



## Composite octet searches with jet substructure

Yang Bai<sup>a</sup> and Jessie Shelton<sup>b</sup>

<sup>a</sup>SLAC National Accelerator Laboratory,  
2575 Sand Hill Road, Menlo Park, CA 94025, U.S.A.

<sup>b</sup>Department of Physics, Sloane Laboratory, Yale University,  
New Haven, CT, 06520, U.S.A.

E-mail: [yangbai@slac.stanford.edu](mailto:yangbai@slac.stanford.edu), [j.shelton@yale.edu](mailto:j.shelton@yale.edu)

**ABSTRACT:** Many new physics models with strongly interacting sectors predict a mass hierarchy between the lightest vector meson and the lightest pseudoscalar mesons. We examine the power of jet substructure tools to extend the 7 TeV LHC sensitivity to these new states for the case of QCD octet mesons, considering both two gluon and two b-jet decay modes for the pseudoscalar mesons. We develop both a simple dijet search using only the jet mass and a more sophisticated jet substructure analysis, both of which can discover the composite octets in a dijet-like signature. The reach depends on the mass hierarchy between the vector and pseudoscalar mesons. We find that for the pseudoscalar-to-vector meson mass ratio below approximately 0.2 a simple dijet analysis with only the jet mass variable provides the best discovery limit; for a ratio between 0.2 and the QCD-like value of 0.3, the full jet substructure analysis has the best discovery potential; for a ratio above approximately 0.3, the standard four-jet analysis is more suitable.

**KEYWORDS:** Technicolor and Composite Models, Beyond Standard Model, Jets

ARXIV EPRINT: [1107.3563](https://arxiv.org/abs/1107.3563)

---

**Contents**

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>A simple model for spin-1 and spin-0 composite octets</b>	<b>2</b>
<b>3</b>	<b>Discovery potential</b>	<b>7</b>
3.1	$\pi_G \rightarrow gg$	9
3.2	$\pi_G \rightarrow b\bar{b}$	13
<b>4</b>	<b>Discussion and conclusions</b>	<b>13</b>
<b>A</b>	<b>Vector-like confinement</b>	<b>16</b>

---

**1 Introduction**

With the advent of the LHC, the electroweak scale is being probed at last. However, the startlingly good agreement of the standard model (SM) with precision flavor and electroweak measurements begs the question: does TeV-scale physics substantially modify the SM story of electroweak symmetry breaking (EWSB)? Hints from low-energy data suggest that physics Beyond the Standard Model (BSM), if it has to do with flavor, EWSB, or leptons, may be heavy, while new physics at the TeV scale may more comfortably be hadrophilic, and not obviously related to the mysteries of EWSB. Such a situation poses theoretical puzzles. However, the possibility is one which should be seriously considered: Nature has a track record of handing us particles for which we have no obvious need. There are many open possibilities for interesting BSM physics at the TeV scale, many of which present large signals for the early LHC.

One such possibility is a new non-Abelian gauge interaction  $G$  which confines at scales  $\Lambda_G \gtrsim \text{TeV}$ . To access this sector, some new particles must couple to both the  $G$  sector and the SM; the simplest possibility is a fermion species  $\Psi$  which transforms under both  $G$  and the SM gauge groups. If  $\Psi$  is in a vector representation ( $\Psi_L$  and  $\Psi_R$  have the same quantum numbers) of the SM gauge group and does not have large mixings with SM fermions, SM precision observables are unaffected. For  $\Psi$  charged under QCD, after the group  $G$  confines, the spectrum of  $G$ -hadrons will include several new colored states [1–3].

This simple scenario naturally gives rise to signals which are particularly well-suited to study at the LHC: new colored states with masses at the TeV scale or below. In particular, the  $G$  sector will typically contain a relatively light color-octet vector meson  $\rho_G^a$  with  $m_{\rho_G} \sim \Lambda_G \gtrsim \text{TeV}$ , and colored pseudo-Nambu-Goldstone bosons (pNGB's)  $\pi_G^a$  with masses  $m_{\pi_G}$  parametrically lighter than  $\Lambda_G$ . These states are particularly important for the LHC phenomenology of a new confining interaction, as the  $\rho_G^a$  can be resonantly

produced through its mixing with the gluon, while the  $G$ -pions  $\pi_G^a$ , as the lightest colored particles in the  $G$  sector, have the largest production cross sections.<sup>1</sup>

While the production cross sections for these light composite states are large, the dominant decays of the  $\pi_G^a$  and the  $\rho_G^a$  yield all-hadronic final states, and large QCD backgrounds can make discovery challenging [5]. This is especially true of the vector resonance,  $\rho_G^a$ , which naturally has large branching fractions to other new states, rather than back to dijets, thus leading to multijet final states which can be challenging to separate from QCD backgrounds. Recent advances in jet substructure have extended LHC sensitivity to both SM [6–13] and BSM [14–32] signals in otherwise challenging multi-jet final states. Here, we will demonstrate the power of simple jet substructure techniques to improve LHC discovery sensitivity to colored resonances. The hierarchy of mass scales  $m_{\pi_G}/m_{\rho_G}$  means that the  $G$ -pions produced in the decay of a  $\rho_G^a$  are boosted. The subsequent decay  $\pi_G \rightarrow gg, b\bar{b}$  can be extracted from the large QCD background using a simple and flexible  $G$ -pion tagger, which distinguishes the perturbative  $G$ -pion decay from the shower structure of a QCD jet. Our main motivation is a confining gauge group, as this scenario naturally generates the hierarchy of mass scales which necessitates a substructure analysis, but the techniques presented here are useful for any theory with a colored vector resonance (“coloron” [33–35]) or colored axial vector resonance (“axigluon” [36–38]) with new colored daughters.

The organization of this paper is as follows. In section 2 we introduce a simplified model capturing the dynamics of interest and discuss the parameter space of the theory. In section 3 we introduce two simple jet substructure searches and show the discovery reach of the 7 TeV LHC for both. Section 4 contains our conclusions, and in appendix A we provide a more extended discussion of how our simplified model fits into the Lagrangian of a generic confining sector.

## 2 A simple model for spin-1 and spin-0 composite octets

One generic possibility for physics above the electroweak scale is a new gauge interaction  $G$  which confines at a scale  $\Lambda_G$  above the electroweak scale. New fermion species  $\Psi$  which transform under  $G$  then are not observed in isolation at colliders, but rather in bound states which are singlets of  $G$ , which we call  $G$ -hadrons. This idea is hardly new: technicolor is one example of such a model. Unlike technicolor, however, we do not necessarily imagine here that the chiral condensation of  $G$ -fermions is responsible for electroweak symmetry breaking.

In general the new confining gauge group will result in a rich spectrum of  $G$ -hadrons with a range of SM quantum numbers. In the spirit of Simplified Models [39], we introduce a simplified model which succinctly captures the most relevant dynamics for discovery at hadron colliders, especially in dijet or multi-jet final states. A discussion of how this simplified model maps onto such well-motivated extensions of the standard model as a new confining interaction is provided in appendix A.

---

<sup>1</sup>Here we take the bare masses of the new fermion species to be negligible in comparison with the confinement scale,  $m_\Psi \ll \Lambda_G$ . One can introduce a Peccei-Quinn symmetry to forbid the bare fermion masses [3]. The opposite situation,  $\Lambda_G \ll m_\Psi$ , leads to quirks, with very different phenomenology [4].

The most relevant degrees of freedom for hadron colliders are first, the lightest colored particles, which will enjoy the largest production cross sections, and second, vector octets, which can be resonantly produced through mixing with the gluon. Matter which is charged under both QCD and the new confining group  $G$  will typically lead to a multiplet of colored pNGB's which will be among the lightest  $G$ -hadrons. We will study here a pseudo-scalar octet of  $G$ -pions,  $\pi_G^a$ , which we will take to be electroweak singlets. Pseudo-scalar octets will always have a minimal pair-production cross-section at hadron colliders through their QCD interactions. However, if a heavier spin-one  $G$ -hadron like a  $G$ -vector meson  $\rho_G^a$  is also present, the  $\pi_G^a$  pair-production cross section can easily be enhanced by the potentially large resonant production of the  $\rho_G^a$  together with a large coupling between  $\rho_G^a$  and  $\pi_G^a$ .

We introduce here a phenomenological Lagrangian capturing the dynamics of the pseudo-scalars  $\pi_G^a$  together with an octet vector  $\rho_G^a$ . We will work with the effective Lagrangian

$$\begin{aligned}
 -\mathcal{L} = & -\frac{1}{2}D_\mu\pi_G^a D^\mu\pi_G^a + \frac{m_{\pi_G}^2}{2}\pi_G^a\pi_G^a - \frac{1}{4}\rho_G^{a\mu\nu}\rho_{G\mu\nu}^a + \frac{m_{\rho_G}^2}{2}\rho_{G\mu}^a\rho_G^{a\mu} \\
 & + \frac{\tan\theta}{2}\rho_G^{a\mu\nu}G_{\mu\nu}^a + g_\rho f^{abc}\rho_G^{a\mu}\pi_G^b D_\mu\pi_G^c,
 \end{aligned}
 \tag{2.1}$$

together with two dimension-five operators allowing the  $\pi_G^a$  to decay, which will be discussed later. This Lagrangian consists of mass and kinetic terms for the  $\rho_G^a$  and  $\pi_G^a$ , kinetic mixing between the  $\rho_G^a$  and QCD gluons  $G^a$ , and a  $\rho_G^a$ - $\pi_G^a$ - $\pi_G^a$  vertex analogous to the familiar  $\rho$ - $\pi$ - $\pi$  vertex in QCD. Here the  $\rho_G^a$  kinetic term is written in terms of

$$\rho_G^{a\mu\nu} \equiv D^\mu\rho_G^{a\nu} - D^\nu\rho_G^{a\mu},
 \tag{2.2}$$

with the covariant derivative  $D^\mu\rho_G^{a\nu} \equiv \partial^\mu\rho_G^{a\nu} + ig_s f^{abc}G^{b\mu}\rho_G^{c\nu}$ . After making the field redefinition  $G_\mu^a \rightarrow G_\mu^a + g_s \tan\theta\rho_{G\mu}^a$ , the kinetic mixing between the  $\rho_G$  and the gluon is removed, while introducing a coupling of the  $\rho^a$  to quarks,

$$-\mathcal{L}_{\rho_G q\bar{q}} = ig_s \tan\theta \rho_{G\mu}^a \bar{q} t^a \gamma^\mu q.
 \tag{2.3}$$

Here  $t^a$  denotes the QCD generators. It is the coupling of eq. (2.3) to quarks which leads to resonant production of the  $\rho_G$  at hadron colliders.<sup>2</sup>

