SIAM J. NUMER. ANAL.
Vol. 56, No. 6, pp. 3534–3579

CONVERGENCE AND ERROR ESTIMATES FOR THE LAGRANGIAN-BASED CONSERVATIVE SPECTRAL METHOD FOR BOLTZMANN EQUATIONS*

RICARDO J. ALONSO†, IRENE M. GAMBA‡, AND SRI HARSHA THARKABHUSHANAM§

Abstract. We develop error estimates for the semidiscrete conservative spectral method for the approximation of the elastic and inelastic space homogeneous Boltzmann equation introduced by Gamba and Tharkabhushanam in [*J. Comput. Phys.*, 228 (2009), pp. 2012–2036]. In addition we study the long time convergence of such semidiscrete solution to the equilibrium Maxwellian distribution that conserves the mass, momentum, and energy associated with the initial data. The numerical method is based on the Fourier transform of the collisional operator and a Lagrangian optimization correction that enforces the collision invariants, namely, conservation of mass, momentum, and energy in the elastic case, and just mass and momentum in the inelastic one. We present a detailed semidiscrete analysis on convergence of the proposed numerical method which includes the $L^1 - L^2$ theory for the scheme. This analysis allows us to present, additionally, convergence in Sobolev spaces and convergence to equilibrium for the numerical approximation. The results of this work answer a long standing open problem posed by Cercignani, Illner, and Pulvirenti in [*Mathematical Theory of Dilute Gases*, Springer, New York, 1994, Chapter 12] about finding error estimates for a numerical scheme associated with the Boltzmann equation, as well as showing the semidiscrete numerical solution converges to the equilibrium Maxwellian distribution associated with the initial value problem.

Key words. nonlinear integral equations, rarefied gas flows, Boltzmann equations, conservative spectral methods

AMS subject classifications. 45E99, 35A22, 65C20

DOI. 10.1137/18M1173332

1. Introduction.

The Boltzmann transport equation is an integro-differential transport equation that describes the evolution of a single point probability density function $f(t, v, x)$ defined as the probability of finding a particle at position x with kinetic velocity v at time t. The mathematical and computational difficulties associated with the Boltzmann equation are due to the nonlocal and nonlinear nature of the binary collision operator, which is usually modeled as a bilinear integral form in d-dimensional velocity space and unit sphere $\mathbb{S}^{d-1}$.

The focus of this manuscript is to provide a complete consistency and error analysis and long time convergence to statistical equilibrium states for the Lagrangian-based conservative spectral scheme proposed in [30] to solve the dynamics of elastic binary collisions. In particular, the results of this work answer a long standing open problem posed by Cercignani, Illner, and Pulvirenti in [19, Chapter 12] about finding error estimates for a consistent nonlinear Boltzmann deterministic scheme for elastic binary interactions in the case of hard potentials.

*Received by the editors March 2, 2018; accepted for publication (in revised form) October 8, 2018; published electronically December 13, 2018.

http://www.siam.org/journals/sinum/56-6/M117333.html

Funding: The first author was supported by ONR grants N000140910290 and NSF-RNMS 1107465. The second author was supported by NSF grant DMS-1413064. The third author was supported by NSF grant DMS-0807712.

†Department of Mathematics, P.U.C. Rio de Janeiro, Rio de Janeiro 22451-900, Brazil (ricardoalonsoplata@gmail.com).

‡ICES and Department of Mathematics, The University of Texas at Austin, Austin, TX 78712 (gamba@math.utexas.edu).

§BJSS - London, London EC2V 6BP, UK (jesmith@fictional.edu).

The problem of computing efficiently the Boltzmann transport equation has interested many authors that have introduced different approaches. These approaches can be classified as stochastic methods known as direct simulation Monte Carlo methods (DSMC) ([8, 47, 49, 29]) and deterministic methods (discrete velocity models [35, 18, 12, 39, 33], Boltzmann approximations–lattice Boltzmann, Bhatnagar–Gross–Krook operator and spectral methods [25, 14, 46, 9, 11, 16, 45, 34, 23, 24, 42]). Spectral-based methods, our choice for this work, have been developed by Gamba and Tharkabhushanam [30] inspired in the work developed a decade earlier by Gabetta, Pareschi, and Toscani [25] and later by Bobylev and Rjasanow [14] and Pareschi and Russo [46]. The practical implementation of these methods is supported by the groundbreaking work of Bobylev [9] using the Fourier transformed Boltzmann equation to analyze its solutions in the case of Maxwell-type interactions. After the introduction of the inelastic Boltzmann equation for Maxwell-type interactions and the use of the Fourier transform for its analysis in Bobylev, Carrillo, and Gamba [11], the spectral-based approach is becoming the most suitable tool to deal with deterministic computations of kinetic models associated with the full Boltzmann collisional integral, both for elastic or inelastic interactions. Recent implementations of spectral methods for the nonlinear Boltzmann are due to Bobylev and Rjasanow [14] who developed a method using the fast Fourier transform (FFT) for Maxwell-type interactions and then for hard-sphere interactions [15] using generalized Radon and X-ray transforms via FFT. Simultaneously, Pareschi and Perthame [45] developed a similar scheme using FFT for Maxwell-type interactions. Using [46, 45], Filbet, Mouhot, and Pareschi [23] and Filbet and Russo in [24] have implemented a scheme to solve the space inhomogeneous Boltzmann equation. We also mention the work of Ibragimov and Rjasanow [34] who developed a numerical method to solve the space homogeneous Boltzmann equation on a uniform grid for variable hard potential interactions with elastic collisions. This particular work has been a great inspiration for the current paper and was one of the first steps in the direction of a new numerical method.

The aforementioned works on deterministic solvers for the nonlinear Boltzmann transport equation have been restricted to elastic, conservative interactions. Mouhot and Pareschi [42] have studied some approximation properties of the schemes. Part of the difficulties in their strategy arises from the constraint that the numerical solution has to satisfy conservation of the initial mass. To this end, the authors propose the use of a periodic representation of the distribution function to avoid aliasing. Closely related to this problem is the fact that spectral methods do not guarantee the positivity of the solution due to the combined effects of the truncation in velocity domain (of the equation) and the application of the Fourier transform (computed for the truncated problem). In addition to this, there is no a priori conservation of mass, momentum, and energy in [23, 24, 42]. In fact, the authors in [22] presented a stability and convergence analysis of the spectral method for the homogeneous Boltzmann equation for binary elastic collisions using the periodization approach proposed in those previous references. In their results, the spectral scheme enforced only mass conservation; as a consequence, the numerical solutions converge to the constant state, hence, destroying the time asymptotic behavior predicted by the Boltzmann $\mathcal{H}$-theorem.

It is shown in this manuscript that the conservative approach scheme proposed in [30] is able to handle the conservation problem in a natural way, by means of Lagrange multipliers, and enjoys convergence and a correct long time asymptote to the Maxwelliam equilibrium. Our approximation by conservative spectral Lagrangian schemes and corresponding computational method is based on an alternative approach to the work in [14, 34]. This spectral approach combined with a constrained

minimization problem works for elastic or inelastic collisions and energy dissipative nonlinear Boltzmann-type models for variable hard potentials. We do not use periodic representations for the distribution function and the only restriction of the current method is that it requires that the distribution function be Fourier transformable at any time step. This requirement is met by imposing L^2-integrability to the initial datum. The required conservation properties of the distribution function are enforced through an optimization problem with the desired conservation quantities set as the constraints. The correction to the distribution function that makes the approximation conservative is very small but crucial for the evolution of the probability distribution function according to the Boltzmann equation.

More recently, this conservative spectral method for the Boltzmann equation was applied to the calculation of the Boltzmann flow for anisotropic collisions, even in the Coulomb interaction regime [26], where the solution of the Boltzmann equation approximates the solution for the Landau equation [37, 38]. It has also been extended to systems of elastic and inelastic hard potential problems modeling of a multienergy level gas [44]. In this case, the formulation of the numerical method accounts for both elastic and inelastic collisions. It was also used for the particular case of a chemical mixture of monatomic gases without internal energy. The conservation of mass, momentum, and energy during collisions is enforced through the solution of the constrained optimization problem to keep the collision invariances associated with the mixtures. The implementation was done in the space inhomogeneous setting (see [44, section 4.3]), where the advection along the free Hamiltonian dynamics is modeled by time splitting methods following the initial approach in [31]. The effectiveness of the scheme applied to these mixtures has been compared with the results obtained by means of the DSMC method and excellent agreement has been observed.

In addition, this conservative spectral Lagrangian method has been implemented in a system of electron-ion in plasma modeled by a 2×2 system of Poisson–Vlasov–Landau equations [52] using time splitting methods, that is, staggering the time steps for advection of the Vlasov–Poisson system and the collisional system including recombinations. The constrained optimization problem is applied to the collisional step in a revised version from [30] where such a minimization problem was posed and solved in Fourier space, using the exact formulas for the Fourier transform of the collision invariant polynomials. The benchmarking for the constrained optimization implementation for the mixing problem was done for an example of a space homogeneous system where the explicit decay difference for electron and ion temperatures is known [52, section 7.1.2]. Yet, the used scheme captures the total conserved temperature, being a convex sum of the ions and electron temperatures, respectively.

The keynote results of the manuscript are stated in Theorem 3.1 in section 3. The proof of this theorem relies on the Lagrangian correction problem that enforces conservation at the numerical level. This is a key idea that shows that the conservative spectral scheme converges to the Gaussian (Maxwellian) distribution in velocity space. Indeed, the enforcement of the collision invariants is sufficient to show the convergence result to the Maxwellian equilibrium in the case of a scalar space homogeneous Boltzmann equation for binary elastic interactions. This is exactly how the Boltzmann $\mathcal{H}$-theorem works [19]; the equilibrium Maxwellian (2.10) is proven to be the stationary state due to the conservation properties combined with the elastic collision law.

In the case of inelastic collisions for either Maxwell type of hard sphere interactions for constant rate of local energy law [11, 27, 40] or visoelastic particle type of interactions [6, 7], where local energy rates depend on the local impact angle, making

them an elastic interaction as the interaction is glancing, the number of collision invariants to be enforced is just $d+1$ polynomials. In addition, trivial stationary states are either a singular distribution or vacuum, and it has been shown that there also are nontrivial attracting self-similar solutions that develop power tail distributions in the self-similar framework, as computed in [30] and references therein for an in-depth discussion of the phenomenon. In particular, it would not be correct to use approximating schemes that enforce local or global Maxwellian behavior as they will eventually generate errors. In fact, in the case of the scalar homogeneous Boltzmann for binary inelastic collisions of Maxwell type, the scheme is able to accurately compute the evolution to self-similar states with power tails, by exhibiting the predicted corresponding moment growth as performed in [30].

The conservative spectral Lagrangian has also been implemented to numerically simulate a gas mixture system for chemically interacting gases, [44, 52], where recombination terms depend on mass ratios, even if the particle-particle interaction is elastic. In particular, while each component of the gas mixture does not conserve energy, the total system does. The resulting conservation scheme, then, enforces the proper collision invariants for the total system by enforcing a convex combination of the thermodynamic macroscopic quantities, but not for the collision invariants of individual components.

Enforcing the system to conserve total quantities by the suitable constrained minimization problem associated with initial data for the mixture will select the correct equilibrium states associated with each system component. A proof of this statement would require us to adjust the *conservation correction estimate* of Lemma 3.4 now extended to the adequate convex combination of collision invariants corresponding to the initial data of the system, as it was computed in [44] for a 2×2 neon argon gas mixture, or a 5×5 multienergy level gas mixture using the classical hard sphere model, as well as in [52] for an electronion plasma mixture using the Landau equation for Coulomb potentials.

The paper is organized as follows. In section 2, the preliminaries and description of the spectral method for the space homogenous Boltzmann equation are presented. In section 3, we introduce the optimization problem proving the basic estimates including spectral accuracy and consistency, results in both elastic and inelastic collisions in Theorem 3.4. In sections 4, 5, and 6 we develop the existence, convergence, and error estimates for the elastic interactions scheme, which heavily relies on the analytical properties of the model for a space homogeneous, monoatomic, single component, elastic interacting gas for hard potentials and integrable angular cross section kernel. Finally, in section 7 we show local stability and long time convergence of the method. In this section we prove that, in fact, all constant in the estimates are uniform in time. We point out that it is possible to carry out this program for the inelastic framework of viscoelastic interactions, as all the necessary analytical tools are already available in [6, 7]. The methodology we follow is summarized in the following steps:

1. In section 4 we prove a priori estimates for the moments and the L^2_k-norms of the scheme under a small negative mass assumption. The analysis involves a coupled estimate on moments and the L^2-norm due to the fact that spectral methods fundamentally need the L^2-theory. An estimate for the amount of the negative mass produced by the scheme along time is proven as well.
2. We use the a priori estimates of section 4 to prove global existence in section 5. The key ingredient is to keep the negative mass formation under the numerical scheme in control. We, then, show propagation of regularity.

3. In section 6 we develop the error estimates of the scheme using the propagation of moments and Sobolev norms provided in sections 4 and 5. The core of the document finishes, in section 7, with a result on the local stability and exponential convergence of the scheme to the thermal equilibrium. This last part helps to make all constants found in previous sections uniform in time.

Finally, some conclusion are drawn in section 8 and a useful toolbox is given in the appendix.

2. Preliminaries.

2.1. The Boltzmann equation and its Fourier representation. The initial value problem associated with the space homogeneous Boltzmann transport equation modeling the statistical evolution of a single point probability distribution function $f(t,v)$ is given by

$$\frac{\partial f}{\partial t}(t,v) = Q(f,f)(t,v) \text{ in } (0,T]\times\mathbb{R}^d \tag{2.1}$$

with initial data $f(0,v)=f_0$. The weak form of the collision integral is given by

$$\int_{\mathbb{R}^d} \mathcal{Q}(f,f)(v)\phi(v)\mathrm{d}v = \int_{\mathbb{R}^{2d}}\int_{\mathbb{S}^{d-1}} f(v,t)f(w,t)[\phi(v')-\phi(v)]B(|u|,\hat{u}\cdot\sigma)\mathrm{d}\sigma\mathrm{d}w\mathrm{d}v\,, \tag{2.2}$$

where the corresponding velocity interaction law exchanging velocity pairs $\{v,w\}$ into postcollisional pairs $\{v',w'\}$ is given by the law

$$v' = v+\frac{\beta}{2}(|u|\sigma-u) \quad \text{and} \quad w' = w-\frac{\beta}{2}(|u|\sigma-u), \tag{2.3}$$

where $\beta\in(1/2,1]$ is the energy dissipation parameter, $u=v-w$ is the relative velocity, and $\sigma\in\mathbb{S}^{d-1}$ is the unit direction of the postcollisional relative velocity $u'=v'-w'$. The parameter β is related to the degree of inelasticity of the interactions with $\beta=1$ being elastic and $\beta<1$ inelastic interactions.

The collision kernel, quantifying the rate of collisions during interactions, carries important properties that are of fundamental importance for the regularity therory of the Boltzmann collisional integral. It is assumed to be

$$B(|u|,\hat{u}\cdot\sigma) = |u|^\lambda\, b(\hat{u}\cdot\sigma) \quad \text{with} \quad 0\le\lambda\le 1\,. \tag{2.4}$$

The scattering angle θ is defined by $\cos\theta=\hat{u}\cdot\sigma$, where the hat stands for unitary vector. Further, we assume that the differential cross section kernel $b(\hat{u}\cdot\sigma)$ is integrable in $\mathbb{S}^{d-1}$, referred to as the *Grad cutoff assumption* [32], and it is renormalized in the sense that

$$\int_{\mathbb{S}^{d-1}} b(\hat{u}\cdot\sigma)\mathrm{d}\sigma = \left|\mathbb{S}^{d-2}\right|\int_0^\pi b(\cos\theta)\sin^{d-2}\theta\mathrm{d}\theta = \left|\mathbb{S}^{d-2}\right|\int_{-1}^1 b(s)(1-s^2)^{(d-3)/2}\mathrm{d}s = 1\,, \tag{2.5}$$

where the constant $\left|\mathbb{S}^{d-2}\right|$ denotes the Lebesgue measure of $\mathbb{S}^{d-2}$. The parameter λ in (2.4) regulates the collision frequency and accounts for interparticle potentials occurring in the gas. These interactions are referred to as variable hard potentials whenever $0<\lambda<1$, Maxwell–molecules-type interactions for $\lambda=0$, and hard spheres

for $\lambda = 1$. In addition, if kernel b is independent of the scattering angle we call the interactions isotropic, otherwise, we refer to them as anisotropic variable hard potential interactions.

It is worth mentioning that the weak form of the collisional form (2.2) also takes the following weighted double mixing convolutional form

$$\int_{\mathbb{R}^d} Q(f,f)(v)\,\phi(v)\,\mathrm{d}v = \int_{\mathbb{R}^{2d}} f(v)f(v-u)\,\mathcal{G}(v,u)\mathrm{d}u\mathrm{d}v\,. \tag{2.6}$$

The weight function defined by

$$\mathcal{G}(v,u) = \int_{\mathbb{S}^{d-1}} \big[\phi(v') - \phi(v)\big]\, B(|u|, \hat{u}\cdot\sigma)\,\mathrm{d}\sigma \tag{2.7}$$

depends on the test function $\phi(v)$, the collisional kernel $B(|u|, \hat{u}\cdot\sigma)$ from (2.4), and the exchange of collisions law (2.3). This is actually a generic form of a Kac master equation formulation for a binary multiplicatively interactive stochastic Chapman–Kolmogorov birth-death rate process, where the weight function $\mathcal{G}(v,u)$ encodes the detailed balance properties, collision invariants as well as existence, regularity, and decay rate dynamics to equilibrium.

We also denote by $'v$ and $'w$ the precollision velocities corresponding to v and w. In the case of elastic collisions (i.e., $\beta = 1$) the pairs $\{'v,'w\}$ and $\{v', w'\}$ agree, otherwise, extra caution is advised.

Collision invariants and conservation properties. The collision law (2.3) is equivalent to the following relation between the interacting velocity pairs:

$$v + w = v' + w' \quad \text{and} \quad |v|^2 + |w|^2 = |v'|^2 + |w'|^2 - \beta(1-\beta)B(|u|, \hat{u}\cdot\sigma).$$

In particular, when testing with the polynomials $\varphi(v) = 1,\ v_j,\ |v|^2$ in $\mathbb{R}^d$, it yields the following conservation relations

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} f \begin{pmatrix} 1 \\ v_j \\ |v|^2 \end{pmatrix} \mathrm{d}v = \int_{\mathbb{R}^{2d}} f(v_*)f(v) \int_{\mathbb{S}^{d-1}} \begin{pmatrix} 0 \\ 0 \\ -\beta(1-\beta) \end{pmatrix} B(|u|, \hat{u}\cdot\sigma)\mathrm{d}\sigma\mathrm{d}v_*\mathrm{d}v\,. \tag{2.8}$$

The polynomials that make the collisional integral vanish are called collision invariants. Clearly, in the elastic case when $\beta = 1$, the homogeneous Boltzmann equation has $d+2$ collision invariants and corresponding conservation laws, namely, mass, momentum, and kinetic energy. For the inelastic case $\beta < 1$, the number of invariants and conserved quantities is $d+1$.

Finally, when testing with $\varphi(v) = \log f(v)$ it yields the inequality ($\mathcal{H}$-theorem holding for the elastic case)

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} f\log f\mathrm{d}v = \int_{\mathbb{R}^d} Q(f)\log f\mathrm{d}v \\
&\quad= \int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} f(w)f(v)\left(\log\left(\frac{f(w')f(v')}{f(w)f(v)}\right) + \frac{f(w')f(v')}{f(w)f(v)} - 1\right) B(|u|, \hat{u}\cdot\sigma)\mathrm{d}\sigma\mathrm{d}w\mathrm{d}v \\
&\qquad+ \int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} f(w)f(v)\left(\frac{1}{(2\beta-1)J_\beta} - 1\right) B(|u|, \hat{u}\cdot\sigma)\mathrm{d}\sigma\mathrm{d}w\mathrm{d}v \\
&\quad\le \int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} f(w)f(v)\int_{\mathbb{S}^{d-1}}\left(\frac{1}{(2\beta-1)J_\beta} - 1\right) B(|u|, \hat{u}\cdot\sigma)\mathrm{d}\sigma\mathrm{d}w\mathrm{d}v = 0 \ \text{ iff } \ \beta = 1\,.
\end{aligned} \tag{2.9}$$

Recall the following fundamental result in elastic particle theory:

The Boltzmann Theorem (for $\beta = 1$).

$$\int_{\mathbb{R}^d} Q(f)\log f = 0 \iff \log f(v) = a + \mathbf{b}\cdot v - c|v|^2\,,$$

where $f \in L^1(\mathbb{R}^d)$ for $c > 0$, where the parameters a, $\mathbf{b}$, and c are determined by the initial state moments given by the $d+2$ collision invariants.

