



# Representation Theory Based Algorithm to Compute Boltzmann's Bilinear Collision Operator in the Irreducible Spectral Burnett Ansatz Efficiently

Andrea Hanke<sup>1</sup> · Manuel Torrilhon<sup>1</sup>

Received: 16 January 2022 / Revised: 13 January 2023 / Accepted: 8 February 2023 /  
Published online: 21 April 2023  
© The Author(s) 2023

## Abstract

Numerically solving the Boltzmann equation is computationally expensive in part due to the number of variables the distribution function depends upon. Another contributor to the complexity of the Boltzmann Equation is the quadratic collision operator describing changes in the distribution function due to colliding particle pairs. Solving it as efficiently as possible has been a topic of recent research, e.g. Cai and Torrilhon (Phys Fluids 31(12):126105, 2019. <https://doi.org/10.1063/1.5127114>), Wang and Cai (J Comput Phys 397:108815, 2019. <https://doi.org/10.1016/j.jcp.2019.07.014>), Cai et al. (Comput Fluids 200:104456, 2020. <https://doi.org/10.1016/j.compfluid.2020.104456>). In this paper we exploit results from representation theory to find a very efficient algorithm both in terms of memory and computational time for the evaluation of the quadratic collision operator. With this novel approach we are also able to provide a meaningful interpretation of its structure.

**Keywords** Gas dynamics · Particle interactions · Moment equations · Symmetric trace-free tensors · Spherical harmonics · Boltzmann equation

## 1 Introduction

Solving the Boltzmann equation numerically is an area of ongoing research. There are already existing results based on Spectral-Fourier Approach by Pareschi and Russo [1] and Gamba and Tharkabhushanam [2], with recent work improving some of its shortcomings such as [3] on the preservation of conservations or [4] on the stability of such methods. Another ansatz is based on the spectral Hermite ansatz introduced by Grad (1949) [5] with early work [6] and more recent work such as [7, 8] on linearising the collision operator. Wang and Cai were

---

✉ Andrea Hanke  
hanke@acom.rwth-aachen.de  
Manuel Torrilhon  
mt@acom.rwth-aachen.de

<sup>1</sup> Applied Computational Mathematics, RWTH Aachen University, Schinkelstr. 2, 52062 Aachen, Germany

recently [9] able to calculate the bilinear collision operator based on the spectral Hermite ansatz. However, their method is computationally expensive. In response, Cai et al. [10] developed and implemented a more efficient algorithm based on the spectral Burnett ansatz. This ansatz has also been looked into analytically [11] and numerically [12]. Inspired by Struchtrup's (2005) tensorial approach to the collision operator in Equation 6.35 in [13], we chose a different approach for calculating the Spectral-Burnett approximation of the Boltzmann collision operator. The main difference lies in the exploitation of features of the basis set, which is adapted to the irreducible subspaces with respect to the orthogonal group of the polynomial space as it consists of real solid spherical harmonics multiplied with Laguerre polynomials. While we also implemented the mostly analytical calculation of collision coefficients for a variety of potentials, the focus of this paper is the algorithm and corresponding numerical code, which uses these coefficients to solve the space-homogeneous Boltzmann equation for any distribution. Utilizing the properties of the irreducible subspaces, compared to the previous works we are able to significantly reduce the memory and computation time required for calculating the numerical solution. We could achieve this by virtue of the uniqueness up to a constant of linear maps between the irreducible subspaces. Representation theory allowed us to isolate each bilinear map, the constant for each map that encodes particle collisions and identify all bilinear maps, that evaluate to zero due to the mathematical properties of the irreducible subspaces in question. Our method provides a deep understanding of the underlying structure of the bilinear collision operator, allows for a very flexible setup for numerical computations and therefore is ideal for a purposeful reduction of computational effort.

The structure of this paper is as follows: The next section provides an overview of the collision operator and the potentials used in this paper. Sections 3 and 4 describe the application of representation theory to the collision operator, while Sect. 5 explains the representation theoretically inspired algorithm and its implementation. In Sect. 6 we apply our algorithm and also compare the results with [9]. For an overview of the applied representation theory refer to the "Appendix".

## 2 Boltzmann's Collision Operator

To describe gas flows, the Boltzmann equation is known to be valid for a wide range of situations. In particular, it describes non-equilibrium gas flows accurately for Knudsen numbers in the transition regime between the range of Navier–Stokes–Fourier equations and the kinetic regime. The Boltzmann equation describes the evolution of the gas distribution function  $f$ , which in general depends on seven variables—time  $t$ , position space  $\mathbf{x}$  and velocity space  $\mathbf{c}$ .

As done in previous papers we look at the space-homogenous Boltzmann equation to be able to focus solely on the collision operator:

$$\frac{\partial f}{\partial t} = \mathcal{S}(f, f) \quad (1)$$

where  $f$  gives the time-dependent velocity distribution of particles,

$$f : [0, T] \times \mathbb{R}^3 \rightarrow \mathbb{R}$$

$$f(t, \mathbf{c}) \, d\mathbf{c} = \text{density of particles with velocity in } [\mathbf{c}, \mathbf{c} + d\mathbf{c}] \text{ at } t, \quad (2)$$

and  $S(f, f)$  is the bilinear collision operator, see e.g. the textbooks [14, 15], modelling the binary collisions of the particles in the gas:

$$S(f, f) = \int_{\mathbb{R}^3} \int_{\mathbf{n} \perp \mathbf{g}} \int_0^\pi (f(t, \mathbf{c}') f(t, \mathbf{c}'_1) - f(t, \mathbf{c}) f(t, \mathbf{c}_1)) B(|\mathbf{g}|, \chi) d\chi d\mathbf{n} d\mathbf{c}_1 \quad (3)$$

where  $\mathbf{n}$  is the unit vector orthogonal to the collisional plane, the relative velocity  $\mathbf{g} = \mathbf{c} - \mathbf{c}_1$ , the average velocity  $\mathbf{h} = \frac{1}{2}(\mathbf{c} + \mathbf{c}_1)$ , and the primed velocities  $\mathbf{c}', \mathbf{c}'_1, \mathbf{g}' = \mathbf{c}' - \mathbf{c}'_1$  with  $\|\mathbf{g}'\| = \|\mathbf{g}\|, \mathbf{h}' = \mathbf{h}$  are the velocities after the collision. Additionally, we assume the initial average velocity of the gas to be 0, and as this is one of the conserved moments it will always be 0. The solution of (1) converges to a Gaussian centered around the average velocity. Our assumption will allow us to exactly represent this solution with our numerical ansatz.

As discussed in several textbooks, [14, 15], they are given by

$$\mathbf{g}' = R_{\mathbf{n}, \chi}(\mathbf{g}) = \cos(\chi)\mathbf{g} + \sin(\chi)\mathbf{n} \times \mathbf{g} \quad \mathbf{c}' = \mathbf{h} + \frac{1}{2}\mathbf{g}' \quad \mathbf{c}'_1 = \mathbf{h} - \frac{1}{2}\mathbf{g}', \quad (4)$$

where  $\mathbf{g}'$  is the result of rotating  $\mathbf{g}$  around  $\mathbf{n}$  by the angle  $\chi$ . According to Cercignani’s work [14], the kernel  $B(|\mathbf{g}|, \chi)$  depends on the assumed potential between the colliding particles. Following [14, 15], different choices are possible and common models [7–10] include

- The hard spheres model, i.e.

$$B(|\mathbf{g}|, \chi) = |\mathbf{g}| \frac{d^2}{4} \quad \text{with } d \text{ a constant for the diameter of the colliding particles.} \quad (5)$$

- Its generalization given by the variational hard spheres (VHS) model,

$$B(|\mathbf{g}|, \chi) = |\mathbf{g}|^{\frac{\eta-5}{\eta-1}} \frac{d^2}{4}. \quad (6)$$

- The inverse power law potentials (IPL), i.e. with  $\eta > 3$  determining the power law exponent

$$B(|\mathbf{g}|, \chi) = |\mathbf{g}|^{\frac{\eta-5}{\eta-1}} W_0 \frac{dW_0}{d\chi} \quad \text{with} \quad \chi = \pi - 2 \int_0^{W_1} \left( 1 - W^2 - \frac{2}{\eta-1} \left( \frac{W}{W_0} \right)^{\eta-1} \right)^{-\frac{1}{2}} dW \quad (7)$$

$$\text{and } W_1 \text{ the positive root of} \quad 1 - W_1^2 - \frac{2}{\eta-1} \left( \frac{W_1}{W_0} \right)^{\eta-1}. \quad (8)$$

We use the change of variables given in [7] to evaluate  $W_0(\chi)$ .

We identify two types of assumed potentials corresponding to whether or not  $B$  depends on the angle  $\chi$ . Additionally,  $\eta$  can be chosen from  $(3, \infty)$ . Note that the IPL potential models can be derived from assuming a specific potential between the colliding particles, whereas the VHS model only allows that for  $\eta = \infty$ , in which case the two models correspond to each other. In this sense, the power potentials can be seen as more realistic models and the VHS models as a simplified version.

In this paper we will work with different examples of these models and compare them. In particular, we work with the classical  $\chi$ -dependent Maxwell Potential with  $\eta = 5$ , the  $\chi$ -independent and hence simplified Maxwell potential, as well as the hard power potential with  $\eta = 10$ . As we will explore in Sect. 4.2.2, these choices reflect different levels of computational difficulty both in a priori calculations as well as in the numerical evaluation.

### 3 Irreducible Burnett Ansatz

#### 3.1 Traditional Setup

##### 3.1.1 Expansion of the Distribution Function

To solve the space homogeneous Boltzmann equation numerically, we follow [16] to expand the distribution function with an orthonormal basis set  $\Upsilon_\alpha(\mathbf{c})$ ,

$$f(t, \mathbf{c}) = \sum_\alpha w_\alpha(t) \Upsilon_\alpha(\mathbf{c}) \quad \text{with} \quad \Upsilon_\alpha(\mathbf{c}) := \Psi_\alpha(\mathbf{c}) \cdot f^{(\text{eq})}(\mathbf{c}) \quad (9)$$

where  $\alpha$  is typically a multi-index in three indices determining the basis polynomial  $\Psi_\alpha$ ,  $f^{(\text{eq})}(\mathbf{c})$  is the global equilibrium or Maxwellian function and the scalar product is chosen for the weighted function space  $W_{\text{eq}} : L^2(\mathbb{R}^3, \mathbb{R}, \frac{1}{f^{(\text{eq})}} \, d\mathbf{c})$ , see e.g. [16]:

$$\begin{aligned} \langle \cdot, \cdot \rangle : W_{\text{eq}} \times W_{\text{eq}} &\rightarrow \mathbb{R} \\ \langle g, h \rangle &:= \int_{\mathbb{R}^3} gh \frac{1}{f^{(\text{eq})}} \, d\mathbf{c}, \end{aligned} \quad (10)$$

to allow us to identify the weights  $w_\alpha$  as moments of the distribution function. Here, we shift the velocity-coordinates such that the medium velocity is zero and rescale the variables so that the temperature is 1, giving us

$$f^{(\text{eq})} := \frac{1}{(\sqrt{2\pi})^3} \exp\left(-\frac{\mathbf{c}^2}{2}\right). \quad (11)$$

Generalizing this approach, the distribution function  $f(t, \mathbf{c})$  is approximated as a time-dependent polynomial in  $\mathbf{c}$  of degree  $M \in \mathbb{N}$  or less, multiplied by the global Maxwellian,

$$f(t, \mathbf{c}) = \underbrace{p_t(\mathbf{c})}_{\substack{\text{time-dependent,} \\ \text{real valued} \\ \text{polynomial in } \mathbf{c} \text{ of degree } \leq M}} \cdot \underbrace{f^{(\text{eq})}(\mathbf{c})}_{\substack{\text{global Maxwellian,} \\ \text{Gaussian-like} \\ \text{equilibrium distribution}}},$$

implicitly choosing the approximation space for the distribution function  $f(t, \mathbf{c}) \in V_{\text{eq}}$  with

$$V_{\text{eq}} := P^M \cdot f^{(\text{eq})}(\mathbf{c}), \quad P^M \subset \mathbb{R}[c_x, c_y, c_z], \quad P^M := \{p \in \mathbb{R}[c_x, c_y, c_z] \mid \text{degree of } p \leq M\} \cong \mathbb{R}^N, \quad (12)$$

where  $\mathbb{R}[c_x, c_y, c_z]$  is the space of all real polynomials in three variables and  $N = \frac{1}{6}(M + 1)(M + 2)(M + 3)$  the dimension of  $P^M$ . We can equip the polynomial space with a scalar product corresponding to the setup (10):

$$\begin{aligned} \langle \cdot, \cdot \rangle_P : \mathbb{R}[c_x, c_y, c_z] \times \mathbb{R}[c_x, c_y, c_z] &\rightarrow \mathbb{R} \\ \langle p, q \rangle_P &:= \int_{\mathbb{R}^3} pq f^{(\text{eq})} \, d\mathbf{c}. \end{aligned} \quad (13)$$

### 3.1.2 Expansion of the Collision Operator

Inserting the expansion of the distribution function (9) into the space homogeneous Boltzmann equation (1) and using the linearity of the bilinear collision operator yields:

$$\sum_{\alpha'} \Upsilon_{\alpha'}(\mathbf{c}) \frac{\partial w_{\alpha'}(t)}{\partial t} = \sum_{\beta\gamma} \mathcal{S}(\Upsilon_{\beta}, \Upsilon_{\gamma}) w_{\beta} w_{\gamma}. \tag{14}$$

We now project this expression onto  $V_{\text{eq}}$  spanned by orthonormal basis functions  $\Upsilon_{\alpha}$ . The left hand side leads to

$$\left\langle \Upsilon_{\alpha}, \sum_{\alpha'} \Upsilon_{\alpha'} \frac{\partial w_{\alpha'}(t)}{\partial t} \right\rangle = \sum_{\alpha'} \underbrace{\langle \Upsilon_{\alpha}, \Upsilon_{\alpha'} \rangle}_{\delta_{\alpha'\alpha}} \frac{\partial w_{\alpha'}(t)}{\partial t} = \frac{\partial w_{\alpha}(t)}{\partial t},$$

while the right hand side evaluates to

$$\sum_{\beta,\gamma} \underbrace{\langle \Upsilon_{\alpha}, \mathcal{S}(\Upsilon_{\beta}, \Upsilon_{\gamma}) \rangle}_{C_{\alpha\beta\gamma}^{(\Psi)}} w_{\beta} w_{\gamma} =: s_{\alpha}. \tag{15}$$

Overall, we find that the rate of change of the distribution function’s weights depends on the collision vector  $s_{\alpha}$ ,

$$\frac{\partial w_{\alpha}(t)}{\partial t} = s_{\alpha}. \tag{16}$$

The collision vector in turn is determined by the bilinear operation (15) between the weights  $w_{\alpha}$  of the distribution function and a basis-dependent collision tensor

$$C_{\alpha\beta\gamma}^{(\Psi)} := \langle \Psi_{\alpha} f^{(\text{eq})}, \mathcal{S}(\Psi_{\beta} f^{(\text{eq})}, \Psi_{\gamma} f^{(\text{eq})}) \rangle. \tag{17}$$

### 3.2 Structure of the Approximation Space

The remaining choice are the  $N$  basis polynomials  $\Psi_{\alpha}(\mathbf{c})$ , with which we span  $P^M$ , which has great influence on the structure and sparsity of the collision tensor and hence on the required computational time and memory to obtain  $s_{\alpha}$ .

For example Grad [5] or more recently Cai [9] choose Hermite polynomials in three coordinates, since these are known to be orthogonal with respect to the corresponding scalar product  $\langle \cdot, \cdot \rangle_P$ . However, structure and sparsity of the resulting collision tensor can still be improved.

We want to use a basis that exploits a key feature of the collision operator: It describes a physical process in the three dimensional velocity space, requiring it to be, like all physical processes, compatible with isometric transformations of the coordinate system, i.e. with  $\mathcal{O}_3$  the group of orthogonal matrices from  $\mathbb{R}^{3 \times 3}$ :

$$\forall r \in \mathcal{O}_3 : r(\mathcal{S}(f)(\mathbf{c})) = \mathcal{S}(r(f)). \tag{18}$$

While this identity is often accepted, we included a proof in the ‘‘Appendix’’ for completeness, as this statement is central for this paper. This restriction allows us to employ results from the representation theory of  $\mathcal{O}_3$  and reduce the bilinear maps with which we express  $s_{\alpha}$  to only those bilinear maps between the  $\mathcal{O}_3$ -irreducible subspaces of  $P^M$ .