The theory described by eq. (2.1) depends on four parameters: the masses  $m_{\rho_G}$  and  $m_{\pi_G}$ , the  $\rho_G^a - G^a$  mixing  $\tan\theta$ , and the  $\rho_G^a - \pi_G^a$  coupling  $g_\rho$ . The coupling  $\tan\theta$  governs the resonant  $\rho_G^a$  production cross section, while  $g_\rho$  controls the relative branching fraction of the  $\rho_G$  into  $G$ -pions or back into  $q\bar{q}$ . Although  $g_\rho$  and  $\tan\theta$  are independent parameters, for simplicity we take them to be related according to

$$g_\rho = \frac{g_s}{\tan 2\theta}.
 \tag{2.4}$$

This choice of relationship between  $g_\rho$  and  $\tan\theta$  is convenient for comparison to a weakly coupled renormalizable coloron model [35], which realizes similar phenomenology. For strong interaction models, one can estimate  $g_\rho \sim \sqrt{4\pi}$ , which corresponds to a small value

---

<sup>2</sup>The field redefinition of the gluon also shifts the value of  $g_\rho$ ; we absorb this shift into the definition of  $g_\rho$ .

of  $\theta \approx 0.14$ . We will concentrate on this portion of parameter space, where the dominant decay of the  $\rho_G$  is to  $G$ -pion pairs.<sup>3</sup> Once we have chosen a relationship between  $\tan\theta$  and  $g_\rho$ , the production times branching ratio  $\sigma(q\bar{q} \rightarrow \rho_G) \times \text{Br}(\rho_G \rightarrow q\bar{q})$  is fixed for a given  $m_\rho$  and  $\tan\theta$ . Results for other choices of  $g_\rho$  at a given  $\tan\theta$  can be obtained by scaling the branching ratio as desired, provided the total  $\rho_G$  width remains narrow.

The mass ratio  $m_{\pi_G}/m_{\rho_G}$  is important for determining the model's signatures at the LHC. Previous LHC studies [1, 2, 5, 40, 41] have focused on the region of parameter space where the *a priori* unknown ratio  $m_{\pi_G}/m_{\rho_G}$  is chosen by scaling from QCD, yielding [1]

$$\frac{m_{\pi_G}^2}{m_{\rho_G}^2} = 3 \left( \frac{\alpha_s}{\alpha} \right) \frac{\delta m_\pi^2|_{\text{EM}}}{m_\rho^2}, \tag{2.5}$$

where the observed electromagnetic contribution to the pion mass splitting is  $\delta m_\pi^2|_{\text{EM}} \simeq \frac{3\alpha}{4\pi} 2 \ln 2 m_\rho^2$  [42–45]. This model for the unknown  $G$  dynamics yields  $m_{\pi_G} \simeq 0.3 m_{\rho_G}$ . However, this scaling relies critically on specific features of the QCD spectral functions, whose genericity is unclear. On general grounds, and avoiding the usual QCD-like  $N_c$  scaling, we may expect the  $G$ -pion mass to scale like

$$m_{\pi_G}^2 \sim \frac{g_s^2}{(4\pi)^2} \Lambda_G^2, \tag{2.6}$$

where  $\Lambda_G$  is the scale where the  $G$  interactions become strong. For  $\rho_G$  with mass of order the cutoff, we can estimate

$$\frac{m_{\pi_G}}{m_{\rho_G}} \sim 0.1. \tag{2.7}$$

Our main interest will be to establish the discovery reach in the region of parameter space where  $m_{\pi_G}$  is sufficiently small compared to  $m_{\rho_G}$  that a flexible treatment of jets allows for better separation of signal from background. As we will see in section 3, our range of interest is therefore from  $m_{\pi_G}/m_{\rho_G} \approx 0.1$  up to the QCD-like value  $m_{\pi_G}/m_{\rho_G} \approx 0.3$ , while for larger mass ratios traditional multi-jet searches become more efficient.

Gauge invariance allows for additional renormalizable interactions beyond those in the simplified model of eq. (2.1) which we will neglect, and which are further discussed in the appendix A. Note that the leading interactions of the new colored degrees of freedom with the SM proceed only through QCD gauge interactions. This ensures agreement with precision electroweak and flavor constraints.

The renormalizable interactions of eq. (2.1) have a  $Z_2$  symmetry  $\pi_G^a \rightarrow -\pi_G^a$  and do not yet allow the  $\pi_G^a$  to decay. At dimension five, we can write down interactions which allow either  $\pi_G^a \rightarrow gg$  or  $\pi_G^a \rightarrow q\bar{q}$ . Pion decay to gluons is mediated by

$$\mathcal{O}_{\pi gg} = -\frac{g_s^2}{16\pi^2 f_{\pi_G}} \text{Tr}[t^a t^b t^c] \pi_G^a \epsilon_{\mu\nu\rho\sigma} G^{b\mu\nu} G^{c\rho\sigma}. \tag{2.8}$$

---

<sup>3</sup>For  $\theta \approx 0.14$ , the relation eq. (2.4) yields a smaller value for  $\rho_G - G$  mixing than would be obtained from scaling the observed  $\rho - \gamma$  mixing in QCD [1]. The only importance of this for our present purposes is in reducing the resonant  $\rho_G$  cross-section relative to the QCD-like expectation, and our analysis is in this sense conservative.

Here  $f_{\pi_G}$  is the  $G$ -pion decay constant,  $4\pi f_{\pi_G} \sim \Lambda_G$ . Using  $\text{Tr}[t^a t^b t^c] = \frac{1}{2}d^{abc} + \frac{i}{2}f^{abc}$ , only the  $d^{abc}$  part has non-vanishing contributions. The operator  $\mathcal{O}_{\pi_{gg}}$  is analogous to the operator mediating  $\pi^0 \rightarrow \gamma\gamma$  in the SM, and is naturally generated in many theories through triangular anomaly diagrams, as we discuss further in the appendix A. In addition, depending on the details of the model, the pion may also decay to quarks through the dimension-five operators

$$\mathcal{O}_{\pi q\bar{q}} = i \frac{c_{ij}^d}{M} \pi_G^a H \bar{Q}_L^i t^a \gamma^5 d_R^j + i \frac{c_{ij}^u}{M} \pi_G^a \tilde{H} \bar{Q}_L^i t^a \gamma^5 u_R^j + \text{h.c.} \quad (2.9)$$

Here  $M$  is an ultraviolet mass scale. Assuming for simplicity that these interactions are proportional to the SM Yukawa couplings,  $c_{ij}^{u,d} \propto Y_{ij}^{u,d}$ , after electroweak symmetry breaking these operators allow  $G$ -pion to decay through

$$\mathcal{O}_{\pi q\bar{q}} = i \frac{m_q}{M} \pi_G^a \bar{q} t^a \gamma^5 q, \quad (2.10)$$

where any order one coefficients have been absorbed into the definition of  $M$ . These couplings will favor  $\pi_G^a \rightarrow b\bar{b}$ .<sup>4</sup> At dimension-five level operators coupling  $D_\mu \pi_G^a$  to quarks also appear,  $\mathcal{O} \propto D_\mu \pi_G^a \bar{q}_{L,R} t^a \gamma^\mu \gamma^5 q_{L,R}$ . The derivative portion of these operators, upon use of the equations of motion, is equivalent to eq. (2.10), while the non-derivative portion contributes only to 3-body  $G$ -pion decays.

Since  $G$ -pion decay to both  $b\bar{b}$  and to gluons proceeds through higher dimension operators, details of the model can dramatically affect the branching ratios of the  $\pi_G^a$ . We will study discovery prospects for either  $\pi_G^a \rightarrow gg$  or  $\pi_G^a \rightarrow b\bar{b}$  as the dominant decay channel. Strikingly, we will find that simple jet substructure tools can dramatically enhance the prospects for discovery even when the gluonic decays dominate.

The decay widths of the  $\pi_G^a$  are given by

$$\Gamma(\pi_G \rightarrow gg) = \frac{5\alpha_s^2}{192\pi^3 f_{\pi_G}^2} m_{\pi_G}^3, \quad (2.11)$$

through eq. (2.8), and

$$\Gamma(\pi_G \rightarrow b\bar{b}) = \frac{m_{\pi_G}}{16\pi} \left( \frac{m_b}{M} \right)^2 \sqrt{1 - 4m_b^2/m_{\pi_G}^2}, \quad (2.12)$$

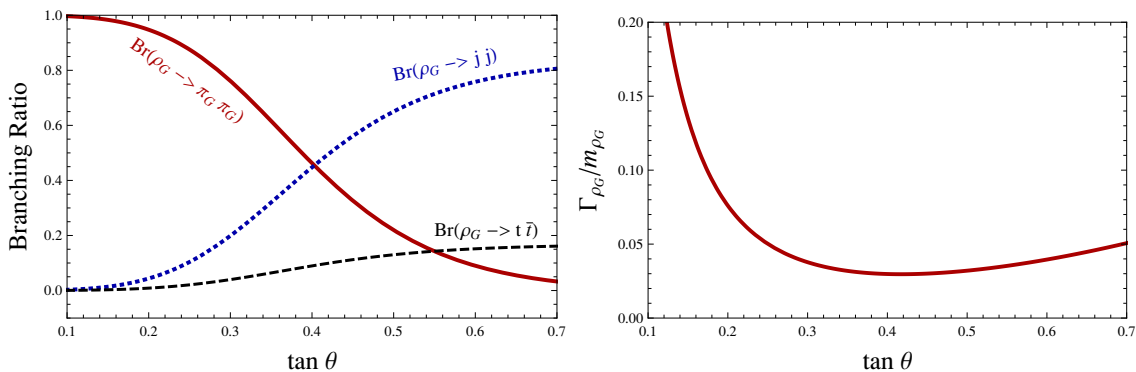
through eq. (2.10). The  $G$ -pions are narrow, and decay within the detector for parameter choices of  $f_{\pi_G}$  and  $M$  in this paper.

