# Binary geometries, generalized particles and strings, and cluster algebras

Nima Arkani-Hamed,<sup>1</sup> Song He,<sup>2</sup> Thomas Lam<sup>®</sup>,<sup>3</sup> and Hugh Thomas<sup>9</sup>

<sup>1</sup>School of Natural Sciences, Institute for Advanced Studies, Princeton, New Jersey 08540, USA and Center of Mathematical Sciences and Applications, Harvard University,

Cambridge, Massachusetts 02138, USA

<sup>2</sup>CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics,

Chinese Academy of Sciences, Beijing 100190, China; School of Fundamental Physics and Mathematical

Sciences, Hangzhou Institute for Advanced Study and ICTP-AP International Centre for Theoretical

Physics Asia-Pacific, UCAS, Hangzhou 310024, China

<sup>3</sup>Department of Mathematics, University of Michigan, 530 Church St, Ann Arbor, Michigan 48109, USA

and Department of Mathematics, Massachusetts Institute of Technology,

77 Massachusetts Ave., Cambridge, Massachusetts 02139, USA

<sup>4</sup>LaCIM, Département de Mathématiques, Université du Québec à Montréal, Montréal, Quebec H3C 3P8, Canada

(Received 7 March 2020; accepted 16 February 2023; published 17 March 2023)

We introduce the notion of "binary" positive and complex geometries, giving a completely rigid geometric realization of the combinatorics of generalized associahedra attached to any Dynkin diagram. We also define open and closed "cluster string integrals" associated with these "cluster configuration spaces". The binary geometry of type  $\mathcal{A}$  gives a gauge-invariant description of the usual open and closed string moduli spaces for tree scattering, making no explicit reference to a world sheet. The binary geometries and cluster string integrals for other Dynkin types provide a generalization of particle and string scattering amplitudes. Both the binary geometries and cluster string integrals enjoy remarkable factorization properties at finite  $\alpha'$ , obtained simply by removing nodes of the Dynkin diagram. As  $\alpha' \rightarrow 0$  these cluster string integrals reduce to the canonical forms of the Arkani-Hamed-Bai-He-Yan (ABHY) generalized associahedron polytopes. For classical Dynkin types these are associated with *n*-particle scattering in the biadjoint  $\phi^3$  theory through one-loop order.

DOI: 10.1103/PhysRevD.107.066015

### **I. BINARY GEOMETRIES**

A remarkable fact about particle scattering in arbitrary spacetime dimension is that the poset of planar singularities form combinatorial polytopes. The associahedron polytope [1,2] encodes combinatorially all planar cubic tree graphs, which has been realized directly in the kinematic space and is known as the ABHY *kinematic associahedron* [3]. Its canonical form [4] computes the tree-level *S*-matrix of biadjoint  $\phi^3$  theory, which makes hidden symmetries of the amplitudes manifest. In [5] the ABHY realization was extended to *generalized associahedra of any finite type cluster algebra* [6,7], whose canonical forms compute biadjoint  $\phi^3$  amplitudes through one loop [8]; the canonical form of the type- $\mathcal{B}$  or type- $\mathcal{C}$  polytope (known as the cyclohedron) contains one-loop tadpole diagrams,

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 SCOAP<sup>3</sup>. while for the type-D polytope it gives the integrand for oneloop biadjoint  $\phi^3$  amplitudes.

For a finite type, rank-*n* cluster algebra, these are *n*-dimensional polytopes whose facets (respectively, vertices) are in one-to-one correspondence with cluster variables (resp., cluster seeds). The boundary structure of such a polytope has interesting factorizations encoded in the corresponding Dynkin diagram, where each facet corresponds to removing a node of the Dynkin diagram. For the associahedron of type  $\mathcal{A}_n$ , the facets factorize as  $\mathcal{A}_m \times \mathcal{A}_{n-1-m}$ ; the type- $\mathcal{B}_n/\mathcal{C}_n$  associahedron has facets of the shape  $\mathcal{B}_m \times \mathcal{A}_{n-1-m}$  or  $\mathcal{C}_m \times \mathcal{A}_{n-1-m}$ ; the type- $\mathcal{D}_n$  associahedron has facets of the shape  $\mathcal{D}_m \times \mathcal{A}_{n-1-m}$ ,  $\mathcal{A}_1 \times \mathcal{A}_1 \times \mathcal{A}_{n-3}$  or simply  $\mathcal{A}_{n-1}$ . Examples of these factorizations are shown below.

It is fascinating that cluster polytopes provide a geometric realization of the combinatorics of "compatibility" for cluster variables. Each cluster variable is associated with a facet of the polytope; compatible variables correspond to facets in the polytope that meet, while incompatible variables correspond to facets that do not touch at all. But there is still freedom in the particular realization of the polytope, and any realization can be continuously deformed in various ways without changing the relevant combinatorics. It is therefore natural to wonder whether any geometric realization of the combinatorics exists that is more rigid and canonical.

This is what we seek to do in this letter. We introduce the idea of binary geometries as a canonical and rigid way of realizing the combinatorics of generalized associahedra for any finite-type cluster algebra. For each facet or cluster variable denoted as a, we assign a  $u_a$  variable, and impose the same number of constraints of the form

$$u_a + \prod_{\text{all } b} u_b^{b||a} = 1, \quad \forall a \tag{1}$$

where the non-negative integer b||a| is called the *compatibility degree* from b to a (originally defined in [6]). It is zero if and only if a and b are compatible, or equivalently if the two facets meet. It is highly nontrivial that for n-dimensional generalized associahedra, these u equations are consistent and the solution space turns out to be *n*-dimensional. Demanding  $u_a \ge 0$ , these equations also force all the  $u_a$  to take values in [0, 1]. Thus when a facet a is reached with  $u_a \rightarrow 0$ , the  $u_b$  variables for all *incompatible* facets (b||a > 0) are forced to go to 1. This is why we call this a "binary" realization of generalized associahedra. Restricting the *u* variables to be positive, the solution space gives a "curvy" realization of the polytope. It is striking that the "binary" property is not restricted to real and positive  $u_a$ , but is a feature of the complex space of solutions of Eq. (1). When  $u_a \rightarrow 0$ , the second term in the equation  $u_b + \prod_{\text{all } c} u_c^{c||b} = 1$  becomes 0 for all incompatible  $u_b$ 's, and so those  $u_b$  are all equal to 1. This is a novel feature of the binary geometry: almost all the connections between geometries with factorizing boundary structures and physics, have involved reality and positivity in a crucial way, whereas here the factorization holds with complex variables.