### 3.2.1 Decomposition of $P^M$ into Its Irreducible Subspaces

With the space of homogeneous and harmonic polynomials of degree  $n \in \{0, 1, \dots, M\}$  and dimension  $2n + 1$

$$\mathcal{H}^n := \left\{ p \in \mathbb{R}[c_x, c_y, c_z] \mid \begin{array}{l} \forall \lambda \in \mathbb{R} : p(\lambda c_x, \lambda c_y, \lambda c_z) = \lambda^n p(c_x, c_y, c_z) \\ \Delta p := (\partial_x^2 + \partial_y^2 + \partial_z^2)p = 0 \end{array} \right\} \cong \mathbb{R}^{2n+1} \tag{19}$$

we have the building blocks for the irreducible subspaces of  $P^M$  [17]. To make it clear when an index describes the degree of a harmonic polynomial, we call  $n$  the *anisotropic index*. We also define the one-dimensional space of homogeneous isotropic polynomials of second degree

$$\Omega := \left\{ p(c_x, c_y, c_z) = \lambda(c_x^2 + c_y^2 + c_z^2) \mid \lambda \in \mathbb{R} \right\} \in P^M, \tag{20}$$

and denote with  $c^2$  its element where  $\lambda = 1$ . Ignoring the chosen scalar product from Eq. (13) for now allows us to decompose  $P^M$  into a direct sum of homogeneous polynomials with respect to some corresponding scalar product,

$$P^M = \bigoplus_{a=0}^{\lfloor \frac{M}{2} \rfloor} \bigoplus_{n=0}^{M-2a} (c^2)^a \cdot \mathcal{H}^n. \tag{21}$$

We can check that this decomposition is complete by counting dimensions:

$$\sum_{a=0}^{\lfloor \frac{M}{2} \rfloor} \sum_{n=0}^{M-2a} \dim((c^2)^a \mathcal{H}^n) = \sum_{a=0}^{\lfloor \frac{M}{2} \rfloor} \sum_{n=0}^{M-2a} 2n + 1 = \frac{1}{6}(M + 1)(M + 2)(M + 3) = \dim(P^M). \tag{22}$$

Figure 1 sorts these subspaces column-wise according to their homogeneous degree  $\deg = 2a + n$  and row-wise according to their anisotropic index  $n$ . Subspaces with same anisotropic index  $n$  are isomorphic to each other and the decomposition is not unique within each set of isomorphic subspaces. Subspaces with different anisotropic indices are non-isomorphic to each other and therefore necessarily pairwise orthogonal with respect to any  $\mathcal{O}_3$  invariant scalar product, see Remark 9.5 in this paper’s “Appendix”.

Now we come back to the chosen scalar product in (13). The isomorphic subspaces shown in Fig. 1 are not orthogonal with respect to  $\langle \cdot, \cdot \rangle_P$ . We therefore orthogonalize each set of isomorphic subspaces with a linear combination of suitable powers of  $(c^2)$  determined by associated Laguerre polynomials:

$$P^M = \bigoplus_{a=0}^{\lfloor \frac{M}{2} \rfloor} \bigoplus_{n=0}^{M-2a} \underbrace{L_a^{(n+\frac{1}{2})} \left( \frac{c^2}{2} \right)}_{=: {}^a\mathcal{H}^n} \cdot \mathcal{H}^n, \tag{23}$$

where  ${}^a\mathcal{H}^n$  are the orthogonalized irreducible subspaces. We call  $a$  the *isotropic index* and are left with the basis choice for  $\mathcal{H}^n$ .

	deg = 0	deg = 1	deg = 2	deg = 3	deg = 4	deg = 5	deg = 6	...	deg = M	
n = 0	$\mathcal{H}^0$		$c^2 \cdot \mathcal{H}^0$		$(c^2)^2 \cdot \mathcal{H}^0$		$(c^2)^3 \cdot \mathcal{H}^0$	...	$(c^2)^{\lfloor \frac{M}{2} \rfloor} \cdot \mathcal{H}^{M-2\lfloor \frac{M}{2} \rfloor}$	
n = 1	$\mathcal{H}^1$			$c^2 \cdot \mathcal{H}^1$		$(c^2)^2 \cdot \mathcal{H}^1$	...			
n = 2	$P^1$			$\mathcal{H}^2$		$c^2 \cdot \mathcal{H}^2$	$(c^2)^2 \cdot \mathcal{H}^2$	...		
n = 3	$P^2$				$\mathcal{H}^3$		$c^2 \cdot \mathcal{H}^3$	...		
n = 4	$P^3$					$\mathcal{H}^4$		$c^2 \cdot \mathcal{H}^4$	...	
n = 5						$\mathcal{H}^5$	...			
n = 6							$\mathcal{H}^6$	...	$c^2 \cdot \mathcal{H}^{M-2}$	
⋮								⋮		
n = M	$P^M$									$\mathcal{H}^M$

**Fig. 1** Decomposition of  $P^1, P^2, P^3$  to  $P^M$  into their homogeneous irreducible subspaces. The irreducible subspaces are sorted column-wise by their homogeneous degree  $\text{deg} = 2a + n$  and row-wise by the degree  $n$  of the harmonic spaces  $\mathcal{H}^n$ . Those that share a row are isomorphic, subspaces that are from different rows are pairwise non-isomorphic. The decomposition is not unique within each row, linear combinations of different orders of  $c^2$  are possible. Here we chose the homogeneous polynomials of  $c^2$  for visual clarity. All non-isomorphic irreducible subspaces are orthogonal to each other for every  $\mathcal{O}_3$  invariant scalar product. To achieve orthogonality with respect to our chosen scalar product within each row of isomorphic irreducible subspaces, we use the linear combinations given by associated Laguerre polynomials of  $c^2, L_a^{(n+1/2)}(c^2/2)$

### 3.2.2 Basis Choice Adapted to the Irreducible Subspaces

Here we employ the real-valued solid spherical harmonics  $Y_n^l : \mathbb{R}^3 \rightarrow \mathbb{R}$  of  $n$ -th degree with the *directional index*  $l \in \{-n, -n + 1, \dots, n - 1, n\}$ , related to the real spherical harmonics  $\mathcal{Y}_n^l : S^2 \rightarrow \mathbb{R}$  by extension  $S^2 \rightarrow \mathbb{R}^3$  via

$$Y_n^l(\mathbf{c}) = \|\mathbf{c}\|^n \cdot \mathcal{Y}_n^l\left(\frac{\mathbf{c}}{\|\mathbf{c}\|}\right). \tag{24}$$

Note, that we can draw parallels to indices in quantum mechanics where the spherical harmonics are complex, and the principal, azimuthal and magnetic quantum numbers have similar mathematical roles to the *isotropic index*  $a$ , *anisotropic index*  $n$  and *directional index*  $l$ , respectively.

This leads to the following choice for  $\Psi_\alpha$ :

$$\Psi_\alpha = \Psi_{anl} = \underbrace{\frac{1}{\mathcal{N}_{anl}}}_{\text{normalization coefficient}} \underbrace{Y_n^l(\mathbf{c})}_{\text{real solid spherical harmonic}} \underbrace{L_a^{(n+\frac{1}{2})}\left(\frac{c^2}{2}\right)}_{\text{associated Laguerre polynomials}}, \tag{25}$$

with  $(a, n, l) \in \mathbb{N}_0 \times \mathbb{N}_0 \times \{-n, \dots, n\}$  specifying the multi-index  $\alpha$ , and the normalization coefficient given by

$$\mathcal{N}_{anl} := \sqrt{\left\langle Y_n^l L_a^{(n+\frac{1}{2})}, Y_n^l L_a^{(n+\frac{1}{2})} \right\rangle_P}.$$

### 3.3 Discussion of the Basis Set and Irreducibility

Instead of immediately choosing a basis set, we first looked at the irreducible subspaces of  $P^M$  with respect to  $\mathcal{O}_3$  and collected them as spaces of homogeneous polynomials in Fig. 1. Upon finding that we have multiple isomorphic subspaces, all row-wise collected in Fig. 1, we linearly combined them via Laguerre polynomials to obtain  ${}^a\mathcal{H}^n$ , the orthogonalized irreducible subspaces. Then we chose basis polynomials for each  ${}^a\mathcal{H}^n$ , but again using the isomorphism between spaces within one row, it was enough to choose a set of basis polynomials for each lowest degree irreducible subspace per row, i.e. for the spaces  $\mathcal{H}^n$  of harmonic polynomials of degree  $n$ .

We have chosen real solid spherical harmonics, since they are easily accessible, orthogonal and adapted to the chain of subgroups  $\mathcal{O}_3 \geq \mathcal{O}_2 \geq \mathcal{O}_1$  with  $c_z$  the special symmetry direction, i.e. for an  $\mathcal{O}_2$ -invariant problem rotated so that  $c_z$  is the rotation axis for  $\mathcal{O}_2$ , all coefficients with  $l \neq 0$  are zero.

The basis set containing all basis polynomials with  $2a + n \leq M$  spans the full space  $P^M$ . Note, that only basis polynomials with  $n = 0$  are rotationally invariant. However, similar to the decomposition of a natural number into its prime factors, this basis is adapted to the decomposition of  $P^M$  into its *smallest invariant subspaces*. Here “smallest” is defined as an invariant subspace  $U$ , that cannot be further decomposed into invariant subspaces other than  $U$  itself and the zero space, similar to a prime number only being divisible by itself and 1. These subspaces are called *irreducible*. Figure 1 shows the decomposition of  $P^M$  into its homogeneous irreducible subspaces, while we use orthogonalized subspaces with respect to  $\langle \cdot, \cdot \rangle_P$  given for a pair of chosen isotropic index  $a$  and anisotropic index  $n$  by

$$\text{span} \{ \Psi_{anl} \mid -n \leq l \leq n \}$$

with a dimension of  $2n + 1$ .

Note that we have recovered the spectral Burnett ansatz with real spherical harmonics. First results with the spectral Burnett basis have been developed previously in [10–12, 18], but we are the first to exploit the underlying structure of the collision operator in our implementation. We also implement the nearly fully analytic calculation of the tensor coefficients and carry out numerical tests for large  $M$ .

## 4 Implications of Representation Theory

### 4.1 Decomposing the Collision Operator

Equation (15) is the orthogonal projection of  $\mathcal{S}(f, f)$  onto the approximation space  $V_{\text{eq}}$ . Equivalently we can identify the bilinear map  $\mathcal{C} : P^M \times P^M \rightarrow P^M$  with coefficients given by the collision tensor in terms of the basis functions of  $P^M$ , now defined as

$$C_{\alpha\beta\gamma} = C_{anl, b\hat{n}\hat{l}, c\hat{n}\hat{l}} := \langle \Psi_{anl}, \mathcal{S}(\Psi_{b\hat{n}\hat{l}} f^{(\text{eq})}, \Psi_{c\hat{n}\hat{l}} f^{(\text{eq})}) \rangle_P, \tag{26}$$

i.e. with  $p, q \in P^M$  and their respective expansion coefficients  $w_{anl}, u_{anl}$  we can express the map  $\mathcal{C}$  via

$$\mathcal{C}(p, q) = \sum_{anl} \sum_{b\hat{n}\hat{l}} \sum_{c\hat{n}\hat{l}} C_{anl, b\hat{n}\hat{l}, c\hat{n}\hat{l}} w_{b\hat{n}\hat{l}} u_{c\hat{n}\hat{l}} \Psi_{anl}. \tag{27}$$



Not only is  $\mathcal{C}$  a bilinear map, it inherits the property of the collision operator  $\mathcal{S}$  that it is linear with respect to  $\mathcal{O}_3$ , compare to Eq. (18):

$$\forall r \in \mathcal{O}_3 : r(\mathcal{C}(p, q)) = \mathcal{C}(r(p), r(q)). \tag{28}$$

Such a map is called *natural*. This restriction has been studied in representation theory and implies a decomposition of  $\mathcal{C}$  into a sum over isotropic and anisotropic indices

$$\mathcal{C} : P^M \times P^M \rightarrow P^M = \sum_{abc} \sum_{n\hat{n}\tilde{n}} \tag{29}$$

with  $S_{abc}^{n\hat{n}\tilde{n}} \in \mathbb{R}$  some constants that, given a set of maps  $\Phi_{abc}^{n\hat{n}\tilde{n}}$ , distinguish between different maps  $\mathcal{C}$ . Each  $\Phi_{abc}^{n\hat{n}\tilde{n}} : P^M \times P^M \rightarrow P^M$  is a non-zero natural bilinear map given by

$$\begin{array}{ccc}
 P^M & & P^M \\
 \downarrow \mathcal{P}(\hat{n}, b) & & \downarrow \mathcal{P}(\tilde{n}, c) \\
 {}^b\mathcal{H}^{\hat{n}} & & {}^c\mathcal{H}^{\tilde{n}} \\
 \downarrow \psi(\hat{n}, b) & & \downarrow \psi(\tilde{n}, c) \\
 \mathcal{H}^{\hat{n}} & & \mathcal{H}^{\tilde{n}} \\
 \underbrace{\hspace{10em}} & & \\
 & \downarrow Q^{n\hat{n}\tilde{n}} & \\
 & \mathcal{H}^n & \\
 & \downarrow \psi_{(n,a)}^{-1} & \\
 & {}^a\mathcal{H}^n & \\
 & \downarrow \mathcal{E}_{(n,a)} & \\
 & P^M & 
 \end{array} \tag{30}$$

with orthogonal projections  $\mathcal{P}_{(n,a)} : P^M \rightarrow {}^a\mathcal{H}^n$ , isomorphisms  $\psi_{(n,a)} : {}^a\mathcal{H}^n \rightarrow \mathcal{H}^n$ , embedding  $\mathcal{E}_{(n,a)} : {}^a\mathcal{H}^n \rightarrow P^M$ , and the *universal* natural bilinear map

$$Q^{n\hat{n}\tilde{n}} : \mathcal{H}^{\hat{n}} \times \mathcal{H}^{\tilde{n}} \rightarrow \mathcal{H}^n. \tag{31}$$

This map  $Q^{n\hat{n}\tilde{n}}$  is either zero or unique up to a constant, depending on the combination of anisotropic indices:

$$Q^{n\hat{n}\tilde{n}} \text{ is zero unless } \left\{ \begin{array}{l} \hat{n} + \tilde{n} - n \in 2\mathbb{N}_0, \text{ and} \\ 0 \leq \hat{n} + \tilde{n} - n \leq 2 \min(\hat{n}, \tilde{n}). \end{array} \right. \tag{32}$$

Note, that a decomposition of a (multi-) linear map between vector spaces into a sum of maps between subspaces as done in Eq. (29) by itself is generally possible. Equation (28) is crucial for the decomposition of  $\mathcal{C}$  in (29), requiring  $\Phi_{abc}^{n\hat{n}\tilde{n}}$  and  $Q^{n\hat{n}\tilde{n}}$  to be natural as well,

and ultimately restricting all possible bilinear and natural maps  $\mathcal{H}^{\hat{n}} \times \mathcal{H}^{\tilde{n}} \rightarrow \mathcal{H}^n$  to be the same up to a constant with the additional condition (32).

### 4.2 Decoding the Structure of the Collision Tensor

Overall, the previous result implies the decomposition of the collision tensor into two tensors:

$$C_{ant,\hat{b}\hat{l},c\tilde{n}\tilde{l}} = S_{n\hat{n}\tilde{n}}^{abc} Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}, \tag{33}$$

mirroring the structure of irreducible subspaces of  $P^M$ .  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$  and  $S_{n\hat{n}\tilde{n}}^{abc}$  have two distinct roles for the overall collision tensor. The *coupling tensor*  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$  collects the coefficients of the map  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$ . It is independent of the potential and encodes the purely mathematical coupling of  $Y_n^{\hat{l}}$  and  $Y_{\tilde{n}}^{\tilde{l}}$  to  $Y_n^l$ . The *impact tensor*  $S_{n\hat{n}\tilde{n}}^{abc}$  encodes the physics of the particle collisions.

#### 4.2.1 The Coupling Tensor

The coupling tensor  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$  encodes the natural bilinear map, unique up to a constant factor, from two harmonic polynomial spaces of degrees  $\hat{n}$  and  $\tilde{n}$  to a third harmonic polynomial space of degree  $n$ , see Eq. (31). Crucially, two natural bilinear maps between different, but isomorphic irreducible subspaces are identical up to a constant factor and hence all can be expressed in terms of one  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$ . Referring to Fig. 1 and knowing that all subspaces of the same row are isomorphic to each other, we can also picture  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$  as the bilinear map between three rows. This universality is the key in our algorithm, allowing us to achieve high efficiency both in terms of computational time and memory usage. We can calculate the values of the coupling tensor for example with

$$Q_{l\tilde{l}}^{n\hat{n}\tilde{n}} = \langle Y_n^l, Y_{\hat{n}}^{\hat{l}} \cdot Y_{\tilde{n}}^{\tilde{l}} \rangle_P = \int_{\mathbb{R}^3} Y_n^l \cdot Y_{\hat{n}}^{\hat{l}} \cdot Y_{\tilde{n}}^{\tilde{l}} f^{(eq)} \, d\mathbf{c}. \tag{34}$$

The coupling tensor is related to the Clebsch–Gordon-Coefficients [19]. While the Clebsch–Gordon-Coefficients are more common in quantum physics, where they are usually used for three harmonic spaces of complex polynomials, the coupling tensor is the equivalent for real harmonic polynomials. It shares the property that it is zero for most combinations of anisotropic and directional indices  $(n, \hat{n}, \tilde{n}, l, \hat{l}, \tilde{l})$ , but does not have as many restrictions as its complex counterpart. In particular, the conditions in Eq. (32) hold for  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$ , and if they are met, then  $Q_{000}^{n\hat{n}\tilde{n}}$  is unequal to zero. Keeping in mind the uniqueness up to a constant of the coupling tensor, we choose one instance of  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$  for each triplet  $(n, \hat{n}, \tilde{n})$ , calculate the values of the coupling tensor once and hard-code them in the implementation as function `calcQ`.

#### 4.2.2 The Impact Tensor

In contrast to  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$ , the impact tensor  $S_{n\hat{n}\tilde{n}}^{abc}$  contains all information about the physics of the collision. While from the mathematical point of view the constant factor distinguishing two maps  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$  with same anisotropic indices  $(n, \hat{n}, \tilde{n})$  may seem irrelevant, it is clear that in the physical application these factors are the main interest. Here the conservation rules come into play as zeroes for particular combinations of anisotropic and isotropic indices  $a, b, c, (n, \hat{n}, \tilde{n})$ , but this is also where different potentials lead to different values in the

**Table 1** Overview of a few choices for  $\eta$  and  $\chi$ -dependence in  $B(|\mathbf{g}|, \chi)$  and their corresponding potential models

	Realistic power potential $\chi$ -dependent	VHS model $\chi$ -independent
$\eta = 5$	Maxwell potential	Simplified Maxwell potential
$\eta = 10$	Hard power potential	Hard VHS model
$\eta = \infty$	Hard spheres potential	

Note, that for  $\eta = \infty$  the power potential corresponds to the VHS model, and therefore the Hard spheres potential is  $\chi$ -independent

coefficients of the impact tensor. Notice, that the impact tensor is independent of the directional indices  $(l, \hat{l}, \tilde{l})$ , so with our chosen map  $Q^{n\hat{n}\tilde{n}}$  we can compute its coefficients with

$$S_{n\hat{n}\tilde{n}}^{abc} = \frac{\langle \Psi_{an0} f^{(eq)}, \mathcal{S}(\Psi_{b\hat{n}0} f^{(eq)}, \Psi_{c\tilde{n}0} f^{(eq)}) \rangle}{Q_{000}^{n\hat{n}\tilde{n}}} \tag{35}$$

for combinations of anisotropic indices  $(n, \hat{n}, \tilde{n})$  fulfilling the conditions in (32). Depending on the potential, we can solve the integration in (35) fully or almost fully analytically with a computer algebra system, see the git-repository [20] for further details.

