Subjet mass constraints on boosted jets

Laís Sarem Schunk

Work currently being finalized with Mrinal Dasgupta and Gregory Soyez, with Gavin Salam collaboration in early stages.

IPhT - CEA Saclay

BOOST2015, Aug. 10-14, Chicago IL.

Presentation Plan

Presentation Plan

Introduction

- Constrain radiation to discriminate different types of jets.
- Focus on 2-pronged signal emissions (W/Z/H vs. QCD).
- Jet-shapes studied: $v = \tau_{21}^{(\beta=2)}$, μ^2 and $C_2^{(\beta=2)}$.
- Understand differences/similarities from a first-principle analytical study.

Presentation Plan

Jet shapes

- **N-subjettiness** with axes $a_1, ..., a_N$ J. Thaler, K. V. Tilburg (2010)

$$\tau_{21}^{(2)} = \frac{\tau_2^{(2)}}{\tau_1^{(2)}}, \qquad \tau_N^{\beta} = \frac{1}{p_{t,jet} R^{\beta}} \sum_{i \in jet} p_{t,i} \min_{a_i \ldots a_N} (\theta_{ia_1}^{\beta}, ..., \theta_{ia_N}^{\beta}).$$

- **Mass-drop** with subjets j_1 and j_2 (non-recursive version)
J. M. Butterworth, A. R. Davison, M. Rubin, G. P. Salam (2008)

$$\mu_p^2 = \max(m_{j1}^2, m_{j2}^2)/m_j^2.$$

- **Energy correlation** I. Moult Talk, BOOST2015
A. J. Larkoski, G. P. Salam, J. Thaler (2013); A. J. Larkoski, I. Moult, D. Neill (2014 and 2015)

$$C_2^{(2)} = e_3^{(2)}/(e_2^{(2)})^2 \qquad or \qquad D_2^{(2)} = e_3^{(2)}/(e_2^{(2)})^3,$$

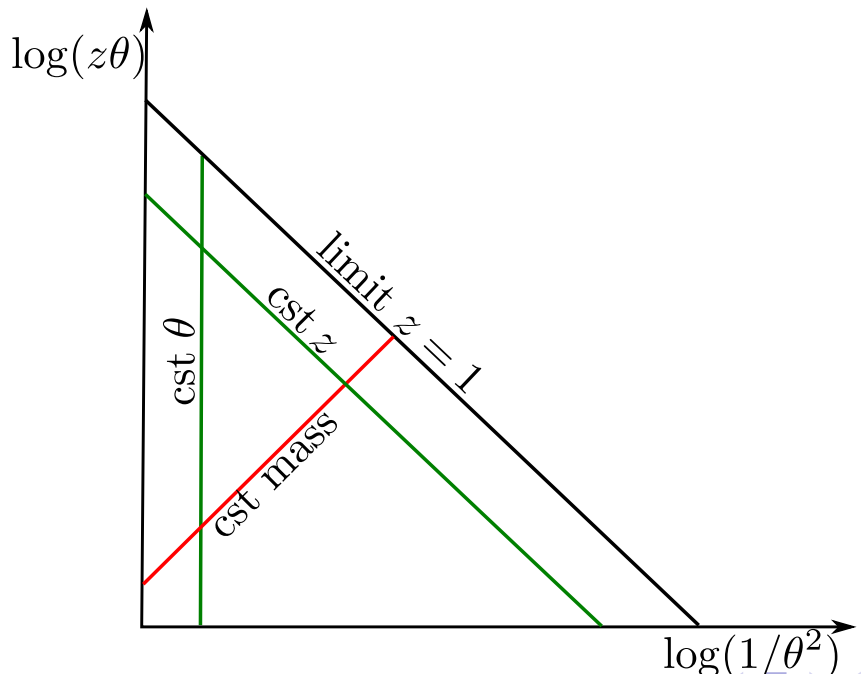
$$e_2^{\beta} = \frac{1}{p_t^2 R^{\beta}} \sum_{i<j \in jet} p_{t,i} p_{t,j} \theta_{ij}^{\beta}, \qquad e_3^{\beta} = \frac{1}{p_t^3 R^{3\beta}} \sum_{i<j<k \in jet} p_{t,i} p_{t,j} p_{t,k} \theta_{ij}^{\beta} \theta_{ik}^{\beta} \theta_{jk}^{\beta}.$$

- **Recursive versions**: decluster until $v < v_{cut}$ (not in this talk).

Presentation Plan

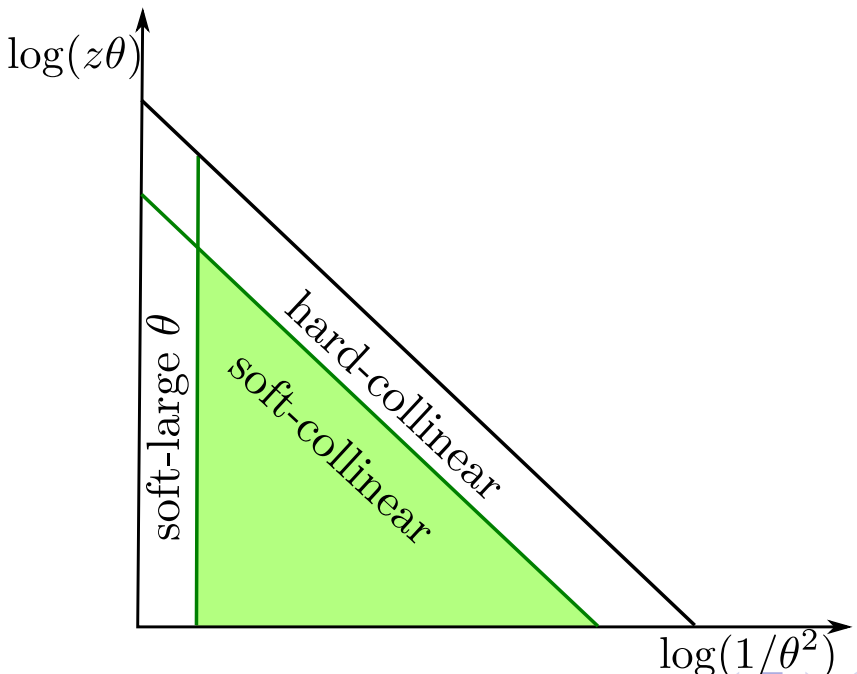
Lund diagrams

- Calculations are performed in the following limits:
  - Boosted jets of a given mass, $\rho = m^2/p_t^2 R^2 \ll 1$.
  - Cut in shape is small, $v_{cut} \ll 1$.
  - Small jet radius R (neglect R^2 corrections).
- Lund diagram : graphical representation of the results in $z\theta$ vs. $1/\theta^2$ coordinates.



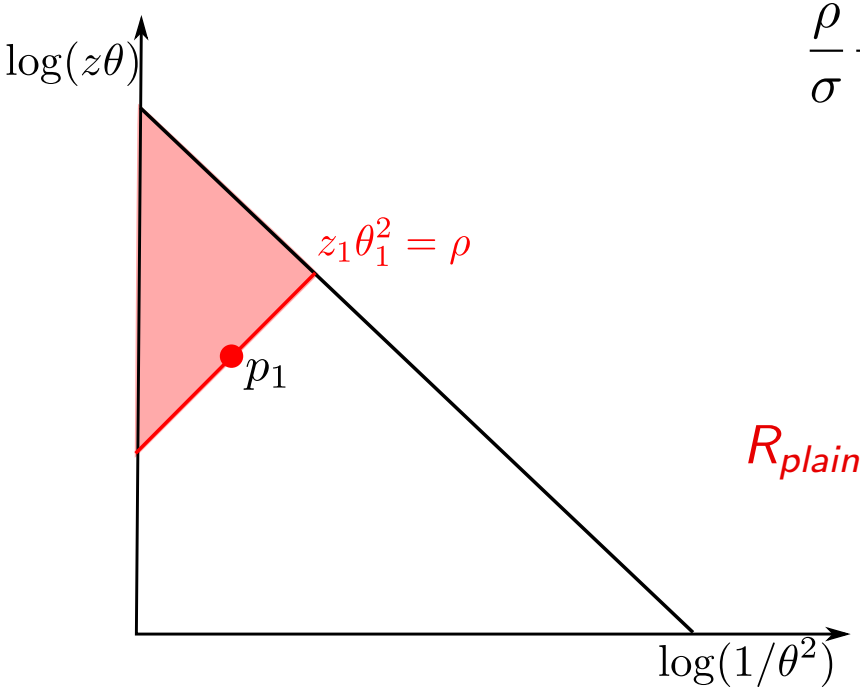
Lund diagrams

- Calculations are performed in the following limits:
  - Boosted jets of a given mass, $\rho = m^2/p_t^2 R^2 \ll 1$.
  - Cut in shape is small, $v_{cut} \ll 1$.
  - Small jet radius R (neglect R^2 corrections).
- Lund diagram : graphical representation of the results in $z\theta$ vs. $1/\theta^2$ coordinates.



