

Tagging a jet from a dark sector with jet substructures at colliders

Myeonghun Park^{1,2,*} and Mengchao Zhang^{3,2,†}

¹*Institute of Convergence Fundamental Studies and School of Liberal Arts, Seoultech, Seoul 01811, Korea*

²*Center for Theoretical Physics of the Universe, Institute for Basic Science (IBS), Daejeon 34051, Korea*

³*Department of Physics and Siyuan Laboratory, Jinan University, Guangzhou 510632, People's Republic of China*



(Received 8 May 2019; published 4 December 2019)

The phenomenology of the dark sector is complicated if the dark sector is charged under a confined hidden gauge group. In such a kind of model, a dark parton produced at a high-energy collider showers and hadronizes to a cluster of dark mesons. Dark mesons then decay to visible particles and produce a jetlike signal, which is called a “dark jet” in this work. Collider signal of a dark jet depend on the property of dark mesons. For example, a finite lifetime of a dark meson would provide a displaced vertex or displaced track; thus, one can use these displaced objects to tag a dark jet. However, if the lifetime of a dark meson is collider negligible (too short to manifest a displaced vertex), it would be difficult to distinguish a dark jet from standard model QCD jets. In this work, we propose a new tagging strategy to identify dark jets from QCD backgrounds. This strategy is based on jet-substructure analysis. We study various jet-substructure variables and find out variables with good discrimination ability. Our result shows that by combining multiple jet-substructure variables one could distinguish dark jets from QCD background and thus enhance the sensitivity of the dark sector search at a collider.

DOI: [10.1103/PhysRevD.100.115009](https://doi.org/10.1103/PhysRevD.100.115009)

I. INTRODUCTION

The existence of dark matter (DM) in our Universe has been confirmed indirectly with its gravitational effects [1]. Still, we have no idea about the nature of DM as we have not found DM “directly” with various DM experiments. Especially, weakly interacting massive particles (WIMPs) as the most popular DM paradigm have been a subject for various experiments including space indirect searches, nucleon scattering direct searches, and collider experiments. However, we have excluded a wide range in the parameter space of WIMPs from null results in the above searches [2–8]. In addition to the WIMP paradigm, another DM scenario called asymmetric DM [9–28], which is motivated by the coincidence of the abundance of visible matter and DM as $\Omega_{\text{DM}} \simeq 5\Omega_B$, has attracted attention. In the asymmetric DM paradigm, DM and its antiparticle anti-dark matter (aDM) are not produced equally in the early Universe. Subsequently, annihilation

between DM and aDM eliminates aDM in the Universe, and the remaining DM particles compose the current relic density.

To stabilize DM and annihilate aDM efficiently, a hidden gauge group is generally introduced in an asymmetric DM model. If some particle charged under this hidden gauge group can be produced at a collider, then it is possible to study the dark sector through final-state radiation of hidden gauge bosons. For example, if DM is charged under a $U(1)'$, then energetic DM produced at a collider will radiate a $U(1)'$ gauge boson (dark photon γ'). A dark photon decays back to the standard model (SM) particles through a kinetic mixing with a SM photon and leads to prompt/long-lived lepton jets or a narrow jet signal at the collider [29–31]. If a hidden gauge group in the dark sector is $SU(N_d)$, which causes confinement at a certain scale Λ_d , then an energetic dark parton, which is the particle charged under $SU(N_d)$, will shower multiple dark partons then hadronize to a cluster of dark hadrons (most of them are the lightest dark meson). Dark hadrons then decay to particles in the SM, through some portals, and produce a jetlike signal. In this work, we will call it a “dark jet.”

The property of dark jets depends on the dark sector setting and the portal between the dark sector and the SM. Previous phenomenology studies of different kinds of dark jets can be found in Refs. [32–45]. These studies generally rely on displaced objects from long-lived dark

*parc.seoultech@seoultech.ac.kr

†Corresponding author.
mczhang@jnu.edu.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

meson decay.¹ For example, in Ref. [41], the authors make the dark meson long lived by introducing a heavy mediator through which the dark meson decays to a SM quark pair. Thus, there will be some displaced tracks inside the dark jet, and one can utilize these displaced tracks to enhance the collider search ability. Recently, a new dark jet study based on a flavor structure in the dark quark sector called the semivisible jet was proposed in Ref. [39]. In their scenario, a certain amount of missing energy, which comes from some stable light dark hadrons, is collimated with the dark jet, and a transverse mass of two leading jets in the final states becomes useful to discriminate the dark jets pair signal from SM background.

All these search methods become ineffective if all or most of the light dark mesons decay to SM particles promptly because in this case there are no displaced objects or missing energy inside the dark jet and the dark jet looks like a SM QCD jet. Inversely, if all or most of the dark hadrons are collider stable, then there is only missing energy in the final state. In this work, we will not study this WIMP-like signal. For illustration, in Fig. 1, we categorize dark jet signatures according to the lifetime of dark hadrons and the fraction of invisible particles inside a jet. Here, we categorize dark jet searches into four categories:

- (i) Exotic (I): One can identify a dark jet via displaced objects (DOs) induced by long-lived dark hadron decay.
- (ii) Exotic (II): Some stable dark hadrons (a dark baryon and also some dark mesons) occupy a non-negligible portion of a dark jet, which makes various kinematic variables useful.
- (iii) WIMP-like: Dark hadrons are stable or collider stable (CS). In this case, the signal is like WIMP.
- (iv) SM QCD-like: There are no displaced objects nor missing energy inside the dark jet, and the dark jet looks like a SM QCD jet.

To distinguish a ‘‘SM QCD-like’’ dark jet from a SM QCD background jet, we suggest looking inside a jet and utilizing various jet-substructure techniques. Because of recent improvements in quark-gluon jet discrimination with jet-substructure and corresponding applications in different new physics searches [46–50], we argue that we are at the stage of discriminating dark QCD jets from SM QCD jets. Jet substructure reveals the underlying structure of a QCD-like model, such as a color factor or confinement scale. Thus, it is possible for us to discriminate between a jet from dark QCD and a jet from SM QCD, provided dark QCD is not the same as SM QCD.

In next section, we briefly introduce our model and discuss which model setting can cause a ‘‘SM QCD-like’’

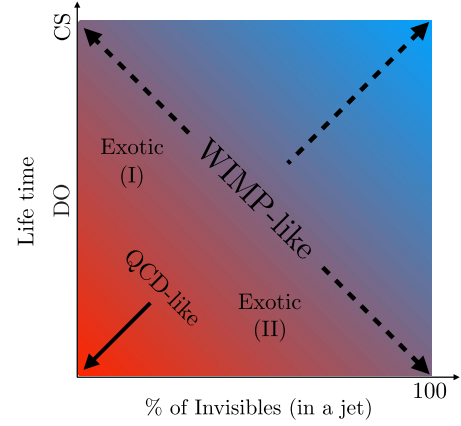


FIG. 1. We present a diagram to illustrate different kinds of dark jets from dark QCD in terms of the (x-axis) percentage of stable (invisible) hadrons in a jet and (y-axis) lifetime of dark mesons. Here, CS means a lifetime enough to be ‘‘collider stable,’’ and DO stands for a sizable lifetime to be tagged with ‘‘displaced objects.’’

dark jet. Section III is dedicated to a comprehensive study of various jet-substructure variables and an exhibition of their discriminant performance. In Sec. IV, we use an example at the LHC to show the effect of our dark jet tagging method. Then, we summarize this work in Sec. V. A brief discussion of theoretical uncertainties from Monte Carlo simulation will be given in Appendix A. In Appendix B, we discuss the feasibility of our method. More details of our benchmark points can be found in Appendix C.

II. BENCHMARK SCENARIOS FOR DARK QCD MODELS

We introduce a new non-Abelian gauge group $SU(N_d)$ which describes dynamics in the dark sector. Several light dark quarks as fundamental representations of $SU(N_d)$ are also required for the constitution of the dark hadron. Here, a light dark quark means a dark quark which is lighter than the dark confinement scale Λ_d . For $SU(N_d)$ to confine, the number of light dark quarks flavors n_f should be smaller than $\frac{11}{2}N_d$. At an energy scale much higher than Λ_d , the Lagrangian of the dark sector can be written as

$$\mathcal{L}_d = \bar{q}'_i (i\not{D} - m_{q'_i}) q'_i - \frac{1}{4} G'^{\mu\nu} G'_{\mu\nu}, \quad (1)$$

with q' and $G'^{\mu\nu}$ denoting the dark quarks and dark gluon field strength, respectively. D_μ corresponds to the covariant derivative of $SU(N_d)$, and i is the flavor index of dark quarks. For minimality, we set the dark quarks to be a SM singlet. To produce dark partons at collider and decay dark hadrons, a mediator between the dark sector and SM sector is required. It could be a bifundamental scalar X which is charged under both $SU(N_d)$ and SM $SU(3)$ [40],

$$\mathcal{L}_{\text{med}} = (D^\mu X)^\dagger (D_\mu X) - M_X^2 X^\dagger X + \kappa_{ij} X \bar{q}'_i q'_j + \text{H.c.}, \quad (2)$$

¹In some scenarios, the lightest dark meson can decay to a bottom quark pair; thus, b tagging can be utilized. But it requires dark mesons to be quite heavy. We do not consider this scenario in this work.

or a heavy vector boson Z' mediator connecting the dark quark pair and SM quark pair [51]²

$$\mathcal{L}_{\text{med}} = -\frac{1}{4}Z'^{\mu\nu}Z'_{\mu\nu} - \frac{1}{2}M_{Z'}^2Z'^{\mu}Z'_{\mu} + Z'_{\mu}(\bar{q}'_i\gamma^{\mu}q'_i + \bar{q}_j\gamma^{\mu}q_j). \quad (3)$$

Here, q_i is a SM quark, and i and j are flavor indices of dark quarks and SM quarks. To make the dark meson decay promptly, one can also extend hidden a gauge group from $SU(N_d)$ to $SU(N_d) \times U(1)'$, with the dark quark also charged under $U(1)'$ [54]. Thus, the light dark meson could decay to a γ' pair immediately, and γ' decays to SM particles through kinetic mixing.

The decay of dark hadrons depends on their spin, mass, and the mediator to the visible sector. Here, we give a comprehensive analysis of different kinds of dark hadrons and point out under which model setting the dark jet is SM QCD-like:

- (i) Dark pion π_d : Generally, spin-0 dark pions π_d are much lighter than other dark hadrons, and they form a large part of particles in a dark jet until they decay. π_d decays to a quark pair through a high-dimensional effective operator. Because of a chiral flipping suppression, π_d tends to decay to a heavy SM quark pair, and its lifetime is closely related to the mass of the dark pion m_{π_d} . We use a formula from Ref. [41] to estimate the partial width of π_d to a SM quark pair:

$$\Gamma(\pi_d \rightarrow q\bar{q}) = \frac{3\kappa^4 f_{\pi_d}^2 m_q^2}{32\pi M_X^4} m_{\pi_d}. \quad (4)$$

Here, κ is the coupling among a mediator X , a SM quark q , and a dark quark q' . f_{π_d} is the decay constant of the dark pion, m_q is the pole mass of the SM quarks, and M_X is the mass of the mediator X . $\kappa \sim 1$ is a natural choice. And approximately, we have $m_{\pi_d}^2 f_{\pi_d}^2 \sim m_{q'} \Lambda_d^3$. Thus, if f_{π_d} and m_{π_d} are several GeV, and the π_d decay channel to the SM K meson is open, then a M_X around several hundred GeV can cause the proper decay length of π_d to be shorter than 1 mm, i.e., a promptly decaying dark pion. This range for a mediator mass is still allowed by previous searches [41]. If the mass of dark pion is greater than twice the mass of D meson, the parameter space for promptly decayed dark pion would be much larger.

Another possibility is the case in which there is an extra $U(1)'$ under which the dark quark is charged [54]. In this case, a dark pion will behave like a SM

pion, and it decays to a dark photon pair promptly. A dark photon can decay into SM particles through a kinetic mixing with SM hypercharge $U(1)_Y$, where the kinetic mixing is parametrized by ϵ . With current limits on the dark photon [55], we find there is still a huge surviving parameter space that can induce a prompt dark photon decay. For instance, a 0.4 GeV dark photon will decay promptly if $\epsilon \gtrsim 10^{-5}$. Thus, it induces the prompt decay of a dark pion into SM particles.

In Ref. [39] and a more recent paper [56], the authors consider a dark meson which is composed of different flavor dark quarks. In this case, a dark meson is stable and results in missing energy along with the dark jet. However, this dark flavor mixing meson can be unstabilized by introducing some dark flavor violating portal. For example, the interaction Lagrangian between two dark quark flavors and a mediator X can be written as

$$\mathcal{L}_{\text{int}} = \kappa_{11}\bar{q}'_1 q_1 X + \kappa_{21}\bar{q}'_2 q_1 X + \text{H.c.} \quad (5)$$

Here, q_1 is a SM quark, and q'_1 and q'_2 are two different flavor dark quarks. By integrating out the heavy mediator X , one can get an effective operator as³

$$\mathcal{L}_{\text{eff}} = \frac{\kappa_{11}\kappa_{21}^*}{M_X^2} (\bar{q}'_1\gamma_{\mu}q'_2)(\bar{q}_1\gamma^{\mu}q_1) + \text{H.c.} \quad (6)$$

So, depending on the parameters, the flavor mixing dark meson π_d ($q'_1\bar{q}'_2$ or $q'_2\bar{q}'_1$) could decay promptly into SM particles through this dark flavor violating operator.

- (ii) Dark rho meson ρ_d : A dark rho meson is a spin-1 bound state made of dark quarks. Generally, there is a mass splitting between a dark pion and a dark rho meson, which depends on the pole mass of dark quarks, because the mass of m_{π_d} and m_{ρ_d} can be approximated as

$$m_{\pi_d}^2 \sim m_{q'} \Lambda_d, \quad m_{\rho_d}^2 \sim \Lambda_d^2. \quad (7)$$

Thus, if $m_{q'} \ll \Lambda_d$, a dark rho meson will decay promptly through decay channel $\rho_d \rightarrow \pi_d\pi_d$. If $m_{q'}$ is comparable with Λ_d , the mass splitting might not be enough to allow double dark pion decay. But due to the spin-1 property of ρ_d , its decay width will not be chiral suppressed. Thus, compared with π_d , it is much easier for ρ_d to decay promptly. Discussion of

²Such a leptophobic Z' will easily induce the chiral anomaly [52,53], but this topic is not so related to our present work, so we ignore the chiral anomaly problem in this work.

³Fiertz transformation is used here. This kind of operator will not cause flavor-changing neutral currents if we only consider one SM quark flavor.