That means that given an initial state $f_0(v) \geq 0$ for a.e. $v \in \mathbb{R}^d$, we have $\int_{\mathbb{R}^d} f_0(v)(1+|v|^2)\,dv < \infty$. In the limit as $t \to +\infty$, we expect that $f(t,v)$ converges to the *equilibrium Maxwellian* distribution, i.e.,

$$f(t,v) \to \mathcal{M}_0[m_0,u_0,\Theta_0](v) := m_0(2\pi\Theta_0)^{-d/2}\exp\left(-\frac{|v-u_0|^2}{2\Theta_0}\right)\,, \tag{2.10}$$

where the density mass, momentum, and energy are defined by

$$m_0 := \int_{\mathbb{R}^d} f_0(v)\,dv\,,\quad u_0 := \frac{1}{m_0}\int_{\mathbb{R}^d} f_0(v)\,dv\,,\quad \Theta_0 := (d\,m_0)^{-1}\int_{\mathbb{R}^d}|v-u_0|^2\,f_0(v)\,dv\,.$$

The Fourier formulation of the collisional form. One of the pivotal points in the success of the spectral numerical method for the computation of the nonlinear Boltzmann equation lies in the simplicity of the representation of the collision integral in Fourier space by means of its weak form. Indeed taking the Fourier multiplier as the test function, i.e.,

$$\psi(v) = \frac{e^{-i\zeta\cdot v}}{(\sqrt{2\pi})^d}$$

in the weak formulation (2.2), where ζ is the Fourier variable, one obtains the Fourier transform of the collision integral

$$\begin{aligned}\widehat{Q(f,f)}(\zeta) &= \frac{1}{(\sqrt{2\pi})^d}\int_{\mathbb{R}^d} Q(f,f)e^{-i\zeta\cdot v}\mathrm{d}v\\ &= \frac{1}{(\sqrt{2\pi})^d}\int_{\mathbb{R}^{2d}}\int_{\mathbb{S}^{d-1}} f(v)f(w)B(|u|,\hat{u}\cdot\sigma)\left(e^{-i\zeta\cdot v'}-e^{-i\zeta\cdot v}\right)\mathrm{d}\sigma\mathrm{d}w\mathrm{d}v\,.\end{aligned}$$

Thus, using (2.4), (2.6), (2.7) yields

$$\begin{aligned}\widehat{Q(f,f)}(\zeta) &= \frac{1}{(\sqrt{2\pi})^d}\int_{\mathbb{R}^{2d}} f(v)f(w)\int_{\mathbb{S}^{d-1}}|u|^\lambda b(\hat{u}\cdot\sigma)e^{-i\zeta\cdot v}\left(e^{-i\frac{\beta}{2}\zeta\cdot(|u|\sigma-u))}-1\right)\mathrm{d}\sigma\mathrm{d}w\mathrm{d}v\\ &= \frac{1}{(\sqrt{2\pi})^d}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f(v)f(v-u)e^{-i\zeta\cdot v}\mathrm{d}v\right)G_{\lambda,\beta}(u,\zeta)\mathrm{d}u\\ &= \frac{1}{(\sqrt{2\pi})^d}\int_{\mathbb{R}^d}\widehat{f\,\tau_{-u}f}(\zeta)\,G_{\lambda,\beta}(u,\zeta)\,\mathrm{d}u\,,\end{aligned} \tag{2.11}$$

where the weight function $G_{\lambda,\beta}(u,\zeta)$ is defined by the spherical integration

$$G_{\lambda,\beta}(u,\zeta) := |u|^\lambda\int_{\mathbb{S}^{d-1}} b(\hat{u}\cdot\sigma)\left(e^{-i\frac{\beta}{2}\zeta\cdot(|u|\sigma-u))}-1\right)\mathrm{d}\sigma\,. \tag{2.12}$$

Note that (2.12) is valid for both isotropic and anisotropic interactions. In addition, the function $G_{\lambda,\beta}(u,\zeta)$ is oscillatory and trivially bounded by $|u|^\lambda$ due to the integrability of $b(\cdot)$ from the Grad's cutoff assumption. Further simplification ensues for the three-dimensional isotropic case where a simple computation gives

$$G^{\text{iso}}(u,\zeta) = |u|^\lambda \left(e^{i\frac{\beta}{2}\zeta\cdot u}\,\text{sinc}\left(\frac{\beta|u||\zeta|}{2}\right) - 1\right). \tag{2.13}$$

In addition, recalling elementary properties of the Fourier transform yields

$$\begin{aligned}\widehat{f\,\tau_{-u}f}(\zeta) &= \frac{1}{(\sqrt{2\pi})^d}\,\hat{f} * \widehat{\tau_{-u}f}(\zeta) = \frac{1}{(\sqrt{2\pi})^d}\int_{\mathbb{R}^d}\hat{f}(\zeta-\xi)\widehat{\tau_{-u}f}(\xi)\mathrm{d}\xi \\ &= \frac{1}{(\sqrt{2\pi})^d}\int_{\mathbb{R}^d}\hat{f}(\zeta-\xi)\hat{f}(\xi)e^{-i\xi\cdot u}\mathrm{d}\xi\,.\end{aligned}$$

Hence, using this last identity in (2.11), we finally obtain the following structure in Fourier space:

$$\widehat{Q(f,f)}(\zeta) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\hat{f}(\zeta-\xi)\hat{f}(\xi)\widehat{G_{\lambda,\beta}}(\xi,\zeta)\mathrm{d}\xi\,, \tag{2.14}$$

where

$$\widehat{G_{\lambda,\beta}}(\xi,\zeta) = \int_{\mathbb{R}^d} G_{\lambda,\beta}(u,\zeta)e^{-i\xi\cdot u}\mathrm{d}u\,. \tag{2.15}$$

That is, the Fourier transform of the collision operator $\widehat{Q(f,f)}(\zeta)$ is a weighted convolution of the inputs in Fourier space with weight $\widehat{G_{\lambda,\beta}}(\xi,\zeta)$.

As an example, we compute the weight for the isotropic case in three dimensions. Assume that f has support in the ball of radius $\sqrt{3}L$, hence, the domain of integration for the relative velocity is the ball of radius $2\sqrt{3}L$. Using polar coordinates $u = r\omega$,

$$\begin{aligned}\widehat{G^{\text{iso}}}(\xi,\zeta) &= \int_0^\infty\int_{\mathbb{S}^2} r^2 G^{\text{iso}}(r\omega,\zeta)e^{-ir\xi\cdot\omega}\mathrm{d}\omega\mathrm{d}r \\ &= 4\int_0^{2\sqrt{3}L} r^{\lambda+2}\left(\text{sinc}\left(\frac{r\beta|\zeta|}{2}\right)\text{sinc}\left(r|\frac{\beta}{2}\zeta-\xi|\right) - \text{sinc}\left(r|\xi|\right)\right)\mathrm{d}r.\end{aligned} \tag{2.16}$$

A point worth noting here is that the numerical calculation of expression (2.14) results in $O(N^{2d})$ number of operations, where N is the number of discretizations in each velocity component (i.e., N counts the total number of Fourier modes for each d-dimensional velocity space). However it may be possible to reduce the number of operations to $O(N^{2d-1}\log N)$ for any anisotropic kernel and any initial state. Due to the oscillatory nature of the weight function (2.16) even in the simple case of three dimensions for the hard sphere case, when $b(\hat{u}\cdot\sigma) = 4\pi$, such a calculation cannot be accomplished by $N\log N$ operations if the initial state is far from a Maxwellian state or has an initial discontinuity, as claimed in [23].

Notation and spaces. Before continuing with the discussion, we recall the definition of the Lebesgue's spaces $L^p_k(\Omega)$ and the Hilbert spaces $H^\alpha_k(\Omega)$. These spaces will be used during the manuscript. The set Ω could be any measurable set in the case of the L^p_k spaces or any open set in the case of the H^α_k spaces, however, for our present purpose Ω is either $(-L,L)^d$ or $\mathbb{R}^d$ most of the time:

$$L^p_k(\Omega) := \Big\{ f \,:\, \|f\|_{L^p_k(\Omega)} := \Big(\int_\Omega \big| f(v)\langle v\rangle^{\lambda k} \big|^p \mathrm{d}v \Big)^{\frac{1}{p}} < \infty \Big\} \quad \text{with } p \in [1,\infty)\,,\ k \in \mathbb{R}\,,$$

$$H^\alpha_k(\Omega) := \Big\{ f \,:\, \|f\|_{H^\alpha_k(\Omega)} := \Big(\sum_{\beta \le \alpha} \|D^\beta f\|^2_{L^2_k(\Omega)} \Big)^{\frac{1}{2}} < \infty \Big\} \quad \text{with } \ \alpha \in \mathbb{N}^d\,,\ k \in \mathbb{R}\,,$$

where $\langle v\rangle := \sqrt{1+|v|^2}$. The standard definition is used for the case $p = \infty$,

$$L^\infty_k(\Omega) := \Big\{ f \,:\, \|f\|_{L^\infty_k(\Omega)} := \operatorname{esssup} \big| f(v)\langle v\rangle^{\lambda k} \big| < \infty \Big\} \quad \text{with } \ k \in \mathbb{R}\,.$$

We will commonly use the following shorthand to ease notation when the domain Ω is clear from the context:

$$\|\cdot\|_{L^p_k(\Omega)} = \|\cdot\|_{L^p_k}\,,$$

and the subindex k may be omitted in the norms for the classical spaces L^p and H^α. In addition, following the notation and language of the classical analysis of the Boltzmann equation, and including the fact that numerical solutions are not nonnegative in general, the *moments of a function* f are denoted by

$$m_k(f) := \int_{\mathbb{R}^d} \big| f(v) \big|\, |v|^{\lambda k}\, \mathrm{d}v. \tag{2.17}$$

2.2. Choosing a computational cutoff domain. In order to make a good approximation to the probability density $f(t,v)$, defined for all $v \in \Omega := \mathbb{R}^d$, the solution of the dynamical homogeneous Boltzmann equation initial value problem, we need to solve the proposed spectral numerical scheme in a computational domain given by the bounded set $\Omega_L \in \mathbb{R}^d$ for a sufficiently large N Fourier modes, as it will be defined next in the beginning of section 3. In particular, the global collision operator $Q(f,f)(t,v)$, defined weakly in (2.2), needs to be approximated in such a computational domain Ω_L that will be carefully chosen below for the specific task of solving the homogeneous Boltzmann equation, for a particularly chosen initial datum being a probability density with a prescribed finite and positive initial mass and kinetic energy (i.e., the choice of the computational domain depends on the initial data, as will be carefully explained).

It will be clear, after the discussion of the approximating numerical scheme for the space homogeneous Boltzmann initial value problem and Theorem 3.1, that there are two sources of error: one due to the mode truncation and the other due to domain truncation. Both are always present due to the global nature of the equation. The key point in the choice of the computational domain $\Omega_L = (-L, L)^d$ is that the time dynamics of the analytical solution remains bounded and decays with Maxwellian tails if initially so following the result in [28]. In particular, it is possible to choose a large enough cutoff length L, depending on the initial data with a Gaussian decay rate, whose approximating $g_0(v)$ satisfies condition (3.8), and $\operatorname{supp}\{g_0(v)\} \subset \Omega_{aL}$ for $0 < a \ll 1$. As a consequence the periodization of the domain is not necessary, since the analytical result from [28] combined with the conservation algorithm, secures that the numerical solution will take values very close to zero (i.e., below machine accuracy) near the boundary $\partial\Omega_L$. That means it is enough to choose Ω_L such that, at least, most of the mass and energy of the true Boltzmann solution f will be contained in it during the simulation time.

One possible strategy for choosing the size of Ω_L is as follows: assume, without loss of generality, a bounded initial datum f_0 with compact support and having zero momentum $\int f_0\, v\, dv = 0$. Then,

$$f_0(v) \le \frac{C_0\, m_0}{(2\pi\Theta_0)^{d/2}}\, e^{-\frac{r_0|v|^2}{2\Theta_0}}\,, \tag{2.18}$$

where $m_0 := \int f_0$ is the initial mass, $\Theta_0 := \int f_0|v|^2$ is the initial temperature, and $r_0 \in (0,1]$ and $C_0 \ge 1$ are the stretching and dilating constants. Since the Boltzmann flow propagates Gaussian weighted Lebesgue norms, refer to [36, 50, 10, 13, 43, 28, 5, 4, 1] for this and more related theoretical facts on the equation. Thus, there are some uniform in the time constants r and c depending on the moments of the initial data as much as the potential rates λ and the angular part b of the collision kernel, namely, $r := r(f_0, \lambda, b) \in (0, r_0]$ and $C := C(f_0, \lambda, b) \ge C_0 \ge 1$ such that

$$f(t,v) \le \frac{C\, m_0}{(2\pi\Theta_0)^{d/2}}\, e^{-\frac{r|v|^2}{2\Theta_0}} =: M(f_0, C, r)\,, \quad t > 0\,. \tag{2.19}$$

Now, choose a small quantity $\delta \ll 1$ being the mass proportion of the tails associated with the Maxwellian $M(f_0, C, r)$ from (2.19) that uniformly controls the solution $f(t,v)$ as in [28]. That is,

$$\int_{\Omega_L^c} f(t,v)\langle v\rangle^2 \mathrm{d}v \le \int_{\Omega_L^c} M(f_0, C, r)\langle v\rangle^2 \mathrm{d}v \le \delta \int_{\Omega_L} f_0(v)\langle v\rangle^2 \mathrm{d}v = \delta(m_0 + \Theta_0)\,.$$

Therefore, the parameter δ, for the solution of the Boltzmann equation , thanks to the L^∞ control of the solution in [28], is interpreted as a domain cutoff error tolerance that *remains uniform in time* and solely depends on the approximated initial state, say, $0 \le g_0(v)$ on the chosen Ω_L so that the magnitude of $\delta(m_0+\Theta_0)$ is well below machine accuracy. Clearly, the mass proportion δ must be small enough for $\mathrm{supp}(g_0) \subset\subset \Omega_L$. Equivalently, one needs to choose the size of L (or the measure of the computational domain Ω_L), such that

$$\frac{\int_{\Omega_L^c} M(f_0, C, r)\langle v\rangle^2 \mathrm{d}v}{m_0 + \Theta_0} \le \delta \approx 0\,. \tag{2.20}$$

In order to minimize the computational effort, one should pick the smallest of such domains, that is Ω_L, such that

$$\text{for a fixed } \ a \ll 1\,, \ \ \mathrm{supp}(f_0) \subset \Omega_{aL} \ \text{ and} \tag{2.21}$$

that Ω_{aL}^c satisfies (2.20) in the sense that the numerical approximated initial datum vanishes in a neighbourhood of the boundary of Ω_L beyond several orders down of machine accuracy. In addition, under this conditions we invoke the restriction operators in Sobolev space arguments, such as (2.27) in the subsection below, which allow us to make rigorous semidiscrete error estimates in Sobolev norms with respect to the solution $f(t,\cdot) \in \mathbb{R}^d$ of the homogeneous Boltzmann–Cauchy problem (2.1)–(2.5) under consideration.

Finally, for such an estimate (2.21) to be of practical use one would need to compute the precise value of the constants C and r. As a general matter, these constants come from available analytical estimates, which, although quantitative, are likely far from optimal. The result is that the choice (2.21) most of the time overestimates the size of the simulation domain. It is reasonable then, for practical purposes, to simply set $r_o = r = 1$ and choose $C = C_o \ge 1$ as the smallest constant satisfying (2.18)

(which always exists for any compactly supported and bounded f_0). That this choice of parameters is natural, is noted from the fact that

$$\max\left\{g_0\,,\, f_\infty := \frac{m_0}{(2\pi\Theta_0)^{d/2}}\, e^{-\frac{|v|^2}{2\Theta_0}}\right\} \le M(f_0, C, 1)$$

with equality if and only if f_0 is the equilibrium Maxwellian as in (2.10) (in such a case $C = 1$).

This propagation property secures a stable numerical simulation of the Boltzmann equation, provided the numerical preservation of the conservation laws or corresponding collision invariants holds. It also secures, as we will see, the convergence of the numerical scheme to the analytic solution of the initial value problem and the correct long time evolution of such a numerical approximation. In this way, the numerical scheme will converge to the equilibrium Maxwellian as defined in (2.10).

We note that the discussion of this section is fairly independent of the choice of computational scheme and applies to new approaches such as that recently developed in [51] for a Galerkin approach to the computation of the space homogeneous Boltzmann equation for binary interactions.

2.3. Fourier series, projections. In the implementation of a spectral method the single most important analytical tool is the Fourier transform. Thus, for $f \in L^1(U)$ with U open in $\mathbb{R}^d$, the Fourier transform is defined by

$$\widehat{f}(\zeta) := \frac{1}{(\sqrt{2\pi})^d} \int_U f(v) e^{-i\zeta\cdot v} \mathrm{d}v\,. \tag{2.22}$$

The Fourier transform allows us to express the Fourier series in a rather simple and convenient way. Indeed, fixing a domain of work $\Omega_L := (-L, L)^d$ for $L > 0$, recall that for any $f \in L^2(\Omega_L)$ one can use the *Fourier series* to express f as

$$f(v) = \frac{1}{(2L)^d} \sum_{k\in\mathbb{Z}^d} \widehat{f}(\zeta_k) e^{i\zeta_k\cdot v}, \tag{2.23}$$

where $\zeta_k := \frac{2\pi k}{L}$ are the spectral modes and $\widehat{f}(\zeta_k)$ is the Fourier transform of f evaluated in such modes.

The mode projection operator is defined as $\Pi_L^N : L^2(\Omega_L) \to L^2(\Omega_L)$ as

$$\left(\Pi_L^N f\right)(v) := \left(\frac{1}{(2L)^d} \sum_{|k|\le N} \widehat{f}(\zeta_k) e^{i\zeta_k\cdot v}\right); \tag{2.24}$$

in other words, it is the *orthogonal projection* on the "first N^d" basis elements. Also observe that for any integer α the derivative operator commutes with the projection operator in $H_o^\alpha(\Omega_L)$. Indeed, note the identity for any $f \in H_o^\alpha(\Omega_L)$,

$$\begin{aligned} \partial^\alpha \left(\Pi_L^N f\right)(v) &= \frac{1}{(2L)^d} \sum_{|k|\le N} (i\zeta_k)^\alpha \widehat{f}(\zeta_k) e^{i\zeta_k\cdot v} \\ &= \frac{1}{(2L)^d} \sum_{|k|\le N} \widehat{\partial^\alpha f}(\zeta_k) e^{i\zeta_k\cdot v} = \left(\Pi_L^N \partial^\alpha f\right)(v)\,. \end{aligned} \tag{2.25}$$

Recall that Parseval's theorem readily shows

1. $\left\|\Pi_L^N f\right\|_{L^2(\Omega_L)} \leq \left\|f\right\|_{L^2(\Omega_L)}$ for any N, and with equality for $N = \infty$. Also,
2. $\left\|(\mathbf{1} - \Pi_L^N) f\right\|_{L^2(\Omega_L)} \searrow 0$ as $N \to \infty$.

Extension operator for Sobolev regularity propagation. The restriction of the original problem posed in $\mathbb{R}^d$ to an approximation problem posed in a bounded domain Ω_L introduces some technical issues at the boundary generated by the truncation. We deal with this problem by introducing the following scaled cutoff function defined by

$$\begin{aligned} &\chi(v) := \chi_L(v) = \phi(v/L) \text{ with } \phi \text{ a smooth nonnegative function,} \\ &\text{such that } \ \operatorname{supp}\{\phi\} \subset 0.99[-1,1]^d \ \text{ with } \ \phi \equiv 1 \text{ in } 0.95[-1,1]^d . \end{aligned} \tag{2.26}$$

The cutoff function χ allows for the scheme propagation of higher Sobolev regularity estimates (it is not necessary for L^2-convergence) as it smooths out the boundary without incurring a meaningful error (provided Ω_L was well chosen as previously discussed in subsection 2.2). Using the product rule, it follows that

$$\|\chi g\|_{H^\alpha(\Omega_L)} \leq \|\chi\|_{\mathcal{C}^\alpha} \|g\|_{H^\alpha(\Omega_L)} \leq C \|g\|_{H^\alpha(\Omega_L)} \tag{2.27}$$

for any function $g \in H^\alpha(\Omega_L)$. Note also that the constant $C := C_\chi$, that controls the operator norm, can be taken independent of $L \geq 1$. It is important to observe that the function χg vanishes near $\partial\Omega_L$, and so it can be considered as a function in $H^\alpha(\mathbb{R}^d)$ after using the extension operator who assigns the zero value to any point in the complement of Ω_L, that is, $E(\chi g) = 0$ in $\mathbb{R}^d \backslash \Omega_L$. In addition the Sobolev norms of such an extension coincide with those of the restricted χg, which takes values in a compactly supported set in Ω_L that vanishes in a neighborhood of the boundary $\partial\Omega_L$, relative to Ω_L. That precisely means

$$\|E(\chi g)\|_{H^\alpha(\mathbb{R}^d)} = \|\chi g\|_{H^\alpha(\Omega_L)} , \qquad g \in H^\alpha(\Omega_L) . \tag{2.28}$$

Therefore our choice of the the cutoff function χ enable us to implement an extension operator by null values to all space (for a full discussion of extension operators, see [48]). These properties will be useful when comparing the continuous and semidiscrete solutions, which lie in different domains. Furthermore, in the case of L^2-convergence one can simply take $\chi \equiv 1$.

Remark 2.1. A common technique found in the literature to deal with the domain truncation is periodization of the initial data. Why do we not periodize the initial data, but rather use the extension method on Sobolev spaces for functions that vanish in a given bounded domain? The answer is that the approximated data and solution in our problem are probability densities that rapidly decay at large values of velocities v. In the particular case of the homogeneous Boltzmann equation approximation for hard potentials and angular integrable collision cross section, as it is developed in this theory, the crucial issue is to choose the computational domain large enough, depending on the initial data as previously discussed. Under this choice, the cutoff function χ effectively implements an extension at the cost of a negligible error, as it will be shown to be of the order $O(1/L^{\lambda k})$ with $k > 0$ depending on the number of moments of the initial data.

Remark 2.2. In this deterministic approach, as much as with Monte Carlo methods like the Bird scheme [8], the x-space inhomogeneous Hamiltonian transport for non-linear collisional forms is performed by time operator splitting algorithms. That means, depending on the problem, the computational v-domain Ω_L can be updated with respect to the characteristic flow associated with underlying Hamiltonian dynamics.

3. Spectral conservation method. We first introduce a formal analytical viewpoint needed to study the convergence, stability, and error estimates for the semi-discrete solution associated with the spectral method derived in [30].

After the cutoff domain Ω_L has been fixed, we applied the projection operator (2.24) to both sides of (2.1) to arrive at

$$\frac{\partial \Pi_L^N f}{\partial t}(t,v) = \Pi_L^N Q(f,f)(t,v) \ \text{ in } \ (0,T]\times\Omega_L\,. \tag{3.1}$$

Then, it is reasonable to expect that for such a domain Ω_L and for a sufficiently large number of modes N the approximation

$$\Pi_L^N Q(f,f) \sim \Pi_L^N Q(\Pi_L^N f, \Pi_L^N f) \ \text{ in } \ (0,T]\times\Omega_L \tag{3.2}$$

will be valid.