For type  $\mathcal{A}$ , the binary realization (with b||a = 0, 1 only) turns out to give a gauge-invariant description of the moduli space of open- and closed-string world sheet at genus zero [9] (with u variables being cross-ratios [10]). The binary positive and complex associahedron geometry directly underpin open- and closed-string amplitudes; we will see that both the Parke-Taylor form and Koba-Nielsen factor [11] in string integrals are naturally written in terms of the u variables. String amplitudes are the stringy canonical form, or  $\alpha'$  extension of the canonical form, for ABHY associahedra [12]; not only does the  $\alpha' \rightarrow 0$  limit

reproduce the  $\phi^3$ -tree amplitude, but it factorizes into products of lower amplitudes at finite  $\alpha'$ .

We will give a concise definition of the compatibility degrees from the ABHY realization [8], and it is remarkable that the *u* equations encode the boundary structures of generalized associahedra from (1). Naturally associated with them are the generalized open- and closed-string amplitudes that we call cluster string integrals [12]; the open cluster string integral is the integral over the binary geometry of its canonical form, and is beautifully regulated by including all the factors of the form  $u^{\alpha' X}$ 

$$\mathcal{I}(\{X\}) = \int_{U^+} \Omega^{(n)}(U^+) \prod_a u_a^{\alpha' X_a}.$$
 (2)

These are the most perfect examples of stringy canonical forms with leading order given by the ABHY polytope [12], and their factorizations at finite  $\alpha'$  again correspond to removing a node of the Dynkin diagram [13]. For example, the  $\alpha' \rightarrow 0$  limit of type  $\mathcal{D}$  integral gives one-loop planar  $\phi^3$  amplitude, and it factorizes as product of lower-point integrals of type  $\mathcal{D}$  and  $\mathcal{A}$ .

Our aim in this paper is to give a self-contained description of the u equations, and to summarize and highlight only a few of the important features of the binary geometry and associated stringy canonical forms. We will present a much more detailed discussion of these in a longer companion article [14], where we will give a cluster algebraic explanation of u variables and u equations and construct these cluster configuration spaces as certain quotients of cluster varieties showing that the space is smooth with normal-crossing boundary divisors.

# II. BINARY ASSOCIAHEDRA AND STRING AMPLITUDES

## A. The U space and the moduli space

We begin by discussing the *u*-variables in the most familiar setting of the usual associahedron. We introduce  $N := \frac{n(n-3)}{2} u_{i,j}$  variables, and consider N u equations [3],

$$1 - u_{i,j} = \prod_{(k,l) \operatorname{cross}(i,j)} u_{k,l},\tag{3}$$

where the product is over all chords (k, l) that cross (i, j), i.e., all *u*'s that are *incompatible* with  $u_{i,j}$ . Equation (3) is the type- $\mathcal{A}$  version of (1) where the only nonzero degree is 1. For example for n = 4 we have  $u_{1,3} + u_{2,4} = 1$  and for n = 5, we have  $u_{1,3} + u_{2,4}u_{2,5} = 1$  and its cyclic images.

Let us denote the solution space of (3) as the U space, and a nontrivial observation is that n-3 of the equations are redundant, and thus  $U_n$  has dimension n-3. The U space has the boundary structure of an associahedron purely algebraically; at a boundary  $u_{i,j} \rightarrow 0$ , we see that for any incompatible (k, l), the rhs of  $1 - u_{k,l}$  vanishes, thus  $u_{k,l} \rightarrow 1$  and all incompatible *u*'s decouple. We are left with *U* spaces for two polygons divided by (i, j), thus the boundary is given by their product,

$$\partial_{u_{i,j} \to 0} U_n = U(i, i+1, ..., j) \times U(j, j+1, ..., i).$$
 (4)

We see that *u* equations encode the boundary structure of  $\mathcal{A}_{n-3}$  even for  $u \in \mathbb{C}$ . For example, as  $u_{1,3} \to 0$ ,  $u_{2,4}, u_{2,5} \to 1$ , which decouple from *u* equations, and we are left with  $u_{1,4} + u_{3,5} = 1$  which defines the n = 4 space U(3, 4, 5, 1) [times the trivial U(1, 2, 3) space].

We can include positivity. By requiring all  $u_{i,j} \ge 0$ , (3) implies all of them satisfy  $0 \le u_{i,j} \le 1$ , which cut out a "curvy" associahedron with *N* facets; we will refer to it as the positive part of *U* space,  $U_n^+$ , which has the same shape as  $\mathcal{A}_{n-3}$ . For n = 5, the positive part  $U_5^+ \sim \mathcal{A}_2$  gives a "curvy" pentagon [3].

As mentioned above, the *U* space turns out to provide a SL(2)-invariant description of the moduli space (or configuration space),  $\mathcal{M}_{0,n}$ , of *n* points on  $\mathbb{P}^1$ , where the *u*'s can be interpreted as *cross-ratios* of such *n* points. Let us see how the moduli space naturally emerges from (3). Consider a pair of disjoint sets of points, *A*, *B* and we can write  $A = \{a + 1, a + 2, ..., b\}$  and  $B = \{c + 1, c + 2, ..., d\}$  where *a*, *b*, *c*, *d* are cyclically ordered. We define  $U_{A,B} = \prod_{i \in A, j \in B} u_{i,j}$  and note the trivial identity  $U_{A,B_1}U_{A,B_2} = U_{A,B}$  for two disjoint sets  $B_1 \cup B_2 = B$ . Furthermore, we define complementary sets  $\overline{A}$ ,  $\overline{B}$  as  $\overline{A} = \{b + 1, ..., c\}$  and  $\overline{B} = \{d + 1, ..., a\}$ . From (3) we can deduce that

$$U_{A,B} + U_{\bar{A},\bar{B}} = 1.$$
 (5)

