Understanding SM Backgrounds

How do we prepare for discovery at 13TeV ?

Frank Würthwein
UCSD

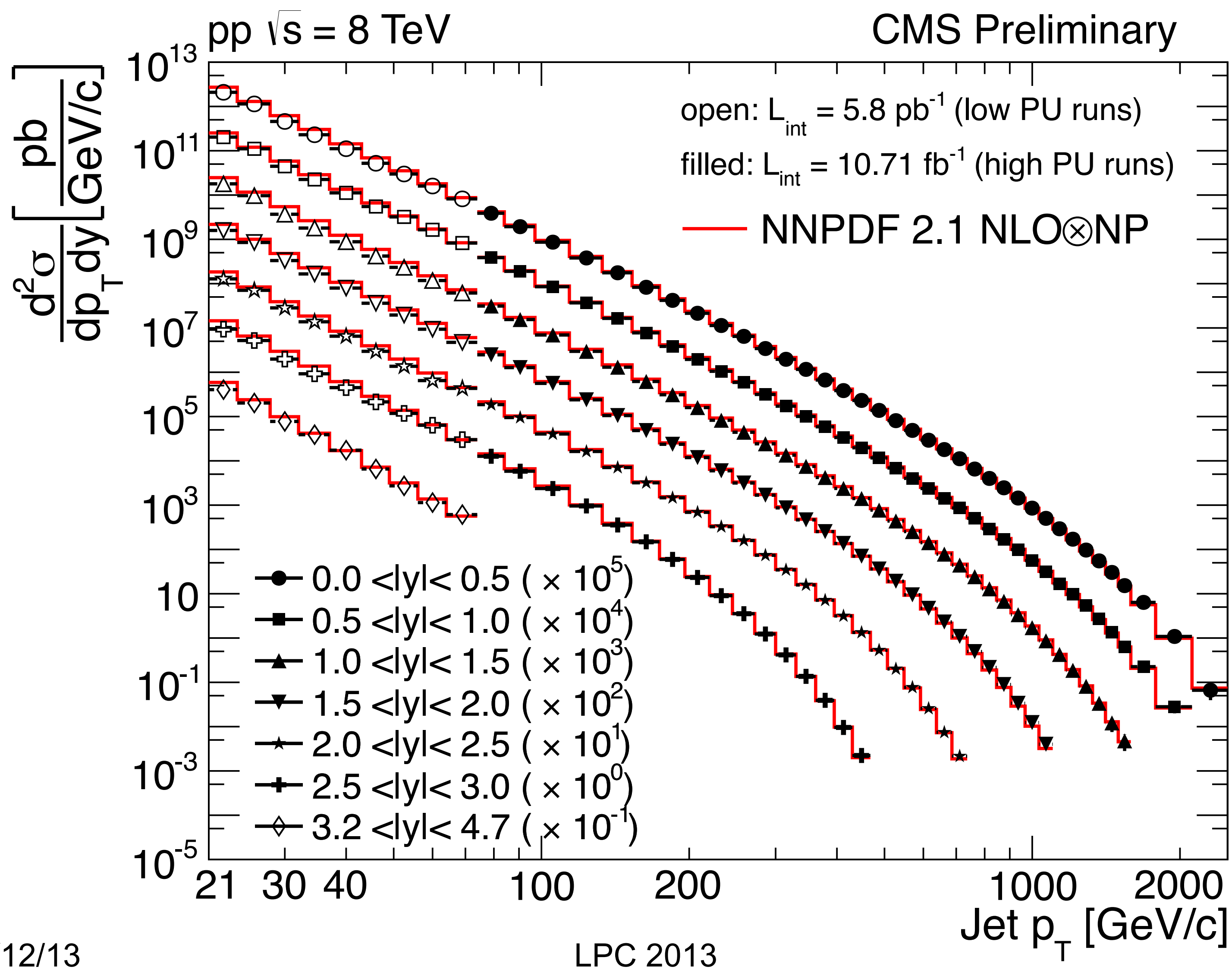
Apology

Fundamentally, the differences in approach between different searches is larger than differences between the two experiments.

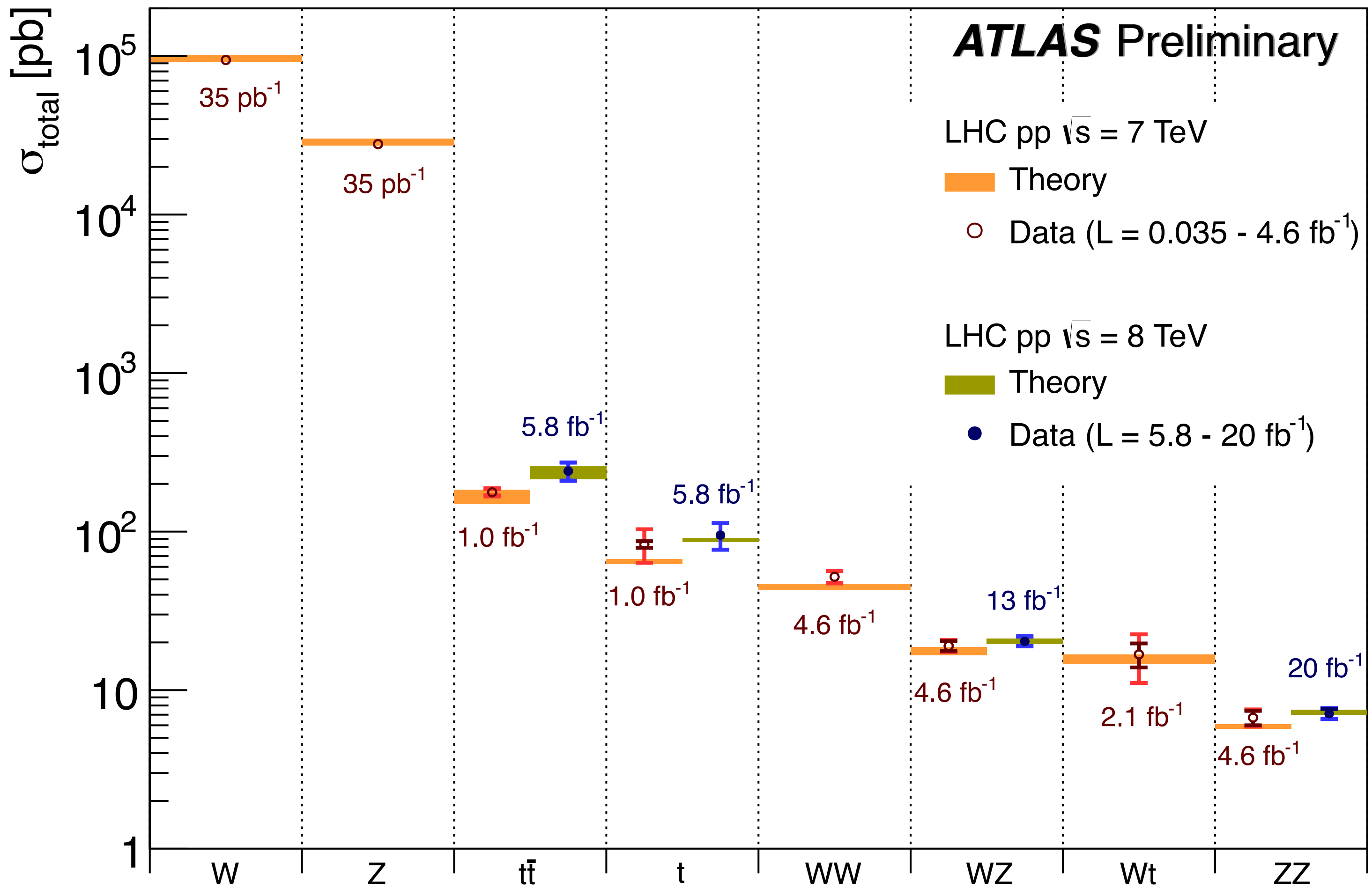
For my own convenience, I am thus taking my examples mostly from CMS.

