Applications of 2D CFT : Entanglement Negativity and Soft Theorem

A Dissertation presented

by

Yihong Wang

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Physics and Astronomy

(String Theory)

Stony Brook University

August 2017

(include this copyright page only if you are selecting copyright through ProQuest, which is optional)

Stony Brook University
The Graduate School

Yihong Wang

We, the dissertation committe for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

Christopher Herzog - Dissertation Advisor
Associate Professor, Department of Physics and Astronomy

George Sterman - Chairperson of Defense
Distinguished Professor and Director, Department of Physics and Astronomy

Martin Rocek
Professor, Department of Physics and Astronomy

Xu Du
Associate Professor, Department of Physics and Astronomy

Oleg Viro
Professor, Mathematics Department Stony Brook University

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Applications of 2D CFT : Entanglement Negativity and Soft Theorem

by

Yihong Wang

Doctor of Philosophy

in

Physics and Astronomy

(String Theory)

Stony Brook University

2017

Conformal field theory (CFT), especially two dimensional conformal field theory, has been an active and fruitful topic of theoretical physics for decades. As a powerful tool and insightful approximation, it has been widely used in different fields of physics such as string theory, statistical physics, and condensed matter physics. In this dissertation I shall discuss two different subjects in which the results of 2D CFT can be applied: entanglement negativity and the soft gluon/graviton theorem. 2D CFT is one of the most important approaches in calculating entanglement of 1D free systems. By applying the replica trick, the entanglement entropy can be expressed as the path integral of the free field on a Riemann surface, or equivalently as correlation functions of corresponding twist operators. The same approach applies to negativity, a well defined measure of entanglement for mixed states. In this dissertation, we discuss the negativity of free fermions, which is a hard problem because individual terms in the sum over spin structures do not respect the replica symmetry. I shall present in detail how some of the terms can be reduced to rational functions by Thomae's formula and how to use these terms to construct upper and lower bounds on free fermion negativity. In the second

part of the dissertation, I shall discuss the soft gluon/graviton theorem of scattering amplitudes. The connection between 4D scattering amplitudes and 2D CFT has been discussed at length in recent years. They share the same $SL(2,\mathbb{C})$ symmetry. By replacing the plane wave external legs with an $SL(2,\mathbb{C})$ wave packet, one can transform from scattering amplitudes to Witten diagrams in AdS_3, which then map to correlation functions on the boundary. The soft gluon and graviton theorems can then be re-expressed as Ward identities of the BMS symmetry. In this dissertation, I shall show how these two soft theorems are related through KLT relations.

Dedication Page

To my mother, my farther, my husband Chihhao, and my cats Coupon Coupino and Earl Grey

Table of Contents

Contents

Acknowledgements

I would first like to express my gratitude to my advisors, Dr. Chris Herzog and Prof. Martin Rocek. I have been working with Chris since the second year of my Ph.D, and we worked on several projects on different subjects. He generously supported me with travel funding, computing devices, valuable time and advice. Many times, he kindly and patiently guided me through areas in which I have no background and into insightful directions whenever I was stuck or confused in research and study. Though, unfortunately, I haven't had the chance to work with Martin, he has been helpful and supportive all along. His door is always open for questions and discussions. It is always helpful talking to Martin. Sometimes my questions could take hours but he is always listening patiently and actively providing me enlightening facts and ideas.

I would also like to thank my defense committee Prof. George Sterman, Prof. Oleg Viro, and Dr. Xu Du. The fact that I'm away in KITP this semester caused some additional inconvenience for the defense, but they are all very understanding and supportive. During my Ph.D study, George had helped me in many different ways: I learned a lot in his QFT classes, received financial support from YITP under his supervision, and enjoyed the parties at his home every year. Xu has been in my committee since my oral exam, and he has helped me with every annual meeting over these years. And it's a happy memory attending Oleg's topology class, I benefited a lot from both his lectures and his talks. I think it was the first time I really enjoyed and understood math.

And I'm grateful for professors who taught me, especially Prof. Peter van Nieuwenhuizen and Prof. Samuel Grushevsky. I took many classes of Peter's. I was amazed by his broad knowledge, fluent techniques and deep understanding of physics and moved by how devoted he is to teaching and helping young students. And I also took many math classes of Sam's. I'm grateful he generously let me register as a physics student everytime and constantly communicated with me to help me catch up with math students along the way. I'm also thankful to my undergraduate advisor Prof. Bo Feng, who always encourages and mentors me every time we meet.

And I would also like to thank my friends and colleges, and staff in Stony

Brook and KITP: Fen Guan, who kept me company from Hangzhou to Stonybrook; Yusong, who helped me a lot during my defense; Sara, Don, Betty Bibi and Lori, who were always there whenever I needed help; my office mates in Stony Brook and KITP and all the old and new friends I met during the Scamp17 program at the KITP.

Last but not least, I would express my profound gratitude to my parents for their warm love and support. I missed them everyday throughout my years away from home. And I would like to thank my husband Chihhao for every happy memory during these years. We were together since I first arrived. And I can still recall many happy, touching, or even awkward scenes of us in Stony Brook when I leave; they are all precious memories for me. I hope we can have more such moments in our life together in the future, and in many other beautiful places in the world

List of Publications

This thesis is based on the following publications

- Y. J. Du, B. Feng, C. H. Fu and Y. Wang, "Note on Soft Graviton theorem by KLT Relation," JHEP **1411**, 090 (2014)
- C. P. Herzog and Y. Wang, "Estimation for Entanglement Negativity of Free Fermions," J. Stat. Mech. **1607**, no. 7, 073102 (2016)

Earlier publications of the author that are not included in this thesis

- B. Feng, J. Wang, Y. Wang and Z. Zhang, "BCFW Recursion Relation with Nonzero Boundary Contribution," JHEP **1001**, 019 (2010)
- Y. Y. Chang, B. Feng, C. H. Fu, J. C. Lee, Y. Wang and Y. Yang, "A note on on-shell recursion relation of string amplitudes," JHEP **1302**, 028 (2013)

Chapter 1

1 Introduction

As a well studied highly symmetric system and one of the few examples of exactly solvable quantum field theories, two dimensional conformal field theory is an essential tool in many areas of physics. For example, in condensed matter theory and statistical mechanics, it is the continuum limit of 2D lattice models such as the critical Ising model. Lattice models are classified by their critical behavior into universality classes which corresponds to different ADE classes in 2D theory [3]. In string theory, 2D CFT is widely used as well: from the world sheet point of view string theory is described by 2D CFT, and the AdS_3/CFT_2 is one of the best known cases in studying the holographic principle [4]. Meanwhile, the algebraic structure and the OPE expansion approach led to the development of vertex algebra in mathematics [5]. Such abundant applications have made 2D CFT a remarkably long-lasting focus of research. Moreover, 2D CFT keeps finding its way in many newly developed areas in physics such as quantum entanglement and scattering amplitudes which I shall discuss in this dissertation.

1.1 Application of 2D CFT in Quantum Entanglement

Entanglement and Negativity

As a distinguished phenomenon that exists only in quantum systems, quantum entanglement was proposed [6] and discussed in the very early years of quantum mechanics (1935) by Einstein and Schrödinger. Afterwards, it remained a crucial topic in studying quantum foundations and in pursuing informational explanations for axioms of quantum mechanics. A more practical viewpoint of quantum entanglement leads to the development of quantum information: entanglement pairs can be used to build quantum computers, and quite a few algorithms that are unique to such systems have been developed since the 1990s. Some of them, such as Shor's algorithm for factoring large integers [9] and Grover's algorithm for searching unordered lists [11], are remarkably more efficient than classical algorithms. Other promising applications of quantum entanglement include quantum cryptography, quantum error correction and quantum teleportation. In recent years, there is an increasing number of gravity and string theorists joining in the study of quantum entanglement, especially after the ER=EPR conjecture by Maldacena and Susskind [7], which implies entanglement could be as fundamental as spacetime in physics. Entanglement entropy was first studied for discrete systems, and only later analyzed in continuous systems. One straightforward application is the continuum limit of spin chains. A less obvious application is an attempt to provide an entanglement interpretation for black hole entropy [8].

To study quantum entanglement quantitatively, a number of entanglement measures have been defined. A good measure of quantum entanglement should be a quantity that does not increase under local operation and classical communication. Among all such quantities, entanglement entropy is the most well-known and well-studied one. It is defined as follows: a state Ψ in a bipartite system with Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ can be written as

$$\Psi = \sum_j c_j \left|\psi_j\right\rangle_A \otimes \left|\psi_j\right\rangle_B$$

and its density matrix as

$$\rho = |\Psi\rangle \otimes \langle\Psi| \ .$$

Then the entanglement entropy of subsystem A is defined as

$$S_A = -\operatorname{tr} \rho_A \log \rho_A$$

where ρ_A is the reduced density matrix: $\rho_A = \operatorname{tr}_B \rho$. One can check explicitly

$$S_A = -\sum_j |c_j|^2 \log |c_j|^2$$

so that we have $S_A = S_B$. Other important properties of entanglement entropy include subadditivity, which says for subsystems $A_1 \cup A_2 = A$

$$S_{A_1} + S_{A_2} \geq S_A$$

and strong subadditivity, for $A_i \cap A_j = \emptyset, i \neq j \in \{1, 2, 3\}$

$$S_{A_1 \cup A_2 \cup A_3} + S_{A_2} \leq S_{A_1 \cup A_2} + S_{A_2 \cup A_3} \ .$$

Since traces of reduced density matrices are path integrals, a large number of tools and ideas developed in calculating path integrals can be brought to bear. In calculating entanglement entropy of quantum field theories, entanglement entropy is often considered as the $n \to 1$ limit of Rényi entropies

$$S_A^{(N)} = \frac{1}{1-n} \operatorname{tr} \rho_A^n \ . \tag{1}$$

The trace $\operatorname{tr} \rho_A^n$ has a direct field theory interpretation as the path integral over n copies of the space-time glued cyclically over region A. The 2D case was first studied in [13], in which Cardy and Calabrese worked out the entanglement entropy for single interval systems, which is a good example of how 2D CFT is used in calculating entanglement entropy. The derivation is based on the replica trick: from the field theory point of view, since the reduced density matrix ρ_A is a path integral traced over region B, the trace $\operatorname{tr} \rho_A^n$ is the path integral over n copies of spacetime glued cyclically along region A, as illustrated in Figure 1.

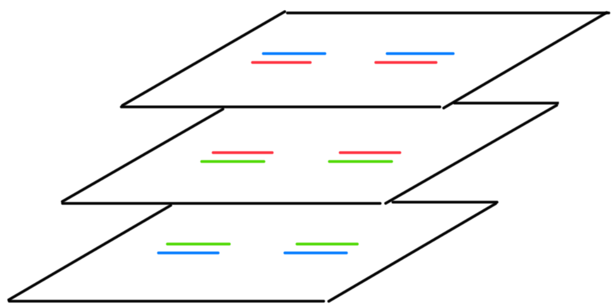


Figure 1: An illustration of the world sheet for the 3rd Rényi entropy in 2D, where the entangling region A consists of two disjoint intervals: each copy of the space time is cut along the two intervals and the different sheets are glued together along the cuts by matching the colors.

By applying the replica trick, a single field path integral over this n-sheeted Riemann surface is reduced to the path integral of n different fields identified cyclicly along the cuts over the complex plane. In CFT language, it can be written as the correlation function of twist operators. For example when A is the interval (u, v), we have $\operatorname{tr}(\rho_A) = \left\langle \mathcal{T}_n(u,0) \bar{\mathcal{T}}_n(v,0) \right\rangle$. The conformal weight of twist operators can be derived from the expectation value of the stress tensor since we have

$$\left\langle T^{(n)}(w) \right\rangle_{\mathcal{R}_{n,1}} = \frac{\left\langle \mathcal{T}_n(u,0) \bar{\mathcal{T}}_n(v,0)\, T^{(n)}(w) \right\rangle_{\mathcal{L}^{(n)},\mathbb{C}}}{\left\langle \mathcal{T}_n(u,0) \bar{\mathcal{T}}_n(v,0) \right\rangle_{\mathcal{L}^{(n)},\mathbb{C}}} \tag{2}$$

in which the subscript $\mathcal{R}_n, 1$ labels quantities evaluated on the n-sheeted Riemann surface, whereas $\mathcal{L}^{(n)}, \mathbb{C}$ labels quantities evaluated on the complex plane under the replicated Lagrangian. Since $\mathcal{R}_n$ can be mapped to the complex plane by the map $z = \left(\frac{w-u}{w-v}\right)^{1/n}$, the left hand side of (2) follows from the Schwarzian of this map. And the conformal weight of the twist operators $d_n$and $\bar{d}_n$, $d_n = \bar{d}_n = \frac{c}{12}\left(n - \frac{1}{n}\right)$, can be read from the Ward identity

$$\left\langle \mathcal{T}_n(u,0)\bar{\mathcal{T}}_n\left(v,0\right)T^{(n)}\left(w\right)\right\rangle_{\mathcal{L}^{(n)},\mathbb{C}} = \left(\frac{1}{w-u}\frac{\partial}{\partial u} + \frac{d_n}{\left(w-u\right)^2} + \frac{1}{w-v}\frac{\partial}{\partial v} + \frac{\bar{d}_n}{\left(w-v\right)^2}\right)\left\langle \mathcal{T}_n(u,0)\bar{\mathcal{T}}_n\left(v,0\right)\right\rangle_{\mathcal{L}^{(n)},\mathbb{C}} \quad (3)$$

With the knowledge of the conformal dimensions, one immediately has the single interval Rényi entropies, $\mathrm{tr}\,\rho_A^{n,1} = c_n\left(\frac{v-u}{a}\right)^{-c(n-1/n)/6}$. For some simple theories, the two disjoint interval Rényi entropy can be derived by applying known results or tricks in CFT. For compactified free bosons, it is the boson partition function on genus-n Riemann surfaces. And for free fermions it can be calculated by bosonization [12]. Due to the Gaussian nature of the reduced density matrix, entanglement entropy of these systems can be verified numerically as well [14].

With the CFT interpretation of entanglement entropy, it is natural to examine its gravity dual under AdS/CFT. It was conjectured in 2006 by Ryu and Takayanagi that the entanglement entropy of a region on the boundary corresponds to the minimal area of hypersurfaces in the bulk anchored on the boundary of the entangling region [10]:

$$S_A = \frac{\mathrm{Area}\left(\gamma_A\right)}{4G_N^{(d+1)}}$$

which is an elegant and inspiring result since it not only confirms the area law of entanglement entropy but also adds a new entry to the AdS/CFT dictionary.

This progress in calculating and understanding entanglement entropy has led to better understanding of information and spacetime and brought together many seemingly disconnected areas of physics such as quantum information, quantum field theory and gravity. However this is not yet the whole picture since the entanglement entropy is not a good measure of entanglement for mixed states, or tripartite systems. One obvious obstacle for entanglement entropy in measuring mixed states or tripartite systems is that $S_A = S_B$ no longer holds for mixed states in general. Therefore to discuss entanglement of mixed states or tripartite systems, one needs a new measure. One such measure, the negativity, was proposed Vidal in 2001 [18], which was later demonstrated to be a good entanglement measure [19]. The negativity is an monotone under local operation classical communication for these mixed states.

The negativity is defined as follows: For a state $|\Psi\rangle$ in a quantum system with bipartite Hilbert space $\mathcal{H} = \mathcal{H}_A \bigotimes \mathcal{H}_B$ and density matrix $\rho = |\Psi\rangle\langle\Psi|$, the reduced density matrix is defined as $\rho_A = \mathrm{tr}_B\, \rho$. If $\mathcal{H}_A$ is factored further into $\mathcal{H}_A = \mathcal{H}_{A_1} \bigotimes \mathcal{H}_{A_2}$, one can define the partial transpose of the reduced density matrix $\rho_A^{T_2}$ as the operator such that the following identity holds for any $e_i^{(1)}$, $e_k^{(1)} \in \mathcal{H}_{A_1}$ and $e_j^{(2)}$, $e_l^{(2)} \in \mathcal{H}_{A_2}$: $\left\langle e_i^{(1)} e_j^{(2)} \left| \rho_A^{T_2} \right| e_k^{(1)} e_l^{(2)} \right\rangle = \left\langle e_i^{(1)} e_l^{(2)} \left| \rho_A \right| e_k^{(1)} e_j^{(2)} \right\rangle$. Negativity is defined as the trace norm[1] of $\rho_A^{T_2}$. Since $\rho_A^{T_2}$ is Hermitian, its trace norm can be written as the following limit

$$\mathcal{E} \equiv |\rho_A^{T_2}| = \lim_{N_e \to 1} \mathrm{tr} \left(\rho_A^{T_2}\right)^{N_e} \tag{4}$$

where N_e is an even integer. This analytic continuation suggests the utility of also defining higher moments of the partial transpose:

$$\mathcal{E}(N) \equiv \mathrm{tr}\left[(\rho_A^{T_2})^N\right] \ . \tag{5}$$

Like the entanglement entropy, the negativity in a quantum field theory can be computed by employing the replica trick [20,21]: like ρ_A^N, $\left(\rho_A^{T_2}\right)^N$ is the partition function over n copies of spacetime glued along the region A but where now for transposed intervals, the gluing order is anti-cyclic rather than cyclic, as illustrated in Figure 2.

In practice, these partition functions can only be computed in special cases [20,21]. For conformal field theories in $1+1$ dimensions, the N-moment negativity $\mathcal{E}(N)$ for a single interval (i.e. B is the empty set) can be mapped to functions of Rényi entropies: the Nth Rényi entropy for odd N; , the square of the $\frac{N}{2}$th Rényi entropy for even N. In more detail, we find

$$\mathrm{tr} \left(\rho_A^{T_2}\right)^{N_e} \propto l^{-\frac{c}{3}\left(\frac{N_e}{2} - \frac{2}{N_e}\right)}, \quad \mathrm{tr} \left(\rho_A^{T_2}\right)^{N_o} \propto l^{-\frac{c}{3}\left(N_o - \frac{1}{N_o}\right)} \tag{6}$$

[1]The trace norm of a matrix M is defined as the sum of its singular values: $|M| \equiv \mathrm{tr}\left[\left(M^\dagger M\right)^{1/2}\right]$. For Hermitian matrices, singular values are absolute values of the eigenvalues.

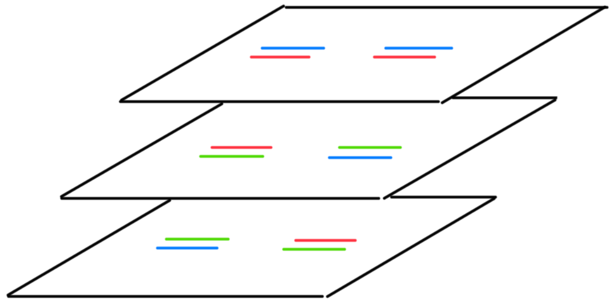


Figure 2: An illustration of the world sheet for the 3rd moment of negativity in 2D, where the entangling region A consists of two disjoint intervals and the second interval is transposed: each copy of the space time is cut along the two intervals, and intervals of the same color are identified

where N_e stands for even N and N_o stands for odd N. Another tractable case is where A is a single interval broken in two pieces by A_1 and A_2. The quantity $\mathcal{E}(N)$ reduces to a three point correlation function of two twist operators and a double twist operator. The result is determined by conformal symmetry:[2]

$$\mathrm{tr}\left(\rho_A^{T_2}\right)^{Ne} \propto (l_1 l_2)^{-\frac{c}{6}\left(\frac{N_e}{2}-\frac{2}{N_e}\right)}(l_1+l_2)^{-\frac{c}{6}\left(\frac{N_e}{2}+\frac{1}{N_e}\right)}, \tag{7}$$

$$\mathrm{tr}\left(\rho_A^{T_2}\right)^{N_o} \propto \left(l_1 l_2 (l_1+l_2)\right)^{-\frac{c}{12}\left(N_o-\frac{1}{N_o}\right)}. \tag{8}$$

For free bosons, moments of negativity for two disjoint intervals can also be derived by mapping to Rényi entropy. And the numerical approach for entanglement entropy can be applied to this case as well. However, it is much more complicated for free fermions, and we shall elaborate in the next chapter.

1.2 Application of 2D CFT in Scattering Ampliudes

Chapter3 of this dissertation is on the soft graviton/gluon theorem, which improves Weinberg's soft photon theorem [42,43]. The soft limit is an import subject in scattering amplitudes because of its universality, the simplicity of its form and the symmetries it reveals. The study of soft limits dates back to the 1960s [42, 43]. In 2013, a new soft theorem for gravity amplitudes was

[2]See ref. [22] for an extension to the massive case.

studied in [44–46]. Using Britto-Cachazo-Feng-Witten (BCFW) recursion [47, 48], Cachazo and Strominger proved the sub- and subsubleading orders in the soft expansion [49], i.e.,[3]

$$M_{n+1}(\{\epsilon\lambda_s, \widetilde{\lambda}_s\}, 1, ..., n) = \left(\frac{1}{\epsilon^3} S_{GR}^{(0)} + \frac{1}{\epsilon^2} S_{GR}^{(1)} + \frac{1}{\epsilon} S_{GR}^{(2)}\right) M_n(1, ..., n) + \mathcal{O}(\epsilon^0). \quad (9)$$

The leading, subleading and subsubleading orders of soft factors are given by

$$S_{GR}^{(0)} = \sum_{a=1}^{n} \frac{\varepsilon^s_{\mu\nu} p_a^\mu p_a^\nu}{p_s \cdot p_a}, \quad S_{GR}^{(1)} = \sum_{a=1}^{n} \frac{\varepsilon^s_{\mu\nu} p_a^\mu (p_{s,\rho} J_a^{\rho\nu})}{p_s \cdot p_a}, \quad (10)$$

$$S_{GR}^{(2)} = \frac{1}{2} \sum_{a=1}^{n} \frac{\varepsilon^s_{\mu\nu} (p_{s,\rho} J_a^{\rho\mu})(p_{s,\sigma} J_a^{\sigma\nu})}{p_s \cdot p_a}, \quad (11)$$

where the $\varepsilon^s_{\mu\nu}$ is the polarization of the soft graviton, p_i are external momenta and $J^{\mu\nu}$ are angular momenta of external legs. Using the BCFW recursion relation, the soft limit of color-ordered tree-level Yang-Mills amplitudes was also studied in [52] and the result is given by

$$A_{n+1}(\{\epsilon\lambda_s, \widetilde{\lambda}_s\}, 1, ..., n) = \left(\frac{1}{\epsilon^2} S_{YM}^{(0)} + \frac{1}{\epsilon} S_{YM}^{(1)}\right) A_n(1, ..., n), \quad (12)$$

where the leading and subleading soft factors are given by

$$S_{YM}^{(0)} = \sum_{a\sim s} \frac{\varepsilon_s \cdot p_a}{p_s \cdot p_a}, \quad S_{YM}^{(1)} = \sum_{a\sim s} \frac{\varepsilon_{s\nu} p_{s\mu} J_a^{\mu\nu}}{p_s \cdot p_a}, \quad (13)$$

with $\varepsilon_{s\nu}$ denoting the polarization of the soft gluon and $a \sim s$ meaning partial a is next to soft particle s.