Indeed, (3) is a special case of (5); choosing  $A = \{i\}$ ,  $B = \{j\}$ , the ranges for k and l in (3) are  $\{\overline{i}\}, \{\overline{j}\}$ , and (3) becomes  $U_{\{i\},\{j\}} + U_{\{\overline{i}\},\{\overline{j}\}} = 1$ . The *extended* u equations (5) nicely lead to the identification of U's as cross-ratios. Denoting  $U_{A,B}$  as [a, b|c, d], we have  $U_{\overline{A},\overline{B}} = [b, c|d, a]$ , and thus we see the appearance of four points and the identities [a, b|c, d] must satisfy. Trivially, we have [a, b|c, e][a, b|e, d] = [a, b|c, d] and more interestingly by (5) we have [a, b|c, d] + [b, c|d, a] = 1. They are precisely identities that invariantly characterize cross-ratios for n points on  $\mathbb{P}^1$ . The solutions are

$$[a,b|c,d] = \frac{(ad)(bc)}{(ac)(bd)},\tag{6}$$

where (ab) denotes a minor of G(2, n) that represents n points on  $\mathbb{P}^1$ . If we further require the n points to be ordered on  $\mathbb{RP}^1$  [equivalently they can be represented by a point in  $G_+(2, n)$ ], then all the cross-ratios are between 0 and 1; 0 < [a, b|c, d] < 1. This is how the open-string moduli space,  $\mathcal{M}_{0,n}^+$ , emerges from u equations.

We have just seen that the  $u_{i,i}$  allow us to describe the open-string moduli space without thinking about *n* ordered points on the boundary of a disk. But this standard picture allows us to see something else as well-the presence of different orderings of the n points, which relate to different color orderings for scattering amplitudes. In contrast, the  $u_{i,j}$  variables appear inexorably linked to a single ordering. So how can we see the other orderings from this point of view? This question has a natural answer. As we have discussed, it is when we restrict all the  $u_{i,i} \ge 0$ , that the u equations force them to lie in the unit interval,  $0 \le u_{i,j} \le 1$ . It is natural to ask whether some of the  $u_{i,i}$ 's might be negative. Without any detailed study of the solutions of the extended *u* equations, we can just ask which sign patterns for the u's are allowed to be compatible with  $U_{A,B} + U_{\bar{A},\bar{B}} = 1$ , i.e., excluding only those sign patterns for which both terms are negative. Quite beautifully, we find that precisely (n-1)!/2 such sign patterns are allowed. For any consistent sign pattern, it is natural define new positive variables  $\hat{u}$  in the obvious way by writing  $\hat{u}_{i,i} = (-1)^{\text{sgn}(u_{i,j})} u_{i,i}$ . We can then rearrange the extended u equations to be in the form of setting the sum of two monomials in the  $\hat{u}$  variables to unity. Remarkably, the set of all the  $\hat{U}$  equations we get in this way is just a a reordering of the extended u equations we started with. This exposes a hidden  $S_n$  symmetry of the  $U_n$  space which we can recognize as permutations of the *n* points once we identify the space with  $\mathcal{M}_{0,n}$ . Taking the  $u_{i,j}$  to be real, we see that  $U_n(\mathbb{R}) \sim \mathcal{M}_{0,n}(\mathbb{R})$  is tiled by (n-1)!/2 connected *components*, and each of them is an associahedron  $U_n^+ \sim$  $\mathcal{A}_{n-3}$  for a cyclic ordering. For example, it is easy to check that there are 12 such consistent sign patterns for n = 5 [for 12 pentagons of  $\mathcal{M}_{0.5}(\mathbb{R})$ ]; similarly there are 60 sign patterns for n = 6, which gives 60  $A_3$  associahedra that tile  $\mathcal{M}_{0,6}(\mathbb{R})$ .

#### **B.** Open string integrals on the U space

It is very natural to define integrals on the U space, which turn out to be usual string integrals with Koba-Nielsen factor. For the open-string case, we are interested in integrating the *canonical form* for the positive part, which can be obtained by a (trivial) pushforward [4]. Given an acyclic quiver of  $A_{n-3}$ , or equivalently a triangulation of an *n*-gon without any internal triangle, it turns out by (3) that one can solve for all the *N u* variables *rationally* in terms of the n-3  $u_{\alpha}$ 's in that cluster seed. This provides a one-to-one map from the space  $\{0 < u_{\alpha} < 1\}$  to  $U_n^+$ , and thus the canonical form can be obtained by pushforward [4]

$$\Omega(U_n^+) = \prod_{\alpha}^{n-3} d \log \frac{u_{\alpha}}{1 - u_{\alpha}},\tag{7}$$

and any acyclic quiver gives the same result. Consider integrating it over  $U_n^+$ , which of course diverges at

boundaries; a natural way to regularize the integral is by putting a factor  $u^{\alpha' X}$  with X > 0 for each boundary,

$$\mathcal{I}_{n}^{U^{+}}(\{X\}) \coloneqq (\alpha')^{n-3} \int_{U_{n}^{+}} \Omega(U_{n}^{+}) \prod_{i,j} u_{i,j}^{\alpha' X_{i,j}}.$$
 (8)

The form  $\Omega(U_n^+)$  is the famous Parke-Taylor form for  $\mathcal{M}_{0,n}^+$ [or  $G_+(2,n)$  mod torus action],  $\Omega(U_n^+) = \frac{d^n z/\mathrm{SL}(2)}{(12)\cdots(n1)}$ , where z's denote the *n* punctures of  $\mathcal{M}_{0,n}^+$  [or inhomogeneous coordinates of  $G_+(2,n)/T$ ]. As shown in [12],  $\mathcal{I}_n^{U^+}$  is nothing but the open-string integral where the regulator is exactly the Koba-Nielsen factor,  $\prod_{i,j} u_{i,j}^{\alpha' X_{i,j}} = \prod_{a,b} (ab)^{\alpha' s_{a,b}}$ with  $\binom{n}{2}$  Mandelstam variables given by  $s_{a,b} = X_{a,b} + X_{a+1,b+1} - X_{a,b+1} - X_{a+1,b}$ ; since we mod out the torus action, there are *n* constraints (momentum conservation)  $\sum_{b\neq a} s_{a,b} = 0$ , and thus only n(n-3)/2 of the  $s_{a,b}$  are linearly independent.