Next, there are two issues worth noting: (1) for functions supported in Ω_L the gain operator Q^+ is supported in $\Omega_{\sqrt{2}L}$, thus, we will consider it, for simplicity, as a function in Ω_{2L}, and (2) the operator Q^- can be exactly computed with a small computational effort since it is a multiplication operator with a standard convolution. As a consequence, one is led to consider the scheme

$$\begin{aligned} \frac{\partial g}{\partial t}(t,v) &= \Pi_{2L}^N Q^+(\chi g, \chi g)(t,v) - Q^-(g,\chi g)(t,v) \\ &=: Q_u(g,g)(t,v) \quad \text{in } (0,T]\times\Omega_L\,, \\ g_0(v) &:= g_0^N = \Pi_L^N f_0(v), \quad \text{initial data,} \end{aligned} \tag{3.3}$$

and expect that it should be a good approximation to $\Pi_L^N f$. Here, Q_u stands for the unconserved collision operator. In other words, we define the numerical solution to be $g_N := g$ and expect to show that this finite mode solution will be a good approximation to the solution of the Boltzmann problem in the cutoff domain, that is $g \approx f$ in Ω_L, provided the number of modes N used is sufficiently large. Classical spectral accuracy theorems would guarantee such an approximation, yet, fixing the number of Fourier modes to say N^* would strip the conservation properties, as Q_u does not preserve the $d+2$ collision invariants after each time step, and that generates a source of cumulative error that heavily constrains the meaningful simulation time of the scheme.

This problem was overcome in [30] with the conservative spectral scheme we are now analyzing. They introduce a conservation correction by solving a Lagrangian constrained minimization problem each time step (with $O(N)$ in computational complexity), where the objective function to be minimized is the $L^2(\Omega_L)$-distance from the unconseved Q_u to the minimizer $X^* =: Q_c$ subject to the constraint of preserving the $d+2$ collision invariants. To be more precise, the following problem is computationally solved in [30].

Minimization elastic problem (E): Consider the Banach space

$$\mathcal{B}^e = \left\{ X \in L^2(\Omega_L) : \int_{\Omega_L} X = \int_{\Omega_L} Xv = \int_{\Omega_L} X|v|^2 = 0 \right\}, \tag{3.4}$$

and the minimization problem

$$X^* := \min_{X\in\mathcal{B}^e} \mathcal{A}^e(X) := \min_{X\in\mathcal{B}^e} \int_{\Omega_L} \big(Q_u(f,f)(v) - X\big)^2 \mathrm{d}v. \tag{3.5}$$

The solution of this problem applied to our semidiscrete framework will be addressed in the next subsection 3.2. It can be solved by an algorithm, described below in (3.32)–(3.36), that delivers a unique explicit algorithm discrete vector form

$$Q_c(f,f) := X^* \in N^d\,, \tag{3.6}$$

associated with any discretization of f on N^d Fourier modes, where the constraint in (3.4) is given by the linear equation $\mathbf{C}^e Q_c = \mathbf{a}^e$, where the vector $\mathbf{a}^e = 0 \in N^{d+2}$ for the elastic problem or $\mathbf{a}^e = 0 \in N^{d+1}$ for the inelastic one. The matrix $\mathbf{C}^e$ is explicitly precomputed depending on the quadrature rule used to compute the integrals associated with the collision invariants.

In the following sections we intend to prove this formalism under reasonable assumptions. In fact, we study a modification of this problem, namely, the convergence towards f of the solution g of the problem

$$\begin{aligned} &\frac{\partial g}{\partial t}(t,v) = Q_c(g,g)(t,v) \quad \text{in } (0,T]\times\Omega_L\,,\\ &g_0(v) := g_0^N = \Pi_L^N f_0(v), \quad \text{initial data}\,, \end{aligned} \tag{3.7}$$

with $Q_c(f,f)$ the solution of the Lagrangian constrained problem (3.5), (3.6), and the initial datum g_0 satisfies the following condition

$$\frac{\int_{\{g_0<0\}} |g_0(t,v)|\langle v\rangle^2 \mathrm{d}v}{\int_{\{g\geq 0\}} g_0(t,v)\langle v\rangle^2 \mathrm{d}v} \;\leq\; \epsilon \quad \text{and} \quad \|g_0\|_{L^2(\Omega_L)} < \infty \tag{3.8}$$

for some fixed $0 < \epsilon \leq 1/4$, where the operator $Q_c(g,g)$ is defined as the $L^2(\Omega_L)$-closest function to $Q_u(g,g)$ having null mass, momentum, and energy.

We summarize the main results on convergence, error estimates, and asymptotic behavior in the following theorem, whose rigorous proof is developed in the rest of the manuscript. As mentioned in the introduction, the following theorem is proved for the classical elastic model $\beta = 1$. A rigorous proof for the inelastic model can be done, at least, for some special regimes such as the viscoelastic particle model [6, 7] with analog arguments. Additional considerations about self-similar scaling are needed to obtain sharp long time behavior associated with the model, which will be properly addressed by the authors in an upcoming manuscript.

THEOREM 3.1 (error estimates and convergence to Maxwellian equilibrium). *Fix a nonnegative initial datum $f_0 \in L^1_k \cap L^2(\mathbb{R}^d)$ with $k \geq k_*(f_0) \geq 2$, and let $f \geq 0$ be the solution of the Boltzmann equation* (2.1) *with* (2.5). *Then, there exist a cutoff domain $L_0(f_0) > 0$ and a number of modes $N_0 := N(L_0, f_0) > 0$ such that*

1. *semidiscrete existence and uniqueness: Taking $g_0 = \Pi_L^N f_0$, the semi-discrete problem* (3.7) *has a unique solution $g \in \mathcal{C}(0,T; L^1_k \cap L^2(\Omega_L))$ for any $T > 0$, $L \geq L_0$, $N \geq N_0$;*
2. *$L^2_{k'}$-error estimates: Taking $f_0 \in L^1_2 \cap L^2_k(\mathbb{R}^d)$, $k'' \geq 0$, $k_*(f_0) \leq k' \leq k - 1 - \frac{d^+}{2\lambda} - k''$, then*

$$\begin{aligned} \sup_{t\geq 0} \|f-g\|_{L^2_{k'}(\Omega_L)} \leq C(f_0)\big(&\|f_0 - g_0\|_{L^2_{k'}(\Omega_L)} \\ &+ O(L^{\lambda k'}/N^{(d-1)/2}) + O(1/L^{d/2+\lambda k''})\big)^{\frac{1}{1+\theta}}\,, \quad L \geq L_0\,,\ N \geq N_0; \end{aligned}$$

3. *$H^\alpha_{k'}$-error estimates: For the smooth case, taking $f_0 \in L^1_2 \cap H^{\alpha_0}_k(\mathbb{R}^d)$, $k'' \geq 0$, $k_*(f_0) \leq k' \leq k - 1 - \alpha/2 - \frac{d^+}{2\lambda} - k''$, it follows for any $0 \leq \alpha \leq \alpha_0$ that*

$$\sup_{t\geq 0}\|f-g\|_{H^\alpha_{k'}(\Omega_L)} \leq C(f_0)\big(\|f_0-g_0\|_{H^\alpha_{k'+\alpha/2}(\Omega_L)} + O(L^{\lambda(k'+\alpha/2)+\alpha_0}/N^{(d-1)/2+\alpha_0}) + O(1/L^{d/2+\lambda k''})\big)^{\frac{1}{1+\theta}}, \quad L\geq L_0\,,\ N\geq N_0\,.$$

In all cases $k, k' \geq k_*(f_0) \geq 2$, where $k_*(f_0)$ is a required threshold that only depends on f_0. Also, the constant $C(f_0) := C(k', \alpha, f_0)$ in items 2 *and* 3 *depends on f_0 by means of its initial regularity, and the constant $\theta := \theta(k', \alpha) > 0$;*

4. *Convergence to the equilibrium Maxwellian: For every $\delta > 0$ there exists a simulation time $T(\delta) \sim \nu^{-1} \ln\big(\|f_0\|_{H^\alpha(\mathbb{R}^d)}/\delta\big)$ such that for any $\alpha \leq \alpha_0$*

$$\sup_{t\geq T(\delta)}\|g-\mathcal{M}_0\|_{H^\alpha(\Omega_L)} \leq \delta\,, \quad L\geq L_0\,,\ N\geq N_0\,,$$

where $\nu > 0$ is the spectral gap of the linearized Boltzmann operator, and $\mathcal{M}_0$ is the equilibrium Maxwellian (2.10) *having the same mass, momentum, and kinetic energy as the initial datum f_0.*

The proof of these statements in Theorem 3.1 is made in the next four sections. Before starting with the details of the proof, we introduce the shorthand notation

$$O_r := O(L^{-r})\,, \quad r>0\,, \tag{3.9}$$

which will be extensively used throughout the manuscript.

3.1. Conservation method: An extended isoperimetric problem. Throughout this section we fix $f \in L^2(\Omega_L)$. Due to the truncation of the velocity domain the unconserved discrete operator $Q_u \in N^d$ defined for N^d Fourier modes, as a function in Ω_L, does not preserve mass, momentum, and energy. Such a conservation property is at the heart of the kinetic theory of the Boltzmann equation, thus, it is desirable for a numerical solution to possess it. In order to achieve this, we enforce these moment conservations artificially by imposing them as constraints in an optimization problem.

Hence, we first focus on the general form of solution of the minimization problem (3.4), (3.5), whose proof is presented next.

LEMMA 3.2 (elastic Lagrange estimate). *The problem* (3.5) *has a unique minimizer given by*

$$Q_c(f,f)(v) := X^\star = Q_u(f,f)(v) - \frac{1}{2}\Big(\gamma_1 + \sum_{j=1}^{d}\gamma_{j+1}v_j + \gamma_{d+2}|v|^2\Big), \tag{3.10}$$

where γ_j, for $1 \leq j \leq d+2$, are Lagrange multipliers associated with the elastic optimization problem. They are given by

$$\begin{aligned} \gamma_1 &= O_d\rho_u + O_{d+2}e_u\,,\\ \gamma_{j+1} &= O_{d+2}\mu_u^j\,, \quad j=1,2,\cdots,d,\\ \gamma_{d+2} &= O_{d+2}\rho_u + O_{d+4}e_u \end{aligned} \tag{3.11}$$

with O_r defined in (3.9) *and the parameters ρ_u, e_u, μ_u^j are the numerical moments of the unconserved numerical collision operator, defined below in* (3.15). *The minimized objective function can be estimated by*

$$\begin{aligned}\mathcal{A}^e\big(Q_c(f,f)(v)\big) &= \big\|Q_u(f,f) - Q_c(f,f)(v)\big\|^2_{L^2(\Omega_L)}\\ &\le C(d)\Big(2\gamma_1^2 L^d + \sum_{j=1}^{d}\gamma_{j+1}^2 L^{d+2} + \gamma_{d+2}^2 L^{d+4}\Big)\\ &\le \frac{C(d)}{L^d}\left(\rho_u^2 + \frac{e_u^2}{L^{d+1}} + \sum_{j=2}^{d+1}\mu_j^2\right).\end{aligned} \tag{3.12}$$

In the particular case of dimension $d = 3$ the estimate becomes

$$\begin{aligned}&\big\|Q_u(f,f)-Q_c(f,f)\big\|^2_{L^2(\Omega_L)}\\ &\quad= 2\gamma_1^2 L^3 + \tfrac{2}{3}\sum_{j=2}^{4}\gamma_j^2 L^5 + 4\gamma_1\gamma_d L^5 + \tfrac{38}{15}\gamma_5^2\, L^7 \le \frac{C}{L^3}\left(\rho_u^2 + \frac{e_u^2}{L^4} + \sum_{j=2}^{4}\mu_j^2\right).\end{aligned} \tag{3.13}$$

Proof. From the calculus of variations when the objective function is an integral equation and the constraints are integrals, the optimization problem can be solved by forming the Lagrangian functional and finding its critical points. Set

$$\begin{aligned}\psi_1(X) &:= \int_{\Omega_L} X(v)\mathrm{d}v\,,\\ \psi_{j+1}(X) &:= \int_{\Omega_L} v_j X(v)\mathrm{d}v \quad \forall j = 1,2,\dots,d,\\ \psi_{d+2}(X) &:= \int_{\Omega_L} |v|^2 X(v)\mathrm{d}v\,,\end{aligned}$$

and define

$$\mathcal{H}(X,X',\boldsymbol{\gamma}) := \mathcal{A}^e(X) + \sum_{i=1}^{d+2}\gamma_i\psi_i(X) = \int_{\Omega_L} h(v,X,X',\boldsymbol{\gamma})\mathrm{d}v.$$

We introduce

$$h(v,X,X',\boldsymbol{\gamma}) := \big(Q_u(f,f)(v) - X(v)\big)^2 + \gamma_1 X(v) + \sum_{j=1}^{d}\gamma_{j+1}v_j X(v) + \gamma_{d+2}|v|^2 X(v).$$

In order to find the critical points one needs to compute $D_X\mathcal{H}$ and $D_{\gamma_j}\mathcal{H}$. The derivatives $D_{\gamma_j}\mathcal{H}$ just retrieve the constraint integrals. For multiple independent variables v_j and a single dependent function $X(v)$ the Euler–Lagrange equations are

$$D_2 h(v,X,X',\boldsymbol{\gamma}) = \sum_{j=1}^{d}\frac{\partial D_3 h}{\partial v_j}(v,X,X',\boldsymbol{\gamma}) = 0\,.$$

We used the fact that h is independent of X'. This gives the following equation for the conservation correction in terms of the Lagrange multipliers:

$$2\big(X(v) - Q_u(f,f)(v)\big) + \gamma_1 + \sum_{j=1}^{d}\gamma_{j+1}v_j + \gamma_{d+2}|v|^2 = 0 \tag{3.14}$$

and, therefore, $Q_c(f,f)(v) = X^\star(v) := Q_u(f,f)(v) - \frac{1}{2}\Big(\gamma_1 + \sum_{j=1}^{d}\gamma_{j+1}v_j + \gamma_{d+2}|v|^2\Big)$.

Let $g(v,\gamma)=\gamma_1+\sum_{j=1}^d \gamma_{j+1}v_j+\gamma_{d+2}|v|^2$. Substituting (3.14) into the constraints $\psi_j(X^\star)=0$ gives

$$
\begin{aligned}
\rho_u &:= \int_{\Omega_L} Q_u(f,f)(v)\mathrm{d}v = \frac{1}{2}\int_{\Omega_L} g(v,\gamma)\mathrm{d}v,\\
\mu_u^j &:= \int_{\Omega_L} v_j Q_u(f,f)(v)\mathrm{d}v = \frac{1}{2}\int_{\Omega_L} v_j g(v,\gamma)\mathrm{d}v, \qquad j=1,2,\cdots,d,\\
e_u &:= \int_{\Omega_L} |v|^2 Q_u(f,f)(v)\mathrm{d}v = \frac{1}{2}\int_{\Omega_L} |v|^2 g(v,\gamma)\mathrm{d}v\,.
\end{aligned}
\tag{3.15}
$$

Identities (3.15) form a system of $d+2$ linear equations with $d+2$ unknown variables that can be uniquely solved. Solving for the critical γ_j,

$$
\begin{aligned}
\gamma_1 &= O_d\rho_u + O_{d+2}e_u\,,\\
\gamma_{j+1} &= O_d\mu_u^j\,, \qquad j=1,2,\ldots,d,\\
\gamma_{d+2} &= O_{d+2}\rho_u + O_{d+4}e_u\,.
\end{aligned}
\tag{3.16}
$$

Hence, relation (3.11) holds. Substituting these values of critical Lagrange multipliers (3.16) into (3.14) gives explicitly the critical $Q_c(f,f)(v) := X^\star(v)$. Moreover, the objective function $\mathcal{A}^e(X)$ can be computed at its minimum as

(3.17)

$$
\begin{aligned}
\mathcal{A}^e(Q_c(f,f)) &= \big\|Q_u(f,f)-Q_c(f,f)\big\|^2_{L^2(\Omega_L)} = \int_{\Omega_L}\big(Q_u(f,f)(v)-X^\star(v)\big)^2\mathrm{d}v\\
&= \tfrac{1}{4}\int_{\Omega_L}\Big(\gamma_1+\sum_{j=1}^d\gamma_{j+1}v_j+\gamma_{d+2}|v|^2\Big)^2\mathrm{d}v\\
&\le \tfrac{d+2}{4}\int_{\Omega_L}\Big(\gamma_1^2+\sum_{j=1}^d(\gamma_{j+1}v_j)^2+\gamma_{d+2}^2|v|^4\Big)\\
&\le C(d)\Big(2\gamma_1^2L^d+(\sum_{j=1}^d\gamma_{j+1}^2)L^{d+2}+\gamma_{d+2}^2L^{d+4}\Big),
\end{aligned}
$$

where $C(d)$ is a universal constant depending on the dimension of the space. Hence, using the relation (3.16) in the right-hand side of (3.17), yields a bound from above to the difference of the conserved and unconserved approximating collision operators

$$
\big\|Q_u(f,f)-Q_c(f,f)\big\|^2_{L^2(\Omega_L)} \le \frac{C(d)}{L^d}\Big(\rho_u^2+\frac{e_u^2}{L^{d+1}}+\sum_{j=2}^{d+1}\mu_j^2\Big)
\tag{3.18}
$$

and, therefore, the Lagrange estimate (3.12) holds. Upon simplification one can obtain a detailed estimate for the three-dimensional case, given by

(3.19)

$$
\begin{aligned}
\big\|Q_u(f,f)-Q_c(f,f)\big\|^2_{L^2(\Omega_L)} &= 2\gamma_1^2L^3+\tfrac{2}{3}(\gamma_2^2+\gamma_3^2+\gamma_4^2)L^5+4\gamma_1\gamma_5L^5+\tfrac{38}{15}\gamma_5^2L^7\\
&\le \frac{C}{L^3}\Big(\rho_u^2+\frac{e_u^2}{L^4}+\sum_{j=2}^4\mu_j^2\Big),
\end{aligned}
$$

which is precisely (3.13). That this critical point is in fact the unique minimizer follows from the strict convexity of $\mathcal{A}^e$. ∎

Similarly, as was also proposed in the simulations of [30], one can form the optimization problem for the inelastic case. The only difference is that now only $(d+1)$-collision invariants are conserved:

Minimization inelastic problem (IE): Minimize in the Banach space

$$\mathcal{B}^i = \left\{ X \in L^2(\Omega_L) : \int_{\Omega_L} X = \int_{\Omega_L} Xv = 0 \right\},$$

the functional

$$\mathcal{A}^i(X) := \int_{\Omega_L} \big(Q_u(f,f)(v) - X\big)^2 \mathrm{d}v. \tag{3.20}$$

As in the elastic case, we state a rather similar analog to the Lagrange estimate for the inelastic collision law. The proof of this statement is similar to the case of elastic interactions, and we leave it to the readers.

LEMMA 3.3 (inelastic Lagrange estimate). *The problem* (3.20) *has a unique minimizer given by*

$$Q_c^{ine}(f,f)(v) := X^\star(v) = Q_u(f,f)(v) - \frac{1}{2}\Big(\gamma_1 + \sum_{j=1}^{d} \gamma_{j+1} v_j\Big). \tag{3.21}$$

The γ_j are Lagrange multipliers associated with the inelastic optimization problem given by

$$\begin{aligned} \gamma_1 &= O_d \rho_u\,, \\ \gamma_{j+1} &= O_{d+2} \mu_u^j\,, \quad j = 1, 2, \ldots, d\,. \end{aligned} \tag{3.22}$$

In particular, for the three-dimensional case the minimized objective function is

$$\mathcal{A}^i(X^\star) = \big\| Q_u(f,f) - Q_c^{ine}(f,f) \big\|^2_{L^2(\Omega_L)} = 2\gamma_1^2 L^3 + \tfrac{2}{3}\big(\gamma_2^2 + \gamma_3^2 + \gamma_4^2\big) L^5\,. \tag{3.23}$$

Conservation correction estimate. We develop here a useful estimate between the unconserved and conserved discrete collisional forms.

Definition. For any fixed $f \in L^2(\Omega_L)$ the *conserved operator* $Q_c(f,f)$ is defined as the minimizer of problem (E) defined by (3.10) (or problem (IE) in the inelastic case defined by (3.21)).

Note that the minimized objective function (3.12) in the elastic optimization problem depends only on the unconserved moments ρ_u, μ_u, and e_u of $Q_u(f,f)$. Since these quantities are expected to be approximations to zero, then the conserved projection operator is a perturbation of $Q_u(f,f)$ by a second order polynomial in the elastic case. Similarly, it is a perturbation by a first order polynomial in the inelastic case.