TABLE I. Benchmark models we considered in this work. All dark mesons are assumed to decay promptly. We mainly consider two cases: the high Λ_d case like A and C and the low Λ_d case like B and D. Parameters in a dark sector for A and C and B and D are the same, except the decay channel of a dark pion π_d . π_d and ρ_d masses obey following two equations: $m_{\pi_d} = 2\tilde{m}_{q'} - \frac{3}{4}\frac{\Omega}{\tilde{m}_{q'}}$ and $m_{\rho_d} = 2\tilde{m}_{q'} + \frac{1}{4}\frac{\Omega}{\tilde{m}_{q'}}$ [57], where Ω is proportional to the binding energy. Thus, m_{ρ_d} can be determined by constituent dark quark masses $\tilde{m}_{q'}$ and m_{π_d} . The ratio of π_d to ρ_d after fragmentation is decided by $\frac{\#\pi_d}{\#\rho_d} = \frac{1}{3}e^{-\frac{m_{\pi_d}-m_{\rho_d}}{\Lambda_d}}$. The branching ratios of their decay modes shown here are all 100%, if we do not give a specific value. Decay modes of a dark photon γ' with different masses can be found in Ref. [30].

	N_d	n_f	Λ_d (GeV)	$\tilde{m}_{q'}$ (GeV)	m_{π_d} (GeV)	m_{ρ_d} (GeV)	π_d Decay mode	ρ_d Decay mode
A	3	2	15	20	10	50	$\pi_d \rightarrow c\bar{c}$	$\rho_d \rightarrow \pi_d\pi_d$
B	3	6	2	2	2	4.67	$\pi_d \rightarrow s\bar{s}$	$\rho_d \rightarrow \pi_d\pi_d$
C	3	2	15	20	10	50	$\pi_d \rightarrow \gamma'\gamma'$ with $m_{\gamma'} = 4.0$ GeV	$\rho_d \rightarrow \pi_d\pi_d$
D	3	6	2	2	2	4.67	$\pi_d \rightarrow \gamma'\gamma'$ with $m_{\gamma'} = 0.7$ GeV	$\rho_d \rightarrow \pi_d\pi_d$

the flavor mixing case is similar to that of a dark pion, so we do not repeat it here.

- (iii) Dark baryon: The lightest dark baryon is stable, and thus it can be a dark matter candidate. In the $SU(3)_d$ case, the population ratio of baryons over mesons in a hadronization process is about 10%, which is negligible. If $N_d > 3$, the population ratio of the baryon will be further suppressed. Only in the $SU(2)_d$ case, a considerable part of the hadron in a dark jet consists of stable dark baryons. Thus, in this work, we choose $N_d = 3$ and neglect dark baryons in the dark hadronization process.
- (iv) Dark glueball: If all the dark quarks are much heavier than the confinement scale of $SU(N)_d$ ($m_{q'} \gg \Lambda_d$), the lightest dark hadron will be made of a dark gluon. Thus, one can call this dark hadron a dark glueball. As a dark gluon and SM gluon belong to different gauge groups, the decay of the dark glueball is loop induced by heavy particles which have charges of both gauge groups. Thus, the lifetime of the dark glueball will be quite long. We will not discuss this scenario in this work.

Based on the above discussion, we only consider dark mesons π_d and ρ_d in our simulation, and we let all the dark mesons decay promptly in an event generator. Detailed benchmark settings are listed in Table I. To cover the diversity of dark QCD models, we consider different confinement scales, spectra, and decay channels. Because of the nonperturbative nature of a QCD-like theory, some of those parameters need to be given by hand. In Table I, constituent quark mass $\tilde{m}_{q'}$ is used to estimate the dark hadron spectrum. And for simplicity, we assume π_d and ρ_d composed by different flavor dark quarks have identical mass and decay channels. In the next section, we will show how one can utilize jet-substructure variables to distinguish a dark jet from SM QCD jets.

III. JET-SUBSTRUCTURE VARIABLES ANALYSIS

Underlying parameters in a dark sector will affect the collider phenomenology of a dark jet. The RGE running of

a dark sector gauge coupling $\alpha_d(\mu)$ is controlled by these parameters,

$$\frac{d}{d \ln \mu^2} \alpha_d^{-1}(\mu) = \frac{1}{12\pi} (11N_d - 2n_f), \quad (8)$$

with boundary condition $\alpha_d^{-1}(\Lambda_d) = 0$. A comparison in a running coupling between SM QCD and various dark QCD models is shown in Fig. 2 (the corresponding dark sector setting can be found in Table I). Running coupling determines the parton shower, which happens at a short distance smaller than $1/\Lambda_d$. Then, those showered partons fragment to dark hadrons. Finally, dark hadrons decay back to SM particles, which are measured by a detector. Combining these three processes, the detector-level measurements of jet-substructure variables, like jet mass or track multiplicity for a dark jet, could be quite different from the expectations for SM QCD jets.

A dark jet originated from a single dark parton can be considered as a one-prong jet. Thus, jet grooming [58–60] methods including the mass dropping algorithm or pruning, which are suitable for reconstructing a boosted heavy object like a gauge boson ($W/Z/H$) or top quark, are

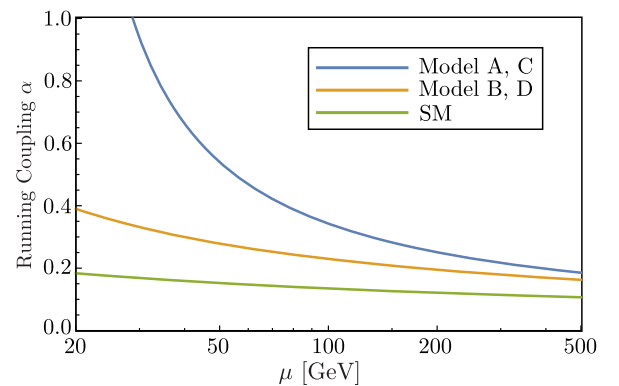


FIG. 2. QCD coupling running in the dark sector and SM QCD. The definition of A (C) and B (D) can be found in Table I. For SM QCD, we choose $n_f = 5$ as we do not include the effects from a top quark.

not expected to be effective in tagging a dark jet. Compared to two or three-prong jet tagging, one-prong jet tagging is easier due to a simpler jet structure. Jet-substructure variables used to tag a one-prong jet roughly fall into two categories: infrared collinear (IRC) safe ones and IRC unsafe ones.

An IRC safe variable is not sensitive to soft or collinear radiations inside the jet; i.e., contributions from extra radiation to an IRC safe variable are negligible if radiations are soft or collinear. Thus, an analytical description of IRC safe variables is possible. We choose jet mass, two-points energy correlation function $C_1^{(\beta)}$ [46], and linear radial geometric moment (girth) [61] as our IRC safe variables. Analytical descriptions of these IRC safe variables have been given in the literature; it helps us to understand our results, which are mainly based on Monte Carlo simulation.

An IRC unsafe variable, for example, the charged track multiplicity, is sensitive to soft and collinear radiations. Besides that, some IRC unsafe variables are also related to the dark meson decay channel. For those variables, we will provide Monte Carlo-based results and give some qualitative arguments.

We choose PYTHIA8 [62] for shower and hadronization simulation. It has been shown that jet substructures obtained from PYTHIA8 fits well with real data [63,64]. The Hidden Valley model [32] implemented in PYTHIA8 can be used to simulate the dark QCD process, and recently the running of the dark gauge coupling has been added to PYTHIA8, which greatly enhances the reliability of dark QCD simulation. We generate three processes at the LHC, $f\bar{f} \rightarrow Z' \rightarrow q'\bar{q}'$, $qg \rightarrow Zq$, and $q\bar{q} \rightarrow Zg$, to study the dark jet, quark jet, and gluon jet, respectively. For realistic analyses, we perform analyses at the detector level with DELPHES3 [65]. We use FASTJET [66] to cluster final-state particles with an anti-kt algorithm [67]. The objects for a jet clustering are energy deposits in an electric calorimeter, a hadronic calorimeter, and muons without an isolation criterion. Muons are also included in our jet clustering because there can be a fraction of dark jet energy carried by a muon, depending on the decay channel of a dark pion.⁴ Examining the discrimination performance of jet-substructure variables with different choices of jet radius (R), jet transverse momentum (p_T), and jet algorithms can be interesting. In our study, we choose $R = 0.4$ as it is a typical jet radius in the LHC experiment analyses for the QCD jet, and this choice was studied in the ATLAS light-quark and gluon jet discrimination [68]. For the choice of jet transverse momentum p_T , we start with the range of $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$ as this p_T range has the minimum

⁴In this case, some dedicated method can be designed to tag a muon-rich jet. But as we study the behavior of general jet substructure variables to cover various types of dark jets, we will not pay special attention to muons in this work.

systematic uncertainties [69] and it overlaps with the p_T range in the ATLAS jet discrimination study [68]. We consider a detector geometry of pseudorapidity $\eta \in (-2.5, 2.5)$.

A. Jet mass

Jet mass, as a simple and intuitive variable which reflects the underlying structure of a jet, has been studied for decades [70–75]. Jet mass originates from the virtuality of the primordial parton of a jet. As we consider the first-order splitting process, a normalized differential cross section of virtuality is

$$\frac{1}{\sigma} \frac{d\sigma}{dp^2} = \frac{C}{2\pi p^2} \int_{\epsilon}^{1-\epsilon} dz \alpha(zp^2) P(z, p^2), \quad (9)$$

where $\sigma = \int (d\sigma/dp^2) dp^2$ is the integrated jet cross section, C is the color factor, p is the 4-momentum of a primordial parton, and p^2 is its virtuality. ϵ is an infrared cut, z is the energy fraction carried by a radiated parton, and $\alpha(\mu)$ and $P(z, p^2)$ are QCD running coupling and splitting kernel, respectively. The above fixed-order result is divergent when a jet mass becomes zero, which is in conflict with experiment data. To get a reasonable distribution, one needs to resum higher-order corrections. In leading-log order, the differential cross section becomes

$$\frac{1}{\sigma} \frac{d\sigma}{dp^2} = \frac{d}{dp^2} S(p^2, Q^2), \quad (10)$$

which is a differential to the Sudakov factor $S(p^2, Q^2)$:

$$S(p^2, Q^2) = \exp \left\{ - \int_{p^2}^{Q^2} dk^2 \frac{C}{2\pi k^2} \int_{\epsilon}^{1-\epsilon} dz \alpha(zk^2) P(z, k^2) \right\}. \quad (11)$$

Here, Q is the energy scale of the corresponding hard process. This leading-order result can roughly reproduce the shape of the real data distribution from the LHC experiments. Obviously, this distribution is determined by running coupling $\alpha(\mu)$ and color factor C . To get an intuition for jet mass distributions, we approximate Eqs. (10) and (11) below. With fixing running $\alpha(\mu)$ as α , $P(z, k^2) = 1/z$, and choosing $\epsilon = p^2/Q^2$, we obtain the following approximation:

$$\frac{1}{\sigma_0} \frac{d\sigma}{d(p^2/Q^2)} \approx \frac{C\alpha}{\pi} \frac{Q^2}{p^2} \ln \frac{Q^2}{p^2} \exp \left\{ - \frac{C\alpha}{2\pi} \left(\ln \frac{Q^2}{p^2} \right)^2 \right\}. \quad (12)$$

As we see in Eq. (12), the peak of a jet mass distribution moves to the right side as $C\alpha$ becomes larger. Thus, the peak of a jet mass distribution for a gluon-initiated jet is on the right side compared to the peak of a distribution from a quark-initiated jet, as the color factor for a gluon

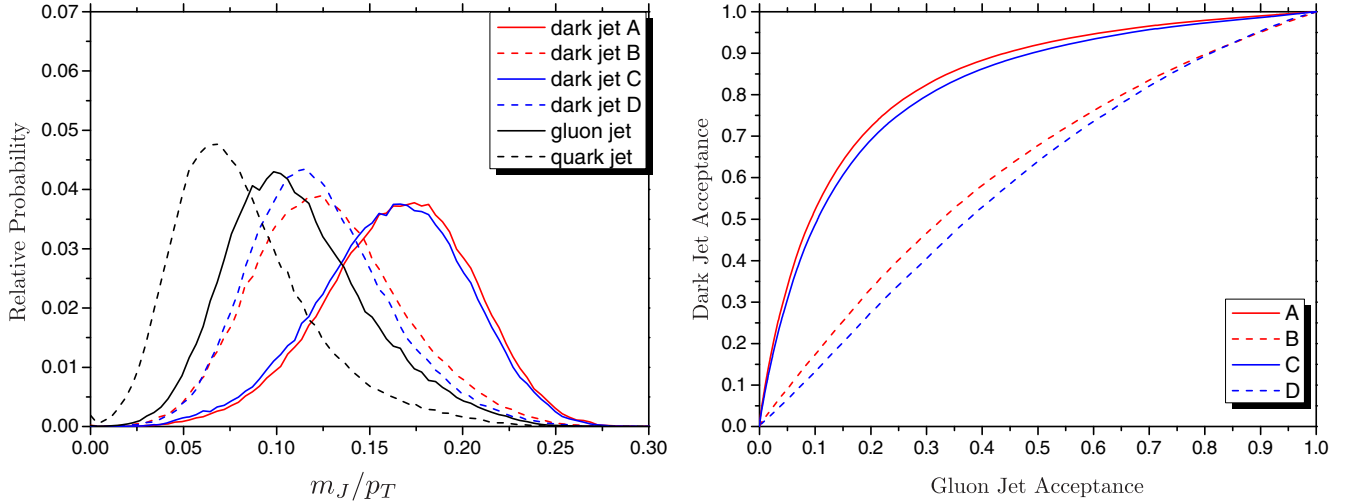


FIG. 3. Left: mass over p_T distribution of the dark jet, quark jet, and gluon jet with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$. Right: ROC curve with m_J/p_T used for the separation between a jet from various dark QCD models and a SM gluon-initiated jet.

$C_A = 3$ is larger than the color factor $C_F = 4/3$ of a quark as in Fig. 3.

In SM QCD, the only difference between a quark jet and gluon jet is the color factors C_F (for a quark) and C_A (for a gluon). Even so, a dimensionless parameter m_J/p_T , the jet mass divided by its p_T , is a good variable used in quark/gluon jet discrimination. For a dark jet, because of a quite different running coupling and a possible different color factor, one could certainly expect a very different distribution of a jet mass compared to the case of SM QCD.

With considering subleading contributions, one can include the effect of a jet size or a hadronization [72,75]. In our study, we will not go further analytically but utilize Monte Carlo simulation (PYTHIA8) to get numerical results. Jet mass distributions from different models in Table I and SM QCD are shown in Fig. 3.