The decay widths of the  $\rho_G^a$  into two quarks and into two  $G$ -pions are given in terms of  $\tan\theta$  as

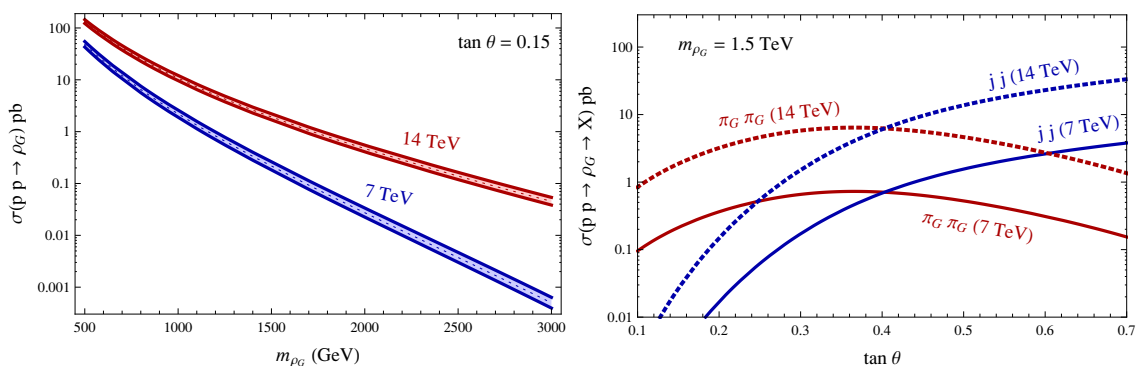
$$\begin{aligned} \Gamma(\rho_G \rightarrow q\bar{q}) &= \frac{\alpha_s}{6} \tan^2\theta m_{\rho_G} \left( 1 - \frac{4m_q^2}{m_{\rho_G}^2} \right)^{1/2}, \\ \Gamma(\rho_G \rightarrow \pi_G \pi_G) &= \frac{\alpha_s}{8 \tan^2 2\theta} m_{\rho_G} \left( 1 - \frac{4m_{\pi_G}^2}{m_{\rho_G}^2} \right)^{3/2}, \end{aligned} \quad (2.13)$$

---

<sup>4</sup>If the pion is sufficiently heavy that the decay to tops is open,  $m_{\pi_G} > 2m_t$ , then  $\pi_G^a \rightarrow t\bar{t}$  will dominate the quark decay modes. The final state will be  $4t$ 's. We will largely be interested in  $\pi_G^a$  below the  $t\bar{t}$  threshold, and will consequently neglect the  $t\bar{t}$  decay mode.



**Figure 1.** Left panel: the branching ratios of  $\rho_G^\mu$  into different modes as a function of the mixing angle. Right panel: the  $\rho_G^\mu$  width over its mass as a function of  $\tan\theta$ . The daughter particle masses are neglected.



**Figure 2.** Left panel: the production cross section of  $\rho_G$  at the LHC for  $\tan\theta = 0.15$ . The range of cross section is for two different renormalization scales:  $\frac{m_{\rho_G}}{2}$  (upper) and  $2m_{\rho_G}$  (lower). Right panel: the production cross section of different decaying modes, where the renormalization scale is fixed to be  $m_{\rho_G}$ .

where the branching fraction into quarks is per flavor. For a heavy  $\rho_G^a$  much above the  $\pi_G^a$  and the top quark masses, we have the branching ratios and the width over mass ratio shown in figure 1.

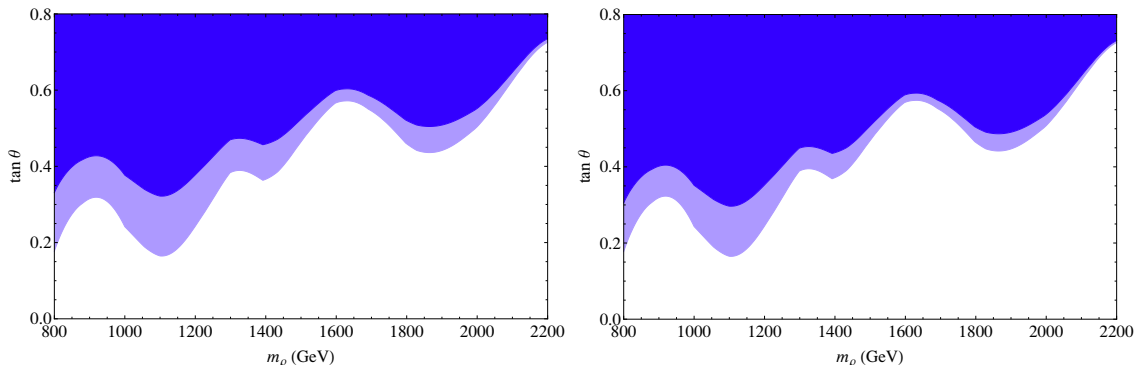
Using the narrow width approximation to estimate the cross section for producing a  $\rho_G^a$  in the  $s$ -channel gives

$$\sigma(q\bar{q} \rightarrow \rho_G^\mu) \approx \frac{8\pi^2\alpha_s \tan^2\theta}{9m_{\rho_G}} \delta(\sqrt{\hat{s}} - m_{\rho_G}). \tag{2.14}$$

Convoluting this partonic cross section with the MSTW [46] PDFs yields the LHC production cross sections shown in figure 2.

The parameter space of  $\tan\theta$  and  $m_{\rho_G}$  is subject to various constraints, most notably  $t\bar{t}$  and dijet resonance searches. The latest  $t\bar{t}$  narrow resonance searches with  $200 \text{ pb}^{-1}$  at Atlas [47] do not constrain our model parameter space, because of the suppressed branching ratio of  $\rho_G \rightarrow t\bar{t}$  seen in figure 1. However, the dijet resonance searches [48–51] do constrain our parameter space, and the limits are shown in figure 3. Applying the dijet limits to





**Figure 3.** Dijet constraints on the model parameters at 95% C.L. from Atlas dijet narrow resonance searches with  $163 \text{ pb}^{-1}$  luminosity [51]. The shaded regions are excluded. The two boundary lines are for (lower) 100% efficiency to detect  $\pi_G^a$  as a single jet and (upper) 0% efficiency to detect  $\pi_G^a$  as a single jet. Branching ratios are given by eq. (2.13). The left panel shows limits for  $m_{\pi_G}/m_{\rho_G} = 0.1$  while the right panel shows limits for  $m_{\pi_G}/m_{\rho_G} = 0.3$ .

our model is not completely straightforward, as there will be some nonzero efficiency for  $\rho_G \rightarrow \pi_G \pi_G$  events to be reconstructed in the  $\rho_G \rightarrow jj$  sample. This efficiency depends on the mass ratio  $m_{\pi_G}/m_{\rho_G}$  and on the jet algorithm used by Atlas. In figure 3, we show the constraints in the  $m_{\rho_G} - \tan \theta$  plane from the dijet resonance search for two fixed ratios of  $m_{\pi_G}/m_{\rho_G}$ . The two boundaries in each plot correspond to (lower) 100% efficiency to detect  $\pi_G^a$  as a single jet and (upper) 0% efficiency to detect  $\pi_G^a$  as a single jet, with the branching ratios as given in eq. (2.13). The limits are sensitive to the mass ratio only through the factor  $(1 - 4 m_{\pi_G}^2/m_{\rho_G}^2)^{3/2}$ , and thus the limits are broadly similar for mass ratios  $\ll 1$ .

Finally, there are few meaningful limits on the  $G$ -pion masses. While light  $G$ -pions can be pair-produced through  $gg \rightarrow \pi_G \pi_G$  at the Tevatron, no multi-jet or multi- $b$  searches have limited their masses [1, 35]. Introducing operators beyond those in the simplified model Lagrangian of eq. (2.1) can potentially lead to indirect limits, as we discuss further in the appendix A.

### 3 Discovery potential

In this section we will detail discovery prospects for  $\pi_G^a$  and  $\rho_G^a$ . The process we will use for discovery is  $\pi_G^a$  pair production from an initial  $\rho_G^a$  resonance,  $q\bar{q} \rightarrow \rho_G \rightarrow \pi_G \pi_G$ , followed by  $\pi_G \rightarrow b\bar{b}, gg$ . This process, in contrast to non-resonant  $G$ -pion pair production, is particularly useful in theories with hierarchical spectra, where the  $G$ -pions coming from the  $\rho_G$  are sufficiently boosted that their daughter partons have a reduced probability to be reconstructed as separate jets. In this regime, the kinematics of the boosted  $\pi_G^a$  allow for easier separation of signal from background than do the non-boosted  $\pi_G^a$  coming from nonresonant QCD pair production. Simple jet substructure analyses then suffice to give excellent discovery reach over much of the simplified model parameter space. Note that as our signal has the same color structure as the background, the utility of jet substructure



techniques is entirely due to their flexible and efficient characterization of the distinctive signal kinematics.

The main task is to distinguish a collimated perturbative two-body decay  $\pi_G^a \rightarrow jj$  from a QCD jet. When the pions are sufficiently boosted that both their daughter jets have a moderate probability to be contained in a standard (here  $R = 0.7$  anti- $k_T$  [60]) jet, the jet mass alone can provide significant improvements over a standard dijet search, as we will discuss below. For slightly less boosted  $G$ -pions, a more involved analysis improves the prospects. We employ a fat jet analysis based on the mass drop procedure pioneered by [6, 7]. Specifically, we cluster the events on a large angular scale ( $R = 1.2$ ) using the Cambridge/Aachen (C/A) algorithm and require two fat jets with  $p_T > p_{T,\text{cut}}$ . The fat jets are then each iteratively decomposed by undoing the clustering sequence step by step in search of a splitting which resembles a perturbative decay. At each splitting of a parent  $J$  to two daughters  $j_1, j_2$  with  $m_{j_1} > m_{j_2}$ , we check whether the splitting

- shows a sudden drop in the jet mass,  $m_{j_1} < \mu m_J$ ,
- and is relatively symmetric,  $\min(p_{T,j_1}^2, p_{T,j_2}^2) \Delta R_{j_1,j_2}^2 / m_j^2 > r_{xy}$ .

Optimal values for the mass drop variable  $\mu$  and the symmetric splitting cut  $r_{xy}$  will be chosen below. If both conditions are satisfied, one identifies  $J$  as the fat jet and  $j_{1,2}$  as the subjets and exits the loop. Otherwise, one replaces  $J$  by  $j_1$  and repeats the previous procedure. In addition to vetoing QCD, the mass drop analysis [7] helps clean up the jets and improves mass resolution.<sup>5</sup>

Another observable which can distinguish a perturbative decay from a QCD branching is the jet shape  $N$ -subjettiness [54]. Given a jet found with initial radius  $R$  and a set of  $N$  subjet centers  $j_k$  found (with some algorithm) inside the jet, the  $N$ -subjettiness of the jet is

$$\tau_N = \frac{\sum_i p_{T,i} \min[\Delta R_{ik}]}{\sum_i p_{T,i} R}, \tag{3.1}$$

where the sum runs over the particles in the jet, and  $\Delta R_{ik}$  is the distance between the  $i^{\text{th}}$  particle and the  $k^{\text{th}}$  subjet axis. Jets with smaller (larger) values of  $\tau_N$  have radiation more (less) concentrated around the subjet axes, and are therefore more (less) amenable to a description in terms of  $N$  subjets. Since the QCD background tends to have larger values of  $\tau_2/\tau_1$  than the signal, the ratio  $\tau_2/\tau_1$  can be used as a good discriminant to reduce QCD backgrounds. We find that while  $N$ -subjettiness and the mass drop procedure are clearly correlated, they are sufficiently distinct that incorporating a cut on  $N$ -subjettiness marginally improves discovery sensitivity.

