Improved phase space treatment
of massive multi-particle final states

B.P. Kersevan[1,2,a], E. Richter-Wąs[3,4,b,★]

[1] Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia
[2] Jozef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia
[3] Institute of Physics, Jagellonian University, 30-059 Krakow, ul. Reymonta 4, Poland
[4] Institute of Nuclear Physics PAS, 31-342 Krakow, ul. Radzikowskiego 152, Poland

Abstract. In this paper the revised Kajantie–Byckling approach and improved phase space sampling techniques for the massive multi-particle final states are presented. The application of the developed procedures to the processes representative for LHC physics indicates the possibility of a substantial simplification of multi-particle phase space sampling while retaining a respectable weight variance reduction and unweighing efficiencies in the event generation process.

1 Introduction

With the advent of the LHC era the need for precise predictions (and subsequently detailed simulation) of many QCD and electroweak processes has arisen. One of the necessary components of an accomplished simulation tool is definitely efficient phase space modeling of multi-particle final states. While the light (massless) particles in the final states and the relevant topologies can quite effectively be simulated using general tools (e.g. RAMBO, SARGE, HAAG [3–5]), further development might be needed for phase space sampling where the massive final state particles are present. While at LEP's 200 GeV a massless approximation for the final state particles was adequate, in most of the studied cases at LHC in contrast the massless approximation becomes less suitable because of crossing the top quark production barrier in conjunction with the possibility of multi-quark final states and the shifting centre-of-mass energy in proton–proton collisions. Furthermore, in the LEP era the number of (hard process) particles in the final states of relevance only rarely rose above four and subsequently the number of possible final state topologies was relatively modest. Typical processes of interest (signal) and their backgrounds at LHC will involve several heavy quarks and/or weak bosons in the intermediate states; thus the number of particles in a typical process under study is generally at least four. Furthermore, the number of Feyn-

man diagrams contributing to a typical process involving heavy quarks in the final state is mostly ranging from a few to hundred(s) (and can steeply rise to many thousands when massless quarks are added). Subsequently, this results in varied particle topologies which prove to be a challenge when trying to adequately describe them using the available statistical approaches and numerical methods. Important steps have already been made in matrix element calculations (e.g. MadGraph [6]) and have so far surpassed the corresponding development of phase space modeling and sampling techniques.

The general objective in the simulation of physics processes for the LHC environment is thus to improve the integration of the differential cross-section using Monte Carlo sampling methods[1]. The sampling method used should aim to minimise the variance of the integral as well as maximise the sampling efficiency given a certain number of iterations and the construction of the sampling method itself should aim to be sufficiently general and/or modular to be applicable to a wide range of processes. Let us write down a (process) cross-section integral for LHC type (hadron–hadron) collisions:

$$\sigma = \int \sum_{a,b} f_a(x_1, Q^2) f_b(x_2, Q^2) \frac{|\mathcal{M}_n|^2}{(2\pi)^{3n-4}(2\hat{s})} \, dx_1 \, dx_2 \, d\Phi_n,$$

(1)

where $f_{a,b}(x, Q^2)$ represent the gluon or (anti) quark parton density functions, $|\mathcal{M}_n|^2$ the squared n-particle matrix element divided by the flux factor $[(2\pi)^{3n-4}2\hat{s}]$, and $d\Phi_n$ denotes the n-particle phase space differential. The quantity $\hat{s} = x_1 x_2 s$ is the effective centre-of-mass energy, and

ᵃ e-mail: borut.kersevan@ijs.si
ᵇ e-mail: elzbieta.richter-was@cern.ch

---

[1] For a nice discussion of this topic see e.g. [1,2]

the sum $\sum_{a,b}$ runs in case of quark–antiquark incident partons over all possible quark–antiquark combinations $(a, b = u, d, s, c, \bar{u}, \bar{d}, \bar{s}, \bar{c})$. In case of the gg initial state the sum has only one term with $a = b = g$.

It is often convenient to re-write the differential cross-section in the form

$$\sigma = \int \sum_{a,b} x_1 f_a(x_1, Q^2)\, x_2 f_b(x_2, Q^2) \frac{|\mathcal{M}_n|^2}{(2\pi)^{3n-4}(2\hat{s}^2)}$$

$$\times \mathrm{d}y\, \mathrm{d}\hat{s}\, \mathrm{d}\Phi_n, \tag{2}$$

with the new (rapidity) variable given by $y = 0.5 \ln(x_1/x_2)$. The n-body phase space differential $\mathrm{d}\Phi_n$ and its integral Φ_n depend only on $\hat{s}$ and particle masses m_i due to Lorentz invariance:

$$\Phi_n(\hat{s}, m_1, m_2, \ldots, m_n) \tag{3}$$

$$= \int \mathrm{d}\Phi_n(\hat{s}, m_1, m_2, \ldots, m_n)$$

$$= \int \delta^4 \left((p_a + p_b) - \sum_{i=1}^{n} p_i \right) \prod_{i=1}^{n} \mathrm{d}^4 p_i \delta(p_i^2 - m_i^2)\Theta(p_i^0),$$

with a and b denoting the incident particles and i running over all outgoing particles $i = 1, \ldots, n$. What one would like to do is to split the n-body phase parameterised by $3n-4$ essential (i.e. non-trivial) independent variables into manageable subsets (modules) to be handled by techniques which reduce the variance of the result and/or the sampling efficiency (e.g. importance sampling [7] or adaptive integration like VEGAS [8] or FOAM [9]). Stating this in formal terms, the above equation, (2), should be transformed into an expression like

$$\sigma = \left(\prod_{i=1}^{n} \int_{s_i^-}^{s_i^+} \mathrm{d}s_i \right) \left(\prod_{j=1}^{m} \int_{t_j^-}^{t_j^+} \mathrm{d}t_j \right) \left(\prod_{k=1}^{l} \int_{\Omega_k^-}^{\Omega_k^+} \mathrm{d}\Omega_k \right) |\mathcal{J}_n|$$

$$\times \int \sum_{a,b} x_1 f_a(x_1, Q^2)\, x_2 f_b(x_2, Q^2)$$

$$\times \frac{|\mathcal{M}_n|^2}{(2\pi)^{3n-4}(2\hat{s}^2)}\, \mathrm{d}y\, \mathrm{d}\hat{s}, \tag{4}$$

where one integrates over Mandelstam type (Lorentz invariant) momenta transfers s_i, t_j and space angles $\Omega_k \equiv (\cos\vartheta_k, \phi_k)$ within the kinematically allowed limits ($3n-4$ variables in total) with the term $|\mathcal{J}_n|$ denoting the Jacobian of the transformation. If one would then decide to introduce importance sampling functions in order to reduce the peaking behavior of the integrand [7], the integrals would take the form

$$\int_{s_i^-}^{s_i^+} \mathrm{d}s_i = \int_{s_i^-}^{s_i^+} \frac{g_i(s_i)}{g_i(s_i)} \mathrm{d}s_i, \tag{5}$$

where the importance sampling function g_i is the probability density function normalised in the integration region $[s_i^-, s_i^+]$:

$$\int_{s_i^-}^{s_i^+} g_i(s_i)\mathrm{d}s_i = 1, \tag{6}$$

which exhibits a similar peaking behavior as the integrand. Formally, one then inserts the identity

$$1 = \int_0^1 \delta \left(r_i - \int_{s_i^-}^{s_i} g_i(s_i)\mathrm{d}s_i \right) \mathrm{d}r_i \tag{7}$$

into the integral and then derives the *unitary* sampling prescription:

$$\int_0^1 \mathrm{d}r_i \int_{s_i^-}^{s_i^+} \delta \left(r_i - \int_{s_i^-}^{s_i} g_i(s_i)\mathrm{d}s_i \right) \frac{g_i(s_i)}{g_i(s_i)} \mathrm{d}s_i$$

$$= \int_0^1 \mathrm{d}r_i \int_{s_i^-}^{s_i^+} \delta \left(s_i - G^{-1}(r_i) \right) \frac{1}{g_i(s_i)} \mathrm{d}s_i$$

$$= \int_0^1 \frac{\mathrm{d}r_i}{g_i(G^{-1}(r_i))}, \tag{8}$$

which formally means that the s_i values are sampled from the interval according to the $g_i(s_i)$ distribution by using the (pseudo-) random variable r_i together with the $g_i(s_i)$ cumulant $G(s_i) = \int_{s_i^-}^{s_i} g_i(s_i)\mathrm{d}s_i$ with its inverse G^{-1}. The unitarity of the algorithm states that each trial (r_i value) produces a result (i.e. a corresponding s_i value distributed according to $g_i(s_i)$).

