A graphic approach to identities induced from multi-trace Einstein-Yang-Mills amplitudes

Yi-Jian Du\textsuperscript{a,b,c}, Linghui Hou\textsuperscript{a}

\textsuperscript{a}Center for Theoretical Physics, School of Physics and Technology, Wuhan University, No.299 Bayi Road, Wuhan 430072, China
\textsuperscript{b}College of Science, Tibet University, No.10 Zangda East Road, Lasa, 850000, China
\textsuperscript{c}Suzhou Institute of Wuhan University, No.377 Linquan Street, Suzhou, 215123, China

E-mail: yijian.du@whu.edu.cn, hlh@whu.edu.cn

ABSTRACT: Symmetries of Einstein-Yang-Mills (EYM) amplitudes, together with the recursive expansions, induce nontrivial identities for pure Yang-Mills amplitudes. In the previous work [1], we have already proven that the identities induced from tree level single-trace EYM amplitudes can be precisely expanded in terms of BCJ relations. In this paper, we extend the discussions to those identities induced from all tree level \textit{multi-trace} EYM amplitudes. Particularly, we establish a refined graphic rule for multi-trace EYM amplitudes and then show that the induced identities can be fully decomposed in terms of graph-based BCJ relations (hence traditional BCJ relations).

KEYWORDS: Amplitude Relation, Gauge invariance
1 Introduction

It has been proven that any tree level multi-trace Einstein-Yang-Mills (EYM) amplitude $A^{(m,s)}$ with $m$ gluon traces and $s$ gravitons can be expanded in terms of the amplitudes $A^{(m',s')}$ with $s' + m' < s + m$ [2] recursively. For $A^{(1,s)}$ with only one trace, these relations turn to the earlier proposed recursive expansions of single-trace EYM amplitudes [3–5]. When applying the recursive expansions repeatedly until there is no graviton and only one gluon trace in each amplitude, we finally express an arbitrary EYM amplitude as a combination of tree level color-ordered Yang-Mills (YM) ones. Such pure-YM expansions, in cases of

- 1 -
amplitudes with only a few gravitons and/or gluon traces, were earlier suggested in [6–9]. In a recent work [10], the pure-YM expansion of single-trace EYM amplitudes into BCJ basis was studied.

A remarkable result of the expansions of multi-trace EYM amplitudes is that nontrivial identities for color-ordered YM amplitudes can be induced when the gauge invariance conditions or cyclic symmetries of traces are imposed [2]. The coefficients in such induced identities generally contain factors of the Lorentz contractions $\epsilon \cdot \epsilon$, $\epsilon \cdot k$ and $k \cdot k$ where $\epsilon^\mu$ are half polarizations of gravitons and $k^\mu$ are external momenta. These identities guaranteed the localities in the Britto-Cachazo-Feng-Witten (BCFW) [11, 12] proof of the recursive expansions [2, 3] and played a crucial role in the proof [13] of the equivalence between distinct approaches [14–16] to nonlinear sigma model amplitudes. On another hand, color-ordered YM amplitudes also satisfy Kleiss-Kuijf (KK) [17] and Bern-Carrasco-Johansson (BCJ) relations [18] which does not involve any $\epsilon \cdot \epsilon$ or $\epsilon \cdot k$ factor.

Although the identities induced from multi-trace EYM amplitudes have quite different forms from the BCJ relations, it seems that they can be related with each other: (i). First, the relationship between the identities induced from single-trace EYM amplitudes and BCJ relations have already been founded [1], while the multi-trace amplitudes can be obtained through replacing gravitons by gluon traces in an appropriate way [2]. (ii). Second, explicit examples in [2] support that the identities induced from multi-trace EYM amplitudes can be expanded in terms of BCJ relations. (iii). Third, the fact that color-kinematic duality [18], which is the structure behind BCJ relations, can be resulted from gauge invariance [19] implies that the identities induced from gauge invariance can also be related with BCJ relations. Nevertheless, the full relationships between identities induced from multi-trace EYM amplitudes and BCJ relations are still unclear yet.

In this paper, we complete our study on identities induced from an arbitrary multi-trace EYM amplitude. We show that all these induced identities can be expanded in terms of BCJ relations. The main idea is sketched as follows. To incorporate more gluon traces, we introduce the refined graphic rule which involves more line styles and more types of components than that of the identities induced from single-trace amplitudes [1]. Then we show that all graphs for the identities induced from multi-trace amplitudes can be obtained by connecting two mutually disjoint connected subgraphs (which are called the final upper and lower blocks) via a line corresponding to $k \cdot k$ factor. All those graphs corresponding to a same configuration of the final upper and lower blocks together contribute a combination of the graph-based BCJ relations [1] which can be further written in terms of the traditional BCJ relations.

The structure of this paper is following. In section 2, we review the expansions of multi-trace EYM amplitudes, BCJ relations and induced identities, which are helpful for the full paper. The refined graphic rule is given in section 3. Two instructive examples are displayed in section 4 and the general pattern is given in section 5. To avoid tedious discussions, we just describe the pattern by emphasizing the new features of identities induced from multi-trace amplitudes without a proof. We claim that the general proof of the pattern in section 5 can be obtained by a parallel discussion with [1]. Conclusions and further

---

1Similar identity was discussed in [20] by an algebraic approach.
discussions are presented in section 6. The proof of a helpful relation is given in appendix A.

2 Expansions of multi-trace EYM amplitudes, induced identities and BCJ relations

In this section, we review the recursive/pure-YM expansions of multi-trace EYM amplitudes, the induced identities and the BCJ relations for YM amplitudes.

2.1 Expansions of multi-trace EYM amplitudes

A tree level multi-trace EYM amplitude with the gluon-trace set $\text{TrTrTr} \equiv \{1, \ldots, m\}$ and the graviton set $\mathcal{H} \equiv \{h_1, \ldots, h_s\}$ is expressed by $A^{(m,s)}(1, 2, \ldots, r|2| \ldots |m||\mathcal{H})$ where the first gluon trace 1 is assumed to consist of $r$ gluons in the cyclic order 1, 2, ..., $r$. We denote the union of the trace set (with the trace 1 excluded) and the graviton set by $\mathcal{HHH} \equiv (\text{TrTrTr} \setminus 1) \cup \mathcal{H} \equiv \{H_1, H_2, \ldots, H_{m+s-1}\}$, in which each gluon trace is treated as a single element. The amplitude $A(1, 2, \ldots, r|2| \ldots |m||\mathcal{H})$ is then written as $A(1, 2, \ldots, r||\mathcal{HHH})$.

With these notations, one can write down the following recursive expansion for tree level multi-trace EYM amplitudes [2]:

$$A(1, 2, \ldots, r||\mathcal{HHH}) = \sum_{\mathcal{HHH} \setminus H_a \to \text{perms } \mathcal{HHH}} \sum_{\mathcal{HHH} A} \sum_{\sigma} C(1, \sigma, r) A(1, \sigma, r||\mathcal{HHH} B)$$ (2.1)

In the expansion (2.1), we have picked out an arbitrary element $H_a$ from $\mathcal{HHH}$, which is called the fiducial element. Apparently, $H_a$ can be either a graviton or a gluon trace. On the RHS of eq. (2.1):

- The first summation is taken over all possible splittings of the set $\mathcal{HHH} \setminus H_a \rightarrow \mathcal{HHH} A |\mathcal{HHH} B$ and all permutations of elements in the set $\mathcal{HHH} A$ for a given splitting.

- For a given splitting of $\mathcal{HHH} \setminus H_a$ and a given permutation of elements in $\mathcal{HHH} A$, the summation $\sum_{\mathcal{HHH} A}$ is defined as follows

$$\text{if } H_a \text{ is a graviton: } \sum_{\mathcal{HHH} A} \rightarrow \sum_{\{a_i, b_i\} \subset t_i \text{ for all } t_i \in \mathcal{HHH} A} (-1)^{|t_i, a_i, b_i|} \sum_{\beta_i} C(1, \sigma, r) A(1, \sigma, r||\mathcal{HHH} B)$$ (2.2)

$$\text{if } H_a = t_0 \text{ is a gluon trace: } \sum_{\mathcal{HHH} A} \rightarrow \sum_{a_0 \in t_0 \text{ for } a_0 \neq b_0 \in t_0} (-1)^{|t_0, a_0, b_0|} \sum_{\beta_0} \sum_{\{a_i, b_i\} \subset t_i \text{ for all } t_i \in \mathcal{HHH} A} (-1)^{|t_i, a_i, b_i|} \sum_{\beta_i}$$ (2.3)

i.e., we sum over all possible choices of the ordered pair of gluons $\{a_i, b_i\} \subset t_i$ for all traces $t_i \in \mathcal{HHH} A$.

If the fiducial element $H_a$ is also a trace, namely $t_0$, we should further fix an arbitrary gluon $b_0 \in t_0$ (we call it fiducial gluon), then sum over all choices of $a_0 \neq b_0$ in this trace. Since a gluon trace has cyclic symmetry, it can always be written in the form $a_i, X_i, b_i, Y_i$ where $X_i$ and $Y_i$ are the two ordered sets of gluons separated by $a_i$ and $b_i$. The sign $(-1)^{|t_i, a_i, b_i|}$ is defined as $(-1)^{|Y_i|}$ where $|Y_i|$
is the number of elements in \( Y_i \). We also sum over permutations \( \beta_i \) (for a given \( \{a_i, b_i\} \)) and \( \beta_0 \) (for a given \( a_0 \neq b_0 \)) which satisfy

\[
\beta_i \in \{a_i, KK|t_i, a_i, b_i|, b_i\} \equiv \{a_i, X_i \sqcup Y_i^T, b_i\}, \quad \text{[Eq:ShuffleInTR]}
\]

\[
\beta_0 \in \{a_0, KK|t_0, a_0, b_0|, b_0\} \equiv \{a_0, X_0 \sqcup Y_0^T, b_0\}, \quad \text{[Eq:ShuffleInTR1]}
\]

where \( Y^T \) denotes the inverse order of elements in \( Y \) and \( A \sqcup B \) for two ordered sets \( A \) and \( B \) stands for the set of all permutations obtained by merging the two sets with keeping the relative order of elements in each set.

- Suppose that the permutation of elements of \( H_A \) in the first summation is \( H_{\rho(1)}, H_{\rho(2)}, \ldots, H_{\rho(|H_A|)} \) and the gluon pairs in the second summation are \( \{a_i, b_i\} \subset t_i \) (i can be 0 if \( H_a \) is a gluon trace \( t_0 \)), we sum over all permutations \( \sigma \) satisfying

\[
\sigma \in \{2, \ldots, r - 1\} \sqcup \{H_{\rho(1)}, H_{\rho(2)}, \ldots, H_{\rho(|H_A|)}, H_a\}. \quad \text{[Eq:sigma]}
\]

Here the traces in \( \{H_{\rho(1)}, H_{\rho(2)}, \ldots, H_{\rho(|H_A|)}, H_a\} \) are no longer considered as single elements but considered as proper permutations of all gluons in them. Particularly, if \( H_{\rho(i)} \) denotes a gluon trace \( t_i \) and the trace can be written as \( a_i, X_i, b_i, Y_i \) (Here, \( X_i \) and \( Y_i \) are the two ordered sets of gluons separated by \( a_i \) and \( b_i \)) for the given choice of pair \( \{a_i, b_i\} \subset t_i \), we should replace \( H_{\rho(i)} \) in eq. (2.6) by a permutation \( \beta_i \) satisfying eq. (2.4). Similarly, if the fiducial element \( H_a \) is also a trace, say \( t_0 \), it must be replaced by a \( \beta_0 \) satisfying eq. (2.5). In other words, each trace in eq. (2.6) is replaced by a permutation of KK basis with fixed two ends. Then the summation over \( \sigma \) means summing over all possible shuffle permutations in eq. (2.6) after this replacement.