We incorporate both the mass drop procedure and  $N$ -subjettiness into a simple and flexible tagger designed to discriminate a boosted  $G$ -pion from a QCD jet. The tagger constructs a fat C/A jet with  $R = 1.2$ . From the constituents of this fat jet, we construct two exclusive subjets using the  $k_T$  algorithm, yielding the two subjet axes we use to evaluate  $\tau_2/\tau_1$ . We require that the fat jet passes a cut on  $\tau_2/\tau_1$  in addition to the mass drop

---

<sup>5</sup>We do not implement filtering [7] or other jet grooming tools [52, 53], nor do we simulate pileup, though in a full analysis both pileup and jet grooming will be necessary.

criterion. Two (C/A) subjects  $j_1$  and  $j_2$  are identified in the mass drop procedure, and the tagger incorporates cuts on both the sum of their transverse momentum  $p_{1T} + p_{2T}$  as well as their invariant mass  $m_{j_1 j_2}$ . The specific values used for the cuts will be discussed further below. The jet mass alone is a useful jet substructure variable [14, 15], and we will also demonstrate the reach of a search which uses only the jet mass.

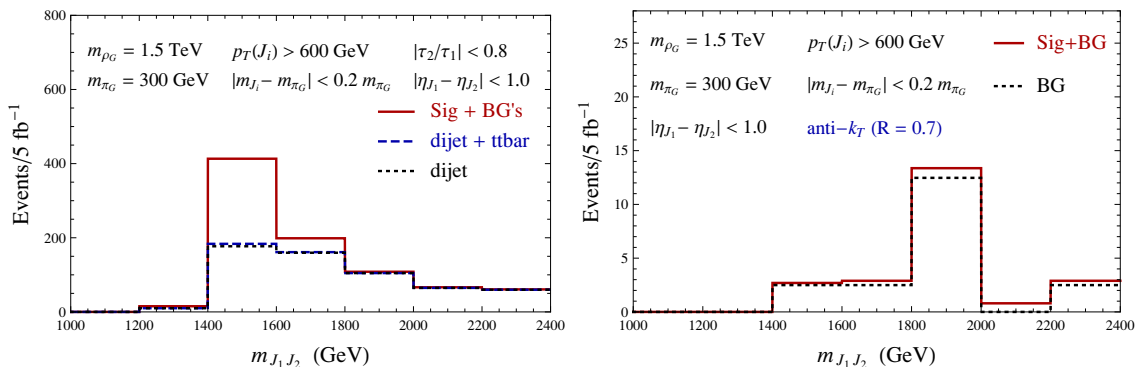
Before presenting results, we describe our simulation procedure. The production cross sections from figure 2 vary from 100 fb to 400 fb at the 7 TeV LHC for the mixing angle  $0.1 < \tan \theta < 0.4$ . To be concrete, we choose  $\tan \theta = 0.15$  or  $\sigma(u\bar{u} \rightarrow \rho_G \rightarrow \pi_G \pi_G) \approx 200$  fb for  $m_{\rho_G} = 1.5$  TeV throughout this section. The backgrounds are dominated by QCD dijets. We use the leading order cross-section as calculated in MadGraph [55], as comparison with measured dijet cross-sections [48, 49] indicates good agreement (i.e.,  $K$ -factors near unity) in the high- $p_T$ , large invariant mass regime of interest. Renormalization and factorization scales are set at  $\mu = m_{\rho_G}$ . The subleading  $W$ + jets and  $t\bar{t}$  backgrounds are negligible compared to the dijet background. Both signal and background events are generated with MadGraph [55] using CTEQ6L PDFs [56] and showered in Pythia 6.4.24 [57]. We then bin visible particles with  $|\eta| < 2.5$  into massless  $0.1 \times 0.1$  calorimeter cells and pass to FastJet [58, 59] for clustering and subsequent jet analysis.

Recent studies have demonstrated that Pythia and Herwig show reasonable agreement both with each other and with the data for jet masses in the range of interest. At high masses, Pythia and Herwig give nearly indistinguishable predictions for large C/A jets put through the mass drop procedure. Results for  $R = 1.0$  anti- $k_T$  jets indicate that while overall agreement is good, Pythia tends to underpredict QCD jet masses by 15–20% in the mass range  $100 \text{ GeV} < m_j < 200 \text{ GeV}$  [63]. We thus conclude that the numbers we will obtain for the analysis built on the full tagger are representative, while the alternate analysis using only anti- $k_T$  jet masses is likely to be slightly optimistic due to the tendency of Pythia to underpredict background QCD jets in the mass range of interest. The performance of the  $G$ -pion tagger on QCD dijets can be validated using dijet events where only one jet has a mass within the  $G$ -pion mass range and the other is light ( $m_j \lesssim 50 \text{ GeV}$ ).

### 3.1 $\pi_G \rightarrow gg$

For the case where  $\pi_G$  dominantly decays into two gluons, the signal is  $q\bar{q} \rightarrow \rho_G \rightarrow \pi_G \pi_G \rightarrow 4g$ 's. We will first illustrate our reconstruction procedures at the specific point  $m_{\rho_G} = 1.5 \text{ TeV}$  and  $m_{\pi_G} = 300 \text{ GeV}$ , and then present the discovery potential for other combinations of  $m_{\rho_G}$  and  $m_{\pi_G}$ .

We find that the final discovery significance is relatively insensitive to varying  $R$ , the mass drop  $\mu$ , and the symmetricity cut  $r_{xy}$ . We fix these parameters at  $R = 1.2$ ,  $\mu = 0.3$  and  $r_{xy} = 0.3$  to generate the left panel in figure 4. After implementing the substructure tagger, we require two tagged  $G$ -pion candidates in the event with  $p_T(j_i) > 600 \text{ GeV}$ . Note the  $p_T$  cut is placed on the final  $G$ -pion candidate rather than the initial fat jet. We also require  $|\eta_{J_1} - \eta_{J_2}| < 1.0$ , as the signal  $\pi_G$ 's are more central than QCD dijet backgrounds. Taking into account the jet energy resolution [62] and jet mass resolution [63], we further require additional mass window cuts  $|m_{J_i} - m_{\pi_G}| < 0.2 m_{\pi_G}$ . We show the histogram distributions of signal and background events in the left panel of figure 4. As demonstrated in this



**Figure 4.** Left panel: signal and backgrounds for a representative parameter point at the 7 TeV LHC using the full jet substructure analysis. Right panel: signal and backgrounds at the 7 TeV LHC for the same parameter point, with the analysis using only jet mass.

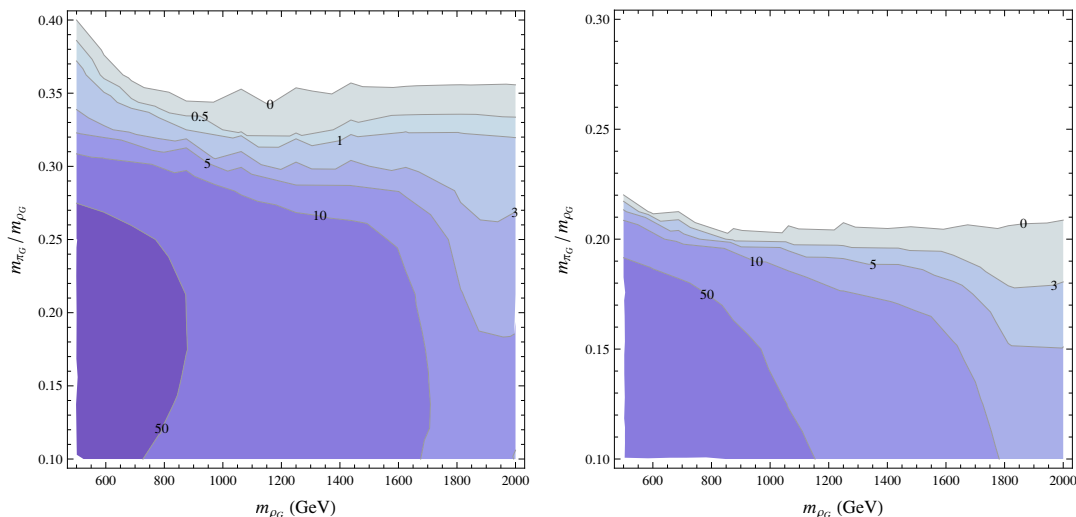
Cuts	Signal ( $m_{\rho_G} = 1.5 \text{ TeV}, m_{\pi_G} = 300 \text{ GeV}$ )	dijet	$t\bar{t}$
$C/A (R = 1.2), p_T^j > 500 \text{ GeV}$	150 fb	31045 fb	221 fb
$\mu = 0.3, r_{xy} = 0.3$ (mass drop)	133 fb	16773 fb	217 fb
$p_T^j > 600 \text{ GeV}$	94.5 fb	7091 fb	64 fb
$ \eta_{J_1} - \eta_{J_2}  < 1.0$	94.0 fb	6987 fb	63 fb
$ m_{J_i} - m_{\pi_G}  < 0.2 m_{\pi_G}$	56 fb	128 fb	1.6 fb
$ m_{J_1 J_2} - m_{\rho_G}  < 0.2 m_{\rho_G}$	54 fb	69 fb	1.6 fb

**Table 1.** Signal and background cross sections after different cuts.

figure, the  $t\bar{t}$  background only contributes a tiny fraction of the total background;  $W + j$  (not shown) is below  $t\bar{t}$ . Further imposing a mass window cut  $|m_{J_1 J_2} - m_{\rho_G}| < 0.2 m_{\rho_G}$ , we find that the discovery significance is  $S/\sqrt{B} \approx 14$  for the 7 TeV LHC with  $5 \text{ fb}^{-1}$  luminosity. To understand the effects of mass window cuts, we show the signal and background cross sections after different cuts in table 1.