Performing such substitutions on all integration parameters would give as the cross-section expression

$$\sigma = \prod_{i=1}^{3n-4} \int_0^1 \mathrm{d}r_i \frac{f(r_1, r_2 \ldots)}{g(r_1, r_2 \ldots)}, \tag{9}$$

where the integrand would (hopefully) have as low a variation as possible at least for a subset of contributing Feynman diagrams[2]. To improve the sampling method further, the r_i (pseudo-) random variables can be sampled from adaptive algorithms of the VEGAS type [8, 10].

A representative Feynman diagram describing a $2 \to 6$ process is shown in Fig. 1. As one can see, the process can be split in several consecutive branchings; this approximation is often used in matrix element (probability amplitude) calculations. It seems rather obvious that any Feynman diagram can be split in a series of horizontal and vertical

---

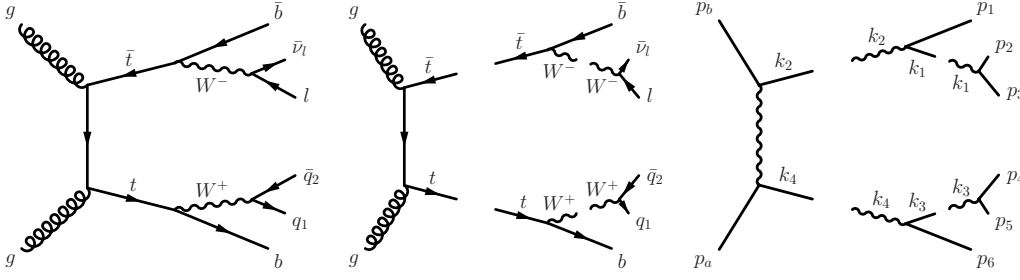[2] The "modularisation" can be performed for several topologies at the same time and multi-channel techniques can be applied.

Fig. 1. A representative Feynman diagram describing a $2 \to 6$ process $gg \to t\bar{t} \to b\bar{b}W^+W^- \to b\bar{b}\ell\bar{\nu}_\ell q_1\bar{q}_2$ and its decomposition into a set of $2 \to 2$ t-channel and s-channel sub-processes

branchings that one can denote as s-type and t-type (u-type) using the analogy with the Mandelstam variables. What one would like to do is thus to modularise the phase space in the form of sequential s- and t-type splits.

The s-splitting of phase space was first examined in the 1950s by several authors [11–14] (and also independently by Raubold and Lynch whose contribution was never published). In itself the s-splitting is relatively easy to implement and has as such been used in many instances of Monte Carlo generation (e.g. FermiSV [15], Excalibur [16], Tauola [17] etc.); the t-type branchings (often tagged as multi-(peri)pheral topologies) have in contrast generally been calculated and used only for specific cases (e.g. for three or four particles in the final state [15, 16, 18]). As it turns out, the problem of several massive particles in the final state has already appeared more than 30 years ago when several hadrons (e.g. pions) have been produced in (comparatively low energy) nuclear interactions. At that time Kajantie and Byckling [19] have derived the formulae for simulating any sequence of s- and specifically t-type branchings which, with some modifications, can also be applied to the EW and QCD processes involving heavy quarks and/or massive bosons at LHC.

In the following sections, Sects. 2 and 3, the revised version of Kajantie–Byckling algorithm (KB) will be discussed and a formulation of the algorithm for the multi-phase space integration will be presented. In Sect. 4 the numerical results for some representative applications as implemented in the AcerMC 2.0 Monte Carlo generator will be presented.

2 Modified Kajantie–Byckling formalism

2.1 The s-type branching algorithms

The s-splits are the simplest method in the KB formalism. For the sake of completeness one should start with the definition of the two-body phase space integral (c.f. Fig. 2):
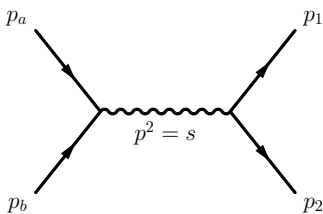
$$\Phi_2(s, m_1, m_2) \tag{10}$$

Fig. 2. A diagram of a generic $2 \to 2$ s-channel process

$$= \int d^4p_1 d^4p_2 \delta(p_1^2 - m_1^2)\delta(p_2^2 - m_2^2)\delta^4(p - p_1 - p_2)$$

$$\times \Theta(p_1^0)\Theta(p_2^0),$$

with the incoming momentum sum $p = (p_a + p_b)$, $p^2 = s$ and outgoing momenta $p_{1,2}$, $p_{1,2}^2 = m_{1,2}^2$. The phase space integral is Lorentz invariant (as one can observe in the above equation where it is written in a manifestly Lorentz invariant form). Subsequently, due to Lorentz invariance, the integral is necessarily a function of the Lorentz scalars s, m_1 and m_2 only. The step function product $\Theta(p_1^0)\Theta(p_2^0)$ is the explicit requirement of the positiveness of the energy terms in $p_{1,2}$ while the delta functions represent the on-shell conditions on $p_{1,2}$ and the total momentum conservation.

The integral can be transformed into a more compact form by integrating out the spurious variables; one thus first integrates over d^4p_2 and chooses the centre-of-mass system (CMS) as the integration system of reference with $p = (\sqrt{s}, 0, 0, 0)$ and then evaluates the integrals over p_1^0 and E_1^*:

$$\Phi_2(s, m_1, m_2)$$

$$= \int d^4p_1 \delta(p_1^2 - m_1^2)\delta\left((p - p_1)^2 - m_2^2\right)\Theta(p_1^0)$$

$$= \int \frac{d^3p_1^*}{2E_1^*}\delta(s + m_1^2 - 2\sqrt{s}E_1^* - m_2^2)$$

$$= \frac{1}{4\sqrt{s}}\int p_*^1 dE_1^* d\Omega_1^* \delta\left(E_1^* - \frac{s + m_1^2 - m_2^2}{2\sqrt{s}}\right)$$

$$= \frac{p_1^*(s, m_1, m_2)}{4\sqrt{s}}\int d\Omega_1^*, \tag{11}$$

with the stars explicitly denoting the values in the centre-of-mass system. The first integration simply sets $p_1^0 = \sqrt{(p_1^*)^2 + m_1^2} = E_1^*$ and the second integral leads to the well known relations for the energy

$$E_1^* = \frac{s + m_1^2 - m_2^2}{2\sqrt{s}}, \quad E_2^* = \sqrt{s} - E_1^* = \frac{s + m_2^2 - m_1^2}{2\sqrt{s}}, \tag{12}$$

and momenta sizes

$$p_1^* = |\mathbf{p}_1^*| = \frac{\sqrt{\lambda(s, m_1^2, m_2^2)}}{2\sqrt{s}}, \quad p_2^* = p_1^* \tag{13}$$

of two-particle production. The $\lambda(s, m_1^2, m_2^2)$ denotes the Lorentz invariant function

$$\lambda(s, m_1^2, m_2^2) = (s - (m_1 + m_2)^2)(s - (m_1 - m_2)^2) \tag{14}$$

and thus explicitly contains the phase space cutoff, i.e. the requirement that the available CMS energy $\sqrt{s}$ should be bigger than the mass sum $\sqrt{s} \geq (m_1 + m_2)$. Note that the integration was so far done only over the spurious parameters, leaving the polar and azimuthal angle of the p_1 particle as the two independent parameters $d\Omega^* = d\cos\theta^* d\varphi^*$. The integral becomes trivial to sample in case the outgoing particles can be approximated as massless (the "boost" factor lambda transforms to unity). As already claimed, the latter approximation is however often unjustified when studying processes representative for the LHC environment.