Referring to Table 1, we can now see where each chosen potential falls on the spectrum of high and low computational difficulty of both a priori evaluation, i.e. the calculation of the impact tensor itself, and runtime evaluation of the collision vector.

Note, that there is some relation between these two types of difficulty. The a priori evaluation only determines a set of numbers that make up the impact tensor, whereas the runtime evaluation mostly depends on the number of non-zero impact tensor entries and not on the specific non-zero entry  $S_{n\hat{n}\tilde{n}}^{abc}$ . As already noted by Grad [5], for the potentials with  $\eta = 5$  the additional condition

$$2a + n \neq 2b + \hat{n} + 2c + c \implies S_{n\hat{n}\tilde{n}}^{abc} = 0 \tag{36}$$

applies, which stems from the orthogonality of the basis functions and the independence of the kernel from  $\mathbf{c}$  and  $\mathbf{c}_1$ . This significantly reduces the amount of non-zero impact tensor coefficients, leading to comparatively low computational difficulty both a priori and during runtime.

The a priori computational difficulty of the impact tensor coefficients additionally is determined by the  $\chi$ -dependency of the kernel  $B$ . All  $\chi$ -independent models are of lower difficulty, with the simplified Maxwell model the lowest difficulty due to (36). On the other side of the spectrum lies the hard power potential with  $\eta = 10$ , since it is  $\chi$ -dependent and condition (36) does not apply, requiring the evaluation of the full set of  $S_{n\hat{n}\tilde{n}}^{abc}$ . This potential also has high runtime computational difficulty, since almost all impact tensor coefficients are non-zero.

### 4.3 Full Generalization of the Collision Operator

Notice, that the decomposition of the collision tensor into the impact and coupling tensor not only implies a particular implementation to calculate the collision vector, but it also gives us a unique understanding of the role of each coefficient. In particular, we can now generalize the collision operator to any symmetric, bilinear map between real polynomials in three variables with physical meaning, i.e. invariance under transformation of the coordinate system: Such a map would not necessarily follow the conservation laws but still follow the

decomposition into coupling and impact tensor. To obtain an upper limit for the computational cost of calculating the collision vector, we can assign the impact tensor random numbers for all combinations of anisotropic and isotropic indices that fulfil Eq. (32). When testing the performance with this setup, we shall call such a potential *general*.

## 5 Implementation

The task is to implement the tensor-vector multiplication between the expansion coefficients of the collision tensor in the form of  $S_{n\hat{n}\tilde{n}}^{abc} Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$  and the expansion coefficients  $w_{anl}$  of the distribution function. The result is saved as a vector with expansion coefficients  $s_{anl}$ , so we want to find an efficient algorithm to compute the collision vector  $s$ , i.e.

$$\text{for all } (a, n, l) : s_{anl} = \sum_{b\hat{b}} \sum_{c\tilde{c}} S_{n\hat{n}\tilde{n}}^{abc} Q_{l\tilde{l}}^{n\hat{n}\tilde{n}} w_{b\hat{b}l} w_{c\tilde{c}l}. \tag{37}$$

### 5.1 Custom Implementation Adapted to the Given Tensor Structure

Looking at the calculation task in (37), we can separate it into operations that need to be carried out for all combinations of anisotropic and isotropic indices. Beginning by setting the collision vector  $s$  to zero, we can rewrite the task as

$$\text{for all } (a, b, c, n, \hat{n}, \tilde{n}) : \text{for all } l : s_{anl} += S_{n\hat{n}\tilde{n}}^{abc} \cdot \sum_{\tilde{l}} Q_{l\tilde{l}}^{n\hat{n}\tilde{n}} w_{b\hat{b}l} w_{c\tilde{c}l}. \tag{38}$$

Note, how the impact tensor is now outside of the sum, only remaining as a constant factor.

#### 5.1.1 Implementing the Coupling Tensor

In fact, we can now isolate the map between the irreducible subspaces of different anisotropic and isotropic indices in terms of an operation that only involves the coupling tensors and the directional indices:

$$\text{for all } l : s_l^{an} += S_{n\hat{n}\tilde{n}}^{abc} \cdot \sum_{\tilde{l}} Q_{l\tilde{l}}^{n\hat{n}\tilde{n}} w_{\hat{b}l}^{b\hat{b}} w_{\tilde{c}l}^{c\tilde{c}}. \tag{39}$$

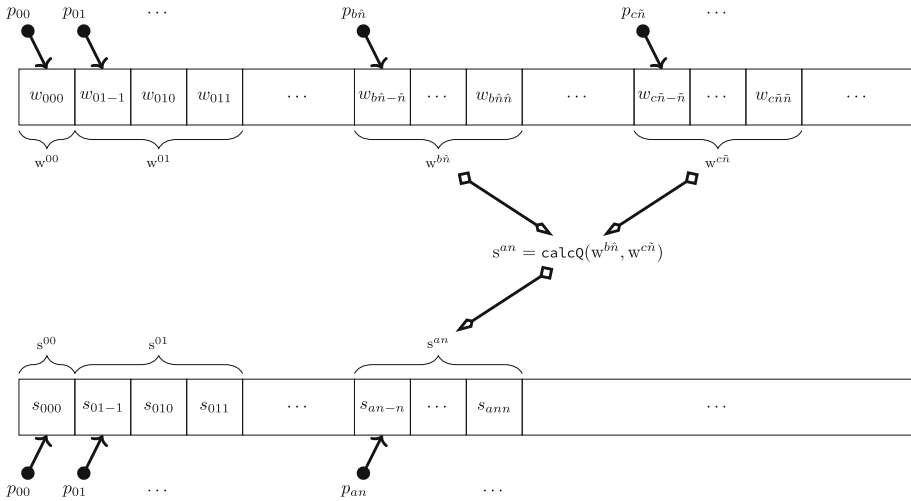
Here the anisotropic indices are only relevant to indicate the degree of the harmonic subspaces and hence the length of the involved vectors. Note, that we changed the positions of the indices and font of the coefficients of the distribution function and collision vector to indicate that this is an operation on a small subsection of the overall vectors.

Apart from the constant factor determined by the impact tensor, this operation is independent of the isotropic indices  $(a, b, c)$ , so we can give it as a general operation between three generic vectors  $s^n, u^{\hat{n}}, v^{\tilde{n}}$  of lengths  $2n + 1, 2\hat{n} + 1, 2\tilde{n} + 1$ , respectively:

$$s_l^n += \beta \cdot \sum_{\tilde{l}} Q_{l\tilde{l}}^{n\hat{n}\tilde{n}} u_{\tilde{l}}^{\hat{n}} v_{\tilde{l}}^{\tilde{n}}. \tag{40}$$

We obtain Eq. (39) by replacing  $\beta$  with the corresponding value of  $S_{n\hat{n}\tilde{n}}^{abc}$ .

Since  $Q_{l\tilde{l}}^{n\hat{n}\tilde{n}}$  are constants, the above operation (40) can be hard-coded without loss of generality of the overall implementation. We do so by giving each triple of anisotropic indices  $(n, \hat{n}, \tilde{n})$  a unique case number  $q$  and saving the expansion coefficients of both the



**Fig. 2** Schematic showing the operation of `calcQ` on the specifically structured arrays of the distribution function  $f$  determined by the weights  $w$  and the collision operator determined by the weights  $s$

distribution function  $w_{anl}$  and the collision vector  $s_{anl}$  in a specific order, so that we separate it into coherent  $2n + 1$  long sections, or sub-vectors,  $w^{an}$ . We also order the  $w_{anl}$  such that the directional index  $l$  starts at  $-n$  and ends at  $n$ , see Fig. 2. This reflects the splitting of the polynomial space into its irreducible subspaces: Each section  $w^{an}$  contains the coordinates of  $w$  in one particular irreducible subspace characterized by the anisotropic and isotropic indices  $a$  and  $n$ . The position of each section  $w^{an}$  is given by a corresponding pointer  $p_{an}$ , which locates the first of the  $2n + 1$  entries. The collision vector is ordered similarly.

Furthermore, we know that the following symmetry relations must hold:

$$S_{\hat{n}\hat{n}\hat{n}}^{abc} = S_{\hat{n}\hat{n}\hat{n}}^{acb} \qquad Q_{\hat{l}\hat{l}\hat{l}}^{n\hat{n}\hat{n}} = Q_{\hat{l}\hat{l}\hat{l}}^{n\hat{n}\hat{n}} \qquad (41)$$

Therefore, we used Mathematica to generate the C++ function `calcQ` that implements Eq. (40) for all combinations of anisotropic indices  $(n, \hat{n}, \tilde{n})$  with  $\hat{n} \leq \tilde{n}$ . Input for this method is  $\beta$  as double `beta`,  $q$  as integer `caseNumber` and the three vectors  $\xi^n, \zeta^{\hat{n}}, \zeta^{\tilde{n}}$  as three double pointers `r, a, b`. The method switches between all relevant triplets of anisotropic indices  $(n, \hat{n}, \tilde{n})$ , where  $q$  indicates the triplet and is thus used in the case distinction. For each case C++-code corresponding to the operation in (40) was generated. Here the three pointers are used as starting positions for the three vectors.

Note, that this code does not check, whether the three vectors actually have been allocated, all input is assumed to be valid in order for the computation to be as time-efficient as possible. Furthermore, it is always assumed that  $\hat{n} \leq \tilde{n}$ , the input vectors should be ordered accordingly (Fig. 3).

### 5.1.2 Implementing the Impact Tensor

In contrast to  $Q$ , the coefficients  $S_{\hat{n}\hat{n}\hat{n}}^{abc}$  can vary depending on the potential. We therefore give the impact tensor in a text file as input for the code. Each line of this text file is an input list containing the necessary information for `calcQ`, compare to Eq. (40) and see Fig. 4:

$$S_{\hat{n}\hat{n}\hat{n}}^{abc} \qquad q \qquad P_{an} \qquad P_{b\hat{n}} \qquad P_{c\tilde{n}}$$

```

1 void calcQ(int caseNumber, double* u, double* v, double* s, double beta){
2     switch(caseNumber){
3         case 1: // n=0  n1=0  n2=0
4             s[0]+= beta*(u[0]*v[0]);
5             break;
6         case 2: // n=0  n1=1  n2=1
7             s[0]+= beta*(u[0]*v[0] + u[1]*v[1] + u[2]*v[2]);
8             break;
9         case 3: // n=1  n1=0  n2=1
10            s[0]+= beta*(u[0]*v[0]);
11            s[1]+= beta*(u[0]*v[1]);
12            s[2]+= beta*(u[0]*v[2]);
13            break;
14            case 4: // n=0  n1=2  n2=2
15            s[0]+= beta*(u[0]*v[0] + u[1]*v[1] + u[2]*v[2]
16                + u[3]*v[3] + u[4]*v[4]);
17            break;
18            case 5: // n=1  n1=1  n2=2
19            s[0]+= beta*(-(u[2]*v[0] + u[1]*v[1]
20                + u[0]*(-0.5773502691896258*v[2] - v[4]));
21            s[1]+= beta*(u[0]*v[1] + 1.1547005383792517*u[1]*v[2]
22                + u[2]*v[3]);
23            s[2]+= beta*(-(u[0]*v[0] + u[1]*v[3]
24                + u[2]*(-0.5773502691896258*v[2] + v[4]));
25            break;
26            /*
27            .
28            .
29            .
30            up to
31            */
32            case 1331: // n=20  n1=20  n2=20
33            /*...*/
34        }
35    }

```

**Fig. 3** Extract from the C++ code for the function `calcQ`

1	-1.3703473031702487	1	2	0	2	// n=0 a=2 n1=0 b=0 n2=0 c=2
2	1.25095021593786	1	2	1	1	// n=0 a=2 n1=0 b=1 n2=0 c=1
3	-0.25019004318757193	4	2	41	41	// n=0 a=2 n1=2 b=0 n2=2 c=0
4	-2.055520954755372	1	3	0	3	// n=0 a=3 n1=0 b=0 n2=0 c=3
5	1.8127990869882546	1	3	1	2	// n=0 a=3 n1=0 b=1 n2=0 c=2
6	-0.4333418663487272	4	3	41	46	// n=0 a=3 n1=2 b=0 n2=2 c=1
7	-2.5244764598556895	1	4	0	4	// n=0 a=4 n1=0 b=0 n2=0 c=4
8	1.624509522645696	1	4	1	3	// n=0 a=4 n1=0 b=1 n2=0 c=3
9	0.5427032405253922	1	4	2	2	// n=0 a=4 n1=0 b=2 n2=0 c=2
10	-0.41944655513081003	4	4	41	51	// n=0 a=4 n1=2 b=0 n2=2 c=2
11	-0.1550580687215395	4	4	46	46	// n=0 a=4 n1=2 b=1 n2=2 c=1
12	0.025843011453588298	16	4	154	154	// n=0 a=4 n1=4 b=0 n2=4 c=0
13	-2.8853228917135803	1	5	0	5	// n=0 a=5 n1=0 b=0 n2=0 c=5
14	1.5450521141320113	1	5	1	4	// n=0 a=5 n1=0 b=1 n2=0 c=4
15	0.8782822629108616	1	5	2	3	// n=0 a=5 n1=0 b=2 n2=0 c=3
16	-0.41666956913861447	4	5	41	56	// n=0 a=5 n1=2 b=0 n2=2 c=3
17	-0.2710437907589509	4	5	46	51	// n=0 a=5 n1=2 b=1 n2=2 c=2
18	0.05778673035354533	16	5	154	163	// n=0 a=5 n1=4 b=0 n2=4 c=1
19	-3.1808091899516264	1	6	0	6	// n=0 a=6 n1=0 b=0 n2=0 c=6
20	1.5066904007148096	1	6	1	5	// n=0 a=6 n1=0 b=1 n2=0 c=5

**Fig. 4** Extract from the impact tensor file, each line containing all necessary information for `calcQ`. These are the first 20 lines corresponding to the Maxwellian potential and  $M = 20$

**Table 2** Overview of the computational requirements to calculate the collision operator with our implementation of the irreducible Burnett ansatz depending on the computational setups

Potential	Theory	Size of impact tensor in		Calculation time per time step
		Number of lines	Memory	
General	$M = 20$	188,176	12.2 MB	200 ms
IPL $\eta = 10$	$M = 20$	186,247	12.7 MB	200 ms
Maxwell	$M = 20$	4522	303 kB	5 ms
Maxwell	$M = 20, n = 0$	34	2kb	$\leq 2 \mu\text{s}$

The average timings are given for one timestep, i.e. they contain 4 evaluations of the collision operator. They have been calculated by measuring the calculation time of 1000 timesteps with a generic input distribution. Notice, that we are able to provide an upper bound for the computational requirements with the general potential that has non-zero entries for every impact tensor coefficient. Note, that the coupling tensor is stored as compiled C++ code and requires for  $M = 20$  roughly 38 MB

To calculate the collision vector, i.e. to compute (38), the C++-code iterates through each line of the impact tensor file and applies `calcQ` with the case  $q, \beta = S_{n\hat{n}\tilde{n}}^{abc}$  and the three vectors determined by  $p_{an}, p_{b\hat{n}}, p_{c\tilde{n}}$ . Utilizing (41), we can shorten the input list by stating twice the value of  $S_{n\hat{n}\tilde{n}}^{abc}$  when  $(b, \hat{n}) \neq (c, \tilde{n})$ , leaving the line with  $\hat{n} < \tilde{n}$  and omitting the corresponding line of  $S_{n\hat{n}\tilde{n}}^{acb}$ .

### 5.1.3 Discussion of the Implementation

Effectively, we have now implemented the map  $\Phi_{abc}^{n\hat{n}\tilde{n}}$  from Eq. (30): Structuring the array of moments  $w_{anl}$  into subsections corresponding to the irreducible subspaces and using the positions given by the pointers  $p_{b\hat{n}}$  and  $p_{c\tilde{n}}$  as starting positions for the operation `calcQ` corresponds to  $\phi_{(\hat{n},b)} \circ \mathcal{P}_{(\hat{n},b)}$  and  $\phi_{(\tilde{n},c)} \circ \mathcal{P}_{(\tilde{n},c)}$ . `calcQ` itself is the implementation of the map  $Q^{n\hat{n}\tilde{n}}$ , only that it operates on the section of the vector  $s_{anl}$  indicated by the pointer  $p_{an}$ , implementing  $\mathcal{E}_{(n,a)} \circ \phi_{(n,a)}^{-1}$ .

This setup is very flexible. It is easily adapted to different potentials, theories and initial conditions, i.e. the initial occupancy of  $w_{anl}$  of the distribution function. Omitting lines belonging to irrelevant combinations of  $(a, b, c, n, \hat{n}, \tilde{n})$  for the given setup, we can build calculation-specific input files for the collision operator. In Table 2 we have listed the required memory space for the impact tensor and the average calculation time of one time step for a selection of different potentials and theories. All computation is done sequentially on a CPU model Intel(R) Core(TM) i7-8650U. We note that the hard power potential with  $\eta = 10$  is very close to the general potential in terms of computational requirements. In Table 3 we have listed the number of coupling coefficients for different  $M$  and the corresponding sizes of the compiled program. We note, that the memory requirement for the compiled program is much larger than that for the text file of the impact tensor.

