DESY 08–102
TTP 08-31
SFB/CPP-08-52
July 2008

Automating dipole subtraction*

K. Hasegawa[a], S. Moch[a], and P. Uwer[b]

[a]Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, D-15738 Zeuthen, Germany

[b]Institut für Theoretische Teilchenphysik, Universität Karlsruhe, D-76128 Karlsruhe, Germany

We report on automating the Catani-Seymour dipole subtraction which is a general procedure to treat infrared divergences in real emission processes at next-to-leading order in QCD. The automatization rests on three essential steps: the creation of the dipole terms, the calculation of the color linked squared Born matrix elements, and the evaluation of different helicity amplitudes. The routines have been tested for a number of complex processes, such as the real emission process $gg \to t\bar{t}ggg$.

1. INTRODUCTION

QCD as the gauge theory of the strong interaction allows to predict cross sections for hard scattering reactions which, at a hadron collider, typically involve high multiplicities of colored partons. In the perturbative approach calculations based on exact QCD matrix elements at leading order (LO) provide first estimates for cross sections and differential distributions. However, in the complicated environment of a hadron collider precision calculations to next-to-leading-order (NLO) in QCD are often needed in order to reliably predict (and separate) the Standard Model background from possible new physics signals. In the era of LHC this has triggered a lot of activity concerning NLO QCD corrections to multi-particle reactions, see e.g. Refs. [1,2].

The salient feature of NLO corrections is the presence of virtual and real emission contributions. Virtual-loop corrections exhibit both ultraviolet (UV) and soft and collinear divergences which we call IR divergencies in the following. The real corrections contain only IR divergencies from soft and collinear emissions. Upon summation of the two parts all IR divergences cancel (for so-called IR safe observables) [3–5]. Since virtual and real corrections have different phase space integrals, these cancellations are not always trivial. In the Catani-Seymour dipole formalism [6–8], the IR divergences of virtual and real corrections are treated separately by subtracting suitable dipole terms so that each of the contributions becomes individually finite. The dipole terms are constructed systematically relying on the universal nature of soft and collinear limits in QCD. Thus the method allows for a general treatment of IR divergences to NLO in QCD.

Current applications in phenomenology consider processes with six or more parton legs [1,2] which require about one hundred dipole terms. These calculations are rather tedious and since the algorithm underlying the Catani-Seymour dipole subtraction is a combinatorial one automatization is favored. The construction of the complete subtraction terms relies to a large extent on squaring color correlated Born amplitudes and dressing them with the corresponding dipoles. To that end, we can use existing software for the automatic evaluation of Born amplitudes [9–16] by means of a suitable interface. Thus, the completely automatic generation of all subtraction terms in the dipole formalism becomes feasible.

In related work Ref. [17] recently reported on details of an automatization, although the code is unpublished. Ref. [18] made code publicly available. However, all process dependent information, i.e. the (color correlated) Born squared matrix elements, still have to be provided by the user. Automating this step is precisely what we are aiming at in the present article.

2. ALGORITHM

In this section we briefly review the algorithm of the dipole subtraction with particular emphasis on the features of the real emission contributions for a given

*Presented by K.H. at *Loops and Legs in Quantum Field Theory*, 20–25 April 2008, Sondershausen (Germany).

process.

1. Choose all possible emitter pairs from the external legs. In the dipole subtraction, the root of the splitting of the quarks and gluons is called emitter. For convenience we call the two fields into which an emitter splits, emitter pair. We use indices, i, j, and k, for fields in a final state and, respectively, indices, a and b, for an initial state field. The quark (anti-quark) and gluon are denoted by f ($\bar{f}$) and g. In case both partons of an emitter pair are in the final state, possible combinations are $(i, j) = (1)(f,g), (2)(g,g), (3)(f,\bar{f})$. In case of one parton in the initial and the other in the final state, we have $(a,i) = (4)(f,g), (5)(g,g), (6)(f,f), (7)(g,f)$. There are also the other combinations where the quarks are replaced by the anti-quarks in the cases, (1), (4), (6), and (7).

