Frascati Physics Series Vol. 73 (2022)
LFC22: Strong interactions from QCD to new strong dynamics at LHC and Future Colliders
August 29 - September 2, 2022

WEIGHING THE TOP WITH ENERGY CORRELATORS

Massimiliano Procura
University of Vienna, Faculty of Physics, Boltzmanngasse 5, A-1090 Vienna, Austria
Jack Holguin
CPHT, CNRS, Ecole polytechnique, IP Paris, F-91128 Palaiseau, France
Ian Moult
Department of Physics, Yale University, New Haven, CT 06511
Aditya Pathak
University of Manchester, School of Physics and Astronomy, Manchester, M13 9PL, United Kingdom

Abstract

Final states in collider experiments are characterized by correlation functions, $\langle \mathcal{E}(\vec{n}_1) \cdots \mathcal{E}(\vec{n}_k) \rangle$, of the energy flow operator $\mathcal{E}(\vec{n}_i)$. We show that the top quark imprints itself as a peak in the three-point correlator at an angle $\zeta \sim m_t^2/p_T^2$, with m_t the top quark mass and p_T its transverse momentum, thereby providing access to one of the most important parameters of the Standard Model in one of the simplest field theoretical observables. Our analysis represents the first step towards a novel precision top mass determination that is, for the first time, highly insensitive to soft physics and underlying event contamination whilst remaining directly calculable from the Standard Model Lagrangian.

1 Introduction

The top quark mass plays a central role both in determining the structure of the electroweak vacuum and in the consistency of precision Standard Model fits. A field theoretic definition of m_t, and its relation to experimental measurements, though, is notoriously subtle [1, 2]. At future e^+e^- colliders, high precision m_t measurements from the threshold lineshape will be possible. At present, the remarkably small quoted uncertainties on m_t from direct extractions at the LHC have been argued to be potentially affected by an additional $\mathcal{O}(1 \text{ GeV})$ contribution stemming from the theoretical interpretation of the measured "Monte Carlo (MC) top mass parameter" (for quantitative estimates, see e.g. refs. [4, 5, 6]). It is thus crucial and timely to explore kinematic top-mass sensitive observables at the LHC where a direct comparison of experimental data with accurate first principles theory predictions can be carried out.

Significant progress has been made in this regard from multiple directions. A unique feature of the LHC is that large numbers of top quarks are produced with enough boosts to decay into single collimated jets on which jet shapes can be measured. Using Soft Collinear Effective Theory and boosted

Heavy Quark Effective Theory, factorization theorems have been derived for event shapes measured on boosted top quarks, enabling these observables to be expressed in terms of m_t in a field theoretically well-defined mass scheme. In this framework, a paradigmatic example is given by the groomed jet mass. While jet grooming significantly improves the robustness of the observable on which it is applied, the complicated residual non-perturbative corrections [3] continue to be a limiting factor in achieving a precision competitive with direct measurements. This motivates the exploration of further m_t-sensitive observables that do not rely on jet grooming.

In recent years, intriguing progress has been made within a program aiming to rethink [7] jet substructure in terms of correlation functions, $\langle \mathcal{E}(\vec{n}_1) \cdots \mathcal{E}(\vec{n}_k) \rangle$, of the energy flow $\mathcal{E}(\vec{n})$ in a direction $\vec{n}$ [8, 9, 10], motivated also by the original work in QCD [11]. These correlators have a number of unique and remarkable properties. Most importantly for phenomenological applications, correlators are mostly insensitive to soft radiation without the application of grooming. Additionally they can also be computed on tracks [7, 12], using the formalism of track functions [13], allowing for higher angular resolution and pile-up suppression. However, so far these applications have been restricted to massless quark or gluon jets.