Measured jet production over 15 Orders of magnitude

Perfect agreement between Theory & Experiment



Measured top, W, Z production



If theory and experiment agree over so many orders of magnitude and so many different physics processes, what's there left to do to prepare for discovery at 13TeV ?

Status of 8TeV published results

Submitted or published

	ATLAS	CMS
SUSY	4	5
SMP	0	1
Top	0	0

The Standard Model Groups have different needs and objectives, and work on a different time scale than the SUSY groups in either experiment !!!

Difference in Objective

- SM & Top:
  - Measure differential distributions.
  - Unfold the detector effects before publishing.
- EXO & SUSY:
  - Compare data with expectations in control regions.
  - Generally focus on tails of SM phase space.
  - Unfolding is impractical and irrelevant

Need to measure SM bkg to SUSY searches,
in data regions specifically designed for doing so.

Example Z pT distribution

SUSY

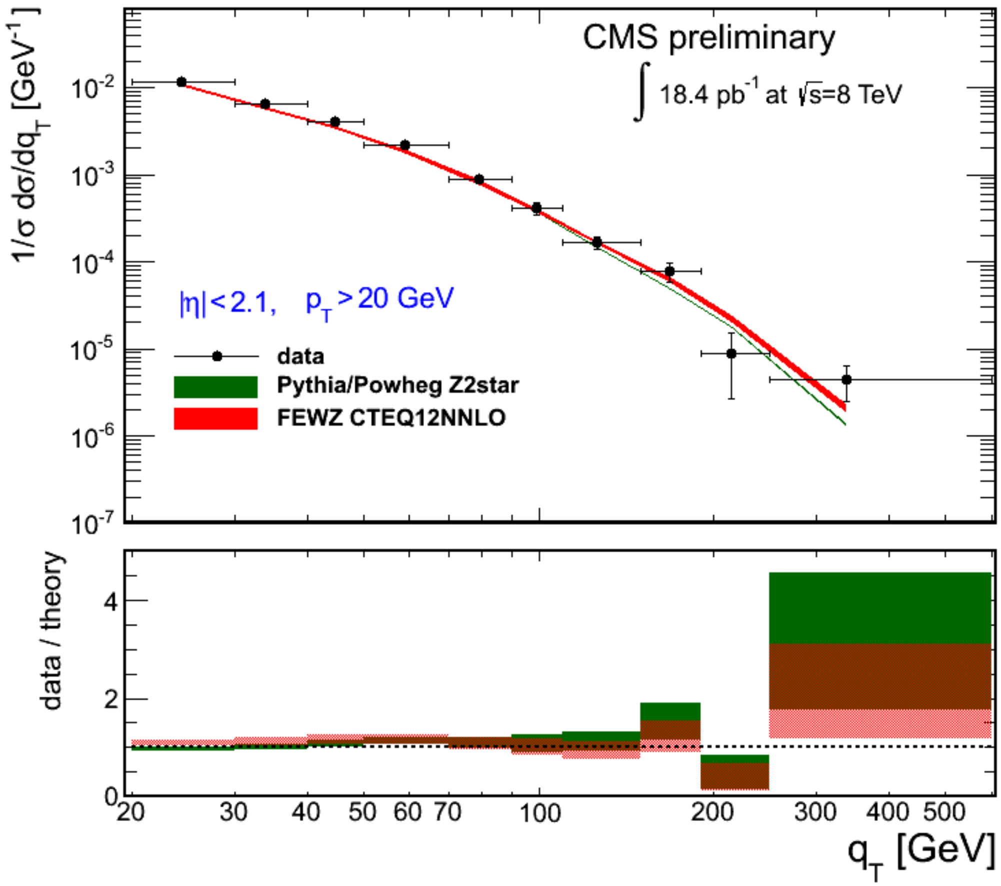
- SUSY signal MC efficiency for compressed spectra depends on ISR recoil to pass MET etc. cuts
- Interested to verify ISR boost for the generator we use vs the data we use.
- Important to reproduce Njet && ISR boost !!!

SMP

- Fundamental test of NNLO calculations.
- Measured in a special 18.4/pb (!!!) data sample with low PU.
- Unfold to compare with theory.