Kajantie and Byckling [19] introduced the *recursion* and *splitting* relations for the n-particle phase space $\Phi_n(s)$ given by (3). The recursion relation can be derived by defining the momentum sum,

$$k_i = \sum_{j=1}^{i} p_j = (k_i^0, \mathbf{k_i}); \quad M_i^2 = k_i^2. \tag{15}$$

Subsequently one can interpret $p = k_n$ and $s = M_n^2$ from (3). One continues by introducing the identities

$$1 = \int dM_{n-1}^2 \delta(k_{n-1}^2 - M_{n-1}^2)\Theta(k_{n-1}^0) \tag{16}$$

and

$$1 = \int d^4 k_{n-1} \delta^4(p - k_{n-1} - p_n) \tag{17}$$

into the integral of (3); separating out the arguments containing k_{n-1} and p_n terms one obtains

$$\Phi_n(M_n^2, m_1, m_2, \ldots, m_n) \tag{18}$$
$$= \int dM_{n-1}^2 \left\{ \int d^4 k_{n-1} d^4 p_n \right.$$
$$\times \delta(k_{n-1}^2 - M_{n-1}^2)\delta(p_n^2 - m_n^2)\delta^4(p - k_{n-1} - p_n)$$
$$\left. \times \Theta(k_{n-1}^0)\Theta(p_n^0) \right\} \Phi_{n-1}(M_{n-1}^2, m_1, m_2, \ldots, m_{n-1}),$$

where the remaining p_i terms form the $(n-1)$-particle phase space integral $\Phi_{n-1}(M_{n-1}^2, m_1, m_2, \ldots, m_{n-1})$ and the terms in curly brackets give a two-particle phase space term (c.f. (11)):

$$\Phi_n(M_n^2, m_1, m_2, \ldots, m_n)$$
$$= \int dM_{n-1}^2$$
$$\times \Phi_2(M_n^2, M_{n-1}, m_n)\Phi_{n-1}(M_{n-1}^2, m_1, m_2, \ldots, m_{n-1})$$
$$= \int dM_{n-1}^2 \frac{p_n^*}{4M_n} \Phi_{n-1}(M_{n-1}^2, m_1, m_2, \ldots, m_{n-1})$$
$$= \int_{\left(\sum_{i=1}^{n-1} m_i\right)^2}^{(M_n - m_n)^2} dM_{n-1}^2 \frac{\sqrt{\lambda(M_n^2, M_{n-1}^2, m_n^2)}}{8M_n^2}$$
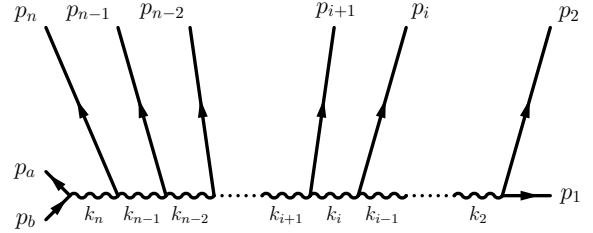$$\times \int d\Omega_n^* \Phi_{n-1}(M_{n-1}^2, m_1, m_2, \ldots, m_{n-1}), \tag{19}$$

Fig. 3. The diagrammatic representation of consecutive s-splits

with the integration limits on M_{n-1}^2 following from its definition in (15). It has to be emphasised that the angles in $d\Omega_i^*$ are each time calculated in the centre-of-mass system of k_i with the invariant mass M_i. The resulting recursion relation is clearly of advantage when describing cascade decays of particles $k_n \to k_{n-1}p_n \to k_{n-2}, p_n, p_{n-1} \to \ldots$; it also proves that the n-particle phase space of (3) can be reduced into a sequence of two-particle phase space terms, as shown in Fig. 3.

It can further prove of advantage to loosen up the splitting terms of (16) and (17) so that instead of summing to $n-1$ one groups an arbitrary set of ℓ particles:

$$1 = \int dM_l^2 \delta(k_l^2 - M_l^2)\Theta(k_l^0), \tag{20}$$

$$1 = \int d^4 k_l \delta^4 \left(p - k_l - \sum_{j=l+1}^{n} p_j \right), \tag{21}$$

which, when repeating the procedure in the recursion relation of (19), results in the expression

$$\Phi_n(M_n^2, m_1, m_2, \ldots, m_n) \tag{22}$$
$$= \int_{\left(\sum_{i=1}^{l} m_i\right)^2}^{(M_{l+1} - m_{l+1})^2} dM_l^2 \Phi_{n-l+1}(M_n^2, M_l, m_{l+1}, \ldots, m_n)$$
$$\times \Phi_l(M_l^2, m_1, m_2, \ldots, m_l),$$

and thus effectively splits the phase space into two subsets, equivalent to introducing an intermediate (virtual) particle with momentum k_l.

The number of splitting relations and the number of particles in each group as given in (21) can be chosen in any possible sequence, thus meaning that the grouping sequence is arbitrary and can be adjusted to fit the topology in question.[3]

At this point some modifications were introduced into the algorithm in order to adapt it to the specifics of the processes expected at the LHC. Kajantie and Byckling namely assumed that the generation sequence would be "down" the cascade (i.e. by sampling first a M_n value, then M_{n-1} value etc. as is indeed most often done in Monte Carlo generators). This might however not be optimal in the LHC

---

[3] Suggestions of [19] on how to pick random number sequences will not be used since one might like to couple this method with an adaptive algorithm to improve the sampling efficiencies.

environment since the available centre-of-mass energy for the hard process ($\hat{s}$) can vary in a wide range of values (c.f. (2)) and has to be sampled from a distribution itself. The shape of the distribution function for $\hat{s}$ is expected to behave as a convolution of the peaking behavior of all participating invariant masses times the parton density functions (c.f. (2)); it subsequently seems to be more natural (and efficient) first to sample the individual propagator peaks and then their subsequent convolutions. Furthermore, by generating the invariant masses "up" the cascade (i.e. first M_2, M_3 ... M_n and finally $\hat{s}$) the kinematic limits on the branchings occur in a more efficient way (bound on the $\sqrt{\lambda}$ values; see (14) and (24)), which is very convenient since in the LHC environment no stringent generation cuts should be made on the inherently non-measurable $\hat{s}$ as it cannot be accounted for by an analogous cut in a physics analysis.

A necessary modification of the algorithm would thus be to reverse the generation steps by starting with the last pair(s) of particles. In terms of integration (i.e. sampling) limits this translates into changing the limits of (19):

$$\Phi_n(M_n^2, m_1, m_2, \ldots, m_n)$$

$$= \int\limits_{\left(\sum_{i=1}^{n-1} m_i\right)^2}^{(M_n-m_n)^2} dM_{n-1}^2 \frac{\sqrt{\lambda(M_n^2, M_{n-1}^2, m_n^2)}}{8M_n^2} \int d\Omega_n^*$$

$$\times \int\limits_{\left(\sum_{i=1}^{n-2} m_i\right)^2}^{(M_{n-1}-m_{n-1})^2} dM_{n-2}^2 \frac{\sqrt{\lambda(M_{n-1}^2, M_{n-2}^2, m_{n-1}^2)}}{8M_{n-1}^2} \int d\Omega_{n-1}^*$$

$$\times \ldots \int\limits_{\left(\sum_{j=1}^{i-1} m_j\right)^2}^{(M_i-m_i)^2} dM_{i-1}^2 \frac{\sqrt{\lambda(M_i^2, M_{i-1}^2, m_i^2)}}{8M_i^2} \int d\Omega_i^* \ldots$$

$$\times \int\limits_{(m_1+m_2)^2}^{(M_3-m_3)^2} dM_2^2 \frac{\sqrt{\lambda(M_3^2, M_2^2, m_3^2)}}{8M_3^2} \int d\Omega_3^*$$

$$\times \frac{\sqrt{\lambda(M_2^2, m_1^2, m_2^2)}}{8M_2^2} \int d\Omega_2^*, \tag{23}$$

which accommodates the mass generation sequence $k_n \to k_{n-1} + p_n \to \ldots$ (i.e. first sample M_{n-1}^2, then M_{n-2}^2, etc.), into

$$\Phi_n(M_n^2, m_1, m_2, \ldots, m_n)$$

$$= \int\limits_{(M_{n-2}+m_{n-1})^2}^{(M_n-m_n)^2} dM_{n-1}^2 \frac{\sqrt{\lambda(M_n^2, M_{n-1}^2, m_n^2)}}{8M_n^2} \int d\Omega_n^*$$

$$\times \int\limits_{(M_{n-3}+m_{n-2})^2}^{(M_n-m_n-m_{n-1})^2} dM_{n-2}^2 \frac{\sqrt{\lambda(M_{n-1}^2, M_{n-2}^2, m_{n-1}^2)}}{8M_{n-1}^2}$$

$$\times \int d\Omega_{n-1}^*$$

$$\times \ldots \int\limits_{(M_{i-1}+m_i)^2}^{\left(M_n-\sum_{j=i+1}^n m_j\right)^2} dM_{i-1}^2 \frac{\sqrt{\lambda(M_i^2, M_{i-1}^2, m_i^2)}}{8M_i^2}$$

$$\times \int d\Omega_i^* \ldots$$

$$\times \int\limits_{(m_1+m_2)^2}^{\left(M_n-\sum_{j=3}^n m_j\right)^2} dM_2^2 \frac{\sqrt{\lambda(M_3^2, M_2^2, m_3^2)}}{8M_3^2} \int d\Omega_3^*$$

$$\times \frac{\sqrt{\lambda(M_2^2, m_1^2, m_2^2)}}{8M_2^2} \int d\Omega_2^*, \tag{24}$$

where one first samples the mass $M_2, M_3, \ldots, M_{n-1}$ in the appropriate limits.