In recent years 2D CFT techniques has been introduced into the study of topics in $3+1$D scattering amplitudes. There have been straightforward applications of CFT approaches in calculating scattering amplitudes in some specific theories such as applying AdS/CFT duality in deriving S-matrix for boundary CFTs [81–84], and using bootstrap to study $1+1$ dimension S-matrix [85]. In fact for $3+1$ dimensional QFTs, the 2D CFT approaches have

[3]That the leading soft factor $S_{GR}^{(0)}$ is not corrected to all loop orders is shown in [43, 50] while the general subleading behavior of soft gluons and gravitons has also been discussed in [51].

a broader application. It is shown in [78] that soft theorems can be written as Ward identities for a conserved current in 2D CFT. The result motivates a discussion of the relation between 2D CFT and 4D scattering amplitudes in general. Strong evidence for the link between these two subjects is that the Lorentz group in 4D $SL(2,\mathbb{C})$ is the global conformal symmetry group in 2D. To be specific, a normalized momentum $\hat{p}$ in $\mathbb{R}^{1,3}$ can be expressed by coordinates in AdS_3:

$$\hat{p}^\mu(y,z,\bar{z}) = \left(\frac{1+y^2+|z|^2}{2y}, \frac{\text{Re}(z)}{y}, \frac{\text{Im}(z)}{y}, \frac{1-y^2-|z|^2}{2y}\right) . \tag{14}$$

In terms of $y, z, \bar{z}$, the Lorentz transformation $SL(2,\mathbb{C})$ reads

$$\begin{aligned} z \to z' &= \frac{(az+b)(\bar{c}\bar{z}+\bar{d})+a\bar{c}y^2}{|cz+d|^2+|c|^2y^2} , \\ \bar{z} \to \bar{z}' &= \frac{(\bar{a}\bar{z}+\bar{b})(cz+d)+\bar{a}cy^2}{|cz+d|^2+|c|^2y^2} , \\ y \to y' &= \frac{y}{|cz+d|^2+|c|^2y^2} , \end{aligned} \tag{15}$$

which are isometries of AdS_3 or conformal transformations of the boundary CFT and preserve the on-shell condition

$$-(\hat{p}^0)^2 + (\hat{p}^1)^2 + (\hat{p}^2)^2 + (\hat{p}^3)^2 = -1 . \tag{16}$$

For scattering amplitudes the only objects with nontrivial transformation under Lorentz transformation are the external legs, which are plane waves. After changing the external states into eigenstates of the Lorentz symmetry, the scattering ampliltude turns into a Lorentz invariant object: the Witten diagram in $\text{AdS}_3/\text{CFT}_2$ [79, 80]. For example, with the knowledge of the scalar bulk to boundary propagator

$$G_\Delta(y,z,\bar{z};w,\bar{w}) = \left(\frac{y}{y^2+|z-w|^2}\right)^\Delta \tag{17}$$

and the map (14), the conformal primary wave function reads

$$\phi^\pm_{\Delta,m}(X^\mu;w,\bar{w}) = \int_0^\infty \frac{dy}{y^3}\int dzd\bar{z}\, G_\Delta(y,z,\bar{z};w,\bar{w}) \exp\left[\pm im\,\hat{p}^\mu(y,z,\bar{z})\,X_\mu\right] \tag{18}$$

Therefore the n-point scattering amplitudes can be related to amplitudes in CFT_2 in the following way

$$\tilde{\mathcal{A}}_{\Delta_1,\cdots,\Delta_n}(w_i,\bar{w}_i) \equiv \prod_{i=1}^{n}\int_0^\infty \frac{dy_i}{y_i^3}\int dz_i d\bar{z}_i\, G_{\Delta_i}(y_i,z_i,\bar{z}_i;w_i,\bar{w}_i)\,\mathcal{A}(m_j\hat{p}_j^\mu)\,. \quad (19)$$

This relation is derived and verified for general n-point amplitudes for massive scalars [80]. And there are on-going efforts to generalize it to higher spin amplitudes. Connection between scattering amplitudes in momentum space $\mathbb{R}^{1,3}$ and objects in CFT_2 are most studied and well understood in the soft limit.

And it is shown in [78] that by treating the soft particle as a conserved current in the boundary CFT, the leading order soft theorem can be written as the Ward identity of the current, for example. Following the map (14), the soft gluon theorem reads

$$\begin{aligned}&\left\langle j(w)^a \mathcal{O}^{b_1}(w_1,\bar{w}_1)\cdots\mathcal{O}^{b_n}(w_1,\bar{w}_1)\right\rangle \\ &=\sum_{i=1}^{n}\frac{f^{ab_ic_i}}{w-w_i}\left\langle \mathcal{O}^{b_1}(w_1,\bar{w}_1)\cdots\mathcal{O}^{b_n}(w_1,\bar{w}_1)\right\rangle \qquad (20)\end{aligned}$$

which is exactly a 2D conformal Ward identity for a conserved current $j(w)$, and it was later shown in [87] that this conserved current corresponds to a generalized BMS symmetry. Meanwhile many other related studies have been achieved including the soft limits from Poincaré symmetry and gauge invariance [53, 54], Feynman diagram approach [55], conformal symmetry approach to the soft limits in Yang-Mills theory [56], the soft limit in arbitrary dimension [57–60], loop correction of the soft limit [61–63, 65], string theory approach to the soft limit [64, 65] and ambitwistor string approach [66, 67]. The relation between scattering amplitudes and 2D correlation functions is yet to be found for most of the theories. And this is a hard problem in general, as we mentioned earlier, such relation is first studied and best known for amplitudes with soft particles. The study of soft theorem would help us understand the whole story. On the other hand some identities between scattering amplitudes, such as the Kawai-Lewellen-Tye (KLT) relations is also helpful. Because it relates graviton ampliudes to gluon amplitudes, which we know how to get from 2D correlation functions. In Chapter 3, we shall show how the soft gravity theorem and the soft gluon theorem are related by the KLT relations.

Chapter 2

2 Estimation for Entanglement Negativity of Free Fermions

2.1 Introduction

As we mentioned in the introduction, there are a few simple systems for which we can calculate their negatitvity, besides these examples another special case where the negativity can be determined, at least for $N > 1$, is a massless free fermion field in 1+1 dimensions. In this case, the N-sheeted partition functions are known in terms of Riemann-Siegel theta functions although it is not known in general how to continue the result away from integer $N > 1$ and in particular to $N = 1$. Since the partial transposed reduced density matrix is Gaussian, the negativity for a free scalar can be checked through a lattice computation by using Wick's Theorem [21, 23].

The case of free fermions in 1+1 dimensions appears to be more difficult than the case of free scalars however. The partial transpose of the reduced density matrix is no longer Gaussian but a sum of two, generically non-commuting, Gaussian matrices [24]:

$$\rho^{T_2} = \frac{1}{\sqrt{2}} \left(e^{i\pi/4} O_+ + e^{-i\pi/4} O_- \right) \ . \tag{21}$$

(We will define $O_\pm$ in section 2.2.) This fact brings additional complication to both the lattice and field theoretical calculations. On the lattice side, eigenvalues of $(\rho^{T_2})^N$ cannot be simply derived from eigenvalues of a covariance matrix as in the Gaussian case. In a field theory setting, one has to sum over partition functions with different spin structure, corresponding to different terms in the expansion of $(e^{i\pi/4}O_+ + e^{-i\pi/4}O_-)^N$. Various efforts have been made to tame the difficulties in deriving the negativity of free fermions: On the lattice side, algebraic simplification and numerical diagonalization of products of these two Gaussian matrices yields the $N > 1$ moments of negativity for the two disjoint interval case [24, 25].[4] (Monte-Carlo and tensor network methods have also been used to calculate negativity for the Ising model [26–28] which, although not identical to the Dirac fermion, is closely

[4] See also ref. [31] for an extension to two spatial dimensions.

related.) The analytical form of such moments are derived by evaluation of the corresponding path integrals [29, 30]. However in the existing results the sheet number N does not appear as a continuous variable; it remains an open problem how to take the $N \to 1$ limit to get the negativity.

In this letter we shall introduce a $\mathbb{Z}_N$-symmetric free fermion with specific choice of spin structure. This fermion has several nice features that we believe will help us explore and understand the features of free fermion negativity. 1) The partition function explicitly reproduces the correct adjacent interval limit. 2) The $N \to 1$ limit of the N sheeted path integral can be easily derived. 3) There exists a natural generalization to multiple interval cases, nonzero temperature, and nonzero chemical potential. 4) While such a partition function is not an N^{th} moment of ρ^{T_2} (except in the special case $N = 2$), it appears to be a useful quantity for bounding these N^{th} moments including the negativity itself.

The rest of this letter is arranged as follows: In section 2.2 we review previous results. Section 2.3 contains a derivation of the partition function for the $\mathbb{Z}_N$-symmetric free fermion system and in particular $\text{tr}(O_+^N)$ and $\text{tr}[(O_+O_-)^{N/2}]$. In section 2.4, we discuss bounds on the negativity and its N^{th} moments. We conclude in section 2.5 with remarks on possible generalizations of our results and future directions. An appendix contains a discussion of a two-spin system.

2.2 Review of Previous Results

We first review the definition of the negativity. For a state $|\Psi\rangle$ in a quantum system with bipartite Hilbert space $\mathcal{H} = \mathcal{H}_A \bigotimes \mathcal{H}_B$ and density matrix $\rho = |\Psi\rangle\langle\Psi|$, the reduced density matrix is defined as $\rho_A = \text{tr}_B\,\rho$. If $\mathcal{H}_A$ is factored further into $\mathcal{H}_A = \mathcal{H}_{A_1} \bigotimes \mathcal{H}_{A_2}$, one can define the partial transpose of the reduced density matrix $\rho_A^{T_2}$ as the operator such that the following identity holds for any $e_i^{(1)}$, $e_k^{(1)} \in \mathcal{H}_{A_1}$ and $e_j^{(2)}$, $e_l^{(2)} \in \mathcal{H}_{A_2}$: $\left\langle e_i^{(1)} e_j^{(2)} \left| \rho_A^{T_2} \right| e_k^{(1)} e_l^{(2)} \right\rangle = \left\langle e_i^{(1)} e_l^{(2)} \left| \rho_A \right| e_k^{(1)} e_j^{(2)} \right\rangle$. Negativity is defined as the trace norm[5] of $\rho_A^{T_2}$. Since

[5] The trace norm of a matrix M is defined as the sum of its singular values: $|M| \equiv \text{tr}\left[\left(M^\dagger M\right)^{1/2}\right]$. For Hermitian matrices, singular values are absolute values of the eigenvalues.

$\rho_A^{T_2}$ is Hermitian, its trace norm can be written as the following limit

$$\mathcal{E} \equiv |\rho_A^{T_2}| = \lim_{N_e \to 1} \mathrm{tr}\left(\rho_A^{T_2}\right)^{N_e} \tag{22}$$

where N_e is an even integer. This analytic continuation suggests the utility of also defining higher moments of the partial transpose:

$$\mathcal{E}(N) \equiv \mathrm{tr}\left[(\rho_A^{T_2})^N\right] \ . \tag{23}$$

We are interested in systems in one time and one spatial dimension. We will assume a factorization of the Hilbert space corresponding to a partition of the real line with A_1 and A_2 each being the union of a collection of disjoint intervals: $A_1 = \cup_{i=1}^{p}(s_i, t_i)$ and $A_2 = \cup_{i=1}^{q}(u_i, v_i)$.

In this paper, we are particularly interested in the case of free, massless fermions in 1+1 dimension with the continuum Hamiltonian

$$H = \mp i \int_0^L \Psi^\dagger(t,x)\partial_x \Psi(t,x)\, \partial x \tag{24}$$

where $\{\Psi^\dagger(t,x), \Psi(t,x')\} = \delta(x-x')$. The sign determines whether the fermions are left moving or right moving. We will take one copy of each to reassemble a Dirac fermion. It will often be convenient to consider the lattice version of this Hamiltonian as well

$$H = \mp \frac{i}{2} \sum_j \left(\Psi_j^\dagger \Psi_{j+1} - \Psi_{j+1}^\dagger \Psi_j\right) \ , \tag{25}$$

and anti-commutation relation $\{\Psi_j^\dagger, \Psi_k\} = \delta_{jk}$, which suffers the usual fermion doubling problem. We choose as our vacuum the state annihilated by all of the Ψ_j.

The authors of ref. [24] were able to give a relatively simple expression for the negativity in the discrete case by working instead with Majorana fermions $a_{2j-1} = \frac{1}{2}(\Psi_j^\dagger + \Psi_j)$ and $a_{2j} = \frac{1}{2i}(\Psi_j^\dagger - \Psi_j)$. Re-indexing, we can write the reduced density matrix as a sum over words made of the a_j:

$$\rho_A = \sum_\tau c_\tau \prod_{j=1}^{2n} a_j^{\tau_j} \tag{26}$$

where τ_j is either zero or one, depending on whether the word τ contains the Majorana fermion a_j, and n is the length of region A. Consider now instead the matrices $O_\pm$ constructed from ρ_A by multiplying all the a_j in region A_2 by $\pm i$:

$$O_\pm = \sum_{\tau,\sigma} c_{\tau,\sigma} \left(\prod_{j=1}^{2n_1} a_j^{\tau_j} \right) \left(\prod_{j=2n_1+1}^{2n_1+2n_2} (\pm i a_j)^{\sigma_j} \right) . \tag{27}$$

Here n_j is the length of region A_j, and we have broken the sum into words τ involving region A_1 and words σ involving region A_2. As we already described in eq. (21), the central result of ref. [24] is that the partial transpose of the reduced density matrix can be written in terms of $O_\pm$.

While the spectrum of ρ_A is not simply related to the spectra of $O_\pm$, it is true that O_+ and O_- are not only Hermitian conjugates but are also related by a similarity transformation and so have the same eigenvalue spectrum. Consider a product of all of the Majorana fermions in A_2,

$$S = i^{n_2} \prod_{j=2n_1+1}^{2(n_1+n_2)} a_j \ , \tag{28}$$

which squares to one, $S^2 = 1$. This operator provides the similarity transformation between O_+ and O_-, i.e. $O_+ = S O_- S$. This similarity transformation means, along with cyclically of the trace, that if we have a trace over a word constructed from a product of O_+ and O_-, the trace is invariant under the swap $O_+ \leftrightarrow O_-$. Employing this similarity transformation, the negativity for the first few even N can be written thus

$$\mathrm{tr}[(\rho_A^{T_2})^2] = \mathrm{tr}(O_+ O_-) \ , \tag{29}$$

$$\mathrm{tr}[(\rho_A^{T_2})^4] = -\frac{1}{2}\,\mathrm{tr}(O_+^4) + \mathrm{tr}(O_+^2 O_-^2) + \frac{1}{2}\,\mathrm{tr}[(O_+ O_-)^2] \ , \tag{30}$$

$$\mathrm{tr}[(\rho_A^{T_2})^6] = -\frac{3}{2}\,\mathrm{tr}(O_+ O_-^5) + \frac{1}{4}\,\mathrm{tr}[(O_+ O_-)^3] + \frac{3}{4}\,\mathrm{tr}(O_+^3 O_-^3) + \frac{3}{2}\,\mathrm{tr}(O_+ O_- O_+^2 O_-^2) \tag{31}$$

To obtain analytic expressions for $\mathrm{tr}[(\rho_A^{T_2})^N]$ from the decomposition (21) of $\rho_A^{T_2}$, a key step [25] is the relation between matrix elements of ρ_A and matrix elements of O_+. Consider arbitrary coherent states $\langle \zeta(x)|$ and $|\eta(x)\rangle$ that further break up into $\langle \zeta_1(x_1), \zeta_2(x_2)|$ and $|\eta_1(x_1), \eta_2(x_)\rangle$ according to the decomposition of A into A_1 and A_2. Then the matrix elements of ρ_A and O_+ are related via

$$\langle \zeta(x)|O_+|\eta(x)\rangle = \langle \zeta_1(x_1), \eta^*(x_2)|U_2^\dagger \rho_A U_2|\eta_1(x_1), -\zeta_2^*(x_2)\rangle \ , \tag{32}$$

where U_2 is a unitary operator (whose precise form [25] does not concern us) that acts only on the part of the state in region A_2.

In pursuit of an analytic expression, let us move now to a path integral interpretation of $\mathrm{tr}[\rho_A^N]$ and $\mathrm{tr}[(\rho_A^{T_2})^N]$. The trace over ρ_A^N becomes a path integral over an N sheeted cover of the plane, branched over A. Now consider instead $\mathrm{tr}\, O_+^N$ given the relation (32). Performing a change of variables, we can replace U_2 acting on ζ_2^* and η_2^* with ζ_2 and η_2 inside the trace, and we see that $\mathrm{tr}(\rho_A^N)$ is related to $\mathrm{tr}\, O_+^N$ by an orientation reversal of region A_2. In terms of the N sheets, fixing a direction, passing through an interval in A_1, we move up a sheet while passing through an interval in A_2 we move down a sheet. Indeed, the trace of any word constructed from the O_+ and O_- has a similar path integral interpretation. Given the sign flip relation $O_- = SO_+S$ however, replacing some of the O_+ by O_- in the word will change the spin structure of the N sheeted cover.

For simplicity, consider the case where A_1 is a single interval bounded by $s < t$ and A_2 a single interval bounded by $u < v$. The trace of a word constructed from $O_\pm$, up to an undetermined over-all normalization c_N, can be written in terms of a Riemann-Siegel theta function [25]

$$\mathrm{tr}\left[\prod_{i=1}^{N} O_{s_i}\right] = c_N^2 \left(\frac{1-x}{(t-s)(v-u)}\right)^{2\Delta_N} \left|\frac{\Theta[\mathbf{e}](\tilde{\tau}(x))}{\Theta(\tilde{\tau}(x))}\right|^2 , \quad \mathbf{e} = \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\delta} \end{pmatrix} , \tag{33}$$

where $\mathbf{0}$ is a vector of $N-1$ zeros and $\boldsymbol{\delta}$ is fixed by the word $\prod_{i=1}^N O_{s_i}$. In particular, if $s_i \neq s_{i+1}$, then $\delta_i = 1/2$ and $\delta = 0$ otherwise. The exponent

$$\Delta_N = \frac{c}{12}\left(N - \frac{1}{N}\right) \tag{34}$$

is the dimension of a twist operator field with $c = 1$ for a Dirac fermion. The cross ratio is defined to be

$$x \equiv \frac{(s-t)(u-v)}{(s-u)(t-v)} \in (0,1) . \tag{35}$$

(The limit in which the intervals become adjacent corresponds to $x \to 1$.) The Riemann-Siegel theta function is defined as

$$\Theta[\mathbf{e}](\mathbf{z}|M) \equiv \sum_{\mathbf{m}\in\mathbb{Z}^{N-1}} e^{i\pi(\mathbf{m}+\boldsymbol{\epsilon})^t\cdot M\cdot(\mathbf{m}+\boldsymbol{\epsilon})+2\pi i(\mathbf{m}+\mathbf{e})^t\cdot(\mathbf{z}+\boldsymbol{\delta})} , \quad \mathbf{e} \equiv \begin{pmatrix} \boldsymbol{\epsilon} \\ \boldsymbol{\delta} \end{pmatrix} , \tag{36}$$

and further $\Theta(\mathbf{z}|M) \equiv \Theta[\mathbf{0}](\mathbf{z}|M)$. The $(N-1)\times(N-1)$ period matrix is then [21,32]

$$\tau_{i,j} = i\frac{2}{N}\sum_{k=1}^{N-1}\sin(\pi k/n)\frac{{}_2F_1(k/N,1-k/N;1;1-x)}{{}_2F_1(k/N,1-k/N;1;x)}\cos[2\pi(k/N)(i-j)] \;, \tag{37}$$

and further $\tilde{\tau}(x) = \tau(x/(x-1))$. There are Riemann-Siegel theta functions that one can write down for multiple interval cases as well, but we shall not need their explicit form.

Among the words that enter in the binomial expansion of $\mathrm{tr}[(\rho_A^{T_2})^N]$, the traces $\mathrm{tr}(O_+^N) = \mathrm{tr}(O_-^N)$ and $\mathrm{tr}[(O_+O_-)^{N/2}]$ are special. Even in the multiple interval case, these two traces can be expressed as rational functions of the endpoints of the intervals. Although we have no proof in general, observationally it seems to be true that among the words of a fixed length $\mathrm{tr}(O_+^N)$ is the smallest in magnitude while $\mathrm{tr}[(O_+O_-)^{N/2}]$ is the largest. These two considerations suggest the utility of trying to bound the negativity using the rational functions $\mathrm{tr}(O_+)^N$ and $\mathrm{tr}[(O_+O_-)^{N/2}]$, as we pursue in section 2.4

In the two interval case, it follows from the result (33) that $\mathrm{tr}(O_+^N)$ and $\mathrm{tr}[(O_+O_-)^N]$ are rational functions. That $\mathrm{tr}(O_+^N)$ reduces to a rational function is obvious since $\boldsymbol{\delta} = \mathbf{0}$. That $\mathrm{tr}[(O_+O_-)^{N/2}]$ reduces as well follows from Thomae's formula [33,34] that when $\delta_i = 1/2$ for all i,

$$\left|\frac{\Theta[\mathbf{e}](\tilde{\tau})}{\Theta(\tilde{\tau})}\right|^2 = |1-x|^{-N/4} \;. \tag{38}$$

To see more generally that these words are rational functions of the endpoints, in the next section we employ bosonization.[6]

2.3 Bosonization and Rationality

Consider the normalized partition function of the free Dirac field on the $\mathbb{Z}_N$-curve defined by the following set:

$$X_N = \left\{ (z,y) \,\middle|\, y^N = \prod_{i=1}^{p}\frac{z-s_i}{z-t_i}\prod_{i=1}^{q}\frac{z-v_i}{z-u_i}, (z,y)\in\mathbb{C}^2 \right\} \;. \tag{39}$$

[6] For an application of Thomae's formula to a multiple interval Rényi entropy computation, see ref. [35].

One can see that X_N, as the set of all points in $\mathbb{C}^2$ satisfying the equation in the set, has N sheets corresponding to N different roots of a nonzero complex number. These N copies of $\mathbb{C}$ are cut open along intervals in A on the real axis. As we choose the ordering $s_i < t_i$ and $u_i < v_i$, such open cuts are glued cyclicly if in A_1 and anti-cyclicly if in A_2.

While the Riemann surface (39) has an explicit $\mathbb{Z}_N$ symmetry, to specify a partition function, we also have to give the spin structure. The spin structure can generically break this symmetry, i.e. we can associate relative factors of minus one to cycles that would otherwise be related by the $\mathbb{Z}_N$ shift symmetry. A generic word $\prod_i O_{s_i}$ will generically have a spin structure that does not respect this symmetry. However, a few words do, namely $\mathrm{tr}(O_+^N) = \mathrm{tr}(O_-^N)$ and $\mathrm{tr}[(O_+O_-)^{N/2}]$. The word $\mathrm{tr}(O_+)^N$ associates a $+1$ to all the fundamental cycles constructed from region A_2, while the word $\mathrm{tr}[(O_+O_-)^{N/2}]$ associates a -1.

If we assume the $\mathbb{Z}_N$ symmetry is preserved by the spin structure, then the bosonization procedure is especially simple. Denote the partition function on X_N by $Z[N]$. Rather than a path integral of a single Dirac field on X_N in (39), $Z\,[N]$ can be considered as a path integral of a vector valued Dirac field $\vec{\Psi}\,(z)$ on $\mathbb{C}$: $\mathbf{\Psi}\,(x) = (\Psi_1\,(z)\,, \cdots, \Psi_N\,(z))$. $\Psi_i\,(x)$ is the value of the original field Ψ at coordinate (z, y_i) on X_N. When going anti-clockwise around a branch point w by a small enough circle C_w, $\mathbf{\Psi}\,(x)$ gets multiplied by a monodromy matrix $T\,(w)$.

Define the matrix

$$T \equiv \begin{pmatrix} 0 & \omega & & & \\ & & \omega & & \\ & & \cdot & \cdot & \\ & & & 0 & \omega \\ \omega & & & & 0 \end{pmatrix} \tag{40}$$

where $\omega = e^{2\pi i \frac{N-1}{N}}$. This value of ω is chosen so that T satises the overall boundary condition $T^N = (-1)^{N-1}\,\mathrm{id}$ where id is the $N \times N$ identity matrix. The reason for the factor $(-1)^{N-1}$ comes from considering a closed loop that circles one of the branch points N times. Such a loop should be a trivial closed loop in the y coordinate and come with an overall factor of -1, standard from performing a 2π rotation of a fermion.[7]

[7] In order to preserve an explicit $\mathbb{Z}_N$ symmetry, we have chosen a slightly different matrix than in ref. [36].

