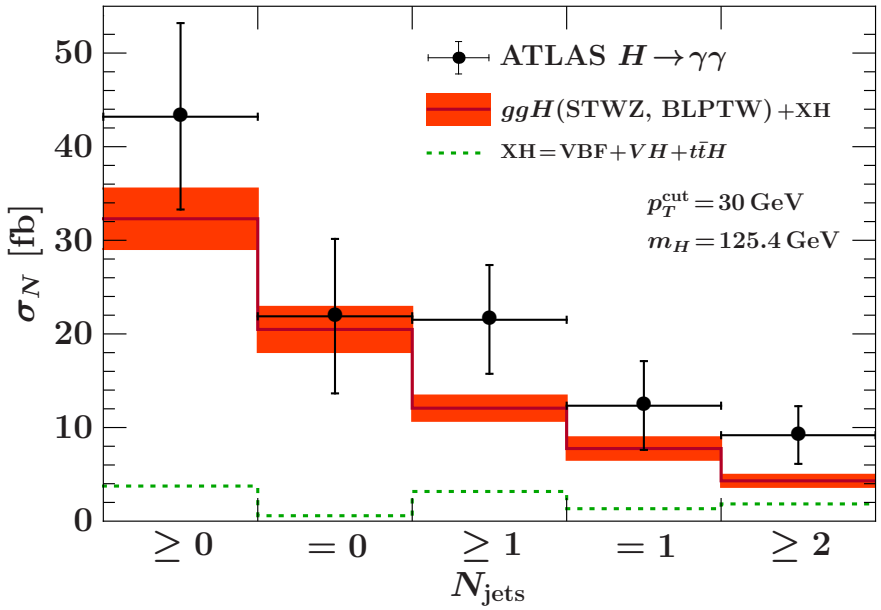
Example: Resolving Jet Bins.

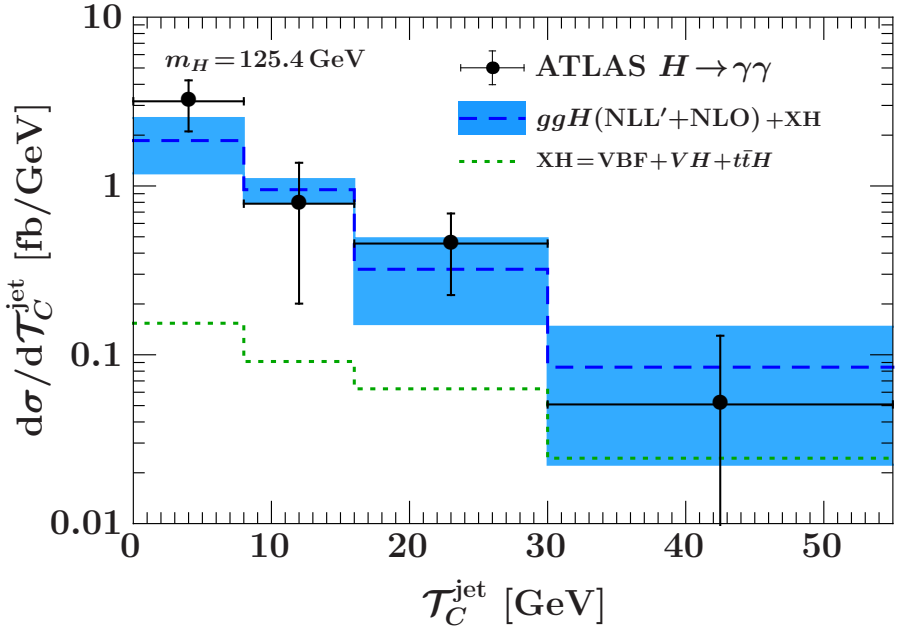
(only using $H \to \gamma\gamma$ here)