- Given a splitting \( H \setminus H_a = H_A \setminus H_B \) and a permutation \( H_{\rho(1)}, H_{\rho(2)}, \ldots, H_{\rho(|H_A|)} \) of elements in \( H_A \), given \( a_i, b_i \) pairs and \( \beta_i \) for all traces in \( \{H_{\rho(1)}, H_{\rho(2)}, \ldots, H_{\rho(|H_A|)}, H_a\} \) and given \( \sigma \) satisfying eq. (2.6), the coefficient \( C(1, \sigma, r) \) in eq. (2.1) is written as

\[
\mathcal{E}_{H_a} \cdot \mathbb{F}_{H_{\rho(|H_A|)}} \cdot \ldots \cdot \mathbb{F}_{H_{\rho(1)}} \cdot Y_{H_{\rho(1)}}(\sigma), \quad \text{[Eq:EFY1]}
\]

where

\[
\mathcal{E}_{H_a}^\mu = \begin{cases} 
\epsilon_{h_a}^\mu & \text{(if } H_a \text{ is a graviton } h_a) \\
-k_{a_0}^\mu & \text{(if } H_a \text{ is a gluon trace } t_0) 
\end{cases} \quad \text{[Eq:EFY1]}
\]

\[
\mathbb{F}_{H_{\rho(l)}}^{\mu \nu} = \begin{cases} 
\epsilon_{h_x}^\mu \epsilon_{h_x}^\nu - k_{h_x}^\mu k_{h_x}^\nu & \text{(if } H_{\rho(l)} \text{ is a graviton } h_x) \\
-k_{h_x}^\mu k_{h_x}^\nu & \text{(if } H_{\rho(l)} \text{ is a gluon trace } t_l) 
\end{cases} \quad \text{[Eq:EFY2]}
\]
\[ Y^\mu_{\mathcal{H}^{(1)}}(\sigma) = \begin{cases} \sum_{\sigma^{-1}(j) < \sigma^{-1}(h_x)} k_j^\mu & \text{(if } H_{\rho(l)} \text{ is a graviton } h_x) \\ \sum_{\sigma^{-1}(j) < \sigma^{-1}(b_i)} k_j^\mu & \text{(if } H_{\rho(l)} \text{ is a gluon trace } t_i) \end{cases} \]  

(2.10)

The notation \( \sigma^{-1}(x) \) denotes the position of \( x \) in the permutation \( \sigma^3 \) and the gluon 1 is always considered as the leftmost element in \( \sigma \).

The recursive expansion eq. (2.1), where the fiducial element \( H_a \) is chosen as a graviton or a gluon trace, corresponds to the type-I or -II relation in [2]. If there is only one trace, the relation eq. (2.1) turns to the expansion of single-trace EYM amplitudes [3]. As examples, we display the two types of expansions for the double-trace EYM amplitude \( A(1, 2 \ldots r|2||h_1) \) with one graviton explicitly:

**Type-I:**

\[
A(1, 2, \ldots, r|2||h_1) = \sum_{\sigma} (\epsilon_{h_1} \cdot Y_{h_1}) A(1, \sigma \in \{2, \ldots, r - 1\} \sqcup \{h_1\}, r|2) \\
+ \sum_{\{a_2, b_2\} \subset 2} \sum_{\sigma} (\epsilon_{h_1} \cdot k_{b_2})(-k_{a_2} \cdot Y_{a_2}) A(1, \sigma \in \{2, \ldots, r - 1\} \sqcup \{a_2, \beta_2, b_2, h_1\}, r),
\]

**Type-II:**

\[
A(1, 2, \ldots, r|2||h_1) = \sum_{a_2 \in 2 \setminus b_2} \sum_{\sigma} (-k_{a_2} \cdot Y_{a_2}) A(1, \sigma \in \{2, \ldots, r - 1\} \sqcup \{a_2, \beta_2, b_2\}, r|h_1) \\
+ \sum_{a_2 \in 2 \setminus b_2} \sum_{\sigma} (-k_{a_2} \cdot F_{h_1} \cdot Y_{h_1}) A(1, \sigma \in \{2, \ldots, r - 1\} \sqcup \{h_1, a_2, \beta_2, b_2\}, r),
\]

where the factor \((-1)^{|2, a_2, b_2}|\) and the sum over \( \beta_2 \in \{a_2, KK|2, a_2, b_2\}, b_2 \} \) are hidden in the summation notations with a tilde.

When the recursive expansion eq. (2.1) is applied repeatedly, one finally expands the multi-trace EYM amplitude in terms of pure YM ones [2]. Instead of collecting the coefficients for amplitudes in each permutation (as shown in [2–4]), we introduce a distinct form where amplitudes corresponding to each graph is collected:

\[
A(1, 2, \ldots, r||\mathcal{H}) = \sum_{\mathcal{F}} \mathcal{C}^\mathcal{F} \left[ \sum_{\sigma^\mathcal{F}} A(1, \sigma^\mathcal{F}, r) \right]. \tag{2.11}
\]

[Eq:PureYMExpansion]

Here, we summed over all possible graphs which will be defined in the next section.

---

\[ ^8 \text{Supposing that the position of } x \text{ in } \sigma \text{ is } i, \text{ we have } x = \sigma(i). \text{ Thus it is reasonable to write } i \text{ as } \sigma^{-1}(x). \]
2.2 Identities induced from multi-trace EYM amplitudes

Symmetries of multi-trace EYM amplitudes, together with the expansions eq. (2.11), induce identities for EYM amplitudes with fewer gravitons and/or traces. There are two symmetries under consideration in this paper: the gauge invariance and the cyclicity of any trace.

Multi-trace EYM amplitudes satisfy the gauge invariance condition, which states that the amplitudes have to vanish once half polarization $\epsilon_{hx}$ of a graviton $h_x$ is replaced by its momentum $k_{hx}$. If this replacement is performed on the RHS of the recursive expansion eq. (2.1), we should consider two distinct situations:

- If the fiducial element $H_a$ is a graviton $h_x$ (i.e. eq. (2.1) is type-I expansion), the gauge invariance condition for $h_x$ induces a nontrivial relation between EYM amplitudes with fewer gravitons:

$$\sum_{H_b} \sum_{\text{perms } H_A/H_B} \sum_{\text{perms } H_A/H_B} \sum_{\sigma} C(1, \sigma, r) A(1, \sigma, r \parallel H_B) = 0.$$  \hspace{1cm} (2.12)

- If $h_x$ is not the fiducial one, it may belong to either $H_A$ or $H_B$ in eq. (2.1). Terms with $h_x \in H_A$ have to vanish due to the antisymmetry of the strength tensor $F_{h_x}^{\mu \nu} \equiv k_{h_x}^\mu \epsilon_{h_x}^\nu - k_{h_x}^\nu \epsilon_{h_x}^\mu$, while terms with $h_x \in H_B$ have to vanish due to gauge invariance of amplitudes with fewer gravitons.

Therefore, the only nontrivial identity induced by the gauge invariance condition of a graviton is eq. (2.12) which is called type-I identity in [1].

Another identity (called type-II identity in [2]) is induced from the type-II expansion where the fiducial element $H_a$ is a gluon trace $t_0$. In particular, we notice that the fiducial gluon $b_0 \in t_0$ can be chosen arbitrarily in the fiducial trace. This arbitrariness is essentially caused by the cyclicity of a trace [2] and implies the following identity:

$$\sum_{H_b} \sum_{\text{perms } H_A/H_B} \sum_{\text{perms } H_A/H_B} \sum_{\sigma'} C(1, \sigma', r) A(1, \sigma', r \parallel H_B) = 0.$$  \hspace{1cm} (2.13)

where $\sigma'$ is defined by

$$\sigma' \in \{2, \ldots, r-1\} \cup \{H_{p(1)}, H_{p(2)}, \ldots, H_{p(|H_A|)}\}, a_0, \beta_0 \in KK|t_0, a_0, b_0, b_0\}. \hspace{1cm} (2.14)$$

Here we remove $b_0$ from the trace $t_0$ first, then shuffle these permutations as eq. (2.6).

When the recursive expansion eq. (2.1) for amplitudes with fewer gravitons and/or gluon traces are applied repeatedly, the two types of relations eq. (2.12) and eq. (2.13) induce nontrivial relations between pure YM amplitudes, which will be studied in detail in the coming sections.
2.3 BCJ relations

Tree level color-ordered YM amplitudes satisfy the following *traditional BCJ relation*\(^4\) [21, 22]:

\[
\sum_{\sigma \in \beta} \sum_{l \in \beta} (k_l \cdot X_l(\sigma)) A(1, \sigma, r) = 0, \tag{2.15} \]

where \(\beta\) and \(\alpha\) are two ordered sets of external gluons, \(X_l(\sigma)\) denotes the sum of all momenta of gluons \(a \in \{1, \sigma\} \cup \beta\) satisfying \(\sigma^{-1}(a) < \sigma^{-1}(l)\).

In [1], the following *graph-based BCJ relation* for YM amplitudes was proposed

\[
\sum_{a \in T} f^a \sum_{\zeta \in T_a} \sum_{\sigma} [k_a \cdot Y_a(\sigma)] A(1, \sigma \in (\zeta \sqcup \gamma), r) = 0. \tag{2.16} \]

Here, \(\gamma\) is an arbitrary permutation of elements in \(\{2, \ldots, r-1\}\) and \(T\) is an arbitrary connected tree graph. When a node \(a\) is chosen as the leftmost element, the tree graph \(T\) establishes permutations \(\zeta \in T_a\) as follows (i). For two adjacent nodes \(x\) and \(y\), if \(x\) is nearer to \(a\) than \(y\), we have \(\zeta^{-1}(x) < \zeta^{-1}(y)\), (ii). If there are subtree structures attached to a same node, we should shuffle the permutations established by these subtrees together. The factor \(f^a\) is a relative sign depending on the node \(a\). This factor is determined as follows: (i). Choose an arbitrary node \(c\) and require \(f^c = 1\), (ii). For arbitrary two adjacent nodes \(c_1\) and \(c_2\), we have \(f^{c_1} = -f^{c_2}\). As already proven in [1], the graph based BCJ relation eq. (2.16) can always be written as combination of the traditional ones eq. (2.15).

3 Refined graphic rule for multi-trace EYM amplitudes

In this section, we introduce the refined graphic rule for the expansion eq. (2.11).

---

\(^4\)The name *traditional BCJ relations* is used to distinguish the BCJ relation proposed before from the *graph-based BCJ relations* proposed in [1].
\[ \beta \in KK[t_0, a_0, b_0] \]

\[ -k_{a_0} = \cdots \bullet \bullet \bullet \bullet \bullet \cdots \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \quad \beta \]

\[ \beta_{|t_0| - 2} \quad b_0 \]

\[ -k_{a_0}^\mu k_{a_i}^\nu = \cdots \bullet \bullet \bullet \bullet \bullet \cdots \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \quad \beta \]

\[ \beta_{|t_i| - 2} \quad b_i \]

Figure 2. The starting trace (a) and an internal trace (b) of a chain. Each element \( \beta(i) \) is written as \( \beta_i \) for short.

\[ a \epsilon_a \cdot \epsilon_b \quad a \epsilon_a \cdot k_b \quad a k_a \cdot k_b \quad a \bullet \bullet \bullet \bullet \bullet \quad b \]

Type-1 Type-2 Type-3 Type-4

(a) (b) (c) (d)

Figure 3. Four types of lines in the refined graphic rule for multi-trace EYM amplitudes

3.1 Refined graphic rule

The pure-YM expansion eq. (2.11) of the tree level multi-trace EYM amplitude \( A(1, 2, \ldots, r|2|\ldots|m||H) \equiv A(1, 2, \ldots, r||H) \) can be achieved by the following refined graphic rule\(^5\):

**Step-1** Define a reference order of elements in \( H \equiv (\mathbf{T}_s \setminus 1) \cup H \):

\[ R = \{ H_{\rho(1)}, H_{\rho(2)}, \ldots, H_{\rho(l=m+s-1)} \} \]  

where the position of each element (i.e. graviton or gluon trace) is called the weight.

