



Matrix Element Monte Carlo generators: dalla teoria all'utilizzo sperimentale ad Atlas e CMS

la Parte

Patrizia Azzi - INFN Padova
MCWS4LHC @Frascati 27/2/2006

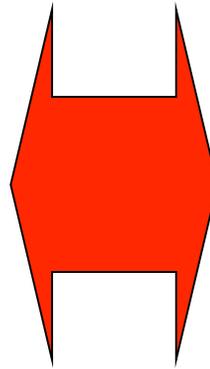
Goal

TEORICI

Stato dell'arte

Ultime novita'

Aiuto nella comprensione
delle problematiche



ESPERIMENTI

Specifiche funzionalita'

Descrizione delle modalita'
di utilizzo negli exp

Costruzione di una sinergia fra i teorici e gli sperimentali di CMS e ATLAS italiani.

Esperienza gia' provata al Tevatron sia a livello italiano che a livello di esperimento: sostanziale miglioramento nella qualita' dei campioni MC

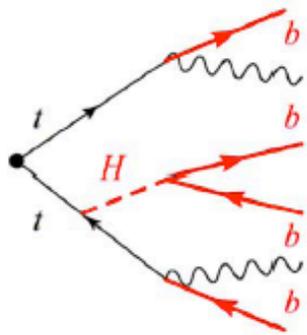
Patrizia Azzi @ MCWS 27/2/2006

Dal Tevatron a LHC

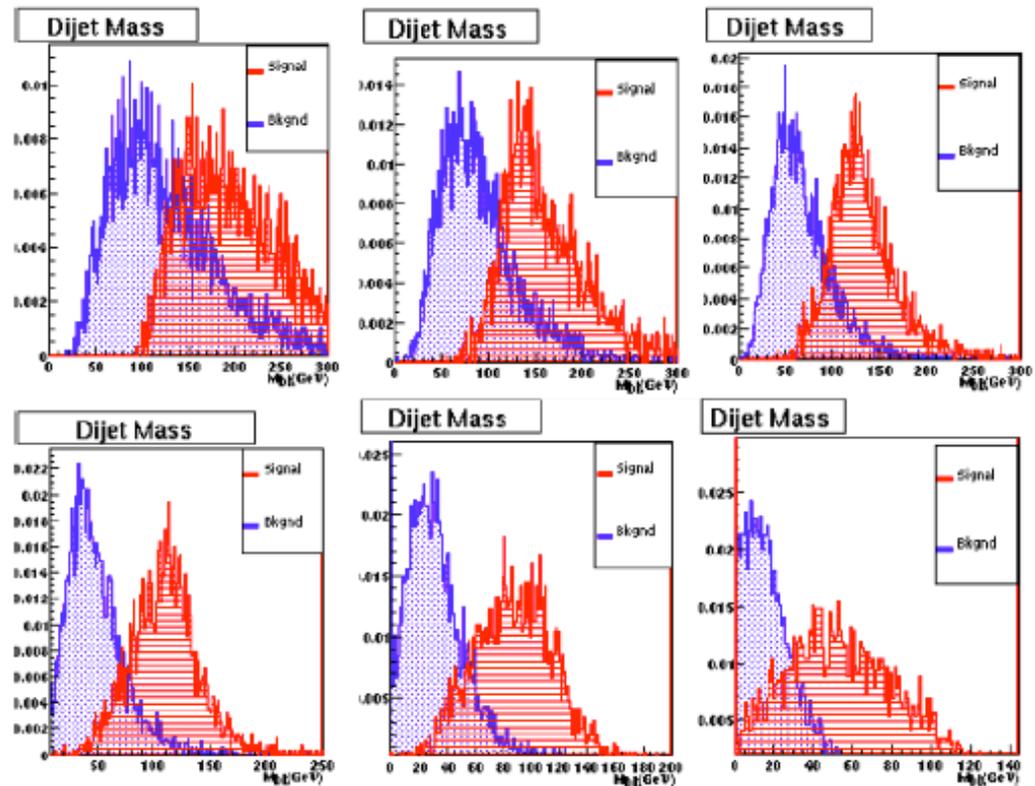
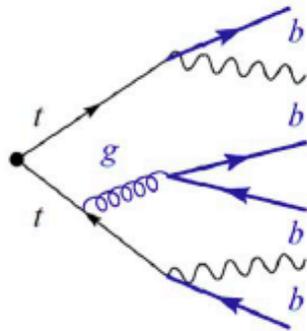
- Confronto e tuning con dati del Tevatron e' essenziale (*vedere talk di Andrea Messina*)
- Ma non basta! La realta' sperimentale sara' comunque ancora diversa a LHC:
 - Molti processi che al Tevatron sono dovuti in modo dominante a processi q-qbar a LHC sono dovuti a produzione tipo glu-glu (es. top): **diverse caratteristiche**
 - L'alta energia a disposizione apre molti nuovi processi iniziati da gluoni che non sono visibili al Tevatron: **nuovi fondi**
- Metodi di analisi e selezione "collaudati" al Tevatron possono rivelarsi inadatti alla nuova situazione sperimentale
- Servono studi di simulazione per comprendere la fisica dei segnali e dei fondi a LHC:
 - sessione ME: descrizione piu' tecnica e pratica di come, cosa serve generare e come validare i campioni
 - sessione Standard Model e BSM: approccio basato sull'analisi (*vedere talk di Marina Cobal*)

