Conformal Colliders Meet the LHC

Kyle Lee,1 Bianka Mečaj,2† and Ian Moul†
1 Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
2 Department of Physics, Yale University, New Haven, CT 06511

The remarkably high energies of the Large Hadron Collider (LHC) have allowed for the first measurements of the shapes and scalings of multi-point correlators of energy flow operators, \( \langle \Psi | \mathcal{E}(\vec{n}_1) \mathcal{E}(\vec{n}_2) \cdots \mathcal{E}(\vec{n}_k) | \Psi \rangle \), providing new insights into the Lorentzian dynamics of quantum chromodynamics (QCD). In this Letter, we use recent advances in effective field theory to derive a rigorous factorization theorem for the light-ray density matrix, \( \rho = |\Psi\rangle \langle \Psi| \), inside high transverse momentum jets at the LHC. Using the light-ray operator product expansion, the scaling behavior of multi-point correlators can be computed from the expectation value of the twist-2 spin-\( J \) light-ray operators, \( \mathcal{O}^{[J]} \), in this state, \( \text{Tr}[\rho \mathcal{O}^{[J]}] \). We compute the light-ray density matrix at next-to-leading order, and combine this with results for the next-to-leading logarithmic scaling behavior of the correlators up to six-points, comparing with CMS Open Data. This theoretical accuracy allows us to resolve the quantum scaling dimensions of QCD light-ray operators inside jets at the LHC. Our factorization theorem for the light-ray density matrix at the LHC completes the link between recent developments in the study of energy correlators and LHC phenomenology, opening the door to a wide variety of precision jet substructure studies.

Introduction.—The Large Hadron Collider (LHC) provides an opportunity to explore quantum field theory in general, and quantum chromodynamics (QCD) in particular, at unprecedented energy scales, and with modern resolution detectors [1, 2]. Due to the phenomenon of asymptotic freedom [3–6], this gain in energy is particularly advantageous, as it enables QCD to be studied in the perturbative regime, where first principles calculations are currently possible.

One of the major achievements, which re-invigorated the study of QCD at the LHC, was the introduction of experimentally robust infrared safe jet algorithms, most notably the anti-\( k_T \) algorithm [7–10], that allow for the identification of high transverse momentum, \( p_T \), jets in hadronic collisions. The inclusive production of such jets can be studied using rigorous factorization theorems [11–17], whose perturbative components have been computed to next-to-next-to-leading accuracy [18–24], allowing for precision studies of QCD in hadron colliders.

However, many of the most fascinating questions about QCD, namely understanding the Lorentzian dynamics of quarks and gluons, and the nature of their real-time confinement into hadrons, are not encoded in the distribution of jets, but rather in the structure of energy flow within jets, known as jet substructure [25, 26]. Jet substructure has been extraordinarily successful as a new way to search for physics beyond the Standard Model [27–29], and provides new opportunities to study the dynamics of QCD both in vacuum and medium [30, 31].

From a theoretical perspective, jet substructure is the study of the statistical properties of the asymptotic energy flux in collider experiments. It was placed in a modern field theoretic context in [32], where it was shown that it can be formulated as the study of correlation functions of \( \langle \Psi | \mathcal{E}(\vec{n}_1) \mathcal{E}(\vec{n}_2) \cdots \mathcal{E}(\vec{n}_k) | \Psi \rangle \) of particular light-ray

\[
\mathcal{E}(i\vec{n}) = \lim_{r \to \infty} \int_0^{\infty} dt \, s^2 \mathcal{T}_0(t, r\vec{n}),
\]

(ANEC) operators [32–39]

measured inside high energy jets at the LHC, as illustrated in Fig. 1. For simplicity, we will occasionally use a shorthand notation \( \mathcal{E}(\vec{n}_i) \to \mathcal{E}_i \).