THEOREM 3.4 (conservation correction estimate/elastic case). *Fix $f \in L^2(\Omega_L)$, then the accuracy of the conservation minimization problem is proportional to the spectral accuracy. That is, for any $k' \geq k \geq 0$ it follows that*

$$\begin{aligned} \big\| \big(Q_c(f,f) - Q_u(f,f)\big) |v|^{\lambda k} \big\|_{L^2(\Omega_L)} &\leq \frac{C\, L^{\lambda k}}{(2\lambda k + d)^{1/2}} \big\| (\mathbf{1} - \Pi_{2L}^N) Q^+(\chi f, \chi f) \big\|_{L^2(\Omega_L)} \\ &\quad + \frac{1}{(2\lambda k + d)^{1/2}} O_{(d/2 + \lambda(k'-k))} \big(m_{k'+1}(f) m_0(f) + Z_{k'}(f) \big)\,, \end{aligned} \tag{3.24}$$

where C is a universal constant and $Z_{k'}(f)$ is defined by

$$Z_{k'}(f) := \sum_{j=0}^{k'-1} \left(\begin{array}{c} k' \\ j \end{array} \right) m_{j+1}(f) m_{k'-j}(f) \tag{3.25}$$

depending on the moments up to order k' (See also Appendix (A.3)*). As before, we are using the shorthand $O_r := O(L^{-r})$.*

Proof. Using Lemma 3.2 for elastic interactions, given a $0 \le k \in \mathbb{R}$, estimate

(3.26)

$$\begin{aligned} \Big\| \left(Q_c(f,f) - Q_u(f,f)\right) |v|^{\lambda k} \Big\|_{L^2(\Omega_L)} &= \Big\| \tfrac{1}{2} \Big(\gamma_1 + \sum_{j=1}^{d} \gamma_{j+1} v_j + \gamma_{d+2} |v|^2 \Big) |v|^{\lambda k} \Big\|_{L^2(\Omega_L)} \\ &\le \frac{C\, L^{\lambda k}}{(2\lambda k + d)^{1/2}} \Big(|\gamma_1| L^{d/2} + |\gamma_j| L^{1+d/2} + |\gamma_{d+2}| L^{2+d/2} \Big). \end{aligned}$$

For any $f \in L^2(\Omega_L)$ the Lagrange multipliers γ_j, $1 \le j \le d+2$, can be estimated by observing that

(3.27)

$$\begin{aligned} \Big| \int_{\Omega_L} Q_u(f,f)(v) \psi(v) \mathrm{d}v \Big| &= \Big| \int_{\Omega_L} \big(Q_u(f,f)(v) \\ &\quad - Q(\chi f, \chi f)(v) \big) \psi(v) \mathrm{d}v - \int_{\mathbb{R}^d \setminus \Omega_L} Q(\chi f, \chi f)(v) \psi(v) \mathrm{d}v \Big| \\ &\le \big\| (\mathbf{1} - \Pi_{2L}^N) Q^+(\chi f, \chi f) \big\|_{L^2(\Omega_L)} \|\psi\|_{L^2(\Omega_L)} + I_\psi \end{aligned}$$

for I_ψ defined by

$$I_\psi := \Big| \int_{\mathbb{R}^d \setminus \Omega_L} Q^+(\chi f, \chi f)(v) \psi(v) \mathrm{d}v - \int_{\mathbb{R}^d \setminus 0.95\Omega_L} Q^-(f(1-\chi), f) \psi(v) \mathrm{d}v \Big|. \tag{3.28}$$

Since

$$\begin{aligned} \| 1 \|_{L^2(\Omega_L)} &\sim L^{d/2}, \\ \| v_j \|_{L^2(\Omega_L)} &\sim L^{d/2+1} \quad \text{for} \quad j = 1, 2, 3, ..., d, \\ \| |v|^2 \|_{L^2(\Omega_L)} &\sim L^{d/2+2}, \end{aligned} \tag{3.29}$$

then, for $\psi = 1, v^j, |v|^2$ with $j = 1, 2, \dots d$, the corresponding estimate (3.27) combined with (3.29) yields the following estimates to the unconserved moments defined in (3.15):

$$\begin{aligned} |\rho_u| &\le C L^{d/2} \big\| (\mathbf{1} - \Pi_{2L}^N) Q^+(\chi f, \chi f) \big\|_{L^2(\Omega_L)} + I_1, \\ |\mu_u^j| &\le C L^{d/2+1} \big\| (\mathbf{1} - \Pi_{2L}^N) Q^+(\chi f, \chi f) \big\|_{L^2(\Omega_L)} + I_{v_j}, \quad j = 1, 2, 3, \dots d, \\ |e_u| &\le C L^{d/2+2} \big\| (\mathbf{1} - \Pi_{2L}^N) Q^+(\chi f, \chi f) \big\|_{L^2 B(\Omega_L)} + I_{|v|^2}. \end{aligned} \tag{3.30}$$

Therefore, using (3.30) in (3.16), Lagrange multipliers are estimated by

$$\begin{aligned} |\gamma_1| &= O_{d/2} \big\| (\mathbf{1} - \Pi_{2L}^N) Q^+(\chi f, \chi f) \big\|_{L^2(\Omega_L)} + O_d I_1 + O_{d+2} I_{|v|^2}, \\ |\gamma_j| &= O_{d/2+1} \big\| (\mathbf{1} - \Pi_{2L}^N) Q^+(\chi f, \chi f) \big\|_{L^2(\Omega_L)} + O_{d+2} I_{v_j}, \quad j = 1, 2, 3, \dots d, \\ |\gamma_{d+2}| &= O_{d/2+2} \big\| (\mathbf{1} - \Pi_{2L}^N) Q^+(\chi f, \chi f) \big\|_{L^2(\Omega_L)} + O_{d+2} I_1 + O_{d+4} I_{|v|^2}. \end{aligned} \tag{3.31}$$

Finally, the Lagrangian critical parameters from (3.26) are estimated by (3.31) to yield

$$\begin{aligned}\big\|\big(Q_c(f,f)-Q_u(f,f)\big)|v|^{\lambda k}\big\|_{L^2(\Omega_L)}=&\frac{C}{(2\lambda k+d)^{1/2}}\Big(L^{\lambda k}\big\|(\mathbf{1}-\Pi_{2L}^N)Q^+(\chi f,\chi f)\big\|_{L^2(\Omega_L)}^2\\&+O_{d/2-\lambda k}\,I_1+O_{d/2+1-\lambda k}\,I_{v_j}\\&+O_{d/2+2-\lambda k}\,I_{|v|^2}\Big).\end{aligned}$$

In order to estimate the second term in the above inequality, the terms I_ψ defined in (3.28) are estimated combining classical moment estimates for binary collisional integrals for elastic interactions with hard potentials as shown in Theorem A.2 in the appendix. In particular, for any $k'\geq 0$ and $\lambda\in[0,2]$

$$\begin{aligned}\max\big\{I_1,L^{-1}I_{v_j},L^{-2}I_{|v|^2}\big\}&\leq CL^{-\lambda k'}\big(m_{k'+1}(\chi f)\,m_0(\chi f)+Z_{k'}(\chi f)\big)\\&\leq CL^{-\lambda k'}\big(m_{k'+1}(f)\,m_0(f)+Z_{k'}(f)\big).\end{aligned}$$

Therefore, a simple calculation shows

$$O_{d/2-\lambda k}I_1+O_{d/2+1-\lambda k}I_{v_j}+O_{d/2+2-\lambda k}I_{|v|^2}=O_{d/2+\lambda(k'-k)}\big(m_{k'+1}(f)m_0(f)+Z_{k'}(f)\big),$$

and so inequality (3.26) holds.

This estimate also follows for the *inelastic collisions* case. Their computations follow in a similar fashion using Lemma 3.3, the Lagrange multipliers (3.22), and the first two inequalities in (3.30). ∎

3.2. Semidiscrete conservation method: Lagrange multiplier method. In this subsection we consider the discrete version of the conservation scheme. For such a discrete formulation, the conservation routine is implemented as a Lagrange multiplier method where the conservation properties of the discrete distribution are set as constraints. Let $M=N^d$, the total number of Fourier modes. For elastic collisions, $\rho=0$, $\mathbf{m}=(m_1,\dots,m_d)=(0,\dots,0)$ and $e=0$ are conserved, whereas for inelastic collisions, $\rho=0$ and $\mathbf{m}=(m_1,\dots,m_d)=(0,\dots,0)$ are conserved. Let $\omega_j>0$ be the integration weights for $1\leq j\leq M$ and define

$$\mathbf{Q}_u=\begin{pmatrix} Q_{u,1} & Q_{u,2} & \cdots & Q_{u,M}\end{pmatrix}^T \tag{3.32}$$

as the distribution vector at the computed time step, and

$$\mathbf{Q}_c=\begin{pmatrix} Q_{c,1} & Q_{c,2} & \cdots & Q_{c,M}\end{pmatrix}^T \tag{3.33}$$

as the corrected distribution vector with the required moments conserved. For the elastic case, let

$$\mathbf{C}^e_{(d+2)\times M}=\begin{pmatrix}\omega_j\\ v_1\,\omega_j\\ \cdots\\ v_d\,\omega_j\\ |v_j|^2\,\omega_j\end{pmatrix},\quad 1\leq j\leq M\,, \tag{3.34}$$

be the integration matrix, where the $w_j, j = 1 \dots M$, are fixed set of quadrature points, and

$$\mathbf{a}^e_{(d+2)\times 1} = \begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t}\rho & \frac{\mathrm{d}}{\mathrm{d}t}m_1 & \cdots & \frac{\mathrm{d}}{\mathrm{d}t}m_d & \frac{\mathrm{d}}{\mathrm{d}t}e \end{pmatrix}^T \tag{3.35}$$

be the vector of conserved quantities. With this notation in mind, the semidiscrete conservation method corresponding to (3.4), (3.5) is written as the constrained optimization problem

(3.36)

find the vector $\mathbf{Q}_c \in \mathbb{R}^M$, *such that it is the unique solution of*

$$\mathcal{A}(\mathbf{Q}_c) = \left\{ \min \|\mathbf{Q}_u - \mathbf{Q}_c\|_2^2 : \mathbf{C}^e\mathbf{Q}_c = \mathbf{a}^e \text{ with } \mathbf{C}^e \in \mathbb{R}^{d+2\times M},\, \mathbf{Q}_u \in \mathbb{R}^M,\, \mathbf{a}^e \in \mathbb{R}^{d+2} \right\}.$$

In order to solve the constrained minimization problem $\mathcal{A}(\mathbf{Q}_c)$, we employ the Lagrange multiplier method proposed by two of the authors [30] in 2009. The proposed algorithm works as follows.

Let $\boldsymbol{\gamma} \in \mathbb{R}^{d+2}$ be the Lagrange multiplier vector. Then the scalar objective function to be optimized is given by

$$\mathrm{L}(\mathbf{Q}_c, \boldsymbol{\gamma}) = \sum_{j=1}^{M} \left| Q_{u,j} - Q_{c,j} \right|^2 + \boldsymbol{\gamma}^T (\mathbf{C}^e\mathbf{Q}_c - \mathbf{a}^e)\,, \tag{3.37}$$

where $\mathbf{C}^e$ is given by the integration matrix that computes the number of collision invariants associated with the conservation problem (i.e., $d+2$ for the elastic case or $d+1$ for the inelastic one). This matrix is independent of the solution and the time parameter. Hence, it can be precomputed and used for different initial data and time steps.

Equation (3.37) can be solved explicitly for the corrected distribution value and the resulting equation of correction be implemented numerically in the code. Indeed, taking the derivative of $\mathrm{L}(\mathbf{Q}_c, \boldsymbol{\gamma})$ with respect to $Q_{c,j}$, for $1 \le j \le M$ and γ_i, for $1 \le i \le d+2$

$$\frac{\partial \mathrm{L}}{\partial Q_{c,j}} = 0\,, \quad j = 1, \dots, M \quad \Rightarrow \quad \mathbf{Q}_c = \mathbf{Q}_u + \frac{1}{2}(\mathbf{C}^e)^T\boldsymbol{\gamma}\,. \tag{3.38}$$

Moreover,

$$\frac{\partial \mathrm{L}}{\partial \gamma_i} = , \quad i = 1, \dots, d+2 \quad \Rightarrow \quad \mathbf{C}^e\mathbf{Q}_c = \mathbf{a}^e$$

retrieves the constraints. Solving for $\boldsymbol{\gamma}$,

$$\mathbf{C}^e(\mathbf{C}^e)^T\boldsymbol{\gamma} = 2(\mathbf{a}^e - \mathbf{C}^e\mathbf{Q}_u)\,. \tag{3.39}$$

Now $\mathbf{C}^e(\mathbf{C}^e)^T$ is symmetric and, because $\mathbf{C}^e$ is an integration matrix, it is also positive definite. As a consequence, the inverse of $\mathbf{C}^e(\mathbf{C}^e)^T$ exists and one can compute the value of $\boldsymbol{\gamma}$ simply by

$$\boldsymbol{\gamma} = 2(\mathbf{C}^e(\mathbf{C}^e)^T)^{-1}(\mathbf{a}^e - \mathbf{C}^e\mathbf{Q}_u)\,.$$

Substituting $\boldsymbol{\gamma}$ into (3.38) and recalling that $\mathbf{a}^e = \mathbf{0}$,

$$\begin{aligned}\mathbf{Q}_c &= \mathbf{Q}_u + (\mathbf{C}^e)^T\big(\mathbf{C}^e(\mathbf{C}^e)^T\big)^{-1}(\mathbf{a}^e - \mathbf{C}^e\mathbf{Q}_u)\\ &= \Big[\mathbb{I} - (\mathbf{C}^e)^T\big(\mathbf{C}^e(\mathbf{C}^e)^T\big)^{-1}\mathbf{C}^e\Big]\,\mathbf{Q}_u\\ &=: \Lambda_N(\mathbf{C}^e)\,\mathbf{Q}_u\,, \end{aligned} \tag{3.40}$$

where $\mathbb{I} = N \times N$ is identity matrix. In the following, we call this conservation routine *Conserve*. Thus,

$$Conserve(\mathbf{Q}_u) = \mathbf{Q}_c = \Lambda_N(\mathbf{C}^e)\,\mathbf{Q}_u\,. \tag{3.41}$$

Define D_t to be any time discretization operator of arbitrary order. Then, the discrete problem that we solve reads

$$D_t\mathbf{f} = \Lambda_N(\mathbf{C}^e)\,\mathbf{Q}_u\,. \tag{3.42}$$

Thus, multiplying (3.42) by $\mathbf{C}^e$ it follows the conservation of observables

$$D_t\big(\mathbf{C}^e\mathbf{f}\big) = \mathbf{C}^e D_t\mathbf{f} = \mathbf{C}^e\Lambda_N(\mathbf{C}^e)\,\mathbf{Q}_u = 0\,, \tag{3.43}$$

where we used the commutation $\mathbf{C}^e D_t = D_t\mathbf{C}^e$ valid since $\mathbf{C}^e$ is independent of time; see [30] for additional comments.

4. A priori estimates, propagation of moments, and L^2_k-norm. In this section we prove L^1_k and L^2_k estimates for the approximation solutions $\{g_N\}$ of the problem (3.7) in the elastic case. For this purpose, we use several well-known results that require different integrability properties for the angular kernel b. Thus, we will work with a bounded b to avoid as many technicalities as possible and remarking that a generalization for $b \in L^1(\mathbb{S}^{d-1})$ can be made at the cost of technical work [1, 5, 43]. For technical reasons this assumption helps since estimates for the gain part of the collision operator become bilinear, that is, the role of the inputs can be interchanged without essentially altering the constants in the estimates. We also restrict ourselves to the case of variable hard potentials and hard spheres $\lambda \in (0, 1]$ and remark that the theory for Maxwell molecules $\lambda = 0$ needs a different approach.

Recall that we have imposed conservation of mass, momentum, and energy by building the operator $Q_c(g, g)$ with a constrained minimization procedure. Thus,

$$\int_{\Omega_L} g(t, v)\psi(v)\mathrm{d}v = \int_{\Omega_L} g_0(v)\psi(v)\mathrm{d}v$$

for any collision invariant $\psi(v) = \{1, v, |v|^2\}$. However, due to velocity-mode truncation, the approximating solution g in general may be negative in some small portions of the domain. This is one of the important technical difficulties that we have to overcome.

Before starting with the calculations recall the smoothing property of the gain collision operator Q^+ given in [17, Theorem 2.1],

$$\|Q^+(f, f)\|_{\dot{H}^{(d-1)/2}(\mathbb{R}^d)} \le C\|b\|_{L^2(\mathbb{S}^{d-1})}\|f\|^2_{L^2_{1+\lambda^{-1}}(\mathbb{R}^d)}\,, \tag{4.1}$$

where C is a universal constant only depending on the space dimension d.

Therefore, recalling that $\mathrm{supp}(Q^+(\chi g, \chi g)) \subset \Omega_{2L}$ and using Parseval's theorem, it follows that (for $a > 0$)

$$\begin{aligned}\left\|(\mathbf{1}-\Pi_{2L}^N)Q^+(\chi g,\chi g)\right\|^2_{L^2(\Omega_{2L})} &= \sum_{|k|\geq N}\left|\widehat{Q^+(\chi g,\chi g)}(\xi_k)\right|^2\\ &= \sum_{|k|\geq N}\frac{1}{|\xi_k|^{2a}}\left|\widehat{(-\Delta)^{a/2}Q^+}(\chi g,\chi g)(\xi_k)\right|^2\\ &\lesssim \frac{1}{N^{2a}}\sum_{|k|\geq N}\left|\widehat{(-\Delta)^{a/2}Q^+}(\chi g,\chi g)(\xi_k)\right|^2\\ &\leq \frac{1}{N^{2a}}\left\|(-\Delta)^{a/2}Q^+(\chi g,\chi g)\right\|_{L^2(\Omega_{2L})}.\end{aligned}$$

As a conclusion of the previous two facts, choosing $a=\frac{d-1}{2}$, we obtain an important estimate used in the following arguments:

$$\left\|(\mathbf{1}-\Pi_{2L}^N)Q^+(\chi g,\chi g)\right\|_{L^2(\Omega_{2L})}\leq\frac{C}{N^{\frac{d-1}{2}}}\|\chi g\|^2_{L^2_{1+\lambda^{-1}}(\Omega_L)}, \tag{4.2}$$

since χg vanishes outside a compactly supported set in Ω_L, so we make use of the extension norm identity (2.28) that asserts $\|\chi g\|_{L_{1+\lambda^{-1}}(\Omega_{2L})}=\|\chi g\|_{L_{1+\lambda^{-1}}(\Omega_L)}$.

4.1. Differential estimates for moments of the scheme. In the analysis of the following two sections, we assume that a semidiscrete solution $g\in\mathcal{C}\big(0,T;L^2(\Omega_L)\big)$ for problem (3.7) where initial condition $g_0\in L^2(\Omega_L)$ exists satisfying condition (3.8). We denote $T_\epsilon\geq 0$ any time that the smallness relation for the negative mass and energy of $g(t,v)$ and its boundedness in L^2 holds:

$$\sup_{t\in[0,T_\epsilon]}\left(\epsilon(t):=\frac{\int_{\{g<0\}}|g(t,v)|\langle v\rangle^2\mathrm{d}v}{\int_{\{g\geq 0\}}g(t,v)\langle v\rangle^2\mathrm{d}v}\right)\leq\epsilon,\qquad\sup_{t\in[0,T_\epsilon]}\|g(t,\cdot)\|_{L^2(\Omega_L)}<\infty \tag{4.3}$$

for some fixed $\epsilon>0$ sufficiently small to be specified below in (4.9). Observe that the *conservation scheme* and this assumption imply that semidiscrete moments up to order 2 are controlled by the initial datum. Indeed, for $k=\{0,2\}$

$$\begin{aligned}\int_{\Omega_L}|g||v|^k &= \int_{\Omega_L}g|v|^k-2\int_{\Omega_L}g^-|v|^k=\int_{\Omega_L}g_0|v|^k-2\int_{\Omega_L}g^-|v|^k\\ &\leq\int_{\Omega_L}g_0|v|^k+2\epsilon\int_{\Omega_L}g^+|v|^k\leq\int_{\Omega_L}g_0|v|^k+2\epsilon\int_{\Omega_L}|g||v|^k.\end{aligned}$$

Indeed, choosing $\epsilon\leq 1/4$ one obtains

$$\int_{\Omega_L}|g(t,v)||v|^k\mathrm{d}v\leq 2\int_{\Omega_L}g_0|v|^k\mathrm{d}v\quad\text{for } t\in[0,T_\epsilon],\quad k=1,2. \tag{4.4}$$

Remark. Conditions (3.8) and (4.3) are a sort of stability condition for the semidiscrete scheme.

Next, we start getting estimates for the discrete conserved form (3.7). Indeed, taking the right-hand side from (3.3) combined with those of (3.1), (3.2), the discrete equation (3.7) for the numerical scheme can be written in $(0,T_\varepsilon]\times\Omega_L$ as

$$\begin{aligned}\frac{\mathrm{d}g}{\mathrm{d}t}=Q_c(g,g)&=Q_c(g,g)-Q_u(g,g)+Q(\chi g,\chi g)\\ &\quad-(\mathbf{1}-\Pi_{2L}^N)Q^+(\chi g,\chi g)-Q^-((1-\chi)g,\chi g),\end{aligned} \tag{4.5}$$

as the second term in this equation is actually null.

In the next lemma we prepare estimates to obtain an ordinary differential inequality that will yield uniform estimates to the numerical moments of the semidiscrete solutions corresponding to the initial value problem (3.7).