As the gauge coupling strength of a dark QCD model A (C) is larger than the gauge coupling strength of B (D), a jet from model A or C has a larger mass than a jet from model B or D. Equivalently, a dark QCD model with a higher confinement scale Λ_d is easier to distinguish from SM QCD. We can check discrimination performance with receiver operating characteristic (ROC) curves in the right column of Fig. 3. We also argue that a jet mass is not sensitive to final states (SM particles from the decay of dark mesons) as jet mass distributions of A (B) almost overlap the distribution of C (D) in Fig. 3.

B. Two-point energy correlation function $C_1^{(\beta)}$

Another variable which is useful to probe properties of a one-prong jet is the two-point energy correlation function [46]

$$C_1^{(\beta)} = \sum_{i < j \in J} z_i z_j (R_{ij})^\beta, \quad (13)$$

with $z_i = p_{Ti} / \sum_{i \in J} p_{Ti}$ being the p_T fraction carried by component i within a jet J and R_{ij} being the distance between components i and j . As studied in Ref. [46], the advantage of an infrared collinear safe variable like $C_1^{(\beta)}$ is that analytical calculation of them is possible. Here, we adopt analytical results from Ref. [46] to see the dependence of $C_1^{(\beta)}$ on the parameters of dark QCD. First, one can consider the simplest case, which is the fixed leading-order distribution with treating α as a constant for simplicity:

$$\frac{1}{\sigma} \frac{d\sigma}{dC_1^{(\beta)}} = \frac{2\alpha C}{\pi} \int_0^{R_0} \frac{d\theta}{\theta} \int_0^1 \frac{dz}{z} \delta(z(1-z)\theta^\beta - C_1^{(\beta)}). \quad (14)$$

Here, R_0 is the size of a jet, which is the upper limit of a splitting angle in the shower process. After integrations, we get

$$\begin{aligned} \frac{1}{\sigma} \frac{d\sigma}{dC_1^{(\beta)}} &= \frac{2\alpha C}{\pi\beta C_1^{(\beta)}} \ln \left(\frac{1 + \sqrt{1 - 4C_1^{(\beta)}/R_0^\beta}}{1 - \sqrt{1 - 4C_1^{(\beta)}/R_0^\beta}} \right) \\ &\equiv \frac{2\alpha C}{\pi\beta C_1^{(\beta)}} L(C_1^{(\beta)}/R_0^\beta). \end{aligned} \quad (15)$$

Similarly to our previous fixed-order calculation for the distribution of a jet mass, $C_1^{(\beta)}$ distribution is also divergent in the soft and collinear regions. With a leading-order resummation, one obtains

$$\begin{aligned}
\frac{1}{\sigma} \frac{d\sigma}{dC_1^{(\beta)}} &= \frac{d}{dC_1^{(\beta)}} \exp\left(-\int_{C_1^{(\beta)}}^{R_0^\beta} d\tilde{C} \frac{2\alpha C}{\pi\beta\tilde{C}} L(\tilde{C}/R_0^\beta)\right) \\
&= \frac{2\alpha C}{\pi\beta C_1^{(\beta)}} L(C_1^{(\beta)}/R_0^\beta) \\
&\quad \times \exp\left(-\int_{C_1^{(\beta)}}^{R_0^\beta} d\tilde{C} \frac{2\alpha C}{\pi\beta\tilde{C}} L(\tilde{C}/R_0^\beta)\right). \quad (16)
\end{aligned}$$

One can notice that the probability in the soft and collinear region will be suppressed by an exponent. As we have seen in jet mass distribution, the peak value of dark jet $C_1^{(\beta)}$ distribution is larger than the peak value of SM QCD jet $C_1^{(\beta)}$ distribution, as dark QCD has a larger coupling compared to SM QCD.

There are two more factors that can enhance the discriminant power of $C_1^{(\beta)}$. First, there is a contribution from the nonperturbative fragmentation process. This effect can be estimated by convolving a resummed perturbative distribution with a so-called shape function [76,77]. The effect of this convolution is shifting the perturbative distribution of $C_1^{(\beta)}$ to a higher value, and the shift from this nonperturbative process is roughly proportional to the corresponding confinement scale. Thus, the fragmentation process will further separate $C_1^{(\beta)}$ distribution of the dark jet and SM QCD jet due to their different confinement scales.

Second, when the mass of a dark meson is much larger than SM QCD confinement scale Λ_{QCD} , the decay of dark mesons inside a jet will strongly affect the distribution of $C_1^{(\beta)}$. This effect can be understood by the following simple estimation. We consider two nearly collinear dark mesons inside a dark jet, with energy fractions z_1 and z_2 and distance θ between these two dark mesons. θ should be small because we assume these two dark mesons to be nearly collinear. In this case, contribution from these two mesons to $C_1^{(\beta)}$ is $z_1 z_2 \theta^\beta$. After both mesons decay to two SM particles with roughly equal energy, this contribution changes to

$$z_1 z_2 \theta^\beta \rightarrow \frac{1}{4} (z_1 + z_2)^2 \left(\frac{m_{\pi_d}}{\overline{p_T}}\right)^\beta. \quad (17)$$

Here, m_{π_d} is the mass of a dark meson, and $\overline{p_T}$ is the average transverse momentum of dark mesons inside a dark jet. As we consider a collinear limit between two dark mesons, an angular distance between dark meson decay products is approximated as $(m_{\pi_d}/\overline{p_T})$. Thus, the mass of a dark meson will increase $C_1^{(\beta)}$ of a dark jet as we consider $\beta > 0$. For a discrimination between a quark-initiated jet and a gluon-initiated jet, β has been chosen as 0.2 [46,49]. In this paper, we also follow this choice of $\beta = 0.2$ to distinguish a dark jet from a SM QCD jet.

Simulation results are shown in Fig. 4. First, we show $C_1^{(\beta)}$ distributions from parton level to detector level on the top row. Here, parton level $C_1^{(\beta)}$ means the objects we used to calculate $C_1^{(\beta)}$ is the dark parton after the dark shower and before dark hadronization; meson level $C_1^{(\beta)}$ comes from dark mesons after dark hadronization; particle level $C_1^{(\beta)}$ comes from all the visible SM particles after dark meson decay; and detector level $C_1^{(\beta)}$ comes from energy deposits at the detector. Top left and top right plots are distributions of $C_1^{(\beta)}$ for model A and model B, respectively. In the top left plot, there is a tall spike at $C_1^{(\beta)} = 0$ in the parton-level distribution. This spike comes from a large angle split where one of daughter partons is located outside the jet cone. At meson level, through convolution with a shape function, this spike at $C_1^{(\beta)} = 0$ becomes lower, and the distribution is shifted to a higher value. Together with this effect, due to the decay of dark mesons, the particle-level distribution of $C_1^{(\beta)}$ is pushed farther to the right side. Finally, the finite resolution of a detector decreases $C_1^{(\beta)}$ to a lower value. For model B, due to a weaker coupling and lower confinement scale compared to the case of model A, there is no tall spike of $C_1^{(\beta)} = 0$ at parton-level distribution, and dark meson decay pushes up $C_1^{(\beta)}$ only a little. In conclusion, jets from a dark QCD model with a high dark confinement scale jet are easier to tag over SM QCD jets compared to the case of a low dark confinement scale. We also observed that tagging efficiency is not sensitive to the decay channel of the dark meson as $C_1^{(\beta)}$ distribution for A (B) is similar to the distribution of C (D).

C. Linear radial geometric moment

Angularity-style variables including jet broadening or width have been studied since the Large Electron-Positron (LEP) collider period [50,78–82]. Here, we choose the linear radial geometric moment (girth) to study, which is known as an effective observable in discriminating between the quark and gluon jet [61]. Girth is defined as

$$\text{girth} = \sum_{i \in J} \frac{p_{T_i}}{P_{T_J}} |r_i|; \quad (18)$$

here, r_i is the distance between a component i of the jet and jet axis. Girth is sensitive to the direction of a jet axis compared to $C_1^{(\beta)}$, which does not require a jet axis. Here, the jet axis is defined as the vector sum of all the constituents inside a jet.

Girth, as a jet width variable, has been analytically analyzed in Ref. [76]. Here, we give a brief description, and readers can check more details in Ref. [76] if they are interested. At parton level, perturbative calculation shows that quark/gluon jet discrimination ability mainly relies on

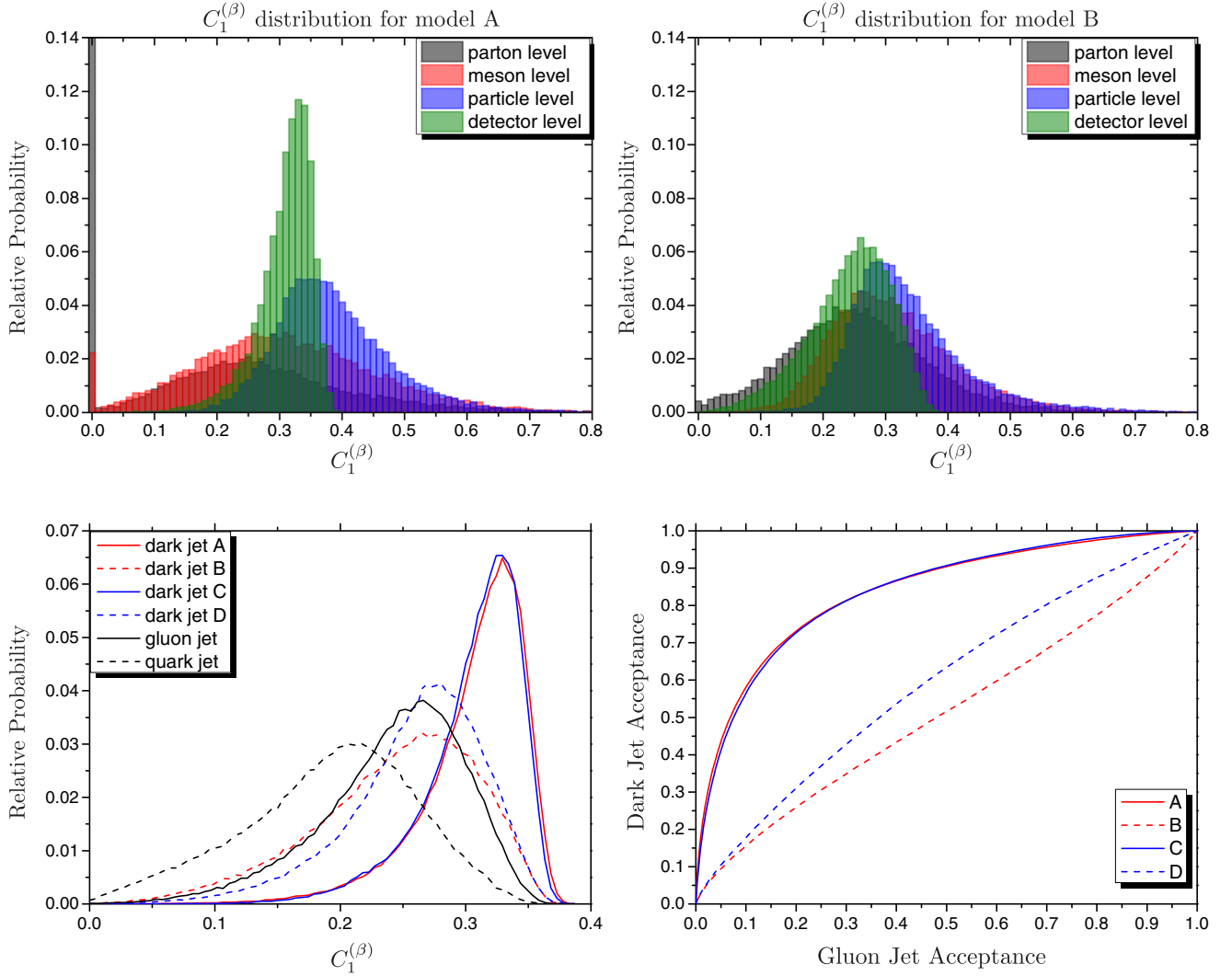


FIG. 4. Top left: $C_1^{(\beta)}$ distribution of a dark jet with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$ at parton level, meson level, final-state particle level, and detector level for a dark QCD model A (corresponding to a high dark QCD confinement scale). Top right: the same as the top left, but for dark QCD model B (corresponding to a low dark QCD confinement scale). Bottom left: $C_1^{(\beta)}$ distribution of different kinds of jets with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$. Bottom right: Corresponding ROC curves for discrimination between dark QCD jets and a SM QCD gluon-initiated jet. Here, β is set to 0.2.

the color factor ratio C_A/C_F , this is called Casimir scaling. For dark jet discrimination, due to a different coupling, the ratio should be replaced by $\alpha_S C_A/\alpha_d C_d$. Thus, one could expect a better discrimination power if α_d is quite different than α_S . Meson-level distribution, as we described in the last subsection, can be obtained by convoluting parton-level distribution with a shape function, which has a mean value proportional to the confinement scale. So, large $\Lambda_d/\Lambda_{\text{QCD}}$ will separate girth distribution of the dark jet and QCD jet further. Finally, decay of the heavy dark meson will push up the girth value of the dark jet.

Our results from simulations are presented in Fig. 5. In these results, we show the distribution of girth from model A and model B from parton level to detector level, as we did for $C_1^{(\beta)}$. The relationship between different levels

is as we expected, but the changes are not very much compared to $C_1^{(\beta)}$. This is because $C_1^{(\beta)}$ is more sensitive to small angular distribution. And unlike the case of the $C_1^{(\beta)}$ variable, which needs to have at least two components for non-zero value, girth has a nonzero value with even one component.⁵ Thus, a large angle parton splitting does not cause a zero-point spike in the distribution of girth as we can find in Fig. 4. We conclude that the performance of girth is dependent on the confinement scale of dark QCD as a dark jet from a higher confinement scale is easier to distinguish than cases from a low confinement scale. With

⁵Here, we use the jet axis obtained from detector level to calculate all girths.