**Step-2** We consider elements (i.e. gravitons and gluon traces) in \( H \equiv (\mathbf{T}_s \setminus 1) \cup H \) and the gluons in the trace \( 1 \) as nodes. A chain, which starts from the highest-weight element \( H_{\rho(l)} \), passes through some other elements \( j_1, \ldots, j_u \) \((j_1, \ldots, j_u \in H \setminus H_{\rho(l)})\) an ends at \( w \in \mathcal{R} = 1 \setminus \{ r \} = \{ 1, \ldots, r - 1 \} \) can be constructed

\[ \mathcal{E}_{H_{\rho}} \cdot \Xi_{j_u} \cdot \cdots \cdot \Xi_{j_1} \cdot k_w. \]

Here, \( H_{\rho} \), the elements \( j_1, \ldots, j_u \) and \( w \) are correspondingly called the starting, internal and ending elements of this chain. Redefine the ordered set \( R \) by removing the elements which have been used:

\[ R \rightarrow R' = R \setminus \{ j_1, \ldots, j_r, H_{\rho(l)} \} \equiv \{ H_{\rho(1')}, H_{\rho(2')}, \ldots, H_{\rho(l')} \}. \]

\(^5\)Graphs for CHY formulas can be found in [23–26], graphs for BCJ numerators can be found in [16] and the refined graphic rule for single-trace EYM amplitudes was established in [1]
and redefine the set $\mathcal{R}$:

$$\mathcal{R} \to \mathcal{R}' = \mathcal{R} \cup \{j_1, \ldots, j_r, \mathcal{H}_{\rho(l)}\},$$

where each gluon trace in $\{j_1, \ldots, j_r, \mathcal{H}_{\rho(l)}\}$ is treated as a single object.

**Step-3** Repeat the above step by using the new defined $\mathcal{R}$ and $\mathcal{R}'$ until the ordered set $\mathcal{R}$ is empty. We then get a tree graph.

**Step-4** Refine all the graphs constructed by the above steps via replacing gravitons and gluon traces by Fig. 1 and Fig. 2. Here, (i). All gravitons and all gluons are treated as nodes. (ii). A half polarization $\epsilon^\mu$ of a graviton is presented by a solid line with no arrow, while the momentum $k^\mu$ of a node (graviton or gluon) is presented by a solid arrow line pointing towards this node. If an arrow points away from gluons (i.e. the roots) in the trace 1, an extra minus should be dressed. Lorentz contractions give rise three types of solid lines Fig. 3 (a), (b) and (c). (iii). For a given choice of $\{a_i, b_i\} \in t_i$ ($t_i \neq 1$) and a given permutation in $\{a_i, \beta \in KK[t_i, a_i, b_i], b_i\}$, two adjacent gluons in trace $t_i$ are connected by a dashed arrow line (the type-4 line Fig. 4 (d)) whose arrow points to $a_i$. We also connect adjacent gluons in the trace $111 = \{1, 2, \ldots, r\}$ by type-4 lines whose arrows point to the gluon 1. (iv). If a trace plays as the ending element of a chain, it means any gluon belonging to this trace can be the ending node.

**Step-5** For a given graph $\mathcal{F}$, the coefficient $C^F$ in eq. (2.11) can be read off as the product of all factors corresponding to the type-1,-2 and -3 lines (see Fig. 3). The sign associated to such a graph gets two distinct contributions (i). $(-1)^N(\mathcal{F})$, where $N(\mathcal{F})$ is the number of arrows pointing away from the root 1; (ii). Each trace $t_i$ contributes a $(-1)^{|t_i,a_i,b_i|}$ for given $a_i$ and $b_i$.

**Step-6** Collect amplitudes $A(1, \sigma^F, r)$ for a given graph $\mathcal{F}$. In any graph, the gluons 1 and $r$ in the trace $1 = \{1, 2, \ldots, r\}$ are always treated as the first and the last elements. The gluon 1 can be considered as the root of any graph. Permutations $\sigma^F$ are determined as follows: (i). If two adjacent nodes $x$ and $y$ are connected by a line (of any style), they must live on a path towards the root 1. If $x$ is nearer to 1 than $y$ on this path, we have $(\sigma^F)^{-1}(x) < (\sigma^F)^{-1}(y)$. (ii). If there are several branches attached to a node, the relative order is defined by shuffling the branches together.

When summing over all possible graphs constructed by the above steps, (i.e., (i). summing over all graphs with given $\{a_i, b_i\}$ pairs and given permutations $\beta \in KK[t_i, a_i, b_i]$ for all traces, (ii). summing over all possible permutations $\beta \in KK[t_i, a_i, b_i]$ for given $\{a_i, b_i\}$ pairs in all traces, (iii). summing over all possible choices of the $\{a_i, b_i\}$ pairs for internal traces and all possible choices of $a_i \in t_i, a_i \neq b_i$ for starting traces), we finally arrive the expansion eq. (2.11).

### 3.2 Examples for the refined graphic rule

Now we take the double-trace amplitude $A(1, 2 \ldots r | 2\parallel h_1)$ with one graviton as an example. For this amplitude, the reference order can be $R = \{2, h_1\}$ or $R = \{h_1, 2\}$ where the highest-weight element is the graviton $h_1$ or the gluon trace 2 respectively.
If $R = \{2, h_1\}$, the typical graphs are shown by Fig. 4 (a), (b) and (c). Correspondingly, these graphs contribute

$$C(a) = (\epsilon_{h_1} \cdot k_{b_2})(-k_{a_2} \cdot k_l),$$  \hspace{1cm} \sigma(a) \in \{2, \ldots, l, \{l+1, \ldots, r-1\} \cup \{a_2, \beta_1, \ldots, \beta_{|2|-2}, b_2, h_1\}\}, \quad [\text{Eq:EGA}] \quad (3.5)$$

$$C(b) = (\epsilon_{h_1} \cdot k_{l'})(-k_{a_2} \cdot k_l),$$  \hspace{1cm} \sigma(b) \in \{2, \ldots, l', \{h_1\} \cup \{l'+1, \ldots, l, \{l+1, \ldots, r-1\} \cup \{a_2, \beta_1, \ldots, \beta_{|2|-2}, c_2\}\}, \quad [\text{Eq:EGb}] \quad (3.6)$$

$$C(c) = (\epsilon_{h_1} \cdot k_l)(-k_{a_2} \cdot k_{h_1}),$$  \hspace{1cm} \sigma(c) \in \{2, \ldots, l, \{l+1, \ldots, r-1\} \cup \{h_1, a_2, \beta_1, \ldots, \beta_{|2|-2}, c_2\}\}, \quad [\text{Eq:EGc}] \quad (3.7)$$

where in $\sigma(a)$, permutation $\beta = \{\beta_1, \beta_2, \ldots, \beta_{|2|-2}\}$ satisfies $\beta \in KK[2, a_2, b_2]$ and in $\sigma(b), (c)$, $\beta \in KK[2, a_2, c_2]$. Then the sum over all graphs in eq. (2.11) is given by

$$
\sum_{l \in \{1, \ldots, r-1\}} \left[ \sum_{\{a_2, b_2\} \subset 2} \sum_{\sigma(a)} C(a) A(1, \sigma(a), r) \right. \\
+ \sum_{a_2 \in 2} \sum_{a_2 \neq c_2} \left( \sum_{l' \in \{1, \ldots, r-1\}} \sum_{\sigma(b)} C(b) A(1, \sigma(b), r) \right) + \sum_{\sigma(c)} C(c) A(1, \sigma(c), r) \right] \quad (3.9)
$$
where the summation over $\beta \in KK[2, a_2, b_2]$ ($\beta \in KK[2, a_2, c_2]$) and the factor $(-1)^{t_1, a_2, b_2} ((-1)^{t_1, a_2, c_2})$ associated to the trace 2 have been absorbed in to the notations $\sum_{\{a_2, b_2\} < 2} (\sum_{a_2 \neq b_2} (\sum_{a_2 \in 2} ))$ for short. The $c_2 \in 2$ in the second term in the square brackets is arbitrarily fixed because the trace 2 plays as starting trace in Fig. 4 (b) and (c).

- If $R = \{h_1, 2\}$, the typical graphs are given by Fig. 4 (b), (c), (d) and (e). The expression of Fig. 4 (b) and (c) are already shown by eq. (3.6) and eq. (3.7). The contributions from Fig. 4 (d) and (e) read

$$C^{(d)} = (-k_{h_1} \cdot k_l)(-\epsilon_{h_1} \cdot k_{a_2}),$$

$$\sigma^{(d)} \in \{2, \ldots, l, \{l + 1, \ldots, r - 1\} \sqcup \{h_1, a_2, \beta_1, \ldots, \beta_2[l - 2, c_2]\}\}$$

$$C^{(e)} = (\epsilon_{h_1} \cdot k_{l'})(-k_{a_2} \cdot k_l),$$

$$\sigma^{(e)} \in \{2, \ldots, l, \{l + 1, \ldots, r - 1\} \sqcup \{a_2, \beta_1, \ldots, \beta_{l'} = \beta_j, \beta_{j+1}, \ldots, \beta_2[l - 2, c_2] \sqcup \{h_1\}\}\}.$$  

where $l'$ can be any element in the trace 2. The sum over all graphs in eq. (2.11) is then written as

$$\sum_{l \in \{1, \ldots, r - 1\}} \sum_{a_2 \neq c_2} \sum_{l' \in \{1, \ldots, r - 1\}} C^{(b)} A(1, \sigma^{(b)}), r) + \sum_{\sigma^{(c)}} C^{(c)} A(1, \sigma^{(c)}), r)$$

$$+ \sum_{\sigma^{(d)}} C^{(d)} A(1, \sigma^{(d)}), r) + \sum_{l' \in 2} \sum_{\sigma^{(e)}} C^{(e)} A(1, \sigma^{(e)}), r) \right].$$  

### 3.3 Expressing the induced identities by refined graphs

Having shown the refined graphic rule, we are now ready to express the induced identities eq. (2.12) and eq. (2.13) as identities between pure YM amplitudes. We just need to replace the highest-weight element (i.e. the fiducial element in the first-step recursive expansion) in

- If the highest-weight element $H_{pl(l)}$ in the reference order eq. (3.1) is a graviton $h_a$, the identity eq. (2.12), which is induced from the gauge invariance condition of $h_a$, implies:

$$0 = \sum_{F} C_{F} |_{h_a \rightarrow k_{h_a}} \sum_{\sigma} A(1, \sigma, r)].$$

The corresponding graphs can be obtained by the replacement Fig. 5 (a).