Structure of the results

- Up to LL, emissions are strongly ordered in mass and angle (the most massive emission p_1 sets the mass).
- Independent emissions $\rightarrow$ constraints as an exponential factor.
- For a jet of a given mass:

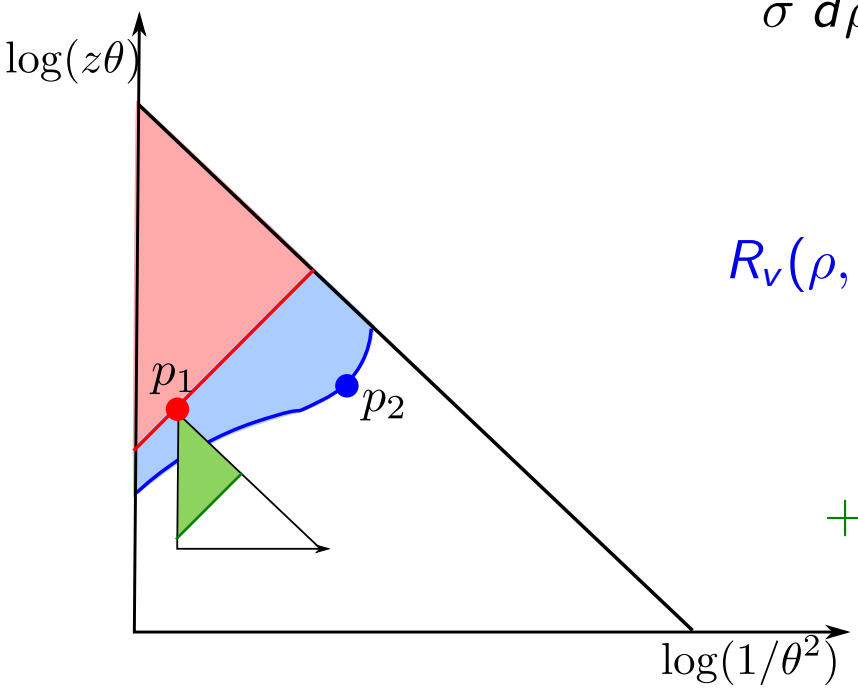


$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho} = \int_\rho^1 dz_1 P_i(z_1)\frac{\alpha_s(z_1\rho p_t^2 R^2)}{2\pi} \times e^{-R_{plain}(\rho)}$$

$$R_{plain}(\rho) = \int_0^1 \frac{d\theta^2}{\theta^2}\int_0^1 dz P_i(z)\frac{\alpha_s(z^2\theta^2 p_t^2 R^2)}{2\pi} \times \Theta(z\theta^2 > \rho)$$

Structure of the results

- The second most massive emission p_2 sets the value of v.
- For a jet of a given mass + a cut in the shape v_{cut}:



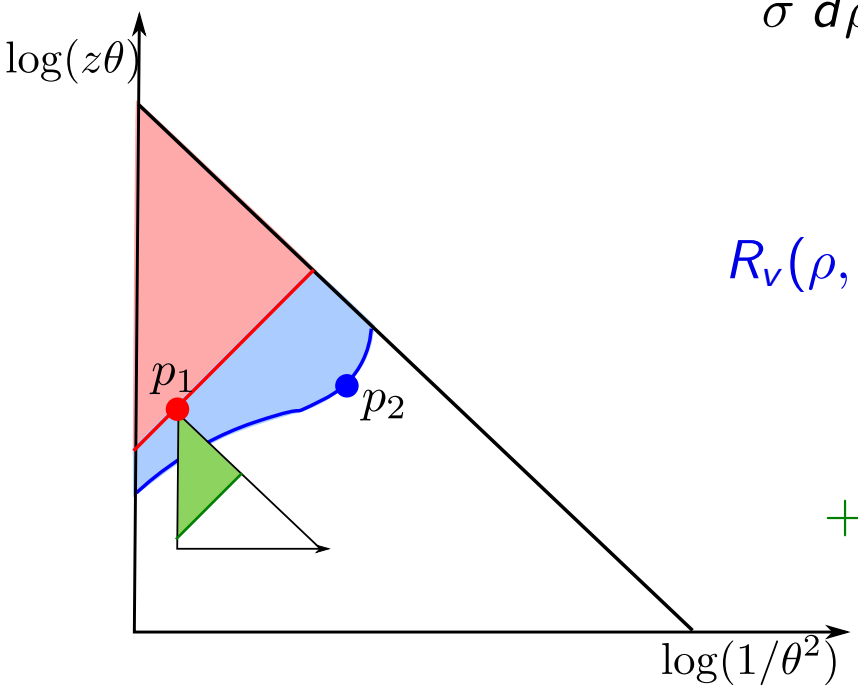
$$\left.\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\right|_{<v} = \int_\rho^1 dz_1 P(z_1)\frac{\alpha_s(z_1\rho p_t^2 R^2)}{2\pi} \times e^{-R_{plain}(\rho)-R_v(\rho,z_1)}$$

$$R_v(\rho,z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2}\int_0^1 dz_2 P_i(z_2)\frac{\alpha_s(z_2^2\theta_2^2 p_t^2 R^2)}{2\pi} \times \Theta(v(\rho,z_1,z_2,\theta_2) > v_{cut})\Theta(z_2\theta_2^2 < \rho)$$

$$+ \int_0^1 \frac{d\theta_{12}^2}{\theta_{12}^2}\int_0^1 dz_2 P_g(z_2)\frac{\alpha_s(z_1\rho z_2^2\theta_{12}^2 p_t^2 R^2)}{2\pi} \times \Theta(v^{sec}(\rho,z_1,z_2,\theta_{12}) > v_{cut})$$

Structure of the results

- The second most massive emission p_2 sets the value of v.
- For a jet of a given mass + a cut in the shape v_{cut}:



$$\left.\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\right|_{<v} = \int_\rho^1 dz_1 P(z_1)\frac{\alpha_s(z_1\rho p_t^2 R^2)}{2\pi} \times e^{-R_{plain}(\rho)-R_v(\rho,z_1)}$$

$$R_v(\rho,z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2}\int_0^1 dz_2 P_i(z_2)\frac{\alpha_s(z_2^2\theta_2^2 p_t^2 R^2)}{2\pi} \times \Theta(v(\rho,z_1,z_2,\theta_2) > v_{cut})\Theta(z_2\theta_2^2 < \rho)$$

$$+ \int_0^1 \frac{d\theta_{12}^2}{\theta_{12}^2}\int_0^1 dz_2 P_g(z_2)\frac{\alpha_s(z_1\rho z_2^2\theta_{12}^2 p_t^2 R^2)}{2\pi} \times \Theta(v^{sec}(\rho,z_1,z_2,\theta_{12}) > v_{cut})$$

Now all we need is to find $v(\rho, z_1, z_2, \theta_2)$.

Presentation Plan

Before we start...

N-subjettiness

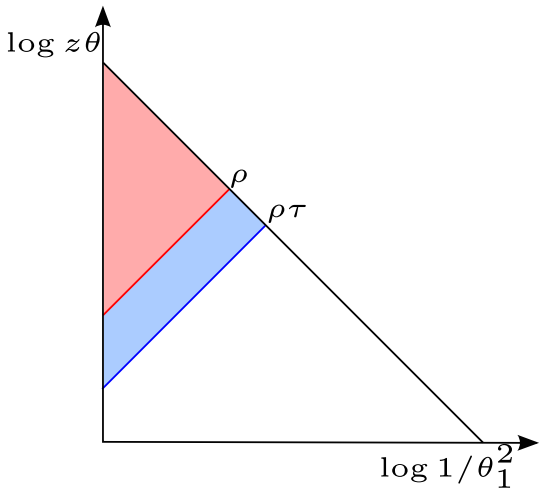
- For τ_2: three possible partitions of p_0 and emissions p_1, p_2.
- Depends on choice of axis:

optimal	$\tau_2 \approx z_2\theta_2^2$		up to our accuracy
excl. gen-$k_t^{p=1/2}$	$\tau_2 \approx z_2\theta_2^2$		simpler than optimal
excl. k_t	$\tau_2 \approx z_2\theta_2^2$	if $z_2\theta_2 < z_1\theta_1$	not in this talk
	$\approx z_1\theta_1^2$	if $z_2\theta_2 > z_1\theta_1$	

Mass drop

- IR safety relies on recursive C/A declustering.
- Solution: combine non-recursive version with gen-$k_t^{p=1/2}$ declustering.