As a comparison, we also estimate the discovery significance obtained by using a simple extension of the traditional dijet resonance searches, which are performed with anti- $k_T$  jets at a fixed  $R$ . On top of the usual cuts, namely jet  $p_T$  cuts and the dijet mass window cut, we also require both jet masses to be within the  $\pi_G^a$  mass window. The efficacy of this search depends on the efficiency for a boosted  $G$ -pion to be contained within a single jet. In contrast to traditional searches, which use *smaller* cone sizes for high- $p_T$  jets, it is thus advantageous to use *larger* cone sizes. There is, however, a trade-off: the mass resolution degrades rapidly with cone size, scaling roughly as  $R^3$ . Therefore going to larger cone sizes without employing additional substructure techniques does not help.

We show results for an anti- $k_T$  cone size of  $R = 0.7$ , which is the largest standard cone size in use at the LHC. Together with the broad pion mass windows, this cone size is a reasonable trade-off between efficiency and resolution. A value of  $R = 1.0$  as studied in [61, 63] will improve the reach, *provided* that jet grooming techniques such as trimming and pruning can recover the mass resolution. As the performance of trimming/pruning has not yet been established in data, we do not pursue this further.



**Figure 5.** Left panel: the discovery significance for different masses of  $\rho_G$  and  $\pi_G$  for  $\pi_G \rightarrow gg$ . We scanned five variables to find the optimized significance: the mass drop variable  $\mu$ , the symmetric splitting cut  $r_{xy}$ , the  $p_T$  cut of the fat jets,  $\Delta\eta$  of the two fat jets, and the  $N$ -subjettiness variable  $\tau_2/\tau_1$ . The numbers besides each contour line are the significance in  $\sigma$ . Right panel: the same as the left panel but instead of using the full jet substructure analysis, only the jet masses,  $p_T$ 's, and  $\Delta\eta$  are used in this plot.

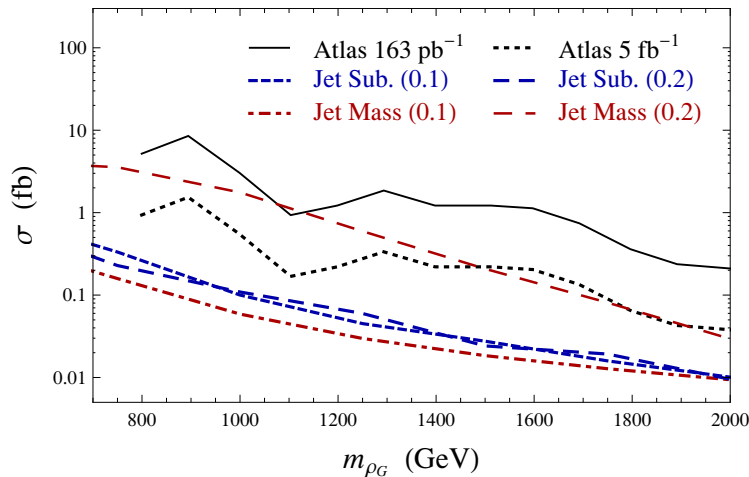
We show the histograms of the signal and background events in the right panel of figure 4. The discovery significance is around  $2\sigma$  for this parameter point, much poorer than the result obtained from the jet substructure analysis. Because  $R = 0.7 < 2m_{\pi_G}/p_{T,\text{cut}}$ , the jet clustering algorithm in the simple dijet search will typically not capture all the signal decay products in a single jet and hence suffers a reduction in the discovery significance. We have checked that for a different mass combination,  $m_{\rho_G} = 1.5 \text{ TeV}$  and  $m_{\pi_G} = 150 \text{ GeV}$ , the dijet resonance search supplemented with jet mass can obtain a discovery sensitivity as good as the jet substructure analysis.

For different mass combinations and especially when there are few signal and background events, we use the Poisson distribution to quantify the discovery significance as

$$\text{significance} \equiv \sqrt{-2 \ln [e^{-S-B}(S+B)^B/\Gamma(B+1)]}. \quad (3.2)$$

For different  $\rho_G^a$  masses and different values of the mass ratio  $m_{\pi_G}/m_{\rho_G}$ , we find the best discovery significance for each mass point in the left panel<sup>6</sup> shown in figure 5 by scanning the cut on  $\mu$  from 0.2 to 0.4 with a step of 0.05, the cut on  $p_T$  from  $m_{\rho_G}/3$  to  $m_{\rho_G}/3 + 300 \text{ GeV}$  with a step of 50 GeV, the cut on  $|\eta_{J_1} - \eta_{J_2}|$  from 0.5 to 5 with a step of 0.5, the cut on  $r_{xy}$  from 0.2 to 0.4 with a step of 0.05, the cut on  $r_{xy}$  from 0.2 to 0.9 with a step of 0.1. We further require the mass window cuts  $|m_{J_i} - m_{\pi_G}| < 0.2 m_{\pi_G}$  and  $|m_{J_1 J_2} - m_{\rho_G}| < 0.2 m_{\rho_G}$ . To obtain the left panel of figure 5, we have scanned 7 different  $\rho_G$  masses from 500 GeV to 2 TeV with a 250 GeV interval and 7 different mass ratios from 0.1 to 0.4 with a 0.05

<sup>6</sup>Strictly, there is an additional trials factor associated with the substructure searches due to the unknown  $m_{\pi_G}$ .

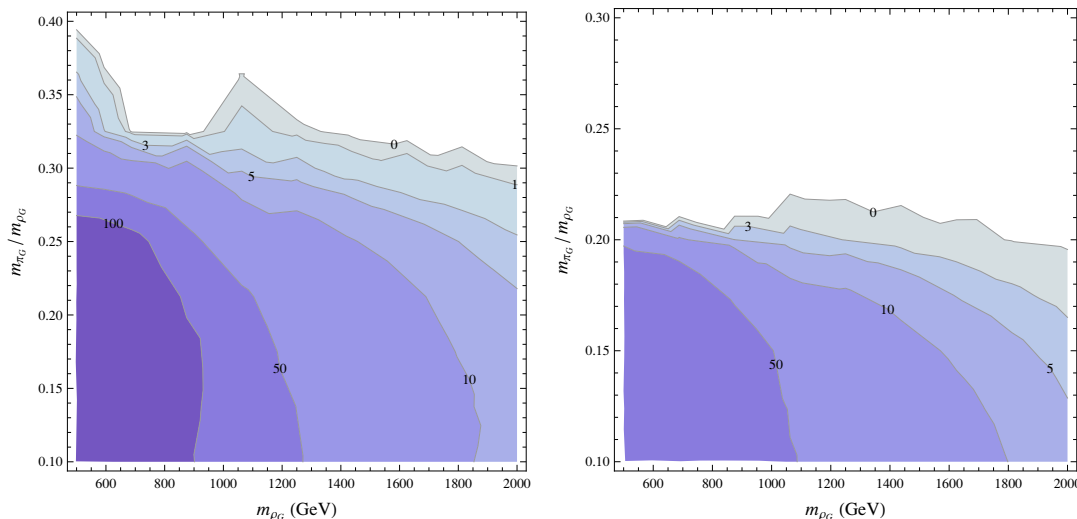


**Figure 6.** The 95% C.L. exclusion limit on the resonance production cross sections times branching ratio from different searches. The dotted black line is the projected Atlas exclusion limit at  $5 \text{ fb}^{-1}$  based on the current limit with  $163 \text{ pb}^{-1}$  luminosity [51]. The numbers in parentheses denote the ratio  $m_{\pi_G}/m_{\rho_G}$ .

interval. As can be seen from figure 5, the jet substructure analysis can discover the composite color octets for a wide range of masses. For smaller mass ratios of  $m_{\pi_G}/m_{\rho_G}$ , the discovery significances are better. Generically for  $m_{\pi_G}/m_{\rho_G} > 0.3$ , the jet substructure analysis loses its effectiveness and one should instead carry out a more traditional multi-jet resonance analysis to cover this region [1, 34, 35, 65].

As a comparison, we show the discovery limit in the right panel of figure 5 by using the ordinary dijet searches (with anti- $k_T$  and  $R = 0.7$ ) and requiring the two jet masses satisfying the mass window cuts  $|m_{J_i} - m_{\pi_G}| < 0.2 m_{\pi_G}$  and  $|m_{J_1 J_2} - m_{\rho_G}| < 0.2 m_{\rho_G}$ . From this plot, one can see that this very simple analysis can discover the  $\rho_G^a$  together with the  $\pi_G^a$  especially for  $m_{\pi_G}/m_{\rho_G} < 0.2$ . Comparing it with the left panel of this figure, one can see that for the light  $\rho_G^a$  mass region the traditional dijet resonance searches with jet mass constraints are even better than the more involved jet-substructure analysis.

Finally, we compare the sensitivities from the traditional dijet searches and from the jet substructure searches in figure 6. We take the current results from narrow resonance searches in dijets at Atlas with  $163 \text{ pb}^{-1}$  [51], and plot the projected 95% C.L. exclusion limit on the production cross section times dijet branching ratio at  $5 \text{ fb}^{-1}$  by assuming statistically dominated errors for the backgrounds. We show results from the full jet substructure analysis as well as the simple jet mass analysis (with  $R = 0.7$  anti- $k_T$ ) at the 7 TeV LHC with  $5 \text{ fb}^{-1}$ : here the vertical axis is cross-section times  $G$ -pion branching ratio. As can be seen from figure 6, for a small mass ratio  $m_{\pi_G}/m_{\rho_G} = 0.1$  the simple jet mass analysis provides the best exclusion limit, while for a small ratio  $m_{\pi_G}/m_{\rho_G} = 0.2$  the full jet substructure analysis is the most sensitive one. To produce this plot, we have neglected the acceptance of the traditional dijet analysis, which is large and close to  $70\% \sim 80\%$ .



**Figure 7.** The same as figure 5 for the discovery significance for different masses of  $\rho_G$  and  $\pi_G$  but for  $\pi_G \rightarrow b\bar{b}$ .

### 3.2 $\pi_G \rightarrow b\bar{b}$

For the case where the main decay channel of  $\pi_G$  is two  $b$ -jets, we repeat the same analysis as the four gluon case except that we now additionally demand two  $b$  tags in the final state. Although the signal contains four  $b$ -quarks, we have found that requiring two  $b$ -tags for the four daughter jets is sufficient to reject the backgrounds.

The backgrounds now come from both two light jets with a double  $b$  mistag, and two  $b$ -jets. After taking into account the  $b$ -tagging efficiency, these two contributions to the background are comparable. We assume a  $b$ -tagging efficiency of 60% and a mistagging efficiency of 2% for light jets (the  $c$ -jet has a larger mistagging efficiency which we compensate for by choosing a larger value of mis-tagging efficiency for all light jets). Improved  $b$ -tagging efficiencies (70% efficiency without increasing mistagging rates) may be possible [64], but as our final state contains more hadronic activity than the  $(0, 1, 2)\ell + 1$  fat jet states where these studies were performed, we conservatively do not use these improved numbers.