ttH (tt dileptonico) al Tevatron

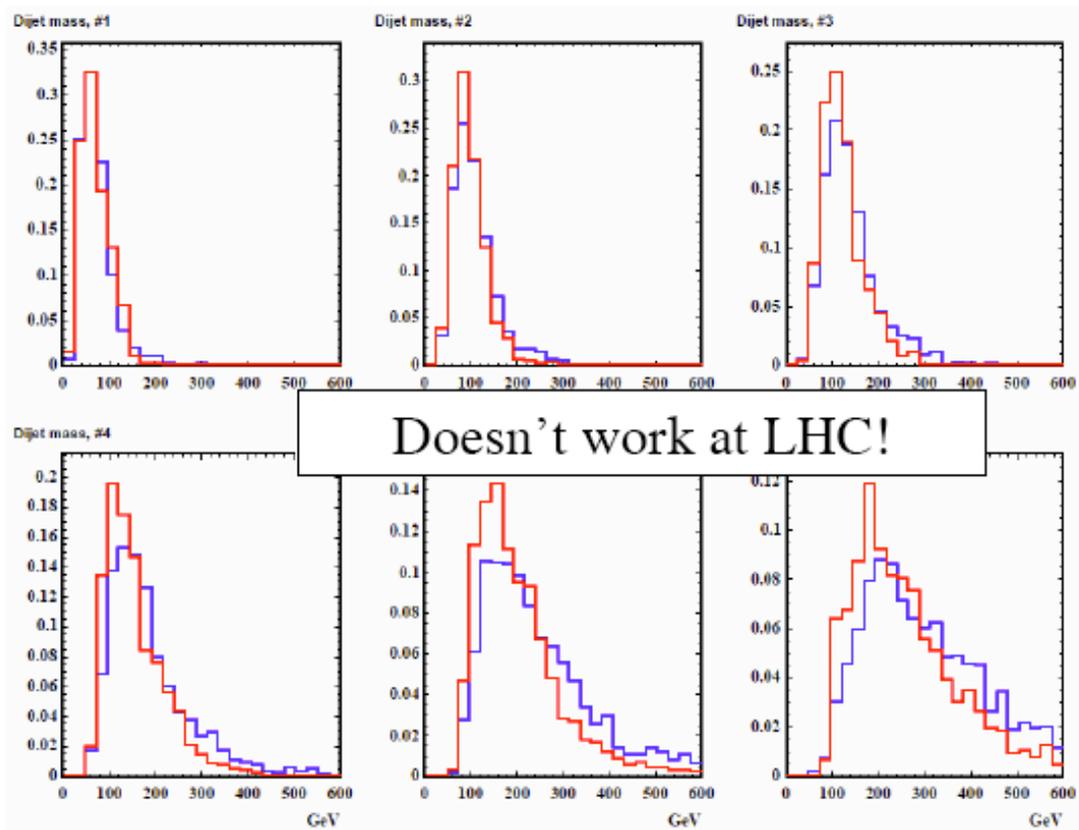
Tevatron Example: 6 ordered mass pairs for each event and plotted for **signal** and **backgrounds**



$$M(g \rightarrow bb) < M(H \rightarrow bb)$$

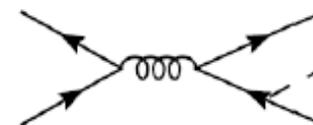
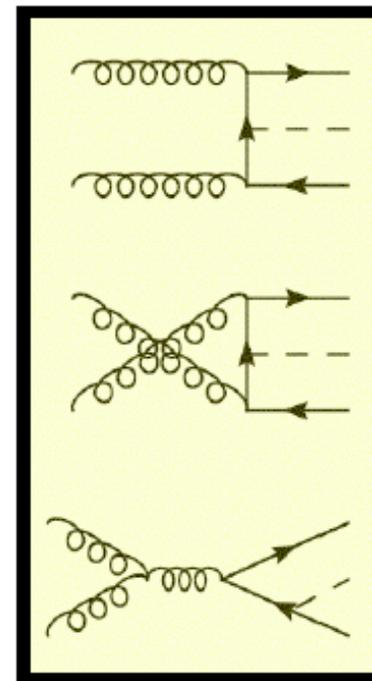


ttH (tt dileptonico) a LHC



— $t\bar{t}H$
 — $t\bar{t}jj + t\bar{t}b\bar{b} + t\bar{t}Z$ (CompHEP)