In [14] we have presented the first steps towards a new precision m_t measurement based on the simple idea of exploiting the mass dependence of the characteristic opening *angle* of the decay products of the boosted top, $\zeta \sim m_t^2/p_T^2$. The motivation for rephrasing the question in this manner is twofold. First, this angle can be accessed via low-point correlators, which are field theoretically drastically more simple than a groomed substructure observable sensitive to ζ. Second, while the jet mass is sensitive to soft contamination and UE, the angle ζ is not, since it is primarily determined by the hard dynamics of the top decay. In the following, we will illustrate a numerical proof-of-principles analysis showing that the three-point correlator in the vicinity of $\zeta \sim m_t^2/p_T^2$ provides a simple, but highly sensitive probe of m_t, free of the typical challenges of jet-shape based approaches. Our goal is to provide the motivation to perform future precision analyses and to find solutions to outstanding theoretical problems concerning low-point correlators relevant to the top mass determination and novel jet substructure studies.

2 The Three-Point Energy Correlator

There has recently been significant progress in understanding the perturbative structure of correlation functions of energy flow operators. This includes the landmark analytical calculation of the two-point correlator at next-to-leading order (NLO) in QCD [15] as well as the first calculation of a three-point correlator [16] at LO. The idea of using the three-point correlator to study the top quark is a natural one, and was considered early on in the jet substructure literature [17]. However, recent theoretical progress enables us now to make concrete steps towards a comprehensive program of using energy correlators as a precision tool for Standard Model measurements [7, 18].

The three-point correlator (EEEC) with generic energy weights is defined as

$$G^{(n)}(\zeta_{12}, \zeta_{23}, \zeta_{31}) = \int \mathrm{d}\sigma \, \widehat{\mathcal{M}}^{(n)}(\zeta_{12}, \zeta_{23}, \zeta_{31}) \,, \tag{1}$$

with the measurement operator given by

$$\widehat{\mathcal{M}}^{(n)}(\zeta_{12}, \zeta_{23}, \zeta_{31}) = \sum_{i,j,k} \frac{E_i^n E_j^n E_k^n}{Q^{3n}} \delta\left(\zeta_{12} - \hat{\zeta}_{ij}\right) \delta\left(\zeta_{23} - \hat{\zeta}_{ik}\right) \delta\left(\zeta_{31} - \hat{\zeta}_{jk}\right) \,. \tag{2}$$

Here $\hat{\zeta}_{ij} = (1 - \cos(\theta_{ij}))/2$, with θ_{ij} the angle between particles i and j, the sum runs over all triplets

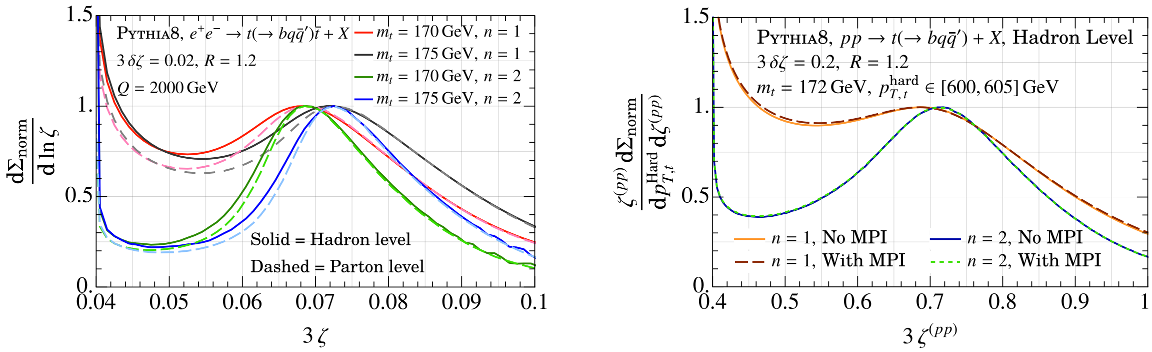

Figure 1: Weighted cross sections from three-point energy correlators using Pythia8. *Left panel:* The $n = 1, 2$ three-point correlators on boosted tops in e^+e^- annihilations showing a clear peak at $\zeta \sim 3m_t^2/Q^2$. All curves are normalized to peak height. *Right panel:* The $n = 1, 2$ three-point correlators on decaying top quarks with a fixed hard p_T in proton-proton collisions, with and without MPI. Here a clear peak can be seen at $\zeta \approx 3m_t^2/p_{T,t}^2$.

of particles in the jet, and Q denotes the hard scale in the measurement. It is worth stressing that the EEEC is not an event-by-event observable, but rather is defined as an ensemble average.