The matrix T is not the only $\mathbb{Z}_N$ symmetric matrix satisfying $T^N = (-1)^{N-1}\,\mathrm{id}$. A relative phase $e^{i2\pi k/N}$, $k = 1, 2, \ldots, N-1$, between monodromy matrices at different branch points is also allowed. Choose the basis of $\mathbf{\Psi}(x)$ so that $T(s_1) = T$ and take into account the constraint that $T(t_i)\,T(s_{i+1}) = \mathrm{id}$, $T(v_i)\,T(u_{i+1}) = \mathrm{id}$. Then, the monodromy matrices are fixed to be

$$T(s_i) = T\ , \qquad T(t_i) = T^{-1}\ , \tag{41}$$

$$T(u_i) = \exp\left(2\pi i\,(N-k)/N\right) T^{-1}\ , \qquad T(v_i) = \exp\left(2\pi i k/N\right) T\ , \tag{42}$$

If we insist on the usual spin structure for fermions, that they can only pick up an overall factor of ± 1 around any closed cycle, then two values of k are singled out, $k = 0$ for all N and $k = N/2$ for even N. The choice $k = 0$ will produce a partition function that computes $\mathrm{tr}(O_+^N)$, while the choice $k = N/2$ will produce a partition function that computes $\mathrm{tr}[(O_+O_-)^{N/2}]$. As we will discuss below, there are a pair of additional special choices, $k = (N \pm 1)/2$ for odd N, which do not have an interpretation as a $\mathrm{tr}[\prod_i O_{s_i}]$, but which nevertheless have some nice properties. For now, we will keep the dependence on k arbitrary.

As introduced in refs. [21, 36–38], a twist operator $\sigma_R^k(w)$ is defined as the field that simulates the following monodromy behavior: $\vec{\Psi}(x)\,\sigma_R^k(w) \to \exp(2\pi k/N)\,T^R\vec{\Psi}(x)\,\sigma_R^k(w)$ when x is rotated counter-clockwise around w. Then $Z[N]$ can be expressed as a correlation function of twist operators on a single copy of $\mathbb{C}$ rather than as a partition function on X_N,

$$Z[N] \sim \left\langle \left(\prod_{i=1}^{p} \sigma_1^0(s_i)\,\sigma_{-1}^0(t_i) \prod_{j=1}^{q} \sigma_{-1}^k(u_j)\,\sigma_1^k(v_j) \right)_{\mathcal{AO}} \right\rangle\ . \tag{43}$$

The subscript $\mathcal{AO}$ means the operators are in ascending order of coordinates. Such correlation functions can be calculated through bosonization (see e.g. ref. [36]). Diagonalization of T leads to N decoupled fields, $\widetilde{\Psi}_l$. Each $\widetilde{\Psi}_l$ is multivalued, picking up a phase $e^{-i\frac{l}{N}2\pi}$, $e^{i\frac{l}{N}2\pi}$, $e^{i\frac{l-k}{N}2\pi}$or $e^{-i\frac{l-k}{N}2\pi}$ when rotated counter-clockwise around s_i, t_i, u_i, or v_i respectively. Then one can factorize each multi-valued field $\widetilde{\Psi}_l$ into a gauge factor that describes this multi-valuedness and a single valued free Dirac field: $\Psi^l = e^{i\int_{x_0}^{x} dx'^{\mu} A_\mu^l(x)}\psi^l(x)$. The gauge field dependent part of the partition function contains the branch point dependence of $Z[N]$ and is moreover straightforward to evaluate. With the notation [37],

$$q_l(R,k) \equiv \frac{1-N}{2N} + \left\{ \frac{lR + k + (N-1)/2}{N} \right\}\ , \tag{44}$$

where the curly braces denote the fractional part of a number and $l \in \ell = \left\{-\frac{N-1}{2}, -\frac{N-1}{2}+1, ..., \frac{N-1}{2}\right\}$, the gauge field $A^l_\mu(x)$ satisfies the contour integrals

$$\oint_{C_{s_i}} dx^\mu A^l_\mu(x) = -\frac{2\pi l}{N} \ , \qquad \oint_{C_{s_i}} dx^\mu A^l_\mu(x) = \frac{2\pi l}{N} \ , \tag{45}$$

$$\oint_{C_{u_i}} dx^\mu A^l_\mu(x) = 2\pi q_l(1, N-k) \ , \quad \oint_{C_{v_i}} dx^\mu A^l_\mu(x) = 2\pi q_l(-1, k) \ . \tag{46}$$

The Lagrangian density[8] in terms of $\psi^l(x)$ becomes $\mathcal{L} = \sum_{l=1}^N \bar{\psi}^l \gamma^\mu \left(\partial_\mu + iA^l_\mu\right)\psi^l$. From eqs. (45) and (46) and Green's theorem we have:

$$\epsilon^{\mu\nu}\partial_\nu A^l_\mu(x) =$$
$$2\pi \sum_{i=1}^p \sum_{j=1}^q \left[\frac{l}{N}\left(\delta(x-s_i) - \delta(x-t_i)\right) - q_l(1, N-k)\left(\delta(x-u_i) - \delta(x-v_i)\right)\right] \ .$$

Since the ψ^l's are decoupled, the partition function becomes a product of expectation values of operators that depend on the gauge field A_μ:

$$\mathcal{T}[N] \equiv \frac{Z[N]}{(Z[1])^N} = \prod_{l\in\ell} \langle e^{i\int A^l_\mu j^\mu_l d^2x}\rangle \ , \tag{47}$$

where j^μ_l is the Dirac current $\bar{\psi}^l\gamma^\mu\psi^l$. After bosonization, it becomes $j^\mu_l = \frac{1}{2\pi}\epsilon^{\mu\nu}\partial_\nu\phi^l$. Then $\mathcal{T}[N]$ can be written as a correlation function of free boson vertex operators $V_e(w) = e^{-i\frac{e}{2}\phi_l(w)}$,

$$\prod_{l=-\frac{N-1}{2}}^{\frac{N-1}{2}} \langle e^{i\int A^l_\mu j^\mu_l d^2x}\rangle = \left\langle \prod_{i=1}^p \prod_{j=1}^q V_{2l/N}(s_i)\, V_{-2l/N}(t_i)\, V_{2q_l(-1,k)}(u_j)\, V_{2q_l(1,N-k)}(v_j) \right\rangle \ . \tag{48}$$

To evaluate the correlation function of twist operators, we use

$$\left\langle \prod_{l_i=1}^m V_{e_i}(w_i) \right\rangle = \prod_{i\neq j} |w_i - w_j|^{-e_i e_j}\, \epsilon^{-m} \tag{49}$$

where ϵ is a UV cut-off to take into account the effect of coincident points in the correlation function. We also need the sums

$$\sum_{l=-\frac{N-1}{2}}^{\frac{N-1}{2}} \frac{l^2}{N^2} = \frac{N^2-1}{12N} \ , \tag{50}$$

[8] Our conventions for the Clifford algebra are that $\{\gamma^\mu, \gamma^\nu\} = 2\delta^{\mu\nu}$. For example, we could choose $\gamma^x = \sigma^3$ and $\gamma^t = \sigma^1$

$$\sum_{l=-\frac{N-1}{2}}^{\frac{N-1}{2}} \frac{l q_l\,(1, N-k)}{N^2} = \frac{N^2-1}{12N} - \frac{(N-k)\,k}{2N}\ . \tag{51}$$

to get an explicit expression for $\mathcal{T}\,[N]$.

To shorten the expressions, we adopt the following notation: $\{s_i\} = S; \{t_i\} = T; \{u_i\} = U; \{v_i\} = V$ along with

$$[Y, Z] = \left| \prod_{y\in Y, z\in Z} (y - z) \right|\ , \quad [Y, Y] = \left| \prod_{y_1,y_2\in Y, y_1\neq y_2} (y_1 - y_2) \right|\ . \tag{52}$$

Then $\mathcal{T}\,[N]$ can be written as:

$$\mathcal{T}[N] = L^{-\frac{N^2-1}{6N}} X^{\frac{N^2-1}{6N} - \frac{(N-k)k}{N}}\ , \tag{53}$$

where we have defined

$$L \equiv \frac{[S,T]\,[U,V]}{[S,S]\,[T,T]\,[U,U]\,[V,V]\,\epsilon^{p+q}}\ , \quad X \equiv \frac{[S,V]\,[T,U]}{[S,U]\,[T,V]}\ . \tag{54}$$

Fixing the appropriate spin structures, we claim then that

$$\mathrm{tr}(O_+^N) = \mathrm{tr}(O_-^N) = \left(\frac{L}{X}\right)^{-\frac{N^2-1}{6N}}\ , \tag{55}$$

$$\mathrm{tr}[(O_+O_-)^{N/2}] = \left(\frac{L}{X}\right)^{-\frac{N^2-1}{6N}} X^{-N/4}\ . \tag{56}$$

Comparing with the two interval case (33), we can absorb c_N into the ϵ dependence of L. A nice feature of these expressions is that it is straightforward to take the $N \to 1$ limit.

2.3.1 Adjacent Limits

Let us consider adjacent limits of the two-interval negativity. We call the single-interval negativity the case when $s = v$ and $t = u$, and there is only one length scale, say $l = t - s$. We call the two-adjacent-interval negativity the case where $t = u$ and we have two length scales, $l_1 = t - s$ and $l_2 = v - u$. The single-interval and two-adjacent-interval negativities are given by a two point function and a three point function of twist fields respectively. They are therefore fully determined by conformal symmetry [20, 21]:

$$\mathcal{E}\left(N_o\right) \sim l^{-\frac{N_o^2-1}{6}} \ , \qquad \mathcal{E}\left(N_e\right) \sim l^{-\frac{N_e^2-4}{6}} \ , \tag{57}$$

$$\mathcal{E}\left(N_o\right) \sim \left(l_1 l_2 \left(l_1+l_2\right)\right)^{-\frac{N_o^2-1}{12}} \ , \qquad \mathcal{E}\left(N_e\right) \sim \left(l_1 l_2\right)^{-\frac{N_e^2-4}{12}} \left(l_1+l_2\right)^{-\frac{N_e^2+2}{12}} \tag{58}$$

While $\mathrm{tr}(O_+^N)$ simply vanishes in these coincident limits, we claim that $\mathrm{tr}[(O_+O_-)^{N/2}]$ reproduces $\mathcal{E}(N_e)$ for even N, in both the single-interval and two-adjacent-interval cases. This agreement provokes the question is there a choice of k for odd N for which $\mathcal{T}[N]$ has the correct adjacent interval limits? The answer is yes. If we choose $k = (N_o \pm 1)/2$, then

$$\mathcal{T}\left[N_o\right] = L^{-\frac{N_o^2-1}{6N_o}} X^{-\frac{N_o^2-1}{12N_o}} \ , \tag{59}$$

and this expression reproduces $\mathcal{E}(N_o)$ in the adjacent interval limits.

To see why the values $k = N_e/2$ and $k = (N_o \pm 1)/2$ are singled out, we consider the merging of twist operators $\sigma_1^k(w_i)\sigma_1^0(w_{i+1}) \to \sigma_2^k(w_i)$. The corresponding constraint on the correlation function is

$$\lim_{w_{i+1}\to w_i} \left\langle \sigma_{R_1}^{k_1}\left(w_1\right) \cdots \sigma_1^k\left(w_i\right) \sigma_1^0\left(w_{i+1}\right) \cdots \right\rangle \left|w_i - w_{i+1}\right|^{-\gamma_{i(i+1)}} = \left\langle \sigma_{R_1}^{k_1}\left(w_1\right) \cdots \sigma_2^k\left(w_i\right) \cdots \right\rangle \tag{60}$$

along with a corresponding constraint from considering $\sigma_{-1}^0\left(w_i\right)\sigma_{-1}^k\left(w_{i+1}\right)$. We have defined

$$\gamma_{ij} \equiv \sum_{l\in\ell} q_l\left(R_i, k_i\right) q_l\left(R_j, k_j\right) \ . \tag{61}$$

These constraints can only be satisfied if the following identities holds for all $l \in \ell$:

$$q_l\left(-2, k\right) = q_l\left(-1, 0\right) + q_l\left(-1, k\right) \ , \qquad q_l\left(2, k\right) = q_l\left(1, 0\right) + q_l\left(1, k\right) \ . \tag{62}$$

The k values $(N_o - 1)/2$, $N_e/2$ and $(N_o + 1)/2$ are the only solutions.

2.4 Bounds on the Negativity

We discuss three types of bounds on $\mathcal{E}\left(N\right)$ in the following subsections. The first, which follows from a triangle inequality on the Schatten p-norm, is an upper bound on the moments of the partially transposed density matrix. The second two are conjectural. We are able to demonstrate these conjectured bounds only for small $N > 1$.

The Schatten p-norm, defined as

$$\|M\|_p \equiv \left(\operatorname{tr}\left(\left(M^\dagger M\right)^{p/2}\right)\right)^{1/p} \quad , \quad p \in [1, \infty) \ , \tag{63}$$

is a generalization of the trace norm. Indeed, the Schatten 1-norm is the trace norm.

Because $\operatorname{tr}[(\rho_A^{T_2})^N]^{1/N}$ is the Schatten N-norm of $\rho_A^{T_2}$, for all even N we have by the triangle inequality that

$$\begin{aligned} \operatorname{tr}[(\rho_A^{T/2})^N] &= \left(\|\rho_A^{T_2}\|_N\right)^N \\ &\leq 2^{-N/2}\left(\|e^{i\pi/4}O_+\|_N + \|e^{-i\pi/4}O_-\|_N\right)^N \\ &= 2^{N/2}\operatorname{tr}[(O_+O_-)^{N/2}] \ . \end{aligned} \tag{64}$$

The $N \to 1$ limit of (64) leads to an upper bound on the negativity in terms of $\operatorname{tr}[(O_+O_-)^{1/2}]$

$$\begin{aligned} \mathcal{E} = \|\rho_A^{T_2}\|_T &\leq \left\|\frac{1+i}{2}O_+\right\|_T + \left\|\frac{1-i}{2}O_-\right\|_T \\ &= \sqrt{2}\operatorname{tr}[(O_+O_-)^{1/2}] = \sqrt{2}X^{-1/4} \ . \end{aligned} \tag{65}$$

We have thus established that $\operatorname{tr}[(O_+O_-)^{N/2}]$ provides a rigorous upper bound on the negativity and its Nth moments, for free fermions.

Conjecture 1: Bounds from Word Order

As we discussed briefly above, for words of a fixed, even length, we conjecture that $\operatorname{tr}(O_\pm^N)$ is the smallest and $\operatorname{tr}[(O_+O_-)^{N/2}]$ is the largest among the traces. In the notation of the previous section, we expect that the trace of an arbitrary word $O_+^{n_1}O_-^{n_2}\cdots$ of length N is bounded above and below by

$$\operatorname{tr}(O_+^N) = \operatorname{tr}[(O_+O_-)^{N/2}]X^{N/4} \leq \operatorname{tr}(O_+^{n_1}O_-^{n_2}\cdots) \leq \operatorname{tr}[(O_+O_-)^{N/2}] \ . \tag{66}$$

We can refine this conjecture on word order further. Define $s = |n_+ - n_-|$ to be the difference between the number of times n_+ that O_+ appears in a word and the times n_- that O_- appears in a word. For two words W_1 and

W_2, we conjecture that if $s(W_1) > s(W_2)$, then $\mathrm{tr}(W_1) < \mathrm{tr}(W_2)$. Indeed, we have checked this conjecture in the two interval case for small N, using the explicit representation of these traces in terms of Riemann-Siegel theta functions. See figure 3.

Given this refined conjecture on word order, we can obtain upper and lower bounds on the negativity. For an upper bound, we first consider all the terms in the binomial expansion of $\mathrm{tr}[(\rho_A^{T_2})^{N/2}]$ that appear with a positive sign such that $s \neq 0$. We replace every word with charge s with a word of charge $s-4$ and hence larger trace. (Note there will be no traces of words of charge $s-2$ with nonzero coefficient.) Because the number of words grows as the charge decreases, we will still have a net negative contribution from words of charge $s-4$. We then replace all the traces of words with negative coefficient by the yet smaller trace $\mathrm{tr}(O_+)^N$. For the words of charge $s=0$, we simply replace all of them by the larger $\mathrm{tr}[(O_+O_-)^{N/2}]$. At the end of this procedure, we find the following upper bound

$$\mathrm{tr}[(\rho_A^{T_2})^{N/2}] \leq \left[1 - \frac{1}{2^{N/2}}\binom{N}{\frac{N}{2}}\right]\mathrm{tr}(O_+^N) + \frac{1}{2^{N/2}}\binom{N}{\frac{N}{2}}\mathrm{tr}[(O_+O_-)^{N/2}] \ . \quad (67)$$

In the large N limit, the right hand side of this expression approaches

$$\sqrt{\frac{2^{N+1}}{\pi N}}\left(\mathrm{tr}[(O_+O_-)^{N/2}] - \mathrm{tr}(O_+^N)\right) \ , \quad (68)$$

which appears to be a somewhat more stringent condition than our rigorous upper bound (64).

We can obtain a lower bound in a similar fashion, reversing the procedure. We consider all the terms in the binomial expansion of $\mathrm{tr}[(\rho_A^{T_2})^{N/2}]$ that appear with negative coefficient. We replace every word with charge s by a word of charge $s-4$. All the traces will then have positive coefficient. Next, except for $\mathrm{tr}[(O_+O_-)^{N/2}]$ itself, we replace all the traces of words with the smaller $\mathrm{tr}(O_+)^N$. In this case, we find the lower bound

$$\left(1 - \frac{1}{2^{N/2-1}}\right)\mathrm{tr}(O_+^N) + \frac{1}{2^{N/2-1}}\mathrm{tr}[(O_+O_-)^{N/2}] \leq \mathrm{tr}[(\rho_A^{T_2})^{N/2}] \ . \quad (69)$$

In comparison with the conjecture we discuss next, this lower bound is not particularly stringent in the large N limit.

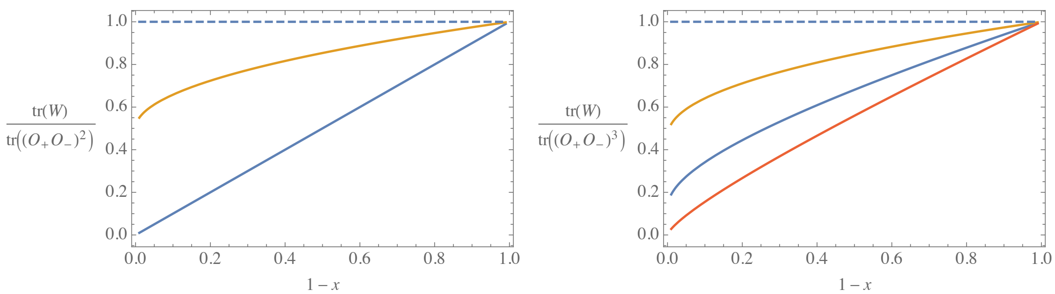
We can establish these bounds rigorously only for small N. Note that for $N=2$, the upper and lower bound reduce to the known equality (29). For

$N = 4$ and $N = 6$, we obtain the constraints

$$\frac{1}{2}(1+X)\,\mathrm{tr}[(O_+O_-)^2] \le \mathrm{tr}[(\rho_A^{T_2})^4] \le \frac{1}{2}(3-X)\,\mathrm{tr}[(O_+O_-)^2]\ , \tag{70}$$

$$\frac{1}{4}(1+3X^{3/2})\,\mathrm{tr}[(O_+O_-)^3] \le \mathrm{tr}[(\rho_A^{T_2})^6] \le \frac{1}{2}(5-3X^{3/2})\,\mathrm{tr}[(O_+O_-)^3]\ . \tag{71}$$

Indeed, in the two interval case, using the explicit representation of the negativity in terms of Riemann-Siegel theta functions, we can verify that these bounds are indeed satisfied. See the insets in figure 4.



(a) $N = 4$ (b) $N = 6$

Figure 3: Plots of ratios of traces of words versus the four point ratio $1-x$ for the two disjoint interval system. In the $N = 4$ case, we compare $\mathrm{tr}(O_+^4)$ and $\mathrm{tr}(O_+^2O_-^2)$ to $\mathrm{tr}[(O_+O_-)^2]$. The lowest curve is the ratio of $\mathrm{tr}(O_+^4)$ to $\mathrm{tr}[(O_+O_-)^2]$. In the $N = 6$, we compare $\mathrm{tr}(O_+^5O_-)$, $\mathrm{tr}(O_+O_-O_+^2O_-^2)$ and $\mathrm{tr}(O_+^3O_-^3)$ to $\mathrm{tr}[(O_+O_-)^3]$. The curve at the bottom corresponds to the ratio of $\mathrm{tr}(O_+^5O_-)$ to $\mathrm{tr}[(O_+O_-)^3]$ and establishes that $\mathrm{tr}(O_+^5O_-)$ is the smallest among the words that appears in the negativity. The dashed line is included as a guide to the eye.

For $N = 4$, we can do better and prove the inequalities in general. That $\mathrm{tr}[|O_+^2 - O_-^2|^2] \ge 0$ implies that $\mathrm{tr}(O_+^4) \le \mathrm{tr}(O_+^2O_-^2)$. Similarly, that $\mathrm{tr}[|O_+O_- - O_-O_+|^2] \ge 0$ implies that $\mathrm{tr}(O_+^2O_-^2) \le \mathrm{tr}[(O_+O_-)^2]$ and the desired inequalities on $\mathrm{tr}[(\rho_A^{T_2})^4]$ follows directly.[9] It is tempting to apply these inequalities to the case $N = 1$.

[9] Alternately, one can employ von Neumann's trace inequality.

Conjecture 2: A Lower Bound from Extremization

The plot of the two disjoint interval system suggests another possible type of lower bound on $\mathcal{E}(N)$. At least for $N = 2$, 4 and 6, and conjecturally for all even N, we find that

$$\mathrm{tr}[(O_+O_-)^{N/2}] \leq \mathcal{E}(N) = \mathrm{tr}[(\rho_A^{T_2})^N] \ . \tag{72}$$

Figure 4 is a comparison of the ratio $\mathrm{tr}[(\rho_A^{T_2})^N]/\,\mathrm{tr}[(O_+O_-)^{N/2}]$ as a function of the four point ratio x to the constant function one. We consider $N = 4$ and $N = 6$ for the two interval case only. For $N = 2$, the inequality is saturated given (29). Given the saturation, we further conjecture that the negativity itself is bounded above,

$$\mathcal{E} = |\rho_A^{T_2}| \leq \mathrm{tr}[(O_+O_-)^{1/2}] \ , \tag{73}$$

further tightening the triangle inequality (65). In the appendix, we compute the Nth moments $\mathrm{tr}[(\rho^{T_2})^N]$ and $\mathrm{tr}[(O_+O_-)^{N/2}]$ explicitly for a two-spin system in a Gaussian state. We are able to show that the bounds (72) and (73) are satisfied in this simple case.

We can try to put more structure behind this conjecture. We begin by introducing some notation. Recalling that $O_+^\dagger = O_-$ and that $O_+ = SO_-S$, we can assume without loss of generality the following block structure for $O_\pm$:

$$O_\pm = \begin{pmatrix} A & \pm B \\ \mp B^\dagger & C \end{pmatrix} \tag{74}$$

where A and C are Hermitian. It will be useful in what follows to consider

$$\alpha \equiv \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix} \ , \quad \beta \equiv \begin{pmatrix} 0 & B \\ -B^\dagger & 0 \end{pmatrix} \ , \tag{75}$$

such that $O_\pm = \alpha \pm \beta$ and, from (21), $\rho_A^{T_2} = \alpha + i\beta$. Finally, we introduce

$$\gamma \equiv S\beta = \begin{pmatrix} 0 & B \\ B^\dagger & 0 \end{pmatrix} \ , \quad \eta_\pm \equiv \begin{pmatrix} A & \pm B \\ \pm B^\dagger & -C \end{pmatrix} \ . \tag{76}$$

Note that the $\eta_\pm$ are Hermitian and that $\eta_+ = SO_+ = O_-S$ while $\eta_- = O_+S = SO_-$.

