RECASTING EXPERIMENTAL SEARCHES

Michele Papucci
LBNL & BCTP

Amherst, November 12th, 2015

BSM at the LHC

- 250-300 analyses SUSY+exotica, CMS+ATLAS, 7+8TeV during run I
- no significant deviation from the Standard Model, but incredibly extensive and valuable information to constrain the Beyond the Standard Model panorama
- Large amount of results brings new challenges in understanding consequences for beyond the Standard Model physics

BSM at the LHC

- A wide variety of searches, in principle covering most of the bases
- Results have been presented in terms of specific models and of simplified models (more on it later)
- Experimental collaborations are limited by computational resources and manpower for constraining all the BSM models out there
- → need for reinterpretation ("recasting") of experimental results outside ATLAS and CMS collaborations

Certain questions force theorists to extrapolate ("recast") experimental results into new territory

- powerful very general statements are contained in ATLAS/CMS results but not immediately available (e.g. what's the limit on particle "X" irrespective of its decay modes?)
- are there "holes" in these searches which have been left out?
- what is the relative performance of two different searches in excluding a specific model?

(often surprises are found)

Simplified models 101

- Simplified models for LHC searches are the equivalent of S,T,U,V… parameters for EW precision data
  - simple models involving only few particles with simple decay modes
  - Idea: break up a full model in terms of simplified models. Full event yield of a model in a search → sum of yields of simplified models = sum of $N_{ev} = L * \sigma *$ BR's $* \varepsilon$.
  - σ and BR's: fast to compute
  - ε: time consuming and needed as function of particle masses → Compute once & reuse for many different models.
- Very powerful method if enough simplified models are available
- Too few simplified models presented by the experimental collaborations (resources limitations) → theorists step in to fill in the gaps → recasting!

Recasting experimental analyses 101

Take search X setting limits for model A

Write code to mock up search X
(not enough info → introduce approximations)

Generate events for model A, use them with mocked-up analysis, compare results with published experimental results

Extrapolation!!

Use mocked-up analysis with model B

Extract approximate limits of search X for model B

Validation (most time consuming part)

Repeat for many many analyses…

The bottomline:

Recasting experimental analyses has been proven successful by 100+ papers…

… but the question about extrapolations is always lurking.
(Few examples of too naive extrapolations)

In principio…

- Until few years ago:
  - PGS4/Delphes for fast detector simulation, but needed to be tuned to ATLAS/CMS
  - Each "practitioner" had her/his own implementation+validation of analyses in some form
  - Rivet: database of unfolded SM measurements for MonteCarlo tuning
  - Recast proposal: protocol to submit BSM event files to experiments for investigation during their spare time

Recasting accessible only to few "practitioners"