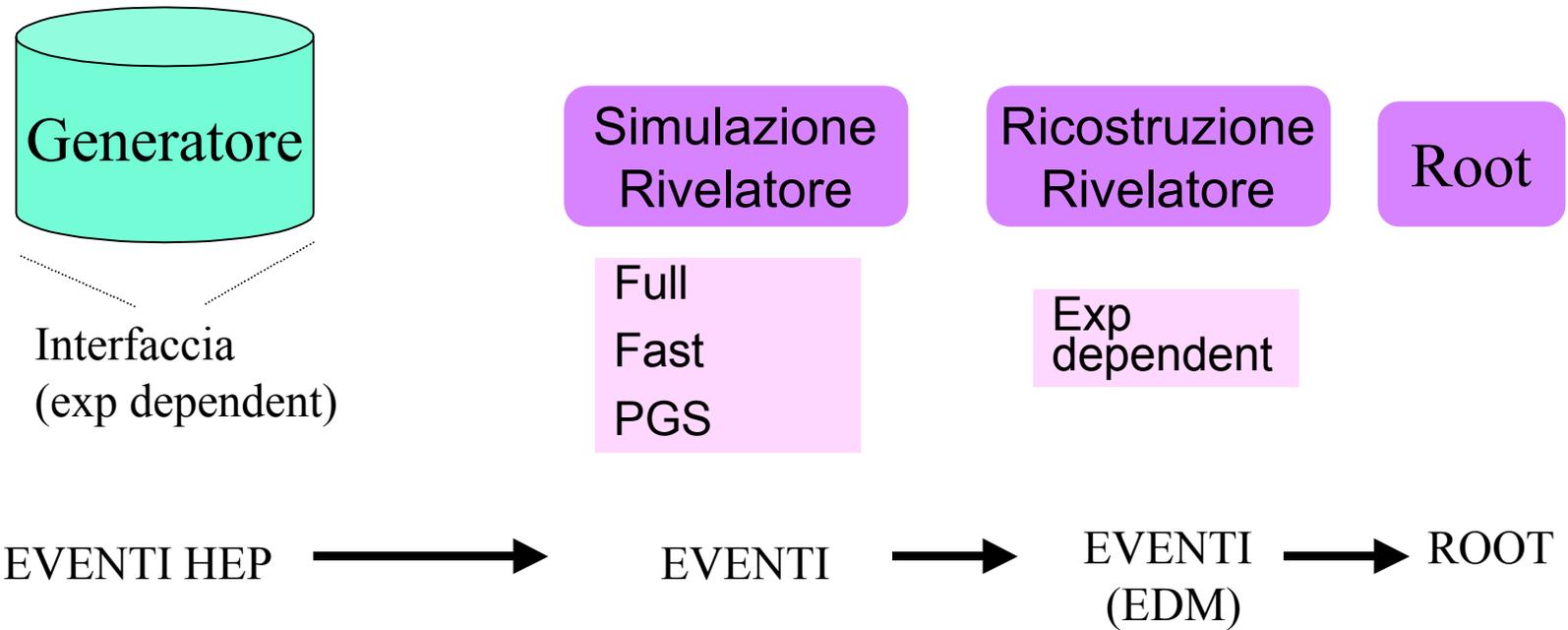
90% @ LHC



I MonteCarlo datasets

- I campioni di MonteCarlo sono essenziali per:
 - Ottimizzazione dei tagli di analisi
 - Valutazione di (alcune) sistematiche
- Il lavoro iniziale e' focalizzato sulla definizione dei:
 - Tagli di generazione
 - Statistica del campione
 - Tipologia dei campioni per la valutazione delle sistematiche
- La generazione, ma soprattutto di simulazione dei campioni per poter essere usati negli esperimenti e' un processo lungo e complesso che richiede grandi risorse di computing e di catalogazione
 - l'intervallo di tempo fra l'inizio della generazione e l'output finale di campione simulato e ricostruito necessario per l'analisi puo' essere anche di mesi!
 - Nonostante farm e procedure automatizzate l'impiego di risorse umane e' considerevole!

Catena di produzione



All'aumentare delle conoscenze del rivelatore la simulazione e la ricostruzione vengono migliorate: ripetizione varie volte del ciclo

In pratica...

- Non sempre (quasi mai...) si riesce ad avere cio' che si vuole.
- Molta attenzione nella scelta dei tagli di generazione. Bisogna bilanciare le due richieste opposte:
 - 1) No bias nelle distribuzioni importanti
 - 2) Efficienza di generazione
 - 3) Generatori "*disponibili*" (i.e. trovare fra i fisici sperimentali coloro che mantengano l'interfaccia con il codice dell'exp e siano in contatto con gli autori (teorici) del generatore)
- *Oggi breve overview del materiale ora disponibile a CMS e Atlas e ai progetti in corso: al prossimo meeting risultati di studi comparativi dedicati*

Atlas: Alpgen, Acermc, Sherpa

CMS: Alpgen , Comphep

W,Z+jets; QQ+jets; tt+jets; N-jets

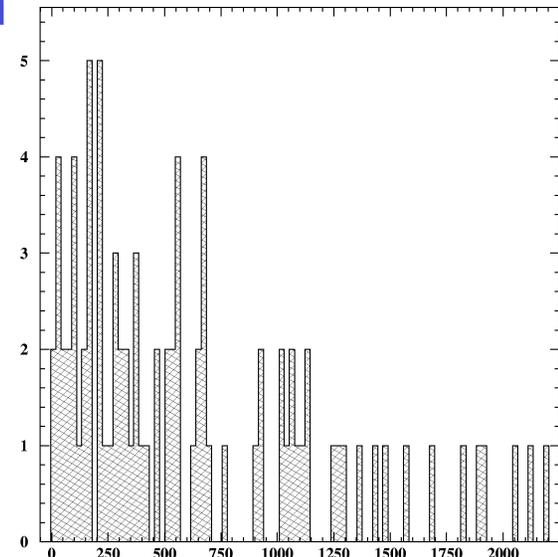
MC@NLO (Atlas, CMS)

W+jets, tt+jets

Esempio 1: fondo per tt (Atlas)

Grazie a Marina Cobal!