2. Choose all possible spectators for each emitter pair. The spectator is one external field which is different from both fields of the emitter pair. For a spectator in the final (initial) state denoted by k (b), this condition means $k \neq i, j$ ($b \neq a$). It emerges from a special feature of the subtraction formalism namely that the color factors of the square terms $|\mathrm{M}_i|^2$ are expressed through the ones of the interference terms $\mathrm{M}_i\mathrm{M}_j^*(i \neq j)$ due to color conservation.

3. Construct the dipole terms from the chosen combinations of emitter and spectator. The previous steps provide all such combinations as pairings (emitter, spectator)= $(ij,k), (ij,b), (ai,k)$, and (ai,b). Each case corresponds to one dipole term, $\mathrm{D}_{ij,k}, \mathrm{D}_{ij}^{a}, \mathrm{D}_{k}^{ai}$, and $\mathrm{D}^{ai,b}$, respectively, and explicit expressions are given in [6,8]. For example, the dipole term $\mathrm{D}_{ij,k}$ in the massless case reads,

$$\mathrm{D}_{ij,k} = \frac{-1}{2p_i \cdot p_j} \langle ij,k| \frac{\mathrm{T}_k \cdot \mathrm{T}_{ij}}{\mathrm{T}_{ij}^2} \mathrm{V}_{ij,k} |ij,k\rangle \,. \tag{1}$$

The quantity, $\langle ij,k \; |\mathrm{T}_k \cdot \mathrm{T}_{ij}| \; ij,k\rangle$, is called color linked Born squared matrix element (CLBS). It is given by the Born amplitude squared with two additional color operator insertions at the emitter and spectator legs. The color operator T denotes either a fundamental t_{ij}^a or an adjoint f^{abc} generator, depending on the parton type (i.e. quark, anti-quark or gluon). The quantity $\mathrm{V}_{ij,k}$ is the so-called dipole splitting function. In case of the emitter being a quark, for example, $(ij,k) = (fg,k)$ the splitting function is diagonal in the spin space of the quark:

$$\langle s|\mathrm{V}_{fg,k}|s'\rangle = 8\pi C_F \alpha_s \left[\frac{2}{1 - z_i(1 - y_{ij,k})} - (1 + z_i)\right] \delta_{ss'} \,, \tag{2}$$

where z_i and $y_{ij,k}$ are functions of the external momenta as,

$$z_i = \frac{p_i \cdot p_k}{p_j \cdot p_k + p_i \cdot p_k} \,, \tag{3}$$

$$y_{ij,k} = \frac{p_i \cdot p_j}{p_i \cdot p_j + p_j \cdot p_k + p_k \cdot p_i} \,. \tag{4}$$

In case of the emitter being a gluon, for example, $(ij,k) = (f\bar{f},k)$, it does exhibit a correlation with the gluon helicity according to

$$\langle \mu|\mathrm{V}_{f\bar{f},k}|\nu\rangle = 8\pi T_R \alpha_s \left[-g^{\mu\nu} - \frac{2}{p_i p_j}(z_i p_i^\mu - z_j p_j^\mu)(z_i p_i^\nu - z_j p_j^\nu)\right] \,, \tag{5}$$

which has to be treated accordingly (see for example Ref. [19]).

In summary, these three steps generate all dipole terms to subtract all IR divergences of a real emission process at NLO. The construction of the integrated dipole terms is less involved and proceeds in complete analogy [6,8], thus we skip a detailed discussion here.

3. CODE STRUCTURE

Here we briefly sketch our implementation of the dipole subtraction formalism. There is a large freedom how such an implementation can be done. In Ref. [20] for example the implementation was done in form of two independent C/C++ libraries providing all the necessary functions to evaluate the dipole terms. The slight disadvantage of this approach is that the produced code is non-local and that there is some redundancy in the calculation. In the present work we follow a different approach. The main idea here is to have code generator which will produce an optimized flat code which can be further optimized by the compiler. To do so we haven chosen to interface a Mathematica program with MadGraph [9,10].

3.1. Mathematica code

We implement the creation of the dipole terms in Mathematica. With a given (real emission) parton scattering process as an input, the Mathematica code automatically writes down all dipole terms needed at NLO. It provides all expressions explicitly except for the CLBS. The code creates the dipole terms in an order according to the kind of the emitter pairs, i.e. the seven combinations of (i,j) or (a,i) listed in Sec. 2.