An improved understanding of the properties of ANEC operators has been central to many recent advances in diverse areas of QFT, ranging from constraining conformal field theory (CFT) data [32, 30, 43], to energy inequalities [44–47], to asymptotic symmetries [48, 49] and the study of entropy in QFT [50–54]. Excitingly, reformulating jet substructure such that it can draw from these diverse areas has led to significant recent progress. See e.g. [32, 37, 49, 55–75]. In particular, this has enabled multi-point correlation functions of ANEC operators to

FIG. 1: An illustration of the two-point correlation function, \( \langle \Psi | \mathcal{E}(i\vec{n}_1) \mathcal{E}(i\vec{n}_2) | \Psi \rangle \), measured inside high-\( p_T \) jets produced in collisions at the LHC.
be studied using the light-ray operator product expansion (OPE) \[52\] \[60\] \[67\] \[77\] \[78\] and celestial blocks \[60\] \[67\] \[73\] \[74\], much in analogy with the study of multi-point correlation functions of local operators.

In the last several years there has been a program \[69\] to bridge the gap between these formal developments and real world phenomenology.\[71\] The goal of this program is to enable new theoretical techniques to be applied to real world collider data, extending our understanding of jet substructure, and conversely, to enable field theoretically interesting multi-point correlators of light-ray operators to be studied experimentally for the first time \[76\] \[78\]. In this Letter, we take the final step, using recent developments in effective field theory (EFT) to derive rigorous factorization theorems allowing calculations of energy correlators to be seamlessly embedded into the complex LHC environment.

The Light-Ray Operator Product Expansion and Light-Ray Densities—The leading scaling behavior of multi-point correlators of energy flow operators in the small angle limit, as is relevant for the study of energy flow within jets at the LHC, is captured by the leading twist operators appearing in the light-ray OPE. At weak coupling, and for unpolarized jets, these are the twist-2 spin-\(J\) light-ray operators \[32\] \[70\],

\[
\mathcal{O}^{[J]} = \left( \mathcal{O}_q^{[J]}(\bar{q}, q) \right)^T = \lim_{r \to \infty} r^2 \int_0^\infty dt \mathcal{O}^{[J]}(t, r\vec{n}) ,
\]

obtained from the light transform \[39\] of the standard twist-2 quark and gluon operators

\[
\mathcal{O}_q^{[J]} = \frac{1}{2J} \bar{\psi} \gamma^+ (iD^+) J^{-1} \psi ,
\]

\[
\mathcal{O}_g^{[J]} = -\frac{1}{27} F_a^{\mu\nu}(iD^+) J^{-3} F_a^{\mu\nu} .
\]

Performing the iterated OPE of light-ray operators, one finds that at leading twist \[32\] \[72\] \[77\],

\[
\mathcal{E}(\vec{n}_1) \mathcal{E}(\vec{n}_2) \cdots \mathcal{E}(\vec{n}_k) = \frac{1}{R_L^2} \left\{ f_q^{[k]}(u_i, v_i) \mathcal{O}_q^{[k+1]}(\vec{n}_1) \right\} + \mathcal{O}(R_L^0) ,
\]

where \(R_L\) is a scaling variable which for concreteness we take to be the largest angle between the energy flow operators, and \(f_q^{[k]}(u_i, v_i)\) are non-trivial functions of the cross ratios for \(k > 2\).

The light-ray OPE reduces the study of multi-point correlation functions to the study of the matrix elements \(\langle \mathcal{O}_q^{[J]}(\vec{n}_1) \rangle = \langle \mathcal{O}_q^{[J]}(\vec{n}_1) | \mathcal{O}_q^{[J]}(\vec{n}_2) | \Psi \rangle\), where \(|\Psi\rangle\) is the state for an identified jet at the LHC. Formally, this state can be described as being sourced by a soft collinear effective theory (SCET) quark or gluon field \[79\] \[83\].

