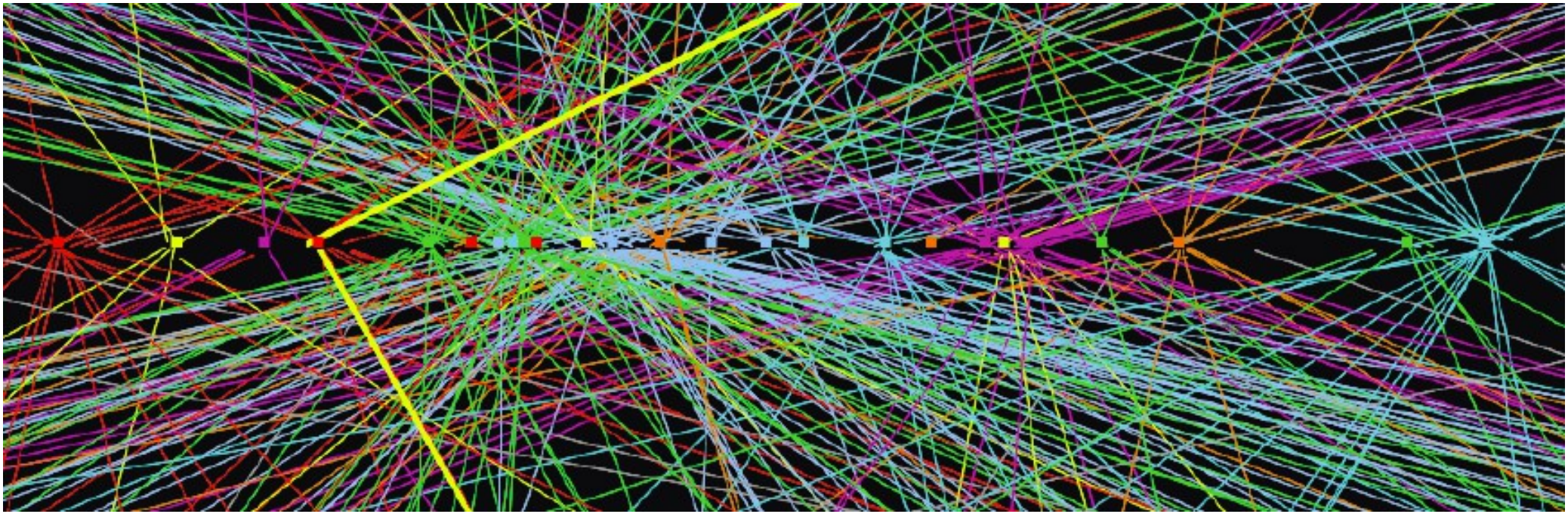
Simulation with Data Overlay

Andy Haas (NYU)

Monte Carlo in ATLAS Tutorial
September 28, 2015



https://indico.cern.ch/event/440423/other-view?view=standard

NEW YORK UNIVERSITY

Introduction

- In addition to hard interaction:
  - Pileup from other collisions in current and surrounding bunch crossings
  - Detector noise
  - Cosmics*, Beam-gas*, Beam-halo*, Cavern bkgd.*, ...

Option 1:
"Pileup MC" (current default)

Simulate all processes in MC* and mix together in proper ratios with realistic timing

Option 2:
"Overlay MC"

Simulate only hard interaction in MC and overlay a "random" data event to include all backgrounds**

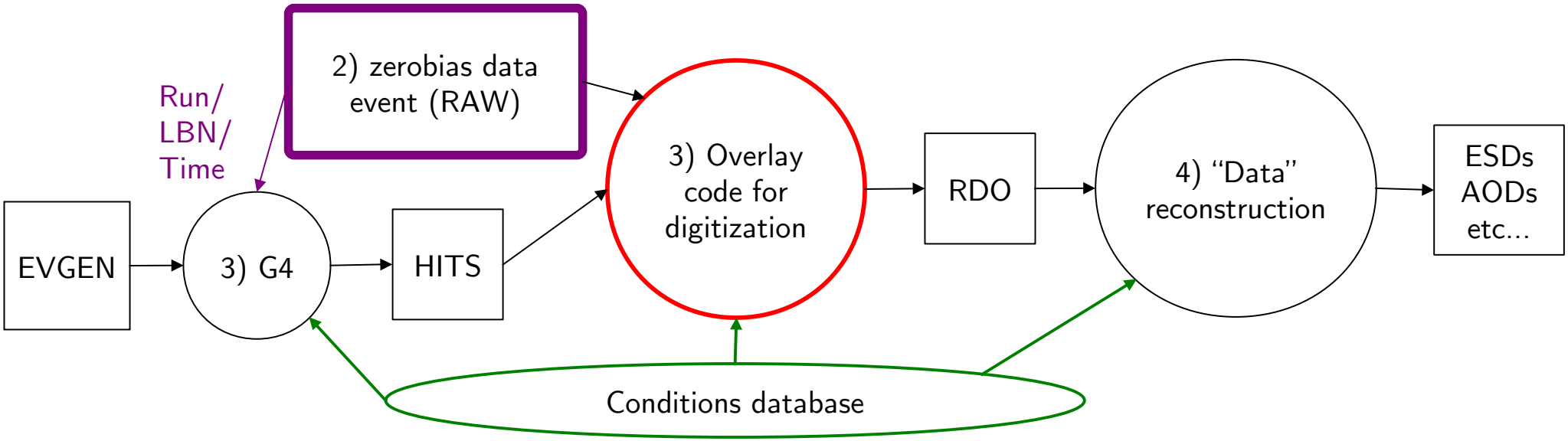
- Why bother to simulate all the backgrounds – just use data!
- Turns out it's not quite so easy - have to mix MC with data...