The first group of dipole terms (dipole 1) are the ones with the emitter pairs, (1),(2),(4), and (5). These emitters reduce the NLO real emission process to a Born amplitude which is the LO contribution to a process with one less gluon in the final state. The second kind of dipole terms (dipole 2) has the emitter pair (3), while the third and fourth (dipole 3 and 4) have (6) and (7) as emitter pairs.

In order to demonstrate the code, let us discuss the example $g(a)g(b) \to u(1)\bar{u}(2)g(3)$. The code starts with the creation of the first dipole $\mathrm{D}_{13,2}$ which belongs to the group dipole 1 and the output is written in the form,

$$\mathrm{D}_{ijkfgk}(132) = \frac{-1}{2p_i \cdot p_j} \mathrm{V}_{ijkfgk}(132) \frac{\mathrm{B1}(132)}{\mathrm{T}_{13}^2}\,, \tag{6}$$

where V_{ijkfgk} is the dipole splitting function and B1 denotes the CLBS. The indices, $ijkfgk(132)$, of D and V mean that (emitter, spectator)=(ij,k)=(quark gluon, something)=(13,2). About the CLBS the code stores only the necessary information for the direct calculation. It writes each CLBS as B'i' corresponding to Dipole 'i', where i can be 1,2,3, or 4. For instance, the output for B1(132) in Eq. (6) is returned in the form,

$$\begin{aligned}\mathrm{B1}(132) = \mathrm{B1}[\{\{\mathrm{g,pa}\},\{\mathrm{g,pb}\}\} - - > \\ \{\{\mathrm{u,pijtil[1,3]}\},\{\mathrm{ubar,pktil[2]}\}\}]\,,\end{aligned} \tag{7}$$

where $gg \to u\bar{u}$ is the reduced Born process. The function pijtil[1,3] is the reduced momenta for the emitter and pktil[2] for the spectator. In general, for a given NLO real emission process with n parton legs, each dipole term has a reduced Born squared matrix element with $(n-1)$ parton legs. The reduced $(n-1)$ external momenta are functions of the original n external momenta. For example, the reduced momentum for an emitter in the dipole term $\mathrm{D}_{ij,k}$ reads

$$\widetilde{p}^{\mu}_{ij} = p^{\mu}_{i} + p^{\mu}_{j} - \frac{y_{ij,k}}{1-y_{ij,k}} p^{\mu}_{k}\,. \tag{8}$$

The Mathematica code provides explicit expressions for the reduced momenta of each dipole in the output according to Eq. (7).

At the end of the run, the total number of generated dipoles is shown. The final output for the complete subtraction term is given as a C- or Fortran routine for numerical evaluation. We have tested the generation of dipoles with our Mathematica code for complex processes, obtaining, e.g. twenty seven dipole terms for the process $gg \to u\bar{u}g$, eighty dipoles for $gg \to u\bar{u}d\bar{d}g$, and one hundred dipoles for $gg \to u\bar{u}ggg$.

3.2. Interface to MadGraph

As we see, the remaining ingredient at this stage is the CLBS appearing in all dipole terms. In case the emitter is a gluon, they also include the different helicity components of the CLBS for the emitter. In order to obtain these quantities in an automatic way, it is advantageous to use a publicly available software for automated LO calculations. We choose MadGraph for this purpose and interface our Mathematica program with the stand-alone version [9,10].

Let us briefly explain our interface to MadGraph to obtain the CLBS. In MadGraph, the color factors are separated from each diagram. In the evaluation everything is expressed in terms of generators of the fundamental representation. A typical example is that the factor f^{abc} of the gluon three point vertex is rewritten in terms of the fundamental generator t^{a}_{ij} due to the identity,

$$f^{abc} = -2i\big(\,\mathrm{Tr}[t^a t^b t^c] - \mathrm{Tr}[t^c t^b t^a]\,\big)\,. \tag{9}$$