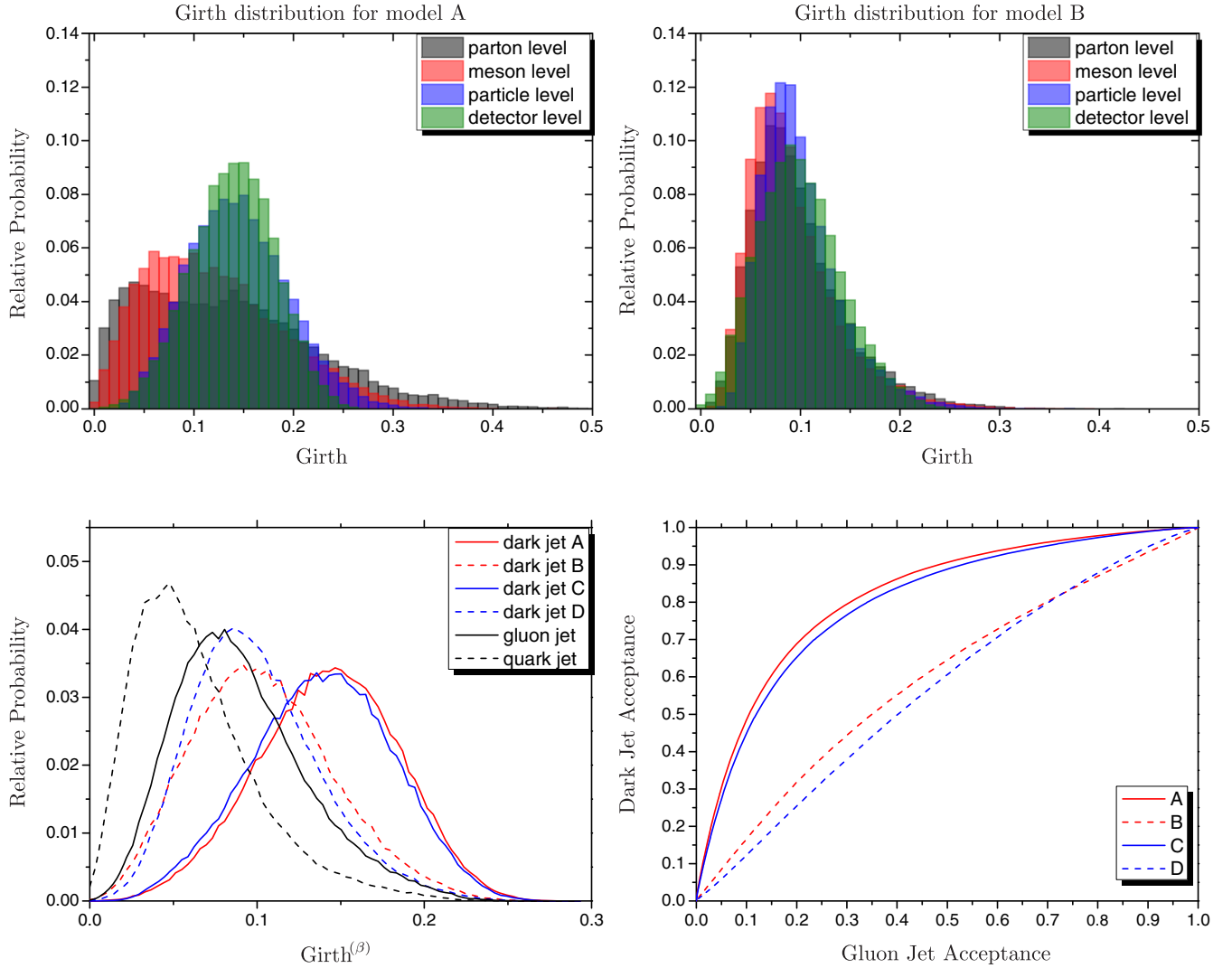


FIG. 5. Top left: girth distribution of a dark jet with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$ at parton level, meson level, final-state particle level, and detector level for dark QCD model A (corresponding to a high dark sector confinement scale). Top right: Same as the top left, but for model B (corresponding to a low dark sector confinement scale). Bottom left: girth distribution at detector level of different kinds of jets with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$. Bottom right: Corresponding ROC curves for discrimination between dark QCD jets and the SM QCD gluon-initiated jet.

comparison between model A and C (also model B and D), we find that girth is not sensitive to the different decay channels of a dark meson. And the discriminant ability of girth is a little weaker than the discriminant ability of $C_1^{(\beta)}$.

D. Charged track multiplicity

Multiplicity-type variables counting the number of hadrons or tracks inside a jet turn out to be useful in discriminating different kinds of one-prong jets. Among them, charged track multiplicity, due to a high resolution and a trigger efficiency of a track reconstruction at the LHC, is the best discriminant variable among various multiplicity-type variables used in quark and gluon jet discrimination [61,83,84]. Unlike jet mass or $C_1^{(\beta)}$, which

are IRC safe, charged track multiplicity does increase its value through soft and collinear radiations. Besides that, it is also closely related to the decay channel of a dark meson. So, we rely on Monte Carlo simulation results to show its property.

For gluon and quark jets, charged tracks inside them are mainly composed of π^\pm , which come from fragmentation directly. But for a dark jet, dark mesons coming from dark fragmentation are neutral, and thus tracks inside a dark jet can only come from dark meson decay. Because of a very different production mechanism, we expect charged track multiplicity to be a useful discriminant variable. The amount of tracks produced through dark meson decay is very model dependent. For our dark sector setting A, B, C, and D, the average amounts of tracks produced by a dark

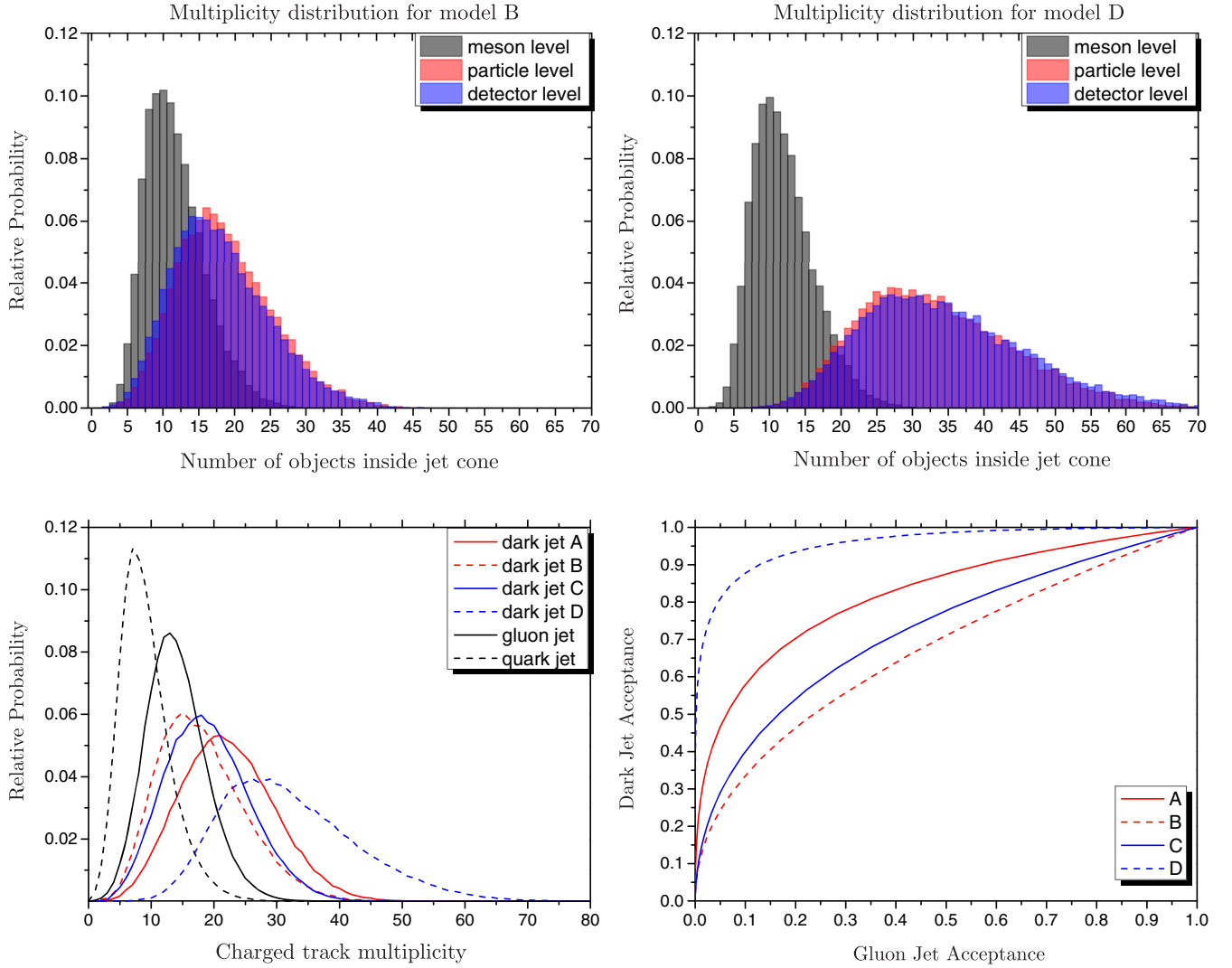


FIG. 6. Top left: Dark meson multiplicity, charged particle multiplicity, and track multiplicity distribution of the dark jet with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$ and setting B. Top right: Same as the top left, but with setting D. Bottom left: Charged track multiplicity distribution of different kinds of jets with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$. Bottom right: Corresponding ROC curves for discrimination between dark QCD jets and the SM QCD gluon-initiated jet.

meson are 4.9, 1.6, 4.4, and 2.8. Dark mesons A and C produce more tracks than dark mesons B and D. This is because the decay of charm mesons and the 4.0 GeV dark photon generally produce multiple tracks.

Figure 6 is our simulation results. To show how the track multiplicity is affected by the dark meson's decay channel, we count the amount of dark mesons, charged particles, and tracks with $p_T > 0.5 \text{ GeV}$ inside a dark jet, which correspond to meson level, particle level, and detector level, respectively, in the first row. With an identical dark sector setting, dark meson multiplicity distributions for model B and model D are almost the same. But different decay channels of the dark meson make their track multiplicities quite different. Thus, compared to the dark jet in model B, the dark jet in model D is much easier to discriminate from

the QCD jet. In general, track multiplicity is a better discriminant variable compared with IRC safe variables.

E. Ratio of energy deposits on different calorimeters

To further reflect final states from dark meson's decay, we suggest utilizing a variable which has a dependency on the detector's response to different final-state particles. At the LHC, most of the SM particles, except muons and neutrinos, will be stopped by calorimeters and deposit their energy on calorimeters. There are two kinds of calorimeters used in the LHC, an electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). Different particles deposit their energy in different calorimeters, as summarized below:

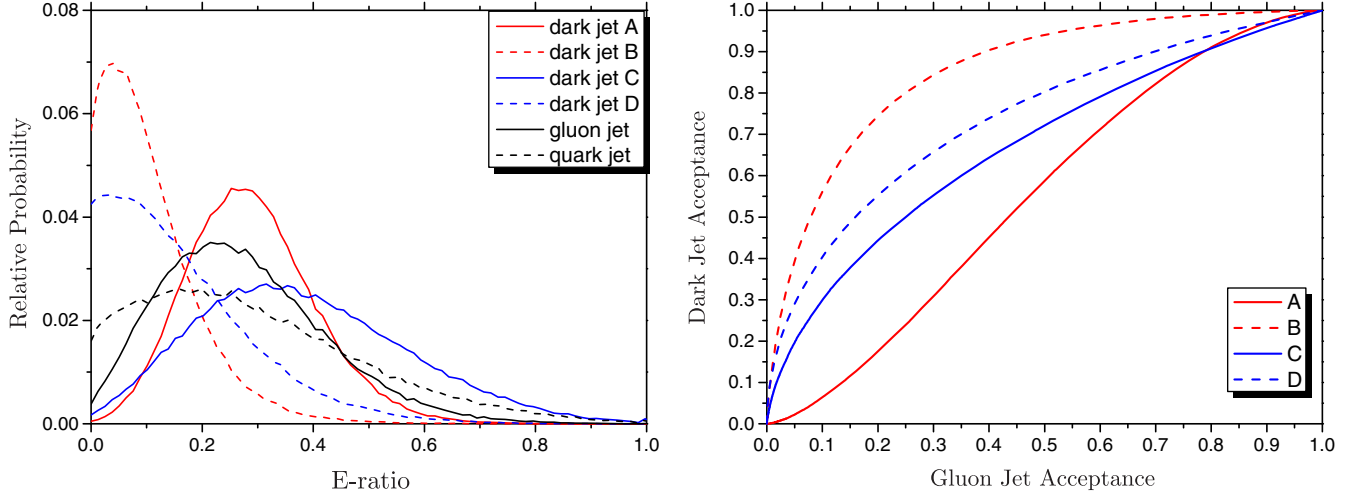


FIG. 7. Left: E-ratio distribution of different kinds of jets with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$. Right: Corresponding ROC curves for discrimination between dark QCD jets and the SM QCD gluon-initiated jet.

- (1) e^\pm, γ : deposit energy in the ECAL.
 - (2) Hadrons: deposit energy in the HCAL. Here, hadrons mainly refer to long-lived or stable hadrons like p, n, π^\pm , and K^\pm . Because hadrons like π^0 decay to $\gamma\gamma$ before they reach the HCAL, they thus deposit energy in the ECAL.
 - (3) μ^\pm, ν : do not deposit energy in the ECAL or HCAL.
- Different kinds of jets might have different particle compositions, and thus for a certain jet energy, the energy deposited in ECAL might be different. So, here we define a variable called the E-ratio:

$$\text{E-ratio} = \frac{\text{Energy deposit on ECAL}}{\text{Jet's } p_T}. \quad (19)$$

The E-ratio is used to reflect how much energy is carried by e^\pm, γ , and π^0 in a jet and thus provides more information in addition to track multiplicity.

Particle composition of the gluon jet and quark jet are controlled by a fragmentation process. Quarks produced in QCD vacuum are mainly u and d . The s quark pair is produced with a smaller probability. Thus, background QCD jets are mainly composed of π^0, π^\pm , and a small quantity of kaons. The precise percentage of each kind of particle is given by the fragmentation function, which needs to be obtained from data fitting. Because of different charge and color factors the, E-ratios of a gluon jet and quark jet are different.⁶ Our simulation result shows that the E-ratio of a gluon jet is larger than that of a quark jet.

For dark jets, the dark meson's decay channel controls particle composition. The dark meson in model A decays to a $c\bar{c}$ pair, and thus most of the energy is carried by D^\pm and D^0

⁶Actually, because of parton distribution functions, some quark jets are initiated by an s quark.

after fragmentation. D^\pm and D^0 decay before reaching the calorimeter. Decay modes of D^\pm and D^0 are very complicated, and they mainly decay to e, μ , and kaons through weak interaction. The dark jet A turns out to deposit energy on the ECAL with a ratio larger than a gluon jet. The dark meson in model B decays to $s\bar{s}$, and thus the main energy carriers are K^\pm, K_L^0 , and K_S^0 . K^\pm and K_L^0 are long lived and deposit energy in the HCAL. Short-lived K_S^0 decays to π^0 (π^\pm) with a probability of 30% (70%). In summary, most of jet energy in model B deposits in the HCAL.