## 5.2 Comparison to an Algorithm for General Tensor-Vector Multiplications and Current State of the Art

Given that the ultimate goal of this work is to produce an implementation of the evaluation of  $S(f, f)$  within a numerical code, that is as fast as possible, the question is, whether the above described implementation is better than an algorithm developed specifically for the computation of tensor-vector multiplications. The efficient implementation of tensor

**Table 3** Overview of the computational requirements for the coupling tensor depending on the computational setups

Theory	Size of coupling tensor in	
	Number of cases	Memory
$M = 20$	1331	38 MB
Maxwell $M = 20$	505	6.6 MB
$M = 15$	612	9.9 MB
$M = 10$	216	1.6 MB
$M = 5$	42	122 MB
$M = 0$	1	2 kB

The memory requirements of the coupling tensor are given by the corresponding sizes of the compiled C++ code

operations is a field of on-going research in high performance computing and currently there are many different options available.

To check how the algorithm we developed compares to other algorithms, we chose the Tensor Algebra Compiler (TACO) [21] as a case study, as it provides code generation in C++ and seems to perform similarly well to other libraries. Here we found that TACO performed best when given the full collision tensor  $C_{\alpha\beta\gamma}$  with every triple of indices  $(a, n, l)$  mapped to one counter index. We found the computational times comparable but slightly slower than our implementation. We also compare memory requirements for the full tensor  $C_{\alpha\beta\gamma}$  with the irreducible Burnett approach and found that for a general potential, the full tensor  $C_{\alpha\beta\gamma}$  is much more demanding than what is needed for the impact tensor and `calcQ` together. However, when measuring the memory requirements for the Maxwell potential, we initially found that the memory requirements for the full tensor  $C_{\alpha\beta\gamma}$  are less than the requirements for `calcQ`. This is because saving `calcQ` as compiled program is more memory intensive than a file with only numbers, and because we still compared to the general coupling tensor which has many more cases than actually necessary—for  $M = 20$  out of the 1331 possible cases only 505 are actually required for the Maxwell potential, leading to a significantly smaller program of only 6.6MB. This reduction is due to the additional condition on the indices (36).

Assessing the performance of the irreducible Burnett algorithm would be incomplete without looking at the performance of the Burnett spectral method in [10] obtained from a single thread computation on a CPU model Intel Xeon E5-2697A V4. In Table 4 we can compare the three algorithms for the Maxwell potential. Additionally, for the hard Maxwell potential with  $\eta = 10$  and  $M = M_0 = 20$ , we find from Figure 15 in [10], that one evaluation of the collision operator takes about 90ms, whereas our algorithm takes about 50ms. Note however, that a different computational setup with different CPU models was used for each computation. We therefore conclude, that the computational time of both methods is similar for the IPL potential with  $\eta = 10$ , while the memory cost is greatly reduced with the irreducible Burnett method, which requires about 51MB, whereas the spectral Burnett method needs 784MB for just the sparse Maxwell potential. Table 4 also allows us to compare performance for the Maxwell potential of both methods and here we can see a significant difference both in computational time and memory usage.

Overall, we are satisfied to conclude that our implementation utilizes the underlying structure of the collision operator well enough to outperform the current state of the art.



**Table 4** Comparison between total runtimes of different algorithms for the Maxwell Potential (top), and a general potential (bottom)

	Algorithm	Runtime (s)	Memory
Maxwell potential	Irreducible Burnett	0.3	7 MB
	TACO [21]	0.6	25 MB
	Spectral Burnett [10]	39.76	784 MB
General potential	Irreducible Burnett	20	50 MB
	TACO [21]	31	1.4 GB

The time is given for  $M = 20$ , 100 iterations and initial function given by the BKW solution for the Spectral Burnett algorithm and a generic input distribution for the other algorithms. Note, that all algorithms use pre-computed coefficients for the collision operator and run sequentially. In these timings the irreducible Burnett algorithm has not been simplified to the homogeneous case, see Table 2 for these specs. The timings and memory usage for the spectral Burnett algorithm are taken from [10] with computation done on a single thread with CPU model Intel Xeon E5-2697A V4, while the other computations ran sequentially on a CPU model Intel(R) Core(TM) i7-8650U

### 5.3 Comparison to the Spectral Burnett Method in [10]

The spectral Burnett method used in the ansatz in [10] is based on complex solid harmonics. Here the presented algorithm exploits the relation for the directional indices—the collision tensor's entries can only be non-zero if  $l = \hat{l} + \tilde{l}$  holds. By contrast, the irreducible Burnett ansatz uses real solid harmonics, for which this condition does not apply. Instead, we find a sparse coupling tensor with band structures, since the real solid harmonics are a simple linear combination of their complex counterparts. Crucially, the main difference between the two approaches lies in the decomposition of the collision tensor into the impact and coupling tensor. The algorithm of the irreducible Burnett ansatz exploits both this decomposition and the sparsity of the coupling tensor. It is possible, that the performance can still be improved by applying the decomposition to the basis with complex spherical harmonics.

### 5.4 Comparison to the Spectral Fourier Approach

The irreducible and the spectral Burnett ansatz share a similar approach to discretize the velocity space through moments. This method is based on picking a finite polynomial degree and using a finite-dimensional approximation space for the velocity-dependency of the distribution function. In contrast, the spectral Fourier approach discretizes the velocity space by choosing some finite 3-dimensional velocity domain and representing that with a grid of equally distant points. The collision operator is a weighted convolution. [2] and [10] compared its computational costs to the presented algorithm and found comparable performance. Fixing the velocity domain introduces problems in the form of aliasing effects, choice of cut off and conservation properties. On the other hand, moment methods have difficulty in approximating some distributions, an example for this will be demonstrated in Sect. 6.3.

Notably, the spectral Fourier approach is connected to the representation theory of abelian groups, such as  $\mathcal{O}_2$  or the group of translations. In fact, the functions  $e^{2\pi i \theta_k}$  used for the expansion each span a one-dimensional irreducible subspace [22].

## 5.5 Using the Irreducible Burnett Ansatz

As we will see in the numerical application, not all initial conditions are computable in the Burnett basis. However, moment equations allow a relatively low number of numerical degrees of freedom to describe physically relevant situations. Moments as variables encode physical quantities, usually those of low order are the most interesting, allowing us to interpret a subset of the numbers representing the distribution function in a meaningful way.

Moment equations come with a level of complexity, and the irreducible Burnett ansatz is no exception. Its implementation is challenging compared to other numerical methods for the Boltzmann equation such as direct simulation monte carlo methods [23] or the aforementioned spectral Fourier approaches. This is due to the low level C++ code of the coupling tensor, which needs to be generated using scripts or similar for a high polynomial degree  $M$  due to the large number of different cases. We generated the function `calcQ` with Mathematica, and while the generating code can be checked, the overall code of `calcQ` cannot be verified by a human, only tested for plausibility in a small number of cases. However, this paper must be seen in the context of moment approximations for the Boltzmann equation, see e.g. [13, 24, 25]. In this context, via one way or the other, the collision operator will always be implemented as some matrix–vector product in the linear case, or a tensor–vector–vector product in the bilinear case. While the implementation of such a (bi-) linear operation can be verified by a human, for large degree  $M$  the matrix or tensor entries themselves similarly cannot be checked due to their number.

Solving just the spatially homogeneous Boltzmann equation is incomplete, the overall goal is to provide a method, with which the collision operator can be evaluated efficiently. However, the application of the irreducible Burnett ansatz to the inhomogeneous Boltzmann equation is out of the scope of this paper and we refer to [26] for further details.

## 6 Numerical Application

The irreducible Burnett ansatz is the real pendant to the spectral Burnett ansatz, additionally mathematically decomposing the collision tensor. While the spectral Burnett ansatz in [10] uses complex solid harmonics, the distribution function is real for their method as well. Therefore we expect its properties to be inherited from the spectral Burnett ansatz, including the rate of convergence. All differences between results of both methods must stem from purely numerical effects resulting from computational accuracy. This section aims to ensure the correct implementation of the method. To do so, we will roughly estimate spectral convergence by comparing lower orders of approximation to the numerical solution with  $M = 20$  for all cases and times. Additionally we compare to the analytical solution where possible.

Let  $f$  denote the exact distribution,  $\hat{f}$  the numerical solution with degree  $M = 20$  and moments  $\hat{w}$ ,  $\tilde{f}$  any other numerical solution with smaller degree and moments  $\tilde{w}$ ,  $\Delta f := \hat{f} - \tilde{f}$  and their difference. Using the  $L^2$ -norm implied by the scalar product on  $V_{\text{eq}}$  from Eq. (10),

$$\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int_{\mathbb{R}^3} f^2 \frac{1}{f^{(\text{eq})}} \, d\mathbf{c}},$$

$\|\Delta f\|$  is simply given by the standard vector norm of the difference between their moments,

$$\|\Delta f\| = \|\hat{w} - \tilde{w}\|.$$

Unlike the exact solution  $f$ , which we can only access in the first test case,  $\Delta f$  is available for all times. It is also a good indicator for spectral convergence, as we find for the residual

$$\|f - \tilde{f}\| = \|f - \hat{f} - (\hat{f} - \tilde{f})\| \leq \|f - \hat{f}\| + \|\Delta f\|.$$

Note that for  $t = 0$ ,  $\hat{f}$  is the result of the orthogonal projection of  $f$  and therefore  $f - \hat{f}$  is in the orthogonal complement of  $V_{eq}$ , implying that  $\Delta f \perp (f - \hat{f})$ . Therefore, at  $t = 0$

$$\|f - \tilde{f}\|^2 = \|f - \hat{f}\|^2 + \|\Delta f\|^2$$

holds, so when  $\|\Delta f\|$  decreases, so does  $\|f - \tilde{f}\|$ . We will calculate  $\|f - \hat{f}\|$  at  $t = 0$  to find the difference's order of magnitude, but show plots of  $\|\Delta f\|$  for both  $t = 0$  and  $t \neq 0$ , as we are mostly interested in the evolution of the coefficients calculated by the irreducible Burnett ansatz.

We consider three initial distributions to demonstrate the range of numerical application of our method:

- BKW-Solution, where the initial function can be visualized as a 3-dimensional hollow sphere.
- Bi-Gaussian distribution, where the initial function consists of two overlapping Gaussians.
- Discontinuous distribution, where the initial function consists of two different states that are separated at a flat plane in velocity space and that haven't interacted with each other yet.

We will look both at the effect of the choice of potential, as well as the difference between various numerical approximations, mainly given by the choice of different maximal polynomial degrees  $M$ . The equilibrium solution to the Boltzmann equation, the Gaussian distribution, lies in the chosen approximation space  $V_{eq}$  for any  $M$ . With our basis the Gaussian distribution corresponds to the state  $w_{000} = 1$  and all other moments equal to zero. Note, that all presented initial distributions evolve towards the Gaussian distribution with time for all presented potentials. Therefore, for large times we expect no difference between any two solutions, regardless of their initial distribution. This also implies that the residual between any two solutions decreases with increasing time. As a result, the focus should be on early time scales, because differences show up in early times and decrease with large time.

Additionally, conservation of mass, momentum and energy imply constraints on the first coefficients, making them irrelevant for further consideration:

$$\forall t \in \mathbb{R}^+ : \quad w_{000} = 1 \quad w_{01-1} = 0 \quad w_{010} = 0 \quad w_{011} = 0 \quad w_{100} = 0. \quad (42)$$

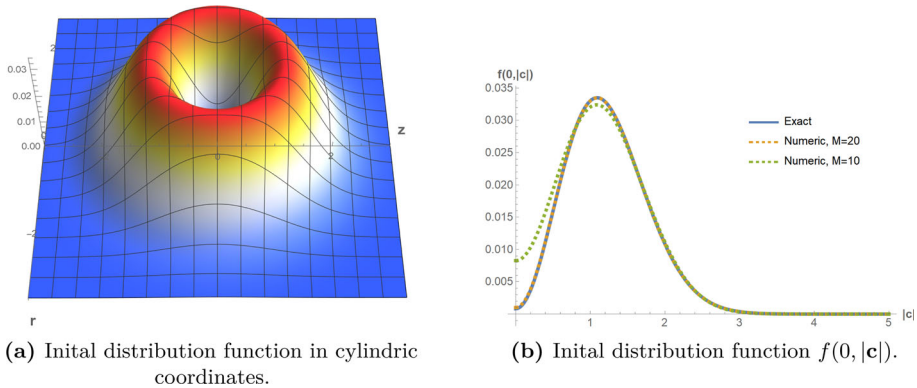
### 6.1 BKW Solution

The BKW solution is one of the few analytical solutions to the homogeneous Boltzmann Equation with Maxwell-potential and has been originally developed by Krupp (1967); Bobylev et al. (1976) [27]. As done in [9], we use the dimensionless formulation from Ernst [27]:

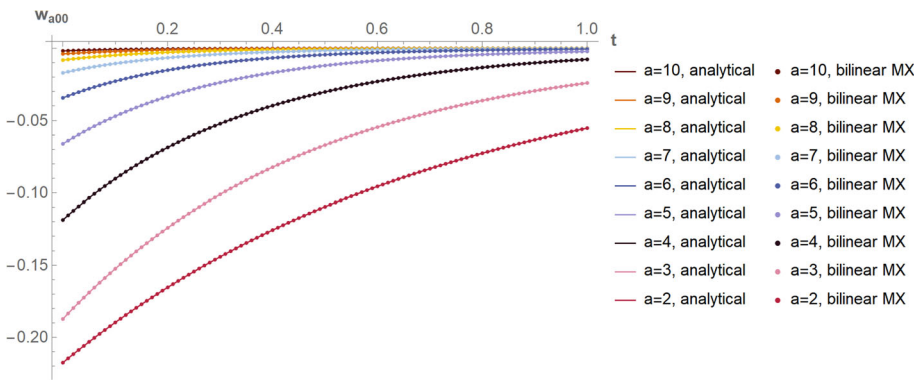
$$f(t, \mathbf{c}) = (2\pi \tau(t))^{-\frac{3}{2}} \exp\left(-\frac{|\mathbf{c}|^2}{2\tau(t)}\right) \left[1 + \frac{1 - \tau(t)}{\tau(t)} \left(\frac{|\mathbf{c}|^2}{2\tau(t)} - \frac{3}{2}\right)\right], \quad (43)$$

where

$$\tau(t) = 1 - \exp\left(\frac{\pi}{3} Bt - T_0\right) \quad \text{with} \quad B = -0.6542926411560946 \quad T_0 = 0.92 \quad (44)$$



**Fig. 5** Graphs of the initial distribution of the BKW solution



**Fig. 6** BKW solution: Analytical (solid lines) vs. numerical (dots) evolution of coefficients  $w_{a00}$ . Numerical coefficients were calculated with the bilinear Maxwell potential

The initial distribution  $f(0, \mathbf{c})$  can be envisioned as a 3D hollow sphere. In particular, it is obvious, that this distribution is rotationally invariant, allowing us to plot it in cylindrical coordinates without losing information. Since our basis is adapted to  $\mathcal{O}_3$ , we can reduce the degrees of freedom to only those coefficients that correspond to rotationally invariant basis functions, i.e. only weights  $w_{anl}$  with  $(a, n, l) = (a, 0, 0)$  can be non-zero (Fig. 5).

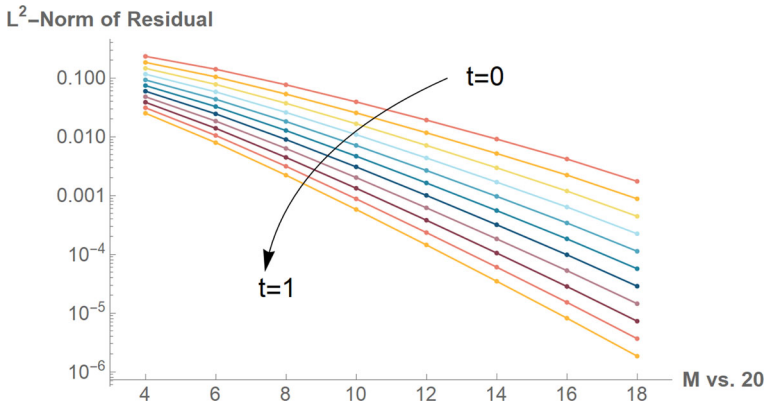
Choosing a maximal degree of  $M = 20$ , we only have 11 relevant coefficients:

$$a \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \quad \text{and} \quad n = l = 0 :$$

$$\begin{matrix} w_{000} & w_{100} & w_{200} & w_{300} & w_{400} & w_{500} \\ w_{600} & w_{700} & w_{800} & w_{900} & w_{1000} & \end{matrix}$$

Combined with the restraints from Eq. (42), this allows us to take a look at the remaining nine time-dependent moments, all of higher order with  $w_{200}$  the lowest one, to judge how well the analytical and numerical solutions match without losing any information.

Given that Cai et al. observed very good correspondence in [9, 10], and that the approximations differ in basis choice at most, we expect to obtain similar results. Looking at Fig. 6, we observe that this expectation is met. We can also see spectral convergence of our method



**Fig. 7** Spectral Convergence of the solution for the BKW distribution. The y-Axis shows  $\|\Delta f\|$ , i.e. the  $L^2$ -Norm given by the weighted scalar product of the difference between the projection of the analytical solution onto the approximation space with  $M = 20$  and the numerical solution with different maximal polynomial degrees  $M$ , the x-Axis shows this degree. The plots are shown for times between 0 (top graph) and 1 (bottom graph) in steps of 0.1 (in order from top to bottom) and the evolution is calculated with the Maxwell potential

in Fig. 7 by comparing results of lower polynomial degrees

$$M \in \{ 4, 6, 8, 10, 12, 14, 16, 18 \}$$

to the projection of the analytical solution onto the approximation space with  $M = 20$  for different times  $t \in [0, 1]$ . We observe faster than exponential convergence, which is demonstrated by the downward sloping curves on the semi-log scale. However, there is no reason to expect that this behaviour translates from the very special BKW test case to the approximation of a general distribution.