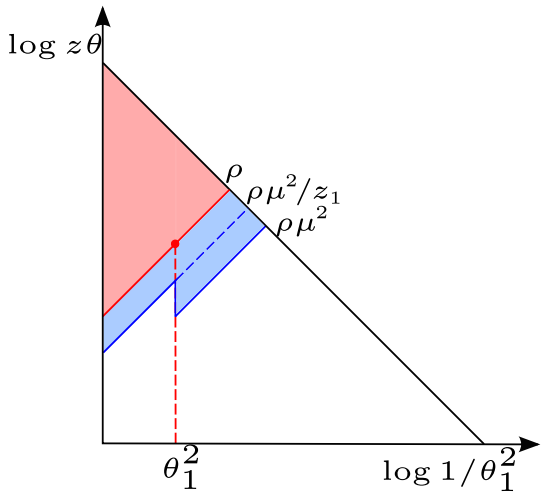
N-subjettiness



$$\tau_{21} = \frac{z_2\theta_2^2}{z_1\theta_1^2}$$

$$\tau_{21}^{sec} = z_2\frac{\theta_{12}^2}{\theta_1^2}$$

Mass drop

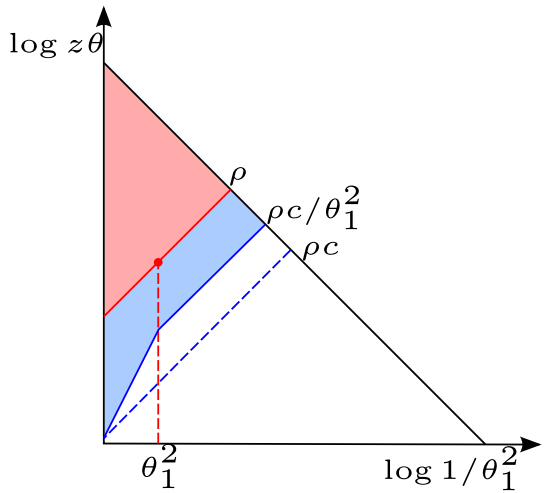


$$\mu^2 \overset{\theta_2\leq\theta_1}{=} \frac{z_2\theta_2^2}{z_1\theta_1^2}$$

$$\overset{\theta_2\geq\theta_1}{=} \frac{z_2\theta_2^2}{z_1\theta_1^2} \text{ or } \frac{z_2\theta_2^2}{\theta_1^2}$$

$$\mu^{2\,sec} = z_1z_2\frac{\theta_{12}^2}{\theta_1^2}$$

Energy correlation

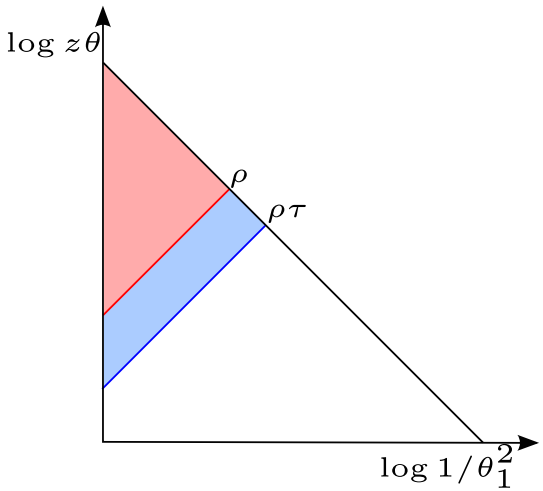


$$C_2 = \frac{z_2\theta_2^2}{z_1\theta_1^2}\max(\theta_1^2,\theta_2^2)$$

$$C_2^{sec} = z_2\frac{\theta_{12}^2}{\theta_1^2}\theta_1^2$$

$$\rho = \tau_1 = e_2 = z_1\theta_1^2$$

N-subjettiness

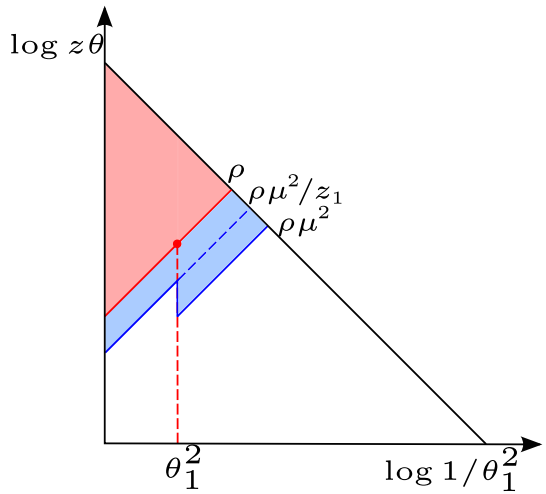


$$R_\tau(z_1) =$$

$$\frac{\alpha_s C_R}{2\pi}\left[\frac{L_\tau^2}{2} + L_\rho L_\tau\right]$$

$$+\frac{\alpha_s C_A}{2\pi}\frac{L_\tau^2}{2}$$

Mass drop



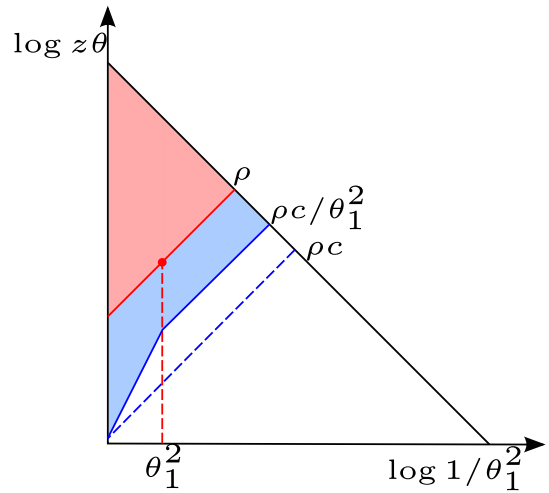
$$R_{\mu_{1/2}^2}(z_1) =$$

$$\frac{\alpha_s C_R}{2\pi}\left[\frac{L_\mu^2}{2} + L_\rho L_\mu\right]$$

$$-\frac{\alpha_s C_R}{2\pi}\frac{L_\rho}{2}(L_\rho - L_1)$$

$$+\frac{\alpha_s C_A}{2\pi}\frac{(L_\mu - L_1)^2}{2}$$

Energy correlation



$$R_{C^2}(z_1) =$$

$$\frac{\alpha_s C_R}{2\pi}\left[\frac{L_e^2}{2} + (L_e - L_\rho + L_1)L_1\right]$$