In some topologies symmetric cases of mass generation can appear (as shown in Fig. 1) where the integration sequence is ambivalent (e.g. in Fig. 1 the ambivalence is which top quark invariant mass to generate first) and after a choice is made (since one of the two cases in the symmetric pair has to take precedence) the procedure itself remains not entirely symmetric. Detailed studies have shown that it proves useful to include all permutations of such ambiguous sequences into the MC algorithm in order to "symmetrise" the solution and thus make it easier to process by further additions (e.g. adaptive algorithms).

2.2 The t-type branching algorithms

The t-splits are a specialty of the KB formalism due to the advanced calculation of the limits on the (massive) t-variable. The formalism can be introduced by observing that in case of $p_a + p_b \to p_1 + p_2$ scattering the momentum transfer is characterised by the (Mandelstam) variable $t = (p_1 - p_a)^2$ (c.f. Fig. 4). It is thus sensible to replace the $d\Omega_1^* = d\cos\theta_1^* d\varphi_1^*$ integration in the two-body phase space integral of (11) with integration over the t-variable. Writing the definition of t in the centre-of-mass system one gets

$$t = q^2 = (p_a - p_1)^2$$
$$= m_a^2 + m_1^2 - 2E_a^* E_1^* + 2p_a^* p_1^* \cos\theta_1^*, \tag{25}$$

and hence

$$dt = 2p_a^* p_1^* d\cos\theta^* \tag{26}$$

Using the latter substitution together with (12) and (13) and the analogue for p_a,

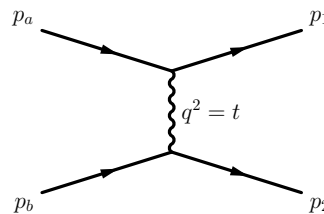$$p_a^* = \frac{\sqrt{\lambda(s, m_a^2, m_b^2)}}{2\sqrt{s}}, \tag{27}$$

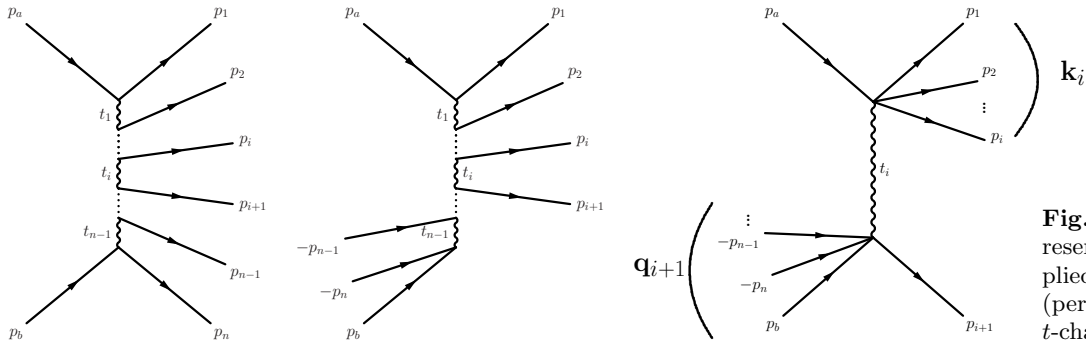**Fig. 4.** A diagram of a generic $2 \to 2$ t-channel process

Fig. 5. The diagrammatic representation of the method applied in translating the multi-(peri)pheral splits into a $2 \to 2$ t-channel configuration

one obtains in place of (11)

$$\Phi_2(s, m_1, m_2) = \frac{p_1^*(s, m_1, m_2)}{4\sqrt{s}} \int \mathrm{d}\Omega_1^*$$

$$= \frac{1}{8\sqrt{s}p_a^*} \int \mathrm{d}t \, \mathrm{d}\varphi^*$$

$$= \frac{1}{4\sqrt{\lambda(s, m_a^2, m_b^2)}} \int\limits_{t^-}^{t^+} \mathrm{d}t \int\limits_{0}^{2\pi} \mathrm{d}\varphi^*. \quad (28)$$

With the integration variable change the integration domain changes from $[-1, 1]$ for $\mathrm{d}\cos\theta^*$ to $[t^-, t^+]$ for the $\mathrm{d}t$ integration. The $t^\pm$ limits are obtained by inserting the $\cos\theta^*$ limits into (25):

$$t^\pm = m_a^2 + m_1^2 - 2E_a^* E_1^* \mp 2p_a^* p_1^*, \quad (29)$$

or in the Lorentz invariant form (c.f. (12) and (13)):

$$t^\pm = m_a^2 + m_1^2 - \frac{(s + m_a^2 - m_b^2)(s + m_1^2 - m_2^2)}{2s}$$

$$\mp \frac{\sqrt{\lambda(s, m_a^2, m_b^2)\lambda(s, m_1^2, m_2^2)}}{2s}. \quad (30)$$

As a step towards generalisation one has to note that the kinematic limits $t^\pm$ can also be derived from *the basic four-particle kinematic function $G(x, y, z, u, v, w)$* [19,20], where the function G can be expressed as a Cayley determinant:

$$G(x, y, z, u, v, w) = -\frac{1}{2} \begin{vmatrix} \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & v & x & z \\ 1 & v & 0 & u & y \\ 1 & x & u & 0 & w \\ 1 & z & y & w & 0 \end{pmatrix} \end{vmatrix}. \quad (31)$$

The kinematic limits on t are in this case given by the condition

$$G(s, t, m_2^2, m_a^2, m_b^2, m_1^2) \leq 0, \quad (32)$$

and it should be noted that the above condition gives either $t^\pm$ limits given a fixed value of s or equivalently $s^\pm$ limits given a fixed t value.

In search of a recursion relation involving t-variables one can note that in (19) the angle in $\cos\theta_n^*$ is equivalent

to the scattering angle in the centre-of-mass system of the reaction $p_a + p_b \to k_{n-1} + p_n$ and thus given by

$$t_{n-1} = (p_a - k_{n-1})^2 \quad (33)$$

$$= m_a^2 + M_{n-1}^2 - 2E_a^* k_{n-1}^{0*} + 2p_a^* k_{n-1}^* \cos\theta_{n-1}^*,$$

with the $t_{n-1}^\pm$ limits expressed by

$$G(M_n^2, t_{n-1}, m_n^2, m_a^2, m_b^2, M_{n-1}^2) \leq 0, \quad (34)$$

and the p_a^* given by (27). In order to produce a more general picture it can further be deduced that the next angle in the recursion θ_{n-1}^*, is the scattering angle of the subsequent process $p_a + (p_b - p_n) \to k_{n-2} + p_{n-1}$ in the centre-of-mass system of k_{n-1}; the $(p_b - p_n) = q_{n-1}$ is in this case considered as a virtual incoming particle with momentum q_{n-1} (c.f. Fig. 5).