…today

Generate & process MC events

CheckMATE, Atom,

Drees et al. 1312.2591

I.W.Kim, M.P., K.Sakurai, A.Weiler, to be released soon

MadAnalysis, …

Conte et al, 1206.1599, 1405.3982, 1407.3278

Use simplified models and spectrum and BR's information from SLHA file

Fastlim, SmodelS, …

M.P., K.Sakurai, A.Weiler, L.Zeune, 1402.0492

Kraml et al. 1412.1745, 1312.4175

+ Recast soon as web interface to (some of) these tools

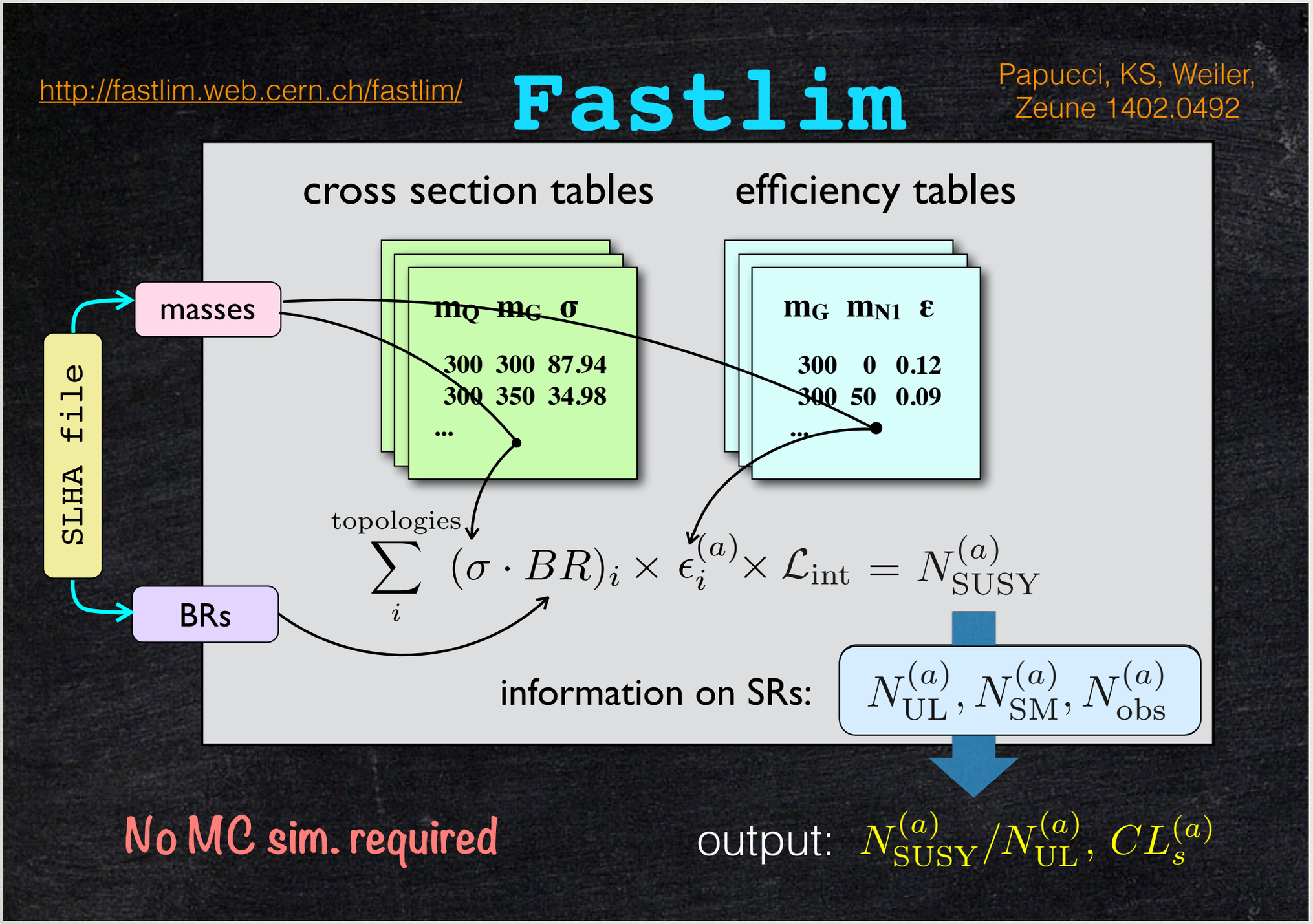
…today: prompt vs. non-prompt

- Both recasting and usage simplified models increasingly straightforward for prompt searches
- Significantly less developed for non-prompt searches
  - No available tools, everyone write her/his own code
  - In some cases event generation requires hacking (dark showers, hadronization, …)

…today: prompt vs. non-prompt

- Simplified model results, when available, are present only for few points in parameter space (1D results as function of lifetime) → recasting needed!
- Less amount of information available for validation of non-prompt analyses (extrapolations??)
- Nevertheless, a few recasting works are out there (see e.g. talks of Cui and Tweedie)