The leading order of  $\mathcal{I}_n^{U^+}$  is given by the biadjoint  $\phi^3$  amplitude  $m_n(\{X\})$ , and the same result can be obtained by summing over saddle points [3], known as Cachazo-He-Yuan (CHY) formulas [15,16]. This connection between  $\alpha' \to 0$  limit and "scattering equations" in the  $\alpha' \to \infty$  limit has been understood as a general phenomenon for any stringy canonical form [12] (see also [17]).

What is special about  $\mathcal{I}_n^{U^+}$  is that it factorizes perfectly even at *finite*  $\alpha'$ , which becomes manifest in the form of (8). It is easy to see that  $\mathcal{I}_n^{U^+}$  has a simple pole at each  $X_{i,j} = 0$ , and the residue is given by the integral at the boundary at  $u_{i,j} \to 0$ ; at this boundary we have  $\partial_{u_{i,j}\to 0}\Omega(U^+) =$  $\Omega(\partial_{u_{i,j}\to 0}U^+)$  by definition, which factorizes into two lower forms, since the boundary factorizes as (4). Remarkably, precisely due to (3), the Koba-Nielsen factor factorizes accordingly, and we have

$$\operatorname{Res}_{X_{i,j}=0}\mathcal{I}_n^{U^+} = \int_{U_L^+ \times U_R^+} (\Omega_L \times \Omega_R) \prod_L (u^{\alpha' X}) \times \prod_R (u^{\alpha' X}),$$

which is nothing but  $\mathcal{I}_L \times \mathcal{I}_R$ , where *L* and *R* denote the two polygons divided by the chord (ij) in the *n*-gon. Let us give an example for n = 5;  $\mathcal{I}_5^{U^+}$  reads

$$\int_0^1 d\log \frac{u_{1,3}}{1-u_{1,3}} d\log \frac{u_{1,4}}{1-u_{1,4}} u_{1,3}^{X_{1,3}} u_{1,4}^{X_{1,4}} u_{2,4}^{X_{2,4}} u_{2,5}^{X_{2,5}} u_{3,5}^{X_{3,5}},$$

whose leading order is the canonical function of  $\mathcal{A}_2$  cut out by 5 facets,  $X_{i,j} > 0$ , with ABHY conditions  $c_{1,3} = X_{1,3} + X_{2,4} - X_{1,4}$ ,  $c_{1,4} = X_{1,4} + X_{2,5} - X_{2,4}$ ,  $c_{2,4} = X_{2,4} + X_{3,5} - X_{2,5}$ . At finite  $\alpha'$ , e.g., as  $X_{13} \rightarrow 0$ , the residue is given by the Veneziano amplitude (times  $\mathcal{I}_3 = 1$ ),  $\mathcal{I}_4^{U^+} = \int_0^1 d \log \frac{u}{1-u} u^s (1-u)^t$  with  $s = X_{14}$ ,  $t = X_{35}$ .

It is also natural to consider closed-string integrals for a pair of orderings,  $\alpha$  and  $\beta$ , in complex U space,

$$\mathcal{I}_{n}^{c}(\alpha|\beta) \coloneqq \int_{U(\mathbb{C})} \Omega(U_{\alpha}^{+}) \prod u^{\alpha' X} \left( \Omega(U_{\beta}^{+}) \prod u^{\alpha' X} \right)^{*},$$

where from the monomial transformation of u's, we see that  $\Omega(U_{\alpha}^{+})$  by (7) is the Parke-Taylor form of ordering  $\alpha$ , and the Koba-Nielsen factor is permutation invariant, thus the integral can be written as  $\int_{U(\mathbb{C})} \Omega(\alpha) \Omega^{*}(\beta) |\prod_{a,b} (ab)^{\alpha' s_{a,b}}|^2$ . Similarly, we can integrate  $\Omega(U_{\alpha}^{+})$  in a different region  $U^{+}(\beta)$ . The leading order of both integrals are general biadjoint  $\phi^3$  amplitudes,  $m_n(\alpha|\beta)$ , given by (the canonical function of) ABHY associahedra with certain facets sent to infinity [12].

We end with some counting regarding  $\mathcal{M}_{0,n}$ . It is wellknown that out of all (n-1)!/2 ordering  $\Omega(\alpha)$ 's, only (n-2)! are linearly independent, which is the number of independent top-dimensional *d* log forms, or Parke-Taylor forms [18]. Moreover, in the presence of the Koba-Nielsen factor, the dimension of the (n-3)th [half the dimension of  $\mathcal{M}_{0,n}(\mathbb{C})$ ] twisted (co)homology group is (n-3)!, which is the number of saddle points [12] (see also [17,19]), thus the matrix  $\mathcal{I}_n^c(\alpha|\beta)$  and  $m_n(\alpha|\beta)$  has rank (n-3)! [20–23]. It is known that the Euler characteristic of  $\mathcal{M}_{0,n}(\mathbb{C})$  is given by  $(-1)^{n-3}(n-3)!$  [24,25].

## **III. CLUSTER CONFIGURATION SPACES**

Now we explain how a binary geometry, called the cluster configuration space, can be constructed for generalized associahedra of finite-type cluster algebras. The only new ingredient in the *u* equations, (1), is that the compatibility degree b||a can take values other than 0, 1. They form a  $N \times N$  matrix for a finite-type cluster algebra with N cluster variables. Let us give a concise definition of the compatibility degree without referring to most of the machinery of cluster algebra, based on a picture of "walk through all cluster variables" which underpins the ABHY realizations of generalized associahedra [5,8].