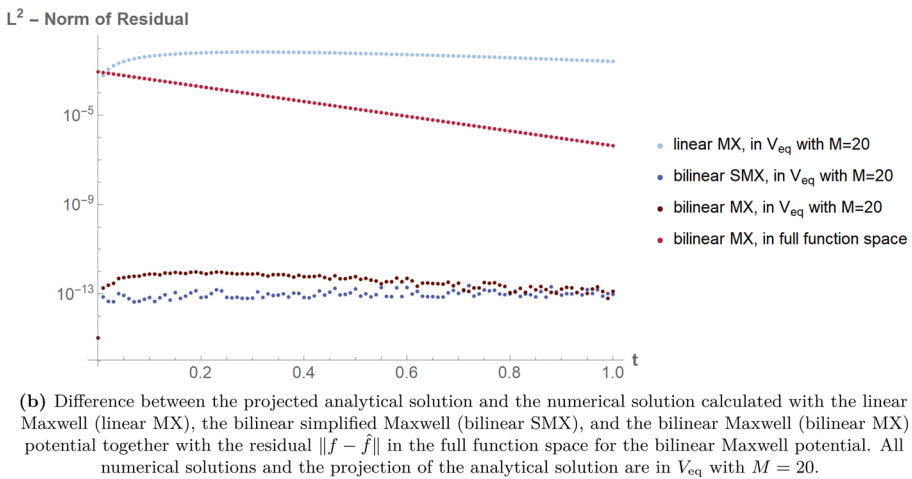
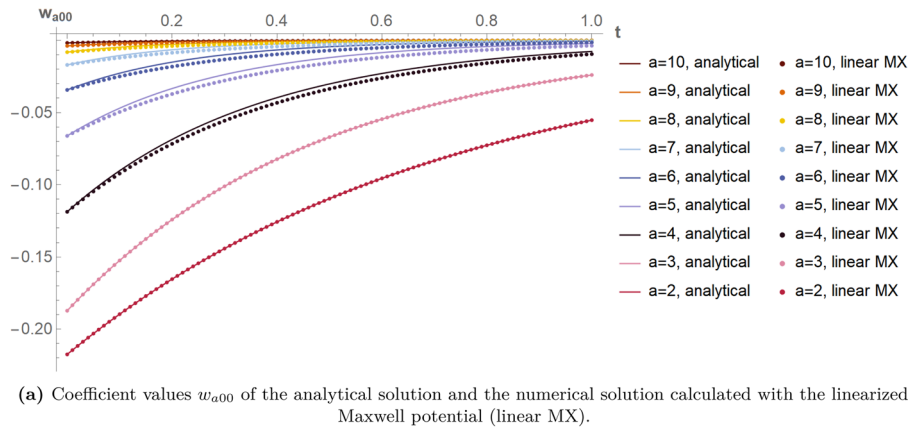
Note, that for various  $M$  the difference between the evolution of the lowest order time-dependent moment  $w_{200}$ , which all compared theories share, is in the order of  $10^{-14}$ . This is due to the restriction in Eq. (36) imposed by the Maxwell potential, which for this particular case simplifies to

$$a = b + c.$$

With  $a = 2$ , the only possible contributions can come from  $(b, c) \in \{ (0, 2), (1, 1) \}$ , which are already taken into account by the lowest order ansatz with  $M = 4$ . A corresponding argument applies to any higher moment—in this specific test case the increase of polynomial degree cannot have an influence on the lower moments, but it does add accuracy to the overall distribution, as shown in Fig. 7.

While the difference in basis choice and understanding of the underlying representation theory seems benign, in this particular case it means we have reduced the roughly 1700 perceived degrees of freedom to just 9. This also affects the number of relevant impact coefficients: For a rotationally invariant distribution and  $M = 20$  we only need 34 impact coefficients  $S_{nnn}^{abc}$  and hence the evaluation time for one time step is only a few ms.

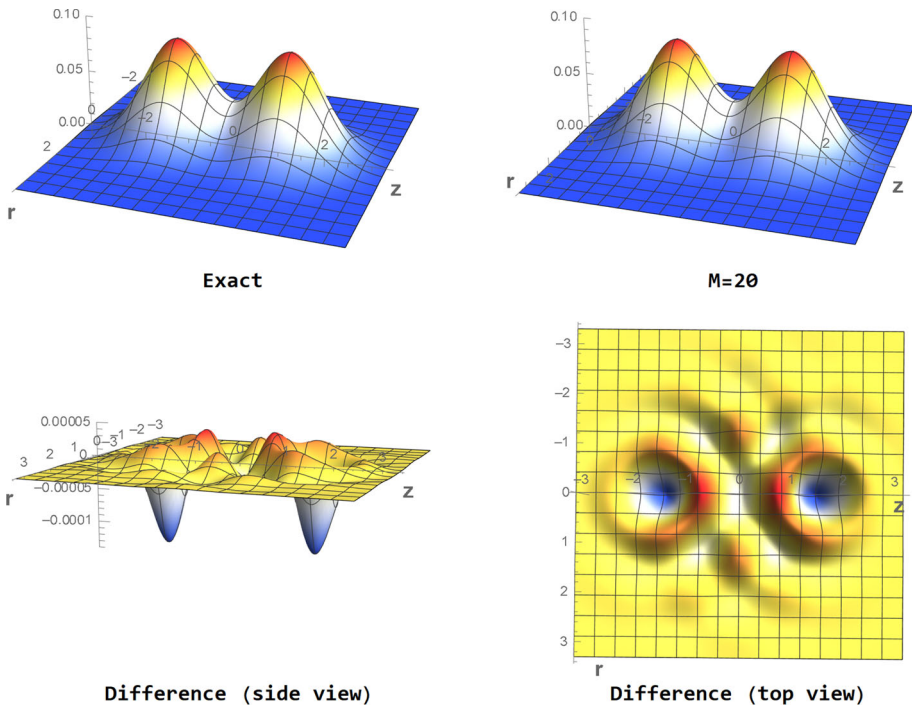
Figure 8 compares the evolution of the moments for different potentials. By omitting any line in the impact tensor file where  $b \neq 0$ , we generated the linear collision operator. In Fig. 8a we can clearly see a difference between the linear Maxwell potential and the analytical solution, in particular in the higher moments. We also calculated the evolution of the coefficients for the simplified Maxwell model and found that it corresponds well to the



**Fig. 8** BKW solution: Analytical (solid lines) vs. numerical (dots) evolution of coefficients for different potentials

analytical solution. Figure 8b compares the residual between the moments obtained from the numerical solution with  $M = 20$  and the moments given by projecting the analytical solution onto the approximation space  $V_{eq}$  with  $M = 20$  for the three different potentials. We can see that both the Maxwell and the simplified Maxwell potential produce very accurate moments with error in the order of  $10^{-13}$ , while the linear potential differs significantly from the analytical solution. The plot also shows  $\|f - \hat{f}\|$ , i.e. the difference between the analytical and numerical solution obtained with the Maxwell potential for the ansatz with  $M = 20$  in the full function space equipped with the same scalar product as  $V_{eq}$ . The residual starts at  $8.9e-4$  for  $t = 0$  and we can see it steadily decreasing with time, as the solutions converge to the same Gaussian.

We conclude, that for rotationally invariant cases, the simplified Maxwell model produces similar results as the exact Maxwell model, and linearising the collision operator introduces



**Fig. 9** Bi-Gaussian distribution: Exact and with  $M = 20$  approximated initial distribution function together with two views of their difference in cylindric coordinates

a noticeable error. We also see even in this simple example the limit of spectral Burnett ansatz—moments can be accurately calculated, but the actual distribution function can only be approximated with some error depending on the degree  $M$ .

We also calculated the evolution of the coefficients for the simplified Maxwellian model and found that it corresponds to the analytical solution as well. The linearised simplified Maxwell model differs to the analytical solution similarly to the linearised Maxwell model. We conclude, that for rotationally invariant cases, the simplified Maxwell model is as good as the exact Maxwell model, and linearising the collision operator introduces a small but noticeable error (Fig. 9).

### 6.2 Bi-Gaussian Distribution

Again we take the initial distribution from [9]

$$f(0, \mathbf{c}) = \frac{1}{2\pi^{3/2}} \left( e^{-c_x^2 - c_y^2 - \left(c_z - \sqrt{\frac{3}{2}}\right)^2} + e^{-c_x^2 - c_y^2 - \left(c_z + \sqrt{\frac{3}{2}}\right)^2} \right). \tag{45}$$

Due to the distribution’s mirror symmetry w.r.t the  $x, y$ -plane, only even anisotropic indices  $n$  can result in non-zero projection coefficients in our basis. This is reflected in the projection of a monomial. Using the computer algebra system *Mathematica*, the projection of a monomial leads to a product from which it is easy to find a corresponding condition—any uneven degree

leads to 0:

$$\langle c_x^a c_y^b c_z^c, f(0, \mathbf{c}) \rangle = ((-1)^a + 1) \left( (-1)^b + 1 \right) \left( (-1)^c + 1 \right) K_1(a, b, c),$$

with some additional function  $K_1(a, b, c)$ . We further see, that the distribution is  $\mathcal{O}_2$ -invariant for rotations around the  $z$ -axis, resulting in the additional constraint that unless the directional index  $l = 0$ , the coefficients are zero. Overall, with  $M = 20$  we find 66 relevant coefficients from combinations of

isotropic index  $a \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ ,

anisotropic index  $n \in \{0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20\}$ , with  $2a + n \leq M$  and  $l = 0$ .

We note, that with all the coefficients of uneven anisotropic indices being zero in the beginning and due to the conditions (32), we again may drop many lines from the impact tensor file, albeit not as many as in the BKW case. Figure 11 shows the residual between the ansatz with  $M = 20$  and lower order approximations with the maximal polynomial degrees  $M \in \{10, 12, 14, 16, 18\}$ .

Here we again see spectral convergence of the solution, indicating that this test case can also be well approximated by the spectral Burnett basis. Figure 10 indicates that the smallest and largest orders  $M = 10$  and  $M = 20$  both have similar looking distributions, with their largest difference located at the tips and overlap of the two Gaussians. We can further see from both Figs. 10 and 11 that the difference between approximations decreases with time as discussed in the beginning of this chapter.  $\|f - \hat{f}\|$  at  $t = 0$  evaluates to  $1.6e-3$ , the same order of magnitude as the first case. Overall, this indicates that the initial distribution is well approximated.

Lastly, we compare the simplified and actual Maxwell potential in Fig. 12, where we look at the evolution of 19 out of the 66 coefficients as examples to demonstrate the range of possibilities. We note that for some moments the Maxwell and simplified Maxwell potential correspond to each other well, while for other moments those calculated with the Maxwell potential converge faster to 0. This trend is observed for most of the not displayed coefficients as well. However, we also note that some coefficients grow first, before converging to zero. Here we see, that the coefficients of the Maxwell potential remain closer to zero than those of the simplified Maxwell potential. Overall, we conclude that the simplified Maxwell potential can lead to a significantly different solution than the Maxwell potential in setups that are not rotationally invariant.

### 6.3 Discontinuous Distribution

Using the dimensionless formulation from [9] for this case we have

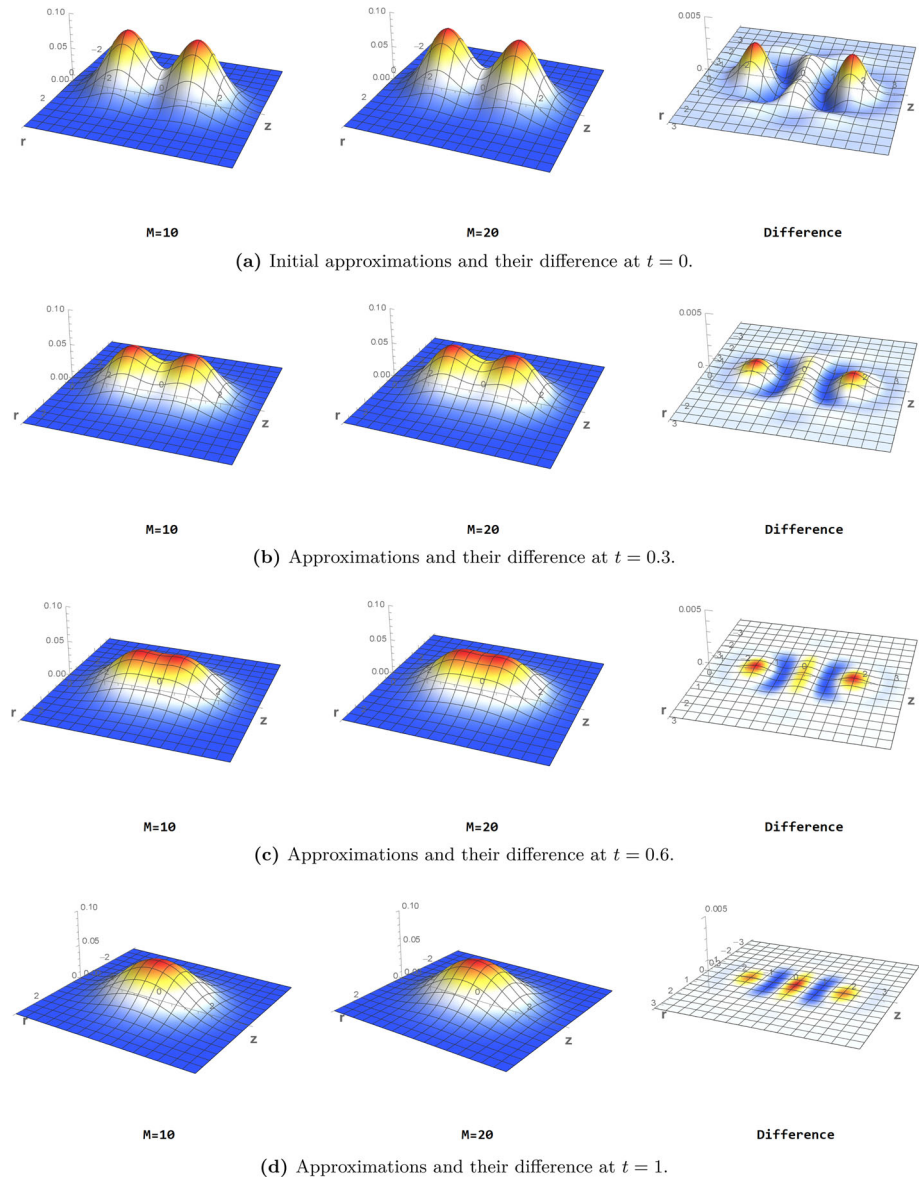
$$f(0, \mathbf{c}) = \begin{cases} \frac{\sqrt[4]{2}(2-\sqrt{2})}{\pi^{3/2}} \cdot \exp\left(-\frac{c^2}{\sqrt{2}}\right) & c_z > 0 \\ \frac{\sqrt[4]{2}(2-\sqrt{2})}{4\pi^{3/2}} \cdot \exp\left(-\frac{c^2}{2\sqrt{2}}\right) & \text{else} \end{cases} \quad (46)$$

Projection of a monomial, evaluated by *Mathematica*, again leads to a product from which we can infer that certain index combinations lead to zero:

$$\langle c_x^a c_y^b c_z^c, f(0, \mathbf{c}) \rangle = ((-1)^a + 1) \left( (-1)^b + 1 \right) \left( (-1)^c 2^{\frac{1}{2}(a+b+c)} + \sqrt{2} \right) K_2(a, b, c),$$

again with some additional function  $K_2(a, b, c)$  However, finding exclusion rules for the projection coefficients is less obvious here. They can only be non-zero if  $n$  is uneven or