Define the function

$$f_N(\theta) \equiv \mathrm{tr}[((\alpha + e^{i\theta}\beta)(\alpha - e^{-i\theta}\beta))^{N/2}] \ . \tag{77}$$

From this definition, it follows that $f_N(\frac{\pi}{2}) = \mathrm{tr}[(\rho_A^{T_2})^N]$ and $f_N(0) = \mathrm{tr}[(O_+O_-)^{N/2}]$. This function has a few other useful properties. It is periodic, with period 2π: $f_N(\theta) = f_N(\theta + 2\pi)$. It also has two reflection symmetries. The first, $f_N(\theta) = f_N(\pi - \theta)$, follows from cyclicity of the trace:

$$\begin{aligned} f(\pi-\theta) &= \mathrm{tr}[((\alpha - e^{-i\theta}\beta)(\alpha + e^{i\theta}\beta))^{N/2}] \\ &= \mathrm{tr}[((\alpha + e^{i\theta}\beta)(\alpha - e^{-i\theta}\beta))^{N/2}] \\ &= f(\theta) \ . \end{aligned}$$

The second, $f_N(\theta) = f_N(-\theta)$, is more subtle. Consider expanding out the product of matrices inside the trace. A generic term in the product will involve n_+ factors of $e^{i\theta}\beta$ and n_- factors $-e^{-i\theta}\beta$. If $n_+ = n_-$, then the θ dependence drops out, and such terms are irrelevant for the argument that follows. Let us therefore assume $n_+ \neq n_-$. Because β is off diagonal, any term that contributes to the trace must have an even number of factors of β. Thus either n_+ and n_- are both odd or both even. For every such term, there will also be a term with n_+ factors of $-e^{-i\theta}\beta$ and n_- factors of $e^{i\theta}\beta$. This second term will always have the same sign and coefficient as the first and the same cyclic ordering of operators. Thus, we can re-express the θ dependence of the combined terms as $\cos((n_+ - n_-)\theta)$, which is an even function of θ.

The two reflection symmetries, $f(\theta) = f(-\theta)$ and $f(\theta) = f(\pi - \theta)$ along with periodicity imply that $f(\pi/2) = f(3\pi/2)$ are extrema of $f(\theta)$ as are $f(0) = f(\pi)$. If we can show that these four extrema are the only extrema in the domain $0 \leq \theta < 2\pi$, and that $f(\pi/2)$ is a local maximum (or alternatively that $f(0)$ is a local minimum), then our conjecture is proven since $f(\theta)$ is a smooth bounded function on this domain.

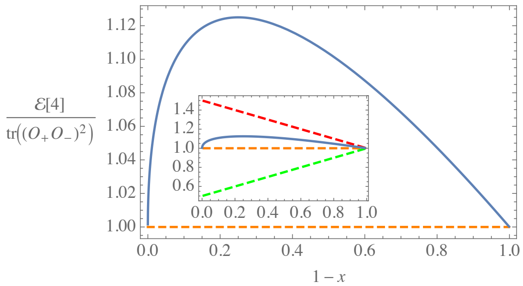
For even N, the difference between the first few $\mathcal{E}(N)$ and $\mathrm{tr}[(O_+O_-)^{N/2}]$ can be written in terms of α and γ:

$$\mathcal{E}(2) - \mathcal{T}[2] = 0 \ , \tag{78}$$

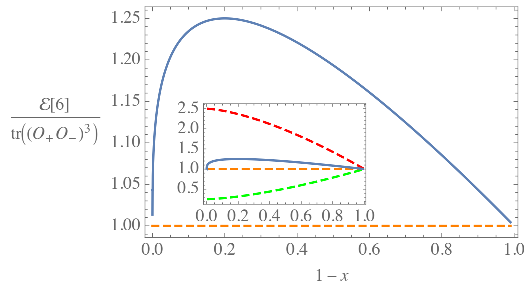
$$\mathcal{E}(4) - \mathcal{T}[4] = 4\,\mathrm{tr}\left((\alpha\gamma)^2\right) \ , \tag{79}$$

$$\mathcal{E}(6) - \mathcal{T}[6] = 6\,\mathrm{tr}\left(\left(\alpha^2+\gamma^2\right)\left((\alpha\gamma)^2 + (\gamma\alpha)^2\right)\right) \ . \tag{80}$$

A sufficient condition for $\mathrm{tr}[(O_+O_-)^{N/2}] \leq \mathcal{E}(N)$ to hold for $N = 4$ and $N = 6$ is that $(\alpha\gamma)^2 + (\gamma\alpha)^2$ be positive definite.



(a) $N = 4$



(b) $N = 6$

Figure 4: Proposed bounds on the negativity. The solid blue line is $\mathrm{tr}[(\rho_A^{T_2})^N]/\,\mathrm{tr}[(O_+O_-)^{N/2}]$. The dashed line is the constant function 1. In the insets, the upper and lower bounds (70) and (71) are included. The horizontal axis is the four point ratio $1-x$.

2.5 Comments and Future Directions

While a determination of the negativity $\mathcal{E}$ for massless free fermions in 1+1 dimensions remains an open problem, we have argued in this paper that $\mathrm{tr}[(O_+O_-)^{N/2}]$ and $\mathrm{tr}(O_+^N)$, which have simple closed form expressions for all real N, can be used to bound $\mathcal{E}$ as well as higher moments of $\rho_A^{T_2}$. One of our main results is that

$$\mathcal{E} \le \sqrt{2}\,\mathrm{tr}[(O_+O_-)^{1/2}] \ , \tag{81}$$

which follows from the triangle inequality. Part of our Conjecture 2 is that the bound can be tightened by removing the $\sqrt{2}$. Also, in the appendix, we demonstrated this tighter upper bound for a two-spin system in a Gaussian state.

For $N > 2$, we have both upper and lower bounds on the moments of $\rho_A^{T_2}$. In their strongest form, our conjectures state that

$$\begin{aligned} \mathrm{tr}[(O_+O_-)^{N/2}] &\le \mathrm{tr}[(\rho_A^{T_2})^N] \\ &\le \left[1 - \frac{1}{2^{N/2}}\binom{N}{\frac{N}{2}}\right]\mathrm{tr}(O_+)^N + \frac{1}{2^{N/2}}\binom{N}{\frac{N}{2}}\mathrm{tr}[(O_+O_-)^{N/2}] \end{aligned} \tag{82}$$

Using the triangle inequality, we were also able to argue rigorously for a somewhat weaker upper bound (64).

An advantage of working with $\mathrm{tr}[(O_+O_-)^{N/2}]$ and $\mathrm{tr}(O_+^N)$ instead of with $\mathrm{tr}[(\rho_A^{T_2})^N]$ is that they are much simpler quantities. In the paper, we discussed how to compute the multiple interval case on the plane. It is straightforward to consider the torus instead, i.e. finite volume and nonzero temperature.[10] One can even introduce a chemical potential. These generalizations require the use of the appropriate torus correlation function in place of eq. (49). See for example refs. [40, 41].

There are many interesting questions that could be asked regarding $\mathrm{tr}[(O_+O_-)^{N/2}]$. What can we deduce about the eigenvalues of $(O_+O_-)^{1/2}$ from the relation $\mathrm{tr}[(O_+O_-)^{N/2}] = L^{-\frac{N^2-1}{6N}} X^{-\frac{N^2+2}{12}}$? Can we prove the two conjectures involving $\mathrm{tr}[(O_+O_-)^{N/2}]$ discussed in the text? Among all such open questions, the most important and intriguing one is whether we can construct both an upper bound and lower bound for $\mathcal{E}\,(N)$ using $\mathrm{tr}[(O_+O_-)^{N/2}]$ that have the same $N \to 1$ limit. If so, then we can extract the value of the negativity $\mathcal{E}$ from these bounds.

[10] See ref. [39] for a discussion of subtleties associated with thermal effects on negativity.

Chapter 3

3 Graviton theorem by KLT relation

3.1 Introduction

In physics, it is very fruitful to study same thing from various angles because it will deepen our understanding and reveal many hidden relations. Now on-shell graviton scattering amplitudes can be calculated using many different ways, such as BCFW recursion relation, the double-copy formula [68], CHY formula [69, 70] and KLT formula [71] (and many more). Since the BCFW recursion relation and CHY formula have been successfully used in the study, in this note we will try to use the KLT formula to investigate the new soft graviton theorem.

Gravity amplitudes at tree level satisfy the famous Kawai-Lewellen-Tye (KLT) relation [71], with which, one can express the **stripped** tree-level gravity amplitudes M_n (i.e., the momentum conservation $\delta^4(\sum p_i)$ has been moved away) in terms of products of tree-level color-ordered **stripped** Yang-Mills amplitudes A_n and $\widetilde{A}_n$

$$M_n(1,2,\ldots,n) = \sum_{\sigma,\rho} A_n(\sigma)\mathcal{S}[\sigma|\rho]\widetilde{A}_n(\rho), \tag{83}$$

where $\mathcal{S}[\sigma|\rho]$ is called **momentum kernel**, which is a function of kinematic factors $s_{ij} = 2p_i \cdot p_j$ and depends on the permutations σ and ρ[11]. KLT relation was firstly proposed in string theory [71] and then was proved in field theory [72, 73] using BCFW recursion. One important feature should be emphasized is that KLT is relation between stripped amplitudes without imposing momentum conservation delta function.

Since KLT relation (83) connects gravity amplitudes to Yang-Mills amplitudes, it is natural to expect that the soft limit of gravity amplitudes can be derived from that of Yang-Mills amplitudes via KLT relation. In this work, we investigate this connection and its consequences. Although the KLT relation holds to general dimension, for simplicity we will focus on the pure 4D. We will show how the leading and sub-leading soft factors of gravity amplitudes can be reproduced by the leading and sub-leading soft

[11] In fact, the momentum kernel can be treated as the metric on the space of $(n-3)!$ BCJ basis.

factors of Yang-Mills amplitudes as it should be. However, to reach such now well established fact, some nontrivial relations among changing matrix of $(n-3)!$ BCJ-basis and momentum kernel $\mathcal{S}[\rho|\sigma]$ must be true. These nontrivial hidden identities are one of our main results.

The structure of this chapter is following. In section 3.2, we provide a brief review of KLT relation. In section 3.3, we recall the soft limit for stripped amplitudes of gravity and Yang-Mills theory. In section 3.4, using results in section 3.3, we present the frame of the proof of the soft graviton soft theorem via KLT relation. In section 3.5, two examples have been given to demonstrate the frame in section 3.4. In section 3.6, we summarize our work with some future directions. In appendix A, we present another more complicated example.

3.2 A review of KLT relation

In this section, we provide a brief review of various formulations of KLT relation for gravity amplitudes (for more details, please refer [72, 73]). The most general formula [50] is given as

$$\begin{aligned} M_n(1,2,\ldots,n) = (-1)^{n+1} \sum_{\sigma\in S_{n-3}} \sum_{\alpha\in S_{j-1}} \sum_{\beta\in S_{n-2-j}} & A_n(1,\sigma_{2,j},\sigma_{j+1,n-2},n-1,n) \\ & \times \mathcal{S}[\alpha_{\sigma(2),\sigma(j)}|\sigma_{2,j}]_{p_1} \mathcal{S}[\sigma_{j+1,n-2}|\beta_{\sigma(j+1),\sigma(n-2)}]_{p_{n-1}} \\ & \times \widetilde{A}_n(\alpha_{\sigma(2),\sigma(j)},1,n-1,\beta_{\sigma(j+1),\sigma(n-2)},n)\,, \end{aligned} \tag{84}$$

where A and $\widetilde{A}$ are two copies of color-ordered Yang-Mills amplitudes and the momentum kernel [72–74] is defined as

$$\mathcal{S}[i_1,i_2,...,i_k|j_1,j_2,...,j_k]_{p_1} = \prod_{t=1}^{k}(s_{i_t1}+\sum_{q>t}^{k}\theta(i_t,i_q)s_{i_ti_q}) \tag{85}$$

where p_1 is the pivot and $\theta(i_t,i_q)$ is zero when pair (i_t,i_q) has same ordering at both set $\mathcal{I}=\{i_1,i_2,...,i_k\}$, $\mathcal{J}=\{j_1,j_2,...,j_k\}$, otherwise it is one[12]. In this definition, the set $\mathcal{J}=\{j_1,j_2,...,j_k\}$ is the reference ordering set, i.e., this set provides the standard ordering. The set $\mathcal{I}=\{i_1,i_2,...,i_k\}$ is the

[12]The function $\mathcal{S}$ is nothing, but the f-function defined in [50] with more symmetric and improved expression

dynamical set which determines the dynamical factor by comparing with set $\mathcal{J}$. A few examples are the following:

$$\begin{aligned}\mathcal{S}[2,3,4|2,4,3]_{p_1} &= s_{21}(s_{31}+s_{34})s_{41},\\ \mathcal{S}[2,3,4|4,3,2]_{p_1} &= (s_{21}+s_{23}+s_{24})(s_{31}+s_{34})s_{41}.\end{aligned}$$

Although it is not so obvious, the momentum kernel, in fact, contains all BCJ-relations by following identities

$$0=\sum_{\alpha\in S_{n-2}}\mathcal{S}[\alpha(i_2,...,i_{n-1})|j_2,j_3,...,j_{n-2}]A_n(n,\alpha(i_2,...,i_{n-1}),1)\ ,\qquad \forall j\in S_{n-2} \tag{86}$$

Using (86) we can derive following relation

$$\begin{aligned}&\sum_{\alpha,\beta}\mathcal{S}[\alpha_{i_2,i_j}|i_2,\ldots,i_j]_{p_1}\mathcal{S}[i_{j+1},\ldots,i_{n-2}|\beta_{i_{j+1},i_{n-2}}]_{p_{n-1}}\\ &\qquad\widetilde{A}_n(\alpha_{i_2,i_j},1,n-1,\beta_{i_{j+1},i_{n-2}},n)\\ &=\sum_{\alpha',\beta'}\mathcal{S}[\alpha'_{i_2,i_{j-1}}|i_2,\ldots,i_{j-1}]_{p_1}\mathcal{S}[i_j,i_{j+1},\ldots,i_{n-2}|\beta'_{i_j,i_{n-2}}]\\ &\qquad\widetilde{A}_n(\alpha'_{i_2,i_{j-1}},1,n-1,\beta'_{i_j,i_{n-2}},n)\,,\end{aligned} \tag{87}$$

Thus we can shift j in (84) all the way to make the left- or right-hand part empty, *i.e.* we can choose $j=1$ or $j=n-2$. These special cases corresponds to the manifest S_{n-3}-symmetric form (88) and its dual form (89), which are given by

$$M_n(1,...,n)=(-)^{n+1}\sum_{\sigma,\widetilde{\sigma}\in S_{n-3}}A_n(1,\sigma_{2,n-2},n-1,n)\mathcal{S}[\widetilde{\sigma}_{2,n-2}|\sigma_{2,n-2}]_{p_1}\widetilde{A}_n(n-1,n,\widetilde{\sigma}_{2,n-2},1). \tag{88}$$

and

$$M_n(1,\ldots,n)=(-1)^{n+1}\sum_{\sigma,\widetilde{\sigma}\in S_{n-3}}A_n(1,\sigma_{2,n-2},n-1,n)\mathcal{S}[\sigma_{2,n-2}|\widetilde{\sigma}_{2,n-2}]_{p_{n-1}}\widetilde{A}_n(1,n-1,\widetilde{\sigma}_{2,n-2},n). \tag{89}$$

3.3 Review of soft limits of gravity and Yang-Mills theory

In this section, we review the soft behavior of gravity and Yang-Mills theory given in [49,52]. Since in KLT formula, amplitudes used are these **stripped** amplitudes, thus we will focus on the soft behaviors of these amplitudes.

We focus on the four dimensional case, thus we can use spinor variables. Under these variables, soft factors in (11) and (13) are given by [49] for gravity theory[13]

$$
\begin{aligned}
S^{(0)}_{GR} &= -\sum_{i=1}^{n}\frac{[s|i]\,\langle x|i\rangle\,\langle y|i\rangle}{\langle s|i\rangle\,\langle x|s\rangle\,\langle y|s\rangle},\\
S^{(1)}_{GR} &= -\frac{1}{2}\sum_{i=1}^{n}\frac{[s|i]}{\langle s|i\rangle}\left(\frac{\langle x|i\rangle}{\langle x|s\rangle}+\frac{\langle y|i\rangle}{\langle y|s\rangle}\right)\widetilde{\lambda}_s^{\dot\alpha}\frac{\partial}{\partial\widetilde{\lambda}_i^{\dot\alpha}}\\
S^{(2)}_{GR} &= -\frac{1}{2}\sum_{i=1}^{n}\frac{[s|i]}{\langle s|i\rangle}\widetilde{\lambda}_s^{\dot\alpha}\widetilde{\lambda}_s^{\dot\beta}\frac{\partial^2}{\partial\widetilde{\lambda}_i^{\dot\alpha}\widetilde{\lambda}_i^{\dot\beta}},
\end{aligned}
\tag{90}
$$

where x,y are two auxiliary spinors used to define the helicity of soft graviton

$$
\epsilon^{+2}=\left(\frac{\lambda_x\widetilde{\lambda}_k}{\langle x|k\rangle}\right)\left(\frac{\lambda_y\widetilde{\lambda}_k}{\langle y|k\rangle}\right)+\{x\leftrightarrow y\}\ , \tag{91}
$$

and by [52] for Yang-Mills theory[14]

$$
\begin{aligned}
S^{(0)}_{YM}(n,s,1,...) &= \frac{\langle n|1\rangle}{\langle n|s\rangle\,\langle s|1\rangle},\\
S^{(1)}_{YM}(n,s,1,...) &= \frac{1}{\langle s|1\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_1}+\frac{1}{\langle n|s\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_n}.
\end{aligned}
\tag{92}
$$

To reach these expressions, we have used the fact that in 4D, angular momentum can be written as spinor form

$$
J_{\mu\nu}\to J_{\alpha\beta}\epsilon_{\dot\alpha\dot\beta}+\widetilde{J}_{\dot\alpha\dot\beta}\epsilon_{\alpha\beta},\quad J_{\alpha\beta}=\lambda_\alpha\frac{\partial}{\partial\lambda^\beta}+\lambda_\beta\frac{\partial}{\partial\lambda^\alpha},
$$

[13] It is worth to emphasize that here we have used the QCD convention, i.e., $2p\cdot q=\langle p|q\rangle\,[q|p]$.

[14] We have assumed the color ordering is $(1,...,n,s)$.

$$\widetilde{J}_{\dot{\alpha}\dot{\beta}} = \widetilde{\lambda}_{\dot{\alpha}} \frac{\partial}{\partial \widetilde{\lambda}^{\dot{\beta}}} + \widetilde{\lambda}_{\dot{\beta}} \frac{\partial}{\partial \widetilde{\lambda}^{\dot{\alpha}}}. \tag{93}$$

We will explain the meaning of differential operators for stripped amplitudes shortly.

For stripped amplitudes, we must impose momentum conservation from beginning. This can be done as given in [49]. Under the **holomorphic soft limit** which is defined as

$$\lambda_s \to \epsilon \lambda_s, \quad \widetilde{\lambda}_s \to \widetilde{\lambda}_s \tag{94}$$

momentum conservation $\sum_{i=1}^{n} k_i + \epsilon k_s = 0$ can be used to solve two arbitrarily chosen anti-spinors $\widetilde{\lambda}_i, \widetilde{\lambda}_j$ as

$$\widetilde{\lambda}_i = -\sum_{k \neq i,j} \frac{\langle j|k\rangle}{\langle j|i\rangle} \widetilde{\lambda}_k - \epsilon \frac{\langle j|s\rangle}{\langle j|i\rangle} \widetilde{\lambda}_s, \quad \widetilde{\lambda}_j = -\sum_{k \neq i,j} \frac{\langle i|k\rangle}{\langle i|j\rangle} \widetilde{\lambda}_k - \epsilon \frac{\langle i|s\rangle}{\langle i|j\rangle} \widetilde{\lambda}_s \tag{95}$$

In other words, for stripped amplitudes, now the independent variables are $\lambda_i(i = 1, ..., n)$, λ_s, $\widetilde{\lambda}_s$ and $\widetilde{\lambda}_k$ ($k = 1, ..., n$ and $k \neq i, j$). With the fixed choice of pair (i, j), when we use the BCFW recursion relation to discuss the soft behavior as was done in [62], for example, for an $(n + 1)$-point color-ordered Yang-Mills amplitude $A(\{\epsilon\lambda_s, \widetilde{\lambda}_s\}, \{\lambda_1, \widetilde{\lambda}_1\}, ..., \{\lambda_n, \widetilde{\lambda}_n\})$ with $h_s = +1$, we will receive contributions to the singular part from the two-particle channel

$$A_{n+1}\left(\{\epsilon\lambda_s, \widetilde{\lambda}_s\}^+, 1, ..., n\right)|_{div} = A_3\left(\widehat{s}^+, 1^{h_1}, -\widehat{P}_{1s}^{-h_i}\right) \frac{1}{P_{1s}^2} A_n\left(\widehat{P}_{1s}^{h_i}, ..., \widehat{n}\right)|_{div} \tag{96}$$

under the (s, n)-shift

$$\epsilon\lambda_s(z) = \epsilon\lambda_s + z\lambda_n, \qquad \widetilde{\lambda}_n(z) = \widetilde{\lambda}_n - z\widetilde{\lambda}_s \ . \tag{97}$$

It is easy to calculate the divergent part and we find

$$\frac{-\langle n|1\rangle}{\epsilon^2 \langle n|s\rangle \langle s|1\rangle} A_n \left(\{\lambda_1, \widetilde{\lambda}_1 + \epsilon \frac{\langle n|s\rangle}{\langle n|1\rangle} \widetilde{\lambda}_s\}^{h_1}, ..., \{\lambda_i, \widetilde{\lambda}_i(\epsilon)\}, ... \right. \tag{98}$$

$$\left. \{\lambda_j, \widetilde{\lambda}_j(\epsilon)\}, ..., \{\lambda_n, \widetilde{\lambda}_n + \epsilon \frac{\langle 1|s\rangle}{\langle 1|n\rangle} \widetilde{\lambda}_s\} \right) \tag{99}$$

where (95) must be used. A compact way to rewrite above expression is to assume $\widetilde{\lambda}_i, \widetilde{\lambda}_j$ to be independent first, so we have

$$\frac{-\langle n|1\rangle}{\epsilon^2\langle n|s\rangle\langle s|1\rangle}\left\{\mathbf{e}^{\epsilon\frac{\langle n|s\rangle}{\langle n|1\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_1}-\epsilon\frac{\langle j|s\rangle}{\langle j|i\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_i}-\epsilon\frac{\langle i|s\rangle}{\langle i|j\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_j}+\frac{\epsilon\langle 1|s\rangle}{\langle 1|n\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_n}}\right.$$
$$\left.A_n\left(\{\lambda_1,\widetilde{\lambda}_1\},\ldots,\{\lambda_j,\widetilde{\lambda}_j\},\ldots,\{\lambda_n,\widetilde{\lambda}_n\}\right)\right\} \quad (100)$$

Only after the action of $\frac{\partial}{\partial\widetilde{\lambda}_i}$ and $\frac{\partial}{\partial\widetilde{\lambda}_j}$, we can replace $\widetilde{\lambda}_i,\widetilde{\lambda}_j$ by (95) with $\epsilon=0$. However, if we insist to use (95) from beginning, $\widetilde{\lambda}_i,\widetilde{\lambda}_j$ will depend on $\widetilde{\lambda}_1,\widetilde{\lambda}_n$ thus the total derivative of $\frac{d}{d\widetilde{\lambda}_1}$ and $\frac{d}{d\widetilde{\lambda}_n}$ must be written as

$$\begin{aligned}\frac{d}{d\widetilde{\lambda}_1} &= \frac{\partial}{\partial\widetilde{\lambda}_1}+\frac{\partial}{\partial\widetilde{\lambda}_i}\left(-\frac{\langle j|1\rangle}{\langle j|i\rangle}\right)+\frac{\partial}{\partial\widetilde{\lambda}_j}\left(-\frac{\langle i|1\rangle}{\langle i|j\rangle}\right)\\ \frac{d}{d\widetilde{\lambda}_n} &= \frac{\partial}{\partial\widetilde{\lambda}_n}+\frac{\partial}{\partial\widetilde{\lambda}_i}\left(-\frac{\langle j|n\rangle}{\langle j|i\rangle}\right)+\frac{\partial}{\partial\widetilde{\lambda}_j}\left(-\frac{\langle i|n\rangle}{\langle i|j\rangle}\right)\ . \end{aligned}\quad (101)$$