LEMMA 4.1. *Let g be the solution of the numerical scheme satisfying* (4.3) *and set $k \geq k_0 \geq 2$. Then,*

(4.6)
$$\frac{d}{dt}m_k(g) \leq C_k\left(m_0(g_0)+m_k(g)\right) - \frac{\mu_{\frac{\lambda}{2}}\, m_0(g_0)}{4}\, m_{k+1}(g) + C\,\frac{L^{\lambda k+d/2}}{N^{\frac{d-1}{2}}}\|g\|^2_{L^2_{1+\lambda^{-1}}(\Omega_L)}$$

for any $g_0(v)$ satisfying the energy ratio condition (3.8). *In addition, $\mu_{\frac{\lambda}{2}}$, k_0 are constants given by* (4.8) *and* (4.12), *respectively, defined in the proof of this lemma.*

Proof. We fix $k>0$ and $L>0$ and keep in mind that g_0 has support in Ω_L and, thus, possesses moments of any order. Multiply (4.5) by $\operatorname{sgn}(g)|v|^{\lambda k}$ and integrate in Ω_L

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_L}\left|g(v)\right||v|^{\lambda k}\mathrm{d}v\\
&\quad=\int_{\Omega_L} Q(\chi g,\chi g)(v)\operatorname{sgn}(g)(v)\,|v|^{\lambda k}\mathrm{d}v - \int_{\Omega_L} Q^-((1-\chi)g,\chi g)(v)\operatorname{sgn}(g)(v)\,|v|^{\lambda k}\mathrm{d}v\\
&\qquad+\int_{\Omega_L}\left(Q_c(g,g)(v)-Q_u(g,g)(v)\right)\operatorname{sgn}(g)\,|v|^{\lambda k}\mathrm{d}v\\
&\qquad-\int_{\Omega_L}(\mathbf{1}-\Pi^N_{2L})Q^+(\chi g,\chi g)(v)\operatorname{sgn}(g)\,|v|^{\lambda k}\\
&\qquad\leq\int_{\Omega_L} Q^+(|\chi g|,|\chi g|)(v)|v|^{\lambda k}\mathrm{d}v - \int_{\Omega_L} Q^-(g,\chi g)(v)\mathrm{sgn}(\chi g)(v)|v|^{\lambda k}\mathrm{d}v\\
&\qquad+\left\|\left(Q_c(g,g)-Q_u(g,g)\right)|v|^{\lambda k}\right\|_{L^1(\Omega_L)}+\left\|(\mathbf{1}-\Pi^N_{2L})Q^+(\chi g,\chi g)|v|^{\lambda k}\right\|_{L^1(\Omega_L)}.
\end{aligned}$$

We estimate each term starting with the loss collision operator. Use $g=|g|-2g^-$ to conclude that

$$\begin{aligned}
\int_{\Omega_L} Q^-(g,\chi g)(v)\mathrm{sgn}(g)(v)|v|^{\lambda k}\mathrm{d}v \geq{}& \int_{\Omega_L}|g(v)||v|^{\lambda k}\int_{\mathbb{R}^d}|\chi g(v_*)||v-v_*|^{\lambda}\mathrm{d}v_*\mathrm{d}v\\
&-C_{d,\lambda}\,\epsilon\,\|g_0\|_{L^1_{2\lambda^{-1}}(\Omega_L)}\left(m_{k+1}(g)+m_k(g)\right),
\end{aligned}$$

where ϵ is the bound from the energy quotient from (4.3). Whence,

(4.7)
$$\begin{aligned}
&\int_{\Omega_L} Q^+(|\chi g|,|\chi g|)(v)|v|^{\lambda k}\mathrm{d}v - \int_{\Omega_L} Q^-(g,\chi g)(v)\mathrm{sgn}(g)(v)|v|^{\lambda k}\mathrm{d}v\\
&\qquad\leq\int_{\Omega_L} Q(|\chi g|,|\chi g|)(v)|v|^{\lambda k}\mathrm{d}v\\
&\qquad-\int_{\Omega_L} Q((1-\chi)|g|,|\chi g|)(v)|v|^{\lambda k}\mathrm{d}v + C_{d,\lambda}\,\epsilon\,\|g_0\|_{L^1_{2\lambda^{-1}}(\Omega_L)}\left(m_{k+1}(g)+m_k(g)\right).
\end{aligned}$$

Using the *conservative property of the scheme* it follows from the discussion in [13, 6] that

$$\int_{\Omega_L} Q(|\chi g|,|\chi g|)(v)|v|^{\lambda k}\mathrm{d}v \le \int_{\mathbb{R}^d} Q(|\chi g|,|\chi g|)(v)|v|^{\lambda k}\mathrm{d}v \le Z_k(g) - \mu_k\, m_0(g_0)\, m_{k+1}(\chi g), \quad \tfrac{2}{\lambda} < k \in \mathbb{Z},$$

where $Z_k(g)$ depends on the moments of g of order *less than or equal to* k and $\mu_k \nearrow 1$ as $k \to \infty$ being a universal parameter given by

$$\mu_k := 1 - \frac{1}{2^k}\int_{\mathbb{S}^{d-1}} (1+\hat{u}\cdot\sigma)^k\, b(\hat{u}\cdot\sigma)\,\mathrm{d}\sigma \in (0,1)\,. \tag{4.8}$$

We refer to [13, Lemma 3] for details and proof. Choose

$$\epsilon \le \min\left\{\tfrac{1}{4}\,,\, \mu_{\frac{\lambda}{2}}\, \frac{m_0(g_0)}{2\,C_{d,\lambda}\,\|g_0\|_{L^1_{2\lambda^{-1}}(\Omega_L)}}\right\} \tag{4.9}$$

in (4.7) to conclude that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}m_k(g) &\le Z_k(g) - \frac{1}{2}\mu_{\frac{\lambda}{2}}\, m_0(g_0)\, m_{k+1}(g)\\ &\quad + \big\|(Q_c(g,g) - Q_u(g,g))|v|^{\lambda k}\big\|_{L^1(\Omega_L)} + \big\|(\mathbf{1}-\Pi^N_{2L})Q^+(\chi g,\chi g)|v|^{\lambda k}\big\|_{L^1(\Omega_L)}\,.\end{aligned} \tag{4.10}$$

Using the Cauchy–Schwarz inequality and (3.24) from Theorem 3.4, it follows, for any $k' \ge k \ge 0$, that

$$\begin{aligned}\big\|(Q_c(g,g) - Q_u(g,g))|v|^{\lambda k}\big\|_{L^1(\Omega_L)} &\le L^{d/2}\big\|(Q_c(g,g) - Q_u(g,g))|v|^{\lambda k}\big\|_{L^2(\Omega_L)}\\ &\le \frac{C\,L^{\lambda k + d/2}}{(2\lambda k + d)^{1/2}}\big\|(\mathbf{1}-\Pi^N_{2L})Q^+(\chi g,\chi g)\big\|_{L^2(\Omega_L)}\\ &\quad + \frac{O(L^{-\lambda(k'-k)})}{(2\lambda k+d)^{1/2}}\big(m_{k'+1}(g)\; m_0(g_0) + Z_{k'}(g)\big)\,.\end{aligned}$$

Therefore, after choosing $k' = k > 2$, one concludes that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}m_k(g) &\le 2\,Z_k(g) - \left(\frac{1}{2}\,\mu_{\frac{\lambda}{2}}\, m_0(g_0) - \frac{\mathbf{C}}{(2\lambda k+d)^{1/2}}\right) m_{k+1}(g)\\ &\quad + CL^{\lambda k + d/2}\big\|(\mathbf{1}-\Pi^N_{2L})Q^+(\chi g,\chi g)\big\|_{L^2(\Omega_L)}\\ &\le C_k\left(m_0(g_0) + m_k(g)\right) - \frac{1}{4}\mu_{\frac{\lambda}{2}}\, m_0(g_0)\, m_{k+1}(g)\\ &\quad + CL^{\lambda k + d/2}\big\|(\mathbf{1}-\Pi^N_{2L})Q^+(\chi g,\chi g)\big\|_{L^2(\Omega_L)}\,,\end{aligned} \tag{4.11}$$

where $\mathbf{C}$ is a constant independent of k and λ.

In the last inequality we used the classical fact that $Z_k \le C_k\left(m_0(g) + m_k(g)\right)$ for some large constant C_k depending only on k. We also chose k sufficiently large to make the largest moment an absorption term,

$$k \ge k_0 := \frac{1}{2\lambda}\left(\frac{\mathbf{C}}{\mu_{\frac{\lambda}{2}} m_0(g_0)}\right)^2 - \frac{d}{2\lambda} \ge 2. \tag{4.12}$$

Finally, we use estimate (4.2) in (4.11) to obtain the semidiscrete moment ordinary differential inequality (4.6). ∎

LEMMA 4.2 (lower bound estimate). *Let $h(v)$ be a function satisfying* (4.3) *for $\epsilon < 1/2$. Assume also that $\int_{\mathbb{R}^d} h(w)\, w\, dw = 0$ and that*

$$m_\mu := \int_{\mathbb{R}^d} |h(w)||w|^{2+\mu} dw < \infty\,, \quad \mu > 0. \tag{4.13}$$

Then,

$$\left(h * |u|^\lambda\right)(v) \geq \frac{C(h)\,\langle v\rangle^\lambda}{\max\left\{1, m_\mu^{(2-\lambda)/\mu}\right\}} \tag{4.14}$$

with $C(h) > 0$ depending only on the mass and energy of h.

Proof. Notice that in the ball $B(0,r)$ one has for any $R > 0$ and $\mu > 0$,

$$\begin{aligned}\int_{|v-w|\leq R} h(w)|v-w|^2 \mathrm{d}w &= \int_{\mathbb{R}^d} h(w)|v-w|^2 \mathrm{d}w - \int_{|v-w|\geq R} h(w)|v-w|^2 \mathrm{d}w \\ &\geq C(h)\,\langle v\rangle^2 - \frac{1}{R^\mu}\int_{|v-w|\geq R} |h(w)||v-w|^{2+\mu} \mathrm{d}w\,.\end{aligned} \tag{4.15}$$

For the last inequality we expanded the square in the integral of the right side and used the fact that the momentum of g is zero. We use in the right side integral of (4.15) the inequality $|v-w| \leq \langle v\rangle\langle w\rangle$ and the fact that $m_\mu < \infty$ to obtain

$$\int_{|v-w|\leq R} h(w)|v-w|^2 \mathrm{d}w \geq C(h)\langle v\rangle^2 - \frac{m_\mu}{R^\mu}\langle v\rangle^{2+\mu} \geq \frac{C(h)}{2}\langle v\rangle^2 \quad \forall\, v \in B(0,r)\,,$$

provided

$$R := \left(2m_\mu / C(h)\right)^{1/\mu} r. \tag{4.16}$$

Therefore, using the control (4.3)

$$\begin{aligned}\int_{\mathbb{R}^d} h(w)|v-w|^\lambda \mathrm{d}w &= \int_{\mathbb{R}^d} |h(w)||v-w|^\lambda \mathrm{d}w - 2\int_{\{h<0\}} |h(w)||v-w|^\lambda \mathrm{d}w \\ &\geq (1-2\epsilon)\int_{\mathbb{R}^d} |h(w)||v-w|^\lambda \mathrm{d}w \geq (1-2\epsilon)\int_{|v-w|\leq R} |h(w)||v-w|^\lambda \mathrm{d}w \\ &\geq \frac{1-2\epsilon}{R^{2-\lambda}}\int_{|v-w|\leq R} |h(w)||v-w|^2 \mathrm{d}w \geq \frac{1-2\epsilon}{2R^{2-\lambda}}\, C(h)\langle v\rangle^2\,,\end{aligned}$$

valid for any $v \in B(0,r)$ and provided $\epsilon < \frac{1}{2}$. Moreover, for any $\lambda \in (0,1]$

$$\begin{aligned}\int_{\mathbb{R}^d} h(w)|v-w|^\lambda \mathrm{d}w &\geq (1-2\epsilon)\int_{\mathbb{R}^d} |h(w)||v-w|^\lambda \mathrm{d}w \\ &\geq (1-2\epsilon)\|h\|_{L^1_{2\lambda-1}}\left(|v|^\lambda - 2\right).\end{aligned}$$

As a consequence,

$$\int_{\mathbb{R}^d} h(w,t)|v-w|^\lambda \mathrm{d}w \geq (1-2\epsilon)\left(\frac{C(h)}{2R^{2-\lambda}}\,\mathbf{1}_{B(0,r)} + \|h\|_{L^1_{2\lambda-1}}\left(|v|^\lambda - 2\right)\mathbf{1}_{B(0,r)^c}\right). \tag{4.17}$$

Inequality (4.14) follows from (4.17) choosing $r = 3^{1/\lambda}$ in definition (4.16) of R. ∎

4.2. Time differential estimates for the L^2_k-norm of the conservative semidiscrete scheme. The lower bound on the collision operator given in Lemma 4.2 will allow us to control the L^2_k-norms of g. Multiplying (4.5) by $g\langle v\rangle^{2\lambda k}$ and integrating on Ω_L one has

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|g\|^2_{L^2_k(\Omega_L)} = \int_{\Omega_L}\langle v\rangle^{2\lambda k} g\, Q^+(\chi g,\chi g)\mathrm{d}v - \int_{\Omega_L}\langle v\rangle^{2\lambda k} g\, Q^-(g,\chi g)\mathrm{d}v\\
&+\int_{\Omega_L}\langle v\rangle^{2\lambda k} g\,\big(Q_c(g,g)-Q_u(g,g)\big)\mathrm{d}v - \int_{\Omega_L}\langle v\rangle^{2\lambda k} g\,\big(\mathbf{1}-\Pi^N_{2L}\big)Q^+(\chi g,\chi g)\mathrm{d}v\\
&\le \int_{\Omega_L}\langle v\rangle^{2\lambda k} g\, Q^+(\chi g,\chi g)\mathrm{d}v - \int_{\Omega_L}\langle v\rangle^{2\lambda k} g\, Q^-(g,\chi g)\mathrm{d}v\\
&+\bigg(\big\|(Q_c(g,g)-Q_u(g,g))|v|^{\lambda k}\big\|_{L^2(\Omega_L)}\\
&+\big\|(\mathbf{1}-\Pi^N_{2L})Q^+(\chi g,\chi g)\,|v|^{\lambda k}\big\|_{L^2(\Omega_L)}\bigg)\|g\|_{L^2_k(\Omega_L)}.
\end{aligned}$$

Using smoothing properties of the gain collision operator (see Theorem A.6 in the appendix or refer to [43, 5]), the lower bound control (4.14), and noticing that $C(g) = C(g_0)$ due to the conservation routine, it follows that

$$\begin{aligned}
&\int_{\Omega_L}\langle v\rangle^{2\lambda k} g\, Q^+(\chi g,\chi g)\mathrm{d}v - \int_{\Omega_L}\langle v\rangle^{2\lambda k} g\, Q^-(g,\chi g)\mathrm{d}v\\
&\qquad\le \bigg(\frac{\max\big\{1, m_\mu^{(2-\lambda)/\mu}\big\}}{C(g_0)}\bigg)^{\theta_1}\|g\|^{\theta_2}_{L^1_k(\Omega_L)}\|g\|^{1+1/d}_{L^2_k(\Omega_L)}\\
&\qquad- C(g_0)\bigg(\frac{1}{\max\big\{1, m_\mu^{(2-\lambda)/\mu}\big\}} - \frac{\mathbf{C}}{L^{2-\lambda}}\bigg)\|g\|^2_{L^2_{k+1/2}(\Omega_L)}
\end{aligned}$$

with constant $C(g_0)$ depending only on mass and energy, m_μ defined in (4.13), and some universal $\theta_1 > 1, \theta_2 > 1$. Meanwhile, again using estimates from Theorem 3.4, the rest of the terms can be controlled by

$$\begin{aligned}
&\big\|(Q_c(g,g)-Q_u(g,g))|v|^{\lambda k}\big\|_{L^2(\Omega_L)} + \big\|(\mathbf{1}-\Pi^N_{2L})Q^+(\chi g,\chi g)\,|v|^{\lambda k}\big\|_{L^2(\Omega_L)}\\
&\qquad\le L^{\lambda k}\big\|(1-\Pi^N_{2L})Q^+(\chi g,\chi g)\big\|_{L^2(\Omega_L)} + O_{d/2}\big(m_{k+1}(g)\,m_0(g) + Z_k(g)\big)
\end{aligned}$$

with O_r defined in (3.9). Therefore, we conclude, provided $L \ge 2\max\big\{1, m_\mu^{1/\mu}\big\}$, that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|g\|_{L^2_k(\Omega_L)} &\le \bigg(\frac{\max\big\{1, m_\mu^{(2-\lambda)/\mu}\big\}}{C(g_0)}\bigg)^{\theta_1}\|g\|^{\theta_2}_{L^1_k(\Omega_L)}\|g\|^{1/d}_{L^2_k(\Omega_L)}\\
&\quad- \frac{C(g_0)}{\max\big\{1, m_\mu^{(2-\lambda)/\mu}\big\}}\|g\|_{L^2_{k+1/2}(\Omega_L)}\\
&\quad+ L^{\lambda k}\big\|(1-\Pi^N_{2L})Q^+(\chi g,\chi g)\big\|_{L^2(\Omega_L)}\\
&\quad+ O_{d/2}\big(m_{k+1}(g)\,m_0(g)+Z_k(g)\big).
\end{aligned}\tag{4.18}$$

Plugging (4.2) into (4.18) proves the first part of the following lemma.

LEMMA 4.3. *Fix $k \ge 0$ and $\mu > 0$ and assume g is a solution of the numerical scheme satisfying* (4.3) *for a small proportion $\epsilon \le \epsilon(g_0)$ and cutoff domain $L \ge 2\max\big\{1, m_\mu^{1/\mu}\big\}$. Then, the following differential inequality holds:*

$$\frac{d}{dt}\|g\|_{L^2_k(\Omega_L)} \le \left(\frac{\max\left\{1, m_\mu^{(2-\lambda)/\mu}\right\}}{C(g_0)}\right)^{\theta_1} \|g\|^{\theta_2}_{L^1_k(\Omega_L)}\|g\|^{1/d}_{L^2_k(\Omega_L)} - \frac{C(g_0)}{\max\left\{1, m_\mu^{(2-\lambda)/\mu}\right\}}\|g\|_{L^2_{k+1/2}(\Omega_L)} + C\,\frac{L^{\lambda k}}{N^{\frac{d-1}{2}}}\|g\|^2_{L^2_{1+\lambda^{-1}}(\Omega_L)} + O_{d/2}\Big(m_{k+1}(g)\,m_0(g) + Z_k(g)\Big) \tag{4.19}$$

for some universal $\theta_1, \theta_2 > 1$ and O_r defined in (3.9)*. Moreover, the negative part of g satisfies*

$$\frac{d}{dt}\|g^-\|_{L^2(\Omega_L)} \le C\|g_0\|_{L^1_2(\Omega_L)}\|g^-\|_{L^2(\Omega_L)} + \frac{C}{N^{\frac{d-1}{2}}}\|g\|^2_{L^2_{1+\lambda^{-1}}(\Omega_L)} + O_{d/2+\lambda(k-1)}m_k(g)m_0(g_0)\,. \tag{4.20}$$

Proof. For the part related to the negative mass, note that, writing $g = g^+ + g^-$, it follows that

$$\begin{aligned} Q^+(g,g)\,g\,\mathbf{1}_{\{g\le 0\}} &= \big(Q^+(g^+,g^+) + Q^+(g^+,g^-) + Q^+(g^-,g^+) + Q^+(g^-,g^-)\big)g\,\mathbf{1}_{\{g\le 0\}} \\ &\le \big(Q^+(g^+,g^-) + Q^+(g^-,g^+)\big)\,g\,\mathbf{1}_{\{g\le 0\}}\,. \end{aligned} \tag{4.21}$$

Thus, using Young's inequality [3, 2, 43] one concludes that

$$\begin{aligned} \int_{\Omega_L} Q^+(g,g)\,g\,\mathbf{1}_{\{g\le 0\}}\,\mathrm{d}v &\le \int_{\Omega_L}\big(Q^+(g^+,g^-) + Q^+(g^-,g^+)\big)\,g\,\mathbf{1}_{\{g\le 0\}}\,\mathrm{d}v \\ &\le C\,\|b\|_\infty\|g^+\|_{L^1_1(\Omega_L)}\|g^-\|^2_{L^2(\Omega_L)} \le C\,\|g_0\|_{L^1_{2\lambda^{-1}}(\Omega_L)}\,\|g^-\|^2_{L^2(\Omega_L)}\,. \end{aligned}$$

In this last inequality it was important the bilinear estimates for Q^+ be valid for $b \in L^\infty$. Recall, additionally, that Lemma 4.2 implies

$$\int_{\Omega_L} Q^-(g,g)\,g\,\mathbf{1}_{\{g\le 0\}}\,\mathrm{d}v \ge \frac{C(g_0)}{\max\left\{1, m_\mu^{(2-\lambda)/\mu}\right\}}\|g^-\|^2_{L^2_{1/2}(\Omega_L)} \ge 0.$$

As a consequence, multiplying (4.5) by g^-, integrating in Ω_L, and invoking Theorem 3.4 with $k' = k-1$ and $k = 0$, one concludes that

$$\begin{aligned} \frac{d}{dt}\|g^-\|_{L^2(\Omega_L)} &\le C\|g_0\|_{L^1_{2\lambda^{-1}}(\Omega_L)}\|g^-\|_{L^2(\Omega_L)} \\ &\quad + C\,\big\|(\mathbf{1} - \Pi^N_{2L})Q^+(\chi g, \chi g)\big\|_{L^2(\Omega_L)} + O_{d/2+\lambda(k-1)}m_k(g)m_0(g_0)\,. \end{aligned}$$

The proof follows after plugging (4.2) into this estimate. ∎

4.3. Uniform propagation of moments and L^2_k-norms. Now we are ready to prove uniform propagation of the scheme provided the requirement on the negative mass (4.3) is met for $0 < \epsilon \le \epsilon(g_0)$. Since Lemmas 4.1 and 4.3 hold for the aforementioned conditions on $\epsilon(g_0)$, one has the following two estimates on the k-moment and the L^2-norm:

$$\frac{\mathrm{d}}{\mathrm{d}t}m_k(g) \le C_k\left(m_0(g_0) + m_k(g)\right) - \frac{1}{4}\mu_{\frac{\lambda}{2}}\,m_0(g_0)\,m_{k+1}(g) + C\,\frac{L^{\lambda k + d/2}}{N^{\frac{d-1}{2}}}\|g\|^2_{L^2_{1+\lambda^{-1}}(\Omega_L)}\,, \quad k \ge k_0\,,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\|g\|_{L^2(\Omega_L)} \le \left(\frac{\max\left\{1, m_\mu^{(2-\lambda)/\mu}\right\}}{C(g_0)}\right)^{\theta_1} \|g\|_{L^1(\Omega_L)}^{\theta_2} \|g\|_{L^2(\Omega_L)}^{1/d} - \frac{C(g_0)}{\max\left\{1, m_\mu^{(2-\lambda)/\mu}\right\}} \|g\|_{L^2_{1/2}(\Omega_L)} + \frac{C}{N^{\frac{d-1}{2}}} \|g\|^2_{L^2_{1+\lambda^{-1}}(\Omega_L)} + O_{d/2}\|g_0\|^2_{L^1_2(\Omega_L)}\,.$$

Note that using Young's inequality,

$$\left(\frac{\max\left\{1, m_\mu^{(2-\lambda)/\mu}\right\}}{C(g_0)}\right)^{\theta_1} \|g\|_{L^1(\Omega_L)}^{\theta_2} \|g\|_{L^2(\Omega_L)}^{1/d} \le C_1(g_0) + C_2(g_0) m_\mu^{\theta_1(1+d'/d)(2-\lambda)/\mu} + \frac{C(g_0)}{2\max\left\{1, m_\mu^{(2-\lambda)/\mu}\right\}} \|g\|_{L^2(\Omega_L)}\,.$$

Now, choose $\mu = \lambda k - 2$, so that $m_\mu = m_k(g)$, and then take $k \ge k_0$ such that $\theta_1(1 + d'/d)(2-\lambda)/\mu \le 1$. Then, adding previous two differential equations, one has

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(m_k(g) + \|g\|_{L^2(\Omega_L)}\big) \le \left(C_k(g_0) - c(g_0) m_k^{1+1/k}(g) - \frac{C(g_0)}{\max\left\{1, m_k^{(2-\lambda)/\mu}\right\}} \|g\|_{L^2(\Omega_L)}\right) + \frac{C\, L^{\lambda(k+2)+d/2+2}}{N^{\frac{d-1}{2}}} \|g\|^2_{L^2(\Omega_L)}\,;$$

thus, defining $X(t) := m_k(g) + \|g\|_{L^2(\Omega_L)}$ and performing some algebra it follows that

$$\frac{\mathrm{d}X}{\mathrm{d}t} \le \max\left\{1, m_k^{(2-\lambda)/\mu}\right\}\left(C_k(g_0) - c(g_0)X + \frac{C\, L^{\lambda(k+2)+d/2+2}}{N^{\frac{d-1}{2}}} X^{2+(2-\lambda)/\mu}\right). \tag{4.22}$$

With this estimate we are in position to prove the following proposition.