Simplified models are useful to quickly "recast" results in more complete models



K.Sakurai, MC4BSM talk

Using simplified models

- SUSY Les Houches input file (SLHA) restricts usage to SUSY models (for the moment, due to lack of a standard for x-section info, workarounds for non-SUSY models in the pipeline)
- Limits on single point in model parameter space can be evaluated in O(1 sec) → amenable for large scans
- σ and ε tables are pre-computed
- Can use ε from:
  - Published experimental results on simplified models
  - Recasting using CheckMATE, Atom, …
- $\sigma > 0$, $\varepsilon \geq 0$: missing search/topology reduces event yield → bounds always conservative!!

Using simplified models

- Shortcomings:
  - neglected:
    - interference, finite widths: negligible in weakly coupled models
    - production mechanism variations, chirality and spin correl': O(20%) in most of the cases Edelhauser et al '14, Sonneveld '15, Wang et al '13, …
  - complexity for generating ε tables:
    - limit topologies to 2-3 steps cascades

For other cases, other tools need to be used…

Simplified models for long-lived particles

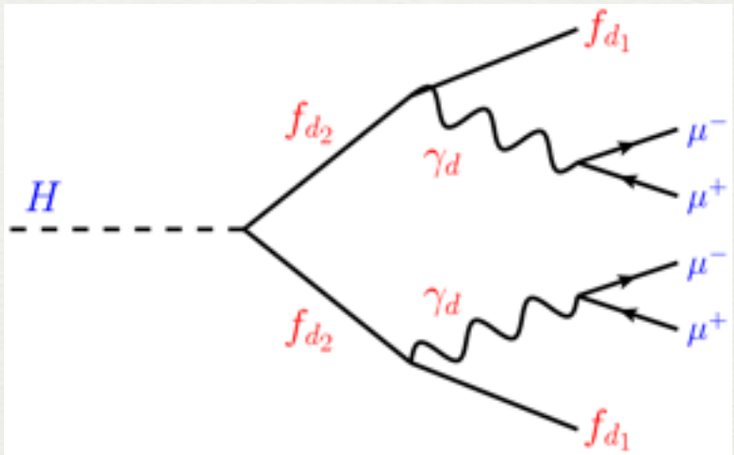
- Naively same paradigm can be utilized for long-lived searches:

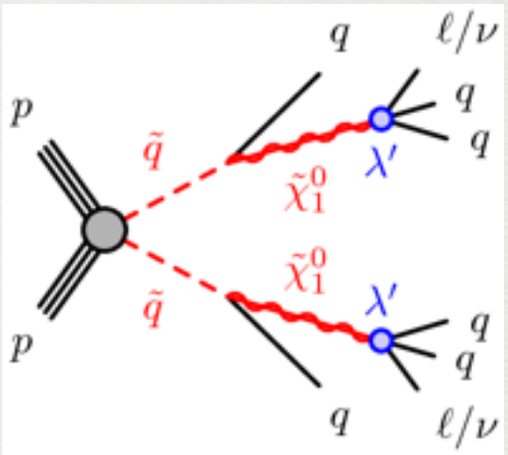
$$\epsilon(m_1, m_2, \ldots) \rightarrow \epsilon(m_1, m_2, \ldots, c\tau)$$

- OK for events with few "well-isolated" long-lived particles (SUSY RPV, "sparse" lepton jets, …)
- introducing lifetime may reduce maximum depth of cascade (complexity)