- If the highest-weight element $H_{pl(l)}$ in the reference order eq. (3.1) is a gluon trace $t_b$, the identity
Figure 5. The type-I induced identity eq. (3.13) is obtained by replacing $\varepsilon_{h_a}$ by $k_{h_a}$ (shown by (a)) in eq. (2.11), where the graviton $h_a$ is the highest-weight element in the reference order. The type-II induced identity eq. (3.14) is obtained by performing the replacement (b) on eq. (2.11). The sum over all $\beta \in \text{KK}[t_0, a_0, b_0]$ with the sign $(-1)^{|t_0|, a_0, b_0}$ can be expressed by (c), where the full trace $t_0$ is written as $d_1, d_2, \ldots, d_{j-1}, a_0 = d_j, d_{j+1}, \ldots, d_{|t_0|-1}, b_0 = d_{|t_0|}$.

**Figure 5.** The type-I induced identity eq. (3.13) is obtained by replacing $\varepsilon_{h_a}$ by $k_{h_a}$ (shown by (a)) in eq. (2.11), where the graviton $h_a$ is the highest-weight element in the reference order. The type-II induced identity eq. (3.14) is obtained by performing the replacement (b) on eq. (2.11). The sum over all $\beta \in \text{KK}[t_0, a_0, b_0]$ with the sign $(-1)^{|t_0|, a_0, b_0}$ can be expressed by (c), where the full trace $t_0$ is written as $d_1, d_2, \ldots, d_{j-1}, a_0 = d_j, d_{j+1}, \ldots, d_{|t_0|-1}, b_0 = d_{|t_0|}$.

**4 Induced identities as combinations of BCJ relations: examples**

In the previous section, we have introduced the refined graphic rule, by which, the induced identities eq. (3.13) and eq. (3.14) are expressed as summations over all possible refined graphs. Now we take the type-I and type-II identities, which are induced from the double-trace EYM amplitude $A(1, \cdots, r|2||h_1)$ (eq. (3.13)) with one graviton, as explicit examples. Through these examples, we will see identities of both types are combinations of graph-based BCJ relations (thus traditional BCJ relations).
Figure 6. Typical graphs for the identity induced from the double-trace amplitude \( A(1, 2, \cdots, r|2\parallel h_1) \) with the reference order \( R = \{2, h_1\} \). The trace 2 plays as an internal trace in (a) and starting traces in (b), (c).

### 4.1 Example-1: the type-I identity induced from \( A(1, 2, \cdots, r|2\parallel h_1) \)

When the reference order is chosen as \( R = \{2, h_1\} \) and \( \epsilon_{h_1} \) is replaced by \( k_{h_1} \), the expansion eq. (2.11) for the double-trace amplitude \( A(1, 2, \cdots, r|2\parallel h_1) \) induces a type-I identity eq. (3.13) for pure YM amplitudes. Typical graphs for this identity Fig. 6 (a), (b) and (c) are correspondingly obtained from Fig. 4 (a), (b) and (c) with the replacement \( \epsilon_{h_1} \rightarrow k_{h_1} \). The total contribution for these graphs then reads

\[
T^{(a)} = \sum_{\sigma^{(a)}} (k_{h_1} \cdot k_{b_2}) (-k_{a_2} \cdot k_{l_1}) A(1, \sigma^{(a)}, r), \quad [\text{Eq:EG1a}]
\]

\[
T^{(b)} = \sum_{\sigma^{(b)}} (k_{h_1} \cdot k_{l_1'}) (-k_{a_2} \cdot k_{l_1}) A(1, \sigma^{(b)}, r), \quad [\text{Eq:EG1b}]
\]

\[
T^{(c)} = \sum_{\sigma^{(c)}} (k_{h_1} \cdot k_{l_1}) (-k_{b_2} \cdot k_{h_1}) A(1, \sigma^{(c)}, r), \quad [\text{Eq:EG1c}]
\]

where \( \sigma^{(a)}, \sigma^{(b)} \) and \( \sigma^{(c)} \) are already given in eq. (3.5), eq. (3.6) and eq. (3.7) respectively. Then the full expression of the RHS of the identity (3.13) induced from \( A(1, \cdots, r|2\parallel h_1) \) is given by

\[
\left[ \sum_{l \in \{1, \ldots, r-1\}} \sum_{\{a_2, h_2\} \subset 2} T^{(a)} \right] + \left[ \sum_{l', l'' \in \{1, \ldots, r-1\}} \sum_{a_2 \in 2} T^{(b)} \right] + \left[ \sum_{l \in \{1, \ldots, r-1\}} \sum_{b_2 \in 2} T^{(c)} \right]. \quad [\text{Eq:SumOfT}]
\]

To investigate the relationship between the induced identity and BCJ relations, we define **standard basis set** for a gluon trace \( i \) by the set of permutations \( \{\{b_i, \beta \in KK[i, b_i, c_i], c_i\} | b_i \neq c_i, b_i \in i\} \) for an arbitrarily fixed \( c_i \in i \). Graphically, the fixed end node \( c_i \) is represented by a hollow node, while other nodes are expressed by solid ones. Any two adjacent nodes for a given permutation in \( \{\{b_i, \beta \in KK[i, b_i, c_i], c_i\} \} \) are connected by dashed arrow line pointing towards the node \( b_i \). Apparently, all those starting traces defined by the refined graphic rule are already expressed by the standard basis (because we have already fixed an end node of this trace). Moreover, all internal traces can be expanded by the standard basis,
Figure 7. When the identity eq. (4.5) is applied, the contribution of Fig. 6 (a) splits into (a1) and (a2) which are already expressed by standard basis. The graph (a1) comes from the second term of eq. (4.5), thus it must be associated with an extra minus. The graph (a3) is a spurious graph which cannot be directly obtained by the refined graphic rule and the identity eq. (4.5). It can be considered as the $a_2 = b_2$ supplement to both terms in eq. (4.5).

according to the following nontrivial property:

\[
\sum_{\{a_i, b_i\} \subseteq i} \left( \sum_{a_i \neq c_i} b_i \right)_{\beta \in KK[i, a_i, b_i]} = \sum_{a_i \in i} \sum_{b_i \in i \text{ and } a_i \neq b_i} \left( \sum_{a_i \neq c_i} b_i \right)_{\beta \in KK[i, a_i, c_i]} - \sum_{b_i \in i} \sum_{a_i \in i \text{ and } a_i \neq b_i} \left( \sum_{a_i \neq c_i} b_i \right)_{\beta \in KK[i, a_i, b_i]}.
\]

where each graph stands for its full contribution and recall that a summation with tilde is defined by, e.g., $\sum_{\{a_i, b_i\} \subseteq i} = \sum_{\{a_i, b_i\} \subseteq i} (-1)^{|i, a_i, b_i|} \sum_{\beta \in KK[i, a_i, b_i]}$. The LHS is an internal trace structure defined by the refined graphic rule. On the RHS, the trace $i$ is expressed by the standard basis with one end $c_i \in i$ fixed. The node $a_i$ ($b_i$) on both sides of eq. (4.5) must be connected to a same node outside the trace $i$ via a line of the same style (i.e., type-2 or type-3 line). We leave the proof of eq. (4.5) in the appendix.

Now we apply the relation eq. (4.5) to the first term of eq. (4.4) for a given $l$. Since the $c_i$ in eq. (4.5) can be chosen arbitrarily, we just choose the $c_i \in i$ (in this example $i = 2$) as the fiducial gluon $c_2$ in the starting trace of Fig. 6 (b) and (c). Then the summation $\sum_{\{a_2, b_2\} \subseteq 2} T^{(a)}$ in the first term of eq. (4.4) splits
The $I_1$ defined by eq. (4.7) can be expanded in terms of graph-based BCJ relations with respect to graphs of the structure $\mathcal{T}_1$. 

\[ \sum_{\{a_2, b_2\} \subset \mathcal{Q}} T^{(a)} = -\sum_{a_2 \not\subset a_2 \not=b_2} \sum_{\{a_2, b_2\} \subset \mathcal{Q}} T^{(a_1)} + \sum_{a_2 \not\subset a_2 \not=b_2} \sum_{\{a_2, b_2\} \subset \mathcal{Q}} T^{(a_2)} \quad [\text{Eq:SumOPT1}] \quad (4.6) \]

where $T^{(a_1)}$ and $T^{(a_2)}$ are the corresponding contributions of Fig. 7 (a1) and (a2). Substituting eq. (4.6) into eq. (4.4) and introducing the contribution of spurious graph Fig. 7 (a3) by $0 = T^{(a_3)} - T^{(a_3)}$, we rewrite eq. (4.4) as the sum of $I_1$ and $I_2$ which are respectively defined by

\[ I_1 \equiv \sum_{a_2 \not\subset a_2 \not=b_2} \left[ \sum_{l \in \{1, \ldots, r-1\}} \left( -\sum_{a_2 \not\subset a_2 \not=b_2} T^{(a_1)} - T^{(a_3)}|_{a_2=b_2} + T^{(c)} \right) \right] \quad [\text{Eq:Sum11}] \quad (4.7) \]

and

\[ I_2 \equiv \sum_{a_2 \not\subset a_2 \not=b_2} \left[ \sum_{l \in \{1, \ldots, r-1\}} \left( \sum_{b_2 \in \mathcal{Q}} T^{(a_2)} + T^{(a_3)} + \sum_{l' \in \{1, \ldots, r-1\}} T^{(b)} \right) \right] \quad [\text{Eq:Sum12}] \quad (4.8) \]

Let us analyze $I_1$ and $I_2$ separately:

- For $I_1$, we find that the expression in the parenthesis in eq. (4.7) can be obtained by summing over all those graphs which are constructed by connecting an arbitrary gluon $l \in \{1, 2, \ldots, r-1\} = 1 \setminus r$ with an arbitrary node $a \in \mathcal{T}_1$ ($\mathcal{T}_1$ is the tree graph shown by Fig. 8) via a type-3 line. The kinematic factor for such a graph is $(k_{a_1} \cdot k_{b_2})(k_{a_2} \cdot k_{l})$. Given $b_2 \in 2 (b_2 \not=c_2)$ in eq. (4.7), there is an overall sign $(-1)^{i(a,b)}$ which has been extracted out and absorbed into the summation notation with a tilde. The sign for any term inside the square brackets is collected as follows: (i). According to the refined graphic rule, any graph $\mathcal{F}$ with $\mathcal{N}(\mathcal{F})$ arrows pointing away from the root 1 is associated by a sign $(-1)^{\mathcal{N}(\mathcal{F})}$. (ii). Each of (a1) and (a3) has an extra sign $(-1)$. Based on this analysis of signs, we have the following observations:

  (1) Once $a \in \mathcal{T}_1$ has been chosen, the sign is independent of the choice of $l \in \{1, 2, \ldots, r-1\} = 1 \setminus r$ and all permutations established by the graph with any $l$ satisfy $\{1, \sigma \in \{2, \ldots, r-1\} \upharpoonright \mathcal{T}_1|_a, r\}$. Hence we can collect the coefficients $(k_{a_2} \cdot k_{l})$ together as $(-k_n \cdot Y_a(\sigma))$ for each $\sigma$. 

Figure 9. Typical graphs for the identity induced from the double-trace amplitude $A(1, 2, \ldots, r|2|h_1)$ with the reference order $R = \{h_1, 2\}$. Here, gluons of the trace 2 are in the cyclic order $\{d_1, d_2, \ldots, d_{|2|}\}$ and the gluon $d_{|2|}$ is assumed to be the removed one in the induced identity eq. (2.13).

(2). On another hand, two graphs with adjacent $a \in T_1$ are associated by opposite signs.