It immediately follows that for a general process $p_a + q_{i+1} \to k_i + p_{i+1}$ with

$$q_i = p_b - \sum_{j=i+1}^{n} p_j = p_a - k_i; \quad q_i^2 = t_i; \quad q_n^2 = t_n = m_b^2,$$
$$\quad (35)$$

a general expression for t_i becomes in the centre-of-mass frame of k_{i+1}

$$t_i = (p_a - k_i)^2 \quad (36)$$

$$= m_a^2 + M_i^2 - 2E_a^{*(i+1)} k_i^{0*(i+1)} + 2p_a^{*(i+1)} k_i^{*(i+1)} \cos\theta_i^*,$$

where the momenta in centre-of-mass frame of k_{i+1}, denoted with the superscript $*(i+1)$, are given by

$$k_i^{*(i+1)} = \frac{\sqrt{\lambda(M_{i+1}^2, M_i^2, m_{i+1}^2)}}{2M_{i+1}}, \quad (37)$$

$$p_a^{*(i+1)} = \frac{\sqrt{\lambda(M_{i+1}^2, m_a^2, t_{i+1})}}{2M_{i+1}}, \quad (38)$$

and the corresponding energies $k_i^{0*(i+1)}$ and $E_a^{*(i+1)}$ can simply be obtained by using the analogues of (12) and (13) or the usual Einstein mass–energy relations directly. The corresponding $t_i^\pm$ limits are given by

$$G(M_{i+1}^2, t_i, m_{i+1}^2, m_a^2, t_{i+1}, M_i^2) \leq 0, \quad (39)$$

and the recursion relation of (19) becomes

$$\Phi_n(M_n^2, m_1, m_2, \ldots, m_n) \qquad (40)$$

$$= \int\limits_{\left(\sum_{i=1}^{n-1} m_i\right)^2}^{(M_n - m_n)^2} \frac{dM_{n-1}^2}{4\sqrt{\lambda(M_n^2, m_a^2, t_n)}}$$

$$\times \int\limits_0^{2\pi} d\varphi_n^* \int\limits_{t_{n-1}^-}^{t_{n-1}^+} dt_{n-1} \, \Phi_{n-1}(M_{n-1}^2, m_1, m_2, \ldots, m_{n-1}),$$

As already argued the resulting set of $(s_i = M_i^2, t_i)$ can again be sampled in any direction with respect to the cascade by applying the appropriate change in the integration limits (c.f. (19) and 24). The recommended approach (i.e. the introduced modification of the algorithm) is again to first sample the invariant masses in the reverse cascade direction (i.e. in the sequence $M_2, M_3, \ldots, M_n$) and then the t_i values within the limits calculated from (34) down the cascade (i.e. in the order of $t_{n-1}, t_{n-2}, \ldots, t_1$).

To sum up, it has been shown that using the Kajantie–Byckling formalism the phase space for any topology can be split in a set of s-type and t-type $2 \rightarrow 2$ branching steps (modules) given by the recursive formulae (24) and (40).

3 Propagator sampling

A well known theoretical issue is that one can expect the most prominent peaks in the differential cross-section of a specific process in the phase space regions of high propagator values in the corresponding probability density. Consequently, in the scope of complementing the modular structure of the derived Kajantie–Byckling based phase space sampling, new approaches were also developed concerning the numerical sampling methods of the relevant kinematic quantities.

In order to get a small variance in the Monte Carlo procedure one would thus like to include the appropriate peaking dependence of the relevant momentum transfers q^2 in the importance sampling function. It however turns out that since the momenta transfers q participate also in the propagator numerators (typically in $p_\mu q^\mu / q^2$) and since in the process of interest one mostly finds several Feynman diagrams contributing to the final probability density, thus causing interferences, it is very difficult or even impossible to estimate the exact power of the momenta transfers in the sampling functions for different propagator peaks. In other words, the probability density dependence on the momentum transfer q^2 can in general be approximated with the dependence $1/(q^2)^\nu$ where the best value of ν must be determined separately (on a process by process basis).

In view of the latter, general formulae have been developed for sampling the $x^{-\nu}$ shape [15,21]: Given a pseudo-random number $r \in [0, 1]$ and limits $x \in [x_-, x_+]$ the value x distributed as $x^{-\nu}$ is obtained from the formulae in (8) as

$$x = \left[x_-^{-\nu+1} \cdot (1 - r) + x_-^{-\nu+1} \cdot r \right]^{-\frac{1}{\nu+1}}; \quad \nu \neq 1; \ (41)$$

$$x = \frac{x_+^r}{x_-^{r-1}}; \quad \nu = 1. \qquad (42)$$

Using the analogous (unitary) approach a recipe for resonant (Breit–Wigner) propagator contributions of the type

$$\mathrm{BW}(s) = \frac{1}{(s - M^2)^2 + M^2 \Gamma^2}, \qquad (43)$$

with $s \in [s_-, s_+]$ and a pseudo-random number $r \in [0, 1]$, is available by the prescription

$$s = M^2 + M\Gamma \cdot \tan\left[(u_+ - u_-) \cdot r + u_-\right], \quad (44)$$

$$u^\pm = \arctan\left(\frac{s^\pm - M^2}{M\Gamma}\right). \qquad (45)$$

Following similar arguments as for the non-resonant propagators one can surmise that the best sampling function for resonant propagators could in general be a Breit–Wigner shape modified by a factor $s^\nu, \nu \in [0, 1]$. In [10] it was found that a shape

$$\mathrm{BW}(s) = \frac{s}{(s - M^2)^2 + M^2 \Gamma^2} \qquad (46)$$

works quite well for a set of processes, and a corresponding sampling recipe was developed. In addition, studies in [22] show that a resonant $\sqrt{s} \times$ Breit–Wigner shape,

$$\mathrm{BW}(s) = \frac{\sqrt{s}}{(s - M^2)^2 + M^2 \Gamma^2}, \qquad (47)$$

should be expected in a range of decay processes. Detailed studies have shown that it is in general better to introduce a $s^\nu, \nu \in [0, 1]$ dependence even if it over-compensates the high mass tails of the corresponding differential cross-section distribution since this provides an overall reduction of the maximal weight fluctuations in the Monte Carlo event generation procedure.

3.1 The inclusion of mass effects in propagator sampling

Studies have shown that the $t^{-\nu} \rightarrow x^{-\nu}$ approximation works quite well for t-channel type propagators since the phase space suppression factor $\sqrt{\lambda}$ participates in the denominator, as shown in (40), and thus contributes only to the $x^{-\nu}$ slope; the sampling of massive propagators can still be achieved by simple variable substitution $(t - m^2)^{-\nu} \rightarrow x^{-\nu}$. On the contrary, while the $x^{-\nu}$ approximation still works reasonably well when describing the s-channel type propagators involving particles with high virtuality and/or decay products with low masses, it can be shown that this is not necessarily the case in the LHC environment, where the presence of massive decay products can significantly affect the invariant mass distributions. As it can be seen in Fig. 6 the shape of the propagator dependence can be strongly suppressed by the phase space $\sqrt{\lambda}$ (boost) factor at low values; thus the sampling function approximation
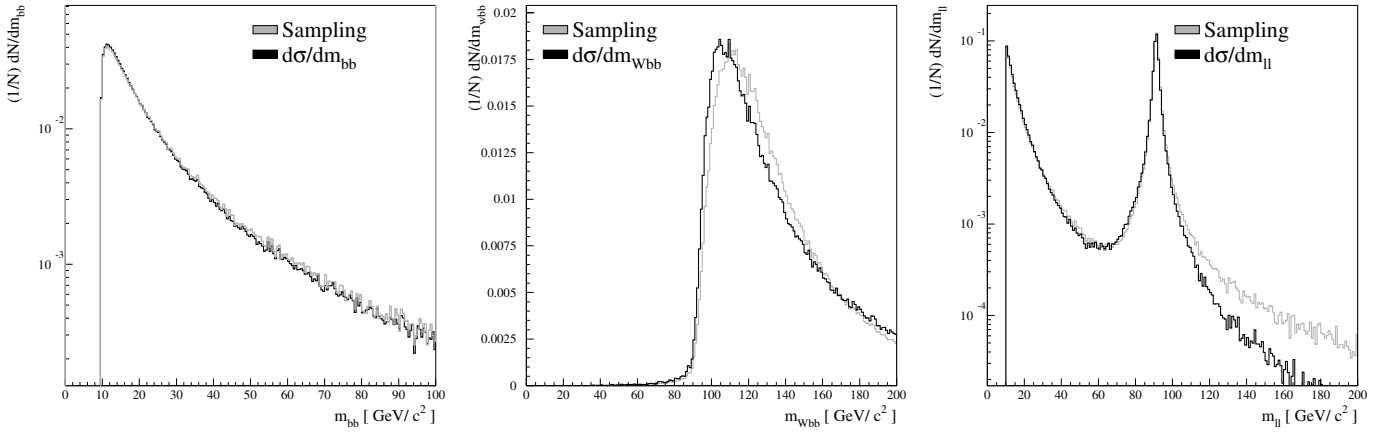
Fig. 6. A few representative invariant mass distribution comparisons between the (normalised) sampling functions and the normalised differential cross-section as obtained with the AcerMC 2.0 Monte Carlo generator. Left: The invariant mass of the $b\bar{b}$ pair in the process $u\bar{d} \to W^+ g^* \to l^+ \nu_l b\bar{b}$. Center: The invariant mass of the $Wb\bar{b}$ system (equivalently the hard centre-of-mass energy $\sqrt{\hat{s}}$) for the same process. Right: The invariant mass of the $\ell\bar{\ell}$ pair in the process $gg \to Z^0/\gamma^* b\bar{b} \to \ell\bar{\ell}b\bar{b}$. All the distributions were obtained using the prescriptions of this paper without the adaptive algorithms also used in the AcerMC 2.0 Monte Carlo generator. As one can see the approximations used seem to work quite well