Using above formula, it is easy to check that

$$\begin{aligned}\frac{\langle n|s\rangle}{\langle n|1\rangle}\widetilde{\lambda}_s\frac{d}{d\widetilde{\lambda}_1} &+ \frac{\langle 1|s\rangle}{\langle 1|n\rangle}\widetilde{\lambda}_s\frac{d}{d\widetilde{\lambda}_n}\\ &= \frac{\langle n|s\rangle}{\langle n|1\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_1}+\frac{\langle 1|s\rangle}{\langle 1|n\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_n}-\frac{\langle j|s\rangle}{\langle j|i\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_i}-\frac{\langle i|s\rangle}{\langle i|j\rangle}\widetilde{\lambda}_s\frac{\partial}{\partial\widetilde{\lambda}_j},\end{aligned}$$

thus (100) becomes

$$\frac{-\langle n|1\rangle}{\epsilon^2\langle n|s\rangle\langle s|1\rangle}\left\{\mathbf{e}^{\frac{\langle n|s\rangle}{\langle n|1\rangle}\widetilde{\lambda}_s\frac{d}{d\widetilde{\lambda}_1}+\frac{\langle 1|s\rangle}{\langle 1|n\rangle}\widetilde{\lambda}_s\frac{d}{d\widetilde{\lambda}_n}}A_n\left(\{\lambda_1,\widetilde{\lambda}_1\},\ldots,\{\lambda_j,\widetilde{\lambda}_j\},\ldots,\{\lambda_n,\widetilde{\lambda}_n\}\right)\right\} \quad (102)$$

Having this new understanding, the meaning of soft factors in (90) and (92) becomes clear: *while there are no variables $\widetilde{\lambda}_i,\widetilde{\lambda}_j$ anymore in stripped amplitudes, all partial derivatives should be considered as a kind of "total derivative" in the sense of (101).*

3.4 KLT relation approach to the soft behavior of gravity amplitude

Having above preparations, now we study the soft behavior of stripped gravity amplitudes using the soft behavior of stripped Yang-Mills amplitudes as input through KLT relation. The total symmetry among the n-particles of gravity amplitudes allows us to choose any leg to be soft leg. We take p_1 to be soft and solve $n-1, n$ as

$$\begin{aligned}\widetilde{\lambda}_{n-1} &= -\sum_{k=2}^{n-2} \frac{\langle n|k\rangle}{\langle n|n-1\rangle}\widetilde{\lambda}_k - \epsilon\frac{\langle n|1\rangle}{\langle n|n-1\rangle}\widetilde{\lambda}_1, \\ \widetilde{\lambda}_n &= -\sum_{k=2}^{n-2} \frac{\langle n-1|k\rangle}{\langle n-1|n\rangle}\widetilde{\lambda}_k - \epsilon\frac{\langle n-1|1\rangle}{\langle n-1|n\rangle}\widetilde{\lambda}_1. \end{aligned} \tag{103}$$

The choice of KLT formula: In section 3.3, we have reviewed various formulations of KLT relation. To make the discussion simpler, we should start with proper choice of KLT formula. Since the leading contribution from two gluon amplitudes is the order $\frac{1}{\epsilon^2}\times\frac{1}{\epsilon^2}$ while the leading contribution of graviton amplitude is $\frac{1}{\epsilon^3}$, we are better to have manifest ϵ-factor from kernel part. Furthermore, since we have solved $\widetilde{\lambda}_{n-1}, \widetilde{\lambda}_n$ in (103), it is more convenient to have formula as less related to p_{n-1}, p_n as possible. Taking these things into consideration, we use the general formula given by (84) with $j=2$

$$\begin{aligned} M_n = (-1)^{n+1}\sum_{t=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}} \quad & A_n(1,t,\sigma,n-1,n)\mathcal{S}[t|t]_{p_1}\mathcal{S}[\sigma|\beta]_{p_{n-1}} \\ & \widetilde{A}_n(t,1,n-1,\beta,n) \end{aligned} \tag{104}$$

In this form, $\mathcal{S}[t|t]_{p_1}\to\epsilon s_{1t}$, while the expansion of the other kernel $\mathcal{S}[\sigma|\beta]_{p_{n-1}}$ can be written as[15]

$$\mathbf{e}^{+\epsilon\frac{\langle n|1\rangle}{\langle n|t\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}}\mathcal{S}[\sigma|\beta]_{p_{n-1}}. \tag{105}$$

[15]From the definition of kernel, the ϵ-expansion should be given by $\mathbf{e}^{-\epsilon\frac{\langle n|1\rangle}{\langle n|n-1\rangle}\widetilde{\lambda}_1|\frac{\partial}{\partial\widetilde{\lambda}_{n-1}}}\mathcal{S}[\sigma|\beta]_{p_{n-1}}$. However, noticing that

$$\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}\mathcal{S}[\sigma|\beta]_{p_{n-1}} = \widetilde{\lambda}_1\left(-\frac{\langle n|t\rangle}{\langle n|n-1\rangle}\right)\frac{\partial}{\partial\widetilde{\lambda}_{n-1}}\mathcal{S}[\sigma|\beta]_{p_{n-1}}$$

where we have used the fact that $\widetilde{\lambda}_s\frac{d}{d\widetilde{\lambda}_t}\mathcal{S}[\sigma|\beta]_{p_{n-1}}$ does not contain momentum p_t, we obtain (105).

For convenience, we use (92) to write down the singular soft limit of two stripped amplitudes in (104) as

$$\begin{aligned} A_n^{(n-1,n)}(1,t,\sigma,n-1,n) &\to \frac{1}{\epsilon^2}\frac{\langle n|t\rangle}{\langle n|1\rangle\langle 1|t\rangle}A_{n-1}^{(n-1,n)}(t,\sigma,n-1,n) \\ &+ \frac{1}{\epsilon}\frac{\langle n|t\rangle}{\langle n|1\rangle\langle 1|t\rangle}\left(\frac{\langle n|1\rangle}{\langle n|t\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}+\frac{\langle t|1\rangle}{\langle t|n\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_n}\right) \\ &A_{n-1}(t,\sigma,n-1,n), \end{aligned} \tag{106}$$

$$\begin{aligned} \widetilde{A}_n^{(n-1,n)}(t,1,n-1,\beta,n) &\to \frac{1}{\epsilon^2}\frac{\langle t|n-1\rangle}{\langle t|1\rangle\langle 1|n-1\rangle}\widetilde{A}_n(t,n-1,\beta,n) \\ &+ \frac{1}{\epsilon}\frac{\langle t|n-1\rangle}{\langle t|1\rangle\langle 1|n-1\rangle}\left(\frac{\langle t|1\rangle}{\langle t|n-1\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_{n-1}}+\frac{\langle n-1|1\rangle}{\langle n-1|t\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}\right) \\ &\widetilde{A}_n(t,n-1,\beta,n). \end{aligned} \tag{107}$$

In the remainder of this section, we discuss the soft behavior of gravity amplitudes by KLT relations order by order.

3.4.1 The leading order part

Substituting the leading part of color-ordered Yang-Mills amplitudes A, $\widetilde{A}$ (given by $\frac{1}{\epsilon^2}$ terms of (106), (107)) as well as the leading part of momentum kernel $\mathcal{S}$ (given by the ϵ term of $\mathcal{S}[t|t]_{p_1}\mathcal{S}[\sigma|\beta]_{p_{n-1}}$) into the KLT expression (104), we get the leading part of gravity amplitude under soft limit

$$\begin{aligned} M_n &= (-1)^{n+1}\sum_{t=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{1}{\epsilon^2}\frac{\langle n|t\rangle}{\langle n|1\rangle\langle 1|t\rangle}A_{n-1}^{(n-1,n)}(t,\sigma,n-1,n)\epsilon s_{1t}\mathcal{S}[\sigma|\beta]_{p_{n-1}^{\epsilon\to 0}} \\ &\quad \frac{1}{\epsilon^2}\frac{\langle t|n-1\rangle}{\langle t|1\rangle\langle 1|n-1\rangle}\widetilde{A}_n^{(n-1,n)}(t,n-1,\beta,n) \\ &= \frac{1}{\epsilon^3}(-1)^{n+1}\sum_{t=2}^{n-2}\frac{[t|1]}{\langle t|1\rangle}\frac{\langle n|t\rangle}{\langle n|1\rangle}\frac{\langle n-1|t\rangle}{\langle n-1|1\rangle} \\ &\quad\times\sum_{\sigma,\beta\in S_{n-4}}A_{n-1}^{(n-1,n)}(t,\sigma,n-1,n)\mathcal{S}[\sigma|\beta]_{p_{n-1}^{\epsilon\to 0}}\widetilde{A}_n^{(n-1,n)}(t,n-1,\beta,n) \\ &= \frac{1}{\epsilon^3}(-)\sum_{t=2}^{n-2}\frac{[t|1]}{\langle t|1\rangle}\frac{\langle n|t\rangle}{\langle n|1\rangle}\frac{\langle n-1|t\rangle}{\langle n-1|1\rangle}M_{n-1}(2,...,n) \end{aligned}$$

$$= \quad \frac{1}{\epsilon^3} S^{(0)}_{GR} M_{n-1}(2,...,n) \tag{108}$$

where, on the third line, we have used the S_{n-3}-symmetric KLT relation (89) for $(n-1)$-point amplitudes. The soft factor of gravity is nothing but the $S^{(0)}_{GR}$ defined in (90) with $x = n$ and $y = n-1$.

3.4.2 The subleading order part

Now let us study the subleading order of stripped gravity amplitudes under the soft limit. We will do it in three steps. In the first step, we act the $S^{(1)}_{GR}$ defined in (90) on the KLT expressions (89) of $(n-1)$-point gravity amplitudes directly. In the second step, we collect contributions of the subleading part from color ordered Yang-Mills amplitudes and momentum kernel in (104). Finally, we compare the two expressions from first two steps to prove (check) the subleading order soft factor $S^{(1)}_{GR}$ of gravity amplitude.

The sub-leading part from direct acting of $S^{(1)}_{GR}$

We use the subleading soft factor given by (90)

$$S^{(1)}_{GR,(n-1)n} \quad = \quad -\sum_{i=2}^{n} \frac{[1|i]}{\langle 1|i\rangle} \frac{\langle n-1|i\rangle}{\langle n-1|1\rangle} \widetilde{\lambda}_1 \frac{d}{d\widetilde{\lambda}_i} = -\sum_{i=2}^{n-2} \frac{[1|i]}{\langle 1|i\rangle} \frac{\langle n-1|i\rangle}{\langle n-1|1\rangle} \widetilde{\lambda}_1 \frac{d}{d\widetilde{\lambda}_i} \tag{109}$$

where we have taken the gauge choice $x = y = n-1$, thus $\frac{d}{d\widetilde{\lambda}_{n-1}} = \frac{d}{d\widetilde{\lambda}_n} = 0$. When acting it with the form (109) on M_{n-1}, for each i, we take different representation of M_{n-1}[16], i.e.,

[16]It is worth to notice that although as a whole, we have the freedom to chose x, y for $S^{(1)}_{GR,(n-1)n}$, when we act it for different i and different part $A, \widetilde{A}$ in (110), we need to stick to a particular gauge choice.

$$
\begin{aligned}
S^{(1)}_{GR,(n-1)n}M_{n-1}(2,...,n) &= -\sum_{i=2}^{n-2}\frac{[1|i]}{\langle 1|i\rangle}\frac{\langle n-1|i\rangle}{\langle n-1|1\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_i}M_{n-1}(2,...,n)\\
= & -\sum_{i=2}^{n-2}\frac{[1|i]}{\langle 1|i\rangle}\frac{\langle n-1|i\rangle}{\langle n-1|1\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_i}\Bigg[(-1)^n\sum_{\sigma,\beta\in S_{n-4}}A_{n-1}(i,\sigma,n-1,n)\\
& \mathcal{S}[\sigma|\beta]_{p_{n-1}}\widetilde{A}_{n-1}(i,n-1,\beta,n)\Bigg]\\
= & (-)^{n+1}\sum_{i=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{[1|i]}{\langle 1|i\rangle}\frac{\langle n-1|i\rangle}{\langle n-1|1\rangle}A_{n-1}(i,\sigma,n-1,n)\\
& \left(\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_i}\mathcal{S}[\sigma|\beta]_{p_{n-1}}\right)\widetilde{A}_{n-1}(i,n-1,\beta,n)\\
+ & (-)^{n+1}\sum_{i=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{[1|i]}{\langle 1|i\rangle}\frac{\langle n-1|i\rangle}{\langle n-1|1\rangle}\left(\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_i}A_{n-1}(i,\sigma,n-1,n)\right)\\
& \mathcal{S}[\sigma|\beta]_{p_{n-1}}\widetilde{A}_{n-1}(i,n-1,\beta,n)\\
+ & (-)^{n+1}\sum_{i=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{[1|i]}{\langle 1|i\rangle}\frac{\langle n-1|i\rangle}{\langle n-1|1\rangle}A_{n-1}(i,\sigma,n-1,n)\mathcal{S}[\sigma|\beta]_{p_{n-1}}\\
& \left(\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_i}\widetilde{A}_{n-1}(i,n-1,\beta,n)\right). \qquad (110)
\end{aligned}
$$

The sub-leading order part from KLT relation

Now we collect the contributions of the subleading part from the KLT relation (104). There are three contributions at this order. The first term is to take kernel to second order of ϵ, while $A,\widetilde{A}$ are the first order (see (106) and (107)). This part is given by

$$
\begin{aligned}
T_1 &= (-1)^{n+1}\sum_{t=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{[1|t]}{\langle 1|t\rangle}\frac{\langle n|t\rangle}{\langle n|1\rangle}\frac{\langle n-1|t\rangle}{\langle n-1|1\rangle}A_{n-1}(t,\sigma,n-1,n)\\
& \left(\frac{\langle n|1\rangle}{\langle n|t\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}\mathcal{S}[\sigma|\beta]_{p_{n-1}}\right)\widetilde{A}_{n-1}(t,n-1,\beta,n)
\end{aligned}
$$

$$
\begin{aligned}
&= (-1)^{n+1}\sum_{t=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{[1|t]}{\langle 1|t\rangle}\frac{\langle n-1|t\rangle}{\langle n-1|1\rangle}A_{n-1}(t,\sigma,n-1,n)\\
&\qquad\left(\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}\mathcal{S}[\sigma|\beta]_{p_{n-1}}\right)\widetilde{A}_{n-1}(t,n-1,\beta,n)
\end{aligned}\tag{111}
$$

For the second term, we keep the leading order of kernel and $\widetilde{A}$ while taking the subleading order of A, thus we have

$$
\begin{aligned}
T_2 &= (-1)^{n+1}\sum_{t=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{[1|t]}{\langle 1|t\rangle}\frac{\langle n|t\rangle\langle n-1|t\rangle}{\langle n|1\rangle\langle n-1|1\rangle}\\
&\quad\times\left[\left(\frac{\langle n|1\rangle}{\langle n|t\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}+\frac{\langle t|1\rangle}{\langle t|n\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_n}\right)A_{n-1}(t,\sigma,n-1,n)\right]\\
&\quad\times\mathcal{S}[\sigma|\beta]_{p_{n-1}}\widetilde{A}_n(t,n-1,\beta,n)\\
&= (-1)^{n+1}\sum_{t=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{[1|t]}{\langle 1|t\rangle}\frac{\langle n-1|t\rangle}{\langle n-1|1\rangle}\left(\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}A_{n-1}(t,\sigma,n-1,n)\right)\\
&\qquad\mathcal{S}[\sigma|\beta]_{p_{n-1}}\left(\widetilde{A}_n(t,n-1,\beta,n)\right),
\end{aligned}\tag{112}
$$

where we have used the fact that $\frac{d}{d\widetilde{\lambda}_n}A_{n-1}(t,\sigma,n-1,n)=0$

For the third term, we keep the leading order of kernel and A while take the subleading order of $\widetilde{A}$, thus we have

$$
\begin{aligned}
T_3 &= (-1)^{n+1}\sum_{t=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{[1|t]}{\langle 1|t\rangle}\frac{\langle n|t\rangle\langle n-1|t\rangle}{\langle n|1\rangle\langle n-1|1\rangle}A_{n-1}(t,\sigma,n-1,n)\mathcal{S}[\sigma|\beta]_{p_{n-1}}\\
&\quad\left(\left\{\frac{\langle t|1\rangle}{\langle t|n-1\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_{n-1}}+\frac{\langle n-1|1\rangle}{\langle n-1|t\rangle}\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}\right\}\widetilde{A}_n^{(n-1,n)}(t,n-1,\beta,n)\right)\\
&= (-1)^{n+1}\sum_{t=2}^{n-2}\sum_{\sigma,\beta\in S_{n-4}}\frac{[1|t]}{\langle 1|t\rangle}\frac{\langle n|t\rangle}{\langle n|1\rangle}A_{n-1}(t,\sigma,n-1,n)\mathcal{S}[\sigma|\beta]_{p_{n-1}}\\
&\qquad\left(\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_t}\widetilde{A}_n^{(n-1,n)}(t,n-1,\beta,n)\right)
\end{aligned}\tag{113}
$$

where again we have used the fact $\widetilde{\lambda}_1\frac{d}{d\widetilde{\lambda}_{n-1}}\widetilde{A}_n^{(n-1,n)}(t,n-1,\beta,n)=0$.

Comparing sub-leading parts

Now we compare (110) with T_1, T_2, T_3. It is easy to see when we identify $i = t$, we have

$$\begin{aligned}
\Delta &= S^{(1)}_{GR,(n-1)n} M_{n-1}(2,...,n) - T_1 - T_2 - T_3 \\
&= (-1)^{n+1} \sum_{t=2}^{n-2} \sum_{\sigma,\beta\in S_{n-4}} \frac{[1|t]}{\langle 1|t\rangle}\left(\frac{\langle n-1|t\rangle}{\langle n-1|1\rangle} - \frac{\langle n|t\rangle}{\langle n|1\rangle}\right) A_{n-1}(t,\sigma,n-1,n) \\
&\qquad \mathcal{S}[\sigma|\beta]_{p_{n-1}} \left(\widetilde{\lambda}_1 \frac{d}{d\widetilde{\lambda}_t} \widetilde{A}^{(n-1,n)}_n(t,n-1,\beta,n)\right) \\
&= (-1)^{n+1} \sum_{t=2}^{n-2} \sum_{\sigma,\beta\in S_{n-4}} \frac{[1|t]\,\langle n|n-1\rangle}{\langle n-1|1\rangle\,\langle n|1\rangle} A_{n-1}(t,\sigma,n-1,n) \\
&\qquad \mathcal{S}[\sigma|\beta]_{p_{n-1}} \left(\widetilde{\lambda}_1 \frac{d}{d\widetilde{\lambda}_t} \widetilde{A}^{(n-1,n)}_n(t,n-1,\beta,n)\right) \\
&= (-1)^{n+1} \sum_{t=2}^{n-2} \sum_{\sigma,\beta\in S_{n-4}} \frac{\widetilde{\lambda}^{\dot\alpha}_1 \widetilde{\lambda}^{\dot\beta}_1\,\langle n|n-1\rangle}{\langle n-1|1\rangle\,\langle n|1\rangle} A_{n-1}(t,\sigma,n-1,n) \\
&\qquad \mathcal{S}[\sigma|\beta]_{p_{n-1}} \left(\widetilde{\lambda}_{t,\dot\alpha} \frac{d}{d\widetilde{\lambda}^{\dot\beta}_t} \widetilde{A}^{(n-1,n)}_n(t,n-1,\beta,n)\right) \\
&= (-1)^{n+1} \sum_{t=2}^{n-2} \sum_{\sigma,\beta\in S_{n-4}} \frac{\widetilde{\lambda}^{\dot\alpha}_1 \widetilde{\lambda}^{\dot\beta}_1\,\langle n|n-1\rangle}{\langle n-1|1\rangle\,\langle n|1\rangle} A_{n-1}(t,\sigma,n-1,n) \\
&\qquad \mathcal{S}[\sigma|\beta]_{p_{n-1}} \left(J_{t,\dot\alpha\dot\beta} \widetilde{A}_n(t,n-1,\beta,n)\right)
\end{aligned} \tag{114}$$

It is obviously that to prove (or check) the subleading soft factor $S^{(1)}_{GR,(n-1)n}$, we need to prove (or check) $\Delta = 0$. Before going to the detail, let us notice that in (114) only the anti-spinor part of angular momentum $J_{t,\dot\alpha\dot\beta}$ appears.