Thus $I_1$ turns to

$$I_1 = \overline{\sum_{a_2 \in 2} \sum_{a \in T_1} (k_{h_1} \cdot k_{b_2})(-1)^{F_{|x_0}} f^a \sum_{\gamma \in T_1} \sum_{\sigma \in \{2, \ldots, r-1\}} (k_a \cdot Y_{h_1}(\sigma)) A(1, \sigma, r)},$$

(4.9)

where $(-1)^{F_{|x_0}}$ denotes the sign for the graph with $a = x_0 \in T_1$. The $f^a$ for any $a \in T_1$ is fixed by (i), $f^{x_0} = 1$, (ii). $f^{x_1} = -f^{x_2}$ ($x_1, x_2 \in T_1$) if $x_1$ and $x_2$ are two adjacent nodes. Therefore, the expression in the square brackets is just the LHS of the graph-based BCJ relation $B(1|T_1, \{2, \ldots, r-1\}|r)$.

Consequently, $I_1$ is a combination of BCJ relations (as proven in [1], all graph-based BCJ relations are combination of traditional BCJ relations eq. (2.15)).

• For $I_2$, all the permutations established by the graphs Fig. 6 (b), Fig. 7 (a2) and (a3) have the form $\{h_1\} \cup \gamma$, where

$$\gamma \in \{2, \ldots, l, \{l+1, \ldots, r-1\} \cup \{a_2, KK\{2, a_2, c_2\}, c_2\}\}.\text{[Eq:gamma]}$$

(4.10)

Coefficient for each permutation $\sigma \in \{h_1\} \cup \gamma$ is collected as $(k_{a_2} \cdot k_l)(k_{h_1} \cdot Y_{h_1}(\sigma))$ (with a minus sign). Hence $I_2$ turns to

$$I_2 = \sum_{\stackrel{a_2 \neq 2}{a_2 \neq c_2}} \sum_{l \in \{1, \ldots, r-1\}} (-k_{a_2} \cdot k_l) \sum_{\gamma} \left[ \sum_{\sigma \in \{h_1\} \cup \gamma} (k_{h_1} \cdot Y_{h_1}(\sigma)) A(1, \sigma, r) \right],$$

(4.11)

in which $\gamma$ satisfies eq. (4.10). Apparently, the expression in the square brackets is just a special case (which is usually called fundamental BCJ relation) of the LHS of BCJ relation eq. (2.15). (If we consider the single node $h_1$ as a tree $T_2$, the expression can also be understood as the LHS of a graph-based BCJ relation $B(1|T_2, \{2, \ldots, r-1\}|r)$).
4.2 Example-2: the type-II identity induced from $A(1, 2, \cdots, r|2| h_1)$

When the reference order for the expansion eq. (2.1) of the double-trace EYM amplitude $A(1, 2, \cdots, r|2| h_1)$ is chosen as $R = \{h_1, 2\}$ (i.e., the trace 2 is chosen as the highest-weight element), the arbitrariness of the choice of fiducial gluon in this trace induces a type-II identity eq. (3.14). All graphs for this identity are constructed by (i) first removing the fiducial gluon from the highest-weight trace 2 (this is equivalent to the replacement Fig. 5 (c)); (ii) then applying the refined graphic rule. Supposing that all gluons in the trace 2 are in the cyclic order $d_1, d_2, \ldots, d_{|2|}$ and the gluon $d_{|2|}$ is chosen as the fiducial one, typical graphs for the type-II induced identity are then presented by Fig. 9 (a), (b), (c), (d) with the following contributions:

$$T^{(a)} = \sum_{\sigma^{(a)}} (\epsilon_{h_1} \cdot k_l) (-k_{a_2} \cdot k_{l'}) A(1, \sigma^{(a)}, r), \quad T^{(b)} = \sum_{\sigma^{(b)}} (\epsilon_{h_1} \cdot k_l) (-k_{a_2} \cdot k_{h_1}) A(1, \sigma^{(b)}, r),$$

$$T^{(c)} = \sum_{\sigma^{(c)}} (-\epsilon_{h_1} \cdot k_{a_2}) (-k_{h_1} \cdot k_l) A(1, \sigma^{(c)}, r), \quad T^{(d)} = \sum_{\sigma^{(d)}} (\epsilon_{h_1} \cdot k_{l'}) (-k_{a_2} \cdot k_l) A(1, \sigma^{(d)}, r),$$

in which,

$$\sigma^{(a)} \in \{2, \ldots, l, \{h_1\}\}$$

$$\cup \{l+1, \ldots, l', \{l' + 1, \ldots, r - 1\} \cup \{a_2 = d_j, \{d_j+1, \ldots, d_{|2|-1}\} \cup \{d_{j-1}, \ldots, d_1\}\}\},$$

$$\sigma^{(b), (c)} \in \{2, \ldots, l, \{l+1, \ldots, r - 1\} \cup \{h_1, a_2 = d_j, \{d_j+1, \ldots, d_{|2|-1}\} \cup \{d_{j-1}, \ldots, d_1\}\},$$

$$\sigma^{(d)} \in \{2, \ldots, l, \{l+1, \ldots, r - 1\}$$

$$\cup \{a_2 = d_k, \{d_k+1, \ldots, d_{|2|-1}\} \cup \{d_{k-1}, \ldots, l' = d_j, \{h_1\} \cup \{d_{j-1}, \ldots, d_1\}\}\}. \quad (4.13)$$

Figure 10. The $I_1$ and $I_2$ in eq. (4.14) can be expanded in terms of graph-based BCJ relations with graphs of the structures $T_3$ and $T_4$ respectively.
Once all graphs are summed over, we arrive the RHS of the type-II induced identity eq. (3.14) for the amplitude $A(1, 2, \ldots, r|2||h_1)$:

$$
\sum_{l,l' \in \{1,\ldots,r-1\}} \sum_{j=1}^{2|-1} T^{(a)} + \sum_{l \in \{1,\ldots,r-1\}} \sum_{j=1}^{2|-1} \left(T^{(b)} + T^{(c)}\right) + \sum_{l \in \{1,\ldots,r-1\}} \sum_{j,k=1}^{2|-1} T^{(d)}, \tag{4.14}
$$

where contributions of all graphs of the form Fig. 9 (a), (b) and (c), (d) were collected as $I_1$ and $I_2$ respectively. Now we analyze $I_1$ and $I_2$ separately.

- For $I_1$, the summation over $j = 1, \ldots, |2| - 1$ is nothing but just the summation over all nodes $a_2$ belonging to the tree graph $T_3$ in Fig. 10 (a). For a graph with a given $a_2 \in T_3$ and a given $l \in \{1, \ldots, r-1\}$, all possible permutations satisfy $\sigma \in (T_3|a_2) \cup \gamma$, where $\gamma$ belongs to $\{2, \ldots, l, \{h_1\} \cup \{l+1, \ldots, r-1\}\}$ ($l$ can be the gluon 1) and $T_3|a_2$ denotes the relative orders of nodes in $T_3$ when $a_2$ is considered as the leftmost one. Following a similar discussion with the example-1, we find that the coefficients $(k_{a_2} \cdot k_l)$ ($l \in \{1, \ldots, r-1\}$ or $l = h_1$) corresponding to a same $\sigma \in (T_3|a_2) \cup \gamma$ are collected as $(k_{a_2} \cdot Y_{a_2}(\sigma))$ and two graphs with choosing adjacent $a_2 \in T_3$ have opposite signs. Then $I_1$ in eq. (4.14) is expressed by

$$
I_1 = \sum_{l \in \{1,\ldots,r-1\}} \sum_{\gamma \in \{2,\ldots,l,h_1\}} (\epsilon_{h_1} \cdot k_l) (k_{h_1} \cdot k_{b_2})(-1)^{|\gamma|\times_0} \times \left[ \sum_{a_2 \in T_3} f^{a_2}_{\sigma} \sum_{\sigma \in (T_3|a_2) \cup \gamma} (-k_{a_2} \cdot Y_{a_2}(\sigma)) A(1, \sigma, r) \right]. \tag{4.15}
$$

Here $(-1)^{|\gamma|\times_0}$ is the sign for choosing $a_2 = x_0 \in T_3$ and $f^{x_0} = 1$. Graphs with two adjacent $a$ are associated with opposite signs. Obviously, the expression in the square brackets is nothing but (upto a total sign) $B(1|T_3, \gamma|r)$. As a result, $I_1$ is a combination of BCJ relations.

- The $I_2$ in eq. (4.14) can be obtained following a parallel discussion with $I_1$ but replacing the tree graph $T_3$ (Fig. 10 (a)) by $T_4$ (Fig. 10 (b)) for $d_j \in \{d_1, \ldots, d_{|2|-1}\}$ and replacing $\gamma$ by $\{1, \ldots, r-1\}$:

$$
I_2 = \sum_{j=1}^{2} (\epsilon_{h_1} \cdot k_{d_j})(-1)^{|\gamma|\times_0} \left[ \sum_{a_2 \in T_4} f^{a_2}_{\sigma} \sum_{\sigma \in (T_4|a_2) \cup \gamma} (-k_{a_2} \cdot Y_{a_2}(\sigma)) A(1, \sigma, r) \right]. \tag{4.16}
$$

where $x_0 \in T_4$ and $f^{x_0} = 1$. Upto a total sign, the expression in the square brackets is just the LHS of graph-based BCJ relation $B(1|T_4, \{2, \ldots, r-1\}|r)$, where the tree graph $T_4$ for a given $j$ is Fig. 10 (b). Again, we conclude that $I_2$ is a combination of BCJ relations.
4.3 Comments on examples in section 4

To close this section, we comment on some crucial features of the examples:

(i). Expanding traces which are not the highest-weight element by standard basis In example-1, the trace $222$ is not the highest-weight element in the reference order $R$. It plays as an internal trace in Fig. 6 (a) and a starting trace in Fig. 6 (b), (c). In the latter case, the trace $222$ was already expressed by standard basis, while in the former case, it is expanded according to the relation eq. (4.5). The fixed node for the expansion of the trace $222$ in Fig. 6 (a) was conveniently chosen as the fiducial node in Fig. 6 (b) and (c).

(ii). Skeletons and components We define skeletons by removing all type-3 lines from the graphs where all traces, except the highest-weight element in $R$ (if it is a trace), are already expressed by the standard basis. Since each graph defined by the refined graphic rule is a connected tree graph, its skeleton must be a disconnected one. Each maximally connected subgraph of a skeleton is called a component. A typical skeleton in example-1 is given by Fig. 11 (a) (for $b_2 \in 222, b_2 \neq c_2$) which consists of three components. Skeletons in example-2 have two distinct structures, Fig. 11 (b) and (c), each of which has two components. From the examples, we can see any skeleton must has at least two components which involve the highest-weight element (graviton or trace) and the trace $1$ respectively.
(iii). **The final upper and lower blocks** Any graph in the examples can be reproduced by connecting a type-3 line between the final upper and lower blocks which are two mutually disjoint connected subgraphs containing the highest-weight element and the trace 1 respectively. In example-2, each of the skeletons Fig. 11 (b) and (c) already consists of only two disjoint connected subgraphs $U$ and $L$ which serve as the final upper and lower blocks. In example-1, there are three components in a skeleton Fig. 11 (a). A typical configuration of the final upper and lower blocks is constructed when we connect any $b_2 \in 2, b_2 \neq c_2$ in Fig. 11 (a) to either (i) $h_1$ (see Fig. 12 (a)) or (ii) an element in $\{1, \ldots, r-1\}$ via a type-3 line (see Fig. 12 (b)).