We are interested in the limit $\zeta_{12}, \zeta_{23}, \zeta_{31} \ll 1$, such that all directions of energy flow lie within a single jet. In the case of a conformal field theory (or massless QCD up to the running coupling), the small-angle limit of the EEEC simplifies due to the rescaling symmetry along the light-like direction defining the jet. In our case, m_t explicitly breaks this rescaling symmetry and appears as a characteristic scale imprinted in the three-point correlator. While the top quark has a three-body decay at leading order, higher-order corrections give rise to additional radiation, which is primarily collinear to the decay products leading to a growth in the distribution at angles $\hat{\zeta}_{ij} \ll m_t^2/p_T^2$. To extract m_t, we therefore focus on the correlator in a specific energy flow configuration sensitive to the hard decay kinematics. In [14] the simplest configuration is studied, that of an equilateral triangle $\hat{\zeta}_{ij} = \zeta$ allowing for a small asymmetry ($\delta\zeta$). Thus the key object of our analysis is the nth energy weighted cross section

$$\frac{d\Sigma(\delta\zeta)}{dQ d\zeta} = \int d\zeta_{12} d\zeta_{23} d\zeta_{31} \int d\sigma \widehat{\mathcal{M}}_{\triangle}^{(n)}(\zeta_{12}, \zeta_{23}, \zeta_{31}, \zeta, \delta\zeta)\,, \tag{3}$$

where the measurement operator $\widehat{\mathcal{M}}_{\triangle}^{(n)}$ is

$$\widehat{\mathcal{M}}_{\triangle}^{(n)}(\zeta_{12}, \zeta_{23}, \zeta_{31}, \zeta, \delta\zeta) = \widehat{\mathcal{M}}^{(n)}(\zeta_{12}, \zeta_{23}, \zeta_{31})\delta(3\zeta - \zeta_{12} - \zeta_{23} - \zeta_{31}) \prod_{l,m,n\in\{1,2,3\}} \Theta(\delta\zeta - |\zeta_{lm} - \zeta_{mn}|)\,. \tag{4}$$

For $\delta\zeta \ll \zeta$,

$$\frac{d\Sigma}{d\zeta} \approx 4(\delta\zeta)^2\, G^{(n)}(\zeta, \zeta, \zeta; m_t)\,, \tag{5}$$

where we have made the dependence on m_t explicit. Three-body kinematics implies that the distribution is peaked at $\zeta_{\rm peak} \approx 3m_t^2/Q^2$, exhibiting quadratic sensitivity to m_t. At the LHC the peak is resilient to collinear radiation since $\ln \zeta_{\rm peak} < 1/\alpha_{\rm s}$, making its properties computable in fixed-order perturbation theory at the hard scale. In the region $\zeta < 2\delta\zeta$ the hard three-body kinematics is no longer identified, leading to a bulge in the distribution, as shown in the *Supplemental Material* in ref. [14].

3 Mass Sensitivity in e^+e^-

To illustrate the mass sensitivity of our observable, we begin with the simplest case of e^+e^- collisions. We simulate the $e^+e^- \to t + X$ process at a center of mass energy of $Q = 2000$ GeV using the PYTHIA8 parton shower and reconstruct anti-k_T jets with $R = 1.2$. Although jet clustering is not required in e^+e^-, this analysis strategy is chosen to achieve maximal similarity with the case of hadron colliders.

In the left panel of Fig. 1 we show the distribution of the three-point correlator in the peak region, both with and without the effects of hadronization. Agreement of the peak position with the leading-order expectation is found, showing that the observed behavior is dictated by the hard decay of the top. In Fig. 1, linear ($n = 1$) and quadratic ($n = 2$) energy weightings are used, see Eq. (2). The latter is not collinear safe, but the collinear IR-divergences can be absorbed into moments of the fragmentation functions or track functions.