#### A. General *u* equations from the walk

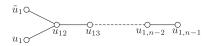
To define a walk, we start with any *acyclic quiver* with *n* nodes, and at each step pick one of the *sources*, i.e., a node with only outgoing arrows, and reverse all these arrows so they become incoming [26]. We assign each node a "linear" variable *X* (which corresponds to a facet of the generalized associahedron), and at each step we have exactly one new variable; when mutating (a source) v to v', we replace  $X_v$  with  $X_{v'}$  and impose a linear relation [5,8]

$$X_v + X_{v'} - \sum_{w \leftarrow v} X_w = c_v, \qquad (9)$$

where  $c_v$  is a positive constant labeled by v as well. Note that for quivers that are not simply laced, there can be integer "weight", so in general we have  $-\sum_w n_w X_w$  instead (with  $n_w$  being the weight for w of the arrow  $w \leftarrow v$ ). We claim that if and only if the quiver is a Dynkin diagram,

we can consistently stop the walk and end up with a polytope which is independent of the constants [8]. For such a case, we walk N - n steps and require all the N variables to be positive, X > 0, which gives a *n*-dim generalized associahedron with N facets; (9) allows us to solve all N variables in terms of the original n X's and N - n positive c's. It is highly nontrivial that given a Dynkin diagram, starting from any acyclic quiver and walking in any order always give the generalized associahedron (with different ABHY realizations).

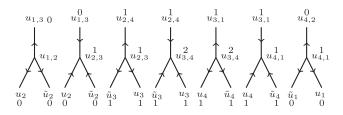
For a given variable *b*, we define b||a for all *a* from a walk with N - n steps as follows. Choose any initial quiver where *b* is a source and for any variable *a*, the degree is defined as the solution to (9),  $b||a := X_a$  with the following conditions on initial *X*'s and the *c*'s. Let us set all the *n* initial variables X = 0 (including  $X_b = 0$ ), and we set all c = 0 except for  $c_b = 1$  (corresponding to the step immediately after  $X_b$ ). By (9) we see  $b||a = X_a$  is an integer "Green's function" from *b* to *a* with the only nonvanishing source from  $c_b = 1$ . To obtain the  $N \times N$  matrix we need to do such walks starting with all *N* of the *a* variables. One can check that in the simply-laced case we always have a||b = b||a, for type  $\mathcal{A}$  we have only a||b = 0, 1, and for other types we can have b||a > 1.



Let us give some examples of u equations from our definition of compatibility degrees. We start with  $\mathcal{D}_n$  with  $N = n^2$  variables, and a natural way to label these variables is to first assign for the Dynkin diagram n initial variables  $u_1, \tilde{u}_1, u_{1,2}, u_{1,3}, \dots, u_{1,n-1}$ , and the rest can be obtained by cyclic rotations. There are three types of u equations, those for  $u_i$  or  $\tilde{u}_i$ , those for  $u_{i,i+1}$  and those for  $u_{i,j}$  for non-adjacent i, j. It suffices to write them explicitly for the first nontrivial example,  $\mathcal{D}_4$ :

4 eps: 
$$1 - u_{1,2} = u_3 \tilde{u}_3 u_4 \tilde{u}_4 u_{3,4}^2 u_{2,3} u_{2,4} u_{4,1} u_{3,1},$$
  
4 eps:  $1 - u_{1,3} = u_4 \tilde{u}_4 u_{4,1} u_{4,2} u_{2,4} u_{3,4},$   
8 eps:  $1 - u_1 = \tilde{u}_2 \tilde{u}_3 \tilde{u}_4 u_{2,3} u_{3,4} u_{2,4}.$  (10)

The compatibility degrees in (10) are obtained from walks as discussed above. Let us show an example of the walk which gives b||a| with b = 12 (when 3 sources can be mutated at some steps we do all of them together).



The nonsimply-laced Dynkin diagrams can be obtained from simply-laced ones by *folding*. Type  $\mathcal{B}_{n-1}$  can be obtained from type  $\mathcal{D}_n$  by identifying  $u_i$  and  $\tilde{u}_i$  for i = 1, ..., n, thus we can obtain  $N = n^2 - n$  *u*'s and equations for  $\mathcal{B}_{n-1}$  (here b||a is not necessarily symmetric) [27]. For example, for  $\mathcal{B}_3$  we have 12 *u* equations:

$$1 - u_{1,2} = u_3^2 u_4^2 u_{4,1} u_{3,1} u_{3,4}^2 u_{2,4} u_{2,3},$$
  

$$1 - u_{1,3} = u_4^2 u_{4,2} u_{4,1} u_{3,4} u_{2,4},$$
  

$$1 - u_1 = u_2 u_3 u_4 u_{2,3} u_{3,4} u_{2,4},$$
(11)

plus cyclic rotations. Similarly,  $C_{n-1}$  can be obtained from  $\mathcal{A}_{2n-3}$ , which corresponds to triangulating a 2n-gon, by identifying  $u_{i,j} = u_{i+n,j+n}$  (the cluster seeds correspond to centrally-symmetric triangulations). Combinatorially, it is also an (n-1)-dim cyclohedron, but its u equations, obtained from folding those of  $\mathcal{A}_{2n-1}$ , are different from those of  $\mathcal{B}_{n-1}$  since some compatibility degrees differ, except for the case  $\mathcal{B}_2 = C_2$ .

Finally let us give the only two 2d cases beyond  $A_2$ ; hexagon,  $B_2 = C_2$  and octagon,  $G_2$ . From identification of  $A_3$ , and denoting the variables cyclically as  $u_1$ ,  $v_1$ ,  $u_2$ ,  $v_2$ ,  $u_3$ ,  $v_3$ , the *u* equations for  $B_2 = C_2$  read

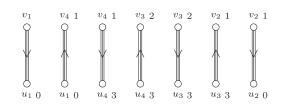
$$1 - u_1 = u_2 v_2 u_3, \qquad 1 - v_1 = v_2 u_3^2 v_3, \qquad (12)$$

plus 2 more cyclic rotations. Here we see the walk that gives the equation for  $v_1$  is

$v_1$	$v_3 1$	$v_3 1$	$v_2 \ 1$	$v_2 \ 1$
Ω	Q	Ω	Ω	Q
$\downarrow$	М	$\downarrow$	М	$\downarrow$
ň	11	Ň	11	Ň
8	8	8	8	8
$u_1 \ 0$	$u_1 \ 0$	$u_3 2$	$u_3 2$	$u_2 \ 0$

From identification of  $\mathcal{B}_3$  and renaming variables as  $u_1, v_1, ..., u_4, v_4$ , we have 8 *u* equations for  $\mathcal{G}_2$  (below we show the walk for  $v_1$ ), which are cyclic rotations of

$$1 - u_1 = u_2 v_2 u_3^2 v_3 u_4, \qquad 1 - v_1 = v_2 u_3^3 v_3^2 u_4^3 v_4. \tag{13}$$



Finally, we note that  $\mathcal{F}_4$  can be obtained by folding  $\mathcal{E}_6$ .