- Main use of background production
- Inputs for sample 1
 - W+4 extra light jets
 - Jet: $P_t > 10$, $|\eta| < 3.0$, $\Delta R > 0.3$
 - No lepton cuts
 - Initial grid: 200000*3
 - Events: 150·106
 - Jobs: 198
- Production:
 - Un-weighting to W lepton (e, μ , τ) decay
 - Effective σ : 4390 pb
 - 108401 events generated (3.6 10^{-6} efficiency)
 - 2.57% (2784) events pass first selection
 - $E_{T\text{miss}} > 20$ GeV, lepton (e, μ) $P_t > 20$ and ≥ 4 jets $P_t > 40$



Alpgen: W+4jets (2)

- Main use of background production
- Inputs for sample 2
 - W+4 extra light jets
 - Jet: $P_t > 10$, $|\eta| < 2.5$, $\Delta R > 0.4$
 - No lepton cuts
 - Initial grid: 200000*3
 - Events: $150 \cdot 10^6$
 - Jobs: 98
- Lower maximum weight by factor 10 (?? Can I do this??)
 - Un-weighting to W lepton (e, μ , τ) decay
 - Effective σ : 2430 pb
 - 380740 events generated ($2.6 \cdot 10^{-5}$ efficiency)
 - 3.41% (13002) events pass first selection
 - $E_{T, \text{miss}} > 20$ GeV, lepton (e, μ) $P_t > 20$ and ≥ 4 jets $P_t > 40$

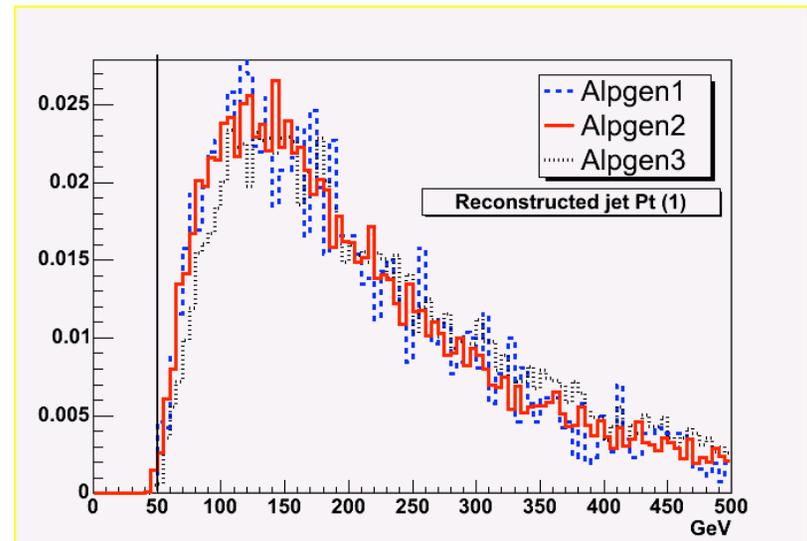
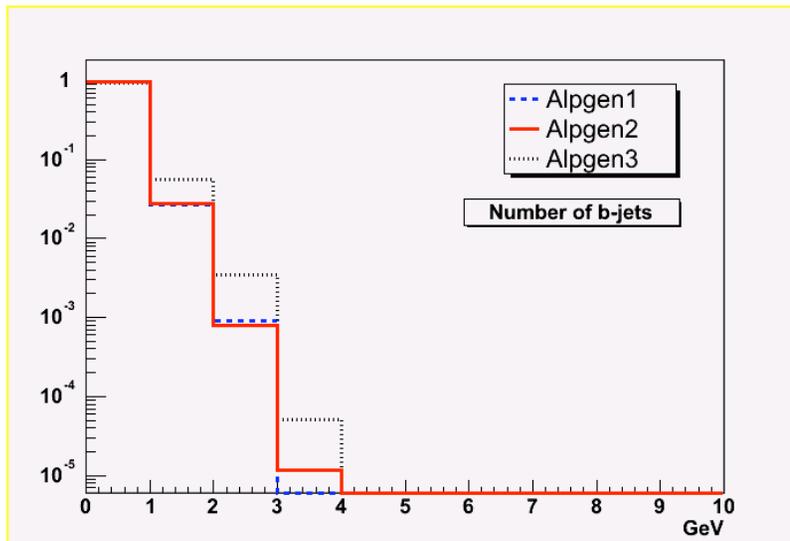
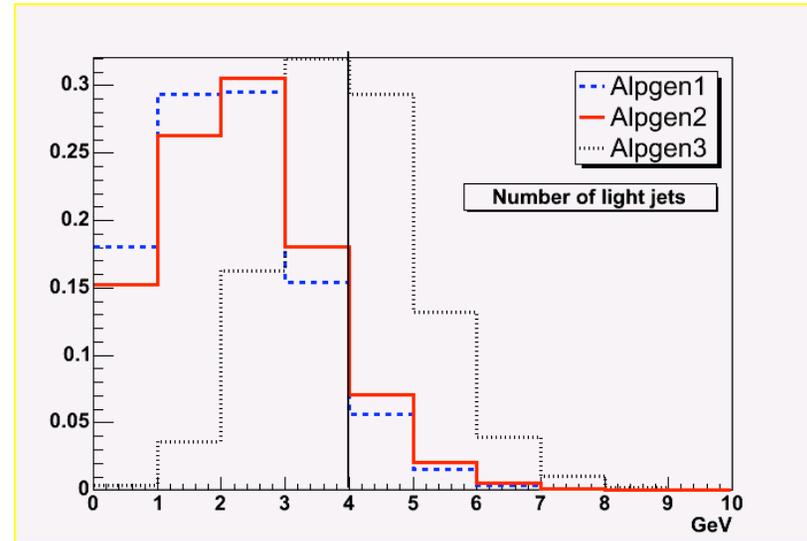
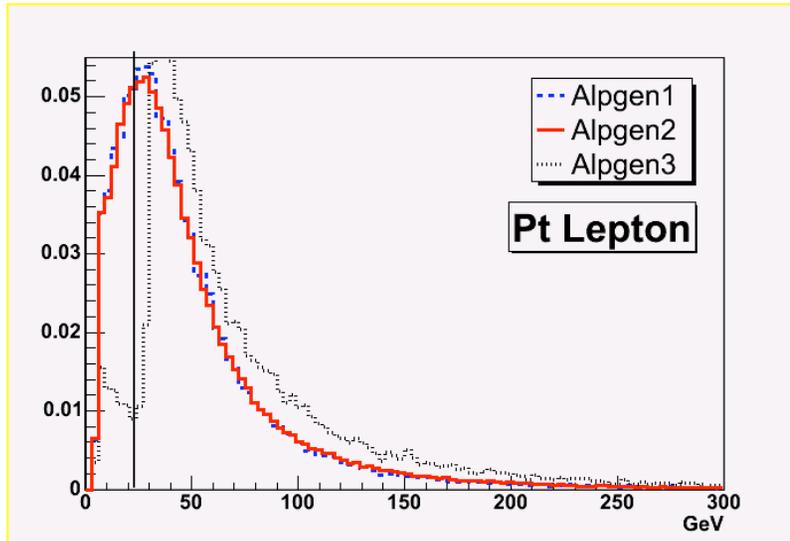
Alpgen: W+4jets (3)