We require each fat jet to contain at least one  $b$ -tagged subjet. Performing the same scan of mass combinations as in the four gluon case, we find the discovery significance shown in figure 7.

## 4 Discussion and conclusions

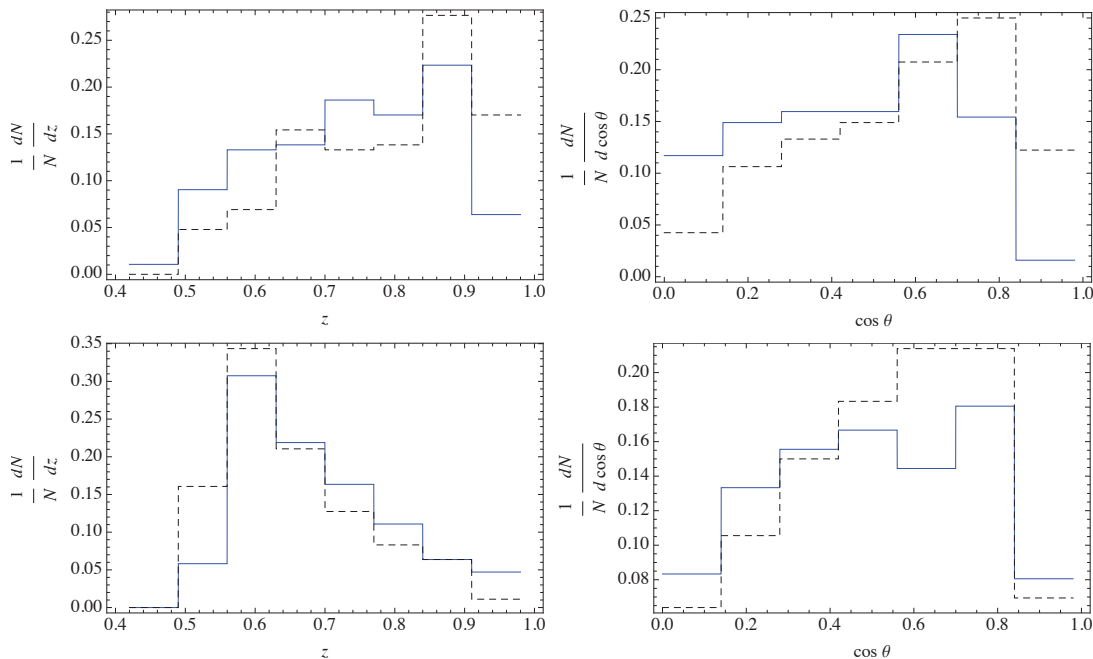
We have demonstrated the excellent potential of the 7TeV LHC to discover composite octets. We emphasize that the search strategies presented here are designed to utilize jet techniques which are being calibrated and validated for other signals. The typically large branching fractions of colored vector resonances to BSM daughters instead of to dijet final states makes their discovery difficult: the clean dijet signature has a suppressed rate, while the multijet signature arising from  $\rho_G \rightarrow \pi_G\pi_G$  can be difficult to reconstruct. We have

demonstrated how jet substructure techniques improve the reconstruction of the  $\rho_G$  and extend the discovery reach of the traditional dijet analysis for colored spin-1 resonances. The topology of the final state in resonant  $\rho_G$  production depends strongly on the mass ratio  $m_{\pi_G}/m_{\rho_G}$ . For large hierarchies,  $m_{\pi_G}/m_{\rho_G} \lesssim 0.2$ , a simple search augmenting dijet resonance search with an additional cut on jet mass works very well. For intermediate hierarchies, including the QCD-like region where  $m_{\pi_G}/m_{\rho_G} \approx 0.3$ , the final state interpolates between the boosted dijet topology and the non-boosted four-jet topology, and a more involved jet substructure analysis using a simple  $G$ -pion tagger gives the best sensitivity.

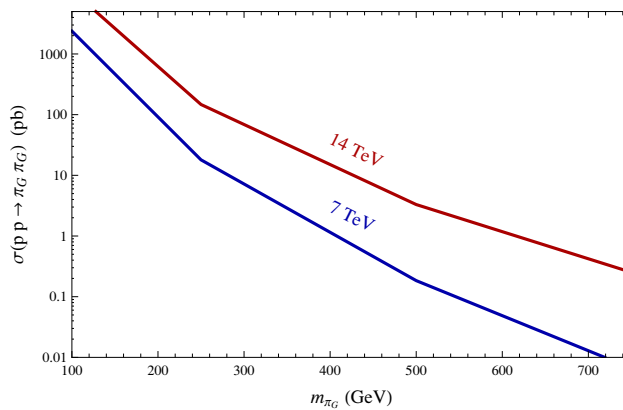
While generally incorporating additional substructure cuts improves discovery reach even in the highly boosted cases, the simple jet mass search has some advantages, as we briefly discuss. Namely, the mass drop and  $N$ -subjettiness cuts used in the full  $G$ -pion tagger shape the angular distributions of the jets coming from  $\pi_G^a \rightarrow jj$ , which makes probing the  $G$ -pion quantum numbers more challenging. Requiring that the jets coming from  $\pi_G^a \rightarrow jj$  be sufficiently hard and symmetric to be distinguished from typical QCD branchings preferentially selects the portion of the angular distribution which is transverse to the axis of the  $G$ -pion boost. This surviving slice of the angular distribution contains reduced information and renders determination of the  $G$ -pion spin difficult. The octets considered here have intrinsically better signal to background than the boosted Higgses considered in [66], and consequently more of the angular distributions can be retained. In the portion of parameter space where jet mass rather than the symmetry of the subjects are driving discovery sensitivity, the subjects identified within the boosted  $G$ -pion jet (via the mass drop procedure, or, for example, by simply resolving at a small angular scale  $R = 0.3$  and selecting the hardest subject) preserve more of the underlying angular distribution and might allow determination of the  $G$ -pion Lorentz quantum numbers. Let  $j_1$  be the hardest (anti- $k_T$ )  $R = 0.3$  subject of a jet  $J$  in the lab frame. In the left and right panels of figure 8, we show the normalized signal and background distributions of  $z = p_T^{j_1}/p_T^J$ , and  $\cos \theta$ , the angle of  $j_1$  in the rest frame of  $J$ , respectively, for the reference point  $m_{\rho_G} = 1500$  GeV,  $m_{\pi_G} = 150$  GeV, after all other cuts have been applied, and compare the results from the jet mass search to the results from the full substructure search. Note that there is less shaping of the signal angular distribution in the jet mass search. While cuts on either  $z$  or  $\cos \theta$  do not substantially improve the discovery significance, the difference in the angular distributions can be useful post-discovery for determining the Lorentz properties of the signal.

We have focused on resonant  $\rho_G \rightarrow \pi_G \pi_G$  production, neglecting the nonresonant QCD pair-production of  $\pi_G$ . For completeness, we show the production cross section of  $pp \rightarrow \pi_G \pi_G$  in figure 9 for the 7 TeV as well as 14 TeV LHC. Nonresonant pair production can allow octet  $G$ -pions to be discovered at the LHC using mass window cuts [5, 34]. We want to emphasize that the studies performed in this paper could lead to the simultaneous discovery of *two* composite color octet particles. Alternatively, if the octet  $G$ -pion is first found in non-boosted multi-jet final states, its mass may be used as an input to the  $G$ -pion tagger in a search for the  $\rho_G$ . The main advantage of the techniques presented here is the improved sensitivity to the  $\rho_G^a$ . Especially for theories with  $m_{\pi_G}/m_{\rho_G} \lesssim 0.2$ , the analyses we propose could be the best way to discover the  $\rho_G^a$  and to understand the detailed properties of a new strong interaction.





**Figure 8.** Angular variables, shown for the point  $m_{\rho_G} = 1500 \text{ GeV}$  and  $m_{\pi_G} = 150 \text{ GeV}$ . The left column shows signal (blue, solid) and background (black, dashed) distributions of the angular variable  $z$  after passing all other kinematical cuts. The right column shows signal and background distributions of the angular variable  $\cos \theta$ . The top row shows results of the jet mass analysis. The bottom row shows results of the full substructure analysis. The angular variables are defined in the text.



**Figure 9.** The pair production cross section of two  $\pi_G$ 's from QCD interactions.

### Acknowledgments

Thanks to W. Goldberger, S. Hoeche, and D. Zerwas for useful discussions. We are grateful to M. Cacciari and G. Salam for assistance with Fastjet and Peter Skands for assistance with Pythia. This material is based upon work supported in part by the National Science Foundation under Grant No. 1066293 and the hospitality of the Aspen Center for Physics.

JS was supported in part by DOE grant DE-FG02-92ER40704. This work was supported in part by the facilities and staff of the Yale High Performance Computing Center, and by the NSF grant CNS 08-21132. SLAC is operated by Stanford University for the US Department of Energy under contract DE-AC02-76SF00515.

## A Vector-like confinement

In this appendix we show how a well-motivated extension of the SM maps on to the simplified model discussed in section 2.

We suppose here that the new gauge sector includes a fermion species  $\Psi$  which transforms as a fundamental under QCD and as  $G_\psi$  under  $G$ ,

$$\Psi_L = (G_\psi, 3), \quad \bar{\Psi}_R = (\bar{G}_\psi, \bar{3}). \tag{A.1}$$

The theory possesses a global  $\Psi$  flavor symmetry  $SU(3)_L \times SU(3)_R$  other than the global baryon symmetry in the  $G$ -sector. When the gauge group  $G$  confines at a scale  $\Lambda_G$ , this chiral symmetry is broken down to the weakly gauged diagonal  $SU(3)_c$ , leaving an octet of pNGB's which we denote  $\pi_G^a$ . In the strong interacting  $G$ -sector, there could also exist vector mesons as well as axivector mesons. As the lightest axivector meson is in principle heavier than the vector meson as is the case in QCD, and has suppressed resonant cross sections, we only consider the vector meson in the following.

We reproduce here the minimal Lagrangian of eq. (2.1),

$$\begin{aligned}
 -\mathcal{L} = & -\frac{1}{2} D_\mu \pi_G^a D^\mu \pi_G^a + \frac{m_{\pi_G}^2}{2} \pi_G^a \pi_G^a - \frac{1}{4} \rho_G^{a\mu\nu} \rho_{G\mu\nu}^a + \frac{m_{\rho_G}^2}{2} \rho_{G\mu}^a \rho_G^{a\mu} \\
 & + \frac{\tan \theta}{2} \rho_G^{a\mu\nu} G_{\mu\nu}^a + g_\rho f^{abc} \rho_G^{a\mu} \pi_G^b D_\mu \pi_G^c,
 \end{aligned} \tag{A.2}$$

and comment on additional possible terms and their consequences.