*not currently included in pileup MC

**Statistics of rare background events, such as high-energy calorimeter deposits from beam-halo, will be very poor. We assume the user will trigger on a signature of the *signal* MC event.

MC+Data Overlay Steps

1. Define data period to simulate (e.g. GRL AllGood_2012.xml), and select "random" (RAW) "zerobias" bkgd. data events from this period
2. Simulate hard-scatter Geant4 events with conditions matching each selected data event (beamspot, alignments, dead modules, etc.)
3. Overlay each zerobias data event with matching G4 event at the detector channel level, then digitize combined signals
4. Reconstruct the combined event as data

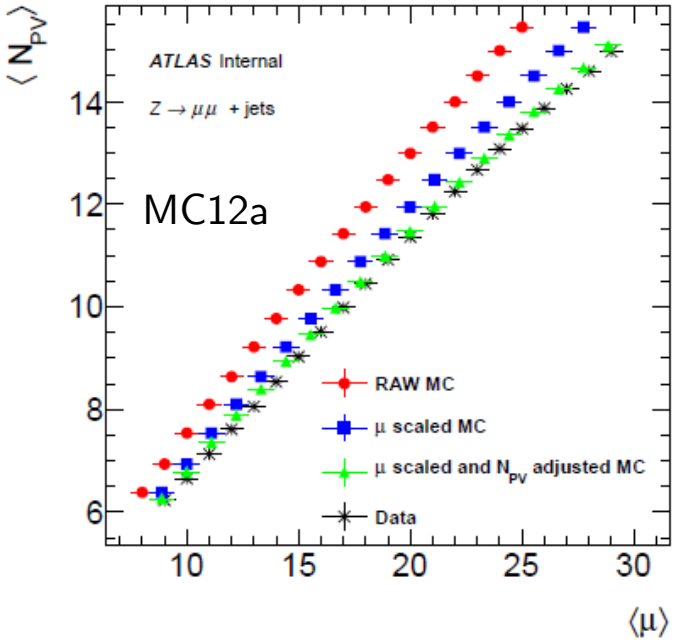


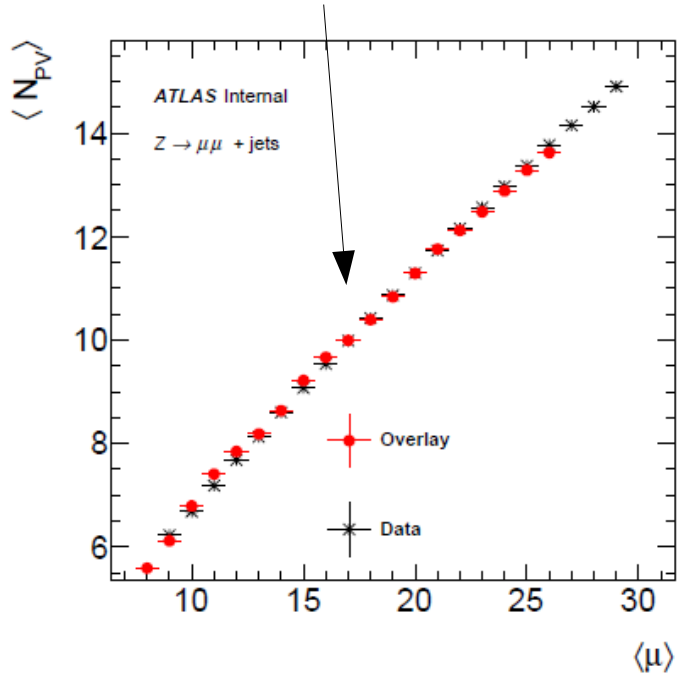
Pileup Simulation vs. Data Overlay

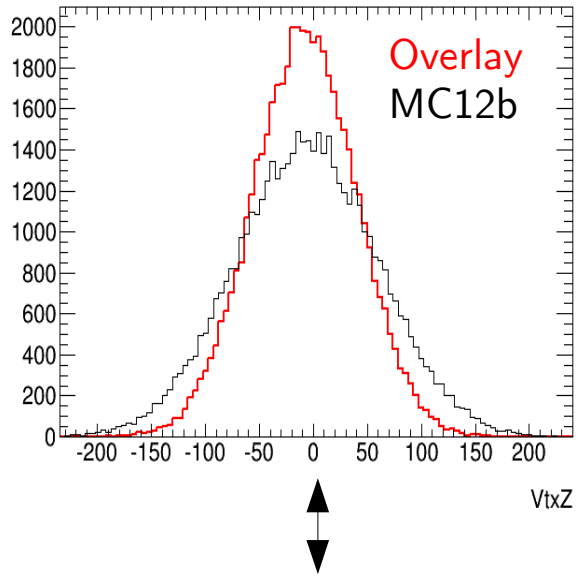
- Drawbacks compared to pileup:
  - Less accurate when combining overlapping background and signal on the same channel for some subdetectors (e.g. silicon) since zerobias data contains less information than background HITS