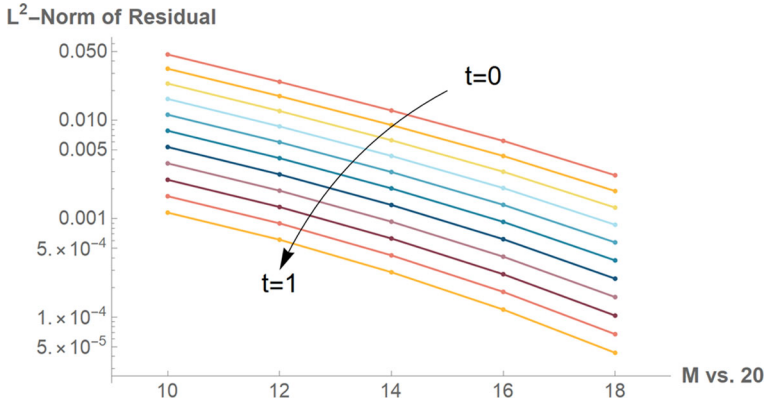




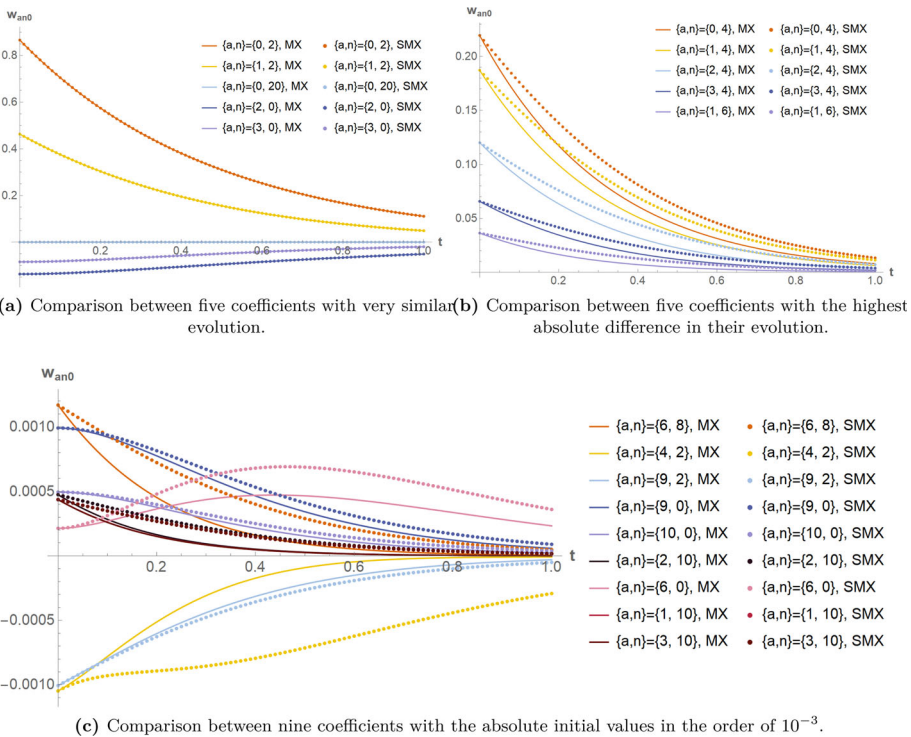
**Fig. 10** Bi-Gaussian distribution: Snapshots of the evolution of the approximation with  $M = 10$ ,  $M = 20$  and their difference in cylindrical coordinates for  $t \in \{0, 0.3, 0.6, 1\}$  calculated with the Maxwell potential

zero. We again find that the distribution function is  $\mathcal{O}_2$ -symmetric, requiring the projection coefficients to be zero unless  $l = 0$ . Note in contrast to the bi-Gaussian case, that even though at the beginning the relevant coefficients have an uneven or zero anisotropic index, two uneven anisotropic indices can couple to an even anisotropic index, see (32). So the total set of relevant coefficients  $w_{anl}$  for  $M = 20$  is given by:

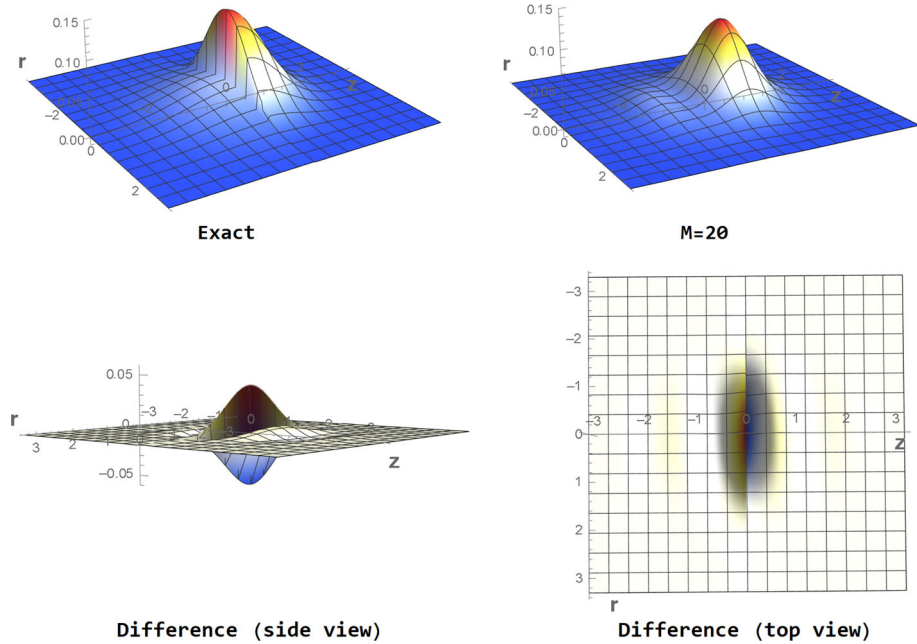
$$\text{isotropic index } a \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\},$$



**Fig. 11** Spectral convergence of the solution for the bi-Gaussian distribution. The y-Axis shows  $\|\Delta f\|$ , i.e. the  $L^2$ -Norm given by the weighted scalar product of the difference between the two solutions of different  $M$  and  $M = 20$  as indicated by the x-Axis. The plots are shown for times between 0 (top graph) and 1 (bottom graph) in steps of 0.1 (in order from top to bottom) and the evolution is calculated with the Maxwell potential



**Fig. 12** Bi-Gaussian distribution: Comparison of the evolution of the coefficients  $w_{an0}$  of the bi-Gaussian distribution function for the Maxwell (MX, solid lines) and simplified Maxwell (SMX, dots) potential



**Fig. 13** Discontinuous distribution: Exact and with  $M = 20$  approximated initial distribution function together with two views of their difference in cylindrical coordinates

anisotropic index  $n \in \{x | x \in \mathbb{N}_0 \wedge 0 \leq x \leq M\}$ , with  $2a + n \leq M$  and  $l = 0$ .

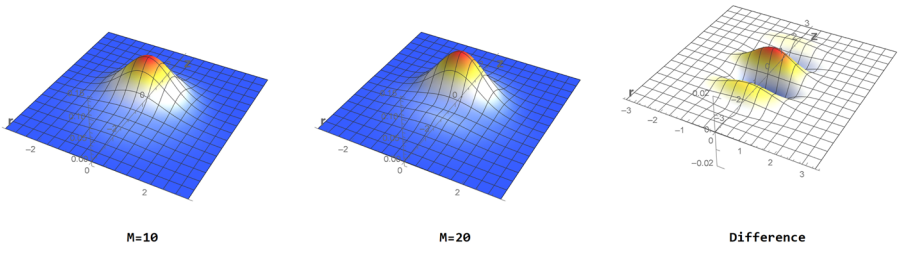
i.e. one coefficient per irreducible subspace, which makes 121 coefficients in total for  $M = 20$ . However, because all combinations of isotropic and anisotropic indices are possible, this case requires the full impact tensor and basis set, which has 1771 moments. Due to the  $\mathcal{O}_2$ -symmetry, most of them are zero and will always be zero for this test case. Looking at Fig. 13, we can clearly see that the discontinuity at  $t = 0$  is not well approximated, even with a large degree of  $M = 20$ . This is reflected in  $\|f - \hat{f}\|$ , which at  $t = 0$  evaluates to  $1.7e-1$ , a much larger value than in the other two cases.

Comparing the evolution of the distribution function in Fig. 14 for approximations of degree  $M = 10$  and  $M = 20$ , we again find that the two approximations get closer to each other with time, as both converge to the same Gaussian.

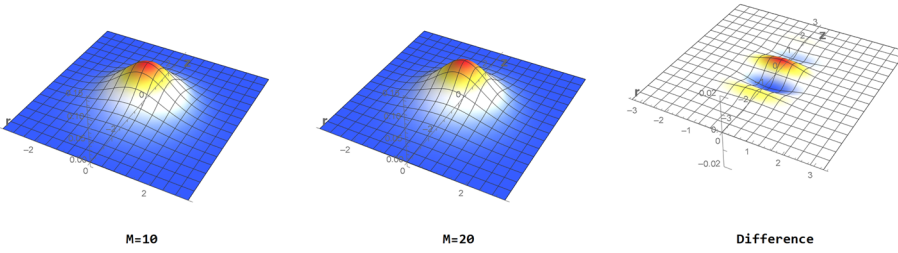
Figure 15 shows the residual between the ansatz with  $M = 20$  and lower order approximations, each denoted by their maximal polynomial degree

$$M \in \{10, 11, 12, 13, 14, 15, 16, 17, 18\}.$$

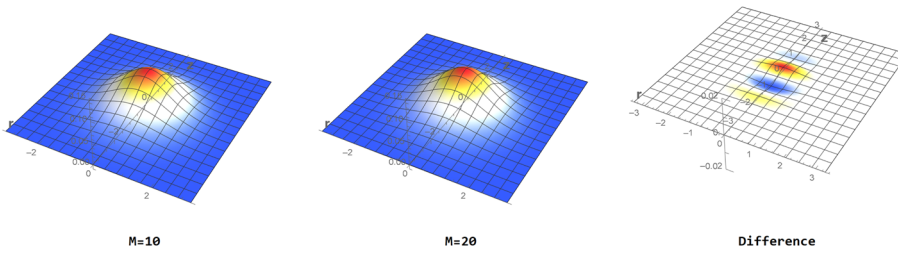
Here we see a rather slow spectral convergence and that there is little gain of accuracy when increasing  $M$  from an uneven to an even degree, which is also why  $M = 19$  was omitted. We propose that this is due to the initial condition, in which most of the coefficients with even  $2a + n$  are zero. Together with the badly approximated discontinuity shown in Fig. 13, this shows one of the limits of the approximation space. This is the first test case where all lines of the impact tensor file are contributing to the overall evolution of the distribution function and as with the previous example we can see clear differences in the evolution of coefficients between the simplified Maxwell and Maxwell potential in Fig. 16.



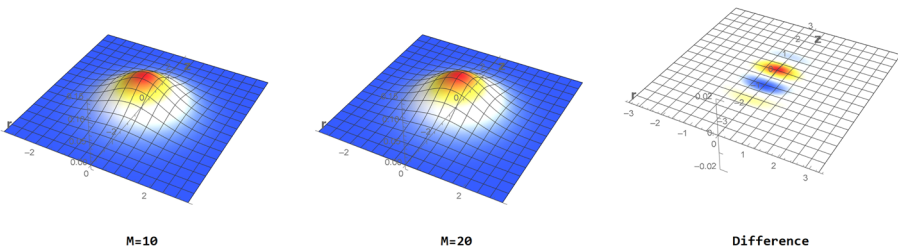
(a) Discontinuous distribution: Initial approximations and their difference at  $t = 0$ .



(b) Discontinuous distribution: Approximations and their difference at  $t = 0.3$ .

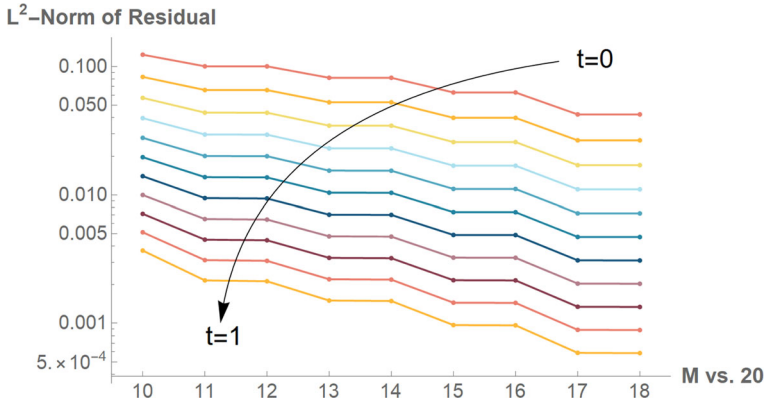


(c) Discontinuous distribution: Approximations and their difference at  $t = 0.6$ .

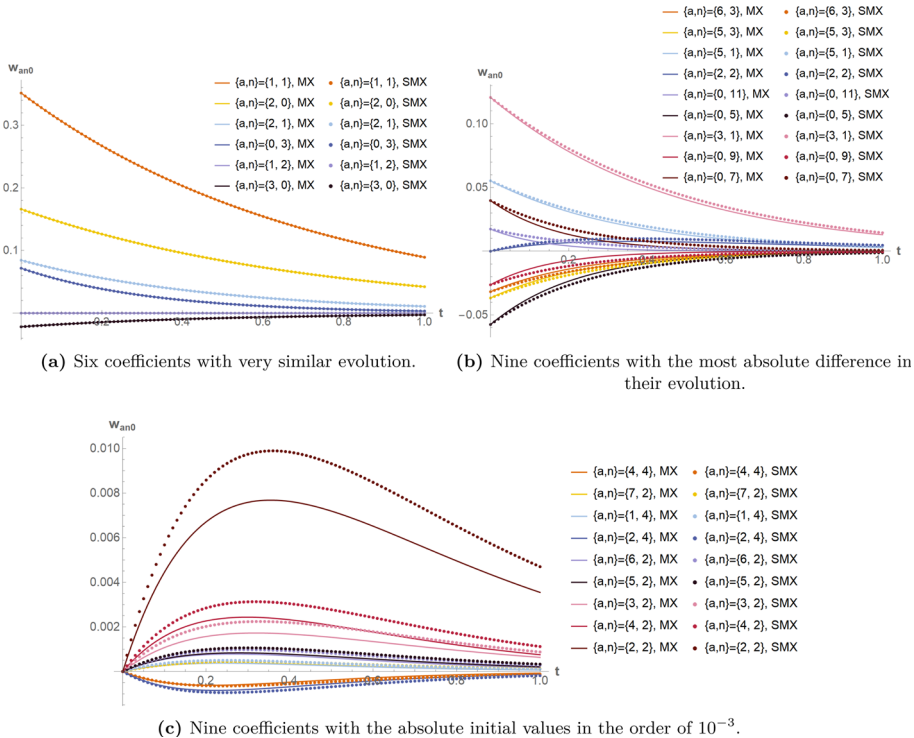


(d) Discontinuous distribution: Approximations and their difference at  $t = 1$ .

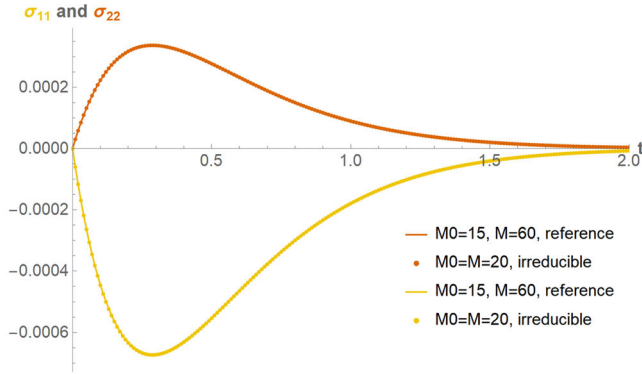
**Fig. 14** Snapshots of the evolution of the approximation of the discontinuous distribution function with  $M = 10$ ,  $M = 20$  and their difference in cylindrical coordinates for  $t \in \{0, 0.3, 0.6, 1\}$  calculated with the Maxwell potential



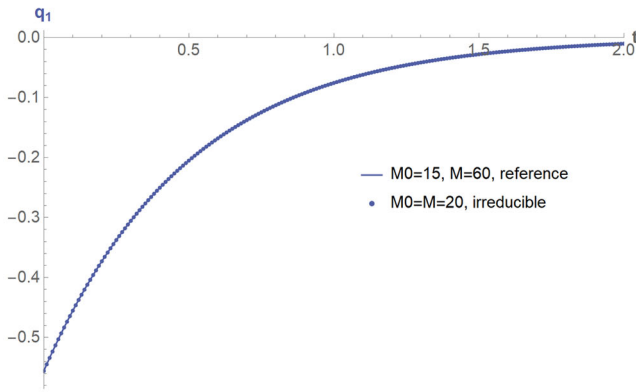
**Fig. 15** Convergence of the solution for the discontinuous distribution. The y-Axis shows  $\|\Delta f\|$ , i.e. the  $L^2$ -Norm given by the weighted scalar product of the difference between the two solutions of different  $M = 20$  and  $M$  as indicated by the x-Axis. The plots are shown for times between 0 (top graph) and 1 (bottom graph) in steps of 0.1 (in order from top to bottom) and the evolution is calculated with the Maxwell potential



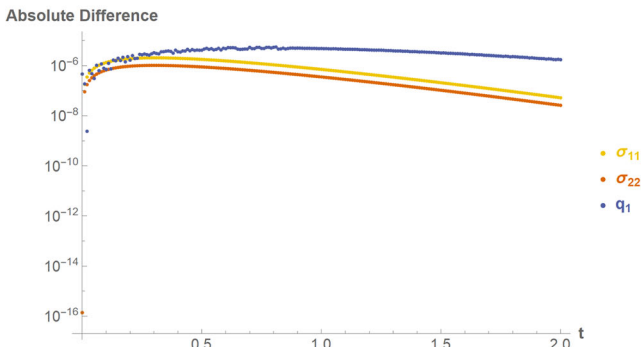
**Fig. 16** Discontinuous distribution: Comparison of the evolution of the coefficients  $w_{an0}$  for the Maxwell (MX, solid lines) and simplified Maxwell (SMX, dots) potential



(a) Stress tensor entries  $\sigma_{11}$  and  $\sigma_{22}$ .



(b) Heat flux entry  $q_1$ .



(c) Absolute difference between the values from [9] and from the irreducible Burnett ansatz. Note that the approximation space differs between the two – [9] use  $M_0 = 15$ ,  $M = 60$ , while our ansatz uses  $M_0 = M = 20$ .

**Fig. 17** Discontinuous distribution: Evolution of the stress tensor entries  $\sigma_{11}$  and  $\sigma_{22}$  and the heat flux entry  $q_1$  for the hard IPL potential with  $\eta = 10$ . The results of the irreducible Burnett ansatz (irreducible, dots) are compared to those from [9] (reference, solid lines), with values kindly provided by the authors. Note, that we transformed our coordinates to match the orientation of the reference values for this plot

When looking at the heat flux and stress tensor, we find good correspondence to previously published results in [9] in Fig. 17, with an absolute difference in the order of  $10^{-6}$ . Note, that the approximation space for this comparison is not the same. Referring to the method demonstrated in [9], our ansatz corresponds to  $M_0 = M = 20$ , while the values provided by the authors have been calculated with  $M_0 = 15, M = 60$ . The good correspondence suggests that this difference has little influence on the moments of this low order and that the irreducible Burnett ansatz produces results very similar to the spectral Burnett ansatz. This should be the case, since the difference between the spectral and the irreducible Burnett ansatz lies solely in the exploitation of the collision tensor's decomposition.