First, terms polynomial in  $\pi_G^a$  will generically be present, but suppressed due to the approximate shift symmetry of the  $G$ -pions. The cubic interaction  $\mu d^{abc} \pi_G^a \pi_G^b \pi_G^c$  is notable as it breaks parity; we set this term to zero. There is also a coupling between the  $G$ -pions and the SM Higgs,  $\lambda_{\pi H} (\pi_G^a \pi_G^a) |H|^2$ , which can lead to indirect limits on  $m_\pi$  through its effect on Higgs production through gluon fusion [69, 70]. In our scenario, the Higgs is not part of the confining gauge sector, and hence  $\lambda_{\pi H}$  is radiatively generated. With  $\lambda_{\pi H} \ll 1$ , the  $\pi_G^a$  may safely have masses in the 100–200 GeV range. The pion number symmetry in the  $G$ -sector is broken by the anomalous coupling among  $\pi_G^a$  and two gluons in eq. (2.8) [71, 72]. Additional higher-dimensional operators can directly couple  $\pi_G$  to SM quarks and mediate  $\pi_G$  decaying into quarks as shown in eq. (2.9).

It is also possible to write additional interactions for the  $\rho_G$ . The renormalizable interaction

$$\mathcal{O}_{\rho\rho g} = \lambda_{\rho\rho g} f^{abc} G_{\mu\nu}^a \rho_G^{b\mu} \rho_G^{c\nu}, \tag{A.3}$$

contributes to  $\rho_G$  pair production. There are two more renormalizable operators containing only the  $\rho_G$  field and we neglect them here. At dimension-6 level, we find the operator

$$\mathcal{O}_{\rho g g} = \frac{i\lambda_{\rho g g}}{4\pi\Lambda_G^2} f^{abc} \rho_{G\nu}^{a\mu} G_\rho^{b\nu} G_\mu^{c\rho}, \tag{A.4}$$

which could be the leading contribution to resonant  $gg \rightarrow \rho_G$  production [67]. As the gluon-gluon luminosity at the LHC is large, this operator can have a noticeable impact on the resonant  $\rho_G$  cross-section despite its high dimension [5]. We conservatively do not include this process when we evaluate  $\sigma(pp \rightarrow \rho_G)$ . Gauge invariance also allows a direct coupling of the  $\rho_G$  to the conserved QCD current, of the form  $\alpha \rho_\mu^a J_\mu^a$ . Through the vector meson dominance calculation, one can absorb this interaction into the kinetic mixing term  $\rho_G^{a\mu\nu} G_{\mu\nu}^a$  in eq. (A.2).

We now comment on the mass ratio  $m_{\pi_G}/m_{\rho_G}$ . The axial SU(3) subgroup of the global chiral flavor symmetry is explicitly broken when the vector subgroup is identified with (gauged) QCD. This ensures that even in the absence of bare masses for  $\Psi$ , QCD interactions will generate a mass  $m_{\pi_G}$  for the  $\pi_G$  octet. The size of the generated  $m_{\pi_G}$  relative to the cutoff, and in particular relative to  $m_{\rho_G}$ , depends on the unknown strong dynamics of  $G$ . Previous studies have used QCD as a model to calculate the mass ratio  $m_{\pi_G}/m_{\rho_G}$ , finding [1]

$$\frac{m_{\pi_G}^2}{m_{\rho_G}^2} = 3 \left( \frac{\alpha_s}{\alpha} \right) \frac{\delta m_\pi^2|_{\text{EM}}}{m_\rho^2} \simeq 0.3, \tag{A.5}$$

based on the observed electromagnetic contribution to the pion mass splitting  $\delta m_\pi^2|_{\text{EM}} \simeq \frac{3\alpha}{4\pi} 2 \ln 2 m_\rho^2$ . Again, this numerical result depends on detailed properties of the QCD spectral functions whose genericity is unclear. A general estimate, not using the simple  $N_c$  counting in QCD, suggests the pNGB mass to scale like

$$m_{\pi_G}^2 \sim \frac{g_s^2}{(4\pi)^2} \Lambda_G^2, \tag{A.6}$$

where  $\Lambda_G$  is the cutoff. For  $\rho_G$  with mass of order the cutoff, we can then estimate

$$\frac{m_{\pi_G}}{m_{\rho_G}} \sim 0.1. \tag{A.7}$$

The above estimation is based on naive dimensional analysis and some order unity numbers can easily modify this relation, which depends on the underlying strong dynamics. Additional explicit sources of chiral symmetry breaking would yield additional contributions to the  $G$ -pion mass. We focus our attention on the regime where  $0.1 \lesssim m_{\pi_G}/m_{\rho_G} \lesssim 0.3$ , where the  $\pi_G$ 's from  $\rho_G$  decay are sufficiently boosted that searches will proceed more profitably with jet substructure techniques.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] C. Kilic, T. Okui and R. Sundrum, *Colored resonances at the Tevatron: phenomenology and discovery potential in multijets*, *JHEP* **07** (2008) 038 [[arXiv:0802.2568](#)] [[INSPIRE](#)].
- [2] C. Kilic, T. Okui and R. Sundrum, *Vectorlike confinement at the LHC*, *JHEP* **02** (2010) 018 [[arXiv:0906.0577](#)] [[INSPIRE](#)].

- [3] Y. Bai and R.J. Hill, *Weakly interacting stable pions*, *Phys. Rev. D* **82** (2010) 111701 [[arXiv:1005.0008](#)] [[INSPIRE](#)].
- [4] J. Kang and M.A. Luty, *Macroscopic strings and ‘quirks’ at colliders*, *JHEP* **11** (2009) 065 [[arXiv:0805.4642](#)] [[INSPIRE](#)].
- [5] C. Kilic, S. Schumann and M. Son, *Searching for multijet resonances at the LHC*, *JHEP* **04** (2009) 128 [[arXiv:0810.5542](#)] [[INSPIRE](#)].
- [6] J.M. Butterworth, B.E. Cox and J.R. Forshaw, *WW scattering at the CERN LHC*, *Phys. Rev. D* **65** (2002) 096014 [[hep-ph/0201098](#)] [[INSPIRE](#)].
- [7] J.M. Butterworth, A.R. Davison, M. Rubin and G.P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001 [[arXiv:0802.2470](#)] [[INSPIRE](#)].
- [8] J. Thaler and L.-T. Wang, *Strategies to identify boosted tops*, *JHEP* **07** (2008) 092 [[arXiv:0806.0023](#)] [[INSPIRE](#)].
- [9] D.E. Kaplan, K. Rehermann, M.D. Schwartz and B. Tweedie, *Top tagging: a method for identifying boosted hadronically decaying top quarks*, *Phys. Rev. Lett.* **101** (2008) 142001 [[arXiv:0806.0848](#)] [[INSPIRE](#)].
- [10] T. Plehn, G.P. Salam and M. Spannowsky, *Fat jets for a light Higgs*, *Phys. Rev. Lett.* **104** (2010) 111801 [[arXiv:0910.5472](#)] [[INSPIRE](#)].
- [11] C. Hackstein and M. Spannowsky, *Boosting Higgs discovery: the forgotten channel*, *Phys. Rev. D* **82** (2010) 113012 [[arXiv:1008.2202](#)] [[INSPIRE](#)].
- [12] Y. Cui, Z. Han and M.D. Schwartz, *W-jet tagging: optimizing the identification of boosted hadronically-decaying W bosons*, *Phys. Rev. D* **83** (2011) 074023 [[arXiv:1012.2077](#)] [[INSPIRE](#)].
- [13] J.L. Hewett, J. Shelton, M. Spannowsky, T.M.P. Tait and M. Takeuchi,  *$A_{FB}^t$  meets LHC*, *Phys. Rev. D* **84** (2011) 054005 [[arXiv:1103.4618](#)] [[INSPIRE](#)].
- [14] W. Skiba and D. Tucker-Smith, *Using jet mass to discover vector quarks at the LHC*, *Phys. Rev. D* **75** (2007) 115010 [[hep-ph/0701247](#)] [[INSPIRE](#)].
- [15] B. Holdom,  *$t'$  at the LHC: the physics of discovery*, *JHEP* **03** (2007) 063 [[hep-ph/0702037](#)] [[INSPIRE](#)].
- [16] J.M. Butterworth, J.R. Ellis and A.R. Raklev, *Reconstructing sparticle mass spectra using hadronic decays*, *JHEP* **05** (2007) 033 [[hep-ph/0702150](#)] [[INSPIRE](#)].
- [17] J.M. Butterworth, J.R. Ellis, A.R. Raklev and G.P. Salam, *Discovering baryon-number violating neutralino decays at the LHC*, *Phys. Rev. Lett.* **103** (2009) 241803 [[arXiv:0906.0728](#)] [[INSPIRE](#)].
- [18] G.D. Kribs, A. Martin, T.S. Roy and M. Spannowsky, *Discovering the Higgs boson in new physics events using jet substructure*, *Phys. Rev. D* **81** (2010) 111501 [[arXiv:0912.4731](#)] [[INSPIRE](#)].
- [19] C.-R. Chen, M.M. Nojiri and W. Sreethawong, *Search for the elusive Higgs boson using jet structure at LHC*, *JHEP* **11** (2010) 012 [[arXiv:1006.1151](#)] [[INSPIRE](#)].
- [20] G.D. Kribs, A. Martin, T.S. Roy and M. Spannowsky, *Discovering Higgs bosons of the MSSM using jet substructure*, *Phys. Rev. D* **82** (2010) 095012 [[arXiv:1006.1656](#)] [[INSPIRE](#)].