(iv). **Physical and spurious graphs** For a given configuration of the final upper and lower blocks $U$ and $L$, we can connect two nodes $x \in U$ and $y \in L$ via a type-3 line. Then a fully connected graph is constructed. In example-2, the graphs Fig. 9 (a), (b) (and Fig. 9 (c), (d)) are reproduced by connecting the final upper and lower blocks in Fig. 11 (b) (and Fig. 11 (c)) via a type-3 line. Similarly, the graphs Fig. 7 (a1), Fig. 6 (c) (and Fig. 7 (a2), Fig. 6 (b)) in example-1 are constructed from Fig. 12 (a) (and (b)). All the graphs Fig. 6 (b), (c), Fig. 7 (a1), (a2) and Fig. 9 (a)-(d) are graphs directly defined by the refined graphic rule (in standard basis). These graphs are called physical graphs. The spurious graph Fig. 7 (a3), which is not defined by refined graphic rule, can be reproduced from either Fig. 12 (a) (connecting $b_2 \in U$ with $l \in L$) or Fig. 12 (b) (connecting $h \in U$ with $b_2 \in L$). The former case is associated with a minus to ensure the cancellation of the same skeleton constructed by distinct ways. Therefore, for any skeleton, the sum of all physical graphs can be given by (1) summing over all possible configurations of the final upper and lower blocks $U$, $L$, (2) for a given $U$ and $L$, connecting two nodes $x \in U$ and $y \in L \setminus \{r\}$ via a type-3 line and summing over all possible choices of $x, y$.

(v). **Induced identities as combinations of graph-based BCJ relations** A crucial observation on the examples is that the sum over all graphs corresponding to a same configuration of the final upper and lower blocks (i.e. the $I_1$ and $I_2$ in either example) are combinations of graph-based BCJ relations (thus traditional BCJ relations).

5 Induced identities as combinations of BCJ relations: the general pattern

In this section, we extend our discussions to (both type-I and type-II) identities induced from arbitrary multi-trace EYM amplitudes. Inspired by the examples, we rewrite the sum over all physical graphs (graphs directly defined by the refined graphic rule) $\mathcal{F}$ in eq. (3.13) and eq. (3.14) by (1) summing over skeletons $\mathcal{F}'$ where traces (except the highest-weight element if it is a trace) are expanded by standard basis, (2) summing over all physical graphs $\mathcal{F}$ for a given skeleton $\mathcal{F}'$. The RHS of the induced identities eq. (3.13) and eq. (3.14) are then uniformly arranged as

$$\sum_{\mathcal{F}'} \mathcal{P}^{[\mathcal{F}']} \left[ \sum_{\mathcal{F} \supseteq \mathcal{F}'} (-)^{\mathcal{F}} \mathcal{K}^{[\mathcal{F} \setminus \mathcal{F}']} \sum_{\sigma} A(1, \sigma, r) \right] \ \text{[Eq:InducedID1]}$$

(5.1)
Figure 13. (a): Type-IA component is a component containing only gravitons. The kernel is defined by the type-1 line between nodes $a$ and $b$. (b): Type-IB component is a component with a gluon trace in it. The kernel in (b) is defined by the type-4 line which is connected to the node $d$. Although a trace in the reference order $R$ is considered as a single object, we let the gluon $c$ in (b) (i.e. the fixed node of a trace in standard basis) carry the weight of the full trace and $c$ is considered as the highest weight element in the full trace. In each graph, we suppose that the highest weight nodes in the regions $\mathcal{A}$ and $\mathcal{B}$ are correspondingly $x$ and $y$. If the the weight of $x$ is higher than that of $y$, $\mathcal{A}$ and $\mathcal{B}$ are respectively the top and the bottom sides. Contrarily, if the weight of $y$ is higher than that of $x$, $\mathcal{A}$ becomes the bottom side while $\mathcal{B}$ the top.

Figure 14. The component with the highest weight element in it is defined as (1) the type-IIA component (the graph (a)) if the highest weight element is a graviton (2) the type-IIB component (the graph (b)) if the highest weight element is a gluon trace. The component with the trace $1\bar{1}1$ is defined as the type-III component (the graph (c)).

where $\mathcal{P}^{[\mathcal{F}]}$ is the kinematic factor (which is a product of factors of the form $\epsilon \cdot \epsilon$ and $\epsilon \cdot k$) for the skeleton $\mathcal{F}'$ and $\mathcal{K}^{[\mathcal{F}', \mathcal{F}]}$ stands for the product of all type-3 lines for the graph $\mathcal{F}$. The overall signs $(-1)^{|t_1, a_1, b_1|}$ which are introduced by all traces $t_i$ (except the fiducial trace $t_0$ if it is the highest-weight element in the reference order because this trace is already replaced according to Fig. 5 (c)) are absorbed by the summation notation $\tilde{\sum}_{\mathcal{F}'}$. The factor $(-)^{\mathcal{F}}$ is the remaining sign associated to this graph. All amplitudes for the graph $\mathcal{F}$ are those with the permutations $1, \sigma^\mathcal{F}, r$, where $\sigma^\mathcal{F}$ denotes permutations established by $\mathcal{F}$. Once the relationship between the expression in the square brackets eq. (5.1) (for any skeleton $\mathcal{F}'$) and BCJ relations have been founded, we know the full relationship between the induced identities eqs.
(3.13), (3.14) and BCJ relations. In the remaining part of this section, we classify components and show the proper construction rule for the final upper and lower blocks (as mentioned in the previous section) $\mathcal{U}$ and $\mathcal{L}$ which are two mutually disjoint connected graphs. We then show that the expression in the square brackets $I[\mathcal{F}']$ for a given skeleton $\mathcal{F}'$ can be written as

$$I[\mathcal{F}'] = \sum_{\mathcal{U} \oplus \mathcal{L}} \kappa(\mathcal{U} \oplus \mathcal{L} | \mathcal{F}') I[\mathcal{F}' \mid \mathcal{U}, \mathcal{L}], \quad \text{[Eq:InducedID4]}$$

where $I[\mathcal{F}' \mid \mathcal{U}, \mathcal{L}]$ for a given configuration of $\mathcal{U}$ and $\mathcal{L}$ is a combination of the graph-based BCJ relations. The $\mathcal{U} \oplus \mathcal{L}$ denotes the union of the mutually disjoint $\mathcal{U}$ and $\mathcal{L}$. For conciseness, the proof of the general pattern will not be presented. We claim that this pattern can be proven following a parallel discussion in the single-trace case [1] but with the enlarged classes of components.

5.1 Classification of components

As shown by our previous work [1], components for identities induced from single-trace amplitudes can be classified according to their internal structures. This classification is extended to multi-trace case when more internal structures are incorporated:

- **Type-I components** Components containing a type-1 line and possible type-2 lines pointing towards the two ends of the type-1 line are called type-IA components (see Fig. 13 (a)). Components containing a gluon trace in the standard basis, with possible type-2 lines pointing to the nodes of this trace, are called type-IB components (see Fig. 13 (b)). The latter is a new feature for identities induced from multi-trace amplitudes, while the former already exists in the discussions on identities induced from single-trace amplitudes (see [1]). The kernel is defined by the type-1 line in a type-IA component and the type-4 line (i.e. the dashed arrow line) that is attached to the ending node of the trace in standard basis in a type-IB component. For a given reference order, any type-IA and -IB component is divided into two parts by the kernel: the part involving the highest-weight node (although a trace is considered as a single object in the reference order, the fixed node $c$ is always considered as the highest-weight node of this trace and it carries the weight of the full trace in the reference order) of this component is called the top side, while the opposite side is called the bottom side.

- **Type-II and Type-III components** In any graph for the induced identities, there always exists a component that contains the highest-weight element (graviton or gluon trace) in the reference order. If the highest-weight element is a graviton (gluon trace) this component is called the type-IIA (type-IIIB) component (see Fig. 14 (a), (b)). The component involving the trace 1 is called the type-III component (see Fig. 14 (c)).

Apparently, both type-II and type-III components appear once and only once in a skeleton. Type-II components may also be involved in a skeleton. All graphs corresponding to a given skeleton are reproduced
5.2 The construction of the final upper and lower blocks

Having classified the components, we are now ready for constructing all configurations of the final upper and lower blocks \( U, L \) in eq. (5.2):

- **Step-1** For any skeleton \( F' \), we define the reference order \( R_C \) of all type-IA and -IB components by the relative order of the highest-weight nodes therein. We also define the upper block \( U \) and lower block \( L \) as the components containing the highest-weight element (graviton or trace) and the trace \( 1 \) respectively. At the beginning, the upper and the lower blocks are nothing but the type-II and the type-III components.

- **Step-2** Supposing the reference order of components is \( R_C = \{ C_1, C_2, \ldots, C_N \} \), pick out the highest-weight component \( C_N \) as well as arbitrary components \( C_{a_1}, C_{a_2}, \ldots, C_{a_i} \) (the relative order of these components is not necessary the same relative order in \( R_C \)). Construct a chain of components towards either the upper block or the lower block as follows

  \[ \mathcal{CH} = [(C_N)_t, (C_N)_b \leftrightarrow (C_{a_i})_{t(a or b)}, (C_{a_i})_{b(or t)} \leftrightarrow \cdots \leftrightarrow (C_1)_{t(a b)}, (C_1)_{b(or t)} \leftrightarrow U \text{ or } L \setminus \{ r \}] \]  

  Here the subscripts \( t \) and \( b \) denote the top and bottom sides of a component respectively. The comma between the top and bottom sides stands for the kernel of a (type-IA or -IB) component. The double arrow line ‘\( \leftrightarrow \)’ between two components stands for the type-3 line (i.e. \( k \cdot k \)), which connects any two nodes belonging to the corresponding regions. For example, if the chain of components has the form \([ (C_N)_t, (C_N)_b \leftrightarrow (C_{a_i})_{t}, (C_{a_i})_{b} \leftrightarrow \cdots ] \), the two ends \( x \) and \( y \) of the type-3 line between the components \( C_N \) and \( C_{a_i} \) must belong to \( (C_N)_b \) and \( (C_{a_i})_{t} \) respectively. After this step, we redefine the reference order of components as well as the upper and lower blocks by:

  \[ R_C \rightarrow R'_C = R_C \setminus \{ C_N, C_{a_1}, \ldots, C_1 \} = \{ C'_1, C'_2, \ldots, C'_N \}, \]

  \[ U \rightarrow U' = U \cup \{ C_N, C_{a_1}, \ldots, C_1 \}, \quad L \text{ (if } \mathcal{CH} \text{ was attached to } U) \]

  \[ L \rightarrow L' = L \cup \{ C_N, C_{a_1}, \ldots, C_1 \}, \quad U \text{ (if } \mathcal{CH} \text{ was attached to } L). \]  

- **Step-3** Repeating step-2 with the new defined \( R'_C, U \) and \( L \) until the ordered set \( R \) becomes empty, we get a graph with only two mutually disjoint subgraphs: the final upper and lower blocks.

All possible configurations of the final upper and lower blocks are produced by the above steps. When we consider identities induced from single-trace amplitudes, we do not have any type-IB or type-IIB component in a skeleton. Then this construction rule becomes the rule in [1]. The final upper and lower blocks of the examples in section 4 are precisely reproduced by this rule.

---

*Notations here are slightly different from those in [1]."
5.3 Induced identities as combinations of graph-based BCJ relations

For a given configuration of the final upper and lower blocks $\mathcal{U}$ and $\mathcal{L}$, a fully connected graph is produced by connecting a type-3 line between arbitrary two nodes $x \in \mathcal{U}$ and $y \in \mathcal{L} \setminus \{r\}$. Such a graph is either a physical or a spurious one. As shown by Fig. 15, a spurious graph has the structure in which the path starting from the highest-weight element and ending at the root passes through single sides of some type-IA or/and type-IB components (for convenience, such a component is called a spurious component).