It is useful to rewrite this expectation value in terms of the so called light-ray density matrix \(|\rho \rangle = |\Psi \rangle \langle \Psi |\), as \(\langle \mathcal{O}_q^{[J]} | \mathcal{O}_q^{[J]} \rangle = \text{Tr}(\rho \mathcal{O}_q^{[J]} \mathcal{O}_q^{[J]} )\). Much like for the light-ray operators, it is convenient to work with states of definite quantum numbers under the Lorentz group, namely the celestial quantum numbers \(\delta\) associated with boosts, and \(J\) associated with transverse rotations \(|\Psi \rangle = \sum_{\delta, J} |\Psi_{\delta, J} \rangle\). A primary complexity of jet physics is that standard observables are sensitive to the infinite sum over states of arbitrary quantum numbers. This has led to different approaches, such as grooming \[54\] \[55\] to attempt to simplify it. The energy correlators allow for a far more dramatic simplification by exploiting the underlying symmetries of the theory. The measurement of \(\mathcal{O}_q^{[J]}\) projects onto the single state with definite quantum numbers \(|\Psi_{\delta, J} \rangle\). Measurements of the low point correlators project onto the particularly simple low \(\delta\) states, ultimately leading to their simple theoretical structure. The OPE has therefore reduced the calculation of multi-point correlators at the LHC, to computing the probability to produce jet states \(|\Psi_{\delta, J} \rangle\). This represents a clear departure from the way in which jets are more traditionally studied at the LHC, heavily motivated by developments in CFT.

\[\text{Factorization Theorem for Light-Ray Densities.—}A \]

\[\text{factorization theorem for the energy correlators in the collinear limit in } e^+ e^- \text{ colliders was derived in } [72], \]

\[\text{allowing them to be systematically resummed at higher}\]
perturbative orders. The main complexity in extending this factorization theorem to the case of hadron colliders is the need to measure the energy correlators inside an identified jet defined with the anti-$k_T$ algorithm used experimentally. The derivation of such a factorization theorem is the main result of this Letter.

Our factorization theorem is derived by combining the seminal work of Collins-Soper-Sterman on factorization for inclusive hadron production, with recent developments in SCET, in particular the fragmenting jets formalism, which allows for the factorization of states produced inside high $p_T$ jets. Combining these approaches, we derive a factorization theorem for generic correlation functions of energy flow operators inside inclusive high-$p_T$ jets at the LHC.

\[
\frac{d\Sigma}{dp_T d\eta d\{\zeta\}} = \sum_i \mathcal{H}_i (p_T/z, \eta, \mu) \otimes \int_0^1 dx x^N J_{ij}(z, x, p_T R, \mu) J_i^{[N]}(\{\zeta\}, x, \mu). \tag{5}
\]

Here, $\mathcal{H}_i$ is the hard function describing the production of parton $i$, which includes the parton distribution functions of the hadron beams. The convolution structure $\otimes$ is over the parton momentum fraction $x$. The matching coefficient $J_{ij}$ incorporates the details of the jet algorithm, and depends explicitly on the jet radius, $R$. Combined, $\mathcal{H}_i$ and $J_{ij}$ the light-ray density. The energy correlator jet function, $J^{[N]}_i$, depends on a momentum fraction $x$, and the set of angles between the detectors, $\{\zeta\}$, which can be written in terms of cross ratios and $R_L$, as in Eq. (4). This factorization theorem enables any calculation of the energy correlators in the collinear limit to be embedded into the LHC environment. A particularly beautiful aspect of this factorization is that each component satisfies a DGLAP evolution, illustrating a natural interplay of the energy correlator and fragmenting jet formalisms.

As compared with previous approaches to jet substructure, we believe that the approach of correlation functions, encapsulated in Eq. (5), presents a number of advantages. First, for an arbitrary $N$-point correlator, both the normalization and the shape dependence are free of soft effects and non-global logarithms, without the application of grooming procedures that complicate higher order calculations. Second, unlike for groomed observables, where the algorithm complicates the non-perturbative structure of the observable

---

4. At leading logarithmic order, one can use the jet calculus. However, this does not allow one to take into account higher order corrections associated with the source.