for non-resonant propagators could be approximated with something like

$$f_{\text{NR}}(s) = \frac{\sqrt{\lambda(s, m_a^2, m_b^2)}}{s} \cdot \frac{1}{s^\nu} = \frac{\sqrt{\lambda(s, m_a^2, m_b^2)}}{s^{\nu+1}} \quad (48)$$

and similarly

$$f_{\text{R}}(s) = \frac{\sqrt{\lambda(s, m_a^2, m_b^2)}}{s} \cdot \frac{\sqrt{s}}{(s - M^2)^2 + M^2\Gamma^2}$$

$$= \frac{\sqrt{\lambda(s, m_a^2, m_b^2)}}{\sqrt{s} \cdot ((s - M^2)^2 + M^2\Gamma^2)} \quad (49)$$

for resonant propagators.

It turns out that the two functions cannot be sampled by the well known unitary algorithms (i.e. the biggest collection of recipes [21] yielded no results); already the integral values of the functions yield complicated expressions which cannot be easily calculated, let alone inverted analytically. The solution was to code numerical algorithms to calculate the integrals (i.e. cumulants) explicitly.

After the integrals are calculated, their inverse and the subsequent sampling value can again be obtained numerically. Namely, resorting to the original definition of the unitarity sampling recipe in (8), by replacing the normalised $g_i(x)$ with

$$g_i(x) \to \frac{f(x)}{\int\limits_{x_-}^{x_+} f(x) \, dx}, \quad (50)$$

which in turn gives

$$\int\limits_{x_-}^{x} f(x) \, dx = r \cdot \int\limits_{x_-}^{x_+} f(x) \, dx, \quad (51)$$

where $f(x)$ is the non-negative function one wants to sample from, $[x_-, x_+]$ is the range of values of the parameter x

we want to sample and r a pseudo-random number $r \in [0, 1]$. As already stated (c.f. (8)), in the case when the integral of the function $f(x)$ is an analytic function, and $F(x) = \int_{x_-}^{x} f(x) \, dx$ and has a known inverse $F^{-1}(x)$, one can construct explicit unitary prescriptions by

$$x = F^{-1}\left(r \cdot [F(x_+) - F(x_-)] + F(x_-)\right), \quad (52)$$

as given for two particular cases in (42) and (45).

In the cases the integral cannot be inverted, the prescription of (51) can directly be transformed into a zero-finding request; thus, since both the integral and the first derivative (i.e. the sampling function and its cumulant) are known, the Newton–Rhapson method is chosen as the optimal one for root finding:

$$g(x) = \left\{ \int\limits_{x_-}^{x} f(x) dx - r \cdot \int\limits_{x_-}^{x_+} f(x) dx \right\} = 0, \quad (53)$$

$$g'(x) = \frac{d}{dx} \left\{ \int\limits_{x_-}^{x} f(x) dx - r \cdot \int\limits_{x_-}^{x_+} f(x) dx \right\}$$

$$= f(x). \quad (54)$$

With a sensible choice of starting points the procedure generally takes on the order of ten cycles until finding the root with adequate numerical precision. The overall generation speed is still deemed quite reasonable.

The integration of the phase space suppressed resonant propagator of (49) yields a rather non-trivial expression:

$$\int\limits_{(m_a+m_b)^2}^{s} f_{\text{R}}(s) \, ds$$

$$= \int\limits_{(m_a+m_b)^2}^{s} \frac{\sqrt{\lambda(s, m_a^2, m_b^2)} \, ds}{\sqrt{s} \cdot ((s - M^2)^2 + M^2\Gamma^2)}$$

$$= \int\limits_{a}^{s} \frac{\sqrt{(s-a)(s-b)}\ \mathrm{d}s}{\sqrt{s}\cdot((s-M^2)^2+M^2\Gamma^2)}$$

$$= \frac{1}{\sqrt{-b}\,\Gamma\,M^2} \times \frac{-2\,\mathrm{i}a\,b\,\Gamma}{(\Gamma^2+M^2)}$$

$$\times\left\{ \mathbf{F}\left[\mathrm{i}\,\mathrm{arcsinh}\left(\frac{\sqrt{-b}}{\sqrt{a}}\right),\frac{a}{b}\right]\right.$$

$$-\mathbf{F}\left[\mathrm{i}\,\mathrm{arcsinh}\left(\frac{\sqrt{-b}}{\sqrt{s}}\right),\frac{a}{b}\right]$$

$$+(\mathrm{i}\Gamma+M)\,(a+\mathrm{i}(\Gamma+\mathrm{i}M)\,M)\,(b+\mathrm{i}(\Gamma+\mathrm{i}M)\,M)$$

$$\times\boldsymbol{\Pi}\left[\frac{M\,(-\mathrm{i}\Gamma+M)}{b},\mathrm{i}\,\mathrm{arcsinh}\left(\frac{\sqrt{-b}}{\sqrt{a}}\right),\frac{a}{b}\right]$$

$$+(\Gamma+\mathrm{i}M)\,(b+(-\mathrm{i}\Gamma-M)\,M)\,(\mathrm{i}a+(\Gamma-\mathrm{i}M)\,M)$$

$$\times\boldsymbol{\Pi}\left[\frac{M\,(\mathrm{i}\Gamma+M)}{b},\mathrm{i}\,\mathrm{arcsinh}\left(\frac{\sqrt{-b}}{\sqrt{a}}\right),\frac{a}{b}\right]$$

$$-(\mathrm{i}\Gamma+M)\,(a+\mathrm{i}(\Gamma+\mathrm{i}M)\,M)\,(b+\mathrm{i}(\Gamma+\mathrm{i}M)\,M)$$

$$\times\boldsymbol{\Pi}\left[\frac{M\,(-\mathrm{i}\Gamma+M)}{b},\mathrm{i}\,\mathrm{arcsinh}\left(\frac{\sqrt{-b}}{\sqrt{s}}\right),\frac{a}{b}\right]$$

$$-(\Gamma+\mathrm{i}M)\,(b+(-\mathrm{i}\Gamma-M)\,M)\,(\mathrm{i}a+(\Gamma-\mathrm{i}M)\,M)$$

$$\left.\times\boldsymbol{\Pi}\left[\frac{M\,(\mathrm{i}\Gamma+M)}{b},\mathrm{i}\,\mathrm{arcsinh}\left(\frac{\sqrt{-b}}{\sqrt{s}}\right),\frac{a}{b}\right]\right\}, \quad (55)$$

where the variables a,b stand for $a=(m_a+m_b)^2$ and $b=(m_a-m_b)^2$ and the functions $\mathbf{F}[\varphi,k]$ and $\boldsymbol{\Pi}[\varphi,k,n]$ are the Legendre incomplete elliptic integrals of the second and third kind with complex arguments. In order to perform the calculations the latter functions had to be coded from scratch since they were not found in any (publicly available) computer libraries or code repositories. The prescriptions for calculating them were found in [23]; the results were checked against the values given by Mathematica$^{\mathrm{TM}}$. In the special case $m_a=m_b$ the above expression simplifies to

$$\int\limits_{(2m_a)^2}^{s} f_{\mathrm{R}}(s)\ \mathrm{d}s$$

$$= \int\limits_{(2m_a)^2}^{s} \frac{\sqrt{\lambda(s,m_a^2,m_a^2)}\ \mathrm{d}s}{\sqrt{s}\cdot((s-M^2)^2+M^2\Gamma^2)}$$

$$= \int\limits_{a}^{s} \frac{\sqrt{(s-a)}\ \mathrm{d}s}{\cdot((s-M^2)^2+M^2\Gamma^2)}$$

$$= \frac{1}{\Gamma\,M\,\sqrt{a+(-\mathrm{i}\Gamma-M)\,M}}$$

$$\times\left\{(\mathrm{i}a+(\Gamma-\mathrm{i}M)\,M)\right.$$

$$\times\,\mathrm{arctan}\left(\frac{\sqrt{-a+z}}{\sqrt{a+(-\mathrm{i}\Gamma-M)\,M}}\right)$$

$$-\mathrm{i}\sqrt{a+(-\mathrm{i}\Gamma-M)\,M}\,\sqrt{a+\mathrm{i}\,(\Gamma+\mathrm{i}M)\,M}$$

$$\left.\times\,\mathrm{arctan}\left(\frac{\sqrt{-a+z}}{\sqrt{a+\mathrm{i}\,(\Gamma+\mathrm{i}M)\,M}}\right)\right\}. \quad (56)$$