## 7 Summary

With the irreducible Burnett ansatz we were able to reduce the computational effort of calculating Boltzmann's collision operator in the spectral Burnett ansatz both in computational time and memory by implementing an algorithm inspired by the representation theoretical decomposition of the collision operator. This decomposition relies on the fact that the used basis is adapted to the irreducible subspaces of the approximation space.

This decomposition not only allowed us to find an efficient algorithm, but also provides an understanding behind the structure of the collision tensor by separating factors that depend on the underlying physics from those that are of pure mathematical nature: The bilinear collision tensor can be expressed as product of the impact tensor with coefficients depending on the potential between the colliding particles and the coupling tensor with coefficients only depending on the basis choice of the irreducible subspaces.

With spherical harmonics as basis, further structure of the coupling tensor could be exploited by looking at  $\mathcal{O}_2$ -irreducible subspaces. For example, the coupling tensor could be split into the case where all anisotropic indices are zero and everything else. This is not necessarily generally applicable, but there may exist interesting use-cases.

Applying the method of this paper to collision operators of different kind, such as polyatomic gases [28] or the Landau operator [29], looks like a promising next step.

**Acknowledgements** This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)-333849990/GRK2379 (IRTG Modern Inverse Problems). The authors have no relevant financial or non-financial interests to disclose. All authors contributed to the conception and design of the presented algorithm. Material preparation, data collection and analysis were mainly performed by Andrea Hanke. The first draft of the manuscript was written by Andrea Hanke and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. Code and other data files concerning this paper are available in the repository [20].

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Data availability** Enquiries about data availability should be directed to the authors.

## Declarations

**Conflict of interest** The authors have not disclosed any competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is

not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## 9 Appendix

### 9.1 Rotational Covariance of the Collision Operator

Repeating the quadratic collision operator  $S(f)$  given in Eq. (3)

$$S(f) = \int_{\mathbb{R}^3} \int_{\mathbf{n} \perp \mathbf{g}} \int_0^\pi (f(t, \mathbf{c}') f(t, \mathbf{c}_1') - f(t, \mathbf{c}) f(t, \mathbf{c}_1)) B(|\mathbf{g}|, \chi) d\chi \, d\mathbf{n} \, d\mathbf{c}_1 \quad (47)$$

we see it is a function in  $\mathbf{c}$ . Any orthogonal coordinate transformation  $r \in \mathcal{O}_3$  acts on any function  $T : \mathbb{R}^3 \rightarrow \mathbb{R}$  by

$$r^{-1}(T(\mathbf{c})) := T(r(\mathbf{c})),$$

which applied to the collision operator is

$$r^{-1}(S(f)(\mathbf{c})) = S(f)(r(\mathbf{c})). \quad (48)$$

Additionally, as a physically motivated map between function spaces, the collision operator is rotationally covariant, i.e. as stated in Eq. (18) it fulfils

$$\forall r \in \mathcal{O}_3 : r^{-1}(S(f)(\mathbf{c})) = S(r^{-1}(f))(\mathbf{c}). \quad (49)$$

This means that rotating or mirroring the input distribution function  $f$  will result in a correspondingly rotated/mirrored collision operator. To see that (49) is true, we will check that

$$S(r^{-1}(f))(\mathbf{c}) = S(f)(r(\mathbf{c})) \quad (50)$$

holds. Since  $r$  is a  $3 \times 3$  orthogonal matrix,  $r(\mathbf{c}) = r \cdot \mathbf{c}$ . Noting that  $r^{-1}(f)(\mathbf{c}) = f(r \cdot \mathbf{c})$ , we find for the left hand side

$$S(r^{-1}(f))(\mathbf{c}) = \int_{\mathbb{R}^3} \int_{\mathbf{n} \perp \mathbf{g}} \int_0^\pi (f(r \cdot \mathbf{c}') f(r \cdot \mathbf{c}_1') - f(r \cdot \mathbf{c}) f(r \cdot \mathbf{c}_1)) B(|\mathbf{g}|, \chi) d\chi \, d\mathbf{n} \, d\mathbf{c}_1, \quad (51)$$

omitting the dependency on time  $t$  for the remainder of this calculation. To obtain the right hand side, we just need to substitute  $\mathbf{c}$  with  $r \cdot \mathbf{c}$ . We will make use of the variable change

$$\tilde{\mathbf{c}}_1 := r \cdot \mathbf{c}_1 \quad d\tilde{\mathbf{c}}_1 = d\mathbf{c}_1 \quad \mathbf{c}_1 = r^{-1}\tilde{\mathbf{c}}_1, \quad (52)$$

which transforms the second pair of distributions in Eq. (51) on the left hand side. Comparing to the same pair of the right hand side we obtain:

$$\text{LHS: } f(r \cdot \mathbf{c}) f(r \cdot \mathbf{c}_1) = f(r \cdot \mathbf{c}) f(\tilde{\mathbf{c}}_1), \quad \text{RHS: } f(r \cdot \mathbf{c}) f(\mathbf{c}_1). \quad (53)$$

Both sides are already very similar—the only difference is the tildation of  $\mathbf{c}_1$ . For the first pair in Eq. (51) we look at the velocities  $r \cdot \mathbf{c}'$  and  $r \cdot \mathbf{c}_1'$  individually. From (4) we obtain

$$r \cdot \mathbf{c}' = r \cdot \left( \mathbf{h} + \frac{1}{2} R_{\mathbf{n}, \chi}(\mathbf{g}) \right)$$



$$\begin{aligned}
 &= \frac{1}{2}r \cdot (\mathbf{c} + \mathbf{c}_1 + R_{\mathbf{n},\chi}(\mathbf{c}) - R_{\mathbf{n},\chi}(\mathbf{c}_1)) \\
 &= \frac{1}{2}r \cdot \mathbf{c} + \frac{1}{2}\tilde{\mathbf{c}}_1 + \underbrace{\frac{1}{2}rR_{\mathbf{n},\chi}r^{-1}r\mathbf{c}}_{=R_{r(\mathbf{n}),\chi}} - \underbrace{\frac{1}{2}rR_{\mathbf{n},\chi}r^{-1}\tilde{\mathbf{c}}_1}_{=R_{r(\mathbf{n}),\chi}}, \tag{54}
 \end{aligned}$$

and similarly for  $r \cdot \mathbf{c}_1$ :

$$r \cdot \mathbf{c}_1 = \frac{1}{2}r \cdot \mathbf{c} + \frac{1}{2}\tilde{\mathbf{c}}_1 - \underbrace{\frac{1}{2}rR_{\mathbf{n},\chi}r^{-1}r\mathbf{c}}_{=R_{r(\mathbf{n}),\chi}} + \underbrace{\frac{1}{2}rR_{\mathbf{n},\chi}r^{-1}\tilde{\mathbf{c}}_1}_{=R_{r(\mathbf{n}),\chi}}. \tag{55}$$

We further note, that  $\mathbf{g}$  and  $\mathbf{n}$  are affected by the variable change:

$$\mathbf{g} = \mathbf{c} - \mathbf{c}_1 = \mathbf{c} - r^{-1}\tilde{\mathbf{c}}_1 = r^{-1}(r\mathbf{c} - \tilde{\mathbf{c}}_1) \implies \begin{cases} |\mathbf{g}| = |r\mathbf{c} - \tilde{\mathbf{c}}_1|, \\ \mathbf{n} \perp \mathbf{g} = \mathbf{n} \perp r^{-1}(r\mathbf{c} - \tilde{\mathbf{c}}_1). \end{cases} \tag{56}$$

Note how the transformation  $r$  acts on the normal vector  $\mathbf{n}$ : The matrix  $R$  now rotates around the axis  $r(\mathbf{n})$ , instead of the axis  $\mathbf{n}$ , while the integration domain is transformed by  $r^{-1}$ . Therefore, we can equivalently integrate over  $\mathbf{n} \perp r\mathbf{c} - \tilde{\mathbf{c}}_1$  and substitute all  $\mathbf{n}$  by  $r^{-1}\mathbf{n}$ . Now we are left with an expression for  $S(r^{-1}(f))$  that has  $r(\mathbf{c})$  instead of  $\mathbf{c}$ , and  $\tilde{\mathbf{c}}_1$  instead of  $\mathbf{c}_1$  everywhere, which is equivalent to the right hand side of (50). This shows covariance of the collision operator.

Note the core reason for this: The collision operator is a concatenation of rotationally covariant operations such as taking the norm of a vector, or integrating over all directions.

### 9.2 Representation Theory: An Overview

Below we describe the core aspects of representation theory that we employed in this paper. This is a use-case restricted overview of the subject and we refer the reader to lecture books, e.g. [30–35], and to the original material such as [17, 36] for further reading and full proofs of the following statements.

**Definition 9.1** (*Representation (Definition 4.1 from [33])*) Let  $G$  be a group and  $F$  a field. A representation of  $G$  over  $F$  consists of

- a  $F$ -vector space  $V$ ,
- a map  $\cdot : G \times V \rightarrow V$ ,  $(g, v) \mapsto g \cdot v$ , called the *operation* of the group on the vector space, that satisfies the axioms in Table 5.

We can say  $V$  is a representation of  $G$  and  $G$  operates on  $V$ .

For our purposes, we are interested in representations of the orthogonal groups, mainly  $\mathcal{O}_3$  given by orthogonal matrices  $r \in \mathbb{R}^{3 \times 3}$ , i.e.  $G = \mathcal{O}_3$ . The field  $F$  is  $\mathbb{R}$  and the vector spaces  $V$  we use are the tensor space of  $\mathbb{R}^3$  and the space of real polynomials in three arguments  $\mathbb{R}[x, y, z]$ . For both the operation of  $\mathcal{O}_3$  is given in terms of the matrix vector-product  $r \cdot v$  with  $v, v_i \in \mathbb{R}^3$ :

$$r \cdot v_1 \otimes \cdots \otimes v_n = (r \cdot v_1) \otimes \cdots \otimes (r \cdot v_n), \tag{57}$$

$$r \cdot p(v) = p(r^{-1} \cdot v). \tag{58}$$

**Table 5** Defining properties of representations and homomorphisms between representations, with  $G$  a group,  $V, U$  two representations of  $G$ ,  $f : U \rightarrow V$  a homomorphism

Axioms of representations	Meaning
Linearity	$\forall g \in G : (V \rightarrow V, v \mapsto g \cdot v)$ is a linear map
Associativity	$\forall g, h \in G \forall v \in V : g \cdot (h \cdot v) = (g \cdot h) \cdot v$
Non-triviality or normalization	$\forall v \in V : 1_G \cdot v = v$
Axioms of homomorphisms	Meaning
$F$ -linearity	$f$ is a linear map $U \rightarrow V$
$G$ -linearity	$\forall g \in G, u \in U : f(g \cdot u) = g \cdot f(u)$

One can generalize this to the tensor product between two vector spaces  $V, U$  each a representation of  $\mathcal{O}_3$ . Then  $V \otimes U$  is also a representation of  $\mathcal{O}_3$  with the action of  $r \in \mathcal{O}_3$  on  $v \otimes u \in V \otimes U, v \in V, u \in U$ , defined according to chapter 1.1 in [37] as

$$r \cdot v \otimes u = (r \cdot v) \otimes (r \cdot u). \tag{59}$$

We are concerned with  $\mathbb{R}[x, y, z] \otimes \mathbb{R}[x, y, z]$ .

**Definition 9.2** (*Homomorphism between two representations (Definition 4.3 from [33])*) Let  $U, T$  be two representations of  $G$  and  $f : U \rightarrow V$  a homomorphism, i.e. a linear map.  $f$  is called  $FG$ -linear map, *intertwining map* or *natural map*, if the two axioms in Table 5 are satisfied. The space of all  $FG$ -linear maps is written as  $\text{Hom}_{FG}(U, V)$ . If the two vector spaces are identical, i.e.  $f : V \rightarrow V$ , the  $FG$ -linear map is called *endomorphism* and the space of such maps is denoted with  $\text{End}_{FG}(V)$ .

If  $f$  is invertible, i.e. there exists a map  $f' : V \rightarrow U$  with  $f \circ f' = f' \circ f = \text{id}$ , then  $f$  is called an *isomorphism* of the representations. We call two representations *isomorphic* if such an  $f$  exists.

For two representations to be isomorphic, their vector spaces also need to be isomorphic. If the dimensions of two vector spaces are not equal, then they cannot be isomorphic representations.

Since the collision operator describes a physical process, one of its fundamental properties is, that it does not depend on the choice of coordinates. We also note, that it is a bilinear map between polynomial spaces,  $C : P^M \times P^M \rightarrow P^M$ , which can be written as a linear map between the tensor product of two polynomial spaces to a polynomial space,  $P^M \otimes P^M \rightarrow P^M$ . In mathematical terms, the collision operator must be a homomorphism from  $\text{Hom}_{\mathbb{R}\mathcal{O}_3}(P^M \otimes P^M, P^M)$ , i.e. a natural map. While this already restricts the degrees of freedom when compared to general linear map  $P^M \otimes P^M \rightarrow P^M$ , we can go further by using the concept of irreducible subspaces.

**Definition 9.3** (*Invariant subspaces (Definition 4.2 from [33])*) A subspace  $U$  of any vector space  $V$  is called *invariant* with respect to the operation of a group  $G$ , or  $G$ -invariant, iff for all  $g \in G$  and for all  $u \in U$ , the operation of  $g$  on  $u$  still is in  $U, g \cdot u \in U$  or equivalently:  $\forall g \in G : g(U) = U$ .

**Definition 9.4** (*Irreducible subspaces (Definition 4.2 from [33])*) A subspace  $U \neq 0$  of any vector space  $V$  is called *irreducible* with respect to the operation of a group  $G$ , or  $G$ -irreducible, iff it is invariant and there is no invariant subspace  $W \subset U$  with the properties

$W \neq U$  and  $W \neq 0$ . If a subspace  $U$  is invariant and  $\dim U = 1$ , then it is also irreducible, however an irreducible subspace can be of dimension larger than 1.

As mentioned in [17], the space of symmetric trace-free tensors of order  $n$ ,  $\text{STF}^n(\mathbb{R}^3)$ , is  $\mathcal{O}_3$ -irreducible. With the isomorphism between  $\text{STF}^n(\mathbb{R}^3)$  and  $\mathcal{H}^n$ , which is also explored in [17],  $\mathcal{H}^n$  is also irreducible.

**Remark 9.5** Two irreducible subspaces  $U_1, U_2$  that are not isomorphic to each other are orthogonal to each other with respect to any scalar product of  $V$  that is compatible with the operation of the group  $G$ , i.e. is itself a natural map:

$$\forall g \in G \forall v_1, v_2 \in V : \langle v_1, v_2 \rangle = \langle g \cdot v_1, g \cdot v_2 \rangle. \tag{60}$$

We can project from the full vector space  $V$  onto an irreducible subspace  $U_1$ . Due to the universal orthogonality, such a projection is always unique up to a constant (or zero) if the irreducible subspace  $U_1$  has a multiplicity of 1, i.e. there is no other irreducible subspace of the vector space isomorphic to  $U_1$ .

This universal orthogonality between non-isomorphic irreducible subspaces is a consequence of one of the fundamental lemmas in representation theory:

**Lemma 9.6** (Schur’s Lemma (Theorem 4.29 from [33])) *Let  $U_1$  and  $U_2$  be two finite-dimensional, irreducible subspaces of a representation. A natural map  $\phi : U_1 \rightarrow U_2$  is only non-zero, iff  $U_1$  and  $U_2$  are isomorphic to each other. All non-zero natural maps  $U_1 \rightarrow U_2$  are the same up to a constant if the field is  $\mathbb{C}$ .*

**Remark 9.7** For real representations, the space of natural maps  $U_1 \rightarrow U_2$  between two isomorphic irreducible subspaces is not necessarily one-dimensional. We shall see with Brauer’s classification of natural maps that for our case ( $F = \mathbb{R}, G = \mathcal{O}_3, V$  the tensor space of  $\mathbb{R}^3$  or  $\mathbb{R}[x, y, z]$ ), we only find one-dimensional spaces of such natural maps.

**Remark 9.8** Given a homomorphism  $h$  in  $\text{Hom}_{FG}(U, V)$  a list of all irreducible subspaces  $U_i$  of  $U$  and  $V_j$  of  $V$  and that for each pair of isomorphic subspaces ( $U_i, V_j$ ) the space of natural maps  $U_i \rightarrow V_j$  is one-dimensional, we can pick non-zero isomorphism  $\phi_{ij} : U_i \rightarrow V_j$ , to express  $h$  as sum over all isomorphism  $\phi_{ij}$  multiplied by free constants, which are the only degrees of freedoms for  $h$ .

$$h : U \rightarrow V = \sum_{ij} U_i \rightarrow V_j \tag{61}$$

$$h = \sum_{ij} c_{ij} \phi_{ij}. \tag{62}$$

This follows immediately from Schur’s Lemma. Note, that the sum does not necessarily include all combinations of  $(i, j)$ , but only those where the corresponding  $U_i$  and  $V_j$  are indeed isomorphic, since all other natural maps are zero. If there are multiple instances of the same irreducible subspace up to isomorphism within  $U$  (or  $V$ ), we can make the splitting unique by choosing some orthogonal decomposition with respect to the chosen scalar product for  $U$  (or  $V$ ).