$$+\frac{\alpha_s C_A}{2\pi}\frac{1}{2}(L_e - L_\rho + L_1)^2$$

$$L_X = \log(1/X)$$

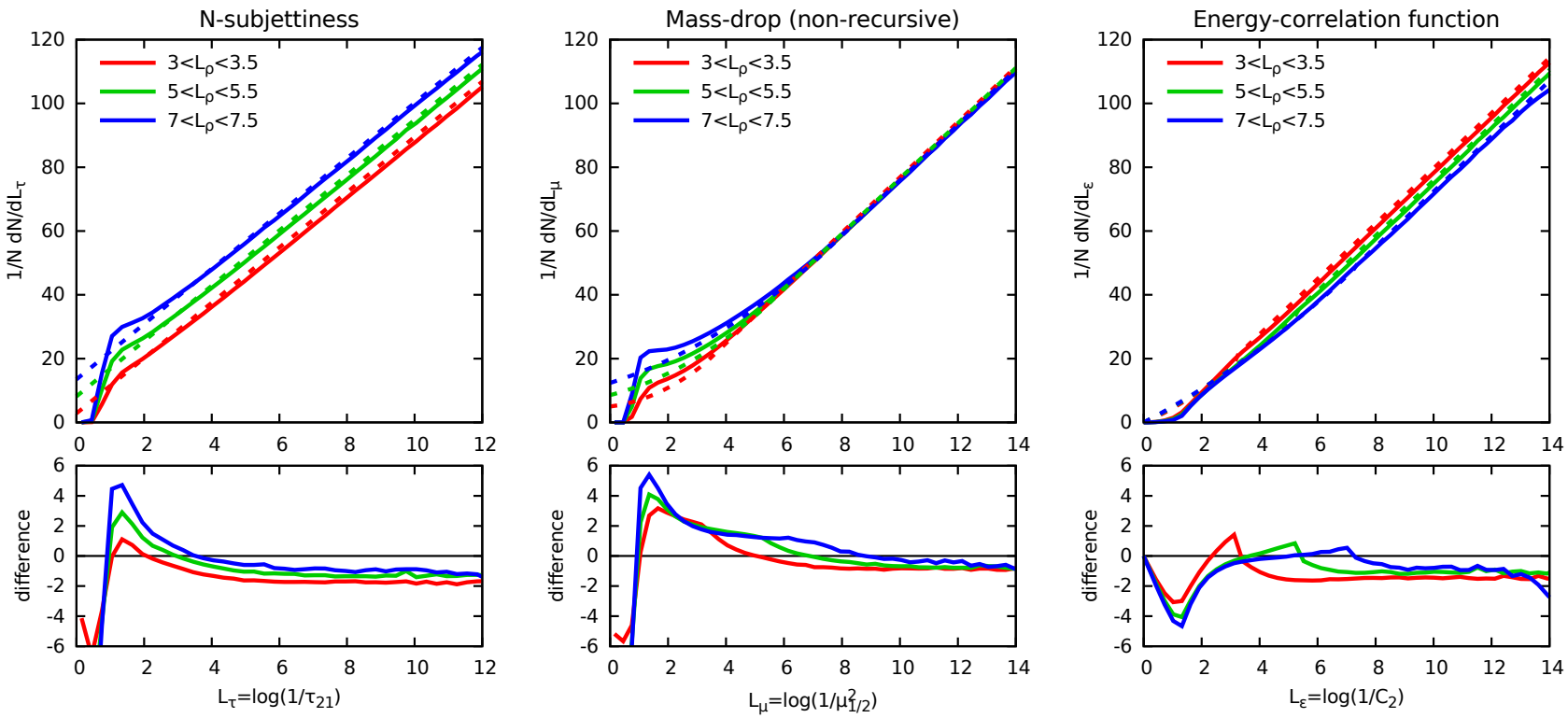
Towards NLL accuracy

- **Hard, collinear radiation**: addition of a term B_i in $P_i(z)$.
- **Multiple emissions** (emissions are no longer ordered):
  - τ_{21} and C_2 are additive shapes, we simply add a factor $\gamma_E R'_v(z_1) + \log[\Gamma(1 + R'_v(z_1))]$
  - For μ^2 or recursive shapes: more delicate, depends on how the emissions cluster.
- **1-loop and 2-loop running coupling**: mostly a technical complication.
- **Finite z_1 contributions**: additional complication in calculations.
- **Soft, large-angle radiation** vanishes in the limit $\theta_1 \ll R$, except for **non-global logs** which are difficult to handle analytically, but can be eliminated by grooming techniques.

Fixed-order Monte-Carlo

- Expressions obtained can be expanded in series of α_S and integrated in z_1 to be compared to Event2.

  [Solid line for simulation, dashed for analytical]

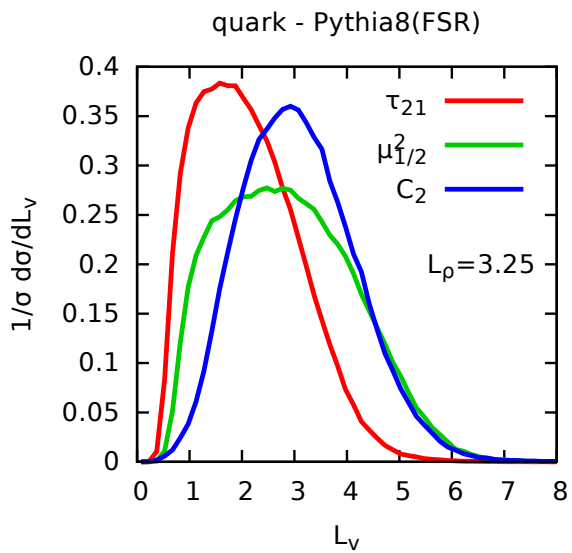


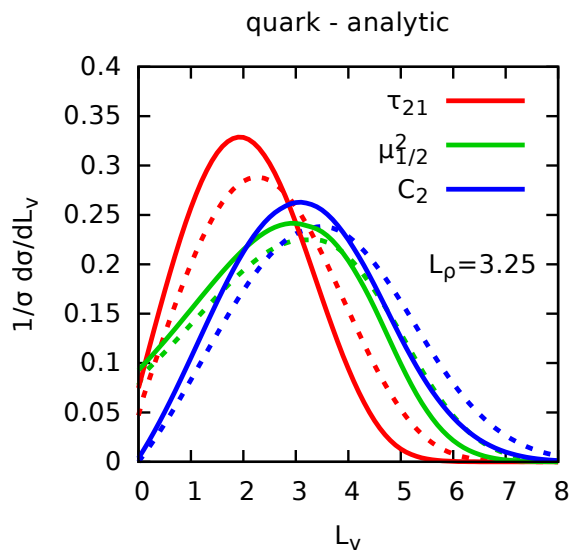
- At large L_v our analytical description agree with simulations.

Parton-shower Monte-Carlo

- Compare results with Monte-Carlo event generators like Pythia. Below we have the L_v distribution.

  [Dashed line is LL only, solid line includes some NLL corrections]



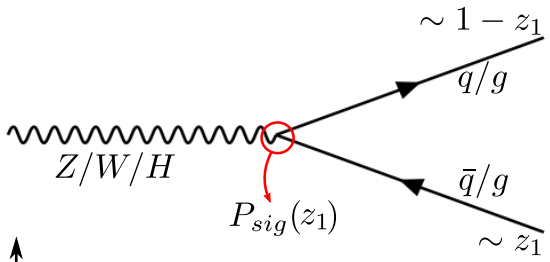


- Difference between shapes qualitatively reproduced (but Pythia has a more "peaked" distribution).
- Including global NLL corrections improves the agreement.

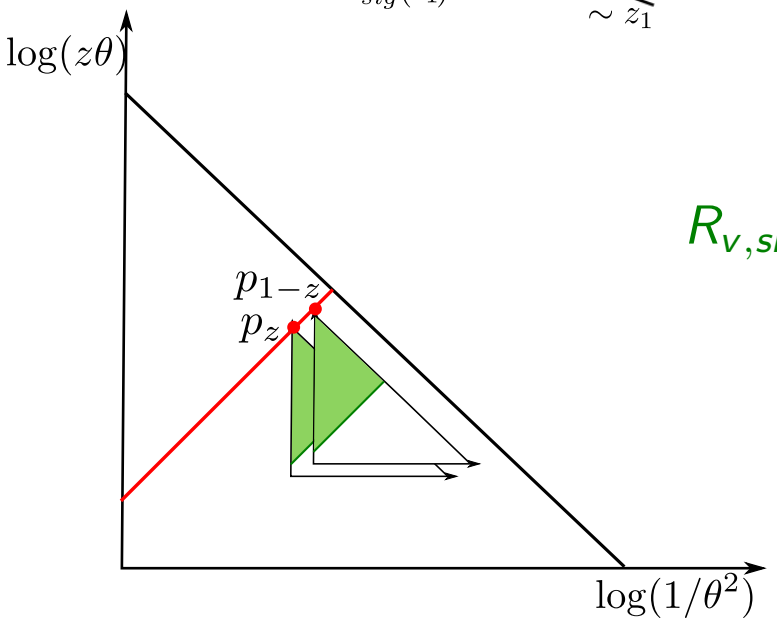
Presentation Plan

Structure of the results

- Decay of a boosted object into a pair $q\bar{q}$ or gg.
- For a signal jet (fixed mass always) + a cut in the shape v:



$$\Sigma_{sig} = \int_{\rho}^{1} dz_1 P_{sig}(z_1) e^{-R_{v,sig}(z_1,\rho) - R_{v,sig}(1-z_1,\rho)}$$