### **B.** Cluster configuration space

The first observation about the *u* equations for a finite type  $\Phi$  is that only N - n of them are independent, thus the solution space, denote as  $U(\Phi)$ , is *n* dimensional. Purely algebraically, the *U* space has the same boundary structure as the generalized associahedron; as  $u_a \to 0$ , all incompatible  $u_b \to 1$  (those with b||a > 0), and the boundary has the same structure as the corresponding facet of the generalized associahedron. Moreover, if we impose that all *u*'s are positive, we have the *positive part* with all *N*  $0 < u_a < 1$ , which cut out a"curvy" generalized associahedron of type  $\Phi$ . For example,  $U^+(\mathcal{B})$ ,  $U^+(\mathcal{C})$  give "curvy" cyclohedra whose facets are of the shape  $\mathcal{A} \times \mathcal{B}$ (or  $\mathcal{A} \times \mathcal{C}$ );  $U^+(\mathcal{D}_4)$  is a curvy polytope with 12  $\mathcal{A}_3 \sim \mathcal{D}_3$ facets and 4  $\mathcal{A}_1^3$  (cubes), and in total 50 vertices (one for each seed when 4 compatible *u*'s approach zero).

We remark that the *u* equations for any finite type can be derived from a set of equations we call *local u* equations, which are associated with the walks. For any walk with N - n steps, at each step we write

$$\frac{1-u_v}{u_v}\frac{1-u_{v'}}{u_{v'}} = \prod_{w \leftarrow v} (1-u_w), \tag{14}$$

for mutating a source v to v'. By considering different walks, we find that there are N such local u equations (14) with rank N - n; we claim that they are equivalent to (1). For example, one can show that (3) are equivalent to the following  $\frac{n(n-3)}{2}$  equations for  $\mathcal{A}_{n-3}$ :

$$\frac{1-u_{i,j}}{u_{i,j}}\frac{1-u_{i+1,j+1}}{u_{i+1,j+1}} = (1-u_{i,j+1})(1-u_{i+1,j}).$$
 (15)

For  $\mathcal{B}_2 = \mathcal{C}_2$  and  $\mathcal{G}_2$ , (12) and (13) are equivalent to

$$\frac{1-u_i}{u_i}\frac{1-u_{i+1}}{u_{i+1}} = 1-v_i, \quad \frac{1-v_i}{v_i}\frac{1-v_{i+1}}{v_{i+1}} = (1-u_{i+1})^p,$$

for p = 2, 3, respectively (note we have i = 1, 2, ..., p + 1). With the identification  $Y \coloneqq \frac{u}{1-u}$ , the local u equations agree with recurrence relations for these Y variables in the Y system for any finite type cluster algebra [6,28].

From (14) we see that only for an acyclic quiver can all the *N u* variables be solved *rationally* using its *n* variables. One can then show that the form  $\prod_{\alpha} d \log \frac{u_{\alpha}}{1-u_{\alpha}}$  is identical (up to a sign) for all such seeds, which in turn gives the canonical cluster form  $\Omega(U^+)$  with logarithmic singularities on the boundaries of the *U* space.

### **IV. OPEN CLUSTER STRING INTEGRALS**

Based on any binary positive geometry, it is natural to write down stringy integrals generalizing usual string amplitudes. To write down the canonical form of  $U^+$ , we apply exactly the same prescription as in the type  $\mathcal{A}$  case. Pick any seed with an acyclic quiver which has  $u_{\alpha}$  with  $\alpha$  varying over the *n* nodes, e.g., for  $\mathcal{D}_n$  we can choose  $u_1, u'_1, u_{12}, ..., u_{1,n-1}$ ; we find a diffeomorphism from  $\{0 < u_{\alpha} < 1\}$  to  $U_n^+(\Phi)$ , thus we have a (trivial) pushforward formula for  $\Omega(U^+(\Phi))$  as in (7). The open cluster string integral over  $U^+(\Phi)$  is defined as

$$\mathcal{I}^{\Phi}(\{X\}) \coloneqq (\alpha')^n \int_{U^+(\Phi)} \prod_{\alpha}^n d\log \frac{u_{\alpha}}{1-u_{\alpha}} \prod_a^N u_a^{\alpha' X_a},$$

where  $X_a > 0$  and we have a meromorphic function of X's which is reminiscent of string amplitudes. Such integrals are exponentially suppressed as  $\alpha' \to \infty$ , and satisfy the analog of channel-duality and Regge behavior [12]. It is the stringy canonical forms of an ABHY generalized associahedron [12] (see Sec. 9.3 for a  $\mathcal{D}_4$  example): if we choose a positive parametrization using e.g., principal coefficients, the regulator  $\prod u^X$  contains the N - n F-polynomials. As  $\alpha' \rightarrow 0$ , we have the ABHY polytope for the dual Dynkin diagram, which is given by the Minkowski sum of the Newton polytopes [14]. Its canonical function gives the leading order of the integral,  $\sum_{\text{seed}} \prod_{v} \frac{1}{X_v}$ , where we sum over all vertices/seeds with each term given by product of  $X_v$  for its d facets. For  $\mathcal{B}_{n-1}/\mathcal{C}_{n-1}$  and  $\mathcal{D}_n$ , it gives planar npoint tadpole diagrams, and one-loop planar integrand for bi-adjoint  $\phi^3$ -theory, respectively [8]. The same result can be obtained from pushforward due to the "scatteringequation map" from  $U^+$  to the ABHY polytope.

These generalized string amplitudes are very special since on any "massless" pole,  $X_a = 0$ , they have factorizations associated with Dynkin diagrams, at finite  $\alpha'$ . The argument is the same as for type  $\mathcal{A}$ —the residue at  $X_a = 0$  is given by the integral over the boundary with  $u_a \rightarrow 0$ , where both the form  $\Omega$  and the regulator  $\prod u^X$  "factorize" by removing a node of the Dynkin diagram.