The result of integrating the phase space suppressed non-resonant propagator (48) yields a similarly non-trivial result:

$$\int\limits_{(m_a+m_b)^2}^{s} f_{\mathrm{NR}}(s)\ \mathrm{d}s$$

$$= \int\limits_{(m_a+m_b)^2}^{s} \frac{\sqrt{\lambda(s,m_a^2,m_b^2)}\ \mathrm{d}s}{s^{\nu+1}}$$

$$= \frac{1}{2\,\sqrt{1-\frac{s}{a}}\,\nu}$$

$$\times\left\{-2\,\sqrt{(a-s)\,(b-s)}\right.$$

$$\times\frac{\mathbf{F_1}\left[-\nu,-\left(\frac{1}{2}\right),-\left(\frac{1}{2}\right),1-\nu,\frac{s}{a},\frac{s}{b}\right]}{s^\nu\,\sqrt{1-\frac{s}{b}}}$$

$$+\sqrt{\pi}\,\sqrt{(-a+b)\,(a-s)}$$

$$\left.\times\frac{\boldsymbol{\Gamma}\left[1-\nu\right]\,\mathbf{F}\left[-\nu,-\left(\frac{1}{2}\right),\frac{3}{2}-\nu,\frac{a}{b}\right]}{a^\nu\,\sqrt{1-\frac{a}{b}}\,\boldsymbol{\Gamma}\left[\frac{3}{2}-\nu\right]}\right\}, \quad (57)$$

where the function $F[\alpha,\beta,\gamma,x]$ is the Gauss hypergeometric function and $F_1[\alpha,\beta,\beta',\gamma,x,y]$ is the two-parameter (Appell) hypergeometric function [24]. Both functions can be calculated by using the prescriptions in [24]; it however turns out that the calculation of the $F_1[\alpha,\beta,\beta',\gamma,x,y]$ to a certain (high) precision is almost two times slower than the explicit numerical calculation of the integral to the same precision. Subsequently the numerical evaluation of the Gauss hypergeometric function $F[\alpha,\beta,\gamma,x]$ was retained since it participates in the $m_a=m_b$ simplification and the calculation of the integral was done by using a 50-point Gauss–Legendre quadrature with $\sqrt{s}$ weight function; the weights were calculated by [25]. The implementation of the (Appell) hypergeometric function calculation was used as a cross-check to confirm the correct implementation and precision of the numerical method.

As already mentioned, the above integral again simplifies for $m_a=m_b$:

$$\int\limits_{(2m_a)^2}^{s} f_{\mathrm{NR}}(s)\ \mathrm{d}s = \int\limits_{(2m_a)^2}^{s} \frac{\sqrt{\lambda(s,m_a^2,m_b^2)}\ \mathrm{d}s}{s^{\nu+1}}$$

$$= \int\limits_{a}^{s} \frac{\sqrt{(s-a)}\ \mathrm{d}s}{s^{\nu+\frac{1}{2}}} \quad (58)$$

$$= \frac{2}{3}\,a^{1-\nu}\,s^{\frac{3}{2}}\,\mathbf{F}\left[\frac{3}{2},\nu+\frac{1}{2},\frac{5}{2},-s\right],$$

and the Gauss hypergeometric function $F[\alpha, \beta, \gamma, x]$ is in this case calculated by the methods described in [24] with some improvements analogous to the ones described e.g. in [26].

4 Application of the method

The phase space "modularisation" described in this paper has successfully been applied for phase space generation in the AcerMC 2.0 Monte Carlo generator [27]. The program uses the multi-channel phase space generation where each channel corresponds to an expected phase space topology as derived from the participating Feynman diagrams. In AcerMC 2.0 this information was obtained from the modified MadGraph [6] program which also supplied the probability amplitudes for the implemented processes. Each channel topology was in turn constructed from the t-type and s-type modules and sampling functions described in this paper together with some additional importance sampling techniques for space angles and rapidity distributions described in detail elsewhere [10,15,16,28]. The unknown slope parameters (denoted ν in the text, c.f. (48)) of the invariant mass sampling functions for non-resonant propagators were obtained by short training runs of the program on a process by process basis.

As a further step the multi-channel self-optimisation procedure was implemented in order to minimise the variance of the event weights further [7]. A few representative invariant mass distribution comparisons between the implemented sampling functions and the actual differential distributions are shown in Fig. 6.

As a further estimate of the success of the methods a comparison of the variance in the differential cross-section determination are presented in Table 1 for AcerMC 2.0 [27] and an earlier version, AcerMC 1.4 [10], which uses the standard phase space sampling techniques [10,15,16,29,30]. The comparison is done for a few representative processes. It has to be stressed that in the AcerMC 1.x versions the phase space was constructed by hand on a process by process basis meticulously tuned while in the new AcerMC 2.0 the "automated" approach sufficed. A further improvement was the inclusion of the ac-VEGAS algorithm [10] which reduces the maximal event weights in order to improve the unweighing efficiencies; the comparison between the unweighing efficiencies reached in the standard and new approaches (AcerMC versions 1.4 and 2.0) is also given in Table 1. The event weight variance V_σ presented in Table 1

is for N measurements of the weights w_i defined as

$$V_\sigma = \frac{\frac{\sum_{i=1}^{N} w_i^2}{N} - \left(\frac{\sum_{i=1}^{N} w_i}{N}\right)^2}{N - 1}. \qquad (59)$$

Likewise, the customary definition is used for the unweighing efficiency estimate:

$$\epsilon = \frac{\frac{\sum_{i=1}^{N} w_i}{N}}{w_{\max}}, \qquad (60)$$

where $w_{\max}$ is the maximal event weight obtained in N trials. It also might be relevant to stress that since the average weight is the total cross-section estimate for the considered process $\sigma \simeq \sum_{i=1}^{N} w_i/N$ the only quantity that is allowed to change in order to improve the unweighing efficiency is the maximal weight $w_{\max}$.

The improved and automated phase space handling provided the means to include the $2 \rightarrow 6$ processes like e.g. $gg \rightarrow t\bar{t} \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell\bar{\nu}_\ell q_1\bar{q}_2$ (c.f. Fig. 1) which would with the very complicated phase space topologies prove to be too much work to be handled manually. The studies show that the overall unweighing efficiency which can be reached in the $2 \rightarrow 6$ processes by using the recommended phase space structuring is on the order of 10 percent. As an example, process cross-sections and variances with their uncertainties and unweighing efficiencies as obtained for two sample $2 \rightarrow 6$ processes implemented in the AcerMC 2.0 Monte Carlo generator are presented in Table 2 and the corresponding weight distributions are shown in Fig. 7. The two processes share the same initial and final state but include different subsets of Feynman diagrams; the $gg \rightarrow t\bar{t} \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ process includes only the three diagrams containing the $t\bar{t}$ intermediate state, while in the second process $gg \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ all 31 diagrams involving the $b\bar{b}W^+W^-$ intermediate state are included. Subsequently, two sampling channels were constructed to describe the topologies in the first and 13 sampling channels in the second process of the two described above. The results show that the sampling procedure scales quite efficiently with the increase in the number of Feynman diagrams and topologies. Detailed studies show that with inclusion of the new propagator modeling described in this paper the sampling of angular distributions becomes a new restraining factor in achieving optimal weight variances and unweighing efficiencies which might therefore be worth investigating further.