Crucially, non-perturbative effects in energy correlators are governed by an additive underlying power law [19, 9], which over the width of the peak has a minimal effect on the normalized distribution. This is confirmed by the small differences in peak position between parton and hadron level distributions in Fig. 1. Taking $m_t = 170, 172$ GeV with $n = 2$ as representative distributions, we find that the shift due to hadronization corresponds to a $\Delta m_t^{\text{Had.}} \sim 250$ MeV shift in m_t. This is in contrast with the groomed jet mass case where hadronization causes peak shifts equivalent to $\Delta m_t^{\text{Had.}} \sim 1$ GeV [20].

4 Hadron Colliders

We now extend our discussion to the more challenging case of proton-proton collisions. This study illustrates the difference between energy correlators and standard jet shape observables, and also emphasizes the irreducible difficulties of jet substructure at hadron colliders.

At variance with the case of e^+e^- annihilations, the hadronic final states in proton-proton collisions on which the energy correlators are computed are necessarily defined through a measurement, e.g. by selecting anti-k_T jets with a specific $p_{T,\text{jet}}$. Due to the insensitivity of the energy correlators to soft radiation, it is in fact the non-perturbative effects on the jet p_T selection that are the only source of complications in a hadron collider environment [14]. This represents a significant advantage of our approach, since it shifts the standard problem of characterizing non-perturbative corrections to infrared jet shape observables, to characterizing non-perturbative effects on a *hard* scale. This enables us to propose a methodology for the precise extraction of m_t in hadron collisions by independently measuring the universal non-perturbative effects on the $p_{T,\text{jet}}$ spectrum. We now illustrate the key features of this approach.

The three-point correlator in hadron collisions is defined as

$$\widehat{\mathcal{M}}^{(n)}_{(pp)}(\zeta_{12}, \zeta_{23}, \zeta_{31}) = \sum_{i,j,k\,\in\,\text{jet}} \frac{(p_{T,i})^n (p_{T,j})^n (p_{T,k})^n}{(p_{T,\text{jet}})^{3n}} \delta\left(\zeta_{12} - \hat{\zeta}^{(pp)}_{ij}\right) \delta\left(\zeta_{23} - \hat{\zeta}^{(pp)}_{ik}\right) \delta\left(\zeta_{31} - \hat{\zeta}^{(pp)}_{jk}\right) , \quad (6)$$

where $\hat{\zeta}^{(pp)}_{ij} = \Delta R^2_{ij} = \sqrt{\Delta\eta^2_{ij} + \Delta\phi^2_{ij}}$, with η, ϕ the standard rapidity, azimuth coordinates.

The peak of the EEEC distribution is determined by the hard kinematics and is found at $\zeta^{(pp)}_{\text{peak}} \approx 3m_t^2/p^2_{T,t}$, where $p_{T,t}$ is the hard top p_T, *not* $p_{T,\text{jet}}$.

To clearly illustrate the distinction between the infrared measurement of the EEEC and the hard measurement of the $p_{T,\text{jet}}$ spectrum, we present a two-step analysis using data generated in PYTHIA8 [14]. First, we generated hard top quark states with definite momentum (like in e^+e^-), but in the more complicated LHC environment including the underlying event (UE). This is shown in the right panel of

Fig. 1, where we see a clear peak that is *completely* independent of the presence of MPI (the PYTHIA8 model for UE), which illustrates that the correlators themselves, on a perfectly characterized top quark state, are insensitive to soft radiation *without* grooming.

In ref. [14] a proof-of-principles analysis was performed to illustrate that a characterization of non-perturbative corrections to the $p_{T,\mathrm{jet}}$ spectrum allows us to extract m_t, with small uncertainties from non-perturbative physics. To extract a value of m_t, we write the peak position as

$$\zeta_{\mathrm{peak}}^{(pp)} = \frac{3F_{\mathrm{pert}}(m_t, p_{T,\mathrm{jet}}, \alpha_s, R)}{(p_{T,\mathrm{jet}} + \Delta_{\mathrm{NP}}(R) + \Delta_{\mathrm{MPI}}(R))^2}\,, \tag{7}$$

where F_{pert} incorporates the effects of perturbative radiation. At leading order, $F_{\mathrm{pert}} = m_t^2$. Corrections from hadronization and MPI are encoded through the shifts $\Delta_{\mathrm{NP}}(R)$ and $\Delta_{\mathrm{MPI}}(R)$. Crucially, in the factorization limit that we consider, these are not a property of the EEEC observable, but can instead be extracted directly from the non-perturbative corrections to the jet p_T spectrum [21]. This is a unique feature of our approach.