The color factors of each diagram are sorted in an unique order and they are expressed in a sum of some terms. When a specific term of a diagram is identical to one of the other diagrams, it is combined as

$$\mathrm{M} = \sum_a C_a \mathrm{J}_a\,, \tag{10}$$

where C_a denotes the independent color factors. Each C_a has fundamental and adjoint color indices corresponding to the external quarks (anti-quarks) and the gluons, respectively. J_a is the joint amplitude, e.g. $\mathrm{J}_1 = +A_1 - A_3 + \cdots$ where A_i is the partial amplitude of i-th diagram (with the color factor stripped off). The invariant matrix element squared is finally expressed in the form,

$$|\mathrm{M}|^2 = \big(\vec{\mathrm{J}}\big)^{\dagger}\, \mathrm{CF}\, \vec{\mathrm{J}}\,, \tag{11}$$

where the color matrix CF is defined as

$$(\mathrm{CF})_{ab} = \sum_{\mathrm{color}} C_a^* C_b\,. \tag{12}$$

For the CLBS we need to evaluate Eq. (11) with an insertion of two additional color operators to the emitter and spectator legs. This is precisely what our interface to MadGraph does.

Let us return to the example of $\mathrm{B1}(123)$ from the previous subsection. The reduced Born process $g(a)g(b) \to u(1)\bar{u}(2)$ has three diagrams and the color factors are combined into two independent ones, $(C_1, C_2) = ((t^a t^b)_{12}, (t^b t^a)_{12})$. The components of the color matrix are written in the traces, $(\mathrm{CF})_{11} = (\mathrm{CF})_{22} = \mathrm{Tr}[t^b t^a t^a t^b]$ and $(\mathrm{CF})_{12} = (\mathrm{CF})_{21} = \mathrm{Tr}[t^b t^a t^b t^a]$. Then the color matrix is calculated as

$$\mathrm{CF} = \begin{pmatrix} 16/3 & -2/3 \\ -2/3 & 16/3 \end{pmatrix}. \tag{13}$$

For the CLBS $\mathrm{B1}(123)$ we need the fundamental operator insertions into the legs 1 and 2. The components of the color matrix CF are modified to $(\mathrm{CF}')_{11} = \mathrm{Tr}[t^b t^a t^c t^a t^b t^c]$ and $(\mathrm{CF}')_{12} = \mathrm{Tr}[t^b t^a t^c t^b t^a t^c]$. Then the modified color matrix is calculated as

$$\mathrm{CF}' = \begin{pmatrix} 1/9 & 10/9 \\ 10/9 & 1/9 \end{pmatrix}. \tag{14}$$

The subroutines of MadGraph for the color factor calculations are well structured and the original routines to add the color factors t^a_{ij} and f^{abc} can be applied to the additional color insertions for CLBS. We have realized the two color insertions in an automatic way and checked that MadGraph with our interface works for rather involved processes. One of the most complex checks consists of the two color insertions into the process $g(a)g(b) \to u(1)\bar{u}(2)g(3)g(4)$. In MadGraph the normal color matrix for the process is a 24 by 24 matrix. Here we show only the first 15 components in the first row as

$$\mathrm{CF} = \frac{1}{54}(512, 8, -64, 80, 8, -10, \\ -1, -64, -64, 8, -1, -10, -1, 62, -10, \cdots). \tag{15}$$

Next, we perform two adjoint operator insertions into the legs 3 and 4, and the extended routines calculate the modified color matrix as

$$\mathrm{CF}' = \frac{1}{4}(8, 0, 8, 16, 0, -2, \\ 0, 8, -1, -1, 1, 2, -8, -7, 1, \cdots). \tag{16}$$

We have checked that the result in Eq. (16) and the sum of all components agree with results of our independent private code.

Next we briefly comment on the different helicity components of the CLBS for the gluon emitter. In MadGraph the gluon polarization vector is calculated by the subroutine 'VXXXXX' of the HELAS library where the polarization vector is taken in the circular polarization representation [21,22]. Then, it is favorable to calculate the dipole terms with the helicity correlation in the circular polarization basis. For example, in the previous process $g(a)g(b) \to u(1)\bar{u}(2)g(3)$, we take one dipole D_1^{a3} and calculate it in the circular polarization basis as $\mathrm{V}_1^{a3}(\lambda, \lambda')\mathrm{B1}(g(\lambda, \lambda')g \to u\bar{u})$ where $\mathrm{V}_1^{a3}(\lambda, \lambda')$ is constructed from the dipole splitting function in the basis of the Lorentz indices as in Eq. (5) by multiplying the circular polarization vectors. The arguments (λ, λ') are the ones for the different helicities of the original and the complex conjugated amplitude for the emitter gluon. We have completed the interface to obtain the CLBS with the different helicities as $\mathrm{Bi}(\lambda, \lambda')$ and checked it in processes as the above $\mathrm{B1}(g(\lambda, \lambda')g \to u\bar{u})$.