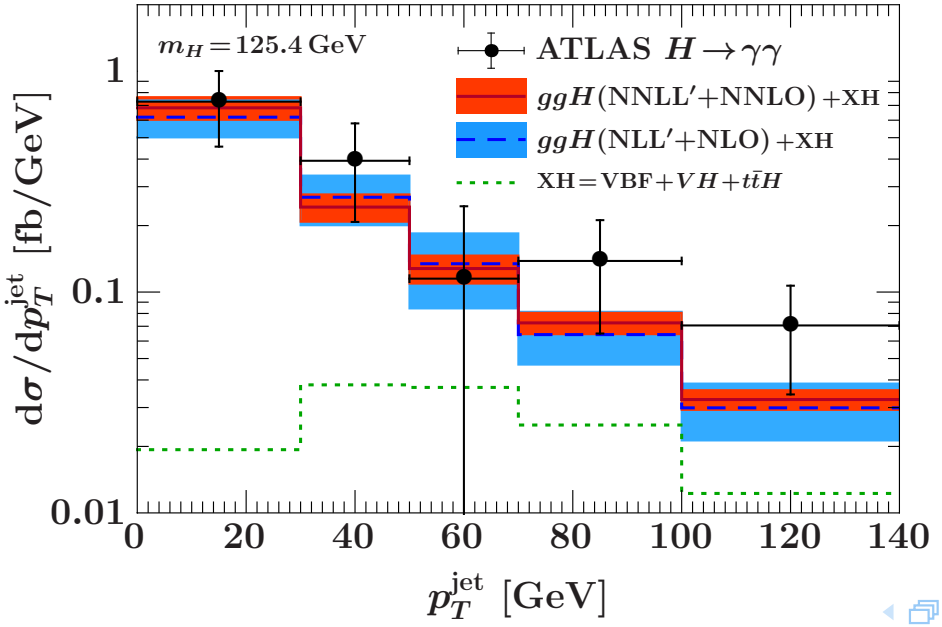
Pretend uncertainties are much smaller and we want to resolve the "excess"

- Different slices of jet phase space adds nontrivial information
- One might conclude that "excess" sits at higher p_T^{jet} at more forward rapidities

⇒ Ultimately measure double-differential in $p_T^{\text{jet}} - \mathcal{T}_f^{\text{jet}}$







Adding Other SM Processes.

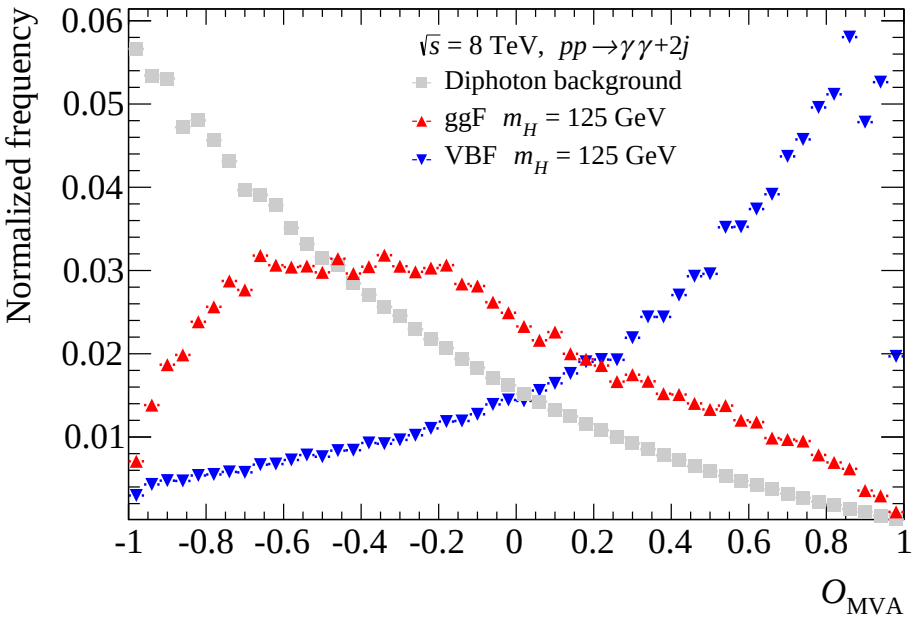
Compare:	$C \ln^2 \frac{p_T}{Q}$	$C \ln \frac{p_T}{Q}$
$gg \to H$: $C = C_A = 3$		
$Q = m_H$, $p_T = 25\,\mathrm{GeV}$	7.8	4.8
$Q = m_H$, $p_T = 30\,\mathrm{GeV}$	6.1	4.3
$q\bar{q} \to V$: $C = C_F = 4/3$		
$Q = m_Z$, $p_T = 10\,\mathrm{GeV}$	6.5	2.9
$Q = 125\,\mathrm{GeV}$, $p_T = 15\,\mathrm{GeV}$	6.0	2.8
$Q = 250\,\mathrm{GeV}$, $p_T = 15\,\mathrm{GeV}$	10.6	3.8