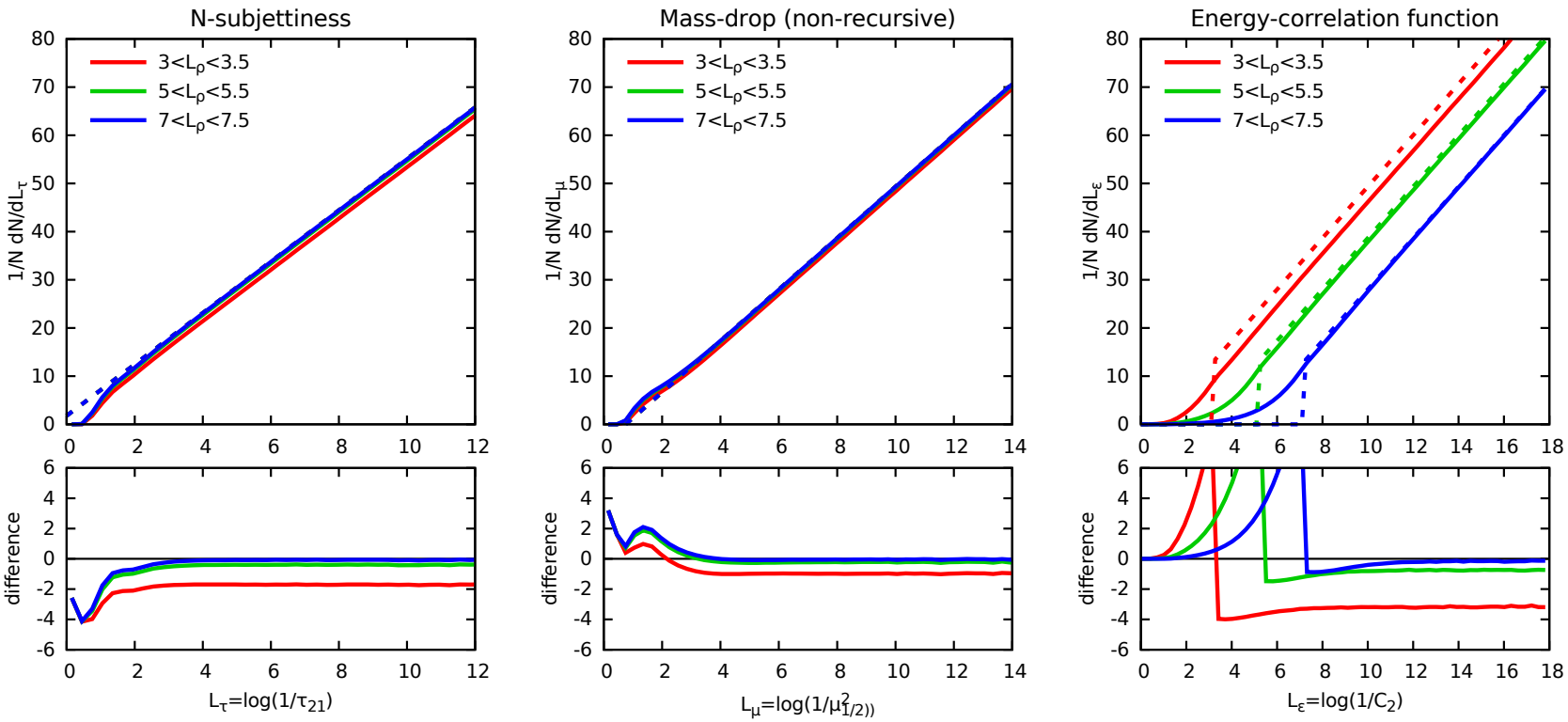


$$R_{v,sig}(z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 P_i(z_2) \frac{\alpha_s(z_1^2\theta_1^2 z_2^2 \theta_2^2 p_t^2 R^2)}{2\pi}$$
$$\times \Theta(v^{sig}(\rho, z_1, z_2, \theta_2) > v_{cut})$$

Fixed-order Monte-Carlo

- Comparison fixed order at α_s to Event2.

  [Solid line for simulation, dashed for analytical]

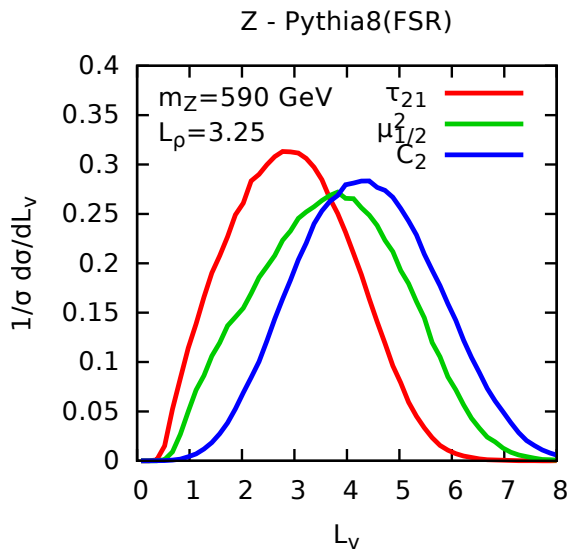


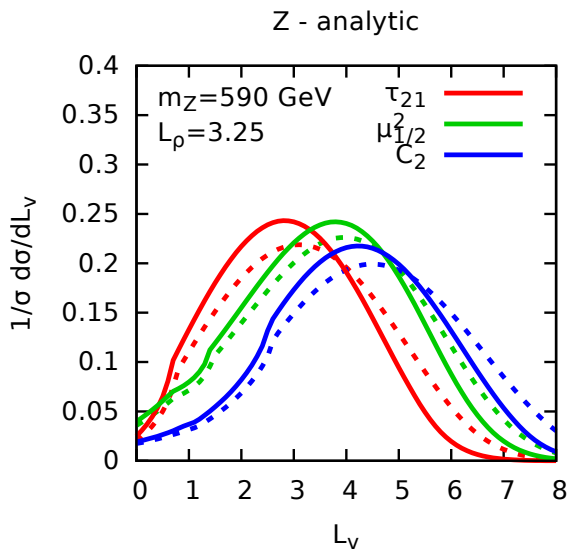
- At large L_v our analytical description shows a good agreement.

Parton-shower Monte-Carlo

- Comparison to a Z-like boson decaying hadronically (varying cut, fixed mass).

  [Dashed line is LL only, solid line includes some NLL corrections]



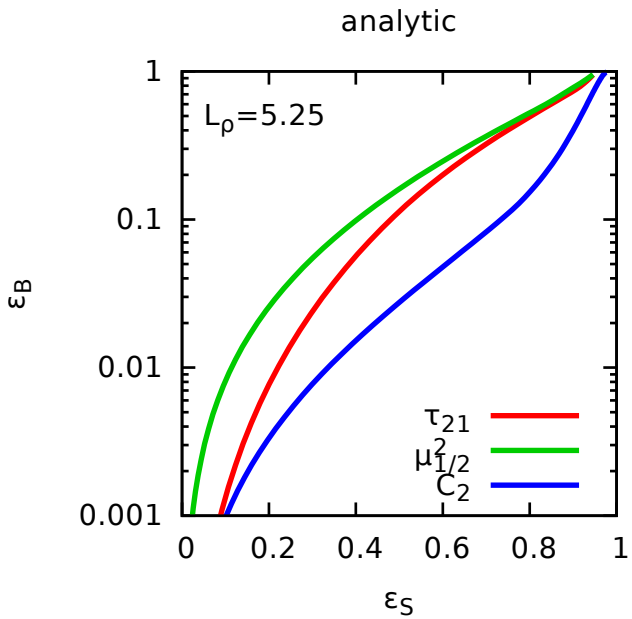


- Good overall description of the shapes.