- Main use of background production
- Inputs for sample 3
 - W+4 extra light jets
 - Jets: $P_t > 10$, $|\eta| < 2.5$, $\Delta R > 0.4$
 - Lepton: $P_t > 30$, $|\eta| < 3.0$, $E_{\text{miss}} > 30$.
 - Initial grid: 200000×3
 - Events: $200 \cdot 10^6$
 - Jobs: 100
- Lower maximum weight by factor 10 (?? Can I do this??)
 - Un-weighting to W lepton (e, μ , τ) decay
 - Effective σ : 106 pb
 - 39810 events generated ($2 \cdot 10^{-6}$ efficiency)
 - 25.8% (10264) events pass first selection
 - $E_{\text{T}}^{\text{miss}} > 20$ GeV, lepton (e, μ) $P_t > 20$ and ≥ 4 jets $P_t > 40$

Chosen too large for fair comparison with other data sets
Data set still included in next plots just for comparison

Confronto

All histograms
normalized to unity



Esempio 2: fondo per ttH (CMS)

Grazie a Susanna Cucciarelli e il gruppo ttH

Generator Level

x-sec	ttH m=120GeV	ttj	ttbb
Atlas	0.520 pb	474 pb	8.6pb
CMS	0.664 pb (NLO)	507.8 pb	3.3 pb

different generators:
PYTHIA (ATLAS)
compHEP+PYTHIA (CMS)

**different generators
AND
different processes:**
tt + IFSR PYTHIA (ATLAS)
(preselection after hadronization)
ttj (matrix element)
compHEP+PYTHIA (CMS)

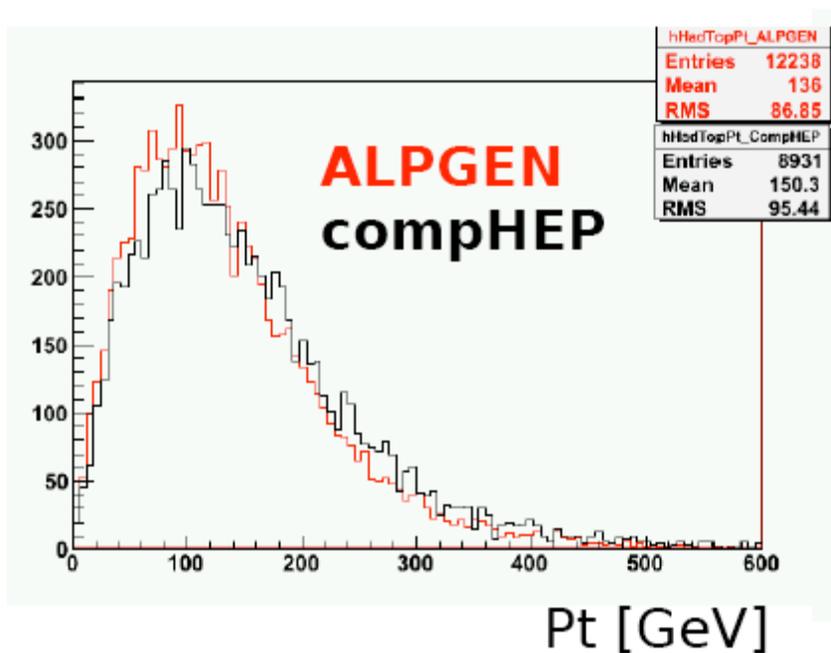
similar generation:
AcerMC (ATLAS)
compHEP (CMS):
generator cuts:
Pt(b)>15
eta<3
deltaR(bb) > 0.3