Final decay products of model C are very involved. In model C, the dark meson decays to a dark photon γ' pair with a mass 4.0 GeV. Main decay modes of 4.0 GeV γ' are $d\bar{d}, u\bar{u}, s\bar{s}, c\bar{c}, e\bar{e}, \mu\bar{\mu}$, and $\tau\bar{\tau}$. Thus, after all prompt decay, decay products of the dark meson are a mixture of π^0, π^\pm , kaons, e, μ , and ν . In all of our models, dark jet C deposits the most energy in the ECAL. Finally, the dark meson D decays to a γ' pair with a mass 0.7 GeV; 0.7 GeV γ' decays to $e\bar{e}$ by a probability of 15%, and other decay products are μ and π^\pm . So, dark jet D does not deposit much energy in the ECAL, and its E-ratio is small.

The distribution of the E-ratio is shown in Fig. 7. As we expected, the E-ratio distributions of models B and D are small compared with other jets, and the corresponding ROC curve shows a good discriminant performance, while for model A and model C, this variable is not so useful.

F. Subjet

Properties of a one-prong jet can also be revealed by measuring observables associated with smaller subsets inside it because different kinds of jets have different energy profiles on a transverse plane. For example, most of the energy of the quark jet concentrates on a small central region, while the energy of a gluon jet will spread to a larger area [61]. Here, we define a subjet by reclustering

constituents of an original jet with the anti- k_T algorithm and a jet radius $R = 0.1$. We require the p_T of these subjects to be larger than 5% of the original jet's p_T . Here, we define $f_{p_T}^{(i)}$ as p_T of the i th hardest subject divided by p_T of an original jet:

$$f_{p_T}^{(i)} = \frac{p_T \text{ of } i\text{th hardest subject}}{\text{original jet's } p_T}. \quad (20)$$

Three variables are used here: 1) the number of subjects, 2) the p_T fraction carried by the hardest subject $f_{p_T}^{(1)}$, and 3) the p_T fraction carried by the second-hardest subject $f_{p_T}^{(2)}$.

Simulation results are shown in Fig. 8. Those distributions show clear physical meaning. A QCD quark jet, with a small coupling and color factor, can only

trigger a large angle shower with a quite low probability. Hence, there is a huge possibility for the quark jet to concentrate most of its energy in a tiny cone with a radius smaller than 0.1. Because of a larger color factor, the QCD gluon jet is a ‘‘broad’’ compared to a ‘‘narrow’’ quark jet, which means the energy of a gluon jet is distributed on a larger area and it is more likely to have more subjects inside the gluon jet. For the dark jets, through a larger coupling, they become even broader, and there are more subjects inside it. The p_T fraction of subjects is a natural expectation of such an argument. Among these three variables, the p_T fraction of the hardest subject $f_{p_T}^{(1)}$ shows the best discriminant ability. Similar to $C_1^{(\beta)}$, girth, and the jet mass, this variable is useful when the dark confinement scale is high.

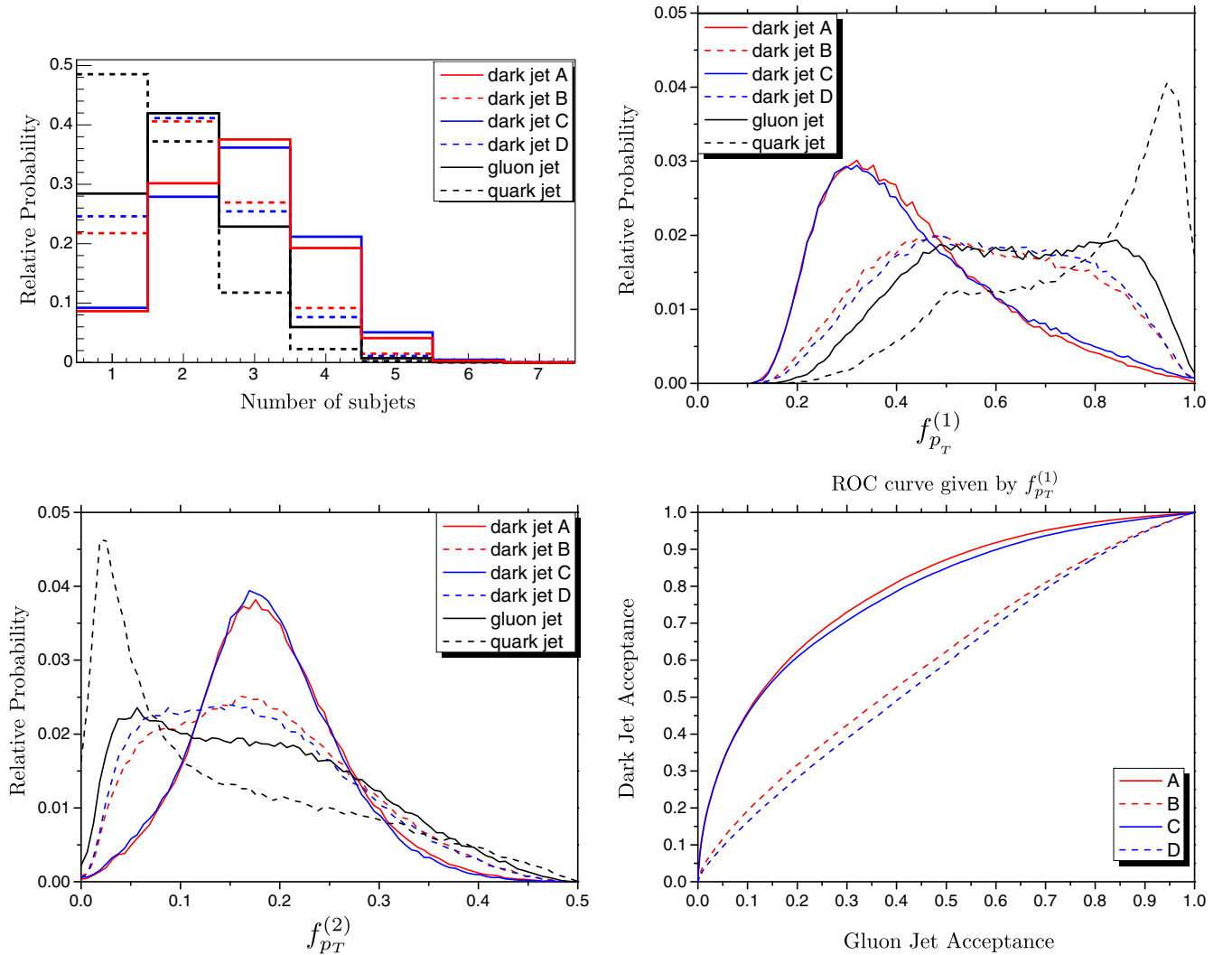


FIG. 8. Top left: number of subjects distribution of different kinds of jets with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$. Top right: hardest subject p_T fraction distribution of different kinds of jets with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$. Bottom left: second-hardest subject p_T fraction distribution of different kinds of jets with $p_T \in (180 \text{ GeV}, 220 \text{ GeV})$. Bottom right: ROC curves given by the hardest subject p_T fraction $f_{p_T}^{(1)}$.

TABLE II. Best discriminant variable for different models.

	Model A	Model B	Model C	Model D
Best discriminant variable	$C_1^{(\beta)}$	E-ratio	$C_1^{(\beta)}$	Track multiplicity

G. Combine multiple variables

One can estimate the discriminant performance of a variable by simply counting the area under the ROC curve, and a bigger area means a better discriminant ability. We show the best discriminant variable for each model in Table II. The discriminant performance can be maximized by combining multiple jet-substructure variables. To start our multiple variables analysis, we consider a set of variables consisting of the best discriminant variables: $\{C_1^{(\beta)}, \text{E-ratio, track multiplicity}\}$.

In principle, a classification can be made by a hyperplane in $\{C_1^{(\beta)}, \text{E-ratio, track multiplicity}\}$ space, but finding such a hyperplane is often difficult. For illustration, in Fig. 9, we show sample distribution of different dark jets and the gluon jet in $\{C_1^{(\beta)}, \text{E-ratio, track multiplicity}\}$ space. It can be seen that the distribution of dark jets is intertwined with the distribution of the gluon jet, and different dark jets populate different regions. Because of these complexities, in this work, we use the boosted decision tree (BDT) [85] in the TMVA-Toolkit [86] to do multiple variables analysis.

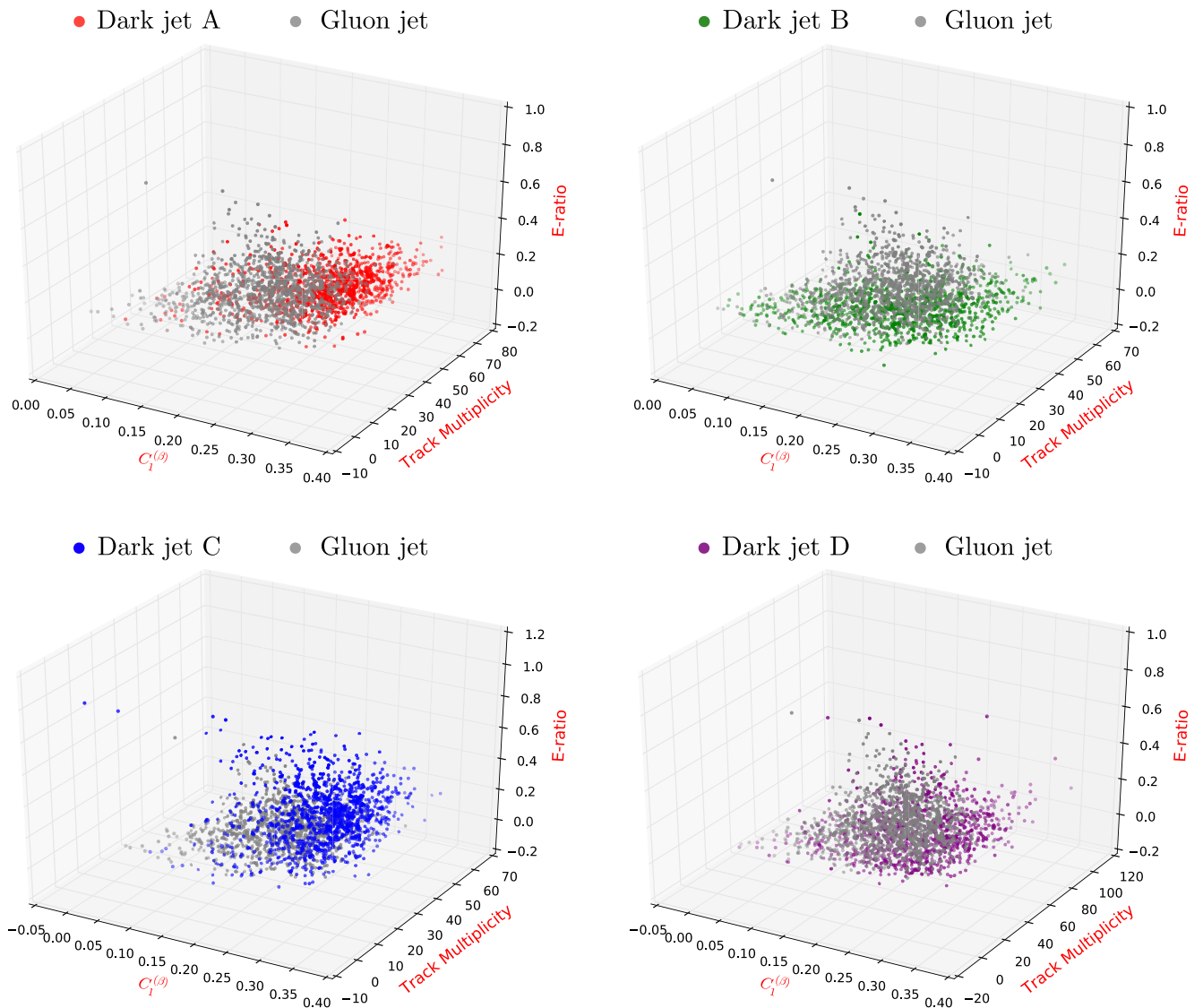


FIG. 9. Variables distribution in $\{C_1^{(\beta)}, \text{E-ratio, track multiplicity}\}$ space. Red, green, blue, purple, and gray points are samples of dark jet A, dark jet B, dark jet C, dark jet A, and the gluon jet, respectively.

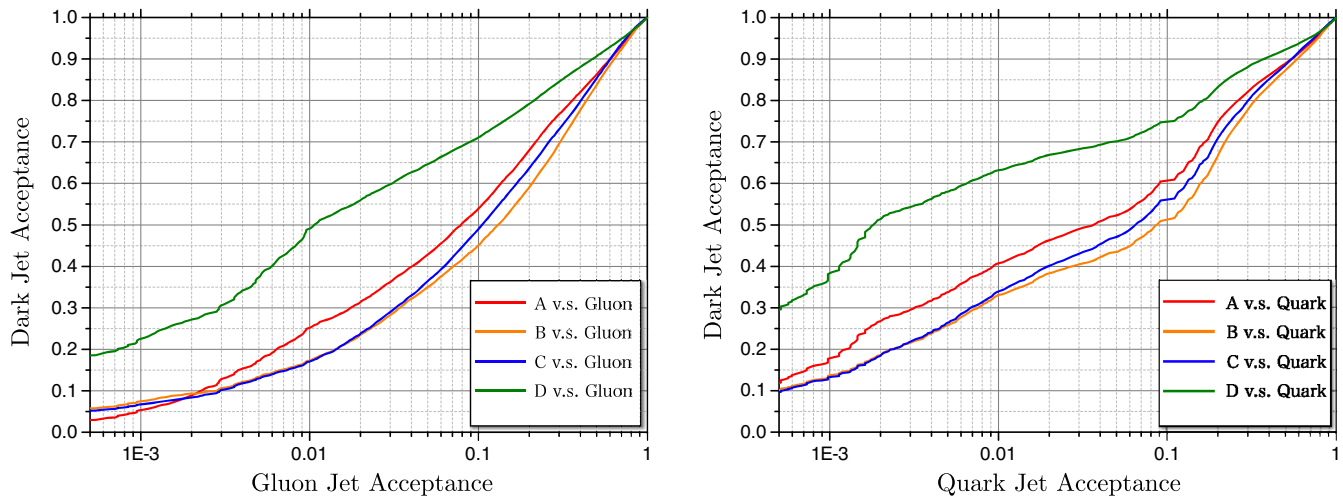


FIG. 10. ROC curves obtained by cutting on the BDT score. Left: dark jet vs gluon jet ROC curves for all four models, with $p_T \in (100 \text{ GeV}, 450 \text{ GeV})$. Right: dark jet vs quark jet ROC curves for all four models, with $p_T \in (100 \text{ GeV}, 450 \text{ GeV})$.