Simplified models for long-lived particles

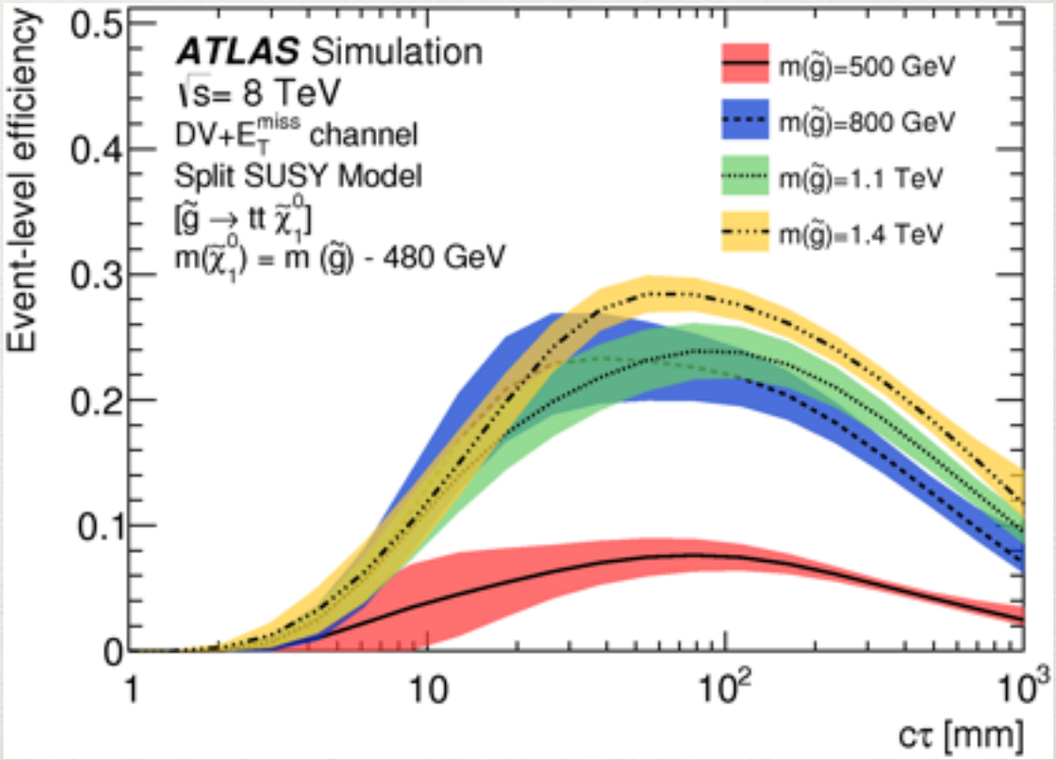
- Various simplified topologies already considered by experiments:





- Results as function of cτ for few mass points

- No full efficiencies for any topologies → need to recast almost everything



Simplified models for long-lived particles

- For hidden valleys with dark forces producing higher multiplicities / FSR radiation / showers parameters easily proliferate

$$\epsilon(m_1, m_2, \ldots) \rightarrow \epsilon(m_1, m_2, \ldots, c\tau, \alpha_D, \Lambda_D, \ldots)$$

- large dimensionality: unless degeneracies of parameters and/or efficiencies factorize, production of efficiency maps for simplified topologies becomes quickly intractable

- Recasting only option in these cases? (less accessible to broader audience: exactly the cases where at the moment more tool hacking is required :( )

Recasting & detectors

- Recasting long-lived searches requires new “object” definitions
- In current recasting tools object are defined via combination of:
  - event-dependent information (e.g. isolation)
  - event-independent truth-vs-detector corrections (e.g. tagging efficiencies, smearing, …) to bring results within O(10-20%) for signal events

Recasting & detectors

Event dependent, truth-level info

- E.g., hadronic taus recipe:
  - take jet
  - look at event decay history to see if any parent of particle in jet was a τ
  - count charged particles inside a smaller cone in the jet to define 1-,(2-,)3-prong
  - apply efficiency/rejection for specific prong-ness as function of pT, η of jet (adapted from τ commissioning paper, validated against few searches/SM measurements)