For example, removing a generic node in type  $\mathcal{D}_n$ corresponds to  $X_{i,i} \to 0$ , and the integral  $\mathcal{I}(\mathcal{D}_n)$  factorizes as  $\mathcal{I}(\mathcal{A}) \times \mathcal{I}(\mathcal{D})$  [there are n(n-3) such facets]. Removing the trivalent node corresponds to  $X_{i,i+1} \rightarrow 0$ , and the integral factorizes as  $\mathcal{I}(\mathcal{A}_{n-3}) \times \mathcal{I}(\mathcal{A}_1) \times \mathcal{I}(\mathcal{A}_1)$  (*n* such facets) and for removing one of the two special nodes,  $X_i \to 0$  (or  $X_i \to 0$ ), the residue is given by the "forward limit" of the (n + 2)-point string integral,  $\mathcal{I}(\mathcal{A}_{n-1})$ . More exotic factorizations happen for type  $\mathcal{E}_m$  with m = 6, 7, 8, e.g., for  $\mathcal{I}(\mathcal{E}_6)$  the polytope has  $6 \times 7 = 42$  facets; for any of the 7 poles associated with the trivalent node, the residue factorizes as  $\mathcal{I}(\mathcal{A}_2) \times \mathcal{I}(\mathcal{A}_2) \times \mathcal{I}(\mathcal{A}_1)$  or "5-pt"  $\times$  "5-pt"  $\times$  "4-pt", and on other facets we have  $\mathcal{I}(\mathcal{A}_4) \times \mathcal{I}(\mathcal{A}_1), \mathcal{I}(\mathcal{D}_5), \text{ or } \mathcal{I}(\mathcal{A}_5).$  It is remarkable that these integrals contain products of string amplitudes as residues on massless poles.

## V. CLOSED CLUSTER STRING INTEGRALS

Finally, let us briefly comment on extended u equations, and the analogs of "orderings" and closed-string integrals for the cluster configuration space. For any finite type  $\Phi$ , one has extended u equations of the form

$$U_I + V_I = 1, \tag{16}$$

where *I* indexes a *mutation relation*—just as for type  $\mathcal{A}$  where (5) corresponds to a mutation from  $(ac) \rightarrow (bd)$ , we show in [14] that Eq. (16) corresponds to mutation relations of the cluster algebra. Here  $U_I$  and  $V_I$  are monomials of *u* variables determined by a mutation, and for a special mutation in our walk, it reduces to one of the usual *u* equations (1) with  $U_I = u_a$  and  $V_I = \prod_b u_b^{b||a}$ . We conjecture that the "orderings", or connected com-

We conjecture that the "orderings", or connected components of  $U(\mathbb{R})$  (with boundaries removed), are in bijection with the sign patterns consistent with all the extended *u* equations (16). A new phenomenon is that in general these regions can have shapes different from the original positive part. For  $U(\mathcal{C}_n)$  we find that the real space is tiled by components that are "curvy" cyclohedra and associahedra, e.g., for  $\mathcal{B}_2 = \mathcal{C}_2$  we find 4 hexagons and 12 pentagons tiling the space  $U(\mathbb{R})$ .

Based on the "orderings", one can also define closed cluster string integrals over  $U(\mathbb{C})$  for finite-type cluster algebras, with canonical forms and regulators for two "orderings"  $\Omega(\alpha) \prod u^X (\Omega(\beta) \prod u^X)^*$  (and similarly open-stringy integrals with forms for two components). Such a complex integral is well-defined if we can write *u*'s in ordering  $\beta$  as monomials of *u*'s in ordering  $\alpha$ , so that exponents of any *u* and *u*<sup>\*</sup> differ only by integer shifts. We leave a detailed discussion of the integrals and possible physical meaning to future work.

#### **VI. POINT COUNTS**

To find the connected components of  $U(\Phi)(\mathbb{R})$  for any finite type, it is useful to study various topological properties of the U space [14]. The latter can be done beautifully by counting the number of points in U over a finite field  $\mathbb{F}_q$ ,  $\mathcal{N}(q)$ , for any prime number q (excluding some bad prime cases). It is remarkable that  $\mathcal{N}(q)$  is a *polynomial* of q for type  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$ , and it is a quasipolynomial for cases including  $\mathcal{D}_4$ ,  $\mathcal{G}_2$  etc. For these cases,  $|\mathcal{N}(q=-1)|$  and  $|\mathcal{N}(q=1)|$  immediately give the number of connected components of  $U(\mathbb{R})$ , and the Euler characteristic of  $U(\mathbb{C})$  (which is the number of saddle points/independent integral functions), respectively. Moreover, the coefficient of each order in q is the dimension of the corresponding (co)homology group, and in particular  $|\mathcal{N}(q=0)|$  gives the number of linearly independent top-dimensional  $d \log$ forms.

Let us spell out the counting for simple cases. For  $A_{n-3}$ , the point count gives  $(q-2)(q-3)\cdots(q-n+2)$ , and

we recover (n-1)!/2, (n-2)! and (n-3)! as mentioned above. For  $U(\mathcal{B}_n)$  we find  $\mathcal{N}(q) = (q-n-1)^n$ , thus the number of components, top-dim *d* log forms and saddle points are  $(n+2)^n$ ,  $(n+1)^n$ , and  $n^n$ , respectively. For  $U(\mathcal{C}_n)$ ,  $\mathcal{N}(q) = (q-n-1)(q-3)(q-5)\cdots(q-2n+1)$ , thus these numbers are (2n)!!(n+2), (2n+1)!!(n+1)and (2n)!!/2.

Let us include two examples with quasipolynomial counting. For  $\mathcal{G}_2$ , we find  $\mathcal{N}(q) = (q-4)^2 + 4\delta_q$  where  $\delta_q = 0$  if  $q = 2 \mod 3$  and 1 if  $q = 1 \mod 3$ ; substituting q = -1 and q = 1 in both cases gives the correct counting—25 connected components and 13 saddle points. For  $\mathcal{D}_4$ ,  $\mathcal{N}(q) = q^4 - 16q^3 + 93q^2 - 231q + 206 + 2\delta_q$ , and indeed we find 547 components and 55 saddle points.