The BDT needs to be trained by a sufficient amount of samples that represent the signal and background characteristics. For both signal and background samples, we require their jets p_T to be evenly distributed from 100 to 450 GeV and pseudorapidity $\eta \in (-2.5, 2.5)$. Our signal sample contains 100,000 events, with 17%, 34%, 45%, and 4% of them coming from settings A, B, C, and D, respectively.⁷ Our background sample contains 30,000 events, and all of them are gluon jets. The background sample do not include quark jets because the distributions in previous subsections indicate that discrimination between the gluon jet and dark jet should be harder than discrimination between the quark jet and dark jet. We use 500 decision trees, choose a minimum in the leaf node as 2.5%, and set a maximum depth of 3. To avoid overtraining, half of the events are chosen as test events, and the Kolmogorov-Smirnov test is required to be larger than 0.01.

After training, the BDT can map an event with a certain $\{C_1^{(\beta)}, \text{E-ratio}, \text{track multiplicity}\}$ value, to a BDT score. A signal-like event tends to get a high BDT score, and a backgroundlike event tends to get a low BDT score. Thus, a discrimination can be performed by simply cutting on the BDT score. In Fig. 10, we show the ROC curve of the dark jet and gluon jet discrimination obtained by cutting on the BDT score. As one can expect, dark jet D is the easiest to distinguish because dark jet D generally produces many more tracks than a gluon jet. And the difficulties of distinguishing dark jets A, B, and C from background

⁷This sample ratio choice comes from the feasibility requirement. We found dark jets A and D are easier to distinguish than dark jets B and C. Thus, in order to obtain a BDT that is useful to all kinds of dark jets, we need to increase the proportion of events from models B and C in the signal sample. After several attempts, we found that the sample ratio used here gives us a BDT with good feasibility.

are close to each other. For comparison, we also present the ROC curve for the dark jet vs the quark jet in Fig. 10. It shows that the BDT trained by the dark jet and gluon jet can also be used to do discrimination between the dark jet and quark jet. And as we expect, distinguishing the dark jet from the quark jet is easier than distinguishing the dark jet from the gluon jet. The trained BDT we used here is also workable to other kinds of dark jets. More discussion on feasibility can be found in Appendix B.

Generally, if we use more variables in the BDT we might get a better discriminant performance. But if a variable is strongly correlated with other variables, then this variable will be redundant and cannot provide more information about a jet. In our training, we also consider an extended set of variables which contains eight variables: $\{m_J/p_T, C_1^{(\beta)}, \text{girth}, \text{track multiplicity}, \text{E-ratio}, \text{number of subjets}, f_{p_T}^{(1)}, f_{p_T}^{(2)}\}$. But the improvement we can get by including eight variables is negligible. It indicates that the main feature of a dark jet is described by $C_1^{(\beta)}$, E-ratio, and track multiplicity.

The BDT we trained in this section is feasible for different kinds of dark jets, but it might have lost some of the characteristics of the samples. Considering more than one BDT may help to comprehensively show the nature of the sample. For a discussion of two BDTs, see Appendix C.

IV. EXAMPLE AT LHC

In this section, we present bifundamental scalar X search at the LHC as an example to utilize our dark jet tagging method. As X is charged under both the SM $SU(3)$ and dark $SU(3)$, a pair of X particles can be produced at the LHC through the QCD process. Once a mediator X is produced, it decays into a SM quark and a dark quark, which evolves to a QCD jet and a dark jet, respectively. If the decay length of a dark meson is around $\mathcal{O}(10)\text{--}\mathcal{O}(100)$ mm, a dark jet will leave displaced

vertices in the detector. By counting the number of displaced vertices, one can obtain a robust limit on the mass of mediator particle X [41,87]. But if the decay length of a dark meson is shorter than 1 mm, analyses with displaced vertices will lose sensitivity. In this section, we will show that tagging the dark jet with jet substructures can be used to enhance a search sensitivity when dark mesons decay promptly.

We consider the dark sector setting A in Table I as an benchmark for the LHC study. Our analysis is based on the search for pair-produced resonances in four-jet final states on ATLAS [88]. Here, we briefly describe the cut flow used in the ATLAS report [88]:

- (i) Events are required to have at least four jets with $p_T > 120$ GeV and $|\eta| < 2.4$.
- (ii) These four jets are paired by minimizing $\Delta R_{\min} = \sum_{i=1,2} |\Delta R_i - 1|$, with ΔR_i the angular distance between two jets in a pair.
- (iii) Define \bar{m} as the average of the invariant masses of this two-jet pair as $\bar{m} = \frac{1}{2}(m_1 + m_2)$. Here, m_1 and m_2 are the invariant masses of two resonances. We veto events of large angular separation

$$\Delta R_{\min} > -0.002(\bar{m} - 225) + 0.72$$

for the case of $\bar{m} < 225$ GeV. If $\bar{m} \geq 225$ GeV, we discard events with

$$\Delta R_{\min} > 0.0013(\bar{m} - 225) + 0.72. \quad (21)$$

- (iv) Boost the system of these two resonances back to their center-of-mass frame, and define $\cos \theta^*$ as the cosine of the angle between one of the resonances and the beam line in the center-of-mass frame. The mass asymmetry \mathcal{A} is defined as

$$\mathcal{A} = \frac{|m_1 - m_2|}{m_1 + m_2}, \quad (22)$$

Events are selected by requiring $\mathcal{A} < 0.05$ and $|\cos \theta^*| < 0.3$. This cut flow defines the inclusive signal region selection.

This analysis utilizes kinetic information of final-state jets, which are p_T , η , and ϕ . However, as we have presented

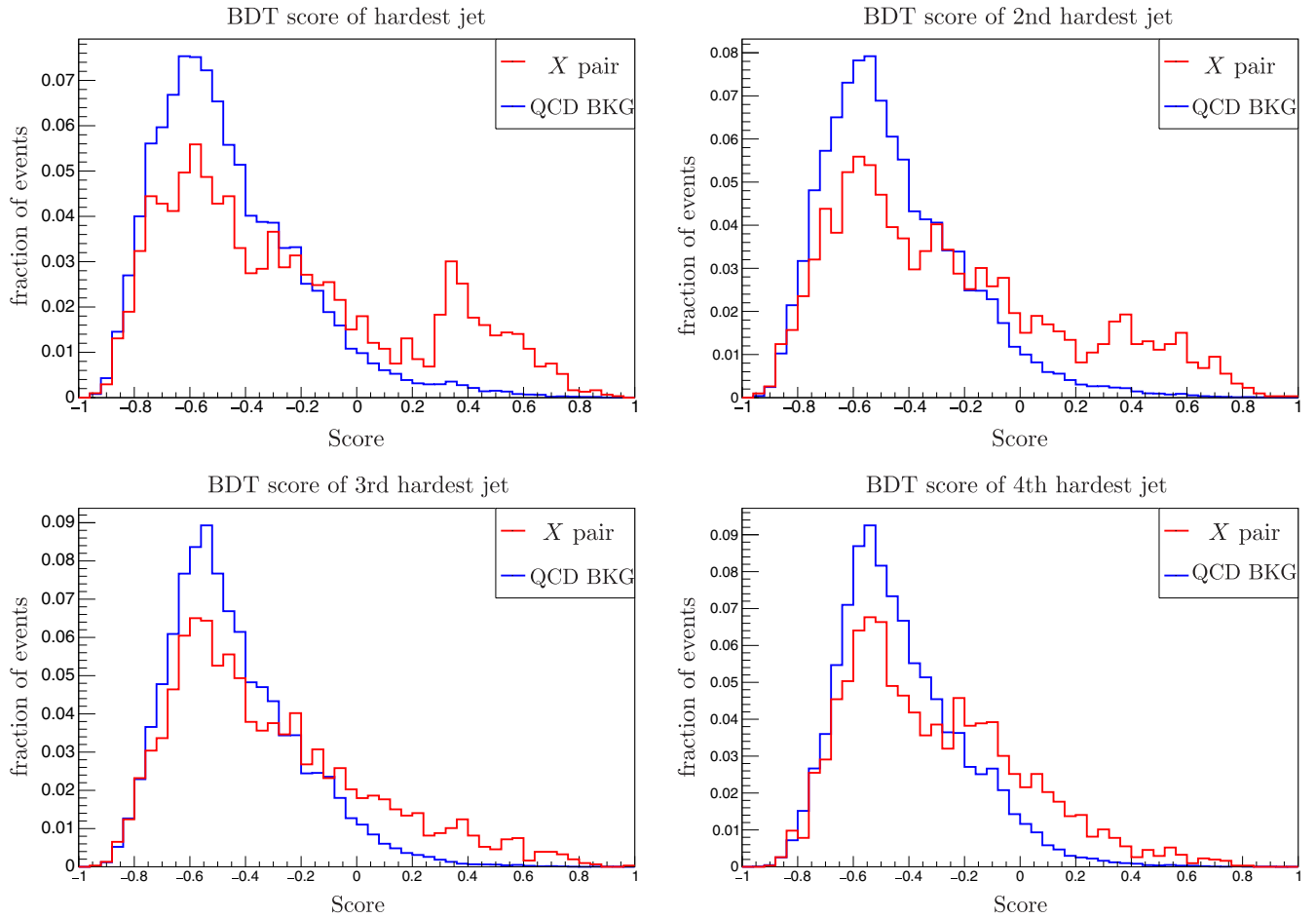


FIG. 11. BDT score distributions of four leading jets of signal and background (BKG). Events used here are required to have at least four jets with $p_T > 120$ GeV and $|\eta| < 2.4$. The BDT score is normalized to region $[-1, +1]$. For the signal process, here we consider $M_X = 400$ GeV for illustration.

TABLE III. The number of signal events and background events after applying the cut. Luminosity is 36.7 fb^{-1} , and the central energy is 13 TeV. The background events number after inclusive selection has been normalized to the data observed in the ATLAS report.

	BKG	$M_X = 300 \text{ GeV}$	$M_X = 500 \text{ GeV}$	$M_X = 700 \text{ GeV}$	$M_X = 900 \text{ GeV}$
Four jets with $ \eta < 2.4$ and $p_T > 120 \text{ GeV}$	2.46×10^7	70701	21723	5127	1267
ΔR_{\min} selection	3.96×10^6	20998	5946	1631	499.6
$\mathcal{A} < 0.05$ and $ \cos\theta^* < 0.3$	154750	2,246	529.7	120.7	30.6
Dark jets ≥ 1	2240	403.0	130.2	41.6	12.3
Dark jets ≥ 2	19.9	30.1	15.3	5.0	1.8
Significance	...	4.03	2.05	0.68	0.25

in Sec. III, one can get more information by looking inside a jet. If the resonance is the mediator particle X , there will be dark jets in the final state. So, by tagging dark jets, search sensitivity can be enhanced. Here, we use the trained BDT we obtained in last section to compute BDT scores of four leading jets in the final state. Thus, the dark jet can be tagged by cutting on the BDT score. A similar method has been performed in the supersymmetry study [49].

Backgrounds from SM QCD processes and signal events from X pair production are generated by PYTHIA8. For background simulation, we generate more than one billion events, and the events number after inclusive cut is normalized to the data observed in the ATLAS report [88]. The production cross section of the X pair is the production cross section of the stop pair multiplied by 3 [89] because X is also charged under a dark $SU(3)$ gauge group. In Fig. 11, we show the BDT score distributions of four leading jets for both background and signal. It can be seen that the BDT score of our signal is larger than the BDT score of the QCD background. Thus, we can suppress the QCD background by cutting on the BDT scores. Here, we

define a jet with a BDT score larger than 0.483^8 to be tagged as dark jet, and we require at least two dark jets in the final state. In Table III, we list the cut flow for both the background and signal. After inclusive selection, the background events number is 3 orders of magnitude larger than the signal events number. By requiring at least two dark jets in the final state, we suppress the background by 4 orders of magnitude, and thus the background and signal become comparable. Significances are estimated by $\frac{S}{\sqrt{B+\epsilon^2 B^2}}$. Here, S and B are the events numbers of the signal and background, respectively, and we assume systematic uncertainty ϵ to be 30% for a conservative estimation.

We give a 95% confidence-level exclusion limit on the cross section of X pair production in Fig. 12. Theoretical uncertainty (TU) from dark jet tagging is estimated by varying renormalization-scale μ in the dark parton shower process. See Appendix A for a detailed discussion about theoretical uncertainty. To compare with the search method without dark jet tagging, in Fig. 12, we also show the up limit obtained by doing a recast of the ATLAS report [88]. In Ref. [88], after the inclusive cut, several mass windows are designed to further increase the search sensitivity. For a certain resonance mass, the average mass m_{avg} is required to be located in a narrow region around it. However, due to the strong shower in dark sector, the average mass obtained by four final-state leading jets distribute in a broad mass region. Thus, the mass window cut discards too many signal events and results in a low sensitivity. Figure 12 shows that the exclusion limit from the ATLAS report recast is weaker than the limit from our dark jet tagging method.

V. CONCLUSION

The dark sector charged under a confined $SU(N_d)$ provides composite states and attractive phenomenologies. At colliders, such a model can produce jetlike signals (called a dark jets), some of which may not be tagged by distinct or exotic signatures including missing energy or the displaced vertex. In this work, inspired by the success of quark/gluon jet discrimination, we try to distinguish the dark jet from the background SM QCD-jet by using

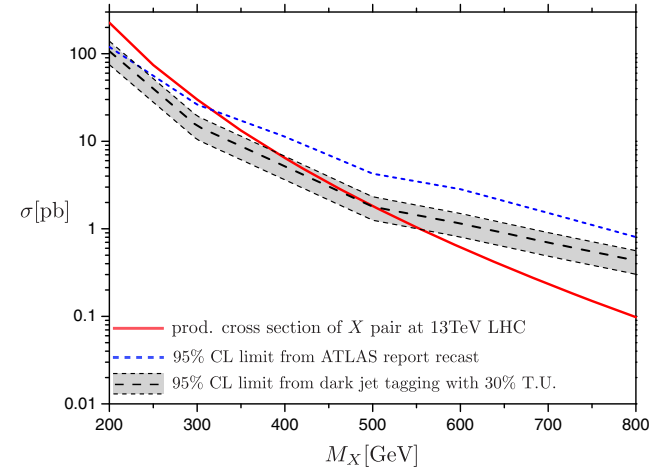


FIG. 12. The 95% C.L. upper limit on the production cross section of the X pair, with X decays to a SM quark and a dark quark. The red line is the production cross section of the X pair at the 13 TeV LHC. The blue dashed line is the up limit obtained by using the cut flow in the ATLAS report [88]. The black dashed line is the up limit obtained by using our dark jet tagging method. The gray band reflects theoretical uncertainty from dark jet tagging.