Event independent info

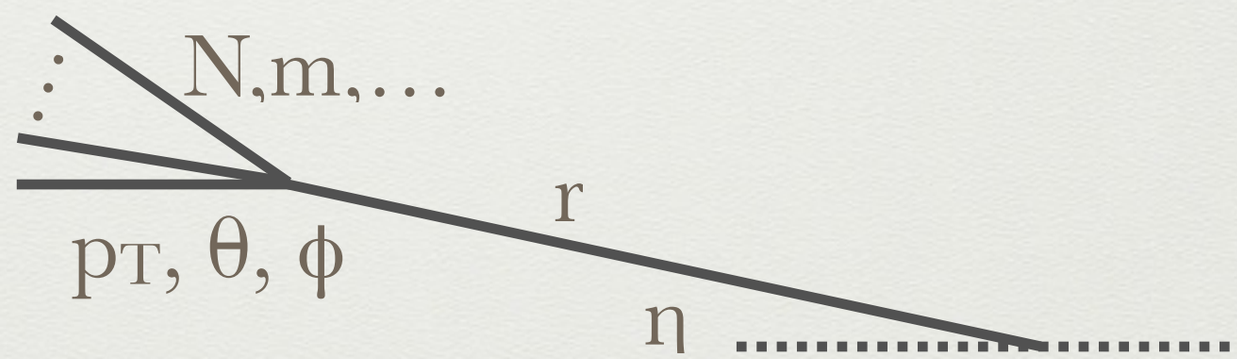
  - implicit assumption: efficiency is uncorrelated among taus in same event (reasonable bc if too close they would likely be merged in same jet both in simulation and real-life)

Recasting for long-lived particles

- Similar procedure could be applied to new “objects” in long-lived searches
  - detector geometry (regions of ID, ECAL, HCAL, muon) easily taken into account
  - easy to take into account properties used in selection, such as impact parameters, kinematic properties and multiplicities of decay products, EM vs. hadronic energy depositions (roughly), …
  - then in principle correct for the discrepancies between truth- and detector- level…

Recasting for long-lived particles

- Many open questions, hard to extrapolate from currently available public info:
  - How "isolated" these objects have to be for this procedure to work?
  - Which parameters are the efficiencies function of? Do they factorize in indep. functions with less arguments? (not feasible to use efficiencies depending on more than 3(-4) correlated parameters…)



  - Can pile-up effects be mostly lumped into these efficiency/smearing functions as in the case of prompt objects?
  - …