Now we present the idea of proof. In (114), for each t, we have used different BCJ-basis for color ordered partial amplitudes. Thus the first step is to translate various basis into a standard basis. In other words, we should do following transformation

$$A_{n-1}(t,\sigma_t,n-1,n) = \sum_{\sigma_{\widetilde{t}}\in S_{n-4}} A_{n-1}(\widetilde{t},\sigma_{\widetilde{t}},n-1,n)\mathcal{D}[\widetilde{t},\sigma_{\widetilde{t}},n-1,n|t,\sigma_t,n-1,n]$$

$$\widetilde{A}_{n-1}(t,n-1,\beta_t,n)=\sum_{\beta_{\widetilde{t}}\in S_{n-4}}\mathcal{C}[t,n-1,\beta_t,n|\widetilde{t},n-1,\beta_{\widetilde{t}},n]\widetilde{A}_{n-1}(\widetilde{t},n-1,\beta_{\widetilde{t}},n).\quad(115)$$

where we have used the σ_t to denote the permutations of $n-4$-elements after deleting particles $1,n,n-1,t$. Inserting above transformation into the extra term (114), when we choose e.g., $\widetilde{t}=2$ in above equations, we obtain

$$\begin{aligned}
&(-1)^{n+1}\Delta\\
=&\sum_{\sigma_2,\beta_2\in S_{n-4}}\frac{\widetilde{\lambda}_1^{\dot\alpha}\widetilde{\lambda}_1^{\dot\beta}\langle n|n-1\rangle}{\langle n-1|1\rangle\langle n|1\rangle}A_{n-1}(2,\sigma_2,n-1,n)\\
&\sum_{\widetilde{t}=2}^{n-2}\sum_{\sigma_{\widetilde{t}},\beta_{\widetilde{t}}\in S_{n-4}}\mathcal{D}[2,\sigma_2,n-1,n|\widetilde{t},\sigma_{\widetilde{t}},n-1,n]\mathcal{S}[\sigma_{\widetilde{t}}|\beta_{\widetilde{t}}]_{p_{n-1}}\\
&J_{\widetilde{t},\dot\alpha\dot\beta}\left\{\mathcal{C}[\widetilde{t},n-1,\beta_{\widetilde{t}},n|2,n-1,\beta_2,n]\widetilde{A}_{n-1}(2,n-1,\beta_2,n)\right\}\\
=&\sum_{\sigma_2,\beta_2\in S_{n-4}}\frac{\widetilde{\lambda}_1^{\dot\alpha}\widetilde{\lambda}_1^{\dot\beta}\langle n|n-1\rangle}{\langle n-1|1\rangle\langle n|1\rangle}A_{n-1}(2,\sigma_2,n-1,n)\\
&\quad\sum_{\widetilde{t}=2}^{n-2}\sum_{\sigma_{\widetilde{t}},\beta_{\widetilde{t}}\in S_{n-4}}\mathcal{D}[2,\sigma_2,n-1,n|\widetilde{t},\sigma_{\widetilde{t}},n-1,n]\mathcal{S}[\sigma_{\widetilde{t}}|\beta_{\widetilde{t}}]_{p_{n-1}}\\
&\quad\mathcal{C}[\widetilde{t},n-1,\beta_{\widetilde{t}},n|2,n-1,\beta_2,n]\left\{\widetilde{J}_{\widetilde{t},\dot\alpha\dot\beta}A_{n-1}(2,n-1,\beta_2,n)\right\}\\
+&\sum_{\sigma_2,\beta_2\in S_{n-4}}\frac{\widetilde{\lambda}_1^{\dot\alpha}\widetilde{\lambda}_1^{\dot\beta}\langle n|n-1\rangle}{\langle n-1|1\rangle\langle n|1\rangle}A_{n-1}(2,\sigma_2,n-1,n)\\
&\quad\sum_{\widetilde{t}=2}^{n-2}\sum_{\sigma_{\widetilde{t}},\beta_{\widetilde{t}}\in S_{n-4}}\mathcal{D}[2,\sigma_2,n-1,n|\widetilde{t},\sigma_{\widetilde{t}},n-1,n]\mathcal{S}[\sigma_{\widetilde{t}}|\beta_{\widetilde{t}}]_{p_{n-1}}\\
&\quad\left\{J_{\widetilde{t},\dot\alpha\dot\beta}\mathcal{C}[\widetilde{t},n-1,\beta_{\widetilde{t}},n|2,n-1,\beta_2,n]\right\}\widetilde{A}_{n-1}(2,n-1,\beta_2,n)
\end{aligned}\quad(116)$$

For the first term in (116), if we have the following identity

$$\boxed{\sum_{\sigma_{\widetilde{t}},\beta_{\widetilde{t}}\in S_{n-4}}\mathcal{D}[t,\sigma_t,n-1,n|\widetilde{t},\sigma_{\widetilde{t}},n-1,n]\mathcal{S}[\sigma_{\widetilde{t}}|\beta_{\widetilde{t}}]_{p_{n-1}}\mathcal{C}[\widetilde{t},n-1,\beta_{\widetilde{t}},n|t,n-1,\beta_t,n]=\mathcal{S}[\sigma_t|\beta_t]_{p_{n-1}}},(1$$

the first term can be simplified as

$$\sum_{\sigma_2,\beta_2\in S_{n-4}}\frac{\widetilde{\lambda}_1^{\dot\alpha}\widetilde{\lambda}_1^{\dot\beta}\langle n|n-1\rangle}{\langle n-1|1\rangle\langle n|1\rangle}A_{n-1}(2,\sigma_2,n-1,n)\sum_{\widetilde{t}=2}^{n-2}\mathcal{S}[\sigma_2|\beta_2]_{p_{n-1}}\left\{J_{\widetilde{t},\dot\alpha\dot\beta}\widetilde{A}_{n-1}(2,n-1,\beta_2,n)\right\}$$

$$
\begin{aligned}
&= \sum_{\sigma_2,\beta_2\in S_{n-4}} \frac{\widetilde{\lambda}_1^{\dot\alpha}\widetilde{\lambda}_1^{\dot\beta}\langle n|n-1\rangle}{\langle n-1|1\rangle\langle n|1\rangle} A_{n-1}(2,\sigma_2,n-1,n)\mathcal{S}[\sigma_2|\beta_2]_{p_{n-1}} \\
&\quad \times \left\{ \left[\sum_{\widetilde{t}=2}^{n-2} J_{\widetilde{t},\dot\alpha\dot\beta} \right] \widetilde{A}_{n-1}(2,n-1,\beta_2,n) \right\} \\
&= 0,
\end{aligned}
$$

where we have used angular momentum conservation

$$
\begin{aligned}
& \widetilde{\lambda}_1^{\dot\alpha}\widetilde{\lambda}_1^{\dot\beta} \left\{ \sum_{\widetilde{t}=2}^{n-2} J_{\widetilde{t},\dot\alpha\dot\beta} \right\} \left\{ \widetilde{A}_{n-1}(2,n-1,\beta_2,n) \right\} \\
&= \widetilde{\lambda}_1^{\dot\alpha}\widetilde{\lambda}_1^{\dot\beta} \left\{ \sum_{\widetilde{t}=2}^{n} J_{\widetilde{t},\dot\alpha\dot\beta} \right\} \left\{ \widetilde{A}_{n-1}(2,n-1,\beta_2,n) \right\} \\
&= (-\widetilde{\lambda}_1^{\dot\alpha}\widetilde{\lambda}_1^{\dot\beta} J_{\widetilde{t}=1,\dot\alpha\dot\beta}) \left\{ \widetilde{A}_{n-1}(2,n-1,\beta_2,n) \right\} = 0 \ . \qquad (118)
\end{aligned}
$$

For the second term in (116), if we have the following identity

$$
\boxed{0 = \sum_{\widetilde{t}=2}^{n-2} \sum_{\sigma_{\widetilde{t}},\beta_{\widetilde{t}}\in S_{n-4}} \mathcal{D}[t,\sigma_t,n-1,n|\widetilde{t},\sigma_{\widetilde{t}},n-1,n]\mathcal{S}[\sigma_{\widetilde{t}}|\beta_{\widetilde{t}}]_{p_{n-1}} J_{\widetilde{t},\dot\alpha\dot\beta} \left\{ \mathcal{C}[\widetilde{t},n-1,\beta_{\widetilde{t}},n|t,n-1,\beta_t,n] \right\}},
$$

for arbitrary $t \in \{2,3,...,n-2\}$ and related $\{\sigma_t,\beta_t\}$, the contribution vanishes also.

Identities (117) and (119) are the consistency requirement of the new soft graviton theorem and the old KLT formula. While the first identity can be understood from the changing of the basis (we will discuss it shortly), the second identity is very nontrivial. Currently, we do not have an analytic proof for them although in our few examples, we have checked them explicitly. We believe the knowledge of these two identities will tell us some important aspects of momentum kernel $\mathcal{S}[\alpha|\beta]$.

Now we present the physical understanding of the first identity (117). Noticing that we have many $(n-3)!$ symmetry KLT forms. They are equivalent to each other, but it is hard to see that from the angle of BCJ relation for color-ordered Yang-Mills theory. In other words, we have

$$
\begin{aligned}
M_{n-1} &= \sum_{\sigma_t,\beta_t\in S_{n-4}} A_{n-1}(t,\sigma_t,n-1,n)\mathcal{S}[\sigma_t|\beta_t]_{p_{n-1}}\widetilde{A}_{n-1}(t,n-1,\beta_t,n) \\
&= \sum_{\sigma_{\tilde{t}},\beta_{\tilde{t}}\in S_{n-4}} A_{n-1}(\widetilde{t},\sigma_{\tilde{t}},n-1,n)\mathcal{S}[\sigma_{\tilde{t}}|\beta_{\tilde{t}}]_{p_{n-1}}\widetilde{A}_{n-1}(\widetilde{t},n-1,\beta_{\tilde{t}},n) \quad (120)
\end{aligned}
$$

where σ_t, β_t is the set of removing element t from $\{2,3,...,n-2\}$. Plugging the transformation of basis (115) back, we have

$$
\begin{aligned}
& \sum_{\sigma_{\tilde{t}},\beta_{\tilde{t}}\in S_{n-4}} A_{n-1}(\widetilde{t},\sigma_{\tilde{t}},n-1,n)\mathcal{S}[\sigma_{\tilde{t}}|\beta_{\tilde{t}}]_{p_{n-1}}\widetilde{A}_{n-1}(\widetilde{t},n-1,\beta_{\tilde{t}},n) \\
= & \sum_{\sigma_{\tilde{t}},\beta_{\tilde{t}}\in S_{n-4}} \left\{ \sum_{\sigma_t\in S_{n-4}} A_{n-1}(t,\sigma_t,n-1,n)\mathcal{D}[t,\sigma_t,n-1,n|\widetilde{t},\sigma_{\tilde{t}},n-1,n] \right\} \\
& \mathcal{S}[\sigma_{\tilde{t}}|\beta_{\tilde{t}}]_{p_{n-1}} \left\{ \sum_{\beta_t\in S_{n-4}} \mathcal{C}[\widetilde{t},n-1,\beta_{\tilde{t}},n|t,n-1,\beta_t,n]\widetilde{A}_{n-1}(t,n-1,\beta_t,n) \right\} \\
= & \sum_{\sigma_t\in S_{n-4}} A_{n-1}(t,\sigma_t,n-1,n) \\
& \left\{ \sum_{\sigma_{\tilde{t}},\beta_{\tilde{t}}\in S_{n-4}} \mathcal{D}[t,\sigma_t,n-1,n|\widetilde{t},\sigma_{\tilde{t}},n-1,n]\mathcal{S}[\sigma_{\tilde{t}}|\beta_{\tilde{t}}]_{p_{n-1}}\mathcal{C}[\widetilde{t},n-1,\beta_{\tilde{t}},n|t,n-1,\beta_t,n] \right\} \\
& \widetilde{A}_{n-1}(t,n-1,\beta_t,n) \qquad (121)
\end{aligned}
$$

Because the independence of the BCJ basis, we should obtain the identity (117).

3.4.3 The sub-sub-leading part from KLT relation

Now we consider the sub-sub-leading order. From the KLT formula, we have

$$
\begin{aligned}
& (\epsilon^{-2}A_{L,0}+\epsilon^{-1}A_{L,1}+\epsilon^{0}A_{L,2}+..)(\epsilon\mathcal{S}_0+\epsilon^2\mathcal{S}_1+\epsilon^3\mathcal{S}_2+...) \\
& (\epsilon^{-2}A_{R,0}+\epsilon^{-1}A_{R,1}+\epsilon^{0}A_{R,2}+..) \\
= & \epsilon^{-3}A_{L,0}\mathcal{S}_0A_{R,0}+\epsilon^{-2}(A_{L,1}\mathcal{S}_0A_{R,0}+A_{L,0}\mathcal{S}_1A_{R,0}+A_{L,0}\mathcal{S}_0A_{R,1}) \\
+ & \epsilon^{-1}(A_{L,2}\mathcal{S}_0A_{R,0}+A_{L,0}\mathcal{S}_2A_{R,0}+A_{L,0}\mathcal{S}_0A_{R,2}+A_{L,1}\mathcal{S}_1A_{R,0}+A_{L,0}\mathcal{S}_1A_{R,1}+A_{L,1}\mathcal{S}_0A_{R,1}) \quad (122)
\end{aligned}
$$

Thus we see that to use this formula to study the sub-sub-leading singularity, we need to get the information of $\epsilon^0 A_{L,2}$, which does not have the universal structure and has not been fully discussed.

3.5 Examples

Having the general frame in previous section, we will present a few examples to demonstrate our ideas. In this section, we will give examples of $n = 5, 6$ while the more complicated example of $n = 7$ will be given in the Appendix.

3.5.1 The case $n = 5$

Following our convention, in the stripped amplitude, $\tilde\lambda_4$ and $\tilde\lambda_5$ should be replaced by

$$\tilde\lambda_4 = -\sum_{k=2,3} \frac{\langle 5|k\rangle}{\langle 5|4\rangle}\tilde\lambda_k - \epsilon\frac{\langle 5|1\rangle}{\langle 5|4\rangle}\tilde\lambda_1,\ \tilde\lambda_5 = -\sum_{k=2,3} \frac{\langle 4|k\rangle}{\langle 4|5\rangle}\tilde\lambda_k - \epsilon\frac{\langle 4|1\rangle}{\langle 4|5\rangle}\tilde\lambda_1. \quad (123)$$

In particular that $\frac{d}{d\tilde\lambda_4^{\dot\beta}}\tilde A = \frac{d}{d\tilde\lambda_5^{\dot\beta}}\tilde A = 0$, and therefore $J_{4\dot\alpha\dot\beta}\tilde A = J_{5\dot\alpha\dot\beta}\tilde A = 0$. At 5-points it is relatively straightforward to write down all of the terms in Δ as

$$\begin{aligned}\Delta_{n=5} &= \frac{\widetilde\lambda_1^{\dot\alpha}\widetilde\lambda_1^{\dot\beta}\langle 5|4\rangle}{\langle 4|1\rangle\langle 5|1\rangle}\Big\{A_4(2,3,4,5)\mathcal{S}[3|3]_{p_4}\Big(J_{2,\dot\alpha\dot\beta}\widetilde A_4(2,4,3,5)\Big)\\ &\quad +A_4(3,2,4,5)\mathcal{S}[2|2]_{p_4}\Big(J_{3,\dot\alpha\dot\beta}\widetilde A_4(3,4,2,5)\Big)\Big\}\\ &= \frac{\widetilde\lambda_1^{\dot\alpha}\widetilde\lambda_1^{\dot\beta}\langle 5|4\rangle}{\langle 4|1\rangle\langle 5|1\rangle}\Big\{A_4(2,3,4,5)s_{34}\Big(J_{2,\dot\alpha\dot\beta}\widetilde A_4(2,4,3,5)\Big)\\ &\quad A_4(3,2,4,5)s_{24}\Big(J_{3,\dot\alpha\dot\beta}\widetilde A_4(3,4,2,5)\Big)\Big\} \end{aligned} \quad (124)$$

Now we do the changing of basis, i.e., using the BCJ relation to write

$$\begin{aligned} A_4(3,2,4,5) &= \frac{s_{34}}{s_{24}}A_4(2,3,4,5)\\ \tilde A_4(3,4,2,5) &= (-)^4\tilde A_4(5,2,4,3) = \tilde A_4(2,4,3,5) \end{aligned} \quad (125)$$

Plugging them back we get

$$\Delta_{n=5} = \frac{\widetilde\lambda_1^{\dot\alpha}\widetilde\lambda_1^{\dot\beta}\langle 5|4\rangle}{\langle 4|1\rangle\langle 5|1\rangle}A_4(3,2,4,5)s_{2\widehat{4}}\Big\{\big(J_{3,\dot\alpha\dot\beta}+J_{2,\dot\alpha\dot\beta}\big)\widetilde A_4(3,4,2,5)\Big\} = 0 \quad (126)$$

by angular momentum conservation $\sum_{i=2}^{5} J_i\widetilde{A} = 0$ (where $J_4\widetilde{A} = J_5\widetilde{A} = 0$ has been used). For this case, two identities (117) and (119) are trivial to check.

3.5.2 The case $n = 6$

For $n = 6$ the difference term $\Delta_{n=6}$ splits into three parts: $t = 2$, 3 and 4,

$$\Delta_{n=6} \;=\; \Delta_{n=6}^{t=2} + \Delta_{n=6}^{t=3} + \Delta_{n=6}^{t=4} \tag{127}$$

where we have solved

$$\begin{aligned}
\tilde{\lambda}_5 &= -\sum_{k=2,3,4} \frac{\langle 6|k\rangle}{\langle 6|5\rangle}\tilde{\lambda}_k - \epsilon\frac{\langle 6|1\rangle}{\langle 6|5\rangle}\tilde{\lambda}_1, \\
\tilde{\lambda}_6 &= -\sum_{k=2,3,4} \frac{\langle 5|k\rangle}{\langle 5|6\rangle}\tilde{\lambda}_k - \epsilon\frac{\langle 5|1\rangle}{\langle 5|6\rangle}\tilde{\lambda}_1.
\end{aligned} \tag{128}$$

For simplicity in the following discussion we suppress a common factor $(-)^{n+1}\frac{\langle n|n-1\rangle}{\langle n-1|1\rangle\langle n|1\rangle}\tilde{\lambda}_1^{\dot{\alpha}}\tilde{\lambda}_1^{\dot{\beta}}$ from the difference terms, thus we can write

$$\begin{aligned}
\Delta_{n=6}^{t=2} \;&=\; A_5(2,3,4,5,6)\mathcal{S}[3,4|3,4]_{p_5}(J_2\widetilde{A}_5(2,5,3,4,6)) \\
&+\; A_5(2,3,4,5,6)\mathcal{S}[3,4|4,3]_{p_5}(J_2\widetilde{A}_5(2,5,4,3,6)) \\
&+\; A_5(2,4,3,5,6)\mathcal{S}[4,3|3,4]_{p_5}(J_2\widetilde{A}_5(2,5,3,4,6)) \\
&+\; A_5(2,4,3,5,6)\mathcal{S}[4,3|4,3]_{p_5}(J_2\widetilde{A}_5(2,5,4,3,6)) \\
\Delta_{n=6}^{t=3} \;&=\; A_5(3,2,4,5,6)\mathcal{S}[2,4|2,4]_{p_5}(J_3\widetilde{A}_5(3,5,2,4,6)) \\
&+\; A_5(3,2,4,5,6)\mathcal{S}[2,4|4,2]_{p_5}(J_3\widetilde{A}_5(3,5,4,2,6)) \\
&+\; A_5(3,4,2,5,6)\mathcal{S}[4,2|2,4]_{p_5}(J_3\widetilde{A}_5(3,5,2,4,6)) \\
&+\; A_5(3,4,2,5,6)\mathcal{S}[4,2|4,2]_{p_5}(J_3\widetilde{A}_5(3,5,4,2,6)) \\
\Delta_{n=6}^{t=4} \;&=\; A_5(4,2,3,5,6)\mathcal{S}[2,3|2,3]_{p_5}(J_4\widetilde{A}_5(4,5,2,3,6)) \\
&+\; A_5(4,2,3,5,6)\mathcal{S}[2,3|3,2]_{p_5}(J_4\widetilde{A}_5(4,5,3,2,6)) \\
&+\; A_5(4,3,2,5,6)\mathcal{S}[3,2|2,3]_{p_5}(J_4\widetilde{A}_5(4,5,2,3,6)) \\
&+\; A_5(4,3,2,5,6)\mathcal{S}[3,2|3,2]_{p_5}(J_4\widetilde{A}_5(4,5,3,2,6))
\end{aligned} \tag{129}$$

Now we translate all amplitudes A into the basis $\{A(6,2,4,3,5), A(6,2,3,4,5)\}$

$$A_5(6,4,2,3,5) \;=\; \frac{(s_{43}+s_{45})A_5(6,2,4,3,5) + s_{45}A_5(6,2,3,4,5)}{s_{46}}$$

$$\begin{aligned} A_5(6,3,2,4,5) &= \frac{(s_{34}+s_{35})A_5(6,2,3,4,5)+s_{35}A_5(6,2,4,3,5)}{s_{36}} \\ A_5(6,4,3,2,5) &= \frac{-s_{24}s_{35}A_5(6,2,4,3,5)-s_{45}(s_{25}+s_{23})A_5(6,2,3,4,5)}{s_{46}s_{25}} \\ A_5(6,3,4,2,5) &= \frac{-s_{23}s_{45}A_5(6,2,3,4,5)-s_{35}(s_{25}+s_{24})A_5(6,2,4,3,5)}{s_{36}s_{25}} \end{aligned} \quad (130)$$

and all amplitudes $\widetilde{A}_5$ into the basis $\{\widetilde{A}_5(2,5,3,4,6), \widetilde{A}_5(2,5,4,3,6)\}$

$$\begin{aligned} \widetilde{A}_5(3,5,2,4,6) &= \frac{-\widetilde{A}_5(2,5,3,4,6)(s_{45}+s_{43})-\widetilde{A}_5(2,5,4,3,6)s_{45}}{s_{24}} \\ \widetilde{A}_5(4,5,2,3,6) &= \frac{-\widetilde{A}_5(2,5,4,3,6)(s_{35}+s_{43})-\widetilde{A}_5(2,5,3,4,6)s_{35}}{s_{23}} \\ \widetilde{A}_5(3,5,4,2,6)) &= \frac{-(s_{43}+s_{46})\widetilde{A}_5(2,5,4,3,6)-s_{46}\widetilde{A}_5(2,5,3,4,6)}{s_{24}} \\ \widetilde{A}_5(4,5,3,2,6)) &= \frac{-(s_{43}+s_{36})\widetilde{A}_5(2,5,3,4,6)-s_{36}\widetilde{A}_5(2,5,4,3,6)}{s_{23}} \end{aligned} \quad (131)$$

Putting it back with some calculation we have

$$\begin{aligned} \Delta_{n=6}^{t=3} &= A_5(2,3,4,5,6) \\ & \left\{ -s_{45}(s_{23}+s_{25})(J_{3,\dot\alpha\dot\beta}\frac{-\widetilde{A}_5(2,5,3,4,6)(s_{45}+s_{43})-\widetilde{A}_5(2,5,4,3,6)s_{45}}{s_{24}}) \right. \\ & \left. +s_{45}s_{26}(J_{3,\dot\alpha\dot\beta}\frac{-(s_{43}+s_{46})\widetilde{A}_5(2,5,4,3,6)-s_{46}\widetilde{A}_5(2,5,3,4,6)}{s_{24}}) \right\} \\ &+ A_5(2,4,3,5,6) \\ & \left\{ -s_{35}s_{24}(J_{3,\dot\alpha\dot\beta}\frac{-\widetilde{A}_5(2,5,3,4,6)(s_{45}+s_{43})-\widetilde{A}_5(2,5,4,3,6)s_{45}}{s_{24}}) \right\} \end{aligned} \quad (132)$$

Further simplification by using (notice that $J_{3,\dot\alpha\dot\beta}s_{24}=0$)

$$\begin{aligned} (s_{23}+s_{25})(J_{3,\dot\alpha\dot\beta}(s_{45}+s_{43}))-s_{26}(J_{3,\dot\alpha\dot\beta}s_{46}) &= s_{24}(J_{3,\dot\alpha\dot\beta}s_{46}) \\ (s_{23}+s_{25})(J_{3,\dot\alpha\dot\beta}s_{45})-s_{26}(J_{3,\dot\alpha\dot\beta}(s_{43}+s_{46})) &= -s_{24}(J_{3,\dot\alpha\dot\beta}s_{45}) \end{aligned} \quad (133)$$

leads

$$\Delta_{n=6}^{t=3} = A_5(2,3,4,5,6)$$

$$\begin{aligned}
&\Big\{\mathcal{S}[3,4|3,4]_{p_5}(J_{3,\dot\alpha\dot\beta}\widetilde{A}_5(2,5,3,4,6))+\mathcal{S}[3,4|4,3]_{p_5}(J_{3,\dot\alpha\dot\beta}\widetilde{A}_5(2,5,4,3,6))\\
&+s_{45}(J_{3,\dot\alpha\dot\beta}s_{46})\widetilde{A}_5(2,5,3,4,6)-s_{45}\widetilde{A}_5(2,5,4,3,6)(J_{3,\dot\alpha\dot\beta}s_{45})\Big\}\\
+\;& A_5(2,4,3,5,6)\\
&\Big\{\mathcal{S}[4,3|3,4]_{p_5}(J_{3,\dot\alpha\dot\beta}\widetilde{A}_5(2,5,3,4,6))+\mathcal{S}[4,3|4,3]_{p_5}(J_{3,\dot\alpha\dot\beta}\widetilde{A}_5(2,5,4,3,6))\\
&-s_{35}\widetilde{A}_5(2,5,3,4,6)(J_{3,\dot\alpha\dot\beta}s_{46})+s_{35}\widetilde{A}_5(2,5,4,3,6)(J_{3,\dot\alpha\dot\beta}s_{45})\Big\} \qquad (134)
\end{aligned}$$

Notice that part of them (i.e., the part with J acting only on $\widetilde{A}$) is exactly the same as $\Delta_{n=6}^{t=2}$ except the $J_{2,\dot\alpha\dot\beta}$ is replaced by $J_{3,\dot\alpha\dot\beta}$. It is nothing, but the explicit checking the identity (117) with $t=2,\widetilde{t}=3$.