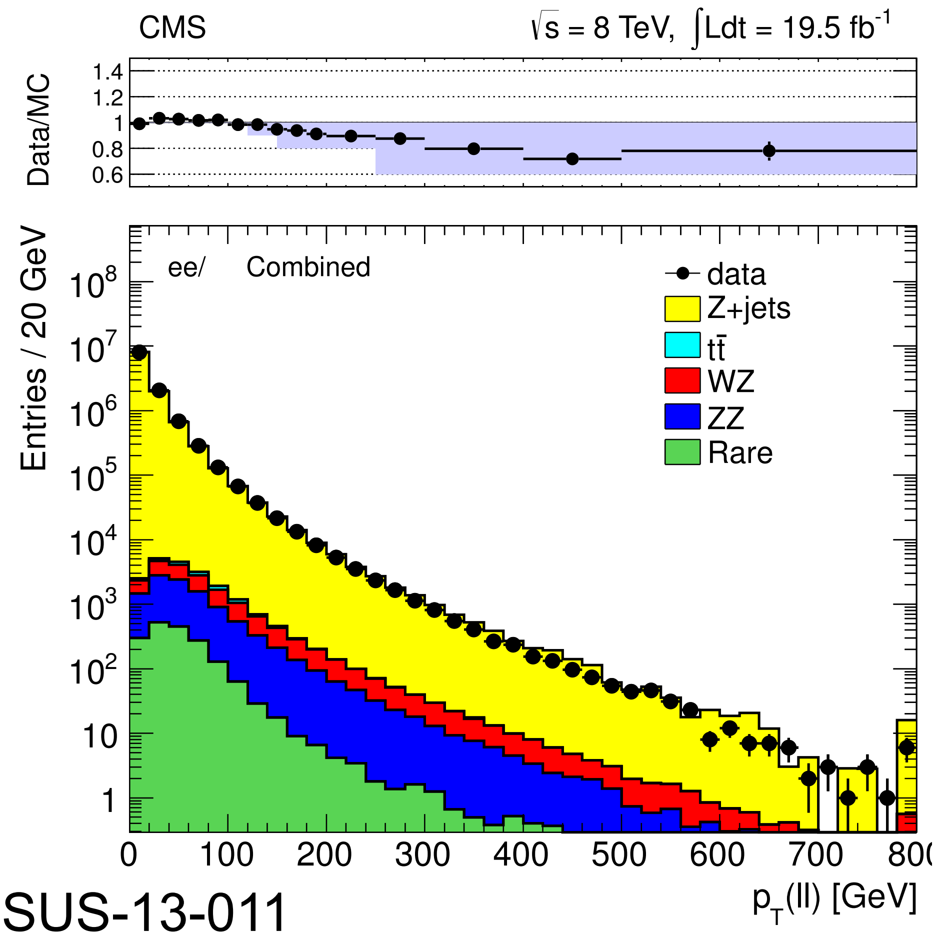
Conclusion SMP

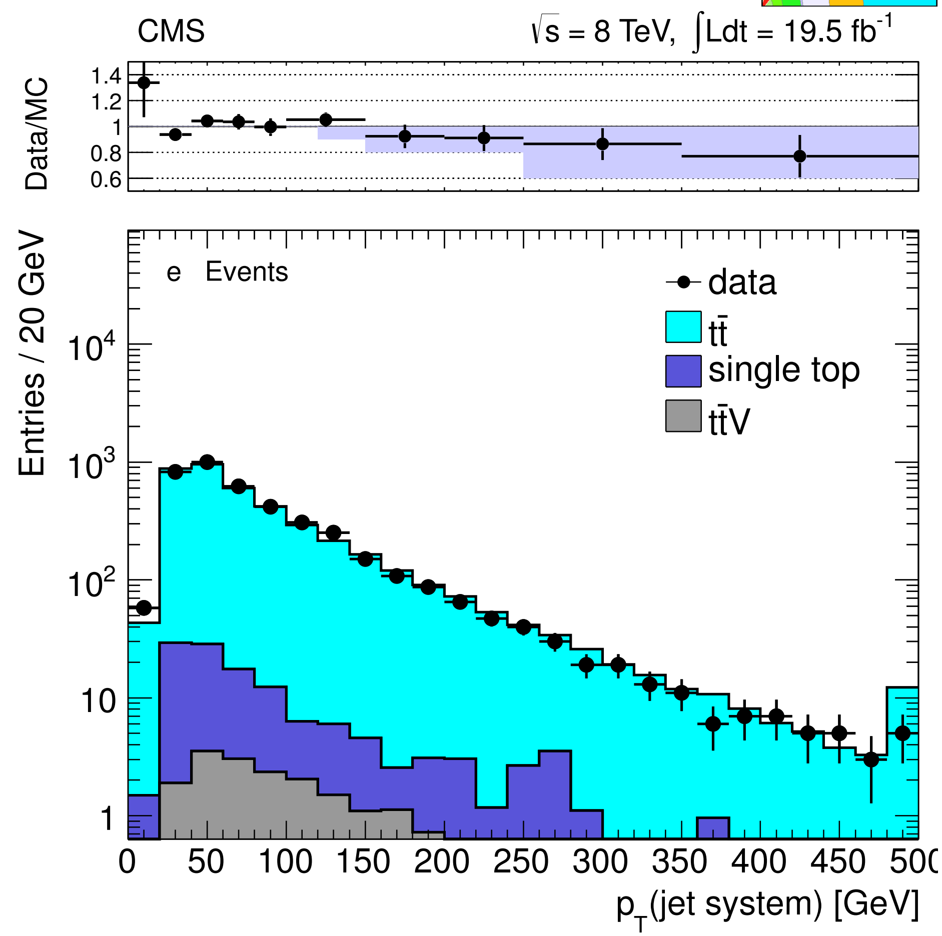
SMP-12-025



Data and Theory agree

Conclusion SUSY (I)