Presentation Plan

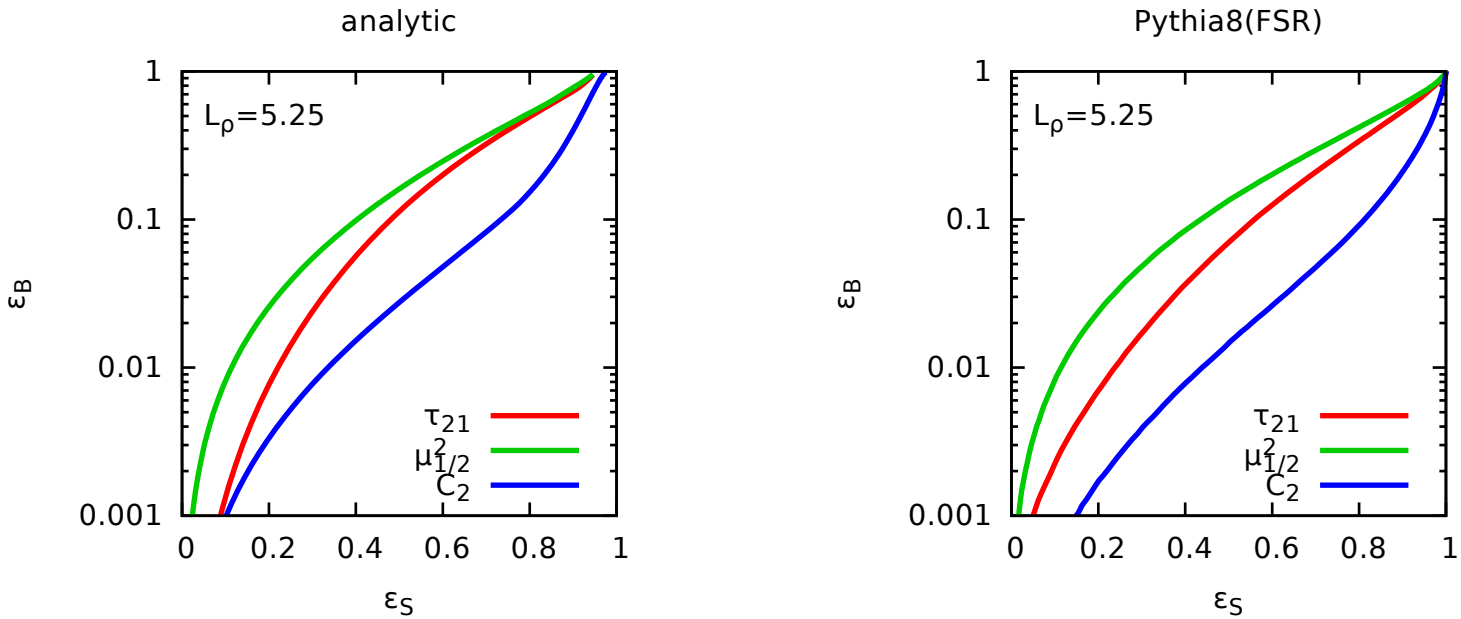
ROC curves

- Efficiency for the QCD background ε_B vs. signal ε_S.
- C_2 is the most efficient due to large-angle constraints, and τ_{21} more efficient than μ^2 (more delicate call).



ROC curves

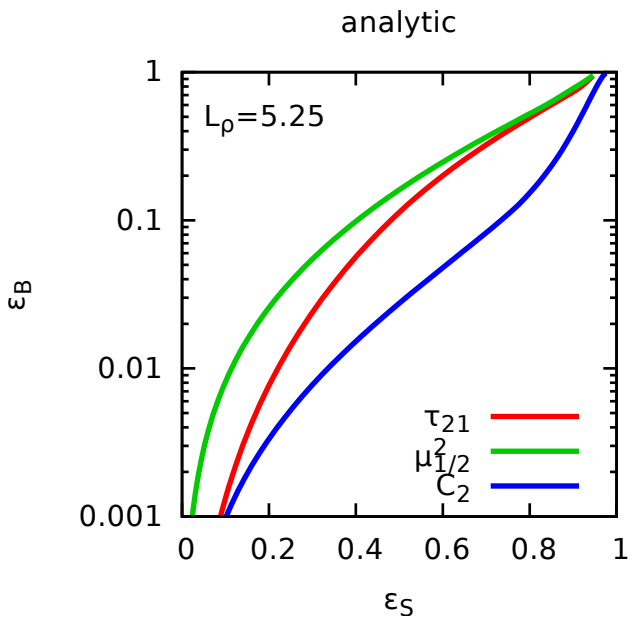
- Efficiency for the QCD background ε_B vs. signal ε_S.
- C_2 is the most efficient due to large-angle constraints, and τ_{21} more efficient than μ^2 (more delicate call).

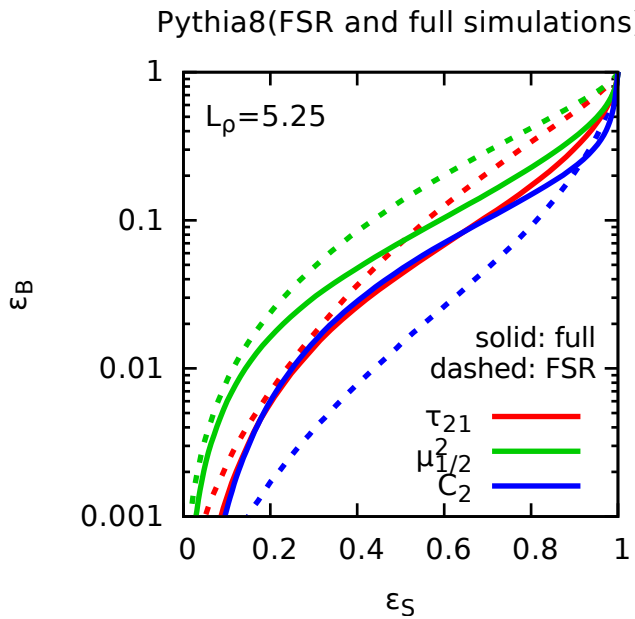


- Good description of the order between shapes.

ROC curves

- Efficiency for the QCD background ε_B vs. signal ε_S.
- C_2 is the most efficient due to large-angle constraints, and τ_{21} more efficient than μ^2 (more delicate call).

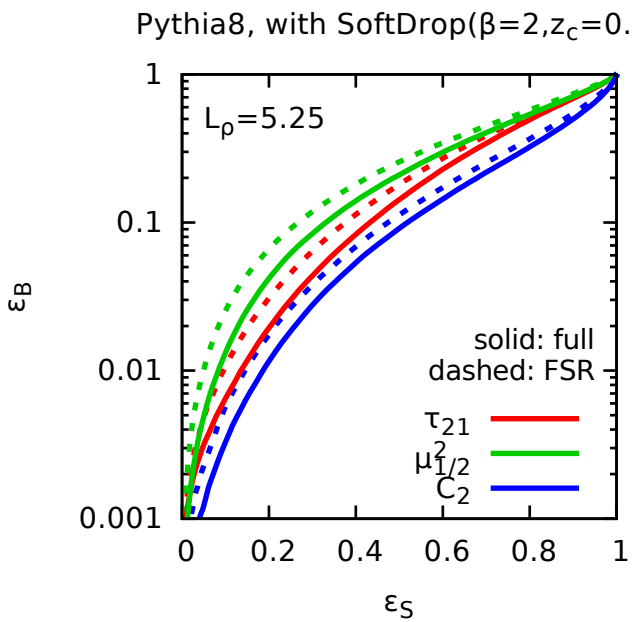


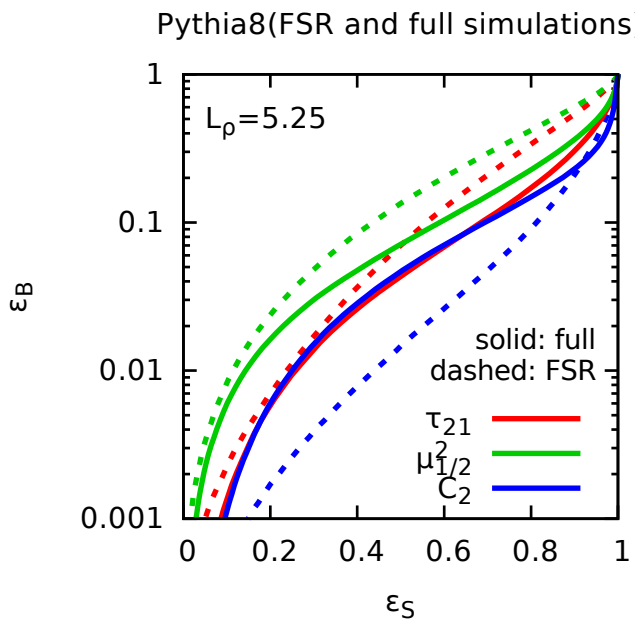


- Constraints at large θ also make C_2 more sensitive to non-perturbative effects.