Conclusions

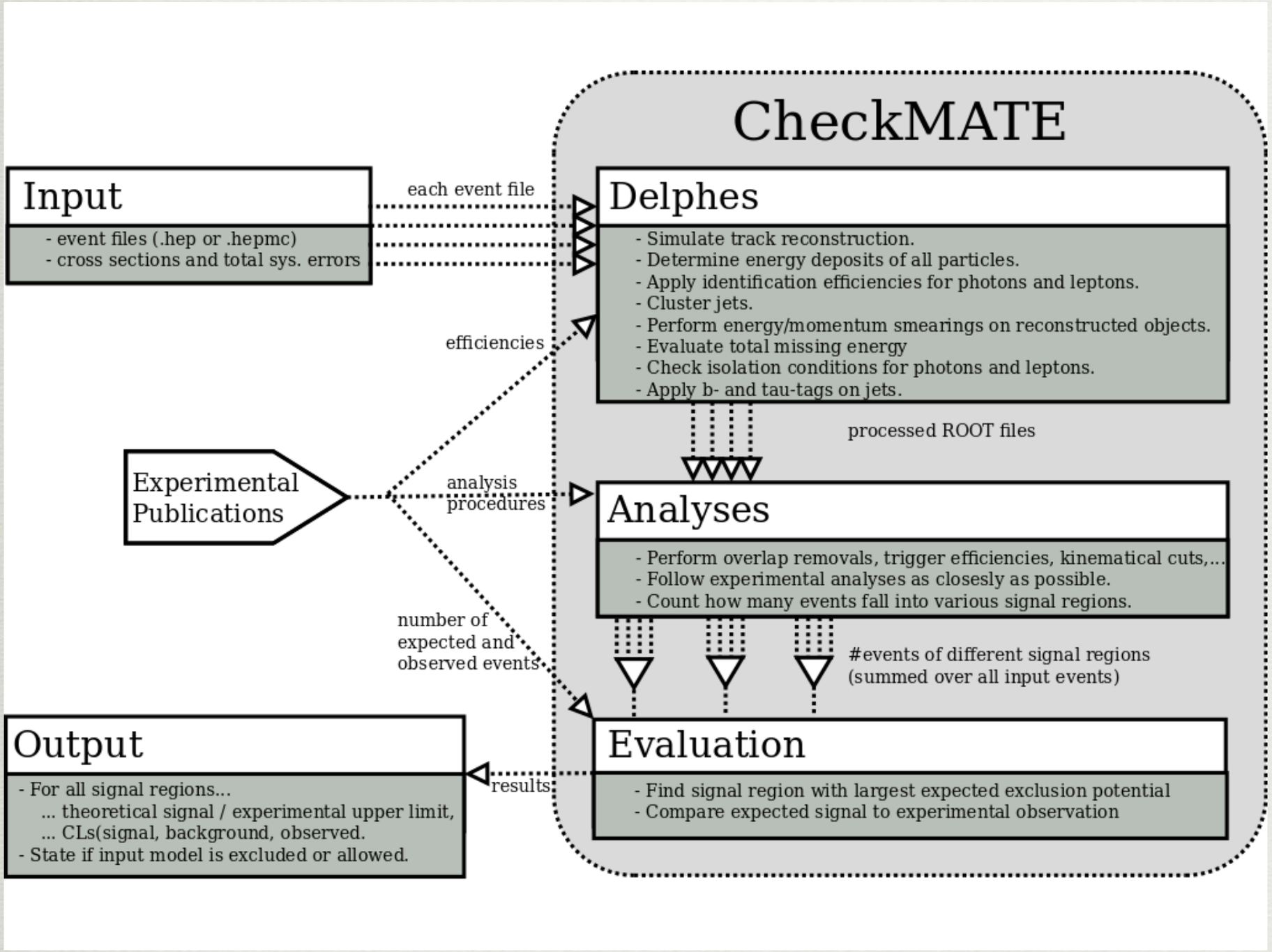
- Recasting analyses can provide feedback to the experimental program by extracting more model-independent results, highlighting "holes" in searches, evaluating the relative strengths of different searches
- For prompt searches:
  - Mature tools for recasting searches for prompt objects using conventional reconstructed objects (leptons, jets, photons, Missing ET, …)
  - Simplified model approach useful for large parameter scans
  - Agreement to release sufficient information to allow recasting helps with analyses' validation

Conclusions

- For long lived searches this program is less mature:
  - No tools for recasting searches -> new code needs to be written and limit efforts to the realm of few practitioners
  - Unclear how further one can push the simplified topologies approach
  - Further work needed to understand how to incorporate detector effects in current tools
  - Further work needed to understand which kind of information is necessary from experiment to allow recasting (implementation and validation)