3.3. Complete structure

Let us finally wrap up by displaying the complete structure of our code shown in the flowchart in Fig. 1. An user specifies a n-parton real emission process as an input to the Mathematica code as well as a set of the external momenta $\{p_n\}$. The Mathematica program then generates all dipole terms in the appropriate order (see Sec. 3.1) along with the CLBS $\mathrm{B1}, \cdots, \mathrm{B4}$. The created dipole terms are written as C- or Fortran routines to a file 'dipole.c' or 'dipole.f'. In this file the explicit expressions for the dipole splitting function V are contained. Together with each dipole, the information on its reduced kinematics as a function of the external momenta is stored (see Sec. 3.1).

The necessary information to calculate each CLBS is transferred to MadGraph through the file 'proc_card_dip.data'. This data file is an imitation of the input file 'proc_card.data' of the original MadGraph. Our interface reads 'proc_card_dip.data' and gets MadGraph to write Fortran routines for the evaluation of each CLBS in a file 'matrix_dip.f'. As we see in Sec. 3.2, the color matrix CF is modified to CF′ due to the additional color insertions.

Finally, the C- and/or Fortran codes in the two files can be used for the numerical evaluation of all dipole terms as functions of the external momenta $\{p_n\}$. The

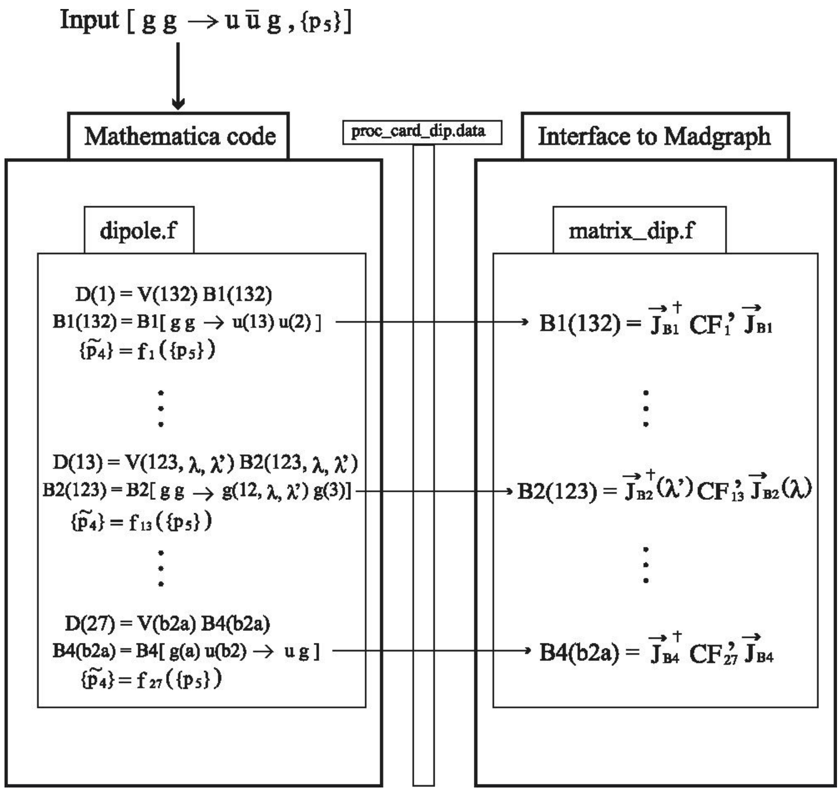

Figure 1.
The complete structure of the code is shown. The Mathematica generates all subtraction terms in the dipole formalism and returns output in C or Fortran for numerical evaluation. The calculation of the CLBS is delegated to MadGraph via an interface.

sum of the dipole subtractions from the invariant matrix element squared of the NLO real emission process reads,

$$|\mathrm{M}(2 \to (n-2)\mathrm{partons})|^2 - \sum_i \mathrm{D}(i)\,. \qquad (17)$$

The invariant matrix element squared can be calculated with the original MadGraph version and i runs over all dipole terms. Eq. (17) is finite upon integration over the phase space of the unresolved parton. Thus, it can finally be integrated over the phase space by using standard Monte Carlo techniques to obtain its contribution to an IR safe cross section.