ROC curves

- Efficiency for the QCD background ε_B vs. signal ε_S.
- C_2 is the most efficient due to large-angle constraints, and τ_{21} more efficient than μ^2 (more delicate call).





- Soft drop reduces NP effects but decreases efficiency.

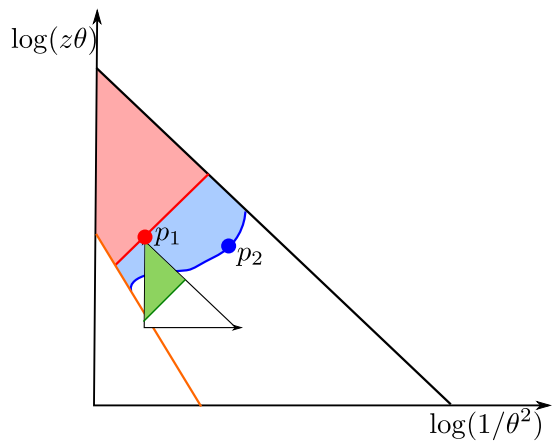
Presentation Plan

Conclusion

- **Efficiency** of the shapes: $C_2 \gtrsim \tau_{21} \gtrsim \mu^2$.
- **Recursive vs. non-recursive versions**
  - Non-recursive version is more efficient.
  - Drawback: larger sensitivity to non-perturbative effects.
- **Future works**
  - Grooming (soft drop or mMDT);

    A. J. Larkoski, S. Marzani, G. Soyez, J. Thaler (2014)

    M. Dasgupta, A. Fregoso, S. Marzani, G.Salam (2013)
  - $\beta \neq 2$ shapes;
  - ISR contributions;
  - Calculations for finite v.



Backup slides

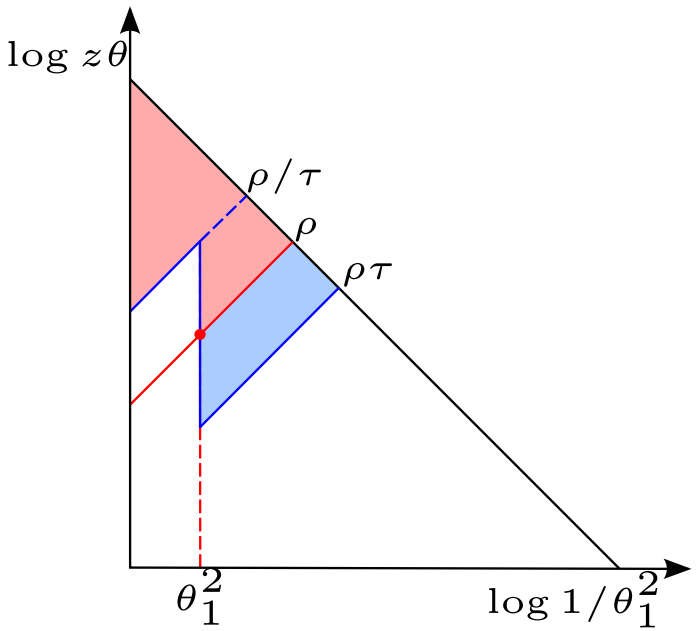
Generalized k_t algorithm

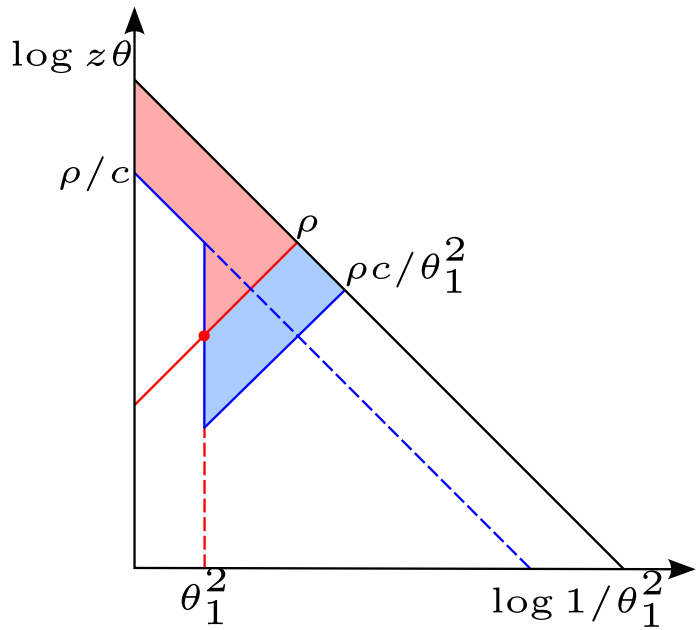
- Variation of the k_t algorithm with additional parameter p.
- Cluster partons by smallest distance $d_{ij} = \min(z_i^{2p}, z_j^{2p})\theta_{ij}^2$.
- Particular cases:
  - p=0 : C/A algorithm, angular ordered;
  - p=-1 : anti-k_t algorithm;
  - p=1/2 : similar to mass measure.

Recursive shapes: results

- Apply recursive C/A declustering until jet cut is met.
- Changes the allowed phase-space, let's suppose p_1 is the particle that determines the mass after the cut is met.
  - For the region $\theta_2 < \theta_1$, declustering reaches p_1 before p_2 leaving the same situation as the non-recursive case ($z_2\theta_2^2 < \rho\tau, \rho\mu^2, \rho c/\theta_1$).
  - For the region $\theta_2 > \theta_1$, emissions p_2 are allowed if they fail the v cut or if they pass the cut but do not change the jet mass ($z_2\theta_2^2 < \rho/\tau, \rho/\mu^2$ or $z < \rho/c$).