Applying this idea to the setup in this paper, we first define the notation for the irreducible and orthogonal subspaces characterized with the isotropic and anisotropic indices:

$$\mathcal{H}_a^n := L_a^{\binom{n+\frac{1}{2}}{}} \left( \frac{c^2}{2} \right) \cdot \mathcal{H}^n. \tag{63}$$

Now we can decompose our bilinear collision operator as the sum over natural maps between irreducible subspaces,

$$\begin{aligned}
 C : P^M \otimes P^M \rightarrow P^M &= \left( \sum_{an} \mathcal{H}_a^n \right) \otimes \left( \sum_{an} \mathcal{H}_a^n \right) \rightarrow \left( \sum_{an} \mathcal{H}_a^n \right) \\
 &= \sum_{abcn\hat{n}\tilde{n}} \mathcal{H}_b^{\hat{n}} \otimes \mathcal{H}_c^{\tilde{n}} \rightarrow \mathcal{H}_a^n,
 \end{aligned} \tag{64}$$

and using the isomorphism between  $\mathcal{H}_a^n$  and  $\mathcal{H}^n$ , we can restrict the different types of maps to

$$\mathcal{H}^{\hat{n}} \otimes \mathcal{H}^{\tilde{n}} \rightarrow \mathcal{H}^n. \tag{65}$$

We now note that symmetric tensors are isomorphic to polynomials of the same degree and same underlying vectorspace, Chapter B.2 in [38], i.e. with  $\mathbb{R}^3$  we can map symmetric tensors of degree  $n$  from the space of symmetric tensors  $\text{Sym}^n(\mathbb{R}^3)$  to homogeneous polynomials of degree  $n$  in  $\mathbb{R}[x, y, z]$ . We identify the tensor space that is isomorphic to the space of harmonic polynomials  $\mathcal{H}^n$  of degree  $n$  as the space of symmetric trace-free tensors  $\text{STF}^n$  of same degree, as is explored in [17]. Anything that we learn from the tensorial view,

$$\text{STF}^{\hat{n}} \otimes \text{STF}^{\tilde{n}} \rightarrow \text{STF}^n, \tag{66}$$

we can apply to the polynomial view. We first note, that the tensor product of two representations is itself a representation. We also note, that the codomain of the natural map is an irreducible representation of  $\mathcal{O}_3$ , so using Schur’s lemma, we know that the natural maps in (66) can only be non-zero if the codomain is isomorphic to one of the irreducible subspaces of the domain. Finding the irreducible subspaces of the tensor product of two representations is not trivial in general and this decomposition is called Clebsch-Gordan theory [39]. For the case of  $\mathcal{O}_3$  and symmetric trace-free tensors over  $\mathbb{R}$ , one can find the decomposition of the maps in (66) relatively easily and we show the argument only for the sake of completeness. We start by considering the types of natural maps between tensors [36]:

- Permuting tensor factors,
- Taking the trace between two factors,
- Taking the tensor product with the equivalent of  $\mathbf{e}^2$ ,  $\omega := \sum_i b_i \otimes b_i$  for any orthonormal basis set  $\{b_1, b_2, b_3\}$  of  $\mathbb{R}^3$ , also known as the  $\delta_{ij}$ -tensor.

Any natural map between tensor spaces can be expressed as linear combination and composition of the above maps. They can be represented graphically with Brauer diagrams [36]. This gives us a search space for the isomorphism  $\text{STF}^{\hat{n}} \otimes \text{STF}^{\tilde{n}} \rightarrow \text{STF}^n$ . We first note, that the result must be a tracefree tensor, so we use the detracer  $\mathcal{D}$  from [17].  $\mathcal{D}$  is an explicitly given projection from a symmetric tensor onto the symmetric tracefree tensor space of same order. As its input is a symmetric tensor, we need the symmetrizer (Part B.2 in [38])

$$q : \mathbb{R}^{3 \otimes n} \rightarrow \text{Sym}^n \tag{67}$$

$$v_1 \otimes \dots \otimes v_n \mapsto \frac{1}{n!} \sum_{\sigma \in S_n} v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(n)}, \tag{68}$$

which gives us the overall projector  $\mathcal{D} \circ q$ , which is applied last. From Remark 9.5 and because the multiplicity of  $\text{STF}^n$  in  $\text{Sym}^n$  is one, we know that this map is unique up to a constant. What remains to be shown is for any natural linear map  $\phi$  that we apply before this projector such that the overall map  $\mathcal{D} \circ q \circ \phi$  is  $\text{STF}^{\hat{n}} \otimes \text{STF}^{\tilde{n}} \rightarrow \text{STF}^n$ , the overall maps

at most differ by a constant factor and thereby fulfilling the promise in Remark 9.7. In other words: We want to show that

$$\text{span} \left\{ \mathcal{D} \circ q \circ \phi : \text{STF}^{\hat{n}} \otimes \text{STF}^{\tilde{n}} \rightarrow \text{STF}^n \mid \phi \text{ is a natural linear map} \right\}$$

is one-dimensional or the space of maps to zero.

Next, we note that the overall tensorial degree of the domain is  $\hat{n} + \tilde{n}$ , and we would like to show the restriction  $n \leq \hat{n} + \tilde{n}$ . To do so, we find all natural maps  $\text{STF}^n \rightarrow (\mathbb{R}^3)^{\otimes(\hat{n}+\tilde{n})}$  and show that they must all be zero, leaving us with the conclusion that a representation isomorphic to  $\text{STF}^n$  cannot be contained in  $(\mathbb{R}^3)^{\otimes(\hat{n}+\tilde{n})}$ . Any natural map  $\text{STF}^n \rightarrow (\mathbb{R}^3)^{\otimes(\hat{n}+\tilde{n})}$  must reduce the tensorial degree  $n$  to  $\hat{n} + \tilde{n}$ , leaving the trace as the only candidate, but since the tensors in  $\text{STF}^n$  are trace-free by definition, the result is zero. First taking the tensor product, i.e. multiplying, with any number of  $\omega$  also does not yield a different result, since in the end more traces need to be taken than the number of  $\omega$  that have been multiplied, inadvertently leading to taking at least one trace of the tensor in  $\text{STF}^n$ . Therefore, for  $n > \hat{n} + \tilde{n}$  there does not exist a subspace in  $(\mathbb{R}^3)^{\otimes(\hat{n}+\tilde{n})}$  isomorphic to  $\text{STF}^n$ , leading us to the desired restriction.

We further note, that the maps listed above all change the degree of tensor product by an even number, so  $\tilde{n} + \hat{n} - n \in 2\mathbb{N}_0$ . It is clear that the traces can only be taken between one factor of  $\text{STF}^{\hat{n}}$  and one factor of  $\text{STF}^{\tilde{n}}$  each, as opposed to a trace of two factors in  $\text{STF}^{\hat{n}}$ , which would map to zero by definition of  $\text{STF}^{\hat{n}}$ . This gives the third condition on the anisotropic indices from (32),  $\tilde{n} + \hat{n} - n \leq 2 \min(\tilde{n}, \hat{n})$ . Finally, we note that the codomain is also tracefree, so if somewhere in the composition the tensor product with  $\omega$  is taken, an additional trace needs to be taken as well to achieve tracefreeness, otherwise the detracer would project onto zero due to the orthogonality between tracefree tensors and tensors with a trace (Part B, section 6 in [40]). Overall, the space of homomorphisms  $\text{STF}^{\hat{n}} \otimes \text{STF}^{\tilde{n}} \rightarrow \text{STF}^n$  is one-dimensional, so each map is unique up to a constant.

Following the isomorphism back to the polynomial spaces, we can apply the results one-to-one, leading to the unique up to a constant natural map given by the coupling tensor. Looking at (64):

$$C = \sum_{abc\hat{n}\tilde{n}} \mathcal{H}_b^{\hat{n}} \otimes \mathcal{H}_c^{\tilde{n}} \rightarrow \mathcal{H}_a^n, \tag{69}$$

and picking one non-zero instance for each of the maps that satisfy (32):

$$\tilde{Q}_{abc}^{\hat{n}\tilde{n}} : \mathcal{H}_b^{\hat{n}} \otimes \mathcal{H}_c^{\tilde{n}} \rightarrow \mathcal{H}_a^n, \tag{70}$$

we apply Remark 9.8 to write any natural map of the form (64) as

$$\sum_{abc\hat{n}\tilde{n}} S_{\hat{n}\tilde{n}}^{abc} \tilde{Q}_{abc}^{\hat{n}\tilde{n}}. \tag{71}$$

Noting that due to the isomorphism between the irreducible subspaces, all  $\tilde{Q}_{abc}^{\hat{n}\tilde{n}}$  with same combination of anisotropic indices work the same, we arrive at the decomposition of the collision tensor from Eq. (33).

## References

1. Pareschi, L., Russo, G.: Numerical solution of the Boltzmann equation I: spectrally accurate approximation of the collision operator. *SIAM J. Numer. Anal.* **37**(4), 1217–1245 (2000). <https://doi.org/10.1137/S0036142998343300>

2. Gamba, I.M., Tharkabhushanam, S.H.: Spectral-Lagrangian methods for collisional models of non-equilibrium statistical states. *J. Comput. Phys.* **228**(6), 2012–2036 (2009)
3. Pareschi, L., Rey, T.: Moment preserving Fourier–Galerkin spectral methods and application to the Boltzmann equation. *SIAM J. Numer. Anal.* (2021). <https://doi.org/10.48550/ARXIV.2105.13158>
4. Hu, J., Qi, K., Yang, T.: A new stability and convergence proof of the Fourier–Galerkin spectral method for the spatially homogeneous Boltzmann equation. *SIAM J. Numer. Anal.* **59**(2), 613–633 (2021). <https://doi.org/10.1137/20M1351813>
5. Grad, H.: On the kinetic theory of rarefied gases. *Commun. Pure Appl. Math.* **2**(4), 331–407 (1949). <https://doi.org/10.1002/cpa.3160020403>
6. Sod, G.A.: Numerical solution of Boltzmann’s equation (1976). <https://doi.org/10.2172/7184312>
7. Cai, Z., Torrilhon, M.: Approximation of the linearized Boltzmann collision operator for hard-sphere and inverse-power-law models. *J. Comput. Phys.* **295**, 617–643 (2015)
8. Cai, Z., Torrilhon, M.: Numerical simulation of microflows using moment methods with linearized collision operator. *J. Sci. Comput.* **74**(1), 336–374 (2018)
9. Wang, Y., Cai, Z.: Approximation of the Boltzmann collision operator based on hermite spectral method. *J. Comput. Phys.* **397**, 108815 (2019). <https://doi.org/10.1016/j.jcp.2019.07.014>
10. Cai, Z., Fan, Y., Wang, Y.: Burnett spectral method for the spatially homogeneous Boltzmann equation. *Comput. Fluids* **200**, 104456 (2020). <https://doi.org/10.1016/j.compfluid.2020.104456>
11. Kumar, K.: Polynomial expansions in kinetic theory of gases. *Ann. Phys.* **37**(1), 113–141 (1966). [https://doi.org/10.1016/0003-4916\(66\)90280-6](https://doi.org/10.1016/0003-4916(66)90280-6)
12. Gamba, I.M., Rjasanow, S.: Galerkin–Petrov approach for the Boltzmann equation. *J. Comput. Phys.* **366**, 341–365 (2018)
13. Struchtrup, H.: Grad’s moment method. In: Struchtrup, H. (ed.) *Macroscopic Transport Equations for Rarefied Gas Flows: Approximation Methods in Kinetic Theory*, pp. 87–107. Berlin, Springer (2005). [https://doi.org/10.1007/3-540-32386-4\\_6](https://doi.org/10.1007/3-540-32386-4_6)
14. Cercignani, C.: The Boltzmann equation. In: Cercignani, C. (ed.) *The Boltzmann Equation and Its Applications*, pp. 40–103. New York, Springer (1988). [https://doi.org/10.1007/978-1-4612-1039-9\\_2](https://doi.org/10.1007/978-1-4612-1039-9_2)
15. Kremer, G.M.: Basic principles of the kinetic theory. In: Kremer, G.M. (ed.) *An Introduction to the Boltzmann Equation and Transport Processes in Gases*, pp. 1–35. Springer, Berlin (2010). [https://doi.org/10.1007/978-3-642-11696-4\\_1](https://doi.org/10.1007/978-3-642-11696-4_1)
16. Cai, Z., Torrilhon, M.: On the Holway–Weiss debate: convergence of the Grad-moment-expansion in kinetic gas theory. *Phys. Fluids* **31**(12), 126105 (2019). <https://doi.org/10.1063/1.5127114>
17. Applequist, J.: Traceless cartesian tensor forms for spherical harmonic functions: new theorems and applications to electrostatics of dielectric media. *J. Phys. A Math. Gen.* **22**(20), 4303–4330 (1989). <https://doi.org/10.1088/0305-4470/22/20/011>
18. Burnett, D.: The distribution of molecular velocities and the mean motion in a nonuniform gas. *Proc. Lond. Math. Soc.* **2**(1), 382–435 (1936). <https://doi.org/10.1112/plms/s2-40.1.382>
19. Alex, A., et al.: A numerical algorithm for the explicit calculation of SU(N) and SL(N, C) Clebsch–Gordan coefficients. *J. Math. Phys.* **52**(2), 023507 (2011). <https://doi.org/10.1063/1.3521562>
20. Hanke, A., Torrilhon, M.: Representation theory based algorithm to compute Boltzmann’s bilinear collision operator in the irreducible spectral Burnett Ansatz efficiently. <https://doi.org/10.5281/zenodo.5848971>
21. Kjolstad, F., et al.: The Tensor Algebra Compiler. In: *Proceeding of the ACM Programming Languages 1.OOPSLA* (2017). <https://doi.org/10.1145/3133901>
22. Group representations. In: *Lie Groups: An Approach through Invariants and Representations*, pp. 195–239. Springer, New York, 2007. [https://doi.org/10.1007/978-0-387-28929-8\\_8](https://doi.org/10.1007/978-0-387-28929-8_8)
23. Babovsky, H., Illner, R.: A convergence proof for Nanbu’s simulation method for the full Boltzmann equation. *SIAM J. Numer. Anal.* **26**(1), 45–65 (1989). <https://doi.org/10.1137/0726004>
24. Müller, I., Ruggeri, T.: Extended thermodynamics of moments. In: Müller, I., Ruggeri, T. (eds.) *Rational Extended Thermodynamics*, pp. 197–220. Springer, New York (1998). [https://doi.org/10.1007/978-1-4612-2210-1\\_9](https://doi.org/10.1007/978-1-4612-2210-1_9)
25. Torrilhon, M., Sarna, N.: Hierarchical Boltzmann simulations and model error estimation. *J. Comput. Phys.* **342**, 66–84 (2017). <https://doi.org/10.1016/j.jcp.2017.04.041>
26. Bünger, J., et al.: Structured derivation of moment equations and stable boundary conditions with an introduction to symmetric, trace-free tensors. *Kinet. Relat. Models* **16**(3), 458–494 (2023). <https://doi.org/10.3934/krm.2022035>
27. Ernst, M.H.: Exact solutions of the nonlinear Boltzmann equation. *J. Stat. Phys.* **34**(5), 1001–1017 (1984). <https://doi.org/10.1007/BF01009454>
28. Djordjic, V., Pavicolic, M., Torrilhon, M.: Consistent, explicit, and accessible Boltzmann collision operator for polyatomic gases. *Phys. Rev. E* **104**, 025309 (2021). <https://doi.org/10.1103/PhysRevE.104.025309>

29. Pennie, C.A., Gamba, I.M.: Convergence and error estimates for the conservative spectral method for Fokker–Planck–Landau equations (2020). [arXiv:2009.10352](https://arxiv.org/abs/2009.10352) [math.NA]
30. Hall, B.C.: Lie groups, lie algebras, and representations. In: Hall, B.C. (ed.) *Quantum Theory for Mathematicians*, pp. 333–366. Springer, New York (2013). [https://doi.org/10.1007/978-1-4614-7116-5\\_16](https://doi.org/10.1007/978-1-4614-7116-5_16)
31. Hall, B.: Matrix lie groups. In: Hall, B. (ed.) *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, pp. 3–30. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-13467-3\\_1](https://doi.org/10.1007/978-3-319-13467-3_1)
32. Hall, B.: Lie algebras. In: Hall, B. (ed.) *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, pp. 49–76. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-13467-3\\_3](https://doi.org/10.1007/978-3-319-13467-3_3)
33. Hall, B.C.: Basic representation theory. In: Hall, B.C. (ed.) *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, pp. 77–107. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-13467-3\\_4](https://doi.org/10.1007/978-3-319-13467-3_4)
34. Weyl, H.: *The Classical Groups: Their Invariants and Representations*. Princeton University Press, Princeton (2016). <https://doi.org/10.1515/9781400883905>
35. Fulton, W., Harris, J.W.: *Representation Theory: A First Course*. Springer, New York (1991)
36. Brauer, R.: On algebras which are connected with the semisimple continuous groups. *Ann. Math.* **38**, 857 (1937)
37. Fulton, W., Harris, J.: Representations of finite groups. In: Fulton, W., Harris, J. (eds.) *Representation Theory: A First Course*, pp. 3–11. New York, NY, Springer, New York (2004). [https://doi.org/10.1007/978-1-4612-0979-9\\_1](https://doi.org/10.1007/978-1-4612-0979-9_1)
38. Fulton, W., Harris, J.: Appendix B on multilinear algebra. In: Fulton, W., Harris, J. (eds.) *Representation Theory: A First Course*, pp. 471–477. Springer, New York (2004)
39. Hall, B.: Clebsch Gordan theory and the Wigner–Eckart theorem. In: Axler, S., Ribet, K. (eds.) *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, pp. 425–434. Springer International Publishing, Cham (2015). <https://doi.org/10.1007/978-3-319-13467-3>
40. Weyl, H.: Chapter V. The orthogonal group. In: Weyl, H. (ed.) *The Classical Groups: Their Invariants and Representations*, pp. 137–164. Princeton University Press, Princeton (2016). <https://doi.org/10.1515/9781400883905-007>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.