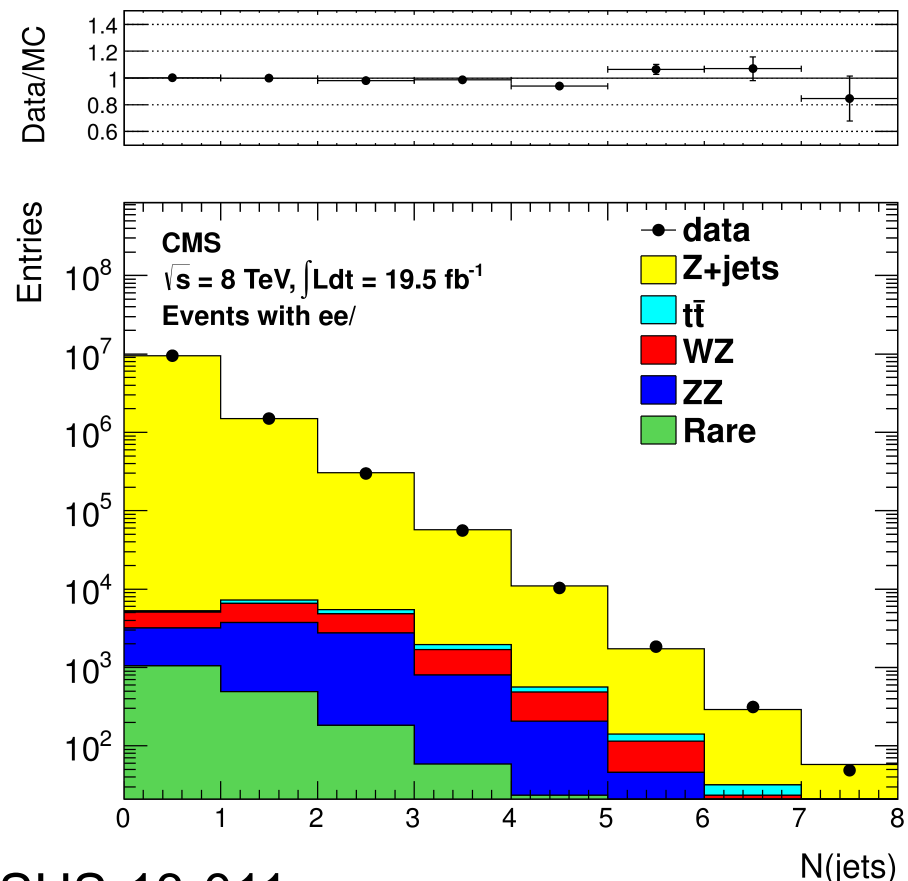


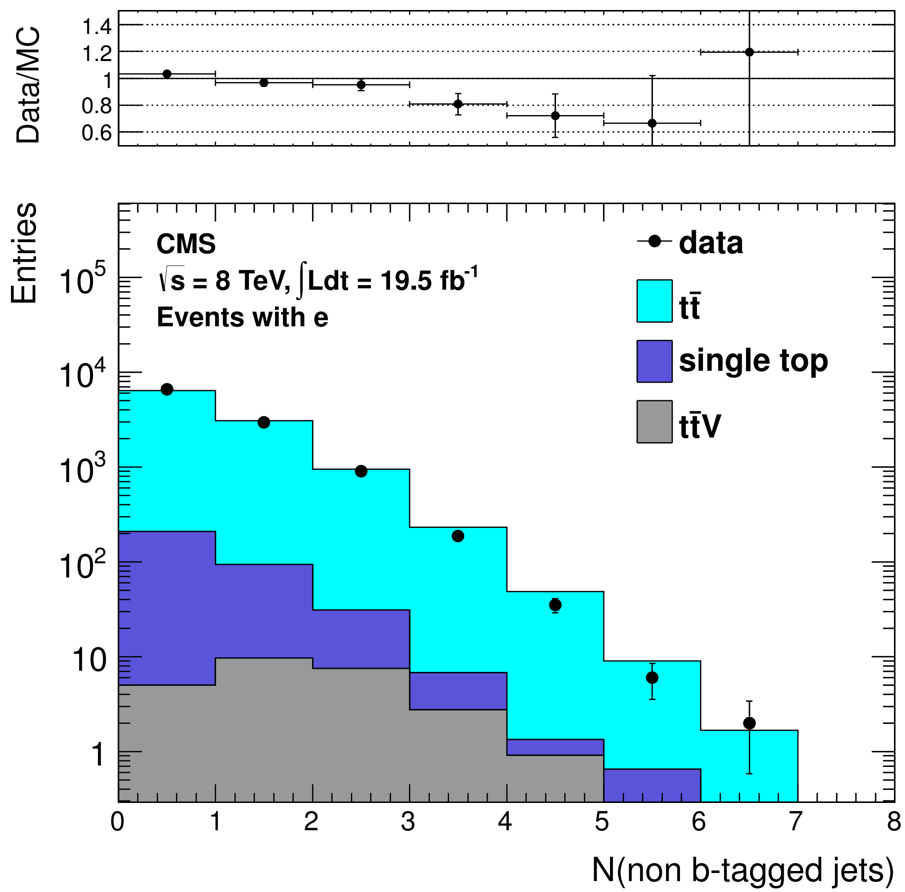


SUS-13-011

Data and MC disagree on ISR boost by ~ same factor in Z and top.

Conclusion SUSY (II)