Studi per il fondo tt+jets

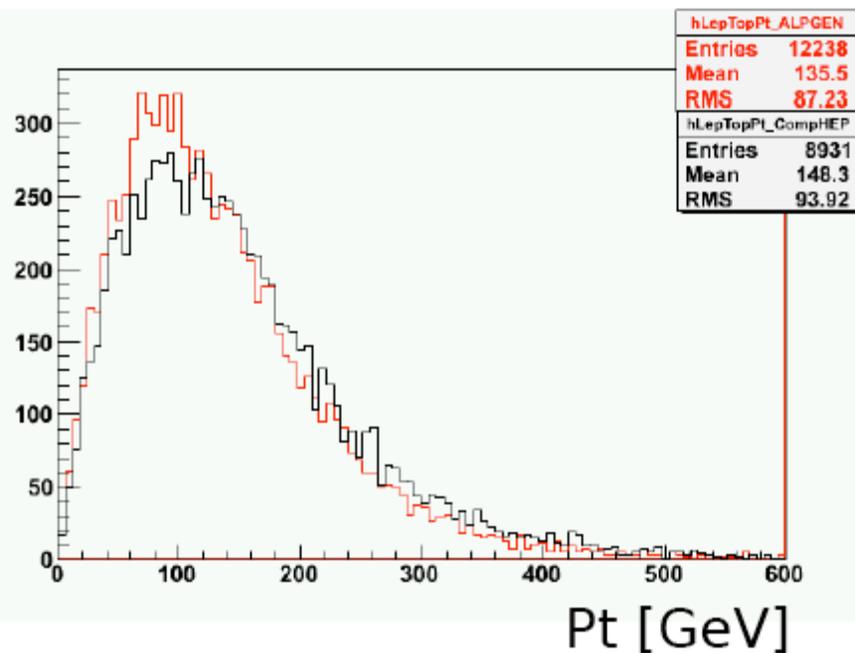
- Campioni generati:
 - COMPHEP: tt+2jj(inclusive) + PYTHIA
 - $\sigma_{\text{prod}} = 507 \text{ pb}$
 - $P_t(j) > 15 \text{ GeV}$, $|\eta(j)| < 3$, $\Delta R(j,j) > 0.3$
 - ALPGEN: tt+2j (exclusive) + PYTHIA
 - $\sigma_{\text{prod}} = 100 \text{ pb}$
 - $P_t(j) > 20 \text{ GeV}$, no eta cut, $\Delta R(j,j) > 0.7$
 - Per fare un confronto bisogna applicare dei tagli:
 - Come si e' visto anche prima, questa non e' la situazione "ideale" ma nella realta' bisogna cercare di estrarre piu' informazioni possibile dai campioni che si hanno a disposizione

Confronto: pt(top)