Doing similar calculation we found

$$\begin{aligned}
\Delta_{n=6}^{t=4} \;=\;& A_5(2,3,4,5,6)\Big\{\mathcal{S}[3,4|3,4]_{p_5}(J_{4,\dot\alpha\dot\beta}\widetilde{A}_5(2,5,3,4,6))\\
&+\mathcal{S}[3,4|4,3]_{p_5}(J_{4,\dot\alpha\dot\beta}\widetilde{A}_5(2,5,4,3,6))\\
&+s_{45}(J_{4,\dot\alpha\dot\beta}s_{35})\widetilde{A}_5(2,5,3,4,6)\\
&+s_{45}\widetilde{A}_5(2,5,4,3,6)(J_{4,\dot\alpha\dot\beta}(s_{34}+s_{35}))\Big\}\\
+\;& A_5(2,4,3,5,6)\Big\{\mathcal{S}[4,3|3,4]_{p_5}(J_{4,\dot\alpha\dot\beta}\widetilde{A}_5(2,5,3,4,6))\\
&+\mathcal{S}[4,3|4,3]_{p_5}(J_{4,\dot\alpha\dot\beta}\widetilde{A}_5(2,5,4,3,6))\\
&-s_{35}\widetilde{A}_5(2,5,3,4,6)(J_{4,\dot\alpha\dot\beta}s_{35})\\
&+s_{35}\widetilde{A}_5(2,5,4,3,6)(J_{4,\dot\alpha\dot\beta}s_{36})\Big\} \qquad (135)
\end{aligned}$$

where again the identity (117) with $t=2,\widetilde{t}=4$ has been checked. Thus when we sum up three terms $\Delta_{n=6}^{t=2},\Delta_{n=6}^{t=3},\Delta_{n=6}^{t=4}$, the part with J acting directly on $\widetilde{A}$ vanishes by angular momentum conservation and we are left with

$$\begin{aligned}
R \;=\;& A_5(2,3,4,5,6)\Big\{+s_{45}(J_{3,\dot\alpha\dot\beta}s_{46})\widetilde{A}_5(2,5,3,4,6)-s_{45}\widetilde{A}_5(2,5,4,3,6)(J_{3,\dot\alpha\dot\beta}s_{45})\\
&+s_{45}(J_{4,\dot\alpha\dot\beta}s_{35})\widetilde{A}_5(2,5,3,4,6)-s_{45}\widetilde{A}_5(2,5,4,3,6)(J_{4,\dot\alpha\dot\beta}s_{36})\Big\}\\
+\;& A_5(2,4,3,5,6)\Big\{-s_{35}\widetilde{A}_5(2,5,3,4,6)(J_{3,\dot\alpha\dot\beta}s_{46})+s_{35}\widetilde{A}_5(2,5,4,3,6)(J_{3,\dot\alpha\dot\beta}s_{45})\\
&-s_{35}\widetilde{A}_5(2,5,3,4,6)(J_{4,\dot\alpha\dot\beta}s_{35})+s_{35}\widetilde{A}_5(2,5,4,3,6)(J_{4,\dot\alpha\dot\beta}s_{36})\Big\} \qquad (136)
\end{aligned}$$

where J acts only on s_{ij}. Using

$$\begin{aligned}
J_{3,\dot{\alpha}\dot{\beta}} s_{i5} &= -\langle 5|i\rangle \frac{\langle 6|3\rangle}{\langle 6|5\rangle} \widetilde{\lambda}_{3,\dot{\alpha}} \widetilde{\lambda}_{i,\dot{\beta}}, \\
J_{3,\dot{\alpha}\dot{\beta}} s_{i6} &= -\langle 6|i\rangle \frac{\langle 5|3\rangle}{\langle 5|6\rangle} \widetilde{\lambda}_{3,\dot{\alpha}} \widetilde{\lambda}_{i,\dot{\beta}}, \\
J_{3,\dot{\alpha}\dot{\beta}} s_{i3} &= +\langle 3|i\rangle \widetilde{\lambda}_{3,\dot{\alpha}} \widetilde{\lambda}_{i,\dot{\beta}}, \qquad i = 2,4 \\
J_{4,\dot{\alpha}\dot{\beta}} s_{i5} &= -\langle 5|i\rangle \frac{\langle 6|4\rangle}{\langle 6|5\rangle} \widetilde{\lambda}_{4,\dot{\alpha}} \widetilde{\lambda}_{i,\dot{\beta}}, \\
J_{4,\dot{\alpha}\dot{\beta}} s_{i6} &= -\langle 6|i\rangle \frac{\langle 5|4\rangle}{\langle 5|6\rangle} \widetilde{\lambda}_{4,\dot{\alpha}} \widetilde{\lambda}_{i,\dot{\beta}}, \\
J_{4,\dot{\alpha}\dot{\beta}} s_{i4} &= +\langle 4|i\rangle \widetilde{\lambda}_{4,\dot{\alpha}} \widetilde{\lambda}_{i,\dot{\beta}}, \qquad i = 2,3
\end{aligned}$$

we see immediately that $R = 0$. In other words, we have explicitly checked the second identity (119) for the special case.

3.6 Conclusion

In this paper, we studied the new soft graviton theorem from the angle of KLT relation. We have demonstrated that how the new soft gluon theorem are glued together by KLT formula to produce the corresponding soft theorem for gravity. In the process, two important identities (117) and (119) has been observed.

There are a lot of open questions deserve to be investigated. First, the two identities need an analytic proof. Secondly, the sub-sub-leading soft factor in KLT relation should be understood. Although at this order, contributions from non-universal soft part of color ordered Yang-Mills amplitudes appear, we guess that their effects will be canceled out by nice property of momentum kernel $\mathcal{S}$. It will be fascinating to see how it happens. Thirdly, in this paper, we have focused on the 4D, it will be interesting to discuss it in general dimension since KLT formula holds in general dimension. Finally, there are also other general formulas for gravity amplitudes (such as these given in [75–77]) and it will be nice to see how the new soft graviton theorem makes its appearance.

A Example with $n = 7$

In this appendix we verify that identities (117) and (119) are holding at $n = 7$. Our strategy used in previous examples applies to 7-points, although the complexity involved increases considerably. As in the previous examples, we choose to work in a convenient minimum basis $\tilde{A}(2, 6, \beta_2, 7), A(2, \sigma_2, 6, 7)$ (Here we use A instead of A_6 for short), i.e., do the following transformation with $t = 3, 4, 5$:

$$\begin{aligned}
\tilde{A}(t, 6, \beta_t, 7) &= \sum_{\beta_t \in S_3} \mathcal{C}\left[t, 6, \beta_t, 7 | 2, 6, \beta_2, 7\right] \tilde{A}(2, 6, \beta_2, 7), \qquad (137)\\
A(t, \sigma_t, 6, 7) &= \sum_{\sigma_t \in S_3} A(2, \sigma_2, 6, 7) \mathcal{D}\left[2, \sigma_2, 6, 7 | t, \sigma, 6, 7\right].
\end{aligned}$$

Our task then amounts to showing that, for both identities, terms associated with each independent product of basis amplitudes $A\tilde{A}$ match accordingly for both sides of the equations (117) and (119). In the discussion below we focus on terms containing $\tilde{A}(2, 6, 3, 4, 5, 7)$, namely when $\beta_2 = \{3, 4, 5\}$. The rest of the coefficients follow similar argument up to permutations of $\{3, 4, 5\}$. In principle it is straightforward to work out all translation coefficients $\mathcal{C}$, $\mathcal{D}$ and check if the identities are holding. However we can perform the calculation in a slightly more organized manner. In particular note that common factors are quite often shared between different translation coefficients.

For the purpose of demonstration let us consider translating a specific amplitude $\tilde{A}(3, 6, 2, 4, 5, 7)$ into minimum basis. This can be done by first expressing the amplitude in the $\tilde{A}(2, \ldots, 7)$ Kleiss-Klein (KK) basis that fixes legs 2 and 7 at both ends, and then subsequently translating to the $\tilde{A}(2, 6, \ldots, 7)$ minimum basis of interest where legs 6 and 2 are adjacent:

$$\begin{aligned}
\tilde{A}(3, 6, 2, 4, 5, 7) &= \tilde{A}(2, 4, 5, 6, 3, 7) + \tilde{A}(2, 4, 6, 5, 3, 7) + \tilde{A}(2, 4, 6, 3, 5, 7) \qquad (138)\\
&\quad + \tilde{A}(2, 6, 4, 5, 3, 7) + \tilde{A}(2, 6, 4, 3, 5, 7) + \tilde{A}(2, 6, 3, 4, 5, 7)\\
&= \left(1 - \frac{(s_{42} + s_{46} + s_{43})}{s_{42}} + E\left[45, 3 | 345\right]\right) \tilde{A}(2, 6, 3, 4, 5, 7)\\
&\quad + \ldots \left(\text{terms not contributing to } \tilde{A}(2, 6, 3, 4, 5, 7)\right),
\end{aligned}$$

where in the third line we used BCJ relation to remove the ill-favored leg 4 between 2 and 6 in the next to adjacent amplitude $\tilde{A}(2, 4, 6, 5, 3, 7)$, and we introduced the shorthand notation $E\,[45,3|345]$ to denote the next-to-next-to adjacent expansion coefficient,

$$\tilde{A}(2,\{4,5\},6,\{3\},7)=\sum_{\sigma}E\,[45,3|\sigma]\,\tilde{A}(2,6,\sigma,7). \tag{139}$$

The coefficient $E\,[45,3|345]$ can be determined from simultaneous equations consisting of BCJ relations, yielding

$$\begin{aligned}E\,[45,3|345] \;=\;& \frac{(-1)}{s_{42}s_{52}-(s_{42}+s_{45})(s_{52}+s_{54})}\\ & \Big[-(s_{42}+s_{45}+s_{46}+s_{43})(s_{52}+s_{56}+s_{53}+s_{54})\\ & +\frac{(s_{42}+s_{45})(s_{52}+s_{54}+s_{56}+s_{53})(s_{42}+s_{46}+s_{43})}{s_{42}}\Big].\end{aligned} \tag{140}$$

All translation coefficients can be determined via similar procedures. Explicitly we have, for the $t=3$ sector,

$$\begin{aligned}
\mathcal{C}\,[3,6,2,4,5,7|2,6,3,4,5,7] &= 1-\frac{(s_{42}+s_{46}+s_{43})}{s_{42}}+E\,[45,3] && (141)\\
\mathcal{C}\,[3,6,2,5,4,7|2,6,3,4,5,7] &= -\frac{(s_{52}+s_{56}+s_{53}+s_{54})}{s_{52}}+E\,[54,3]\\
\mathcal{C}\,[3,6,4,2,5,7|2,6,3,4,5,7] &= \frac{(s_{42}+s_{46}+s_{43})}{s_{42}}-E\,[45,3]-E\,[54,3]\\
\mathcal{C}\,[3,6,4,5,2,7|2,6,3,4,5,7] &= E\,[54,3]\\
\mathcal{C}\,[3,6,5,2,4,7|2,6,3,4,5,7] &= \frac{(s_{52}+s_{56}+s_{53}+s_{54})}{s_{52}}-E\,[45,3]-E\,[54,3]\\
\mathcal{C}\,[3,6,5,4,2,7|2,6,3,4,5,7] &= E\,[45,3]\,. && (142)
\end{aligned}$$

For $t=4$ we have

$$\begin{aligned}
\mathcal{C}\,[4,6,2,3,5,7|2,6,3,4,5,7] &= 1-\frac{(s_{32}+s_{36})}{s_{32}}+E\,[35,4] && (143)\\
\mathcal{C}\,[4,6,2,5,3,7|2,6,3,4,5,7] &= -\frac{(s_{52}+s_{56}+s_{53}+s_{54})}{s_{52}}+E\,[53,4]\\
\mathcal{C}\,[4,6,3,2,5,7|2,6,3,4,5,7] &= \frac{(s_{32}+s_{36})}{s_{32}}-E\,[35,4]-E\,[53,4]\\
\mathcal{C}\,[4,6,3,5,2,7|2,6,3,4,5,7] &= E\,[53,4]
\end{aligned}$$

$$\mathcal{C}[4,6,5,2,3,7|2,6,3,4,5,7] = \frac{(s_{52}+s_{56}+s_{53}+s_{54})}{s_{52}} - E[35,4] - E[53,4]$$
$$\mathcal{C}[4,6,5,3,2,7|2,6,3,4,5,7] = E[35,4],$$

and similarly for $t = 5$,

$$\mathcal{C}[5,6,2,3,4,7|2,6,3,4,5,7] = 1 - \frac{(s_{32}+s_{36})}{s_{32}} + E[34,5] \tag{144}$$
$$\mathcal{C}[5,6,2,4,3,7|2,6,3,4,5,7] = -\frac{(s_{42}+s_{46}+s_{43})}{s_{42}} + E[43,5]$$
$$\mathcal{C}[5,6,3,2,4,7|2,6,3,4,5,7] = \frac{(s_{32}+s_{36})}{s_{32}} - E[34,5] - E[43,5] \tag{145}$$
$$\mathcal{C}[5,6,3,4,2,7|2,6,3,4,5,7] = E[43,5] \tag{146}$$
$$\mathcal{C}[5,6,4,2,3,7|2,6,3,4,5,7] = \frac{(s_{42}+s_{46}+s_{43})}{s_{42}} - E[34,5] - E[43,5] \tag{147}$$
$$\mathcal{C}[5,6,4,3,2,7|2,6,3,4,5,7] = E[34,5], \tag{148}$$

whereas the next-to-next-to adjacent expansion coefficients are given by

$$E[54,3] = \frac{(-1)}{s_{52}s_{42} - (s_{52}+s_{54})(s_{42}+s_{45})} \Big[-(s_{52}+s_{54}+s_{56}+s_{53})(s_{42}+s_{46}+s_{43}) + \frac{(s_{52}+s_{54})(s_{42}+s_{45}+s_{46}+s_{43})(s_{52}+s_{56}+s_{53}+s_{54})}{s_{52}}\Big] \tag{149}$$

$$E[35,4] = \frac{(-1)}{s_{32}s_{52} - (s_{32}+s_{35})(s_{52}+s_{53})} \Big[-(s_{32}+s_{35}+s_{36})(s_{52}+s_{56}+s_{53}+s_{54}) + \frac{(s_{32}+s_{35})(s_{52}+s_{53}+s_{56}+s_{54})(s_{32}+s_{36})}{s_{32}}\Big] \tag{150}$$

$$E[53,4] = \frac{(-1)}{s_{52}s_{32} - (s_{52}+s_{53})(s_{32}+s_{35})} \tag{151}$$
$$\Big[-(s_{52}+s_{53}+s_{56}+s_{54})(s_{32}+s_{36}) + \frac{(s_{52}+s_{53})(s_{32}+s_{35}+s_{36})(s_{52}+s_{56}+s_{53}+s_{54})}{s_{52}}\Big] \tag{152}$$

$$E[34,5] = \frac{(-1)}{s_{32}s_{42} - (s_{32}+s_{34})(s_{42}+s_{43})} \tag{153}$$
$$\Big[-(s_{32}+s_{34}+s_{36})(s_{42}+s_{46}+s_{43})$$

$$+\frac{(s_{32}+s_{34})(s_{42}+s_{43}+s_{46})(s_{32}+s_{36})}{s_{32}}\Bigg]$$

$$E\,[43,5] = \frac{(-1)}{s_{42}s_{32}-(s_{42}+s_{43})(s_{32}+s_{34})} \Big[-(s_{42}+s_{43}+s_{46})(s_{32}+s_{36}) + \frac{(s_{42}+s_{43})(s_{32}+s_{34}+s_{36})(s_{42}+s_{46}+s_{43})}{s_{42}}\Bigg] \tag{154}$$

Identity (117)

Let us first verify identity (117) for the case when $\beta_2 = \{345\}$, or equivalently that

$$\sum_{\sigma_{\tilde{t}},\beta_{\tilde{t}}} A(\tilde{t},\sigma_{\tilde{t}},67)\mathcal{S}\left[\sigma_{\tilde{t}}|\beta_{\tilde{t}}\right]\mathcal{C}\left[\tilde{t},6,\beta_{\tilde{t}},7\right] = \sum_{\sigma_2\in perm\{345\}} A(2,\sigma_2,67)\mathcal{S}\left[\sigma_2|345\right]_6 \tag{155}$$

for each $\tilde{t}$. To keep the derivation simple we introduce the following shorthand notation for repeatedly occurring factors

$$T_{\tilde{t}}(\beta_{\tilde{t}}) = \sum_{\sigma_{\tilde{t}}} A(\tilde{t},\sigma_{\tilde{t}},67)\mathcal{S}\left[\sigma_{\tilde{t}}|\beta_{\tilde{t}}\right]. \tag{156}$$

so that for example when $\tilde{t}=5$, equation (155) reads

$$\sum_{\beta_5\in perm\{234\}} T_5(\beta_5)\mathcal{C}\left[5,6,\beta_5,7\right] = T_2(345). \tag{157}$$

Substituting the explicit expressions for translation coefficients $\mathcal{C}$s, the left hand side of the above equation becomes

$$\begin{aligned} & T_5(234)+(-T_5(234)+T_5(324))\,\frac{s_{32}+s_{36}}{s_{32}} \\ & +(-T_5(243)+T_5(423))\,\frac{s_{42}+S_{46}+s_{43}}{s_{42}} \\ & +(T_5(234)-T_5(324)-T_5(423)+T_5(432))\,E[34,5] \\ & +(T_5(243)-T_5(324)+T_5(342)-T_5(423))\,E[43,5]. \end{aligned} \tag{158}$$

With a little bit more effort, we find that the left hand side of (157) boils down to the following linear combination of amplitudes.

$$-s_{36}s_{46}(s_{56}+s_{54}+s_{53}+s_{52})\,A(523467) \tag{159}$$

$$-s_{36}(s_{34}+s_{46})(s_{56}+s_{54}+s_{53}+s_{52})\,A(524367).$$

On the other hand the right hand side of (157) reads

$$\begin{aligned} T_2(345) &= s_{36}s_{46}s_{56}A(234567)+s_{36}s_{46}(s_{46}+s_{56})A(235467) \\ &+s_{36}s_{46}(s_{53}+s_{54}+s_{56})A(253467)+s_{36}(s_{43}+s_{46})s_{56}A(243567) \\ &+s_{36}(s_{43}+s_{46})(s_{56}+s_{53})A(245367) \\ &+s_{36}(s_{43}+s_{46})(s_{53}+s_{54}+s_{56})A(254367), \end{aligned} \tag{160}$$

We see that the first line of (159) matches the sum of the first three terms of equation (160), and similarly the second line of (159) matches the sum of the last three terms of (160) because of BCJ relation, thereby proving the identity (157). The situations when $\tilde{t}=3$ and 4 can be proved in a likewise manner.

Identity (119)

At 7-points the difference term $\Delta_{n=7}$ splits into four parts, $\Delta_{n=7} = \sum_{t=2,3,4,5} \Delta^t_{n=7}$, where

$$\Delta^t_{n=7} = \sum_{\sigma,\beta\in S_3} A(t,\sigma,6,7)\mathcal{S}[\sigma|\beta]_6 J_t\tilde{A}(t,6,\beta,7) \tag{161}$$

Substituting the above expressions into equation (161) and collecting terms, we find as in the previous examples that terms where angular momentum operate on basis amplitudes $\tilde{A}$ add up to zero because of angular momentum conservation $\sum_t J_t\tilde{A}(2,6,3,4,5,7)=0$, leaving us with the collection of terms that J_t operate on expansion coefficients $\mathcal{C}$, which are functions of kinematic variables. Contributions from the three respective sectors are given by

$$
\begin{aligned}
\Delta_{t=3} &= (-T_3(245) + T_3(425))\, J_3\left(\frac{s_{42}+s_{46}+s_{43}}{s_{42}}\right) \\
&+ (-T_3(254) + T_3(524))\, J_3\left(\frac{s_{52}+s_{56}+s_{53}+s_{54}}{s_{52}}\right) \\
&+ (T_3(245) - T_3(425) - T_3(524) + T_3(542))\, J_3\left(E\,[45,3]\right) \\
&+ (T_3(254) - T_3(425) + T_3(452) - T_3(524))\, J_3\left(E\,[54,3]\right), \qquad (162)
\end{aligned}
$$

$$
\begin{aligned}
\Delta_{t=4} &= (-T_3(235) + T_3(325))\, J_3\left(\frac{s_{32}+s_{36}}{s_{32}}\right) \\
&+ (-T_3(253) + T_3(523))\, J_3\left(\frac{s_{52}+s_{56}+s_{53}+s_{54}}{s_{52}}\right) \\
&+ (T_3(235) - T_3(325) - T_3(523) + T_3(532))\, J_3\left(E\,[35,4]\right) \\
&+ (T_3(253) - T_3(325) + T_3(4352) - T_3(523))\, J_3\left(E\,[53,4]\right), \qquad (163)
\end{aligned}
$$

$$
\begin{aligned}
\Delta_{t=5} &= (-T_3(234) + T_3(324))\, J_3\left(\frac{s_{32}+s_{36}}{s_{32}}\right) \\
&+ (-T_3(243) + T_3(423))\, J_3\left(\frac{s_{42}+s_{46}+s_{43}}{s_{42}}\right) \\
&+ (T_3(234) - T_3(324) - T_3(423) + T_3(432))\, J_3\left(E\,[34,5]\right) \\
&+ (T_3(243) - T_3(324) + T_3(342) - T_3(423))\, J_3\left(E\,[43,5]\right), \qquad (164)
\end{aligned}
$$

Generically the operation of J_t on kinematic variables must fall into one of the following categories:

- $t = 3$,

$$
\begin{aligned}
J_{3\,\dot\alpha\dot\beta} s_{i6} &= \tilde\lambda_{3\,(\dot\alpha}\tilde\lambda_{i\,\dot\beta)}(-)\frac{\langle i6\rangle\langle 73\rangle}{\langle 76\rangle}, \qquad i = 2,4,5 \qquad (165) \\
J_{3\,\dot\alpha\dot\beta} s_{i7} &= \tilde\lambda_{3\,(\dot\alpha}\tilde\lambda_{i\,\dot\beta)}(-)\frac{\langle i7\rangle\langle 63\rangle}{\langle 67\rangle} \\
J_{3\,\dot\alpha\dot\beta} s_{i3} &= \tilde\lambda_{3\,(\dot\alpha}\tilde\lambda_{i\,\dot\beta)}\,\langle i3\rangle \qquad (166) \\
J_{3\,\dot\alpha\dot\beta} s_{i\,i'} &= 0, \qquad i, i' = 2,4,5
\end{aligned}
$$

- $t = 4$,

$$\begin{aligned}
J_{4\,\dot\alpha\dot\beta} s_{i6} &= \tilde\lambda_{4\,(\dot\alpha}\tilde\lambda_{i\,\dot\beta)}(-)\frac{\langle i6\rangle\,\langle 74\rangle}{\langle 76\rangle}, \qquad i = 3,4,5 \\
J_{4\,\dot\alpha\dot\beta} s_{i7} &= \tilde\lambda_{4\,(\dot\alpha}\tilde\lambda_{i\,\dot\beta)}(-)\frac{\langle i7\rangle\,\langle 64\rangle}{\langle 67\rangle} \\
J_{4\,\dot\alpha\dot\beta} s_{i4} &= \tilde\lambda_{4\,(\dot\alpha}\tilde\lambda_{i\,\dot\beta)}\,\langle i4\rangle \\
J_{4\,\dot\alpha\dot\beta} s_{i\,i'} &= 0, \qquad i, i' = 3,4,5
\end{aligned} \tag{167}$$

- $t = 5$,

$$\begin{aligned}
J_{5\,\dot\alpha\dot\beta} s_{i6} &= \tilde\lambda_{5\,(\dot\alpha}\tilde\lambda_{i\,\dot\beta)}(-)\frac{\langle i6\rangle\,\langle 75\rangle}{\langle 76\rangle}, \qquad i = 2,4,5 \\
J_{5\,\dot\alpha\dot\beta} s_{i7} &= \tilde\lambda_{5\,(\dot\alpha}\tilde\lambda_{i\,\dot\beta)}(-)\frac{\langle i7\rangle\,\langle 65\rangle}{\langle 67\rangle} \\
J_{5\,\dot\alpha\dot\beta} s_{i3} &= \tilde\lambda_{5\,(\dot\alpha}\tilde\lambda_{i\,\dot\beta)}\,\langle i5\rangle \\
J_{5\,\dot\alpha\dot\beta} s_{i\,i'} &= 0, \qquad i, i' = 2,4,5
\end{aligned} \tag{168}$$

Suppose if we are interested in checking terms carrying $\tilde\lambda_{3\,(\dot\alpha}\tilde\lambda_{4\,\dot\beta)}$. Before we commence an explicit calculation, note that because all of the $\mathcal{C}$s do not depend explicitly on leg 7, from the list above such a term can only be produced through $J_3(s_{46})$, $J_3(s_{43})$, $J_4(s_{36})$, $J_4(s_{34})$, which allows us to ignore the $t = 5$ sector entirely. Additionally since s_{34} happen to be absent from the $t = 4$ translation coefficient $\mathcal{C}$s, this leaves only $J_3(s_{46})$, $J_3(s_{43})$, $J_4(s_{36})$. Considering the explicit forms given by equations (165), (166) and (167) we further note that (again) because of the absence of the leg 7 dependence in all $\mathcal{C}$s, the contributions from $J_3(s_{46})$, $J_3(s_{43})$, $J_4(s_{36})$ together can only cancel through Jacobi identity $\langle 43\rangle + \frac{\langle 73\rangle\langle 64\rangle}{\langle 76\rangle} + \frac{\langle 74\rangle\langle 36\rangle}{\langle 76\rangle} = 0$. For that to happen, the contribution associated with $J_3(s_{46})$, $J_3(s_{43})$, $J_4(s_{36})$ must be exactly in the ratio $1 : 1 : -1$, in other words they must add up to

$$J_3(s_{46})X + J_3(s_{43})X + J_4(s_{36})(-X) = 0 \tag{169}$$

for some factor X. In the following discussion we shall see that indeed this is the case.