PROPOSITION 4.4. *Fix $k \ge k_*$ and assume g is a solution of the numerical scheme satisfying* (4.3) *for $0 < \epsilon \le \epsilon(g_0)$ with cutoff domain $L \ge 2\max\{1, m_k^{1/(\lambda k-2)}\}$. Then, there exists a threshold $\eta(g_0) > 0$ depending only on g_0 such that if*

$$L^{\lambda(k+2)+d/2+2} N^{(-d+1)/2} \le \eta(g_0),$$

then

$$\sup_{t\ge 0} m_k(g) \le \max\left\{C_k(g_0), m_k(g_0), \|g_0\|_{L^2(\Omega_L)}\right\} =: \mathfrak{c}_1^k(g_0)\,, \quad \text{and}$$

$$\sup_{t\ge 0} \|g\|_{L^2_{k'}(\Omega_L)} \le \max\left\{C_{k'}(g_0), m_{k'+1}(g_0), \|g_0\|_{L^2_{k'}(\Omega_L)}\right\} =: \mathfrak{c}_2^{k'}(g_0) \qquad \forall\, 0 \le k' \le k-1\,.$$

Here $k_* \ge k_0$ is such that $\theta_1(1 + d'/d)(2-\lambda)/(\lambda k_* - 2) \le 1$ and $C_k(g_0)$ a constant depending on mass and energy of g_0 and k.

Proof. Consider the polynomial $p(x) = C_k(g_0) - c(g_0)x + C\,\eta\, x^{2+(2-\lambda)/\mu}$. Note that for sufficiently small η, depending only on $k \ge k_* \ge 2$ and the mass and energy of

g_0, this polynomial has two positive roots r_1 and r_2. As η vanishes, $r_1 \searrow C_k(g_0)/c(g_0)$ and $r_2 \nearrow \infty$. Thus, choose $0 < \eta$ sufficiently small such that

$$m_k(g_0) + \|g_0\|_{L^2(\Omega_L)} < r_2\,,$$

then, the differential inequality (4.22) written as

$$\frac{\mathrm{d}X}{\mathrm{d}t} \le \max\left\{1, m_k^{(2-\lambda)/\mu}\right\} p(X)$$

for $\frac{L^{\lambda(k+2)+d/2+2}}{N^{(d-1)/2}} \le \eta$ implies that

$$\sup_{t\ge 0} X(t) \le \max\left\{C_k(g_0), X(0)\right\}.$$

This proves the first inequality of the statement and the propagation of $\|g\|_{L^2(\Omega_L)}$. Provided the latter, we use Lemma 4.3 to conclude the second statement. ∎

5. Existence and regularity of the scheme.

5.1. Existence. Now we are ready, thanks to the estimates of the previous section, to produce a proof of existence and uniqueness of the numerical scheme. We assume that $f_0 \in L^2(\mathbb{R}^d)$ is supported in Ω_L, where the choice of the cutoff domain Ω_L was discussed in section 2.2, and that $g_0 = \Pi_L^N f_0$ satisfies

$$\|g_0^-\|_{L^2(\Omega_L)} \approx 0 \tag{5.1}$$

for $N \ge N_0(g_0)$ sufficiently large. Observe also that defining the metric space $\mathcal{X} \subset \mathcal{C}(0,T;L^2(\Omega_L))$ as

$$\mathcal{X} := \left\{ f \in \mathcal{C}(0,T;L^2(\Omega_L)) : \sup_{t\in[0,T]} \|f(t)\|_{L^2(\Omega_L)} \le 2\mathfrak{c}_2^0(g_0)\,,\ \sup_{t\in[0,T]} m_k(f) \le 2\mathfrak{c}_1^k(g_0) \right\},$$

$$\text{and the operator } \mathcal{T} : \mathcal{X} \to \mathcal{C}(0,T;L^2(\Omega_L)) \text{ as } \mathcal{T}(f)(t) = g_0 + \int_0^t Q_c(f)(s)\mathrm{d}s\,,$$

where $k \ge k_* \ge 2$ and $\mathfrak{c}_1^k, \mathfrak{c}_2^0$ are those from Proposition 4.4, one has the estimates for some $a, b_k > 0$,

$$\sup_{t\in[0,T]} \|\mathcal{T}(f) - \mathcal{T}(\tilde f)\|_{L^2(\Omega_L)} \le C(c_1^k, c_2^0)\, L^a\, T \sup_{t\in[0,T]} \|f - \tilde f\,\|_{L^2(\Omega_L)}\,,$$
$$\sup_{t\in[0,T]} m_k\big(\mathcal{T}(f)\big) \le m_k(g_0) + C(c_1^k, c_2^0)\, L^{b_k}\, T\,, \qquad f\,, \tilde f \in \mathcal{X}\,.$$

As a consequence, choosing $T_L := 1/L^{a+b_k}$ for $L \ge L_0(g_0)$ sufficiently large, it follows that $\mathcal{T}$ is a contraction with $\mathcal{T}(\mathcal{X}) \subset \mathcal{X}$. Using the Banach fix point theorem, the scheme has a unique solution in $[0, T_L]$.

THEOREM 5.1. *Set $g_0 = \Pi^N f_0 \in L^1_k \cap L^2(\Omega_L)$ with $k \ge k_* \ge 2$. For any time $T > 0$ and domain cutoff $L \ge L_0(T, g_0) > 0$ there exists a number of modes $N_0(T, L, g_0) > 0$ such that the Problem (3.7) has a unique solution $g \in \mathcal{C}(0,T;L^2(\Omega_L))$ for any $N \ge N_0$ satisfying the estimates*

$$\sup_{t\in[0,T]} \|g\|_{L^2(\Omega_L)} \le c_k^0(g_0), \qquad \sup_{t\in[0,T]} m_k(g) \le \mathfrak{c}_1^k(g_0)\,,$$

and negative mass estimate

$$\begin{aligned}\|g^-(t)\|_{L^2(\Omega_L)} \ &\le C(\mathfrak{c}_1^k, \mathfrak{c}_2^0)\, e^{C\|g_0\|_{L^1_{2/\lambda}(\Omega_L)}t} \\ &\times \Big(\|g_0^-\|_{L^2(\Omega_L)} + O\big(L^{2(1+\lambda)}/N^{(d-1)/2}\big) \\ &+ \|g_0\|_{L^1_2(\Omega_L)}\, O\big(1/L^{d/2+\lambda(k-1)}\big)\Big)\,.\end{aligned}$$

Furthermore, the sequence $\{g = g_N\}$ *formed with initial condition* g_0 *converges strongly in* $\mathcal{C}(0,T;L^2(\Omega_L))$ *to* $\bar{g}$*, the solution of problem*

(5.2)
$$\frac{\partial \bar{g}}{\partial t} = Q^+(\chi\bar{g},\chi\bar{g}) - Q^-(\bar{g},\chi\bar{g}) - \tfrac{1}{2}\Big(\bar{\gamma}_1 + \sum_{j=1}^{d} \bar{\gamma}_{j+1} v_j + \bar{\gamma}_{d+2}|v|^2\Big)\,, \quad (t,v) \in [0,T]\times\Omega_L\,,$$

with initial condition $g_0 = f_0$*. Above, the coefficients* $\bar{\gamma}$ *are given in Lemma* 3.2 *with parameters* (3.11)–(3.15) *evaluated at* $Q^+(\chi\bar{g},\chi\bar{g}) - Q^-(\bar{g},\chi\bar{g})$*.*

Proof. We start with $T > 0$ given, $L > 2\max\{1, (2\,c_k^1(g_0))^{1/\lambda k-2}\}$, and $N > 0$ such that $L^{\lambda(k+2)+d/2+2}N^{(-d+1)/2} \le \eta(g_0)$. We discussed that there exists a unique solution $g \in \mathcal{X}$ in the interval $I_1 := [0, 1/L^{a+b_k}]$. Note that the negative mass of such a solution increases continuously in time. Indeed, multiplying the scheme (4.5) by g^-, it readily follows that

(5.3)
$$\frac{\mathrm{d}}{\mathrm{d}t}\|g^-\|_{L^2(\Omega_L)} \le C(c_1^k, c_2^0)\, L^a \to \|g^-(t_1)\|_{L^2(\Omega_L)} \le \|g^-(t_0)|_{L^2(\Omega_L)} + C(c_1^k, c_2^0)\, L^a (t_1 - t_0)\,.$$

Since $g^-(0) \approx 0$, it means that the requirement on the negative mass of Proposition 4.4 is satisfied in some interval $[0, t_*] \subset I_1$,

$$0 < \epsilon(t) \le \epsilon(g_0)\,, \quad t \in [0, t_*]\,. \tag{5.4}$$

Moreover, $L > 0$ and $N > 0$ were chosen to satisfy the requirements as well, therefore, estimate (4.20) holds in $[0, t_*]$. Recalling the notation O_r, as defined in (3.9) and integrating estimate (5.3), it follows that

$$\begin{aligned}\|g^-(t)\|_{L^2(\Omega_L)} &\le e^{C\|g_0\|_{L^1_{2/\lambda}(\Omega_L)}t}\Big(\|g_0^-\|_{L^2(\Omega_L)} + \frac{4}{N^{(d-1)/2}}(\mathfrak{c}_2^0(g_0))^2 \\ &\quad + O_{d/2+\lambda(k-1)}\, c_k^1(g_0)\, m_0(g_0)\Big) \\ &=: \varepsilon(t, L, N) \le \varepsilon(T, L, N)\,.\end{aligned}$$

Now, note that

$$\int_{\{g<0\}} |g(t,v)|\langle v\rangle^2 \mathrm{d}v \le L^{d/2+2}\|g^-(t)\|_{L^2(\Omega_L)} \le L^{d/2+2}\varepsilon(t,L,N)\,;$$

as a consequence, we can increase L and N, if necessary, so that

$$\begin{aligned}\epsilon(t) :=& \frac{\int_{\{g<0\}} |g(t,v)|\langle v\rangle^2 \mathrm{d}v}{\int_{\{g\ge 0\}} g(t,v)\langle v\rangle^2 \mathrm{d}v} = \frac{\int_{\{g<0\}} |g(t,v)|\langle v\rangle^2 \mathrm{d}v}{\int_{\Omega_L} g(t,v)\langle v\rangle^2 \mathrm{d}v - \int_{\{g<0\}} |g(t,v)|\langle v\rangle^2 \mathrm{d}v} \\ &\le \frac{L^{d/2+2}\varepsilon(T,L,N)}{\int_{\Omega_L} g_0(t,v)\langle v\rangle^2 \mathrm{d}v - L^{d/2+2}\varepsilon(T,L,N)} < \epsilon(g_0)\,.\end{aligned}$$

Observe that we used the fact that the scheme conserves mass and energy and assumed that $k > 1 + 2/\lambda$, so that $L^{d/2+2}\varepsilon(T, L, N)$ vanishes as both, L and then N are chosen sufficiently large. Therefore, for this choice of parameters $L \geq L_0(T, g_0)$ and $N \geq N_0(T, L, g_0)$, a continuation argument shows that the negative mass condition (5.4) holds, in fact, on the whole interval I_1. Thus, the a priori estimates of Proposition 4.4 hold in I_1 so that

$$\|g(t)\|_{L^2(\Omega_L)} \leq c_k^0(g_0), \qquad m_k(g(t)) \leq \mathfrak{c}_1^k(g_0) \qquad \forall\, t \in I_1\,. \tag{5.5}$$

Estimate (5.5) shows that the set $\mathcal{X}/2$ is a stable set for the dynamics, thus, it allows us to uniquely extend the solution, by repeating the argument made for I_1 to the intervals $I_i := [(i-1)/L^{a+b_k}, i/L^{a+b_k}]$ with $i = 1, 2, \ldots$, until $[0, T] \subset \cup I_i$. This proves global existence and uniqueness of the scheme.

Now, in the limit $N \to \infty$ one has that the sequence $\{g := g^N\} \subset \mathcal{X}$. Since

$$\big\|Q_c(f,f)(t) - Q_c(\tilde f, \tilde f)(t)\big\|_{L^2(\Omega_L)} \leq C(c_1^k, c_2^0)\, L^a \|f(t) - \tilde f(t)\,\|_{L^2(\Omega_L)} \quad \forall\, f\,, \tilde f \in \mathcal{X}\,,$$

it follows from

$$g(t) = g_0 + \int_0^t Q_c(g,g)(s)\mathrm{d}s$$

that for any $N, M \geq N_0$ and $t \in [0, T]$

$$\begin{aligned}\|g^N(t) - g^M(t)\|_{L^2(\Omega_L)} &\leq \|g_0^N - g_0^M\|_{L^2(\Omega_L)} \\ &\quad + C(c_1^k, c_2^0)\, L^a \int_0^t \|g^N(s) - g^M(s)\,\|_{L^2(\Omega_L)}\mathrm{d}s\,.\end{aligned}$$

Thus, using Gronwall's lemma,

$$\|g^N(t) - g^M(t)\|_{L^2(\Omega_L)} \leq \|g_0^N - g_0^M\|_{L^2(\Omega_L)} e^{C(c_1^k, c_2^0)\, L^a T} \to 0 \quad \text{as} \quad N, M \to \infty\,.$$

Thus, $\{g^N\}$ is Cauchy and converges strongly to $\bar g$, the solution of the problem (5.2) with initial condition $f_0 = \lim_{N\to\infty} \Pi_L^N f_0$. ∎

5.2. Uniform H_k Sobolev regularity propagation. In this section we work with functions in $H^{\alpha_0}(\Omega_L)$ and take multi-index α with $|\alpha| \leq \alpha_0$. Recall that derivatives commute with the projection operator Π_{2L}^N (see (2.25)), for functions in $H_0^{\alpha_0}(\Omega_{2L})$. Therefore, distributing the derivatives in the arguments of the operator and using the estimates (2.27), (2.28), and (4.2), yields

$$\begin{aligned}&\big\|\partial^\alpha(\mathbf{1} - \Pi_{2L}^N)Q^+(\chi g, \chi g)\big\|_{L^2(2\Omega_L)} \\ &\qquad = \big\|(\mathbf{1} - \Pi_{2L}^N)\partial^\alpha Q^+(\chi g, \chi g)\big\|_{L^2(2\Omega_L)} \leq \frac{C\, L^{2(1+\lambda)}}{N^{(d-1)/2}} \|g\|^2_{H^\alpha(\Omega_L)}\,,\end{aligned} \tag{5.6}$$

where, we recall, that the constant $C := C_\chi$ can be taken independent of $L \geq 1$.

Next, in order to prove propagation of regularity let us fix $k \geq k_* \geq 2$ and $0 \leq k' \leq k - 1 - \alpha_0(1+\lambda)$, and use an induction argument on the derivative order $|\alpha|$. The initial step of the induction follows thanks to the propagation of $L^2_{k'}$-norms in Proposition 4.4. For the case $|\alpha| \geq 1$, assume the propagation of the $H^{|\alpha|-1}_{k'+(1+\lambda)}$-norms and differentiate (4.5) w.r.t. velocity. We arrive at

$$\frac{\partial(\partial^\alpha g)}{\partial t} = \partial^\alpha Q^+(\chi g,\chi g) - \partial^\alpha Q^-(g,\chi g) + \partial^\alpha\big(Q_c(g,g)-Q_u(g,g)\big) - \partial^\alpha\big(\mathbf{1}-\Pi_{2L}^N\big)Q^+(\chi g,\chi g)\,.$$

Multiply by $\partial^\alpha g\langle v\rangle^{2\lambda k'}$ and integrate in the velocity domain Ω_L to obtain

$$\begin{aligned}\tfrac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\partial^\alpha g\|^2_{L^2_{k'}(\Omega_L)} &\le \int_{\Omega_L}\big(\partial^\alpha Q^+(\chi g,\chi g)-\partial^\alpha Q^-(g,\chi g)\big)\,\partial^\alpha g\langle v\rangle^{2\lambda k'}\\ &+\|\partial^\alpha g\|_{L^2_{k'}(\Omega_L)}\big\|\partial^\alpha\big(Q_c(g,g)-Q_u(g,g)\big)\big\|_{L^2_{k'}(\Omega_L)}\\ &+\|\partial^\alpha g\|_{L^2_{k'}(\Omega_L)}\big\|\partial^\alpha(\mathbf{1}-\Pi_{2L}^N)Q^+(\chi g,\chi g)\big\|_{L^2_{k'}(\Omega_L)} =: I_1+I_2+I_3\,.\end{aligned}\tag{5.7}$$

Recall from Lemma 3.2 that the term $Q_c(g,g)-Q_u(g,g)$ is a second order polynomial, therefore, its derivatives are at most a second order polynomial, thus Theorem 3.4 implies

$$I_2 \le \|\partial^\alpha g\|_{L^2_{k'}(\Omega_L)}\Big(L^{\lambda k'}\big\|(\mathbf{1}-\Pi_{2L}^N)Q^+(\chi g,\chi g)\big\|_{L^2(\Omega_L)} + O_{d/2}\big(m_{k'+1}(g)m_0(g)+Z_{k'}\big)\Big).\tag{5.8}$$

Additionally, the term I_3 is controlled using (5.6),

$$I_3 \le \frac{L^{\lambda k'+2(1+\lambda)}}{N^{(d-1)/2}}\|\partial^\alpha g\|_{L^2_{k'}(\Omega_L)}\|g\|^2_{H^\alpha(\Omega_L)}\,.\tag{5.9}$$

The term I_1 defined in (5.7) can be controlled implementing the estimate introduced in [17] and used for the control of $H_{k'}$-norms in [43, Theorem 3.5]:

$$\begin{aligned}I_1 &\le C_1\,\|\partial^\alpha g\|_{L^2_{k'}(\Omega_L)}\,\|g\|^2_{H^{|\alpha|-1}_{k'+(1+1/\lambda)}(\Omega_L)} - C(g_0)\,\|\partial^\alpha g\|^2_{L^2_{k'+1/2}(\Omega_L)}\\ &\le C_2\,\|\partial^\alpha g\|_{L^2_{k'}(\Omega_L)} - C(g_0)\,\|\partial^\alpha g\|^2_{L^2_{k'+1/2}(\Omega_L)}\,,\\ &\text{where } C_1\|g\|^2_{H^{|\alpha|-1}_{k'+(1+1/\lambda)}(\Omega_L)} \le C_2 \text{ by induction.}\end{aligned}\tag{5.10}$$

We obtain from inequalities (5.7), (5.8), (5.9), (5.10), and (5.6),

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial^\alpha g\|_{L^2_{k'}(\Omega_L)} \le C_2 - \frac{C(g_0)}{2}\|\partial^\alpha g\|_{L^2_{k'+1/2}(\Omega_L)} + \frac{L^{\lambda k'+2(1+\lambda)}}{N^{(d-1)/2}}\|g\|^2_{H^\alpha(\Omega_L)}.$$

The same inequality is valid for $\alpha=0$, therefore, it is concluded that

$$\frac{\mathrm{d}X}{\mathrm{d}t} \le C_2 - \frac{C(g_0)}{2}X + \frac{L^{\lambda k'+2(1+\lambda)}}{N^{(d-1)/2}}X^2,$$

where $X(t) := \|g\|_{H^\alpha_{k'}(\Omega_L)}$. From here, after taking $N \ge N_0(L,g_0)$ sufficiently large, it follows that

$$X(t) \le \max\big\{X(0), 4C_1/C_2\big\}\,, \qquad t\in[0,T]\,.$$

Note that in each step of the induction one needs to add $(1+1/\lambda)$ moments, so that $\|g\|_{H^{|\alpha|-1}_{k'+(1+1/\lambda)}}$ is finite. Having this in mind, let us state the result we just proved.