hadronic



leptonic

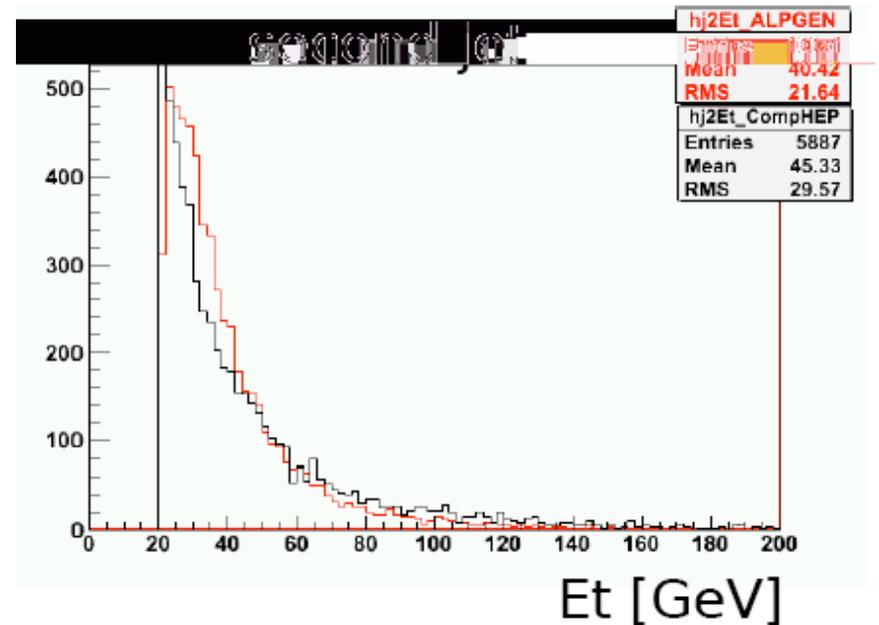
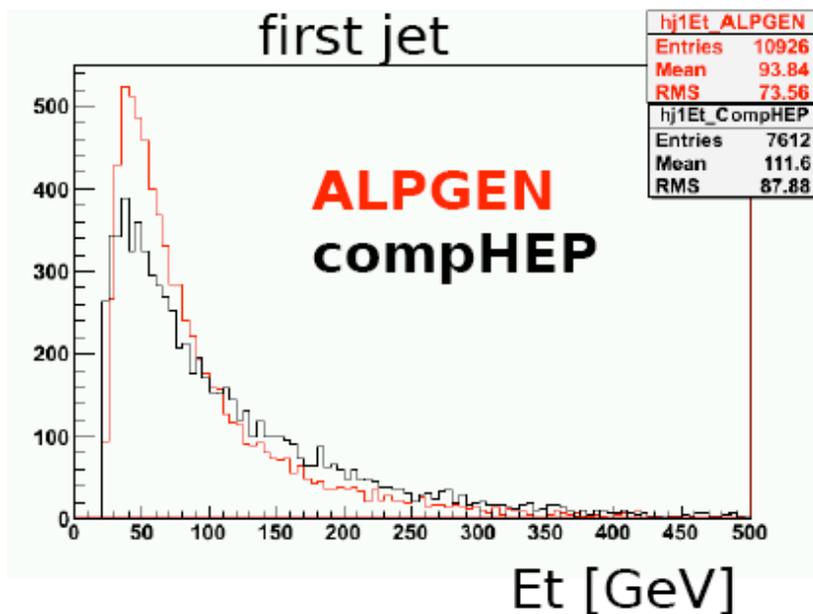
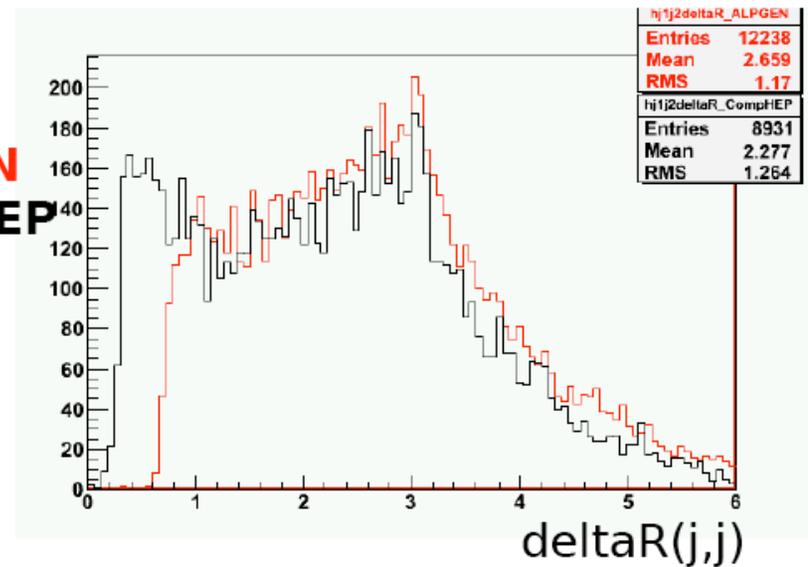