First we note that it is relatively easy to confirm that the ratio between the contributions from $J_3(s_{46})$ and $J_3(s_{43})$ is1 : 1. This can be seen by observing that the kinematic factors s_{46} and s_{43} always show up together through the combination $s_{46}+s_{43}$ in all of the translation coefficients $\mathcal{C}$ in the $t=3$ sector (see equations from (141) to (142) as well as (162)). The only part of the argument that requires explicit calculation is the ratio between $J_3(s_{46})$ and $J_4(s_{36})$. For the purpose of discussion let us tentatively call them respectively as X and Y. From equation (162) and the definition of $E\,[45,3]$ and $E\,[54,3]$, the contribution associated with $J_3(s_{46})$ reads

$$\begin{aligned}
X &= \frac{1}{s_{42}s_{52}(s_{45}+s_{42}+s_{52})}\left[s_{52}(s_{45}+s_{42}+s_{52})\left(-T_3(245)+T_3(425)\right)\right.\\
&\quad +s_{52}(s_{52}+s_{56}+s_{53}+s_{54})\left(T_3(245)-T_3(425)-T_3(524)+T_3(542)\right)\\
&\quad \left.+s_{42}(s_{52}+s_{56}+s_{53}+s_{54})\left(T_3(254)-T_3(425)+T_3(452)-T_3(524)\right)\right]\\
&= \left[s_{36}s_{56}A\left(4,2,3,5,6,7\right)+s_{36}\left(s_{35}+s_{56}\right)A\left(4,2,5,3,6,7\right)\right.\\
&\quad -s_{56}\left(s_{25}+s_{26}+s_{56}+s_{35}+s_{45}\right)A\left(4,3,2,5,6,7\right)\\
&\quad +s_{26}\left(s_{35}+s_{45}\right)A\left(4,3,5,2,6,7\right)\\
&\quad \left.-s_{36}s_{45}A\left(4,5,2,3,6,7\right)+s_{26}s_{45}A\left(4,5,3,2,6,7\right)\right] \qquad (170)
\end{aligned}$$

and similarly,

$$\begin{aligned}
Y &= \frac{1}{s_{32}s_{52}(s_{52}+s_{32}+s_{35})}\left[s_{52}(s_{52}+s_{32}+s_{35})\left(-T_4(235)+T_4(325)\right)\right.\\
&\quad +s_{32}(s_{52}+s_{53}+s_{54}+s_{56})\left(T_4(253)-T_4(325)+T_4(352)-T_4(523)\right)\\
&\quad \left.+s_{52}(s_{52}+s_{53}+s_{54}+s_{56})\left(T_4(235)-T_4(325)-T_4(523)+T_4(532)\right)\right]\\
&= \left[s_{36}s_{56}A\left(4,2,3,5,6,7\right)+s_{36}\left(s_{35}+s_{56}\right)A\left(4,2,5,3,6,7\right)\right.\\
&\quad -s_{56}\left(s_{25}+s_{26}+s_{56}+s_{35}+s_{45}\right)A\left(4,3,2,5,6,7\right)\\
&\quad +s_{26}\left(s_{35}+s_{45}\right)A\left(4,3,5,2,6,7\right)\\
&\quad \left.-s_{36}s_{45}A\left(4,5,2,3,6,7\right)+s_{26}s_{45}A\left(4,5,3,2,6,7\right)\right] \qquad (171)
\end{aligned}$$

Now that we have the explicit formulas of the $J_3(s_{46})$ and $J_4(s_{36})$ term contributions, it is evident from (170) and (171) that they are related by an exchange of legs 3 and 4, $Y=X|_{3\leftrightarrow 4}$. Therefore to prove $X=-Y$ it suffices to show that Y is antisymmetric with respect to indices 3 and 4. This antisymmetric structure will become manifest after some nontrivial manipulations, which we perform in the following.

First of all note that BCJ relation allows us to write

$$
\begin{aligned}
&s_{36}s_{56}A\left(4,2,3,5,6,7\right)+s_{36}\left(s_{35}+s_{56}\right)A\left(4,2,5,3,6,7\right)\\
&=-s_{36}\left(s_{52}+s_{53}+s_{56}\right)A\left(4,5,2,3,6,7\right)\\
&-s_{36}\left(s_{52}+s_{53}+s_{56}+s_{54}\right)A\left(4,2,3,6,7,5\right)
\end{aligned}
\tag{172}
$$

and

$$
\begin{aligned}
&s_{26}\left(s_{35}+s_{45}\right)A\left(4,3,5,2,6,7\right)+s_{26}s_{45}A\left(4,5,3,2,6,7\right)\\
&=-s_{26}\left(s_{35}+s_{45}+s_{25}\right)A\left(4,3,2,5,6,7\right)\\
&-s_{26}\left(s_{35}+s_{45}+s_{25}+s_{65}\right)A\left(4,3,2,6,5,7\right).
\end{aligned}
\tag{173}
$$

Plugging the above two identities into the expression for Y, we have

$$
\begin{aligned}
Y &= -s_{36}\left(s_{52}+s_{53}+s_{56}+s_{54}\right)\left[A\left(4,5,2,3,6,7\right)+A\left(4,2,3,6,7,5\right)\right]\\
&\quad -\left[s_{26}\left(s_{35}+s_{45}+s_{25}\right)+s_{56}\left(s_{25}+s_{26}+s_{56}+s_{35}+s_{45}\right)\right]A\left(4,3,2,5,6,7\right)\\
&\quad -s_{26}\left(s_{35}+s_{45}+s_{25}+s_{65}\right)A\left(4,3,2,6,5,7\right)\\
&= -\left(s_{52}+s_{53}+s_{56}+s_{54}\right)\\
&\quad \times\left[s_{36}A\left(4,5,2,3,6,7\right)+s_{36}A\left(4,2,3,6,7,5\right)+\left(s_{26}+s_{56}\right)A\left(4,3,2,5,6,7\right)\right.\\
&\quad \left.+s_{26}A\left(4,3,2,6,5,7\right)\right]\\
&= s_{57}\left[s_{36}A\left(4,5,2,3,6,7\right)+s_{36}A\left(4,2,3,6,7,5\right)+\left(s_{26}+s_{56}\right)A\left(4,3,2,5,6,7\right)\right.\\
&\quad \left.+s_{26}A\left(4,3,2,6,5,7\right)\right]
\end{aligned}
\tag{174}
$$

Further using BCJ relation identifies the sum of last two terms above with

$$
\begin{aligned}
&\left(s_{26}+s_{56}\right)A\left(4,3,2,5,6,7\right)+s_{26}A\left(4,3,2,6,5,7\right)\\
&=-\left(s_{26}+s_{56}+s_{76}\right)A\left(4,3,2,5,7,6\right)-\left(s_{26}+s_{56}+s_{76}+s_{46}\right)A\left(4,6,3,2,5,7\right)\\
&=\left(s_{36}+s_{46}\right)A\left(4,3,2,5,7,6\right)+s_{36}A\left(4,6,3,2,5,7\right)
\end{aligned}
\tag{175}
$$

Therefore Y simplifies as

$$
\begin{aligned}
Y &= s_{57}s_{36}\left[A\left(4,5,2,3,6,7\right)+A\left(4,2,3,6,7,5\right)+A\left(4,3,2,5,7,6\right)+A\left(4,6,3,2,5,7\right)\right]\\
&\quad +s_{57}s_{46}A\left(4,3,2,5,7,6\right)\\
&= s_{57}\left[s_{46}A\left(4,3,2,5,7,6\right)-s_{36}A\left(3,4,2,5,7,6\right)\right]
\end{aligned}
\tag{176}
$$

where we used $U(1)$ decoupling identity to substitute the summation in the first line with a single amplitude. The final simplified formula of Y is manifestly antisymmetric under the exchange of indices 3 and 4, and we conclude that $X+Y=0$ as claimed.

References

[1] C. P. Herzog and Y. Wang, J. Stat. Mech. **1607**, no. 7, 073102 (2016) doi:10.1088/1742-5468/2016/07/073102 [arXiv:1601.00678 [hep-th]].

[2] Y. J. Du, B. Feng, C. H. Fu and Y. Wang, JHEP **1411**, 090 (2014) doi:10.1007/JHEP11(2014)090 [arXiv:1408.4179 [hep-th]].

[3] P. Mathieu P. Di Francesco and D. Senechal. Conformal Field Theory. Springer, 1997.

[4] J. Polchinski. String theory, volume 1 and 2. Cambridge University Press, 1998.

[5] Kac, Victor (1998), Vertex algebras for beginners, University Lecture Series, 10 (2nd ed.), American Mathematical Society, ISBN 0-8218-1396-X

[6] A. Einstein, B. Podolsky and N. Rosen, Can quantum mechanical description of physical reality be considered complete?, Phys. Rev. 47 (1935) 777 780.

[7] J. Maldacena and L. Susskind, Fortsch. Phys. **61**, 781 (2013) doi:10.1002/prop.201300020 [arXiv:1306.0533 [hep-th]].

[8] G. 't Hooft, Nucl. Phys. B **256**, 727 (1985). doi:10.1016/0550-3213(85)90418-3

[9] D. Beckman, A. N. Chari, S. Devabhaktuni and J. Preskill, Phys. Rev. A **54**, 1034 (1996) doi:10.1103/PhysRevA.54.1034 [quant-ph/9602016].

[10] S. Ryu and T. Takayanagi, Phys. Rev. Lett. **96**, 181602 (2006) doi:10.1103/PhysRevLett.96.181602 [hep-th/0603001].

[11] L. K. Grover, quant-ph/9605043.

[12] H. Casini and M. Huerta, "Entanglement entropy in free quantum field theory," J. Phys. A **42**, 504007 (2009) [arXiv:0905.2562 [hep-th]].

[13] P. Calabrese and J. Cardy, "Entanglement entropy and conformal field theory," J. Phys. A **42**, 504005 (2009) [arXiv:0905.4013 [cond-mat.stat-mech]].

[14] I. Peschel and V. Eisler, "Reduced density matrices and entanglement entropy in free lattice models," J. Phys. A **42**, 504003 (2009) [arXiv:0906.1663 [cond-mat]].

[15] A. Peres, "Separability criterion for density matrices," Phys. Rev. Lett. **77**, 1413 (1996) [quant-ph/9604005].

[16] M. Horodecki, P. Horodecki and R. Horodecki, "On the necessary and sufficient conditions for separability of mixed quantum states," Phys. Lett. A **223**, 1 (1996) [quant-ph/9605038].

[17] [Lee et al.(2000)]Lee, J., Kim, M. S., Park, Y. J., & Lee, S. 2000, Journal of Modern Optics, 47, 2151

[18] G. Vidal and R. F. Werner, "Computable measure of entanglement," Phys. Rev. A **65**, 032314 (2002).

[19] M. B. Plenio, "Logarithmic Negativity: A Full Entanglement Monotone That is not Convex," Phys. Rev. Lett. **95**, 090503 (2005).

[20] P. Calabrese, J. Cardy and E. Tonni, "Entanglement negativity in quantum field theory," Phys. Rev. Lett. **109**, 130502 (2012) [arXiv:1206.3092 [cond-mat.stat-mech]].

[21] P. Calabrese, J. Cardy and E. Tonni, "Entanglement negativity in extended systems: A field theoretical approach," J. Stat. Mech. **1302**, P02008 (2013) [arXiv:1210.5359 [cond-mat.stat-mech]].

[22] O. Blondeau-Fournier, O. A. Castro-Alvaredo and B. Doyon, "Universal scaling of the logarithmic negativity in massive quantum field theory," arXiv:1508.04026 [hep-th].

[23] K. Audenaert, J. Eisert, M. B. Plenio and R. F. Werner, "Entanglement Properties of the Harmonic Chain," Phys. Rev. A **66**, 042327 (2002).

[24] V. Eisler and Z. Zimboras, "On the partial transpose of fermionic Gaussian states," New J. Phys. **16**, 123020 (2014) [arXiv:1502.01369 [cond-mat.stat-mech]].

[25] A. Coser, E. Tonni and P. Calabrese, "Partial transpose of two disjoint blocks in XY spin chains," arXiv:1503.09114 [cond-mat.stat-mech].

[26] H. Wichterich, J. Molina-Vilaplana, and S. Bose, "Scaling of entanglement between separated blocks in spin chains at criticality," Phys. Rev. A **80**, 010304(R) (2009) [arXiv:0811.1285 [quant-ph]].

[27] V. Alba, "Entanglement negativity and conformal field theory: a Monte Carlo study," J. Stat. Mech. P05013 (2013) [arXiv:1302.1110 [cond-mat.stat-mech]].

[28] P. Calabrese, L. Tagliacozzo, and E. Tonni, "Entanglement negativity in the critical Ising chain," J. Stat. Mech., P05002 (2013) [arXiv:1302.1113 [cond-mat.stat-mech]].

[29] A. Coser, E. Tonni and P. Calabrese, "Towards entanglement negativity of two disjoint intervals for a one dimensional free fermion," arXiv:1508.00811 [cond-mat.stat-mech].

[30] A. Coser, E. Tonni and P. Calabrese, "Spin structures and entanglement of two disjoint intervals in conformal field theories," arXiv:1511.08328 [cond-mat.stat-mech].

[31] V. Eisler and Z. Zimboras, "Entanglement negativity in two-dimensional free lattice models," arXiv:1511.08819 [cond-mat.stat-mech].

[32] P. Calabrese, J. Cardy and E. Tonni, "Entanglement entropy of two disjoint intervals in conformal field theory," J. Stat. Mech. **0911**, P11001 (2009) [arXiv:0905.2069 [hep-th]].

[33] A. Nakayashiki, "On the Thomae Formula for $\mathbb{Z}_N$ Curves," Publ. RIMS, **33** 987 (1997).

[34] H. M. Farkas and S. Zemel, "Generalizations of Thomae's Formula for Z_n Curves," Dev. Math. **21**, 1 (2011).

[35] A. Coser, L. Tagliacozzo and E. Tonni, "On Rényi entropies of disjoint intervals in conformal field theory," J. Stat. Mech. **2014**, P01008 (2014) [arXiv:1309.2189 [hep-th]].

[36] H. Casini, C. D. Fosco and M. Huerta, "Entanglement and alpha entropies for a massive Dirac field in two dimensions," J. Stat. Mech. **0507**, P07007 (2005) [cond-mat/0505563].

[37] M. Bershadsky and A. Radul, "Fermionic Fields On Z(n) Curves," Commun. Math. Phys. **116**, 689 (1988).

[38] L. J. Dixon, D. Friedan, E. J. Martinec and S. H. Shenker, "The Conformal Field Theory of Orbifolds," Nucl. Phys. B **282**, 13 (1987).

[39] P. Calabrese, J. Cardy and E. Tonni, "Finite temperature entanglement negativity in conformal field theory," J. Phys. A **48**, no. 1, 015006 (2015) [arXiv:1408.3043 [cond-mat.stat-mech]].

[40] N. Ogawa, T. Takayanagi and T. Ugajin, "Holographic Fermi Surfaces and Entanglement Entropy," JHEP **1201**, 125 (2012) [arXiv:1111.1023 [hep-th]].

[41] C. P. Herzog and T. Nishioka, "Entanglement Entropy of a Massive Fermion on a Torus," JHEP **1303**, 077 (2013) [arXiv:1301.0336 [hep-th]].

[42] F.E. Low, Phys. Rev. **96**, 1428 (1954), Phys. Rev. **110**, 974 (1958); M. Gell-Mann and M.L. Goldberger, Phys. Rev. **96**, 1433 (1954); S. Saito, Phys. Rev. **184**, 1894 (1969);

[43] S. Weinberg, "Photons and Gravitons in s Matrix Theory: Derivation of Charge Conservation and Equality of Gravitational and Inertial Mass," Phys. Rev. **135** (1964) B1049.

S. Weinberg, "Infrared photons and gravitons," Phys. Rev. **140** (1965) B516.

[44] A. Strominger, "On BMS Invariance of Gravitational Scattering," arXiv:1312.2229 [hep-th].

[45] T. He, V. Lysov, P. Mitra and A. Strominger, "BMS supertranslations and Weinberg's soft graviton theorem," arXiv:1401.7026 [hep-th].

[46] D. Kapec, V. Lysov, S. Pasterski and A. Strominger, "Semiclassical Virasoro Symmetry of the Quantum Gravity S-Matrix," arXiv:1406.3312 [hep-th].

[47] R. Britto, F. Cachazo and B. Feng, "New recursion relations for tree amplitudes of gluons," Nucl. Phys. B **715** (2005) 499 [hep-th/0412308].

[48] 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].

[49] F. Cachazo and A. Strominger, “Evidence for a New Soft Graviton Theorem,” arXiv:1404.4091 [hep-th].

[50] Z. Bern, L. J. Dixon, M. Perelstein and J. S. Rozowsky, Nucl. Phys. B **546**, 423 (1999) [hep-th/9811140].

[51] E. Laenen, G. Stavenga and C. D. White, JHEP **0903**, 054 (2009) [arXiv:0811.2067 [hep-ph]].

E. Laenen, L. Magnea, G. Stavenga and C. D. White, JHEP **1101**, 141 (2011) [arXiv:1010.1860 [hep-ph]].

C. D. White, JHEP **1105**, 060 (2011) [arXiv:1103.2981 [hep-th]].

[52] E. Casali, “Soft sub-leading divergences in Yang-Mills amplitudes,” arXiv:1404.5551 [hep-th].

[53] J. Broedel, M. de Leeuw, J. Plefka and M. Rosso, “Constraining sub-leading soft gluon and graviton theorems,” arXiv:1406.6574 [hep-th].

[54] Z. Bern, S. Davies, P. Di Vecchia and J. Nohle, “Low-Energy Behavior of Gluons and Gravitons from Gauge Invariance,” arXiv:1406.6987 [hep-th].

[55] C. D. White, “Diagrammatic insights into next-to-soft corrections,” arXiv:1406.7184 [hep-th].

[56] A. J. Larkoski, “Conformal Invariance of the Subleading Soft Theorem in Gauge Theory,” arXiv:1405.2346 [hep-th].

[57] B. U. W. Schwab and A. Volovich, “Subleading soft theorem in arbitrary dimension from scattering equations,” arXiv:1404.7749 [hep-th].

[58] N. Afkhami-Jeddi, “Soft Graviton Theorem in Arbitrary Dimensions,” arXiv:1405.3533 [hep-th].

[59] M. Zlotnikov, “Sub-sub-leading soft-graviton theorem in arbitrary dimension,” arXiv:1407.5936 [hep-th].

[60] C. Kalousios and F. Rojas, “Next to subleading soft-graviton theorem in arbitrary dimensions,” arXiv:1407.5982 [hep-th].

[61] Z. Bern, S. Davies and J. Nohle, “On Loop Corrections to Subleading Soft Behavior of Gluons and Gravitons,” arXiv:1405.1015 [hep-th].

[62] S. He, Y. -t. Huang and C. Wen, “Loop Corrections to Soft Theorems in Gauge Theories and Gravity,” arXiv:1405.1410 [hep-th].

[63] F. Cachazo and E. Y. Yuan, “Are Soft Theorems Renormalized?,” arXiv:1405.3413 [hep-th].

[64] B. U. W. Schwab, “Subleading Soft Factor for String Disk Amplitudes,” arXiv:1406.4172 [hep-th].

[65] M. Bianchi, S. He, Y. -t. Huang and C. Wen, “More on Soft Theorems: Trees, Loops and Strings,” arXiv:1406.5155 [hep-th].

[66] T. Adamo, E. Casali and D. Skinner, arXiv:1405.5122 [hep-th].

[67] Y. Geyer, A. E. Lipstein and L. Mason, arXiv:1406.1462 [hep-th].

[68] Z. Bern, J. J. M. Carrasco and H. Johansson, Phys. Rev. D **78**, 085011 (2008) [arXiv:0805.3993 [hep-ph]].

[69] F. Cachazo, S. He and E. Y. Yuan, arXiv:1307.2199 [hep-th].

[70] F. Cachazo, S. He and E. Y. Yuan, JHEP **1407**, 033 (2014) [arXiv:1309.0885 [hep-th]].

[71] H. Kawai, D. Lewellen and H. Tye, ”A Relation Betwwen Tree Amplitudes of Closed and Open Strings”, Nucl.Phys.B269 (1986)1.

[72] N. E. J. Bjerrum-Bohr, P. H. Damgaard, B. Feng and T. Sondergaard, “Gravity and Yang-Mills Amplitude Relations,” Phys. Rev. D **82** (2010) 107702 [arXiv:1005.4367 [hep-th]].

[73] N. E. J. Bjerrum-Bohr, P. H. Damgaard, B. Feng and T. Sondergaard, “Proof of Gravity and Yang-Mills Amplitude Relations,” JHEP **1009** (2010) 067 [arXiv:1007.3111 [hep-th]].

[74] N. E. J. Bjerrum-Bohr, P. H. Damgaard, T. Sondergaard and P. Vanhove, "The Momentum Kernel of Gauge and Gravity Theories," JHEP **1101** (2011) 001 [arXiv:1010.3933 [hep-th]].

[75] F. Cachazo and Y. Geyer, arXiv:1206.6511 [hep-th].

[76] F. Cachazo and D. Skinner, Phys. Rev. Lett. **110**, no. 16, 161301 (2013) [arXiv:1207.0741 [hep-th]].

[77] F. Cachazo, arXiv:1301.3970 [hep-th].

[78] C. Cheung, A. de la Fuente and R. Sundrum, JHEP **1701**, 112 (2017) doi:10.1007/JHEP01(2017)112 [arXiv:1609.00732 [hep-th]].

[79] S. Pasterski, S. H. Shao and A. Strominger, arXiv:1706.03917 [hep-th].

[80] C. Cardona and Y. t. Huang, arXiv:1702.03283 [hep-th].

[81] L. Susskind, AIP Conf. Proc. **493**, 98 (1999) doi:10.1063/1.1301570 [hep-th/9901079].

[82] J. Polchinski, hep-th/9901076.

[83] M. Gary, S. B. Giddings and J. Penedones, Phys. Rev. D **80**, 085005 (2009) doi:10.1103/PhysRevD.80.085005 [arXiv:0903.4437 [hep-th]].

[84] J. Penedones, JHEP **1103**, 025 (2011) doi:10.1007/JHEP03(2011)025 [arXiv:1011.1485 [hep-th]].

[85] M. F. Paulos, J. Penedones, J. Toledo, B. C. van Rees and P. Vieira, arXiv:1607.06110 [hep-th].

[86] S. Pasterski, S. H. Shao and A. Strominger, arXiv:1701.00049 [hep-th].

[87] A. Strominger and A. Zhiboedov, JHEP **1601**, 086 (2016) doi:10.1007/JHEP01(2016)086 [arXiv:1411.5745 [hep-th]].