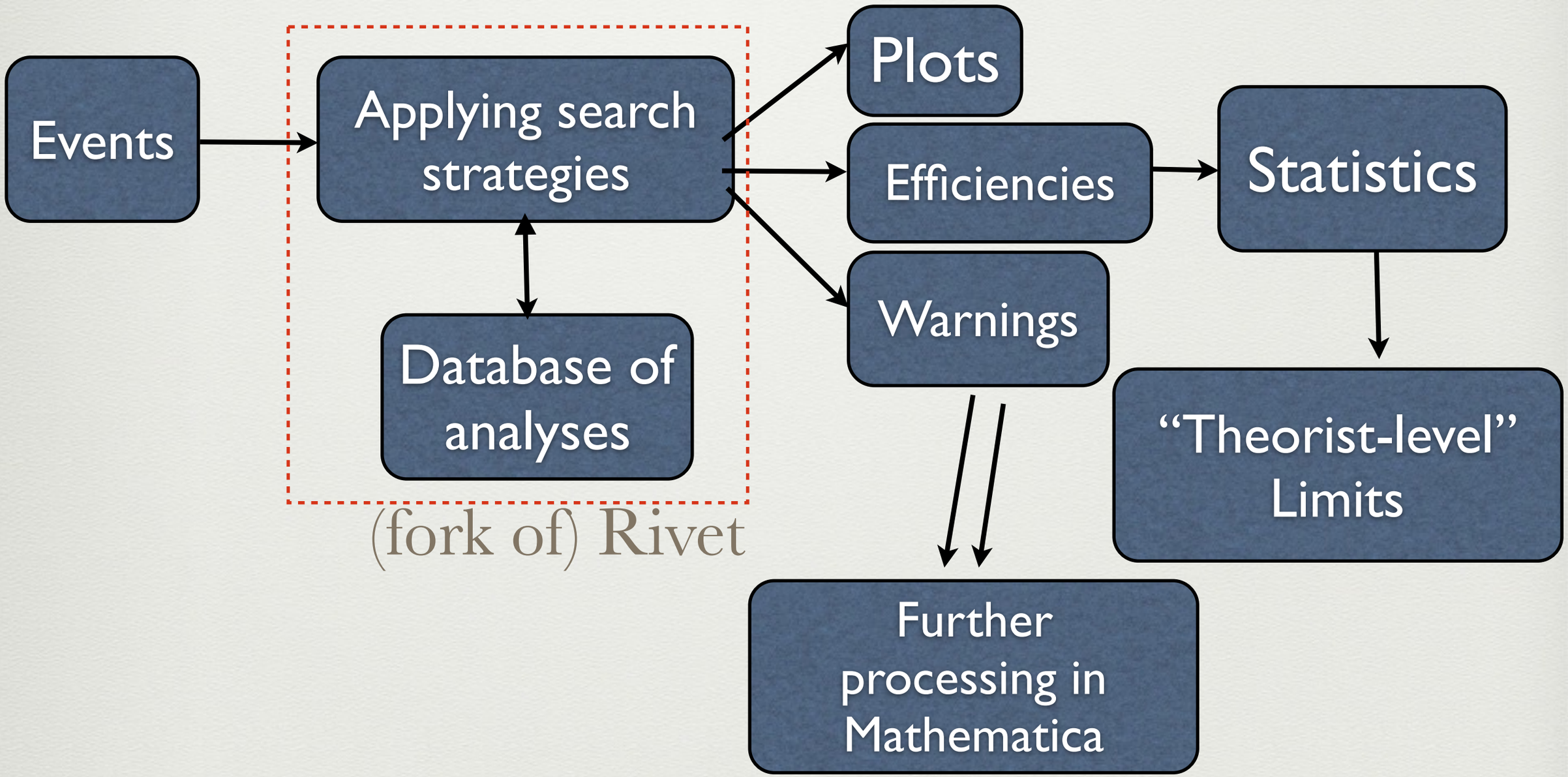
Backup

CheckMATE in a nutshell…



J.S.Kim, talk at MC4BSM 2015

Atom (Automatic Test of Models) in a nutshell…



Atom vs CheckMATE

- Built around Delphes
- Uses re-tuned Delphes for reco
- O(30) analyses
- Tools for helping implementing new analyses
- Output: limits + ROOT file from Delphes

- Fork of Rivet (~backward compatible for analyses)
- Uses truth + eff + smear for reco (via simple param cards)
- 100+ analyses
- Tools for helping implementing new analyses
- Tool for automatic validation
- Warnings of potential extrapolation problems
- Output: limits, distribution plots, warnings