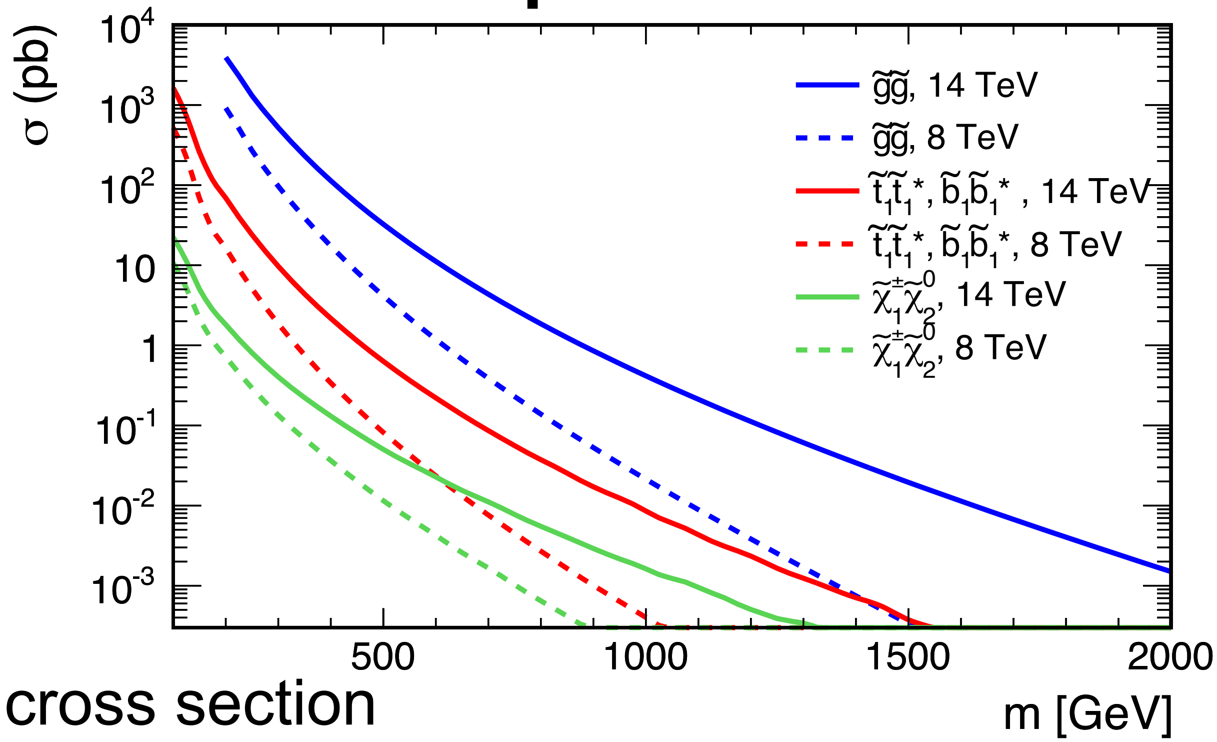




SUS-13-011

of "extra" jets agree well enough to justify boost correction from previous page.

What to expect at 13TeV



increase in cross section

1350GeV gluinos: x30
1000GeV gluinos: x20
750GeV squarks: x9
350GeV $X^{+-}X^0$: x3
top pairs: x4

Reach new territory with 1-6/fb of 13TeV luminosity

Signal grow much faster than SM bkg
-> will need data driven techniques.

Strategy for 13TeV Start-up

- No new physics expected for first ~1/fb
  - Plenty enough time and data for basic object and detector response validation using phase space explored at 8TeV.
- New physics in tails of SM phase space for which SM measurements do not exist.
  - Data driven techniques for all the dominant bkg's.
- Expect interesting results with < 5/fb of 13TeV data.

Let's review the main sources of SM bkg to (RPC) SUSY searches

Disclaimer:

This is a rather superficial review! Details differ between different searches.

“Dominant” Sources of SM bkg

- Lost leptons in 0, 1 lepton analyses
  - Mostly from W jets, ttbar
- $Z \rightarrow \nu\nu$ in 0 lepton analyses
  - Irreducible source of MET
- MET resolution tails, and catastrophic mis-measurements of jets.
  - In 0 lepton from QCD multijet
  - In Z+MET searches
  - In tails of mT, mT2 etc.
- Fake leptons
  - Only relevant in same-sign and >=3 lepton analyses

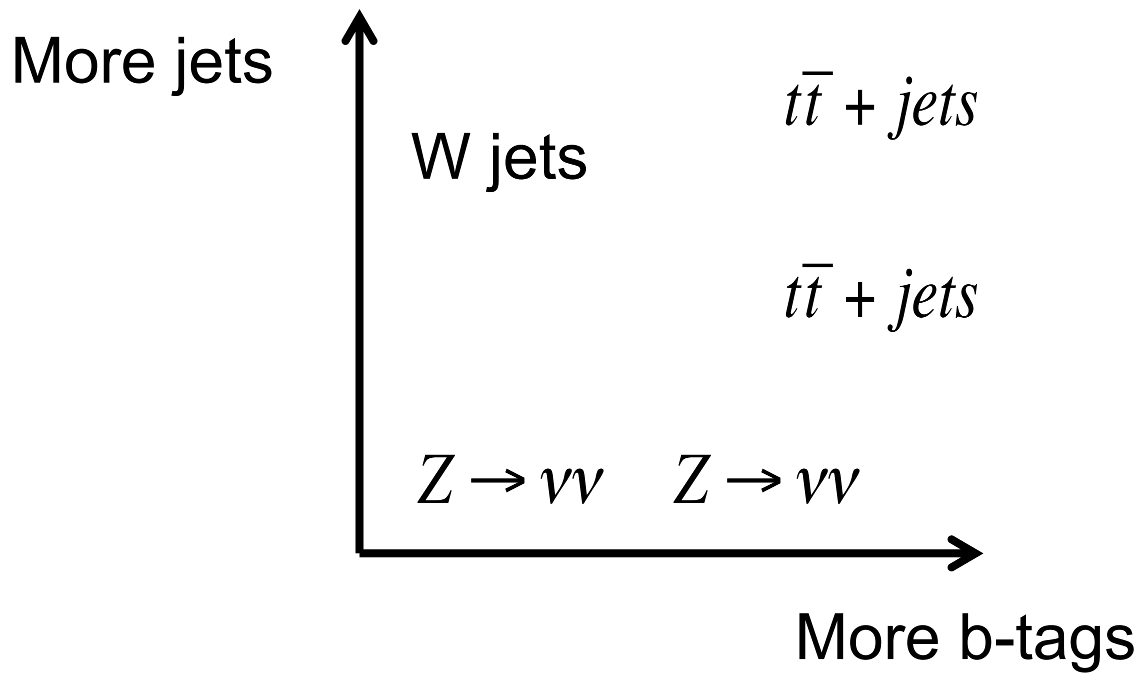
"Dominant" Sources of SM bkg

- Lost leptons in 0, 1 lepton analyses
  - Mostly from W jets, ttbar
- $Z \rightarrow \nu\nu$ in 0 lepton analyses
  - Irred
- MET r
  measu
  - In 0 l
  - In Z+MET searches
  - In tails of mT, mT2 etc.
- Fake leptons
  - Only relevant in same-sign and >=3 lepton analyses