5. For other phenomenological applications of the fragmenting jet function formalism see e.g. [93–100].

FIG. 3: Ratios of the multi-point projected correlators compared with CMS Open Data points extracted from [76]. These ratios probe the quantum scaling associated with higher spin light-ray operators, $\mathcal{O}^{[J]}$, providing a quantitative test of the light-ray OPE.
all algorithm dependence lies in the matching coefficient, $f_{ij}$, simplifying the structure of non-perturbative corrections \[^{[55,109,112]}\]. Third, since the inclusive hadron hard function is known at next-to-next-to-leading order (NNLO) \[^{113}\], and the fragmenting jet functions can be computed at NNLO using the approach of \[^{114}\], the path to higher perturbative orders is clear. Finally, correlation functions of energy flow operators can be computed on tracks using the track function formalism for energy correlators \[^{115–118}\], allowing one to exploit the exceptional angular resolution of tracking detectors at the LHC, and push towards higher-point and higher-precision correlators. For recent progress towards understanding the structure of higher-point correlation functions of energy correlators in perturbation theory, see \[^{[61,73–75,78]}\].

The Two-Point Correlator $\langle \mathcal{E}(\vec{n}_1) \mathcal{E}(\vec{n}_2) \rangle$.—The simplest possible jet substructure observable is the two-point correlator of energy flow operators, $\langle \mathcal{E}(\vec{n}_1) \mathcal{E}(\vec{n}_2) \rangle$, also referred to as the energy-energy correlator (EEC) \[^{[119–122]}\]. As the most basic correlator, it has been studied extensively in perturbation theory \[^{[55,58–60]}\], and the OPE has been understood non-perturbatively in CFTs \[^{[66,67]}\]. In addition to controlling the structure of asymptotic energy flow in this theory, this OPE also controls modular flow \[^{[62]}\]. While the two-point correlator was measured early on in $e^+e^-$ colliders \[^{[123–132]}\], the high energies of the LHC allowed for the first studies of the scaling behavior in the small angle limit \[^{[76]}\].

In Fig. 2 we plot our analytic calculation of the two-point energy correlator at both leading logarithm (LL) and next-to-leading logarithm (NLL) \[^{[62]}\]. We use the NLO hard functions of \[^{[138–143]}\], the NLO fragmenting jet functions from \[^{[90,91]}\] and the NLO energy correlator jet constants from \[^{[62]}\]. To focus on the convergence of the perturbative series, as illustrated by the decreasing scale variations from LL to NLL, we have chosen to normalize the analytic results to the data within the perturbative regime. Comparisons without this normalization will be given in the next section. We compare our results with CMS Open Data extracted from \[^{[76]}\] observing excellent agreement. The beautiful scaling behavior governed by the light-ray OPE, which is reminiscent of critical phenomena, is clearly visible!

As emphasized above, contrary to the standard lore of the field, no grooming algorithms were needed even in the complex LHC environment. This is a remarkable illustration of the universality of the OPE limit in QFT, and highlights the role of formal theory in understanding the correct field theoretic observables for quantifying correlations in energy flux at colliders.

In the comparison between theory and data in Fig. 2 no unfolding has been done on the CMS Open Data. The theoretical advances described in this Letter motivate precise experimental measurements of the energy correlator observables. The non-standard nature of these observables suggests that new unfolding procedures may be required, such as the recently introduced machine learning (ML) based approaches \[^{[151]}\], which would be a beautiful union of new ideas from data science and formal theory to better understand QCD at colliders.

Higher-Point Projected Correlators.—In addition to the two-point correlator, we can probe the scaling behavior of higher point correlation functions. We consider the projected correlators, $\text{ENC}(R_L)$, defined by integrating out the non-trivial shape dependence of the $N$ point correlator, keeping only the longest side $R_L$ fixed \[^{[69]}\]. The projected correlators were first studied in LHC data in \[^{[76]}\]. To isolate the quantum scaling, and suppress non-perturbative corrections, we consider the ratio to the two-point correlator, $\text{ENC}(R_L)/\text{E2C}(R_L)$ \[^{[69]}\].