- [21] A. Falkowski, D. Krohn, L.-T. Wang, J. Shelton and A. Thalapillil, *Unburied Higgs boson: jet substructure techniques for searching for Higgs' decay into gluons*, *Phys. Rev. D* **84** (2011) 074022 [[arXiv:1006.1650](#)] [[INSPIRE](#)].
- [22] T. Plehn, M. Spannowsky, M. Takeuchi and D. Zerwas, *Stop reconstruction with tagged tops*, *JHEP* **10** (2010) 078 [[arXiv:1006.2833](#)] [[INSPIRE](#)].
- [23] B. Bhattacharjee, M. Guchait, S. Raychaudhuri and K. Sridhar, *Boosted top quark signals for heavy vector boson excitations in a universal extra dimension model*, *Phys. Rev. D* **82** (2010) 055006 [[arXiv:1006.3213](#)] [[INSPIRE](#)].
- [24] A. Katz, M. Son and B. Tweedie, *Jet substructure and the search for neutral spin-one resonances in electroweak boson channels*, *JHEP* **03** (2011) 011 [[arXiv:1010.5253](#)] [[INSPIRE](#)].
- [25] A. Katz, M. Son and B. Tweedie, *Ditau-jet tagging and boosted Higgses from a multi-TeV resonance*, *Phys. Rev. D* **83** (2011) 114033 [[arXiv:1011.4523](#)] [[INSPIRE](#)].
- [26] B. Bellazzini, C. Csáki, J. Hubisz and J. Shao, *Discovering a Higgs boson decaying to four jets in supersymmetric cascade decays*, *Phys. Rev. D* **83** (2011) 095018 [[arXiv:1012.1316](#)] [[INSPIRE](#)].
- [27] G.D. Kribs, A. Martin and T.S. Roy, *Higgs boson discovery through top-partners decays using jet substructure*, *Phys. Rev. D* **84** (2011) 095024 [[arXiv:1012.2866](#)] [[INSPIRE](#)].
- [28] P. Bandyopadhyay and B. Bhattacharjee, *Boosted top quarks in supersymmetric cascade decays at the LHC*, *Phys. Rev. D* **84** (2011) 035020 [[arXiv:1012.5289](#)] [[INSPIRE](#)].
- [29] J. Fan, D. Krohn, P. Mosteiro, A.M. Thalapillil and L.-T. Wang, *Heavy squarks at the LHC*, *JHEP* **03** (2011) 077 [[arXiv:1102.0302](#)] [[INSPIRE](#)].
- [30] D.E. Kaplan and M. McEvoy, *Associated production of non-standard Higgs bosons at the LHC*, *Phys. Rev. D* **83** (2011) 115004 [[arXiv:1102.0704](#)] [[INSPIRE](#)].
- [31] T. Plehn, M. Spannowsky and M. Takeuchi, *Boosted semileptonic tops in stop decays*, *JHEP* **05** (2011) 135 [[arXiv:1102.0557](#)] [[INSPIRE](#)].
- [32] C. Englert, T.S. Roy and M. Spannowsky, *Ditau jets in Higgs searches*, *Phys. Rev. D* **84** (2011) 075026 [[arXiv:1106.4545](#)] [[INSPIRE](#)].
- [33] C.T. Hill and E.H. Simmons, *Strong dynamics and electroweak symmetry breaking*, *Phys. Rept.* **381** (2003) 235 [*Erratum ibid.* **390** (2004) 553] [[hep-ph/0203079](#)] [[INSPIRE](#)].
- [34] B.A. Dobrescu, K. Kong and R. Mahbubani, *Massive color-octet bosons and pairs of resonances at hadron colliders*, *Phys. Lett. B* **670** (2008) 119 [[arXiv:0709.2378](#)] [[INSPIRE](#)].
- [35] Y. Bai and B.A. Dobrescu, *Heavy octets and Tevatron signals with three or four b jets*, *JHEP* **07** (2011) 100 [[arXiv:1012.5814](#)] [[INSPIRE](#)].
- [36] J.C. Pati and A. Salam, *Are the new particles color gluons?*, *Phys. Rev. Lett.* **34** (1975) 613 [[INSPIRE](#)].
- [37] L.J. Hall and A.E. Nelson, *Heavy gluons and monojets*, *Phys. Lett. B* **153** (1985) 430 [[INSPIRE](#)].
- [38] P.H. Frampton and S.L. Glashow, *Chiral color: an alternative to the standard model*, *Phys. Lett. B* **190** (1987) 157 [[INSPIRE](#)].
- [39] LHC NEW PHYSICS WORKING GROUP collaboration, D. Alves et al., *Simplified models for LHC new physics searches*, [arXiv:1105.2838](#) [[INSPIRE](#)].

- [40] C. Kilic and T. Okui, *The LHC phenomenology of vectorlike confinement*, *JHEP* **04** (2010) 128 [[arXiv:1001.4526](#)] [[INSPIRE](#)].
- [41] J. Sayre, D.A. Dicus, C. Kao and S. Nandi, *Searching for colorons at the Large Hadron Collider*, *Phys. Rev. D* **84** (2011) 015011 [[arXiv:1105.3219](#)] [[INSPIRE](#)].
- [42] B.W. Lee and H.T. Nieh, *Phenomenological Lagrangian for field algebra, hard pions and radiative corrections*, *Phys. Rev.* **166** (1968) 1507 [[INSPIRE](#)].
- [43] I.S. Gerstein, B.W. Lee, H.T. Nieh and H.J. Schnitzer, *Pion electromagnetic mass difference for physical pions*, *Phys. Rev. Lett.* **19** (1967) 1064 [[INSPIRE](#)].
- [44] H. Harari, *Superconvergent dispersion relations and electromagnetic mass differences*, *Phys. Rev. Lett.* **17** (1966) 1303 [[INSPIRE](#)].
- [45] P. Langacker and H. Pagels, *Pion and kaon electromagnetic masses in chiral perturbation theory*, *Phys. Rev. D* **8** (1973) 4620 [[INSPIRE](#)].
- [46] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J. C* **63** (2009) 189 [[arXiv:0901.0002](#)] [[INSPIRE](#)].
- [47] ATLAS collaboration, *A search for  $t\bar{t}$  resonances in the lepton plus jets channel in 200 pb<sup>-1</sup> of pp collisions at  $\sqrt{s} = 7$  TeV*, *ATLAS-CONF-2011-087* (2011).
- [48] ATLAS collaboration, G. Aad et al., *Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector*, *Eur. Phys. J. C* **71** (2011) 1512 [[arXiv:1009.5908](#)] [[INSPIRE](#)].
- [49] CMS collaboration, *Measurement of the inclusive jet cross section in pp collisions at 7 TeV*, *CMS-PAS-QCD-10-011* (2010).
- [50] ATLAS collaboration, G. Aad et al., *Search for new physics in dijet mass and angular distributions in pp collisions at  $\sqrt{s} = 7$  TeV measured with the ATLAS detector*, *New J. Phys.* **13** (2011) 053044 [[arXiv:1103.3864](#)] [[INSPIRE](#)].
- [51] ATLAS collaboration, *Update of the search for new physics in the dijet mass distribution in 163 pb<sup>-1</sup> of pp collisions at  $\sqrt{s} = 7$  TeV measured with the ATLAS detector*, *ATLAS-CONF-2011-081* (2011).
- [52] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, *Phys. Rev. D* **80** (2009) 051501 [[arXiv:0903.5081](#)] [[INSPIRE](#)].
- [53] D. Krohn, J. Thaler and L.-T. Wang, *Jet trimming*, *JHEP* **02** (2010) 084 [[arXiv:0912.1342](#)] [[INSPIRE](#)].
- [54] J. Thaler and K. Van Tilburg, *Identifying boosted objects with N-subjettiness*, *JHEP* **03** (2011) 015 [[arXiv:1011.2268](#)] [[INSPIRE](#)].
- [55] J. Alwall et al., *MadGraph/MadEvent v4: the new web generation*, *JHEP* **09** (2007) 028 [[arXiv:0706.2334](#)] [[INSPIRE](#)].
- [56] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, *JHEP* **07** (2002) 012 [[hep-ph/0201195](#)] [[INSPIRE](#)].
- [57] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026 [[hep-ph/0603175](#)] [[INSPIRE](#)].
- [58] M. Cacciari, G.P. Salam and G. Soyez, *FastJet*, <http://fastjet.fr/>.

- [59] M. Cacciari and G.P. Salam, *Dispelling the  $N^3$  myth for the  $k_t$  jet-finder*, *Phys. Lett. B* **641** (2006) 57 [[hep-ph/0512210](#)] [[INSPIRE](#)].
- [60] M. Cacciari, G.P. Salam and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, *JHEP* **04** (2008) 063 [[arXiv:0802.1189](#)] [[INSPIRE](#)].
- [61] ATLAS collaboration, *Prospects for top anti-top resonance searches using early ATLAS data*, [ATL-PHYS-PUB-2010-008](#) (2010).
- [62] ATLAS collaboration, G. Aad et al., *Search for new particles in two-jet final states in 7 TeV proton-proton collisions with the ATLAS detector at the LHC*, *Phys. Rev. Lett.* **105** (2010) 161801 [[arXiv:1008.2461](#)] [[INSPIRE](#)].
- [63] ATLAS collaboration, *Measurement of jet mass and substructure for inclusive jets in  $\sqrt{s} = 7$  TeV  $pp$  collisions with the ATLAS experiment*, [ATLAS-CONF-2011-073](#) (2011).
- [64] ATLAS collaboration, *ATLAS sensitivity to the standard model Higgs in the  $HW$  and  $HZ$  channels at high transverse momenta*, [ATL-PHYS-PUB-2009-088](#) (2009).
- [65] T. Plehn and T.M.P. Tait, *Seeking sgluons*, *J. Phys. G* **36** (2009) 075001 [[arXiv:0810.3919](#)] [[INSPIRE](#)].
- [66] C. Englert, C. Hackstein and M. Spannowsky, *Measuring spin and CP from semi-hadronic  $ZZ$  decays using jet substructure*, *Phys. Rev. D* **82** (2010) 114024 [[arXiv:1010.0676](#)] [[INSPIRE](#)].
- [67] R.S. Chivukula, A. Grant and E.H. Simmons, *Two gluon coupling and collider phenomenology of color octet technirho mesons*, *Phys. Lett. B* **521** (2001) 239 [[hep-ph/0109029](#)] [[INSPIRE](#)].
- [68] A.R. Zerwekh, C.O. Dib and R. Rosenfeld, *A new signature for color octet pseudoscalars at the CERN LHC*, *Phys. Rev. D* **77** (2008) 097703 [[arXiv:0802.4303](#)] [[INSPIRE](#)].
- [69] R. Boughezal and F. Petriello, *Color-octet scalar effects on Higgs boson production in gluon fusion*, *Phys. Rev. D* **81** (2010) 114033 [[arXiv:1003.2046](#)] [[INSPIRE](#)].
- [70] R. Boughezal, *Constraints on heavy colored scalars from Tevatron's Higgs exclusion limit*, *Phys. Rev. D* **83** (2011) 093003 [[arXiv:1101.3769](#)] [[INSPIRE](#)].
- [71] Y. Bai and A. Martin, *Topological pions*, *Phys. Lett. B* **693** (2010) 292 [[arXiv:1003.3006](#)] [[INSPIRE](#)].
- [72] A. Freitas and P. Schwaller, *Multi-photon signals from composite models at LHC*, *JHEP* **01** (2011) 022 [[arXiv:1010.2528](#)] [[INSPIRE](#)].