  - Background reconstructed with MC conditions in a few places, so slight resolution degradation for background tracks and muons
  - Potential Geant4 geometry overlaps when using data alignments
  - Harder to simulate future high luminosity (multiple overlay possible...)
  - More challenging to produce – lots of steps, DB access, ...
  - Don't have the background truth information – it's data!
- Overlay advantages:
  - Real pileup data events – no generator tuning
  - NVtx and mu match data – no event weighting
  - Realistic mix of BCID variation, in-time/out-of-time pileup
  - True detector noise, occupancy – including cavern background
  - Conditions (beamspot, dead channels, etc.) from data
  - Faster (and less memory) at high luminosity than standard pileup digi

Pileup vs. Overlay

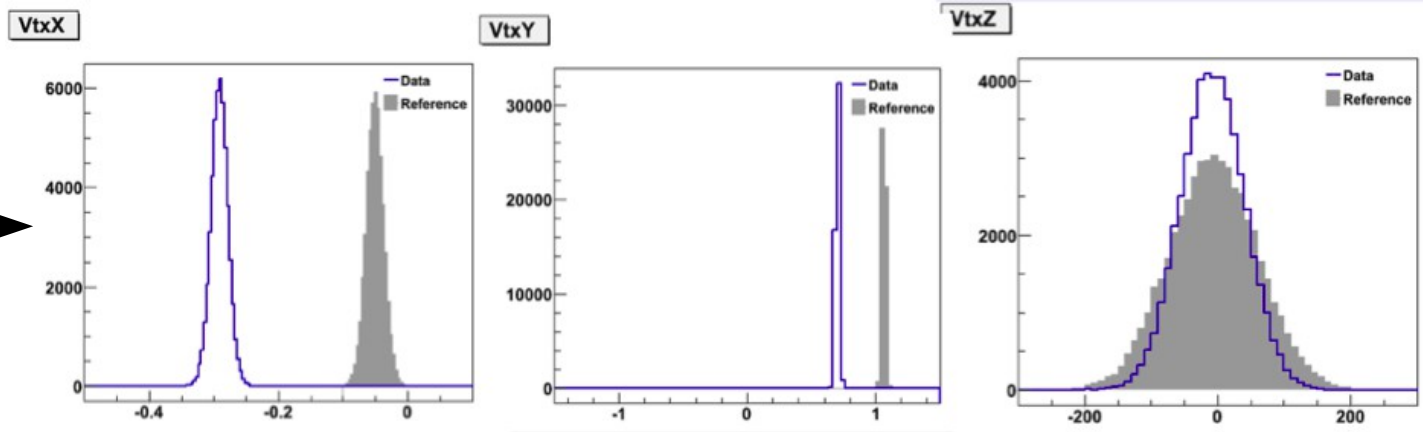
- Overlay does a good job of modeling the primary vertex and luminosity distributions seen in data