The next step would be to calculate F_{pert} at NLO in perturbative QCD within a well-defined short-distance top mass scheme (such as the MSR [22]) and use the result to extract m_t according to the procedure described below. However, since the computation of F_{pert} has not been performed yet, in order to illustrate the feasibility of our approach, we have used PYTHIA8 (including hadronization and MPI) to extract $\zeta_{\mathrm{peak}}^{(pp)}$ as a function of $p_{T,\mathrm{jet}}$, over an energy range within the expected reach of the high luminosity LHC. As a proxy for the perturbative calculation, we used parton-level simulations to extract F_{pert}. To the accuracy we are working, F_{pert} is independent of the jet p_T, and can just be viewed as an effective top mass $\sqrt{F_{\mathrm{pert}}}(m_t)$. We also extract $\Delta_{\mathrm{NP}}(R) + \Delta_{\mathrm{MPI}}(R)$ independently from the $p_{T,\mathrm{jet}}$ spectrum.

Using Eq. (7) we fit $\zeta_{\mathrm{peak}}^{(pp)}$ as a function of $p_{T,\mathrm{jet}}$ for an effective value of $F_{\mathrm{pert}}(m_t)$. With a perfect characterization of the non-perturbative corrections to the EEEC observable, the value of $F_{\mathrm{pert}}(m_t)$ extracted when hadronization and MPI are included should exactly match its extraction at parton level. This would lead to complete control over m_t. In Table 1 we show the extracted value of $F_{\mathrm{pert}}(m_t)$ from our parton level fit, and from our hadron+MPI level fit for two values of the PYTHIA8 m_t. The errors quoted are the statistical errors on the parton shower analysis. The Hadron+MPI fit is quoted with two errors: the first originates from the statistical error on the EEEC measurement, the second stems from the statistical error on the determination of $\Delta_{\mathrm{NP}}(R) + \Delta_{\mathrm{MPI}}(R)$ from the $p_{T,\mathrm{jet}}$ spectrum. A more detailed discussion of this procedure can be found in the *Supplemental Material* in [14]. Thus we find promising evidence that theoretical control of m_t, with conservative errors $\lesssim$ 1GeV, is possible with an EEEC-based measurement. We stress that systematically improvable calculations of $F_{\mathrm{pert}}(m_t)$ within our approach are made feasible by a factorization formula for the weighted cross section discussed in ref. [14]. Theory errors are contingent upon currently unavailable NLO computations, see the discussion in [14]. However, we expect observable-dependent NLO theory errors on m_t to be better than those in other inclusive measurements wherein in the dominant theory errors are from PDFs+α_{s} [23, 24] and which mostly affect the normalization of the observable. By contrast the EEEC is also inclusive but the extracted m_t is only sensitive to the observable's shape.

Our promising results motivate developing a deeper theoretical understanding of the three-point correlator of boosted tops in the hadron collider environment. Nevertheless, there remain many areas in which our methodology could be improved to achieve greater statistical power and bring it closer to experimental reality. These include the optimization of $\delta\zeta$, the binning of $p_{T,\mathrm{jet}}$ and $\zeta^{(pp)}$, and including other shapes on the EEEC correlator. Regardless, our analysis does demonstrate the observable's potential for

Pythia8 m_t	Parton $\sqrt{F_{\rm pert}}$	Hadron + MPI $\sqrt{F_{\rm pert}}$
172 GeV	172.6 ± 0.3 GeV	$172.3 \pm 0.2 \pm 0.4$ GeV
173 GeV	173.5 ± 0.3 GeV	$173.6 \pm 0.2 \pm 0.4$ GeV
175 GeV	175.5 ± 0.4 GeV	$175.1 \pm 0.3 \pm 0.4$ GeV
$173 - 172$	0.9 ± 0.4 GeV	1.3 ± 0.6 GeV
$175 - 172$	2.9 ± 0.5 GeV	2.8 ± 0.6 GeV