Table 1. The process cross-section variances with their uncertainties and unweighing efficiencies obtained for a few implemented processes basing on the old and new phase space sampling techniques as implemented in AcerMC 1.4 and AcerMC 2.0 respectively. The variances are given for a sequence of 10^5 weighted events (i.e. algorithm iterations) obtained by using the procedure described in the text. The unweighing efficiencies were estimated from samples containing $\sim 10^6$ weighted events

Process	AcerMC 2.0 V_σ [pb^2]	AcerMC 1.4 V_σ [pb^2]	AcerMC 2.0 ϵ	AcerMC 1.4 ϵ
$gg \rightarrow Z/(\rightarrow \ell\ell)b\bar{b}$	$0.129 \cdot 10^{-2} \pm 0.52 \cdot 10^{-5}$	$0.159 \cdot 10^{-2} \pm 0.61 \cdot 10^{-5}$	37%	33%
$q\bar{q} \rightarrow W(\rightarrow \ell\nu)b\bar{b}$	$0.390 \cdot 10^{-2} \pm 0.15 \cdot 10^{-4}$	$0.533 \cdot 10^{-2} \pm 0.18 \cdot 10^{-4}$	35%	33%
$gg \rightarrow t\bar{t}b\bar{b}$	$0.522 \cdot 10^{-4} \pm 0.19 \cdot 10^{-6}$	$0.972 \cdot 10^{-4} \pm 0.44 \cdot 10^{-6}$	36%	20%

Table 2. The process cross-sections and variances with their uncertainties and unweighing efficiencies as obtained for two sample $2 \rightarrow 6$ processes implemented in AcerMC 2.0 Monte Carlo generator. The results show that the sampling procedure scales quite efficiently with the increase in the number of Feynman diagrams and sampling channels. Detailed studies show that with inclusion of the new propagator sampling the angular distributions become the restraining factor. The cross-sections and variances are given for a sequence of $2 \cdot 10^5$ weighted events (i.e. algorithm iterations) obtained by using the procedure described in the text. The unweighing efficiencies were estimated from samples containing $\sim 10^6$ weighted events

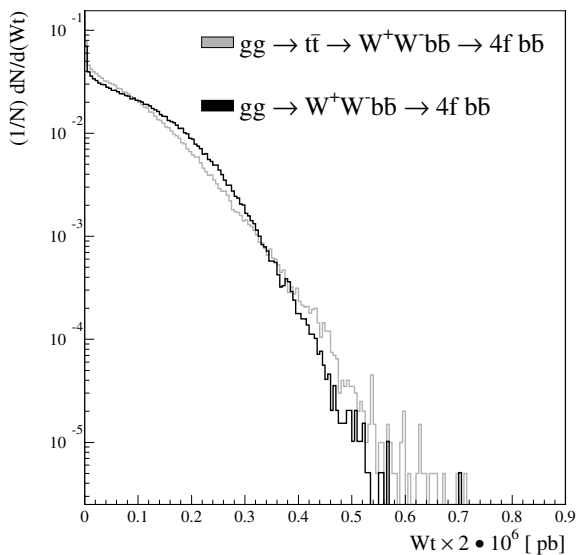| AcerMC 2.0 Process	σ [pb]	V_σ [pb^2]	ϵ
$gg \rightarrow t\bar{t} \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ (3 Feyn. diag./2 sampl. chan.)	4.49	$0.80 \cdot 10^{-4} \pm 0.39 \cdot 10^{-6}$	14%
$gg \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ (31 Feyn. diag./13 sampl. chan.)	4.77	$0.77 \cdot 10^{-4} \pm 0.29 \cdot 10^{-5}$	17%

Fig. 7. The weight distribution of the sampled events for the $gg \rightarrow t\bar{t} \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ (light gray histogram) and $gg \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ (black histogram) processes as obtained with the AcerMC 2.0 Monte Carlo generator. One can observe the well defined weight range for the two processes; as it turns out the weight distribution is even marginally better for the (more complex) second process, possibly because the higher number of sampling channels manage to cover the event topologies in phase space to a better extent

5 Conclusion

In this paper the revised Kajantie–Byckling approach and some improved phase space sampling techniques for the massive multi-particle final states were presented. In order to adapt the procedure to the LHC environment the modifications necessary for reversing the sampling order were introduced and new invariant mass sampling methods, which attempt to describe the propagator dependence of the probability density together with the phase space suppression due to the presence of massive particles, were developed.

The developed procedures have been implemented in the AcerMC 2.0 Monte Carlo generator [27]. Based on the

encouraging evidence provided by the AcerMC 2.0 implementation of the approach it seems reasonable to argue that the methods presented in this paper should substantially simplify and automate the phase space integration (sampling) techniques while retaining a respectable weight variance reduction and unweighing efficiencies provided by the most advanced phase space sampling techniques developed so far [10, 15, 16, 29, 30]. Furthermore, the authors believe that these techniques could easily be combined with algorithms of the type Sarge [4] or HAAG [5] to provide successful sampling of the ($\sim$ massless) final state particles as e.g. final state gluon radiation.

References

1. S. Jadach, physics/9906056 (1999)
2. M. Skrzypek, Z. Was, Comput. Phys. Commun. **125**, 8 (2000)
3. W.J. Stirling, R. Kleiss, S.D. Ellis, Comput. Phys. Commun. **40**, 359 (1986)
4. P.D. Draggiotis, A. van Hameren, R. Kleiss, Phys. Lett. B **483**, 124 (2000)
5. A. van Hameren, C.G. Papadopoulos, Eur. Phys. J. C **25**, 563 (2002)
6. T. Stelzer, W.F. Long, Comput. Phys. Commun. **81**, 357 (1994)
7. R. Kleiss, R. Pittau, Comput. Phys. Commun. **83**, 141 (1994)
8. G.P. Lepage, J. Comput. Phys. **27**, 192 (1978)
9. S. Jadach, Comput. Phys. Commun. **130**, 244 (2000)
10. B. Kersevan, E. Richter-Was, Comput. Phys. Commun. **149**, 142 (2003); the latest version of the manual and code can be found at http://cern.ch/Borut.Kersevan/AcerMC.Welcome.html

11. P.A. Srivastava, G. Sudarshan, Phys. Rev. **110**, 765 (1958)
12. B. Almgren, Arkiv för Fysik **38**, 161 (1968)
13. M.M. Block, Phys. Rev. **101**, 796 (1956)
14. G.I. Kopylov, Zh. Eksperim. i Teor. Fiz. **39**, 1091 (1960); translation in Soviet Physics JETP **12**, 761 (1961)
15. J. Hilgart, R. Kleiss, F. Le Dibider, Comput. Phys. Comm. **75**, 191 (1993)
16. F.A. Berends, R. Kleiss, Comput. Phys. Commun. **85**, 11 (1996)
17. S. Jadach, J.H. Kuhn, Z. Was, Comput. Phys. Commun. **64**, 275 (1990); M. Jezabek, Z. Was, S. Jadach, J.H. Kuhn, Comput. Phys. Commun. **70**, 69 (1992); R. Decker, S. Jadach, J.H. Kuhn, Z. Was, Comput. Phys. Commun. **76**, 361 (1993)
18. P. Nyborg, Phys. Rev. B **140**, 921 (1965)
19. E. Byckling, K. Kajantie, Nucl. Phys. B **9**, 568 (1969); Particle kinematics (Wiley & Sons, London 1973)
20. P. Nyborg, H.S. Song et al., Phys. Rev. B **140**, 914 (1965)
21. C.J. Everett, E.D. Cashwell, LA-9721-MS (1983) 150p
22. P. Lichard, Acta Phys. Slov. **49**, 215 (1999)
23. B.C. Carlson, math.CA/9409227
24. P. Appell, J. Kampé de Fériet, Fonctions hypergéométriques et hypersphériques: polynomes d'Hermite (Gauthier-Villars, Paris 1926)
25. D. Zwillinger (ed.), CRC Standard Mathematical Tables and Formulae (CRC Press, 2000), 30th edition; code derived from http://www.csit.fsu.edu/~burkardt/f_src/quadrule/quadrule.f90
26. W.H. Press et al., Numerical recipes in Fortran 77 (Cambridge Univ. Press, 1992) second edition
27. B. Kersevan, E. Richter-Was, preprint TPJU-6/2004; hep-ph/0405247; the latest version of the manual and code can be found at http://cern.ch/Borut.Kersevan/AcerMC.Welcome.html
28. T. Sjöstrand et al., Comput. Phys. Commun. **135**, 238 (2001)
29. F.A. Berends, C.G. Papadopoulos, R. Pittau, Comput. Phys. Commun **136**, 148 (2001)
30. S. Jadach, E. Richter-Was, B.F.L. Ward, Z. Was, Comput. Phys. Commun. **70**, 305 (1992)