4. OUTLOOK

We have reported on ongoing work to automate the Catani-Seymour dipole formalism in order to calculate the subtracted invariant matrix element squared Eq. (17) in an automatic way.

The automatization essentially requires three ingredients: the automatic generation of all dipole terms, the calculation of the CLBS, and the evaluation of different helicity amplitudes. The implementation of each of these tasks either in our Mathematica program or in an interface to MadGraph has been completed and the respective routines have undergone sufficient checks. We are now finalizing the user interface and the output format to achieve full automatization. At the same time we are checking our code for various massless real emission processes, like $gg \to u\bar{u}g$, $gg \to u\bar{u}gg$, and $gg \to u\bar{u}ggg$, as well as for massive processes, like $gg \to t\bar{t}g$, $gg \to t\bar{t}gg$, and $gg \to t\bar{t}ggg$ to obtain finite results for $|\mathrm{M}|^2 - \sum \mathrm{D}$ in all soft and collinear limits (and IR safe contributions to cross sections). Once the reliability of our software has been fully established and the code has been optimized for speed, it will be made publicly available.

REFERENCES

1. C. Buttar *et al.*, hep-ph/0604120.
2. Z. Bern *et al.* [NLO Multileg Working Group], arXiv:0803.0494 [hep-ph].
3. F. Bloch and A. Nordsieck, Phys. Rev. 52 (1937) 54.
4. T. Kinoshita, J. Math. Phys. 3 (1962) 650.
5. T.D. Lee and M. Nauenberg, Phys. Rev. 133 (1964) B1549.
6. S. Catani and M.H. Seymour, Nucl. Phys. B485 (1997) 291, hep-ph/9605323.
7. L. Phaf and S. Weinzierl, JHEP 04 (2001) 006, hep-ph/0102207.
8. S. Catani, S. Dittmaier, M.H. Seymour, and Z. Trocsanyi, Nucl. Phys. B627 (2002) 189, hep-ph/0201036.
9. T. Stelzer and W. F. Long, Comput. Phys. Commun. **81** (1994) 357, hep-ph/9401258.
10. F. Maltoni and T. Stelzer, JHEP 02 (2003) 027, hep-ph/0208156.
11. T. Hahn, Comput. Phys. Commun. 140 (2001) 418, hep-ph/0012260.
12. M.L. Mangano et al., JHEP 07 (2003) 001, hep-ph/0206293.
13. F. Krauss, R. Kuhn and G. Soff, JHEP 02 (2002) 044, hep-ph/0109036.
14. E. Boos *et al.* [CompHEP Collaboration], Nucl. Instrum. Meth. A **534** (2004) 250, hep-ph/0403113.
15. A. Kanaki and C.G. Papadopoulos, Comput. Phys. Commun. 132 (2000) 306, hep-ph/0002082.
16. C.G. Papadopoulos, Comput. Phys. Commun. 137 (2001) 247, hep-ph/0007335.
17. T. Gleisberg and F. Krauss, Eur. Phys. J. C53 (2008) 501, arXiv:0709.2881 [hep-ph].
18. M.H. Seymour and C. Tevlin, (2008), arXiv:0803.2231 [hep-ph]
19. S. Weinzierl, SACLAY-SPH-T-98-083.
20. S. Dittmaier, P. Uwer and S. Weinzierl, Phys. Rev. Lett. **98**, 262002 (2007), hep-ph/0703120.
21. H. Murayama, I. Watanabe and K. Hagiwara, KEK-91-11.
22. K. Hagiwara, H. Murayama and I. Watanabe, Nucl. Phys. B **367** (1991) 257.