⁸This number is chosen for maximizing the exclusion limit.

jet-substructure variables. A series of jet-substructure variables, like the jet mass, $C_1^{(\beta)}$, or track multiplicity, is studied in this work. We use simulated events to exhibit the discriminant ability of all these jet-substructure variables. A BDT is trained as a classifier to discriminate the dark jet and QCD jet. For all of our dark sector settings, we can use this trained BDT to exclude 99% background gluon jets with more than 15% signal dark jets reserved or exclude 99% background quark jets with more than 30% signal dark jets reserved. A concrete example is used to show that our dark jet tagging method can enhance the search sensitivity of models which produce dark jets in the final state. Theoretical uncertainty from Monte Carlo simulation is also discussed, which shows that our method is robust against theoretical uncertainty.

ACKNOWLEDGMENTS

The authors appreciate Doctor Sung Hak Lim for discussions at the earlier stage of this project. M. Z. appreciates help from PYTHIA authors, Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. M. Z. thanks Yasuhito Sakaki, Jinmian Li, and Zhuoni Qian for useful discussions with them. This work is supported by IBS under Project No. IBS-R018-D1. M. P. is supported by the Basic Science Research Program through the National Research Foundation of Korea Research Grants No. NRF-2017R1C1B5075677 and No. NRF-2018R1C1B6006572.

APPENDIX A: UNCERTAINTY DISCUSSION

The discriminant ability shown in Sec. IV might be quite sensitive to TU from a Monte Carlo event generator. In analyses of quark-gluon jet tagging, one can tune parameters in the Monte Carlo event generator from real data to reduce systematics and enhance predictability. Generally, one can simulate a quark jet very well with public event generators. And for a gluon jet, it is also known that the real distribution lies in between PYTHIA and HERWIG [90] expectation. More information can be found in a recent review [76].

But for a dark jet, we cannot estimate systematics as we do not have the real data of the dark jet. Thus, parameters in simulating dark QCD hadronization and showering leaves unfixed systematics in our analyses. On top of this difficulty, as we do not have various Monte Carlo generators for dark jet simulation except PYTHIA8, we do not have a choice to compare different event generators to get an estimation about uncertainty depending on different showering and hadronization schemes. Alternatively, we do some simple estimation in this work. Changing the

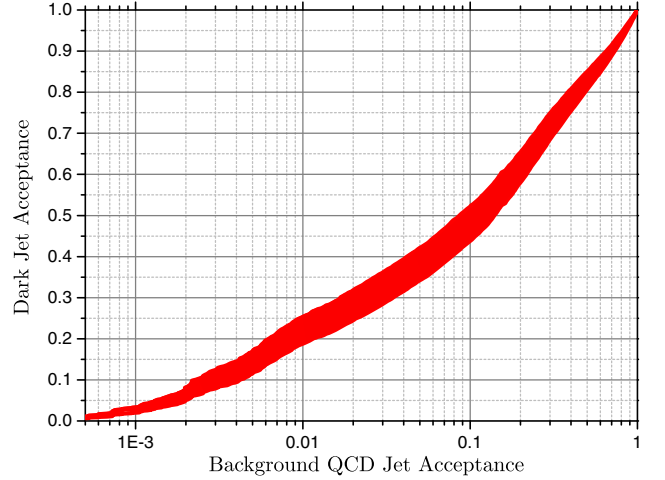


FIG. 13. ROC discriminant curve of dark jet vs SM QCD jet, with renormalization scale μ varying from $0.5p_{\perp}$ to $2.0p_{\perp}$. Here, we use the dark sector, setting A in Table I.

renormalization scale in the parton shower process has been proven to be a good method to estimate theoretical uncertainty in PYTHIA [91]. So, following this method, we also rescale the renormalization scale μ in the dark sector shower process.

For background and signal events simulation, we use the same setting as we used in Sec. IV. But we simulate signal events several times with different renormalization scale μ , varying from $0.5p_{\perp}$ to $2.0p_{\perp}$. Here, p_{\perp} is the order parameter used in the PYTHIA8 shower process, and $\mu = p_{\perp}$ is the default setting in PYTHIA8. So, after utilizing the trained BDT to do discrimination, we obtain several ROC curves. Results are shown in Fig. 13. In the region where the acceptance of the background QCD jet is around 1%, the acceptance of the signal dark jet changes from 19% to 26%. So, if we want to tag one dark jet in the final state, the corresponding theoretical uncertainty is about 15%. In Sec. IV, we want to tag two dark jets in the final state; thus, the theoretical uncertainty is about 30%.

APPENDIX B: FEASIBILITY DISCUSSION

To further test the feasibility of the BDT trained in Sec. III, we consider another two benchmark models, E and F. The detailed setting can be found in Table IV. We choose these two settings because they have confinement scales different than setting A and B. In the left plot of Fig. 14, we present the coupling running of E and F. It can be seen that the coupling running curves of setting E

TABLE IV. Benchmark models used to show the feasibility of the BDT we trained in Sec. III.

	N_d	n_f	Λ_d (GeV)	$\tilde{m}_{q'}$ (GeV)	m_{π_d} (GeV)	m_{ρ_d} (GeV)	π_d Decay mode	ρ_d Decay mode
E	3	3	8	10	5	25	$\pi_d \rightarrow c\bar{c}$	$\rho_d \rightarrow \pi_d\pi_d$
F	3	3	5	6	5	14.33	$\pi_d \rightarrow \gamma'\gamma'$ with $m_{\gamma'} = 0.8$ GeV	$\rho_d \rightarrow \pi_d\pi_d$

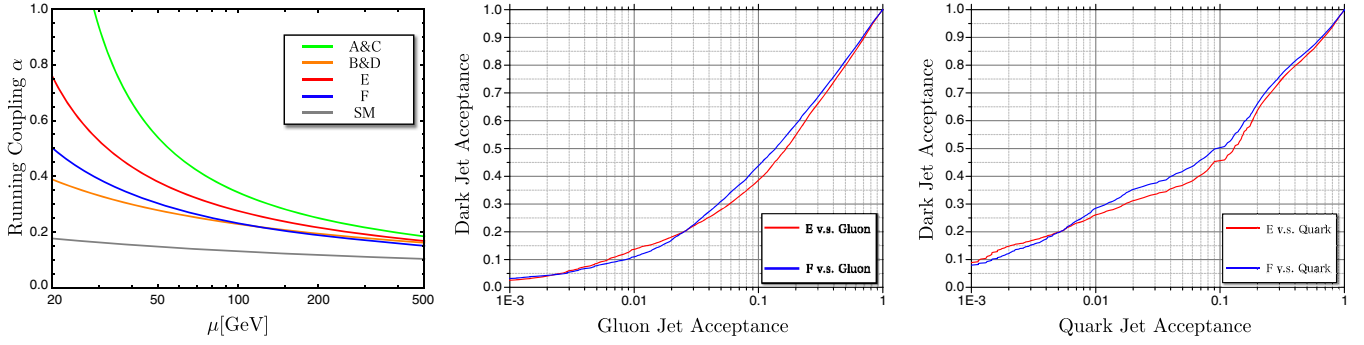


FIG. 14. Left: QCD coupling running in the dark sector and SM QCD. Middle: dark jet vs gluon jet ROC curves for benchmark points E and F, with $p_T \in (100 \text{ GeV}, 450 \text{ GeV})$. Right: dark jet vs quark jet ROC curves for benchmark points E and F, with $p_T \in (100 \text{ GeV}, 450 \text{ GeV})$.

and F lie between the coupling running curves of setting A and B. For model E and F, we generate 50,000 events with jet p_T ranging from 100 to 450 GeV. As we did in Sec. III, we use the trained BDT to map the value of $\{C_1^{(\beta)}, E\text{-ratio}, \text{track multiplicity}\}$ to a BDT score. By cutting on the BDT score, we obtain ROC curves, which are shown in Fig. 14. The discriminant performance is quite good. For models E and F, we can exclude 99% background gluon jets with more than 11% signal dark jets reserved or exclude 99% background quark jets with more than 26% signal dark jets reserved.

Here, we emphasize that we train the BDT with samples from benchmark points A, B, C, and D to cover generic

features of the dark jet. More specifically, our BDT captures a feature of a large coupling by studying sample A, and it also learns different decay modes from C. To prepare small coupling scenarios, we provide a chance of learning with benchmark points B and D for each case. As we test its performance on different parameter points (E and F), our method can be applied to a wide range of parameters in dark QCD physics.

APPENDIX C: DISCRIMINATION BY TWO BDTs

In Sec. III G, we use a mixture of the events from setting A, B, C, and D as signal sample to train a BDT, and this

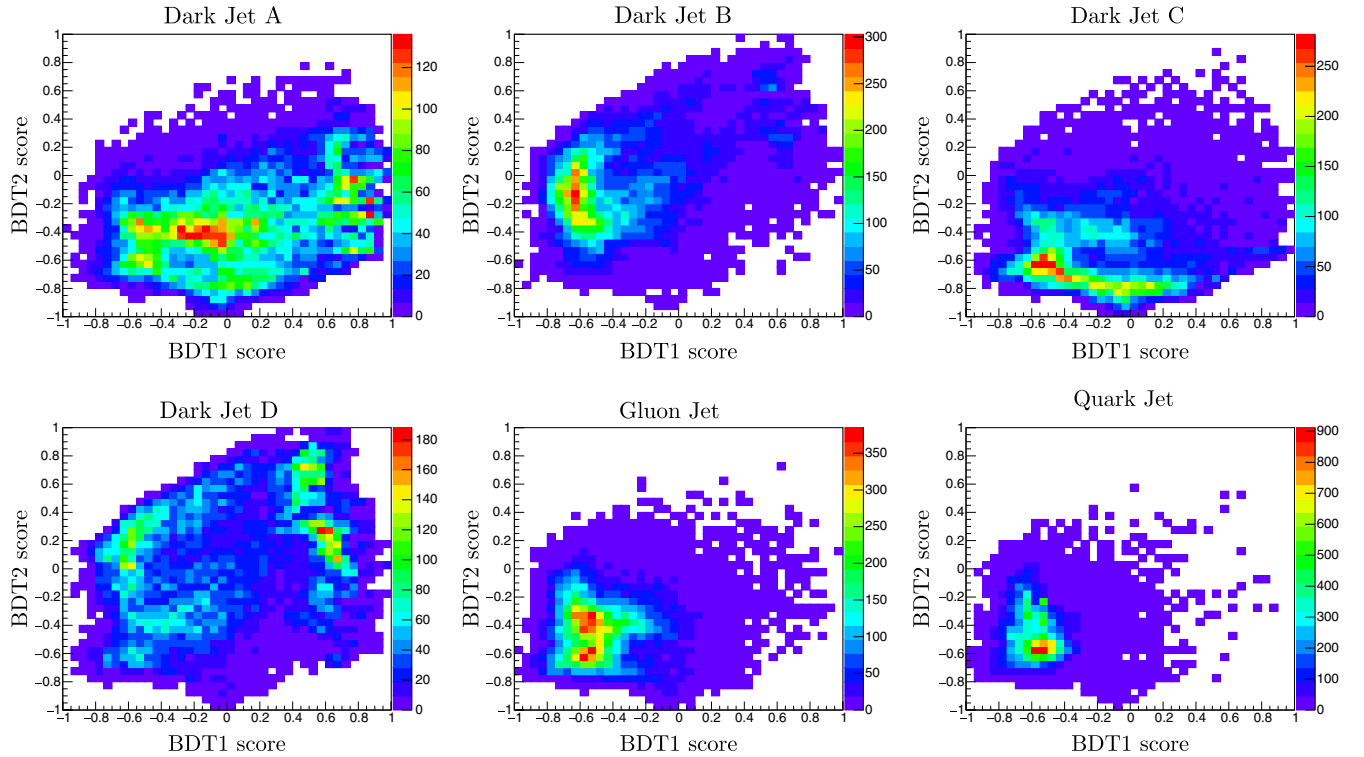


FIG. 15. Distributions on the “BDT1 score”-“BDT2 score” plane of different kinds of dark jets and background QCD jets. Distributions in each plot are obtained from 30,000 events.

BDT shows good feasibility in distinguishing different kinds of dark jets from QCD background. But such a BDT might lose some unique characteristics of a certain kind of dark jets. This problem can be alleviated if we use two BDTs. In this Appendix, we consider a BDT1, which is trained by choosing dark jet A as signal sample, and a BDT2, which is trained by choosing dark jet B as a signal sample. Thus, for each jet, we will get two BDT scores, and more information about a jet can be reflected. We cannot easily obtain a ROC curve by cutting on the BDT scores in this case. So, instead of presenting a ROC curve, we simply

show the distribution of different kinds of jets on the BDT scores plane. Figure 15 is our result. It shows that both gluon and quark jets concentrate on the lower left quarter, but dark jets can spread to a larger region. Dark jet A and dark jet D populate more than half of the BDT score plane. Dark jet B and dark jet C also show distributions which are different with QCD background. Discrimination in this case can be performed by some image identification technique. But a study on utilizing image identification is beyond the scope of current work, and we stop our discussion here.