Now let us determine the sign $(-)^F$ for a graph $F$. 

(i). As proposed in [1], we associate a sign $(-1)^S(\mathcal{U}_x)$ to a graph where $S(\mathcal{U}_x)$ is the number of spurious components living in the final upper block $\mathcal{U}$. Apparently, $S(\mathcal{U}_x)$ in general depends on the choice of the node $x \in \mathcal{U}$. Following a similar proof with [1], we find that each spurious graph must appear twice (corresponding to distinct configurations of the final upper and lower blocks) with opposite signs (This fact is also confirmed by the example-1). Thus all spurious graphs cancel in pairs. 

(ii). Another sign should be taken into account is introduced by the relation eq. (4.5). Particularly, if the arrow of the kernel of a type-IB component is pointing away from the root (i.e. the second term of eq. (4.5)), this component should be dressed by an extra minus. We use $\text{Tr}(\mathcal{U}_x, \mathcal{L})$ to denote the number of such type-IB components, this sign becomes $(-1)^{\text{Tr}(\mathcal{U}_x, \mathcal{L})}$. Since the choice of $y \in \mathcal{L} \setminus \{r\}$ does not affect the direction of arrows in the final lower block $\mathcal{L}$, this number only depends on $\mathcal{U}$, $\mathcal{L}$ and the choice of $x \in \mathcal{U}$. 

(iii). For any graph, the total number $\mathcal{N}(\mathcal{U}_x) + \mathcal{N}(\mathcal{L}) + 1$ of arrows pointing away from the root induces the third sign $(-1)^{\mathcal{N}(\mathcal{U}_x) + \mathcal{N}(\mathcal{L}) + 1}$. Here, $\mathcal{N}(\mathcal{U}_x)$ and $\mathcal{N}(\mathcal{L})$ count the numbers in $\mathcal{U}$ (for $x \in \mathcal{U}$) and $\mathcal{L}$ respectively. The extra minus is caused by the type-3 line between $\mathcal{U}$ and $\mathcal{L}$. To sum up, the sign for any (physical or spurious) graph is given by

$$(-)^F = (-1)^S(\mathcal{U}_x) + \text{Tr}(\mathcal{U}_x, \mathcal{L}) + \mathcal{N}(\mathcal{U}_x) + \mathcal{N}(\mathcal{L}) + 1. \quad [\text{Eq:Sign}]$$

For a physical graph, there is no spurious component, we have $S(\mathcal{U}_x) = 0$. For identities induced from
When all factors $\mathcal{P}[F']$ are extracted from (both physical and spurious) graphs $F$ with the same skeleton $F'$, the remaining contributions of all $F \supset F'$ then becomes eq. (5.2). If all the $(k \cdot k)$ factors (i.e. the factor $\mathcal{K}[\mathcal{P}, F']$) for a given configuration of the final upper and lower blocks are also extracted out, we further get $I[F' \mid \mathcal{U}, \mathcal{L}]$ (in eq. (5.2)), which is expressed by combination of amplitudes associated with a factor $k_x \cdot k_y$ ($x \in \mathcal{U}, y \in \mathcal{L} \setminus \{r\}$) and a sign eq. (5.5). Permutations in amplitudes are those established by the tree graphs. Therefore, $I[F' \mid \mathcal{U}, \mathcal{L}]$ can be written as

$$I[F' \mid \mathcal{U}, \mathcal{L}] = \sum_{x \in \mathcal{U}, y \in \mathcal{L} \setminus \{r\}} (-)^{F}(k_x \cdot k_y) \left[ \sum_{\gamma \in \mathcal{L} \setminus \{1,r\}} \sum_{\zeta} \sum_{\sigma} A \left( 1, \sigma \in (\zeta \cup \gamma) \right) \left| _{\sigma^{-1}(y) < \sigma^{-1}(x), r} \right. \right]$$

where the sum over all permutations established by a tree graph $F$ (for the final upper and lower blocks $\mathcal{U}, \mathcal{L}$ and given $x \in \mathcal{U}, y \in \mathcal{L}$) split into three summations: (i). the sum over all relative orders $\gamma$ established by the final lower block $\mathcal{L}$ where 1 is considered as the leftmost node, (ii). the sum over all
relative orders $\zeta$ established by the final upper block $\mathcal{U}$ (when $x \in \mathcal{U}$ is considered as the leftmost node),
(iii). the sum over all those permutations $\sigma \in \zeta \sqcup \gamma$ such that $\sigma^{-1}(y) < \sigma^{-1}(x)$ is satisfied.

To relate the $I[\mathcal{F}^I | \mathcal{U}, \mathcal{L}]$ with the graph-based BCJ relation eq. (2.16), we rearrange eq. (5.6) based on the following observations: (i). For a fixed $x \in \mathcal{U}$ and a given permutation $\sigma$, only those graphs with $\sigma^{-1}(y) < \sigma^{-1}(x)$ contribute nonzero factors $k_x \cdot k_y$ (with the same sign $(-)^F$). Thus the total factor for permutation $\sigma$ is collected as $k_x \cdot Y_x(\sigma)$ with the same sign. (ii). As shown by Fig. 16, two graphs with $x = x_1$ and $x = x_2$, where $x_1, x_2 \in \mathcal{U}$ are two adjacent nodes, must be associated with opposite signs. Then eq. (5.6) is finally written as

$$I[\mathcal{F}^I | \mathcal{U}, \mathcal{L}] = (\sigma^{-1})^{x_0} \sum_{\gamma \in \mathcal{L}[1 \setminus \{1,r\}] \setminus x \in \mathcal{U}} \sum_{\zeta \subseteq \mathcal{V} | x} \sum_{\sigma \in \zeta \sqcup \gamma} (k_x \cdot Y_x(\sigma)) A(1, \sigma \in \zeta \sqcup \gamma, r),$$

where the summation over $\gamma$ was extracted out because it is independent of the choice of $x \in \mathcal{U}$. The factor $(-)^{\mathcal{F}^I x_0}$ is the sign eq. (5.5) for an arbitrary $x_0 \in \mathcal{U}$. The sign $f_x$ is defined by (i). 1 for $x = x_0$, (ii). $f_{x_1} = -f_{x_2}$ for adjacent $x_1$ and $x_2$ ($x_1, x_2 \in \mathcal{U}$). Therefore, the expression in the square brackets in eq. (5.7) is nothing but the LHS of the graph-based BCJ relation: $B(1 | \mathcal{U}, \gamma | r)$ where $\mathcal{U}$ is the final upper block.

6 Conclusions

In this paper, we have shown that the identities eq. (3.13) and eq. (3.14) induced from arbitrary tree level multi-trace EYM amplitude can be expressed as a combination of graph based BCJ relations (thus traditional BCJ relations). To extend the discussions on identities induced from single-trace amplitudes [1] to arbitrary multi-trace case, we enlarged the set of graphs by including more structures (i.e., more line styles and more types of components) in the refined graphic rule. The relation eq. (4.5) which expands internal traces by standard basis played a key role in this approach.

There are several directions that deserve further study: (i). How to understand the induced identities from string theory? A relationship between string theory and the Cachazo-He-Yuan [27–30] version of field theory was established in [31, 32]. This work may provide an effective method for the study at string theory level. (ii). In [33], a unified web of expansions of amplitudes was established, which inspires that the induced identities exist in various theories. It is worth studying identities induced from other theories. (iii). Loop-level extension is another interesting direction.

Acknowledgments

The authors are grateful to Chih-Hao Fu, Song He, Xiaodi Li, Hui Luo, Gang Yang, Yihong Wang and Yong Zhang for helpful discussions or/and valuable comments. This work is supported by NSFC under Grant Nos. 11875206, 11847309, Jiangsu Ministry of Science and Technology under contract BK20170410 as well as the “Fundamental Research Funds for the Central Universities”.

- 26 -
A Proof of eq. (4.5)

To prove eq. (4.5), we focus on a term with a given choice of \( \{a_i, b_i\} \subset i \) in eq. (4.5) and prove the following stronger relation:

\[
\frac{(A.1)}{\textit{SplittingTraces2}}
\]

where the \( \mathcal{T} \) denotes a tree structure which is attached to the node \( b_i \). In fact, all these nodes in the trace can be attached by arbitrary tree structures and the relation eq. (A.1) still holds. Once all \( \{a_i, b_i\} \subset i \) pairs are summed over, we arrive the relation eq. (4.5).

Suppose the gluons in trace \( i \) are in the cyclic order \( d_1, d_2, \ldots, d_{|i|} \). Without loss of generality, \( a_i \) and \( b_i \) are respectively chosen as \( a_i = d_j, \ b_i = d_{|i|} \). Then the permutations \( \{a_i, b_i \in \mathbb{K}[i, a_i, b_i]\} \) on the LHS of eq. (A.1) are explicitly displayed by

\[
\frac{(A.2)}{\textit{APPLHS}}
\]

with a sign \((-1)^{|i| a_i, b_i|} = (-1)^{j-1}\) (the overall sign coming from \( \mathcal{T}|_x \) has been neglected). Here \( x \) denotes the nearest to \( b_i \) node in \( \mathcal{T} \) and \( \mathcal{T}|_x \) are the permutations established by the tree \( \mathcal{T} \). The notation \( b_i \prec x \) means the position of \( b_i \) in the permutation is less than that of \( x \). The second line of eq. (A.2), where we shuffled \( \mathcal{T}|_x \) with the full trace and required the node \( b_i \) is always on the left of the node \( x \), is apparently equivalent to the first line.

On the RHS of eq. (A.1), we assume that the \( c_i \) is chosen as \( d_l (l > j) \). The case with \( l < j \) follows from a similar discussion. Then permutations in the first term on the RHS of eq. (4.5) is given by

\[
\frac{(A.3)}{\textit{APPRHS1}}
\]
with the sign $(-1)^{i_1 a_1 c_1} = (-1)^{|i|-l+j-1}$. The permutations in the second term of the RHS of eq. (4.5) are given by the following two steps:

- First shuffle the gluons in the trace $i$ so that $b_i = d_{|i|}$ and $c_i = d_l$ become the two ends of the trace. Then the permutations $\beta \in \mathbb{K}\mathbb{K}[i, b_i, c_i]$ are explicitly given by

$$\beta \in \{d_{|i| - 1}, \ldots, d_{l+1}\} \sqcup \{d_1, \ldots, d_{l-1}\}. \quad \text{[Eq:AppBeta1]} \quad (A.4)$$

which is associated by a sign $(-1)^{|i| b_i c_i} = (-1)^{|i|-l-1}$. These permutations can be classified according to the relative orders between the $a = d_j \ (d_j \in \{d_1, \ldots, d_{l-1}\})$ and elements in the set $\{d_{|i| - 1}, \ldots, d_{l+1}\}$:

$$\beta^{(1)} \in \{d_1, \ldots, d_{j-1}, d_j, d_{j+1}, \ldots, d_{l-1}\} \sqcup \{d_{|i| - 1}, \ldots, d_{l+1}\}$$

$$\beta^{(2)} \in \{d_{j+1}, \ldots, d_{|i| - 1}, d_j, d_{j+1}, \ldots, d_{l-1}\} \sqcup \{d_{|i| - 2}, \ldots, d_{l+1}\}$$

$$\ldots$$

$$\beta^{(q)} \in \{d_{j+1}, \ldots, d_{|i| - 1}, d_j, d_{j+1}, \ldots, d_{l-1}\} \sqcup \{d_{|i| - q}, \ldots, d_{l+1}\}$$

$$\ldots$$

$$\beta^{(|i|-l)} \in \{d_{j}, d_{j+1}, \ldots, d_{l-1}\} \sqcup \{d_{|i| - l + 1}, d_j, d_{j+1}, \ldots, d_{l-1}\}. \quad \text{[Eq:AppBeta2]} \quad (A.5)$$

- For any given permutation $\beta^{(q)}$ in eq. (A.5), the permutations $\sigma^{(q)}$ in second term on the RHS of eq. (4.5) are obtained by shuffling the two branches attached to node $a_i = d_i$ together:

$$\sigma^{(q)} \in \left\{ a_i = d_j, \{d_{j+1}, \ldots, d_{l-1}\} \sqcup \{d_{|i| - q}, \ldots, d_{l+1}\}, c_i = d_l \right\}$$

$$\sqcup \left\{ d_{j-1}, \ldots, d_1 \right\} \sqcup \{d_{|i| - q}, \ldots, d_{|i| - 1}\}, b_i = d_{|i|} \right\} \sqcup T_{|i|}. \quad \text{[Eq:AppSigma]} \quad (A.6)$$

with a sign $(-1)^{|i|-l-1}(-1)^{N_{a_i}^{(q)}}$. Here $N_{a_i}^{(q)}$ denotes the number of arrows, which point away from root, in the trace $i$.