To test resummation regime relevant for $gg \to H$ in other processes

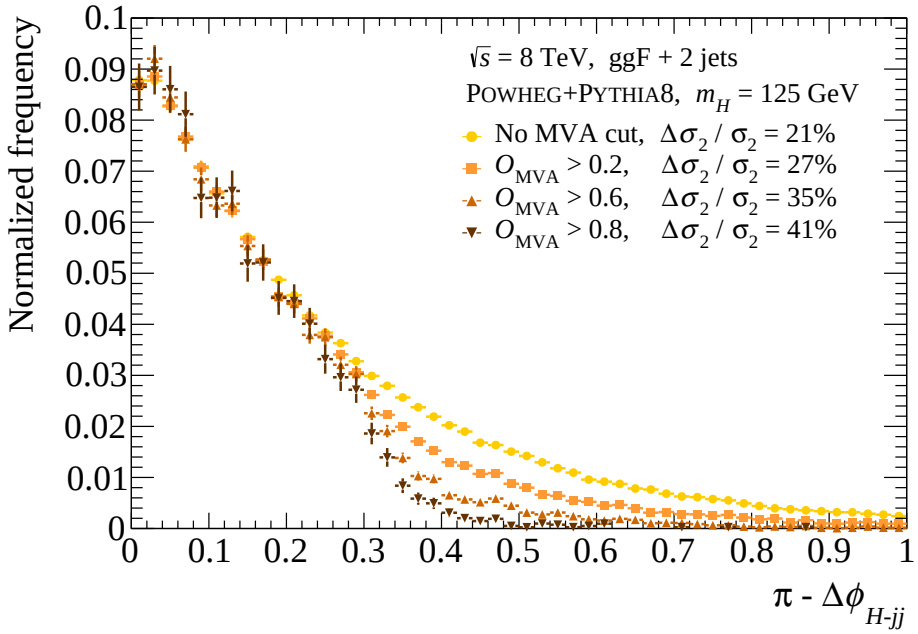
- Drell-Yan: Need to go to (very) low p_T and/or high Q
- $pp \to \gamma\gamma$: $m_{\gamma\gamma} \sim Q$, try get access to $gg \to \gamma\gamma$
- $V+j$: $p_T^{\mathrm{jet1}} \sim Q$ and $p_T \equiv p_T^{\mathrm{jet2}}, p_T^{Vj}$

$\Rightarrow$ It would actually be very valuable to have enough low pile-up data to enable precise measurements of these at low p_T^{jet}

Multivariate Selections and VBF.