-
- [1] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2015 results. XIII. Cosmological parameters, *Astron. Astrophys.* **594**, A13 (2016).
- [2] A. Tan *et al.* (PandaX-II Collaboration), Dark Matter Results from First 98.7 Days of Data from the PandaX-II Experiment, *Phys. Rev. Lett.* **117**, 121303 (2016).
- [3] D. S. Akerib *et al.* (LUX Collaboration), Results from a Search for Dark Matter in the Complete LUX Exposure, *Phys. Rev. Lett.* **118**, 021303 (2017).
- [4] V. Khachatryan *et al.* (CMS Collaboration), Search for dark matter particles in proton-proton collisions at $\sqrt{s} = 8$ TeV using the razor variables, *J. High Energy Phys.* **12** (2016) 088.
- [5] G. Aad *et al.* (ATLAS Collaboration), Search for Dark Matter in Events with a Hadronically Decaying W or Z Boson and Missing Transverse Momentum in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector, *Phys. Rev. Lett.* **112**, 041802 (2014).
- [6] D. Abercrombie *et al.*, Dark matter benchmark models for early LHC run-2 searches: Report of the ATLAS/CMS dark matter forum, *Phys. Dark Universe* **26**, 100371 (2019).
- [7] M. Aaboud *et al.* (ATLAS Collaboration), Search for dark matter at $\sqrt{s} = 13$ TeV in final states containing an energetic photon and large missing transverse momentum with the ATLAS detector, *Eur. Phys. J. C* **77**, 393 (2017).
- [8] M. Aaboud *et al.* (ATLAS Collaboration), Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector, *Phys. Rev. Lett.* **119**, 181804 (2017).
- [9] S. Nussinov, Technoc cosmology: Could a technibaryon excess provide a ‘natural’ missing mass candidate?, *Phys. Lett.* **165B**, 55 (1985).
- [10] D. B. Kaplan, A Single Explanation for Both the Baryon and Dark Matter Densities, *Phys. Rev. Lett.* **68**, 741 (1992).
- [11] S. M. Barr, R. S. Chivukula, and E. Farhi, Electroweak Fermion number violation and the production of stable particles in the early Universe, *Phys. Lett. B* **241**, 387 (1990).
- [12] S. M. Barr, Baryogenesis, sphalerons and the cogeneration of dark matter, *Phys. Rev. D* **44**, 3062 (1991).
- [13] S. Dodelson, B. R. Greene, and L. M. Widrow, Baryogenesis, dark matter and the width of the Z, *Nucl. Phys.* **B372**, 467 (1992).
- [14] M. Fujii and T. Yanagida, A solution to the coincidence puzzle of Omega(B) and Omega (DM), *Phys. Lett. B* **542**, 80 (2002).
- [15] R. Kitano and I. Low, Dark matter from baryon asymmetry, *Phys. Rev. D* **71**, 023510 (2005).
- [16] G. R. Farrar and G. Zaharijas, Dark Matter and the Baryon Asymmetry, *Phys. Rev. Lett.* **96**, 041302 (2006).
- [17] R. Kitano, H. Murayama, and M. Ratz, Unified origin of baryons and dark matter, *Phys. Lett. B* **669**, 145 (2008).
- [18] S. B. Gudnason, C. Kouvaris, and F. Sannino, Towards working technicolor: Effective theories and dark matter, *Phys. Rev. D* **73**, 115003 (2006).
- [19] D. E. Kaplan, M. A. Luty, and K. M. Zurek, Asymmetric dark matter, *Phys. Rev. D* **79**, 115016 (2009).
- [20] J. Shelton and K. M. Zurek, Darkogenesis: A baryon asymmetry from the dark matter sector, *Phys. Rev. D* **82**, 123512 (2010).
- [21] H. Davoudiasl, D. E. Morrissey, K. Sigurdson, and S. Tulin, Hylogenesis: A Unified Origin for Baryonic Visible Matter and Antibaryonic Dark Matter, *Phys. Rev. Lett.* **105**, 211304 (2010).
- [22] M. R. Buckley and L. Randall, Xogenesis, *J. High Energy Phys.* **09** (2011) 009.
- [23] T. Cohen, D. J. Phalen, A. Pierce, and K. M. Zurek, Asymmetric dark matter from a GeV hidden sector, *Phys. Rev. D* **82**, 056001 (2010).
- [24] M. T. Frandsen, S. Sarkar, and K. Schmidt-Hoberg, Light asymmetric dark matter from new strong dynamics, *Phys. Rev. D* **84**, 051703(R) (2011).
- [25] K. Petraki and R. R. Volkas, Review of asymmetric dark matter, *Int. J. Mod. Phys. A* **28**, 1330028 (2013).
- [26] K. M. Zurek, Asymmetric dark matter: Theories, signatures, and constraints, *Phys. Rep.* **537**, 91 (2014).
- [27] M. Ibe, A. Kamada, S. Kobayashi, and W. Nakano, Composite asymmetric dark matter with a dark photon portal, *J. High Energy Phys.* **11** (2018) 203.
- [28] M. Ibe, A. Kamada, S. Kobayashi, T. Kuwahara, and W. Nakano, Ultraviolet completion of a composite asymmetric

- dark matter model with a dark photon portal, *J. High Energy Phys.* **03** (2019) 173.
- [29] Y. Bai, J. Bourbeau, and T. Lin, Dark matter searches with a mono- Z' jet, *J. High Energy Phys.* **06** (2015) 205.
- [30] M. Buschmann, J. Kopp, J. Liu, and P. A. N. Machado, Lepton jets from radiating dark matter, *J. High Energy Phys.* **07** (2015) 045.
- [31] M. Kim, H. S. Lee, M. Park, and M. Zhang, Examining the origin of dark matter mass at colliders, *Phys. Rev. D* **98**, 055027 (2018).
- [32] M. J. Strassler and K. M. Zurek, Echoes of a hidden valley at hadron colliders, *Phys. Lett. B* **651**, 374 (2007).
- [33] M. J. Strassler and K. M. Zurek, Discovering the Higgs through highly-displaced vertices, *Phys. Lett. B* **661**, 263 (2008).
- [34] M. J. Strassler, Possible effects of a hidden valley on supersymmetric phenomenology, [arXiv:hep-ph/0607160](https://arxiv.org/abs/hep-ph/0607160).
- [35] T. Han, Z. Si, K. M. Zurek, and M. J. Strassler, Phenomenology of hidden valleys at hadron colliders, *J. High Energy Phys.* **07** (2008) 008.
- [36] M. Verducci (ATLAS Collaboration), Hidden valley search at ATLAS, *J. Phys. Conf. Ser.* **335**, 012068 (2011).
- [37] Y. F. Chan, M. Low, D. E. Morrissey, and A. P. Spray, LHC signatures of a minimal supersymmetric hidden valley, *J. High Energy Phys.* **05** (2012) 155.
- [38] M. J. Strassler, On the phenomenology of hidden valleys with heavy flavor, [arXiv:0806.2385](https://arxiv.org/abs/0806.2385).
- [39] T. Cohen, M. Lisanti, and H. K. Lou, Semivisible Jets: Dark Matter Undercover at the LHC, *Phys. Rev. Lett.* **115**, 171804 (2015).
- [40] Y. Bai and P. Schwaller, Scale of dark QCD, *Phys. Rev. D* **89**, 063522 (2014).
- [41] P. Schwaller, D. Stolarski, and A. Weiler, Emerging jets, *J. High Energy Phys.* **05** (2015) 059.
- [42] H. Beauchesne, E. Bertuzzo, G. Grilli Di Cortona, and Z. Tabrizi, Collider phenomenology of hidden valley mediators of spin 0 or 1/2 with semivisible jets, *J. High Energy Phys.* **08** (2018) 030.
- [43] A. Pierce, B. Shakya, Y. Tsai, and Y. Zhao, Searching for confining hidden valleys at LHCb, ATLAS, and CMS, *Phys. Rev. D* **97**, 095033 (2018).
- [44] H. Beauchesne, E. Bertuzzo, and G. Grilli Di Cortona, Dark matter in hidden valley models with stable and unstable light dark mesons, *J. High Energy Phys.* **04** (2019) 118.
- [45] S. Renner and P. Schwaller, A flavoured dark sector, *J. High Energy Phys.* **08** (2018) 052.
- [46] A. J. Larkoski, G. P. Salam, and J. Thaler, Energy correlation functions for jet substructure, *J. High Energy Phys.* **06** (2013) 108.
- [47] D. Ferreira de Lima, P. Petrov, D. Soper, and M. Spannowsky, Quark-gluon tagging with shower deconstruction: Unearthing dark matter and Higgs couplings, *Phys. Rev. D* **95**, 034001 (2017).
- [48] I. Moulton, L. Necib, and J. Thaler, New angles on energy correlation functions, *J. High Energy Phys.* **12** (2016) 153.
- [49] B. Bhattacherjee, S. Mukhopadhyay, M. M. Nojiri, Y. Sakaki, and B. R. Webber, Quark-gluon discrimination in the search for gluino pair production at the LHC, *J. High Energy Phys.* **01** (2017) 044.
- [50] J. Gallicchio, J. Huth, M. Kagan, M. D. Schwartz, K. Black, and B. Tweedie, Multivariate discrimination and the Higgs +W/Z search, *J. High Energy Phys.* **04** (2011) 069.
- [51] A. Boveia *et al.*, Recommendations on presenting LHC searches for missing transverse energy signals using simplified s -channel models of dark matter, [arXiv:1603.04156](https://arxiv.org/abs/1603.04156).
- [52] J. Ellis, M. Fairbairn, and P. Tunney, Anomaly-free dark matter models are not so simple, *J. High Energy Phys.* **08** (2017) 053.
- [53] A. Ismail, W. Y. Keung, K. H. Tsao, and J. Unwin, Axial vector Z' and anomaly cancellation, *Nucl. Phys.* **B918**, 220 (2017).
- [54] S. Knapen, S. Pagan Griso, M. Papucci, and D. J. Robinson, Triggering soft bombs at the LHC, *J. High Energy Phys.* **08** (2017) 076.
- [55] J. Alexander *et al.*, Dark sectors 2016 workshop: Community report, [arXiv:1608.08632](https://arxiv.org/abs/1608.08632).
- [56] T. Cohen, M. Lisanti, H. K. Lou, and S. Mishra-Sharma, LHC searches for dark sector showers, *J. High Energy Phys.* **11** (2017) 196.
- [57] D. Griffiths, *Introduction to Elementary Particles* (Wiley-VCH, Weinheim, Germany, 2008), p. 454.
- [58] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Recombination algorithms and jet substructure: Pruning as a tool for heavy particle searches, *Phys. Rev. D* **81**, 094023 (2010).
- [59] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Techniques for improved heavy particle searches with jet substructure, *Phys. Rev. D* **80**, 051501(R) (2009).
- [60] D. Krohn, J. Thaler, and L. T. Wang, Jet trimming, *J. High Energy Phys.* **02** (2010) 084.
- [61] J. Gallicchio and M. D. Schwartz, Quark and Gluon Tagging at the LHC, *Phys. Rev. Lett.* **107**, 172001 (2011).
- [62] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to PYTHIA 8.2, *Comput. Phys. Commun.* **191**, 159 (2015).
- [63] G. Aad *et al.* (ATLAS Collaboration), Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment, *J. High Energy Phys.* **05** (2012) 128.
- [64] G. Aad *et al.* (ATLAS Collaboration), Performance of jet substructure techniques for large- R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector, *J. High Energy Phys.* **09** (2013) 076.
- [65] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3 Collaboration), DELPHES 3, A modular framework for fast simulation of a generic collider experiment, *J. High Energy Phys.* **02** (2014) 057.
- [66] M. Cacciari, G. P. Salam, and G. Soyez, FastJet user manual, *Eur. Phys. J. C* **72**, 1896 (2012).
- [67] M. Cacciari, G. P. Salam, and G. Soyez, The anti- $k(t)$ jet clustering algorithm, *J. High Energy Phys.* **04** (2008) 063.
- [68] G. Aad *et al.* (ATLAS Collaboration), Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, *Eur. Phys. J. C* **74**, 3023 (2014).
- [69] M. Aaboud *et al.* (ATLAS Collaboration), Jet energy scale measurements and their systematic uncertainties in

- proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, *Phys. Rev. D* **96**, 072002 (2017).
- [70] S. Catani, L. Trentadue, G. Turnock, and B.R. Webber, Resummation of large logarithms in e^+e^- event shape distributions, *Nucl. Phys.* **B407**, 3 (1993).
- [71] U. Ellwanger, The jet mass in leading log QCD, *Z. Phys. C* **8**, 261 (1981).
- [72] R. Kelley, M. D. Schwartz, and H. X. Zhu, Resummation of jet mass with and without a jet veto, [arXiv:1102.0561](https://arxiv.org/abs/1102.0561).
- [73] H. N. Li, Z. Li, and C.-P. Yuan, QCD Resummation for Jet Substructures, *Phys. Rev. Lett.* **107**, 152001 (2011).
- [74] H. N. Li, Z. Li, and C.-P. Yuan, QCD resummation for light-particle jets, *Phys. Rev. D* **87**, 074025 (2013).
- [75] Y. T. Chien, R. Kelley, M. D. Schwartz, and H. X. Zhu, Resummation of jet mass at hadron colliders, *Phys. Rev. D* **87**, 014010 (2013).
- [76] P. Gras, S. Höche, D. Kar, A. Larkoski, L. Lönnblad, S. Plätzer, A. Siódmok, P. Skands, G. Soyez, and J. Thaler, Systematics of quark/gluon tagging, *J. High Energy Phys.* **07** (2017) 091.
- [77] G. P. Korchemsky and G. F. Sterman, Power corrections to event shapes and factorization, *Nucl. Phys.* **B555**, 335 (1999).
- [78] P. E. L. Rakow and B. R. Webber, Transverse momentum moments of hadron distributions in QCD jets, *Nucl. Phys.* **B191**, 63 (1981).
- [79] C. F. Berger, T. Kucs, and G. F. Sterman, Event shape/energy flow correlations, *Phys. Rev. D* **68**, 014012 (2003).
- [80] L. G. Almeida, S. J. Lee, G. Perez, G. F. Sterman, I. Sung, and J. Virzi, Substructure of high- p_T Jets at the LHC, *Phys. Rev. D* **79**, 074017 (2009).
- [81] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig, and C. Lee, Jet shapes and jet algorithms in SCET, *J. High Energy Phys.* **11** (2010) 101.
- [82] S. Catani, G. Turnock, and B. R. Webber, Jet broadening measures in e^+e^- annihilation, *Phys. Lett. B* **295**, 269 (1992).
- [83] K. Akerstaff *et al.* (OPAL Collaboration), Multiplicity distributions of gluon and quark jets and tests of QCD analytic predictions, *Eur. Phys. J. C* **1**, 479 (1998).
- [84] J. Gallicchio and M. D. Schwartz, Quark and gluon jet substructure, *J. High Energy Phys.* **04** (2013) 090.
- [85] B. P. Roe, H. J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor, Boosted decision trees, an alternative to artificial neural networks, *Nucl. Instrum. Methods Phys. Res., Sect. A* **543**, 577 (2005).
- [86] A. Hocker *et al.*, TMVA—toolkit for multivariate data analysis, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- [87] CMS Collaboration, Search for new particles decaying to a jet and an emerging jet, *J. High Energy Phys.* **02** (2019) 179.
- [88] M. Aaboud *et al.* (ATLAS Collaboration), A search for pair-produced resonances in four-jet final states at $\sqrt{s} = 13$ TeV with the ATLAS detector, *Eur. Phys. J. C* **78**, 250 (2018).
- [89] C. Borschensky, M. Kramer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell, Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13, 14, 33$ and 100 TeV, *Eur. Phys. J. C* **74**, 3174 (2014).
- [90] M. Bahr *et al.*, Herwig++ physics and manual, *Eur. Phys. J. C* **58**, 639 (2008).
- [91] S. Mrenna and P. Skands, Automated parton-shower variations in Pythia 8, *Phys. Rev. D* **94**, 074005 (2016).