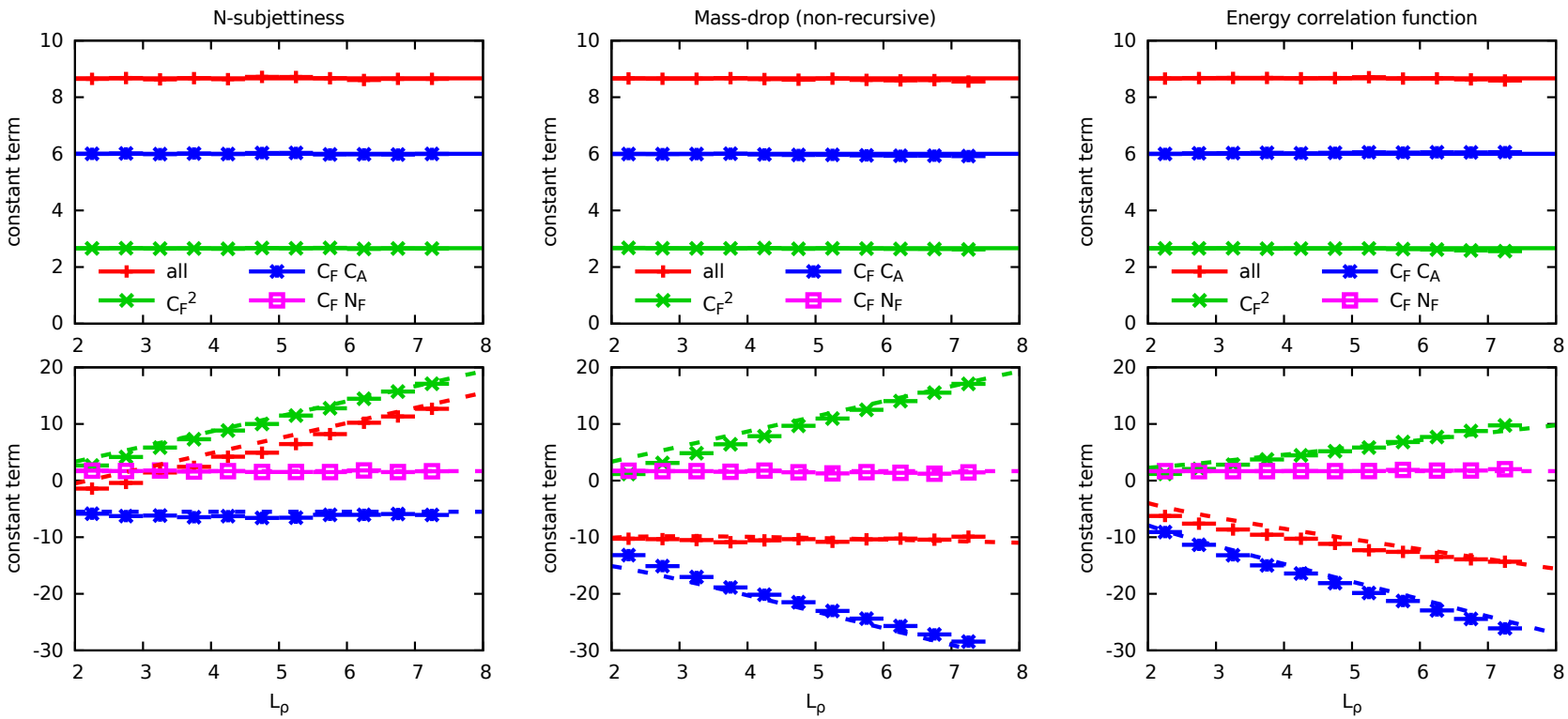
Recursive shapes: diagrams





QCD : Fixed-order Monte-Carlo

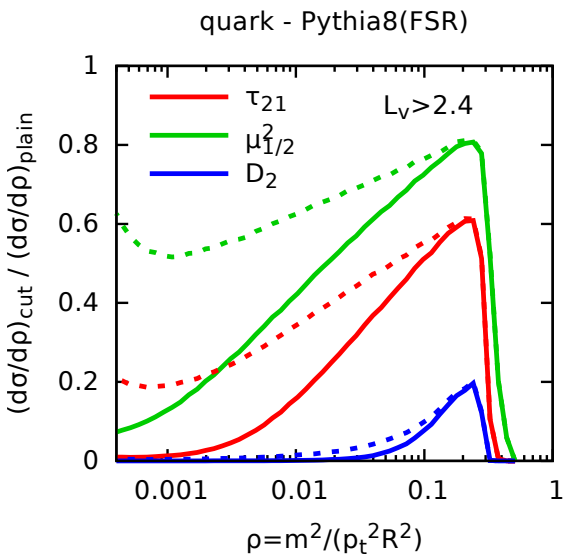
- Fit of the coefficients of the constant term and of the L_v term.

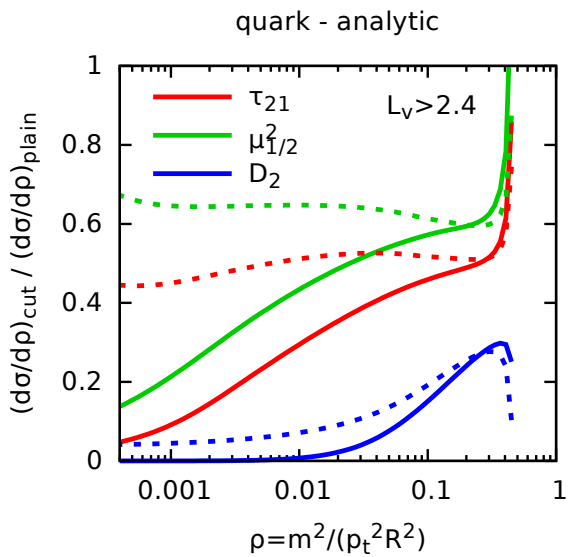


QCD : Parton-shower Monte-Carlo

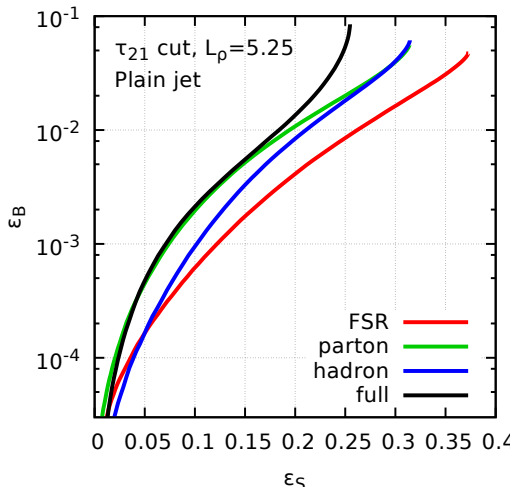
- Analysis for mass distribution at a fixed cut:

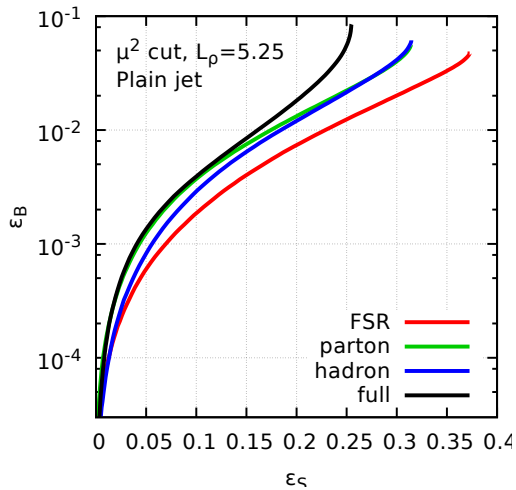
  [Dashed lines are recursive shapes, solid lines are non-recursive]

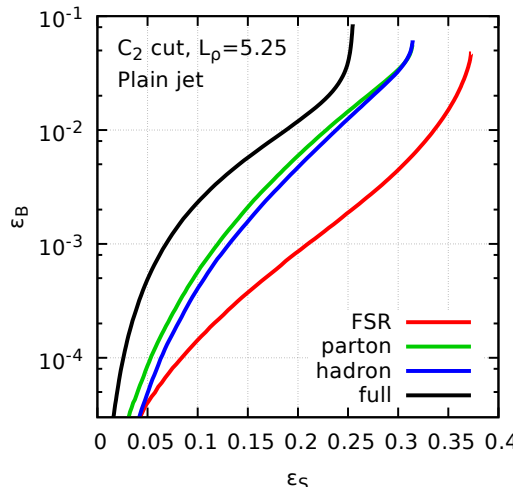


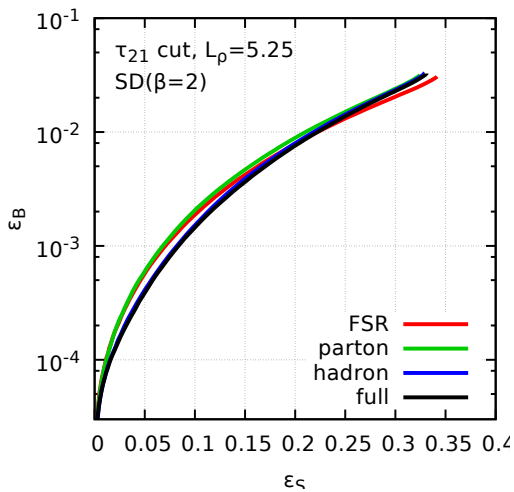


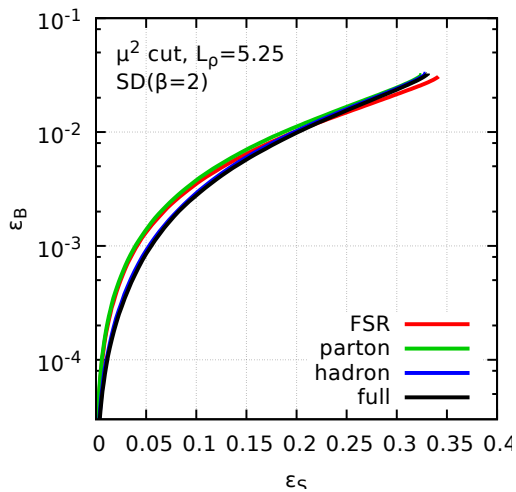
Additional ROC curves

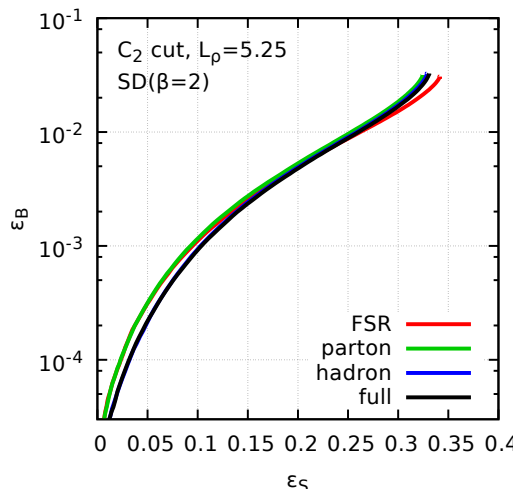












QCD fake rate for signal 50% efficiency

[Dashed line is LL only, solid line includes some NLL corrections]