Plot normalizzati allo
stesso numero di eventi

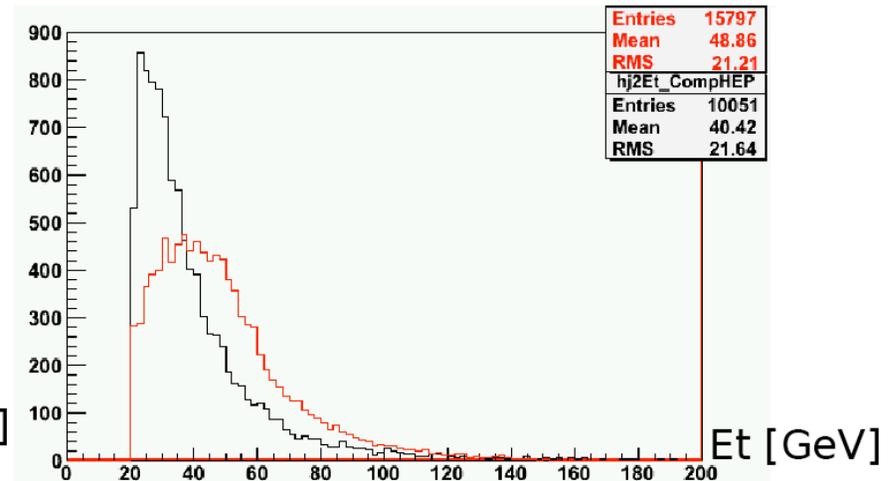
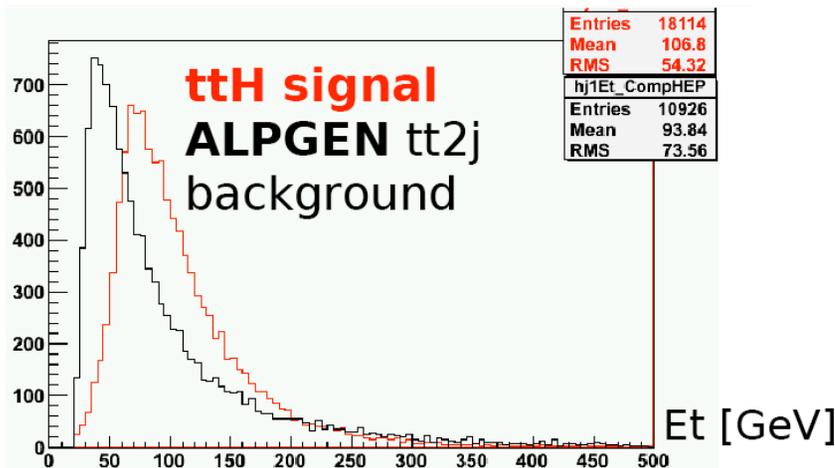
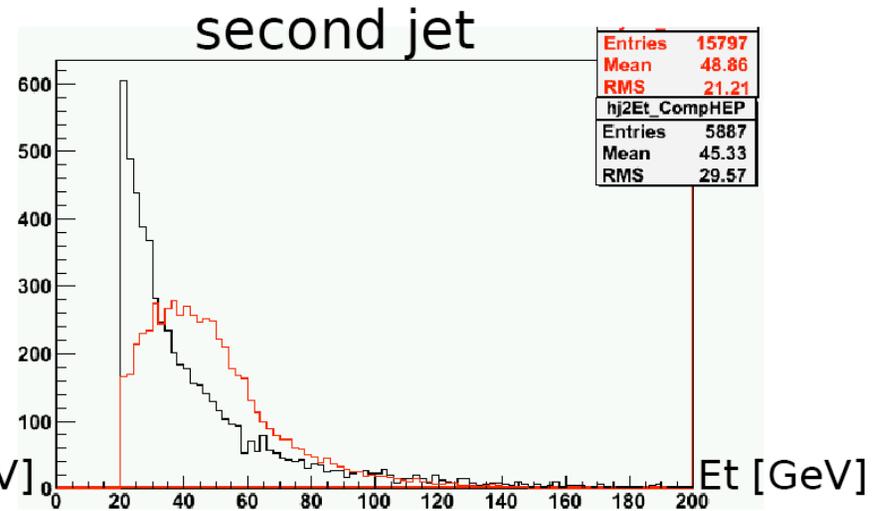
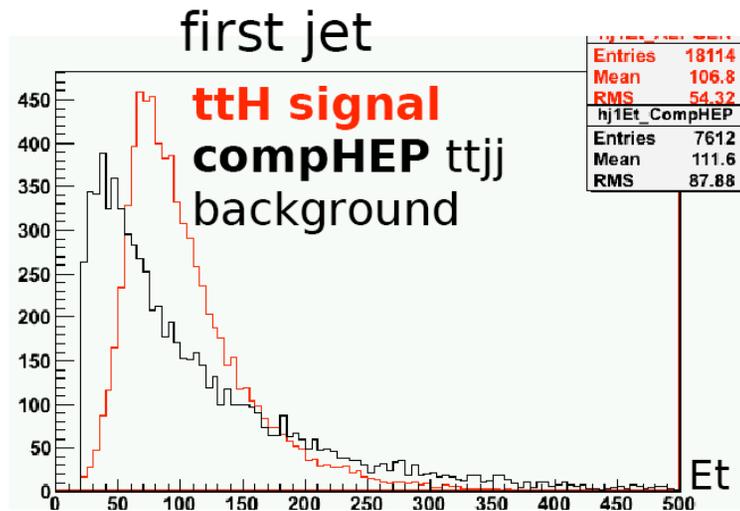
Confronto: extra jets

Dopo aver applicato tagli in modo da rendere i campioni (piu') confrontabili

ALPGEN
compHEP



Confronto segnale-fondo



Spunti per la discussione

- Da quanto visto finora la produzione di MC agli esperimenti, in particolare ME, e' un processo "costoso" e spesso ha come risultato prodotti non pienamente soddisfacenti per gli scopi dell'analisi
- Due fatti da considerare:
 - Data la fatica per generare campioni si cerca di essere sempre molto inclusivi
 - Questo comporta inefficienza
 - Scelta dei parametri di generazione?
 - Scelta di prioritá' di processi
 - Diversa statistica raccolta corrisponde ad analisi diverse (soprattutto per i primi dati a bassa luminosita')
 - Valutazione delle sistematiche
 - solitamente la statistica di questi campioni e' minore
 - solitamente sono gli ultimi ad essere disponibili
 - Sistematiche teoriche possono essere affrontate fino dall'inizio

Piani per il futuro

- Massiva generazione e simulazione di campioni MC e' in corso:
 - CMS ha un (nuovo) programma completo di produzione massiccia con Alpgen
 - Atlas sta creando nuovi campioni con MC@NLO
- Presentazione dettagliata al prossimo incontro del workshop (22 maggio) di studi con questi nuovi campioni:
 - Ovviamente questi campioni sono utilizzati in molte analisi e sono argomento di discussione per vari (tutti) gruppi di questo WS.
 - Nel gruppo ME seguiremo in particolare i dettagli relativi a W/Z +jets e tt +jets a titolo di esempio
- Questi nuovi campioni sono particolarmente importanti da capire e validare perche' sono la base per gli studi di fisica di SM con i primi dati di LHC