3 out of 4 depend on instrumental effects

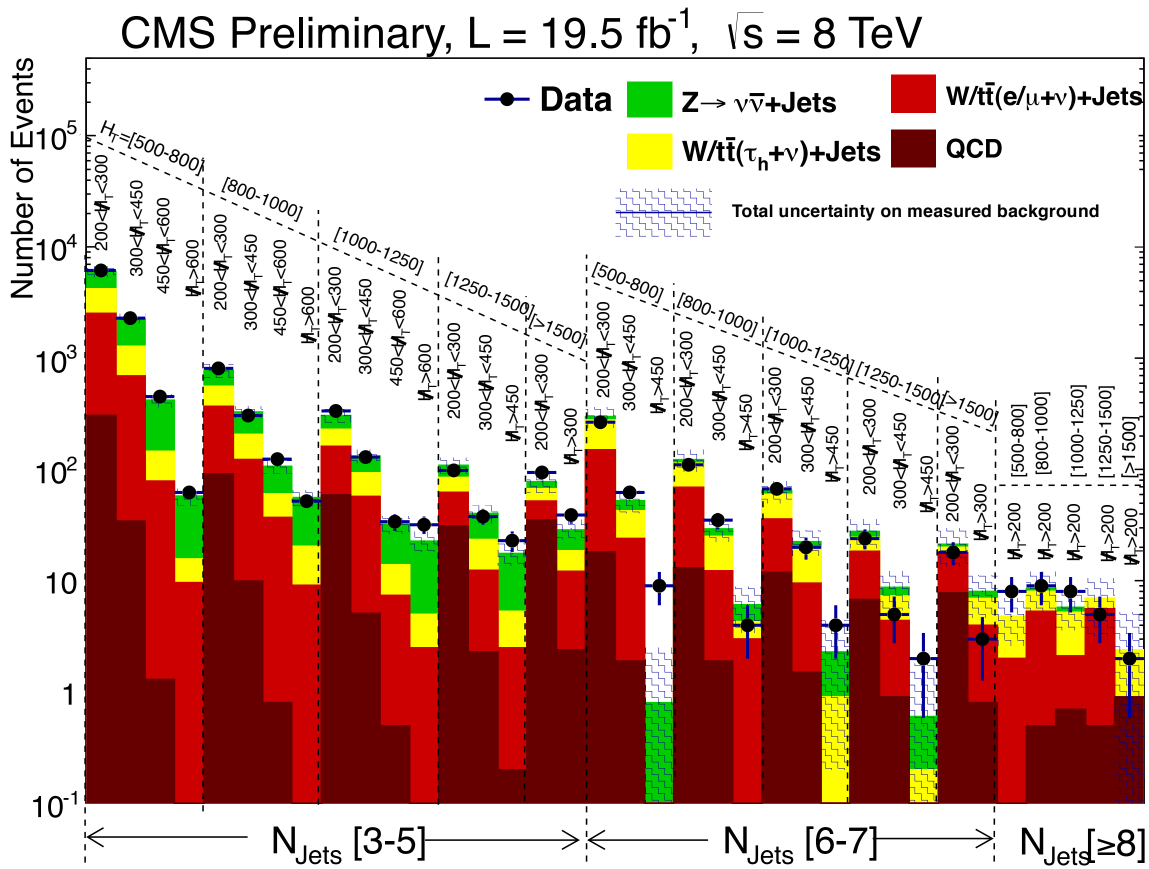
Relative Size of SM bkg's

For searches requiring large MET, H_T, m_{eff}



E.g., 3rd generation SUSY searches are dominated by top bkg except for low jet multiplicity.

QCD Multijet bkg

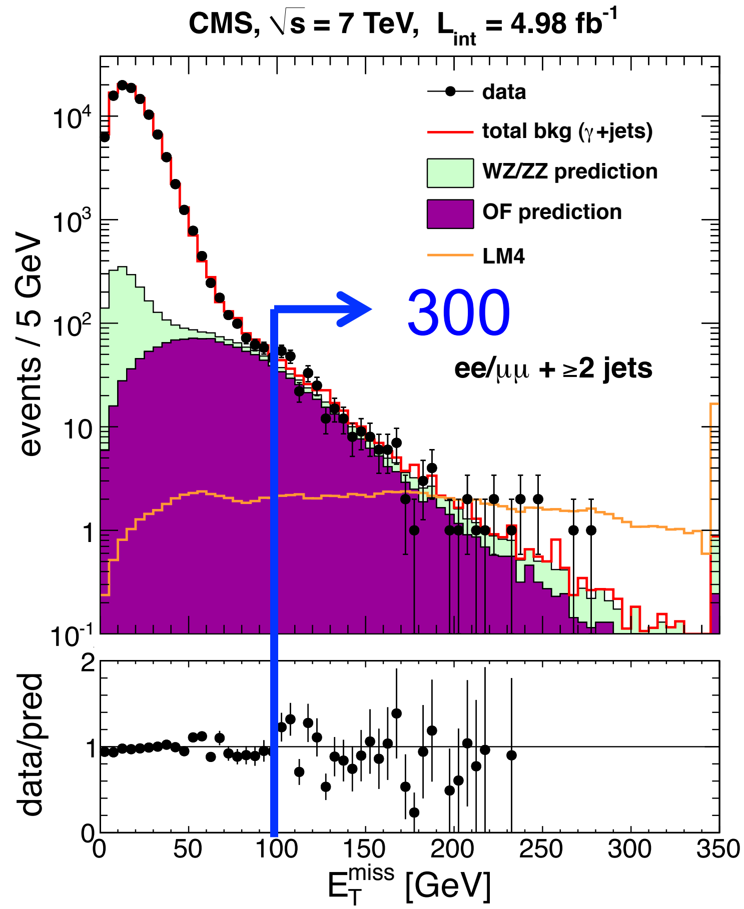


The hardest problem with QCD multijet bkg is to convince yourself that it is not a problem.