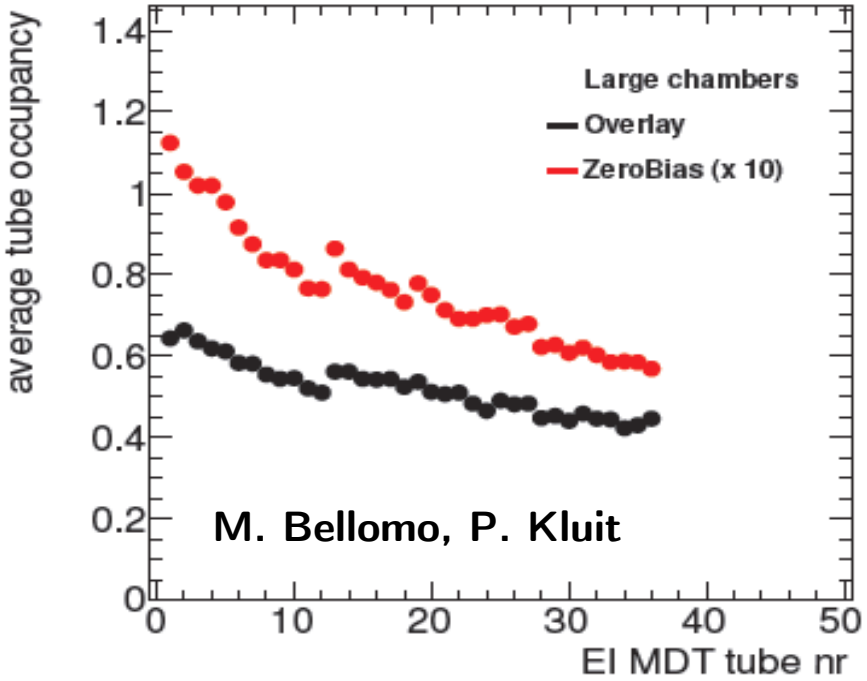


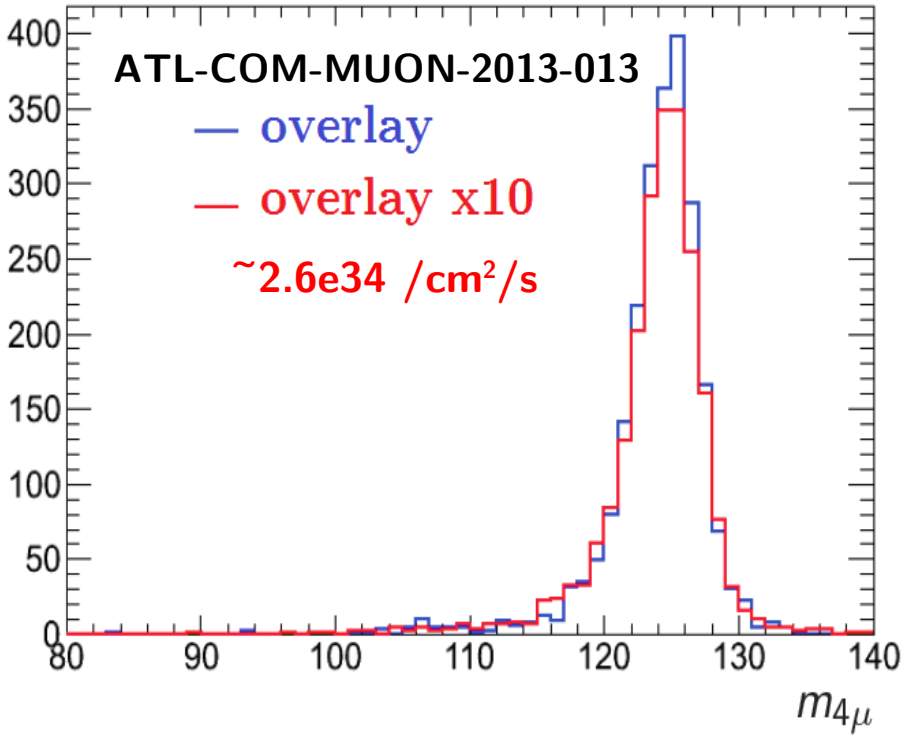
Vertex position also nicely modeled via run-dependent conditions



ATLAS Upgrade Studies

- Cavern background is nicely modeled in overlay
- Overlay *multiple* zerobias data events to simulate higher luminosity!
  - Noise is double-counted, but negligible in the muon system
  - Only implemented for the muon system so far...
- Saturation of MDT EI tubes confirmed → 4% loss for H→ZZ→4μ
  - Helped to motivate New Small Wheel Upgrade





Conclusions

- Overlay is an alternate method for including backgrounds in MC
  - Used for some performance studies*, some pp physics analyses*, HI physics analyses*, and some detector upgrade studies
  - *Chris will discuss these next!
- Lots of work done to support overlay in the official production system
  - You can request central production of overlay MC samples, just like for standard MC – but not the default choice for most at the moment
  - 2015 zerobias data will be "skimmed" soon – will then be ready to make MC15 overlay samples
- You should consider overlay MC if you are concerned that pileup modeling will have a large impact on your study/analysis
  - Analyses sensitive to pileup and/or cavern background should benefit!
  - Comparisons of data to overlay and pileup simulation will help us improve them, by better understanding the sources of mis-modelings