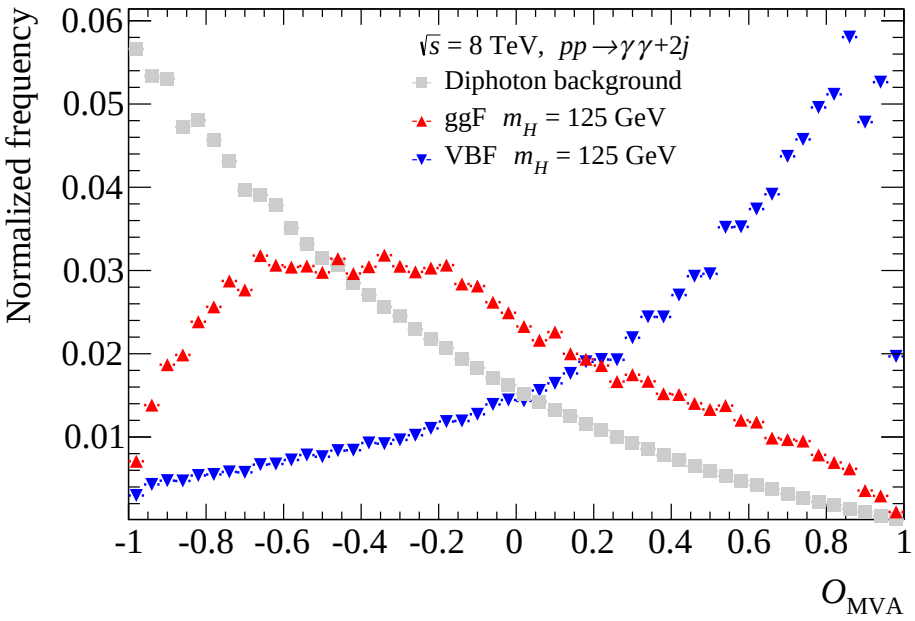
$\Rightarrow$



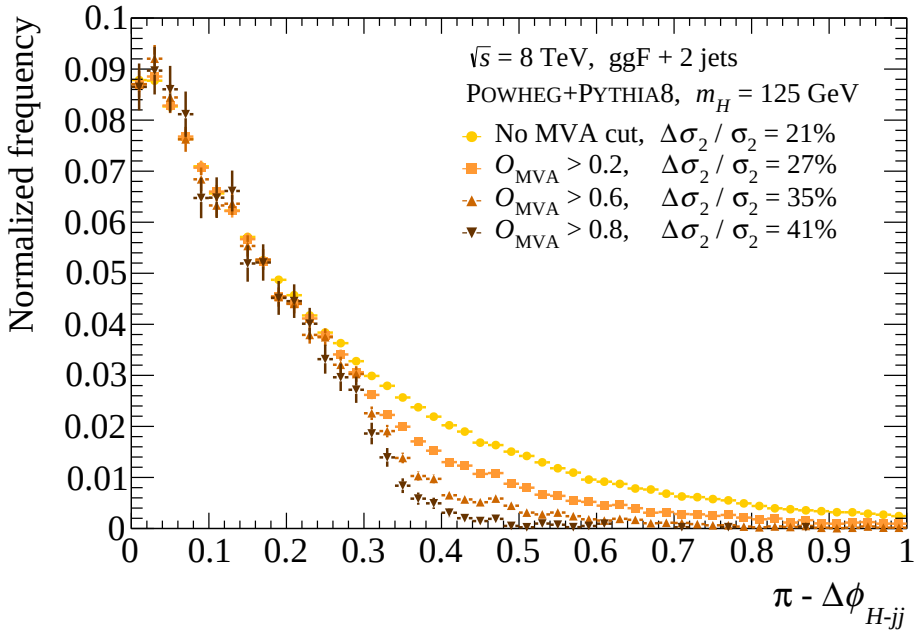
Special scrutiny is required when the observed cross section is defined by multivariate selection techniques

- Theorists will be less scared of MVAs once MVAs are more scared of theory uncertainties
  - MVA can easily construct itself a theory-sensitive variable from seemingly theory-safe inputs, e.g., $E_T^{Hjj} = p_T^H + p_T^{\mathrm{jet1}} + p_T^{\mathrm{jet1}}$
  - Must teach MVA not to cut arbitrarily into exclusive resummation regions (is done e.g. in ATLAS $H \to \gamma\gamma$ VBF categories)

Multivariate Selections and VBF.



$\Rightarrow$



Alternatively, with enough data you can

1.) Measure the ggF vs. VBF sensitive spectra (background-subtracted)
   - m_{jj} or Δy_{jj} spectrum for different jet-veto-like cuts
   - Spectrum in some 3rd-jet-sensitive variable for fixed m_{jj} cut
   - Double-differential

2.) Perform combined fit to ggF + VBF theory predictions
   - The only way really to have control of the ggF theory systematics

Summary.

At high precision

- Run1-type analyses will be theory systematics limited long before 3000/fb
  - Theory unc. that can be crude or ignored now will become relevant

To keep theory systematics under control

- It must be possible to scrutinize and improve them
  - Be careful to separate measurement and interpretation
  - If at all possible, try to avoid MVAs for signal selection
- Be redundant
  - Measure spectra and cut dependence over entire accessible range
  - Use complementary observables
  - Use other processes as standard candles

High pile-up can have negative implications also for theory systematics, e.g.,

- if it prevents SM reference measurements at low p_T^{jet} and high rapidity
- if it prevents getting control of hadronization and UE