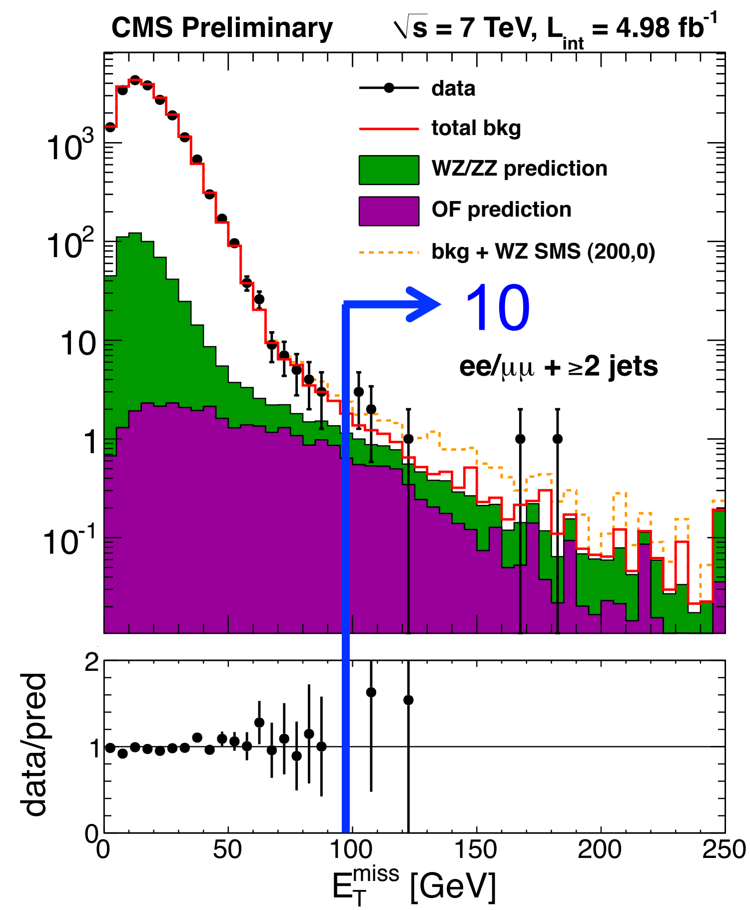
Suppressing top via b-veto

MET for Z + jets selection with and without b-veto

SUS-11-021



SUS-12-006



Selection differs by b-veto and dijet mass cut consistent with W,Z

Lost Lepton bkg

- Veto isolated e,mu
- Veto isolated track
  - Track isolation only to be sensitive to 1-prong tau decays
- Veto isolated tau
- Control region are found lepton samples
  - Get lost/found transfer ratio from mix of MC and data.

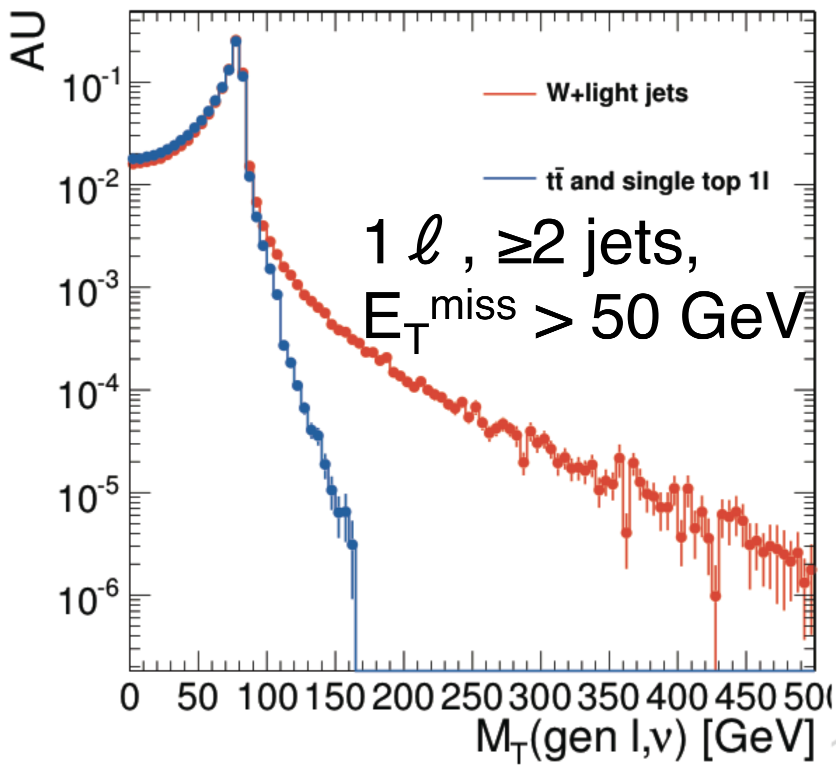
$Z \to \nu\nu$ bkg

- Measure Z to ll and/or W to lnu
- Derive transfer factor from MC
- Often non-trivial extrapolations to overcome statistical limitations.

MET resolution tails

- a not so obvious example -

- Backgrounds with $W \to \ell\nu$ have a kinematic endpoint at $M_T < M_W$
- **The M_T tail has contributions from W+jets (off-shell W production) and $t\bar{t} \to \ell$+jets (MET resolution effects only)**



Generator level

The lepton+jets bkg in the M_T tail thus depends on:

-> width of W mass

-> MET resolution

-> relative Xsection of Wjets and top in the phase space selected for the search.

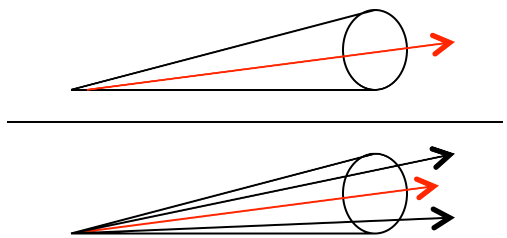
“Data Driven” Techniques

An Attempt at a simple minded Categorization

Distinguish 3 Types of bkg Estimates

- "Everything" from data
  - Only validation of technique is done with MC
  - Bias and MC stats determine syst. error
- Normalization from data
  - "transfer factor" from control to signal region is measured using MC.
  - Validation in control region determines syst. error.
- "Everything" from Monte Carlo

Fake Leptons



= # of “tightly” isolated leptons / # of “loosely” isolated leptons = “fake rate”

Fake rate then applied to kinematically identical region as signal region but with loosely isolated leptons.