PROPOSITION 5.2. *Fix $T>0$, $\alpha\ge 0$, $k\ge k_*\ge 2$, and $0\le k'\le k-1-\alpha(1+1/\lambda)$ and assume $g_0\in H^\alpha_{k'+\alpha(1+1/\lambda)}(\Omega_L)$. Then, for any lateral size $L\ge L_0(T,g_0)$ there exists $N_0(T,L,g_0)>0$ such that*

$$\sup_{t\in[0,T]}\|g\|_{H^\alpha_{k'}(\Omega_L)} \le \max\big\{\|g_0\|_{H^\alpha_{k'+\alpha|(1+1/\lambda)}(\Omega_L)}, C_{k'}(g_0)\big\}, \quad N\ge N_0\,.$$

6. L^2_k and H^α_k error estimates. We are now in position to write the error estimates for the spectral conservation scheme. We start with errors in the $L^2_{k'}$-norm and, then, extend it to Sobolev norms. Again, we start fixing for $T > 0$, the cutoff domain $L \geq L_0(T, g_0)$ and $N \geq N_0(T, L, g_0)$ sufficiently large so that g exists in the interval $[0, T]$. Here $k \geq k_* \geq 2$ and $0 \leq k' \leq k-1$ in order to meet the assumptions of Proposition 4.4. From the identity

$$\begin{aligned} Q(g,g) &= Q(\chi g, \chi g) \\ &\quad + \big(Q((1-\chi)g, g) + Q(g,(1-\chi)g) + Q((1-\chi)g,(1-\chi)g)\big) \\ &=: Q(\chi g, \chi g) + E_0(g,g)\,, \end{aligned} \tag{6.1}$$

and the definition of Q_u, one finds that

$$\begin{aligned} Q_u(g,g) &= Q(g,g) - \big(E_0(g,g) + Q^-((1-\chi)g, \chi g)\big) \\ &=: Q(g,g) - E(g,g)\,. \end{aligned} \tag{6.2}$$

Now, observe that subtracting the Boltzmann equation (2.1) and its conserved projection approximation (3.7) in Ω_L one obtains

$$\begin{aligned} \partial_t(f-g) &= Q(f,f) - Q_c(g,g) = \big(Q(f,f) - Q_u(g,g)\big) + \big(Q_u(g,g) - Q_c(g,g)\big) \\ &= \big(Q(f,f) - Q(g,g)\big) + \big(Q_u(g,g) - Q_c(g,g)\big) + E(g,g)\,. \end{aligned} \tag{6.3}$$

Multiplying this equation by $(f-g)\langle v\rangle^{2\lambda k'}$ and integrating in Ω_L,

$$\begin{aligned} \tfrac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|f-g\|^2_{L^2_{k'}(\Omega_L)} &= \int_{\Omega_L} \langle v\rangle^{2\lambda k'}(f-g)\big(Q(f,f) - Q(g,g)\big)\big)\,\mathrm{d}v \\ &\quad + \int_{\Omega_L} \langle v\rangle^{2\lambda k'}(f-g)\big(Q_u(g,g) - Q_c(g,g)\big)\,\mathrm{d}v \\ &\quad + \int_{\Omega_L} \langle v\rangle^{2\lambda k'}(f-g)E(g,g)\,\mathrm{d}v \\ &=: I_1 + I_2 + I_3\,. \end{aligned}$$

The error term I_3, from the error term $E(g,g)$ in (6.2), is simply controlled as

$$I_3 \leq \|g\|_{L^1_{k'+1}(\mathbb{R}^d)}\|(1-\chi)g\|_{L^2_{k'+1}(\mathbb{R}^d)}\|f-g\|_{L^2_{k'}(\Omega_L)} \leq O_{d/2+\lambda k''}\|f-g\|_{L^2_{k'}(\Omega_L)}, \tag{6.4}$$

where the last inequality holds provided the $L^2_{k'+1+d/2\lambda+k''}$ uniformly propagate. Moreover, using Theorem 3.4 it follows that

$$\begin{aligned} \big\|Q_u(g,g) - Q_c(g,g)\big\|_{L^2_{k'}(\Omega_L)} &\leq C_5\, L^{\lambda k'}\big\|(1-\Pi^N_{2L})Q^+(\chi g, \chi g)\big\|_{L^2(\Omega_L)} \\ &\quad + O_{d/2+\lambda k''} m_{k'+1+k''}(g) m_0(g_0)\,. \end{aligned}$$

Therefore, using the Cauchy–Schwarz inequality and the (4.2) inequality one controls the term I_2 as

$$\begin{aligned} I_2 &\leq \|f-g\|_{L^2_{k'}(\Omega_L)}\left(\frac{C_5\, L^{\lambda k'}}{N^{(d-1)/2}}\|g\|^2_{L^2_{1+1/\lambda}(\Omega_L)} + O_{d/2+\lambda k''} m_{k'+1+k''}(g) m_0(g_0)\right) \\ &= \|f-g\|_{L^2_{k'}(\Omega_L)}\left(O\big(L^{\lambda k'}/N^{(d-1)/2}\big) + O_{d/2+\lambda k''}\right). \end{aligned} \tag{6.5}$$

The term I_1 is more involved. However, it is classical from the Boltzmann theory that the Dirichlet form of the linearized collision operator with polynomial weights is essentially nonpositive in the sense that

$$
\begin{aligned}
I_1 &= \tfrac{1}{2}\int_{\mathbb{R}^d} - \int_{\mathbb{R}^d\setminus\Omega_L} \langle v\rangle^{2\lambda k'}(f-g)\big(Q(f+g,f-g)+Q(f-g,f+g)\big)\,\mathrm{d}v \\
&\le C_{k'}\|f-g\|^2_{L^2_{k'}(\mathbb{R}^d)} + \big(\frac{c_1}{k'} + 2\|g_0\|_{L^1_{2\lambda^{-1}}}\epsilon - c_2\big)\|f-g\|^2_{L^2_{k'+1/2}(\mathbb{R}^d)} \\
&\quad + O_{d/2+\lambda k''}\|f+g\|_{L^1_{k'}(\mathbb{R}^d)}\|f-g\|_{L^2_{k'}(\mathbb{R}^d)}\|f-g\|_{L^2_{k'+1+d/2+k''}(\mathbb{R}^d)} \\
&\le C_{k'}\|f-g\|^2_{L^2_{k'}(\Omega_L)} + O_{d/2+\lambda k''}\big(\|f-g\|_{L^2_{k'}(\Omega_L)} + O_{d/2+\lambda k''}\big)\,.
\end{aligned}
\tag{6.6}
$$

The ϵ-term, with $0<\epsilon<\epsilon(g_0)$, is added in the absorption (second) term to account for the fact that there may be a set where $f+g$ is negative. This is not a problem since this set is small, $\{f+g<0\}\subset\{g<0\}$. Here, $C_{k'}$ is a constant that depends on the moments $L^1_{k'+2}$ and $c_i := c_i(f_0,g_0)$ depends only on the initial mass and some moment $2^+/\lambda$; see, for instance, [20, Proposition 2.1]. In the last inequality, we are taking k' and L sufficiently large so that $c_1/k' + C_k/L^\lambda + 2\|g_0\|_{L^1_{2/\lambda}}\epsilon - c_2 \le 0$, which is achieved for any ϵ in the aforementioned range. This estimate holds, of course, provided the $L^2_{k'+1+d/2\lambda+k''}$-norms propagate uniformly on $[0,T]$. Defining $X(t) := \|f(t)-g(t)\|^2_{L^2_{k'}(\Omega_L)}$ and combining the estimates (6.6), (6.5), and (6.4),

$$
\tfrac{1}{2}\frac{\mathrm{d}X}{\mathrm{d}t}(t) \le C_{k'}X(t) + \Big(O\big(L^{\lambda k'}/N^{(d-1)/2}\big) + O_{d/2+\lambda k''}\Big)\sqrt{X} + O_{d+2\lambda k''}\,.
$$

Thus, Gronwall's lemma implies

$$
\sup_{t\in[0,T]}\|f-g\|^2_{L^2_{k'}(\Omega_L)} \le e^{C_{k'}T}\big(\|f_0-g_0\|^2_{L^2_{k'}(\Omega_L)} + O\big(L^{2\lambda k'}/N^{(d-1)}\big) + O_{d+2\lambda k''}\big)\,. \tag{6.7}
$$

This proves the following theorem.

THEOREM 6.1 (L^2_k-error estimate). *Fix $k\ge k_*\ge 2$, $k''\ge 0$, and $k(f_0)<k'\le k-1-\frac{d^+}{2\lambda}-k''$ with $0\le f_0\in L^1_2\cap L^2_k(\mathbb{R}^d)$ an initial datum and f the solution of the Boltzmann equation* (2.1). *For any $T>0$ and cutoff domain $L(T,f_0)\ge L_0$ there exists $N_0(T,L,f_0)$ such that*

$$
\sup_{t\in[0,T]}\|f-g\|_{L^2_{k'}(\Omega_L)} \le e^{C_{k'}T}\big(\|f_0-g_0\|_{L^2_{k'}(\Omega_L)} + O\big(L^{\lambda k'}/N^{(d-1)/2}\big) + O_{d/2+\lambda k''}\big), \qquad N\ge N_0\,.
$$

The factor $O_{d/2+\lambda k''}$ is defined by (3.9). *The constants depend on $C_{k'} := C_{k'}\big(\|f_0\|_{L^2_{k'}}\big)$. In particular, the strong limit $\bar g$ of the sequence $\{g_N\}$ in $\mathcal{C}(0,T;L^2_k(\Omega_L))$ satisfies the same estimate.*

We study next the improvement in the rate of convergence with respect to the number of modes N of the approximating solutions towards the Boltzmann solution provided that the initial configuration is smooth and has at least initial mass and energy bounded.

THEOREM 6.2 (H^α-error estimates). *Fix $k\ge k_*\ge 2$, $k''\ge 0$, $\alpha_0\ge\alpha\ge 0$, and $k(f_0)\le k'\le k-1-\alpha/2-\frac{d^+}{2\lambda}-k''$ and let $0\le f_0\in L^1_2\cap H^\alpha_k(\mathbb{R}^d)$ be an initial datum*

and f be the solution of the Boltzmann equation (2.1). *Fix $T>0$ and cutoff domain $L \geq L_0(T,f_0)$. Then, there exists $N_0(T,L,f_0)$ such that*

(6.8)
$$\sup_{t\in[0,T]} \|f-g\|_{H^\alpha_{k'}(\Omega_L)} \leq e^{\alpha C_{k'} T}\big(\|f_0-g_0\|_{H^\alpha_{k'+\alpha/2}(\Omega_L)} + O\big(L^{\lambda(k'+\alpha/2)+\alpha_0}/N^{(d-1)/2+\alpha_0}\big) + O_{d/2+\lambda k''}\big)\,, \quad N\geq N_0\,.$$

with the factor $O_{d/2+\lambda k''}$ defined as in (3.9).

Proof. Fix $\alpha_0\geq 0$, $k\geq k_*\geq 2$, $k''\geq 0$, and $k(f_0)\leq k'\leq k-1-\alpha_0/2-\frac{d+1}{2\lambda}-k''$. Now, we perform similar computations to those of the error estimates for $L^2_{k'}$, though, avoiding to resort to the values of g near $\partial\Omega_L$. Thus, we write, for Q_u and Q_c defined in (3.3) and (3.7), respectively,

$$\begin{aligned} Q(f,f)-Q_c(g,g) &= Q(\chi f,\chi f)-Q(\chi g,\chi g)-Q^-((1-\chi)(f-g),\chi g) \\ &\quad + \big(Q_u(g,g)-Q_c(g,g)\big)+\tilde{E}(f,f)\,. \end{aligned} \tag{6.9}$$

Here $\tilde{E}(f,f):=E_0(f,f)+Q^-((1-\chi)f,\chi g)$. Thus, fixing any multi-index α with $|\alpha|\leq\alpha_0$, we apply the operator ∂^α to (6.9), multiply it by $\partial^\alpha(f-g)\langle v\rangle^{2\lambda k'}$, and integrate in Ω_L to obtain

$$\tfrac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\partial^\alpha(f-g)\|^2_{L^2_{k'}(\Omega_L)} = I_1^\alpha + I_2^\alpha + I_3^\alpha\,,$$

where,

$$\begin{aligned} I_1^\alpha &:= \int_{\Omega_L}\langle v\rangle^{2\lambda k'}\partial^\alpha(f-g)\,\big(\partial^\alpha Q(\chi f,\chi f)-\partial^\alpha Q(\chi g,\chi g)-\partial^\alpha Q^-((1-\chi)(f-g),\chi g)\big)\mathrm{d}v\,,\\ I_2^\alpha &:= \int_{\Omega_L}\langle v\rangle^{2\lambda k'}\partial^\alpha(f-g)\,\partial^\alpha\big(Q_u(g,g)-Q_c(g,g)\big)\mathrm{d}v\,,\\ I_3^\alpha &:= \int_{\Omega_L}\langle v\rangle^{2\lambda k'}\partial^\alpha(f-g)\,\partial^\alpha\tilde{E}(f,f)(v)\mathrm{d}v\,. \end{aligned}$$

Regarding the term I_2^α, we directly use Theorem 3.4 to have

$$\begin{aligned} \big\|\partial^\alpha\big(Q_u(g,g)-Q_c(g,g)\big)\big\|_{L^2_{k'}(\Omega_L)} &\lesssim \big\|Q_u(g,g)-Q_c(g,g)\big\|_{L^2_{k'}(\Omega_L)}\\ &\leq C_5\,L^{\lambda k'}\big\|(1-\Pi^N_{2L})Q^+(\chi g,\chi g)\big\|_{L^2(\Omega_L)} + O_{d/2+\lambda k''}m_{k'+1+k''}(g)m_0(g_0)\,. \end{aligned}$$

Therefore, using the Cauchy–Schwarz inequality, inequality (4.2), and Lemma A.1 implies

$$\begin{aligned} I_2^\alpha &\leq \|\partial^\alpha(f-g)\|_{L^2_{k'}(\Omega_L)}\bigg(\frac{C_5\,L^{\lambda k'+\alpha_0}}{N^{(d-1)/2+\alpha_0}}\|g\|^2_{H^{\alpha_0}_{1+1/\lambda}(\Omega_L)}+O_{d/2+\lambda k''}m_{k'+1+k''}(g)m_0(g_0)\bigg)\\ &= \|\partial^\alpha(f-g)\|_{L^2_{k'}(\Omega_L)}\bigg(O\big(L^{\lambda k'+\alpha_0}/N^{(d-1)/2+\alpha_0}\big)+O_{d/2+\lambda k''}\bigg)\,. \end{aligned}$$

The term I_3^α, containing the error term $\tilde{E}(f,f)$, is simply controlled as

$$\begin{aligned} I_3^\alpha &\leq C_\alpha\sum_{\alpha'+\beta'=\alpha}\Big(\|\partial^{\alpha'}(\chi g)\|_{L^1_{k'+1}(\mathbb{R}^d)}+\|\partial^{\alpha'}f\|_{L^1_{k'+1}(\mathbb{R}^d)}\Big)\\ &\quad\times\|\partial^{\beta'}((1-\chi)f)\|_{L^2_{k'+1}(\mathbb{R}^d)}\|\partial^\alpha(f-g)\|_{L^2_{k'}(\Omega_L)}\\ &\leq C_\alpha\,\|(1-\chi)f\|_{H^\alpha_{k'+1}(\mathbb{R}^d)}\|\partial^\alpha(f-g)\|_{L^2_{k'}(\Omega_L)} \leq O_{d/2+\lambda k''}\|\partial^\alpha(f-g)\|_{L^2_{k'}(\Omega_L)}\,, \end{aligned}$$

where the last inequality holds provided the $H^\alpha_{k'+1+d/2\lambda+k''}$-norm of f uniformly propagates. Finally, for the term I_1^α one checks that

$$\begin{aligned}\partial^\alpha\big(Q(\chi f,\chi f)-Q(\chi g,\chi g)\big) &= \tfrac{1}{2}\partial^\alpha\big(Q(\chi(f-g),\chi(f+g))+Q(\chi(f+g),\chi(f-g))\big)\\ &= \tfrac{1}{2}\big(Q\big(\partial^\alpha(\chi(f-g)),\chi(f+g)\big)+Q\big(\chi(f+g),\partial^\alpha(\chi(f-g))\big)\big)+\sum_{\alpha'+\beta'<\alpha}\Gamma^\alpha_{\alpha',\beta'},\end{aligned}$$

where

$$\Gamma^\alpha_{\alpha',\beta'} := \tfrac{C^\alpha_{\alpha',\beta'}}{2}\big(Q\big(\partial^{\alpha'}(\chi(f-g)),\partial^{\beta'}\chi(f+g)\big)+Q\big(\partial^{\beta'}\chi(f+g),\partial^{\alpha'}(\chi(f-g))\big)\big).$$

Observe that

$$\begin{aligned}\big\|\Gamma^\alpha_{\alpha',\beta'}\big\|_{L^2_{k'-1/2}(\Omega_L)} &\le C\,\|\partial^{\alpha'}\chi(f-g)\|_{L^2_{k'+1/2}(\mathbb{R}^d)}\|\partial^{\beta'}\chi(f+g)\|_{L^1_{k'+1/2}(\mathbb{R}^d)}\\ &\le C\,\|\partial^{\alpha'}\chi(f-g)\|_{L^2_{k'+1/2}(\mathbb{R}^d)} = C\,\|\partial^{\alpha'}(f-g)\|_{L^2_{k'+1/2}(\Omega_L)}+O_{d/2+\lambda k''},\end{aligned}$$

provided the $H^\alpha_{k'+1+d/2\lambda+k''}$-norms are propagated. Therefore,

$$\begin{aligned}&\int_{\Omega_L}\langle v\rangle^{2\lambda k'}\partial^\alpha(f-g)\sum_{\alpha'+\beta'<\alpha}\Gamma^\alpha_{\alpha',\beta'}\mathrm{d}v\\ &\qquad\le \|\partial^\alpha(f-g)\|_{L^2_{k'+1/2}(\Omega_L)}\big(C_\alpha\sum_{|\alpha'|<|\alpha|}\|\partial^{\alpha'}\chi(f-g)\|_{L^2_{k'+1/2}(\Omega_L)}+O_{d/2+\lambda k''}\big).\end{aligned}$$

Now, the leading order term in I_1^α is the Dirichlet form of the linearized Boltzmann operator with $\partial^\alpha\chi(f-g)$. Thus, similarly to what was done in the $L^2_{k'}$ error estimate, it follows that

$$\begin{aligned}I_1^\alpha \le C_{k'}&\|\partial^\alpha(f-g)\|^2_{L^2_{k'}(\mathbb{R}^d)}\\ &+\big(\frac{c_1}{k'}+\epsilon-c_2\big)\|\partial^\alpha(f-g)\|^2_{L^2_{k'+1/2}(\mathbb{R}^d)}+O_{d/2+k''}\|\partial^\alpha(f-g)\|_{L^2_{k'}(\mathbb{R}^d)}\\ &+C_\alpha\sum_{|\alpha'|<|\alpha|}\|\partial^{\alpha'}\chi(f-g)\|^2_{L^2_{k'+1/2}(\Omega_L)}+O_{d+2k''}\le C_k\|f-g\|^2_{L^2_{k'}(\Omega_L)}\\ &+O_{d/2+\lambda k''}\big(\|\partial^\alpha(f-g)\|_{L^2_{k'}(\Omega_L)}+O_{d/2+\lambda k''}\big)\\ &+C_\alpha\sum_{|\alpha'|<|\alpha|}\|\partial^{\alpha'}(f-g)\|^2_{L^2_{k'+1/2}(\Omega_L)}+O_{d+2\lambda k''}.\end{aligned}$$

Accordingly, this holds provided the $H^\alpha_{k'+1/2+d/2\lambda+k''}$-norms propagate uniformly on $[0,T]$. Also to obtain this estimate we have used the term with $\partial^\alpha Q^-((1-\chi)(f-g),\chi g)$ to complete the $L^2_{k'+1/2}$- absorbing norm in the whole $\mathbb{R}^d$. As a consequence, defining $X^\alpha(t) := \|\partial^\alpha(f(t)-g(t))\|^2_{L^2_{k'}(\Omega_L)}$ and combining the estimates for I_1^α, I_2^α, and I_3^α

$$\begin{aligned}\tfrac{1}{2}\frac{\mathrm{d}X^\alpha}{\mathrm{d}t}(t) \le C_kX^\alpha(t)+O\big(L^{\lambda k'+\alpha_0}/N^{(d-1)/2+\alpha_0}\big)\sqrt{X^\alpha}+O_{d/2+\lambda k''}\\ +C_\alpha\sum_{|\alpha'|<|\alpha|}\|\partial^{\alpha'}(f-g)\|^2_{L^2_{k'+1/2}(\Omega_L)}.\end{aligned}$$

Thus, Gronwall's lemma implies

$$X^{\alpha}(t) \le e^{2C_k T}\Big(X^{\alpha}(0) + O\big(L^{2\lambda k'+2\alpha_0}/N^{d-1+2\alpha_0}\big) + O_{d+2\lambda k''} + C_\alpha \sum_{|\alpha'|<|\alpha|} \sup_{t\in[0,T]} \|\partial^{\alpha'}(f-g)(t)\|^2_{L^2_{k'+1/2}(\Omega_L)}\Big).$$

Estimate (6.8) follows by iteration of this formula on the multi-index order $|\alpha| = 1, 2, \ldots, \alpha_0$, using Theorem 6.1 as the starting point. ∎

7. Long time behavior. In this final section we address the long time behavior for the semidiscrete problem given by the conservative spectral scheme approximating the space homogeneous elastic Boltzmann equation for hard potentials with integrable angular cross section.