Table 1: The effective parameter $F_{\rm pert}(m_t)$ extracted at parton level, and hadron+MPI level. The consistency of the two simulations provides a measure of our uncertainty due to uncontrolled non-perturbative corrections. Statistical errors are shown.

a precision m_t extraction when measured on a sufficiently large sample of boosted tops. We are optimistic that such a sample will be accessible at the HL-LHC where it is forecast that $\sim 10^7$ boosted top events with $p_T > 500\,\text{GeV}$ will be measured [25]. Our results support the possibility of achieving complete theoretical control over an observable with top mass sensitivity competitive with direct measurements whilst avoiding the ambiguities associated with the usage of MC event generators.

References

1. G. Corcella, Front. in Phys. **7**, 54 (2019).
2. A.H. Hoang, Ann. Rev. Nucl. Part. Sci. **70**, 225 (2020).
3. A.H. Hoang, S. Mantry, A. Pathak and I.W. Stewart, JHEP **12**, 002 (2019).
4. J. Kieseler, K. Lipka and S.-O. Moch, Phys. Rev. Lett. **116**, 162001 (2016).
5. M. Butenschoen, B. Denhadi, A.H. Hoang, V. Mateu, M. Preisser and I.W. Stewart, Phys. Rev. Lett. **117**, 232001 (2016).
6. A.H. Hoang, S. Plätzer and D. Samitz, JHEP **10**, 200 (2018).
7. H. Chen, I. Moult, X.Y. Zhang and H.X. Zhu, Phys. Rev. D **5**, 054012 (2020).
8. N.A. Sveshnikov and F.V. Tkachov, Phys. Lett. B **382**, 403 (1996).
9. G.P. Korchemsky and G.F. Sterman, Nucl. Phys. B **555**, 335 (1999).
10. D. Hofman and J. Maldacena, JHEP **05**, 012 (2008).
11. C.L. Basham, L.S. Brown, S.D. Ellis and S.T. Love, Phys. Rev. Lett. **41**, 1585 (1978).
12. Y. Li, I. Moult, S. Schrijnder van Velzen, W.J. Waalewijn and H.X. Zhu, Phys. Rev. Lett **18**, 182001 (2022).
13. H.-M. Chang, M. Procura, J. Thaler and W.J. Waalewijn, Phys. Rev. Lett. **111**, 102002 (2013).
14. J. Holguin, I. Moult, A. Pathak and M. Procura, arXiv:2201.08393 [hep-ph].
15. L.J. Dixon, M.-X. Luo, V. Shtabovenko, T.-Z. Yang and H.X. Zhu, Phys. Rev. Lett. **120**, 102001 (2018).
16. H. Chen, M.-X. Luo, I. Moult, T.-Z. Yang, X. Zhang and H.X. Zhu, JHEP **08**, 028 (2020).

17. M. Jankowiak and A.J. Larkoski, JHEP **06**, 057 (2011).

18. P.T. Komiske, I. Moult, J. Thaler and H.X. Zhu, arXiv:2201.07800 [hep-ph].

19. G.P. Korchemsky and G.F. Sterman, Nucl. Phys. B **437**, 415 (1995).

20. A.H. Hoang, S. Mantry, A. Pathak and I.W. Stewart, Phys. Rev. D **100**, 074021 (2019).

21. M. Dasgupta, L. Magnea and G.P. Salam, JHEP **02**, 055 (2008).

22. A.H. Hoang, A. Jain, I. Scimemi and I.W. Stewart, Phys. Rev. Lett **101**, 151602 (2008).

23. S. Chatrchyan et al., Phys. Lett. B **728**, 496 (2014).

24. A.M. Sirunyan et al., JHEP **09**, 051 (2017).

25. P. Azzi et al., CERN Yellow Rep. Monogr. **7**, 1 (2019).