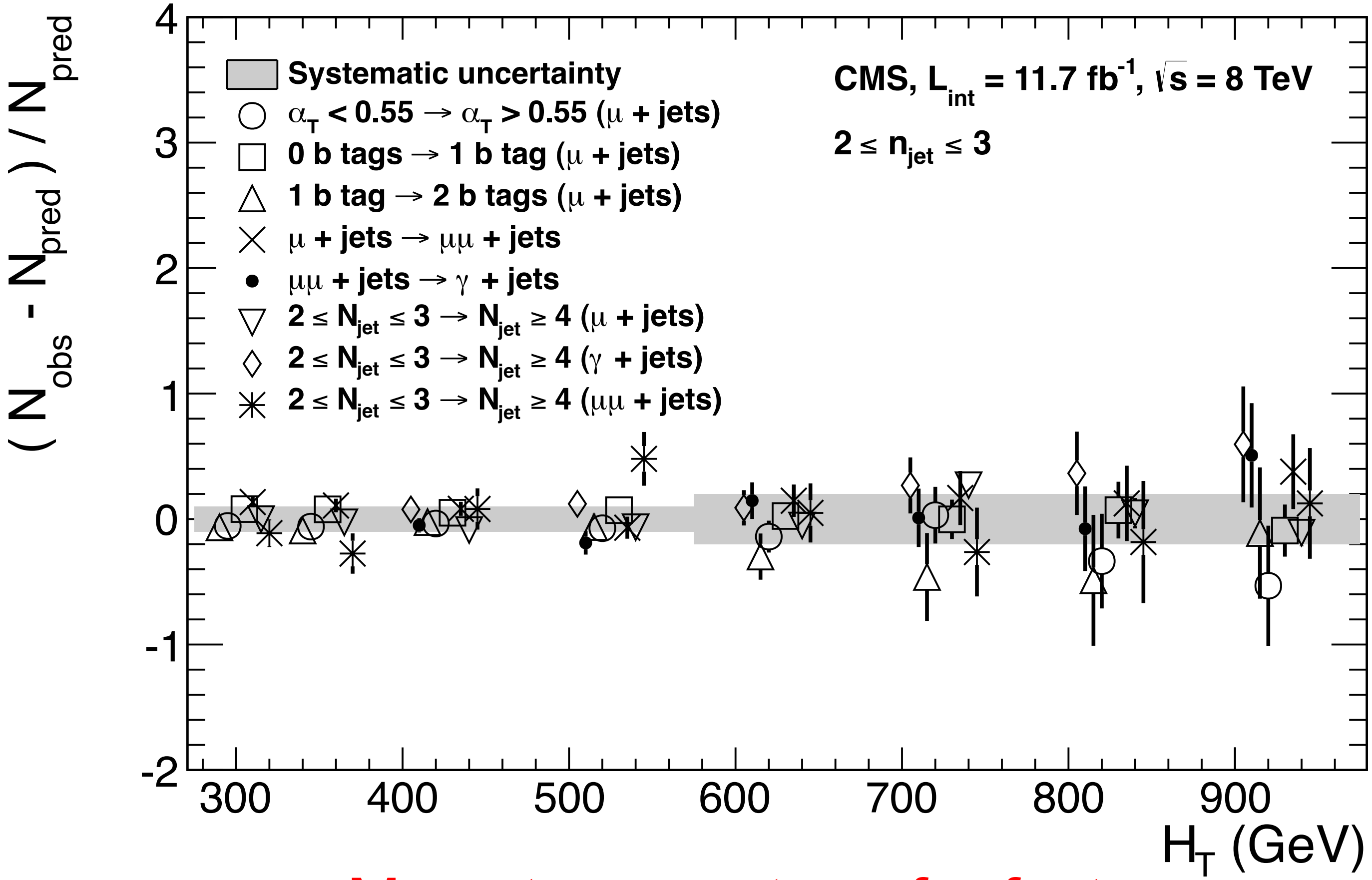
Fundamental Problems:

1. Probability for jet to become a single track depends on non-perturbative QCD.
2. “fake rate” for a given lepton pT depends strongly on parton pT that is fundamentally not measurable.

=> Difficult to reduce systematic errors

"alphaT" analysis

SUS-12-028



Many to many transfer factors.
Systematics based on overall agreement.

Same-sign dileptons

	SR0b	SR1b	SR3b
Observed events	5	8	4
Expected background events	7.5 ± 3.3	3.7 ± 1.6	3.1 ± 1.6
Expected $t\bar{t} + V$ events	0.5 ± 0.4	2.2 ± 1.0	1.7 ± 0.8
Expected diboson events	3.4 ± 1.0	0.7 ± 0.4	0.1 ± 0.1
Expected fake lepton events	3.4 ± 3.1	$0.3^{+1.1}_{-0.3}$	$0.9^{+1.4}_{-0.9}$
Expected charge mis-measurement events	0.1 ± 0.1	0.5 ± 0.2	0.4 ± 0.1

ATLAS-CONF-2013-007

	$N_{\text{b-jets}}$	
SR0b	0	$N_{\text{jets}} \geq 3$, $E_T^{\text{miss}} > 150$ GeV $m_T > 100$ GeV, $m_{\text{eff}} > 400$ GeV
SR1b	≥1	$N_{\text{jets}} \geq 3$, $E_T^{\text{miss}} > 150$ GeV $m_T > 100$ GeV, $m_{\text{eff}} > 700$ GeV
SR3b	≥3	$N_{\text{jets}} \geq 4$

"Irreducible" bkg estimated straight from MC.
MC estimate validated in data control regions

Conclusion

- “The more things change, the more they stay the same.”
- While we expect an exciting jump in sensitivity to new physics at 13TeV, the basic philosophy for estimating SM bkg’s will stay the same.
- 8TeV data analysis, including the Higgs discovery was an excellent “warm-up” for things to come.