According to the relative orders between $d_{|i| - q}$ and $d_{|i| - q + 1}$, permutations $\sigma^{(q)}$ in eq. (A.6) splits into $\sigma^{(q)}_A = \sigma^{(q)}|_{d_{|i| - q} \prec d_{|i| - q + 1}}$ and $\sigma^{(q)}_B = \sigma^{(q)}|_{d_{|i| - q} \succ d_{|i| - q + 1}}$. It is easy to see

$$\sigma^{(q)}_A = \sigma^{(q+1)}_B, \quad N_{a_i}^{(q)} = N_{a_i}^{(q+1)} + 1. \quad \text{[Eq:AppSigma2]} \quad (A.7)$$

Hence the contributions from $\sigma^{(q)}_A$ and $\sigma^{(q+1)}_B$ for all $q = 1, \ldots, |i| - l - 1$ must cancel one another. The
remaining nonzero terms are the two boundaries

\[ \sigma^{(1)}_{B} \in \left\{ \{a_i = d_j, \{d_{j+1}, \ldots, d_{l-1}\} \cup \{d_{|i|-1}, \ldots, d_{l+1}\}, c_i = c_l \right\} \]

\[ \cup \{ d_{j-1}, \ldots, d_1, b = d_{|i|} \} \right|_{d_{|i|-1} > d_{|i|} \cup \mathcal{T}|x} \bigg|_{b_i < x} \]

\[ = \left\{ \{a_i = d_j, \{d_{j+1}, \ldots, d_{l-1}\} \cup \{d_{j-1}, \ldots, d_1, b_i = d_{|i|} \} \right\} \cup \{ d_{|i|}, d_{|i|-1}, \ldots, d_{l+1}\}, c_i = c_l \} \cup \mathcal{T}|x \bigg|_{d_{|i|-1} > d_{|i|} \cup \mathcal{T}|x} \bigg|_{b_i < x} . \]

\[ \sigma^{(|i|-l)}_{A} \in \left\{ \{a_i = d_j, \{d_{j+1}, \ldots, d_{l-1}\} \cup \{d_{j-1}, \ldots, d_1, b_i = d_{|i|} \} \right\} \cup \{ d_{|i|}, d_{|i|-1}, \ldots, d_{l+1}\}, c_i = c_l \} \cup \mathcal{T}|x \bigg|_{b_i < x} . \]

\[ = \left\{ \{a_i = d_j, \{d_{j-1}, \ldots, d_1, b_i = d_{|i|} \} \right\} \cup \{ d_{j+1}, \ldots, d_{|i|-1}, c_i = c_l \} \cup \mathcal{T}|x \bigg|_{d_{|i|-1} < d_{|i|} \cup \mathcal{T}|x} \bigg|_{b_i < x} . \]

(\text{Eq:AppSigma1}) \quad (A.8)

with the signs \((-1)^{|i|-l-1}(-1)^{\lambda_a^{(i)}} = (-1)^{|i|-l-1+j} \) and \((-1)^{|i|-l-1}(-1)^{\lambda_a^{(|i|-l)}} = (-1)^j \) respectively. An extra minus which is introduced from the second term of eq. (4.5) must also be taken into account. Hence the permutations \(\sigma^{(1)}_{B}\) are same with the permutations in eq. (A.3), with an opposite sign. As a result, they cancel with each other. The remaining permutations \(\sigma^{(|i|-l)}_{A}\) are nothing but the permutations eq. (A.2) with the corrected sign. Therefore, we have proven the relation eq. (4.5) for a given \(\{a_i, b_i\}\). After summing over all possible choices of the \(\{a_i, b_i\}\) pairs, the proof of eq. (4.5) is completed.

**Comments on the proof:** In the above proof, the tree structure \(\mathcal{T}|x\) is always treated separately from the trace in each step (i.e. it is shuffled with the full trace with a proper constraint, as shown in eqs. (A.2), (A.3), (A.6), (A.8)). Thus manipulations on nodes inside the trace are independent of \(\mathcal{T}|x\). This observation allows us to generalize eq. (4.5) to cases where more tree structures are attached to nodes in the trace \(i\) straightforwardly.

**References**

[1] L. Hou and Y.-J. Du, *A graphic approach to gauge invariance induced identity*, JHEP **05** (2019) 012, [arXiv:1811.12653].

[2] Y.-J. Du, B. Feng, and F. Teng, *Expansion of All Multitrace Tree Level EYM Amplitudes*, arXiv:1708.04514.

[3] C.-H. Fu, Y.-J. Du, R. Huang, and B. Feng, *Expansion of Einstein-Yang-Mills Amplitude*, JHEP **09** (2017) 021, [arXiv:1702.08158].

[4] M. Chiodaroli, M. Gunaydin, H. Johansson, and R. Roiban, *Explicit Formulae for Yang-Mills-Einstein Amplitudes from the Double Copy*, JHEP **07** (2017) 002, [arXiv:1703.00421].

[5] F. Teng and B. Feng, *Expanding Einstein-Yang-Mills by Yang-Mills in CHY frame*, JHEP **05** (2017) 075, [arXiv:1703.01269].
[6] S. Stieberger and T. R. Taylor, New relations for Einstein-Yang-Mills amplitudes, Nucl. Phys. B913 (2016) 151–162, [arXiv:1606.09616].

[7] D. Nandan, J. Plefka, O. Schlotterer, and C. Wen, Einstein-Yang-Mills from pure Yang-Mills amplitudes, JHEP 10 (2016) 070, [arXiv:1607.05701].

[8] L. de la Cruz, A. Kniss, and S. Weinzierl, Relations for Einstein-Yang-Mills amplitudes from the CHY representation, Phys. Lett. B767 (2017) 86–90, [arXiv:1607.06036].

[9] O. Schlotterer, Amplitude relations in heterotic string theory and Einstein-Yang-Mills, JHEP 11 (2016) 074, [arXiv:1608.00130].

[10] B. Feng, X. Li, and K. Zhou, Expansion of EYM theory by Differential Operators, arXiv:1904.05997.

[11] R. Britto, F. Cachazo, and B. Feng, New recursion relations for tree amplitudes of gluons, Nucl. Phys. B715 (2005) 499–522, [hep-th/0412308].

[12] R. Britto, F. Cachazo, B. Feng, and E. Witten, Direct proof of tree-level recursion relation in Yang-Mills theory, Phys. Rev. Lett. 94 (2005) 181602, [hep-th/0501052].

[13] Y.-J. Du and Y. Zhang, Gauge invariance induced relations and the equivalence between distinct approaches to NLSM amplitudes, JHEP 07 (2018) 177, [arXiv:1803.01701].

[14] Y.-J. Du and C.-H. Fu, Explicit BCJ numerators of nonlinear sigma model, JHEP 09 (2016) 174, [arXiv:1608.05846].

[15] J. J. M. Carrasco, C. R. Mafra, and O. Schlotterer, Abelian Z-theory: NLSM amplitudes and α' -corrections from the open string, JHEP 06 (2017) 093, [arXiv:1608.02569].

[16] Y.-J. Du and F. Teng, BCJ numerators from reduced Pfaffian, JHEP 04 (2017) 033, [arXiv:1703.05717].

[17] R. Kleiss and H. Kuijf, Multi-Gluon Cross-sections and Five Jet Production at Hadron Colliders, Nucl. Phys. B312 (1989) 616–644.

[18] Z. Bern, J. J. M. Carrasco, and H. Johansson, New Relations for Gauge-Theory Amplitudes, Phys. Rev. D78 (2008) 085011, [arXiv:0805.3993].

[19] N. Arkani-Hamed, L. Rodina, and J. Trnka, Locality and Unitarity from Singularities and Gauge Invariance, arXiv:1612.02797.

[20] G. Chen, H. Johansson, F. Teng, and T. Wang, On the kinematic algebra for BCJ numerators beyond the MHV sector, arXiv:1906.10683.

[21] N. E. J. Bjerrum-Bohr, P. H. Damgaard, and P. Vanhove, Minimal Basis for Gauge Theory Amplitudes, Phys. Rev. Lett. 103 (2009) 161602, [arXiv:0907.1425].

[22] Y.-X. Chen, Y.-J. Du, and B. Feng, A Proof of the Explicit Minimal-basis Expansion of Tree Amplitudes in Gauge Field Theory, JHEP 02 (2011) 112, [arXiv:1101.0009].

[23] R. Huang, Y.-J. Du, and B. Feng, Understanding the Cancelation of Double Poles in the Pfaffian of CHY-formalism, arXiv:1702.05840.

[24] X. Gao, S. He, and Y. Zhang, Labelled tree graphs, Feynman diagrams and disk integrals, JHEP 11 (2017) 144, [arXiv:1708.08701].
[25] S. He, G. Yan, C. Zhang, and Y. Zhang, *Scattering Forms, Worldsheet Forms and Amplitudes from Subspaces, JHEP 08* (2018) 040, [arXiv:1803.11302].

[26] C. S. Lam, *Pfaffian Diagrams for Gluon Tree Amplitudes*, [arXiv:1808.07575].

[27] F. Cachazo, S. He, and E. Y. Yuan, *Scattering equations and Kawai-Lewellen-Tye orthogonality, Phys. Rev. D90* (2014), no. 6 065001, [arXiv:1306.6575].

[28] F. Cachazo, S. He, and E. Y. Yuan, *Scattering of Massless Particles in Arbitrary Dimensions, Phys. Rev. Lett. 113* (2014), no. 17 171601, [arXiv:1307.2199].

[29] F. Cachazo, S. He, and E. Y. Yuan, *Scattering of Massless Particles: Scalars, Gluons and Gravitons, JHEP 07* (2014) 033, [arXiv:1309.0885].

[30] F. Cachazo, S. He, and E. Y. Yuan, *Scattering Equations and Matrices: From Einstein To Yang-Mills, DBI and NLSM, JHEP 07* (2015) 149, [arXiv:1412.3479].

[31] S. He, F. Teng, and Y. Zhang, *String amplitudes from field-theory amplitudes and vice versa, Phys. Rev. Lett. 122* (2019), no. 21 211603, [arXiv:1812.03369].

[32] S. He, F. Teng, and Y. Zhang, *String Correlators: Recursive Expansion, Integration-by-Parts and Scattering Equations, JHEP 09* (2019) 085, [arXiv:1907.06041].

[33] K. Zhou, *Unified web for expansions of amplitudes, arXiv:1908.10272.*