In Fig. 3 we show a comparison of our NLL analytic calculations of the projected correlators, again compared with CMS Open data \[^{[76]}\]. We find excellent agreement in the perturbative regime. The deviation as $R_L \rightarrow 0.5$ is due to boundary effects from the jet clustering. In the case that the anomalous dimensions vanished, these curves would be flat. These ratios probe in detail the anomalous dimensions of the higher spin light-ray op-

\[^{6}\] See \[^{[122,127]}\] for studies in the Sudakov regime.
\[^{7}\] Here we use the logarithm counting appropriate for single logarithmic collinear observables, as opposed to that for double logarithmic observables.
\[^{8}\] This data was made publicly available \[^{[144]}\] by the CMS collaboration \[^{[115]}\]. The CMS \[^{[2011A Open Data}][147]\ was simplified into the “MIT Open Data” (MOD) format in Refs. \[^{[148,149]}\]. It was analyzed using the Energy-Energy Correlators package \[^{[150]}\] in \[^{[76]}\]. We thank Patrick Komiske for his development and maintenance of this package.
substructure is its simplicity and lack of grooming algorithms, and the iterated light-ray OPE! The increasing slope occurs due to the monotonic increase of the twist-2 anomalous dimensions as a function of $N$. This is well reproduced in the NLL calculation showing sensitivity to the scaling at the quantum level.

The NLL accuracy achieved in this Letter is essential for clearly resolving the different scaling behaviors of the $N$-point correlators. In Fig. 1 we show all of our NLL predictions in a single plot, and to focus on the scaling, we have identified them at $R_L = 0.01$. With the reduced uncertainty bands at NLL, we are able to begin to clearly separate the scaling behaviors. This is not possible at LL due to the much larger uncertainty bands (see e.g. Fig. 2 as representative). This strongly motivates extending our calculation to NNLL to further reduce the theoretical uncertainties, as well as performing a proper experimental analysis with robust error bars. Both of these goals are achievable in the near future, and will allow us to resolve the structure of the quantum scaling dimensions of QCD within high-$p_T$ jets!

Our results are particularly interesting for comparing with implementations of higher order DGLAP in parton showers [155-159], since the projected correlators cleanly probe the low spin DGLAP anomalous dimensions in a way that has not been possible previously. Such “reference resummations” will play an essential role in the development of improved parton showers (see e.g. [155-163] for recent work), which will in turn extend our ability to search for new phenomena at the LHC.

Conclusions.—In this Letter we have bridged the gap between a wealth of developments in the study of conformal collider physics and real world LHC phenomenology. This was achieved by deriving a rigorous factorization theorem for the light-ray density matrix at the LHC, in which the energy correlators observables can be evaluated. We computed this density matrix at NLO, and combined it with the NLL scaling of the light-ray operators, matching the state of the art accuracy for jet substructure at the LHC, but for a family of observables. To our knowledge, this is the first factorization theorem for an infrared safe jet substructure observable at the LHC, that is based on a rigorous QCD factorization theorem in QCD, and absent non-global logarithms, including in the normalization.

Although we have focused in this Letter on applications to the LHC, the interface between energy correlators and fragmenting jet functions, described by the factorization theorem in Eq. (5), can also be applied in heavy ion colliders, or at the future electron-ion collider. This will be of interest for precision studies of hadronization, cold nuclear matter, and the quark-gluon plasma using jet substructure.

We believe that our results are of interest both phenomenologically and theoretically. On the phenomenological side, a key advantage of our formulation of jet substructure is its simplicity and lack of grooming algorithms, clearing the path towards higher order calculations and the ultimate hope of precision QCD phenomenology at the LHC. On the theoretical side, our results enable the experimental study of correlation functions of light-ray operators beyond the leading order. We hope that this combination can drive genuine progress in our understanding of jet substructure, leading to new insights into QCD, and ever more sophisticated ways to search for new physics and new interactions at the LHC.

Acknowledgements.—We thank Cyuan-Han Chang, Hao Chen, Patrick Komiske, Jingjing Pan, David Simmons-Duffin, Jesse Thaler, Feng Yuan, Hua Xing Zhu for helpful comments, discussions, questions, and collaboration on related work. We thank Felix Ringer for useful discussions and collaboration at the initial stages of the project. K.L. is supported by the LDRD program of LBNL, the U.S. DOE under contract number DE-AC02-05CH11231. B.M. and I.M. are supported by start-up funds from Yale University. I.M. thanks the KITP Santa Barbara for hospitality while parts of this work were completed.