### VII. OUTLOOK

There are a large number of obvious mathematical questions associated with the binary geometries we have introduced in this letter, some of which will be taken up in [14]. But the most urgent and interesting question is a physical one: is there any natural physical meaning to the cluster string integrals beyond the usual string amplitudes associated with type  $\mathcal{A}$ ? We know that these objects strikingly generalize the remarkable factorization properties of string amplitudes, and we also know that in the  $\alpha' \rightarrow 0$  limit, at least for the classical  $\mathcal{B}_n$ ,  $\mathcal{C}_n$ ,  $\mathcal{D}_n$  types, they reduce to field theory integrands at one loop. Are they somehow related to real string amplitudes at one loop? If so the string-loop amplitudes are clearly being represented in a completely different way. If not, what is the physical interpretation of these functions?

Let us close by remarking that the existence of the uequations giving a binary realization of the compatibility for cluster polytopes is rather miraculous. Of course we can write down such *u* equations for any *n*-dimensional polytope; for any facet f and all facets f' that are not adjacent to f, we ask  $u_f + \prod_f' u_f' = 1$ . Naively there are as many equations as unknowns here, and we would find only a discrete set of solutions, rather than an *n*-dimensional space. It is highly nontrivial that these equations are consistent with each other and give an n-dimensional solution space. Indeed, working with the simplest case of n = 2 dimensional polygons, one can attempt to write e.g., the u equations with all unit exponents for any N-gon, and remarkably find that they are only consistent for N = 4, 5, which correspond to  $A_1 \times A_1$ ,  $A_2$ . By allowing nontrivial exponents we then find N = 6, 8 corresponding to  $\mathcal{B}_2 = \mathcal{C}_2$ and  $\mathcal{G}_2$ . Preliminary investigations [29] suggest that one way to obtain such binary geometries is to consider "degenerations" of generalized associahedra with certain c's set to zero, including products of lower-dimensional ones but many more nontrivial cases. It is fascinating that this simple and natural mathematical question, about binary realizations of the combinatorics of polytopal geometry, can be realized ARKANI-HAMED, HE, LAM, and THOMAS

by incredibly special polytopes with a factorizing boundary structures, including the ones associated with unitary particle scattering processes in spacetime.

# ACKNOWLEDGMENTS

We thank Francis Brown, Zhenjie Li, Lecheng Ren, Giulio Salvatori, and Zhang for discussions. S. H. and H. T. would like to thank the Institute for Advanced Study,

- Princeton and Center for Mathematical Sciences and Applications, Harvard, for hospitality during various stages of the work. S. H. was supported in part by National Natural Science Foundation of China under Grant No. 11935013, 11947301, 12047502, 12047503, 12225510. T. L. was supported by NSF Grant No. DMS-1464693 and by a von Neumann Fellowship from the Institute for Advanced Study. H. T. was partially supported by an NSERC Discovery Grant and the Canada Research Chairs program.
- [1] J. D. Stasheff, Trans. Am. Math. Soc. 108, 275 (1963).
- [2] J. D. Stasheff, Trans. Am. Math. Soc. 108, 293 (1963).
- [3] N. Arkani-Hamed, Y. Bai, S. He, and G. Yan, J. High Energy Phys. 05 (2018) 096.
- [4] N. Arkani-Hamed, Y. Bai, and T. Lam, J. High Energy Phys. 11 (2017) 039.
- [5] V. Bazier-Matte, G. Douville, K. Mousavand, H. Thomas, and E. Yildirim, arXiv:1808.09986.
- [6] S. Fomin and A. Zelevinsky, Ann. Math. 158, 977 (2003).
- [7] F. Chapoton, S. Fomin, and A. Zelevinsky, Can. Math. Bull. 45, 537 (2002).
- [8] N. Arkani-Hamed, S. He, G. Salvatori, and H. Thomas, J. High Energy Phys. 11 (2022) 049.
- [9] P. Deligne and D. Mumford, Publ. Math. l'Inst. Hautes Étud. Sci. 36, 75 (1969).
- [10] F. C. S. Brown, Ann. Sci. Ecole Norm. Sup. 42, 371 (2009).
- [11] Z. Koba and H. B. Nielsen, Nucl. Phys. B12, 517 (1969).
- [12] N. Arkani-Hamed, S. He, and T. Lam, J. High Energy Phys. 02 (2021) 069.
- [13] As will be discussed in [14], for non-simply laced case we actually obtain the generalized associahedron of the dual Dynkin diagram.
- [14] N. Arkani-Hamed, S. He, and T. Lam, J. High Energy Phys. 02 (2021) 069.
- [15] F. Cachazo, S. He, and E. Y. Yuan, Phys. Rev. Lett. 113, 171601 (2014).

- [16] F. Cachazo, S. He, and E. Y. Yuan, J. High Energy Phys. 07 (2014) 033.
- [17] S. Mizera, Phys. Rev. Lett. 120, 141602 (2018).
- [18] R. Kleiss and H. Kuijf, Nucl. Phys. B312, 616 (1989).
- [19] F. Cachazo, S. He, and E. Y. Yuan, Phys. Rev. D 90, 065001 (2014).
- [20] Z. Bern, J. J. M. Carrasco, and H. Johansson, Phys. Rev. D 78, 085011 (2008).
- [21] N. E. J. Bjerrum-Bohr, P. H. Damgaard, and P. Vanhove, Phys. Rev. Lett. 103, 161602 (2009).
- [22] S. Stieberger, arXiv:0907.2211.
- [23] F. Cachazo, S. He, and E. Y. Yuan, J. High Energy Phys. 07 (2015) 149.
- [24] V. I. Arnol'd, Math. Notes Acad. Sci. USSR 5, 138 (1969).
- [25] S. Mizera, Aspects of Scattering Amplitudes and Moduli Space Localization (Springer, Cham, 2019).
- [26] Each step of the walk is a special kind of *mutation*, and it is always possible since the quiver stays *acyclic* and has at least one source at each step.
- [27] Note that the arrowheads here are cluster-theoretic, parallelling those in the simply-laced case, and should not be confused with the Lie-theoretic arrowheads which are sometimes drawn in non-simply laced types.
- [28] A. Zamolodchikov, Phys. Lett. B 253, 391 (1991).
- [29] S. He, Z. Li, P. Raman, and C. Zhang, J. High Energy Phys. 10 (2020) 054.