Thus, we start by setting $g = \mathcal{M}_0 + h$, where $h := g - \mathcal{M}_0$ is the perturbation from the global Maxwellian equilibrium defined in (2.10). Note that under this linearization

$$Q_c(g,g) = Q_c(\mathcal{M}_0,\mathcal{M}_0) + Q_c(\mathcal{M}_0,h) + Q_c(h,\mathcal{M}_0) + Q_c(h,h)\,.$$

Introduce then the linear operators

$$\mathcal{L}_c(h) := Q_c(\mathcal{M}_0,h) + Q_c(h,\mathcal{M}_0)\,,$$
$$\mathcal{L}(\chi h) := Q(\mathcal{M}_0,\chi h) + Q(\chi h,\mathcal{M}_0)\,.$$

The reader recognizes the latter $\mathcal{L}$ as the linearized Boltzmann operator. With the estimations we have performed in the previous section, it is clear that

$$\|\chi Q_c(\mathcal{M}_0,\mathcal{M}_0)\|_{H^\alpha_k(\mathbb{R}^d)} \le \mathrm{O}\big(L^{\lambda k}/N^{\frac{d-1}{2}}\big) + \mathrm{O}\big(1/L^{\lambda k}\big)\,,$$
$$\|\chi\mathcal{L}_c(h) - \mathcal{L}(\chi h)\|_{H^\alpha_k(\mathbb{R}^d)} \le \mathrm{O}\big(L^{\lambda k}/N^{\frac{d-1}{2}}\big) + \mathrm{O}\big(1/L^{\lambda k}\big)\,,$$
$$\|\chi Q_c(h,h) - Q(\chi h,\chi h)\|_{H^\alpha_k(\mathbb{R}^d)} \le \mathrm{O}\big(L^{\lambda k}/N^{\frac{d-1}{2}}\big) + \mathrm{O}\big(1/L^{\lambda k}\big)\,.$$

For the last two estimates we need h, thus g, having α derivatives and $2k$-moments in Ω_L. This, of course, is guaranteed by the results of section 5 as long as the negative mass in g is small, $\epsilon \le \epsilon(g_0)$. As a consequence,

$$\frac{\mathrm{d}}{\mathrm{d}t}\chi h = \mathcal{L}(\chi h) + Q(\chi h,\chi h) + \mathcal{R}(h)\,, \tag{7.1}$$

where the remainder is of size $\|\mathcal{R}(h)\|_{H^\alpha_k(\mathbb{R}^d)} \le \mathrm{O}\big(L^{\lambda k}/N^{\frac{d-1}{2}}\big) + \mathrm{O}\big(1/L^{\lambda k}\big)$. Now, classical estimates on the Boltzmann operator and interpolation shows that

$$\|Q(\chi h,\chi h)\|_{H^\alpha_k(\mathbb{R}^d)} \le C_k \|\chi h\|^{3/2}_{H^\alpha_k(\mathbb{R}^d)}\,,$$

where the constant C_k depends on k'-moments and the $H^\alpha_{k'}$-norm of h for some $k' \ge k + 2\lambda + d$. Furthermore, the linearized Boltzmann operation has spectral gap, say $\nu > 0$, in H^α_k. See, for example, reference [21, 41]. Thus, we can integrate (7.1) to obtain that

$$\chi h(t) = \chi h_0 + \int_0^t e^{\mathcal{L}(t-s)} Q(\chi h,\chi h)(s)\mathrm{d}s + \int_0^t e^{\mathcal{L}(t-s)}\mathcal{R}(h)(s)\mathrm{d}s\,.$$

Since, the remainder $\mathcal{R}(h)$ may not have zero mass, momentum, and energy, we apply the operator $1-\pi$, where π is the standard projection on the Boltzmann null space in $H^\alpha_k(\mathbb{R}^d) \subset L^1_{2/\lambda}(\mathbb{R}^d)$, which is given by

$$\pi h = \sum_{\phi\in\{1,v_1,\cdots,v_d,|v|^2\}} \int_{\mathbb{R}^d} h(v)\phi(v)\mathrm{d}v\, \phi(v)\mathcal{M}(v)\,, \quad \mathcal{M} \text{ is the normalized Maxwellian.}$$

Using the fact that the semigroup and π commutes, one has

$$(1-\pi)\chi h(t) = (1-\pi)\chi h_0 + \int_0^t e^{\mathcal{L}(t-s)} Q(\chi h, \chi h)(s)\mathrm{d}s + \int_0^t e^{\mathcal{L}(t-s)}(1-\pi)\mathcal{R}(h)(s)\mathrm{d}s\,,$$

where we used that $(1-\pi)Q(\cdot,\cdot) = Q(\cdot,\cdot)$. Thus, applying the H^α_k-norm we conclude that

(7.2)
$$\begin{aligned}\|(1-\pi)\chi h(t)\|_{H^\alpha_k(\mathbb{R}^d)} \le \|(1-\pi)\chi h_0\|_{H^\alpha_k(\mathbb{R}^d)} + \frac{1}{\nu}\Big(\mathrm{O}\big(L^{\lambda k}/N^{\frac{d-1}{2}}\big) + \mathrm{O}\big(1/L^{\lambda k}\big)\Big)\\ + C_k\int_0^t e^{-\nu(t-s)}\|\chi h(s)\|^{3/2}_{H^\alpha_k(\mathbb{R}^d)}\mathrm{d}s\,.\end{aligned}$$

Now, the conservation routine grants that $\pi h(t) = 0$ for any $t \ge 0$. Then,

$$\|\pi\chi h\|_{H^\alpha_k(\mathbb{R}^d)} = \|\pi(1-\chi)h\|_{H^\alpha_k(\mathbb{R}^d)} \le C_{k,\alpha}\|(1-\chi)h\|_{L^1_k(\mathbb{R}^d)} = O\big(1/L^{\lambda k}\big)\,.$$

As a consequence,

(7.3)
$$\|(1-\pi)\chi h(t)\|_{H^\alpha_k(\mathbb{R}^d)} = \|\chi h(t)\|_{H^\alpha_k(\mathbb{R}^d)} + O\big(1/L^{\lambda k}\big) = \|h(t)\|_{H^\alpha_k(\Omega_L)} + O\big(1/L^{\lambda k}\big)\,.$$

Thus, estimates (7.2), (7.3), and (2.27) leads to the control

(7.4)
$$\begin{aligned}\|h(t)\|_{H^\alpha_k(\Omega_L)} \le \|h_0\|_{H^\alpha_k(\Omega_L)} + \frac{1}{\nu}\Big(\mathrm{O}\big(L^{\lambda k}/N^{\frac{d-1}{2}}\big) + \mathrm{O}\big(1/L^{\lambda k}\big)\Big)\\ + C_k\int_0^t e^{-\nu(t-s)}\|h(s)\|^{3/2}_{H^\alpha_k(\Omega_L)}\mathrm{d}s =: Y(t)\,.\end{aligned}$$

Observing that

$$Y'(t) = C_k\,\|h(t)\|^{3/2}_{H^\alpha_k(\Omega_L)} - \nu\, C_k\int_0^t e^{-\nu(t-s)}\|h(s)\|^{3/2}_{H^\alpha_k(\Omega_L)}\mathrm{d}s\,,$$

one concludes, using (7.4), that

$$Y'(t) + \nu Y(t) \le C_k Y^{3/2}(t) + \nu\|h_0\|_{H^\alpha_k(\Omega_L)} + \mathrm{O}\big(L^{\lambda k}/N^{\frac{d-1}{2}}\big) + \mathrm{O}\big(1/L^{\lambda k}\big)\,. \tag{7.5}$$

This estimate tell us that if

$$C_k\sqrt{Y_0} = C_k\sqrt{\|h_0\|_{H^\alpha_k(\Omega_L)} + \frac{1}{\nu}\Big(\mathrm{O}\big(L^{\lambda k}/N^{\frac{d-1}{2}}\big) + \mathrm{O}\big(1/L^{\lambda k}\big)\Big)} \ll \nu\,, \tag{7.6}$$

then

$$\|h(t)\|_{H^\alpha_k(\Omega_L)} \le Y(t) \lesssim \|h_0\|_{H^\alpha_k(\Omega_L)} + \frac{1}{\nu}\Big(\mathrm{O}\big(L^{\lambda k}/N^{\frac{d-1}{2}}\big) + \mathrm{O}\big(1/L^{\lambda k}\big)\Big)\,, \quad t > 0\,. \tag{7.7}$$

This proves the following local stability estimate for the conservative semidiscrete solution.

PROPOSITION 7.1 (local stability for the semidiscrete scheme). *Fix $\alpha_0 \geq 0$ and let $g_0 \in H^{\alpha_0}_{2k}(\Omega_L)$ with $k \geq k_* > 1 + \frac{d}{2\lambda}$, an initial datum for the semidiscrete problem. Assume that $\|g_0 - \mathcal{M}_0\|_{H^\alpha_k(\Omega_L)} \leq \delta/2$ for $0 < \delta \ll \min\{\nu, \epsilon(g_0)\}$. Then, there exist a lateral size $L_0(g_0, \nu) > 0$ and a number of modes $N_0(g_0, L_0, \nu)$ such that for any $\alpha \leq \alpha_0$*

$$\sup_{t\geq 0} \|g - \mathcal{M}_0\|_{H^\alpha_k(\Omega_L)} \leq \delta\,, \quad L \geq L_0\,,\ N \geq N_0\,,$$

where $\mathcal{M}_0$ is the Maxwellian having the same mass, momentum, and energy as the initial configuration g_0.

Proof. The result follows from the aforementioned discussion noticing that (7.7) is valid provided L is taken first sufficiently large and then $N := N(L)$, in a way that (7.6) is satisfied. Since the constant C_k depends on propagations of moments and the norm $H^{\alpha_0}_k$, the validity of (7.7) holds provided the negative mass of g is small. However, this is clear since

$$\|g^-\|_{L^2_k(\Omega_L)} \leq \|(g - \mathcal{M}_0)1_{\{g\leq 0\}}\|_{L^2_k(\Omega_L)} \leq \delta \ll \epsilon(g_0)\,. \qquad \square$$

As a corollary of the error estimates and the local stability of the scheme, exponential relaxation to the Maxwellian equilibrium follows in Lebesgue and Sobolev norms. Indeed, using the classical asymptotic Boltzmann theory [21, 41] for variable hard potentials,

$$\|f - \mathcal{M}_0\|_{H^\alpha_k(\mathbb{R}^d)} \leq C_k\, \|f_0\|_{H^\alpha_k(\mathbb{R}^d)}\, e^{-\nu t}\,,$$

where $\nu > 0$ is the spectral gap of the linearized Boltzmann equation. Thus, for any $\delta > 0$ we can choose

$$T(\delta) := \ln\left(\frac{4\, C_k\, \|f_0\|_{H^\alpha_k(\mathbb{R}^d)}}{\delta}\right)^{1/\nu}, \quad \text{so that} \quad \sup_{t\geq T(\delta)/2} \|f - \mathcal{M}_0\|_{H^\alpha_k(\mathbb{R}^d)} \leq \delta/4\,. \tag{7.8}$$

THEOREM 7.2 (convergence to the Maxwellian equilibrium). *Fix $\alpha_0 \geq 0$ and let $f_0 \in H^{\alpha_0}_{2k}(\mathbb{R}^d)$ with $k \geq k_* > 1 + \frac{d}{2\lambda}$, an initial datum. Then, for every $0 < \delta \ll \min\{\nu, \epsilon(g_0)\}$ there exist a lateral size $L_0(f_0) > 0$ and a number of modes $N_0(L, f_0)$ such that for any $\alpha \leq \alpha_0$*

$$\sup_{t\geq T(\delta)/2} \|g - \mathcal{M}_0\|_{H^\alpha_k(\Omega_L)} \leq \delta\,, \quad L \geq L_0\,,\ N \geq N_0\,,$$

where $\mathcal{M}_0$ is the Maxwellian having the same mass, momentum, and energy as the initial configuration f_0.

Proof. Letting $T = T(\delta)$ in Theorem 6.1 for the case $\alpha_0 = 0$ or Theorem 6.2 for the case $\alpha_0 > 0$, one concludes that there exist a lateral size $L_0(T(\delta), f_0)$ and number of modes $N_0(T(\delta), L, f_0)$ such that

$$\sup_{t\in[0,T]} \|f - g\|_{H^\alpha_k(\Omega_L)} \leq \delta/4\,, \quad L \geq L_0\,,\ N \geq N_0\,. \tag{7.9}$$

Using the triangle inequality with (7.8) and (7.9) one has that

$$\sup_{t\in[T(\delta)/2, T(\delta)]} \|g - \mathcal{M}_0\|_{H^\alpha_k(\Omega_L)} \leq \delta/2\,, \quad L \geq L_0\,,\ N \geq N_0\,.$$

The result follows after invoking the local stability result of Proposition 7.1. $\square$

Remark 7.3. Since the relaxation of the Boltzmann solution $f(t,v)$ is exponentially fast for variable hard potentials, the simulation times are relatively short as noticed in the previous proof. This makes conservative schemes very stable even when using relatively small working domains and number of modes.

Completion of proof of Theorem 3.1. We just need to discuss the time uniform nature of the constants appearing in the error estimates. We first observe that the conservative spectral scheme follows the nonlinear dynamics of the Boltzmann equation in the time range $[0, T(\delta)/2]$. Next, for $t \geq T(\delta)/2$, the dynamics is relaxed around the thermal equilibrium, so that, it is controlled by the linear evolution. Hence,

$$\|f-g\|_{H^\alpha_k(\Omega_L)} = \|f-\mathcal{M}_0\|_{H^\alpha_k(\Omega_L)} + \|g-\mathcal{M}_0\|_{H^\alpha_k(\Omega_L)} \leq 2\delta \quad \text{for} \quad t \geq T(\delta)/2\,.$$

As a consequence, in the long run, $f-g$ is estimated by the minimum between estimate (6.8) evaluated at $T(\delta)/2$ and 2δ. As a consequence, we conclude

$$\sup_{t\geq 0} \|f-g\|_{H^\alpha_k(\Omega_L)} \leq e^{\alpha C_k T(\delta)} \Big(\|f_0-g_0\|_{H^\alpha_{k+\alpha/2}(\Omega_L)} + O\big(L^{\lambda(k+\alpha/2)+\alpha_0}/N^{(d-1)/2+\alpha_0}\big) + O_{d/2+\lambda k}\Big)\,,$$

for $L \geq L_0(T(\delta), f_0)$, $N \geq N_0(T(\delta), L_0, f_0)$, and the term $O_{d/2+\lambda k}\big)$ as defined in (3.9).

Recalling (7.8), note that

$$e^{\alpha C_k T(\delta)} \sim \left(\frac{4C_k \|f_0\|_{H^\alpha_k(\Omega_L)}}{\delta} \right)^{\alpha C_k/\nu}.$$

The proof of Theorem 3.1 is concluded after minimizing in $\delta > 0$, which gives $\theta = \alpha C_k/\nu$ in items 2 and 3.

8. Conclusion. We have studied the global existence and error estimates for the homogeneous Boltzmann spectral method imposing conservation of mass, momentum, and energy by Lagrange constrained optimization. The methods and estimates presented in the document show that imposing conservation of these quantities stabilizes the long time behavior of the discrete problem because it enforces the collisional invariants. In some sense, this in turn enforces the numerical approximation of the linearized collisional operator to have the same null space as the true linearized collision operator, which is the one in charge of the long time dynamics. In particular, the work domain and the number of modes can be chosen such that the discrete solution approximates with any desired accuracy the stationary state of the original Boltzmann problem in the long run. Although spurious tail behavior is experienced when the optimization is imposed due to the addition of a quadratic polynomial corrector, the natural property of creation of moments remains in the semidiscrete problem. This allows one to minimize such spurious behavior by appropriate choice of simulation parameters. We point out here that other correctors, such as Gaussian, might be more suitable in this respect. Furthermore, conservation of mass and energy limits the negative mass produced by the numerical scheme which is essential for long time accurate simulations.

Appendix A.

A.1. Shannon sampling theorem. The following result is an extension of the standard approximation estimate for regular functions by Fourier series expansions,

Shannon sampling theorem, to the $H^\alpha(\Omega_L)$ space. We include here the result for completeness of the reading.

LEMMA A.1 (Fourier approximation estimate). *Let $g \in H^\alpha(\Omega_L)$, then*

$$\|(\boldsymbol{1} - \Pi_L^N)g\|_{L^2(\Omega_L)} \leq \frac{1}{(\sqrt{2\pi})^d} \left(\frac{L}{2\pi N}\right)^\alpha \|g\|_{H^\alpha(\Omega_L)}\,. \tag{A.1}$$

Proof. Parseval's relation gives

$$\|(\mathbf{1} - \Pi_L^N)g\|_{L^2(\Omega_L)} = \sqrt{\sum_{k>N} |\widehat{g}(\zeta_k)|^2}\,.$$

Furthermore, properties of the Fourier transform imply

$$|\widehat{g}(\zeta_k)| = \frac{1}{(\sqrt{2\pi})^d} \frac{\left|\widehat{D^\alpha g}(\zeta_k)\right|}{\prod_{j=1}^d |(\zeta_k^j)^{\alpha_j}|}\,.$$

Therefore,

$$\sum_{k>N} |\widehat{g}_N(\zeta_k)|^2 = \frac{1}{(2\pi)^d} \sum_{k>N} \frac{\left|\widehat{D^\alpha g}(\zeta_k)\right|^2}{\prod_{j=1}^d |(\zeta_k^j)^{\alpha_j}|^2} \leq \frac{1}{(2\pi)^d} \frac{\sum_{k>N} \left|\widehat{D^\alpha g}(\zeta_k)\right|^2}{\prod_{j=1}^d |(\zeta_N^j)^{\alpha_j}|^2}\,.$$

Observe that the sum in the last inequality equals the L^2-norm square of $D^\alpha g - \Pi^N D^\alpha g$; therefore,

$$\sum_{k>N} |\widehat{g}_N(\zeta_k)|^2 \leq \frac{1}{(2\pi)^d} \frac{\left\|D^\alpha g - \Pi^N D^\alpha g\right\|_{L^2(\Omega_L)}^2}{\prod_{j=1}^d |(\zeta_N^j)^{\alpha_j}|^2} \leq \frac{1}{(2\pi)^d} \frac{\|D^\alpha g\|_{L^2(\Omega_L)}^2}{\prod_{j=1}^d |(\zeta_N^j)^{\alpha_j}|^2}\,.$$

Conclude by recalling the definition of $\zeta_N = \frac{2\pi N}{L}$. ∎

A.2. Estimate on the decay of the collision operator.

THEOREM A.2. *The following estimate holds for any $k \geq 0$ and $\lambda \in [0,2]$:*

$$\left| \int_{\mathbb{R}^d \setminus \Omega_L} Q(f,f)(v)\, dv \right| \leq O_k\big(m_{k+1}(f) m_0(f) + Z_k(f)\big).$$

The term $Z_k(f)$ is defined below in (A.3) *and only depends on moments up to order k. In particular one has*

$$Z_k(f) \leq 2^k\, m_1(f)\, m_k(f). \tag{A.2}$$

Proof. For the negative part,

$$\begin{aligned}\left| \int_{\mathbb{R}^d \setminus \Omega_L} Q^-(f,f)(v) \mathrm{d}v \right| &\leq L^{-\lambda k} \int_{\{|v| \geq L\}} Q^-(|f|,|f|)(v) |v|^{\lambda k} \mathrm{d}v \\ &\leq L^{-\lambda k} \big(m_{k+1} m_0 + m_k m_0\big)\,.\end{aligned}$$

For the positive part,

$$\begin{aligned}\left| \int_{\mathbb{R}^d \setminus \Omega_L} Q^+(f,f)(v) \mathrm{d}v \right| &\leq L^{-\lambda k} \int_{\{|v| \geq L\}} Q^+(|f|,|f|)(v) |v|^{\lambda k} \mathrm{d}v \\ &= L^{-\lambda k} \int_{\mathbb{R}^{2d}} |f(v)||f(v_*)||u|^\lambda \int_{\mathbb{S}^{d-1}} |v'|^{\lambda k} b(\hat{u} \cdot \sigma) \mathrm{d}\sigma \mathrm{d}v_* \mathrm{d}v.\end{aligned}$$

Note,

$$\int_{\mathbb{S}^{d-1}} |v'|^{\lambda k} b(\hat{u}\cdot\sigma)\mathrm{d}\sigma \le \|b\|_{L^1(S^{d-1})}\left(|v|^2+|v_*|^2\right)^{\lambda k/2} \le \|b\|_{L^1(S^{d-1})}\sum_{j=0}^{k}\left(\begin{array}{c} k \\ j \end{array}\right)|v|^{\lambda j}|v_*|^{\lambda(k-j)}\,.$$

Use the inequality $|u|^\lambda \le |v|^\lambda + |v_*|^\lambda$ with the previous expressions to obtain

$$\left|\int_{\mathbb{R}^d\setminus\Omega_L} Q^+(f,f)(v)\mathrm{d}v\right| \le 2\|b\|_{L^1(S^{d-1})}L^{-\lambda k}\big(m_{k+1}(f)m_0(f)+Z_k(f)\big)\,,$$

where

$$Z_k(f) := \sum_{j=0}^{k-1}\left(\begin{array}{c} k \\ j \end{array}\right) m_{j+1}(f)m_{k-j}(f). \tag{A.3}$$

Furthermore, note that interpolation implies, for $0\le j\le k-1$,

$$m_{j+1}(f) \le m_1(f)^{\frac{k-1-j}{k-1}}\, m_k(f)^{\frac{j}{k-1}}\,, \quad m_{k-j}(f) \le m_1(f)^{\frac{j}{k-1}}\, m_k(f)^{\frac{k-1-j}{k-1}}\,.$$

Therefore,

$$m_{j+1}(f)\, m_{k-j}(f) \le m_1(f)\, m_k(f)\,, \quad 0\le j\le k-1\,.$$

This implies that

$$Z_k(f) \le m_1(f)\, m_k(f)\sum_{j=0}^{k-1}\left(\begin{array}{c} k \\ j \end{array}\right) \le 2^k\, m_1(f)\, m_k(f)\,. \qquad \square$$

A.3. L^2-theory of the collision operator. The following theorems follow from the arguments in [27, 2, 3]

THEOREM A.3 (collision integral estimate for elastic/ inelastic collisions). *For $f,g\in L^1_{k+1}(\mathbb{R}^d)\cap L^2_{k+1}(\mathbb{R}^d)$ one has the estimate*

$$\|Q(f,g)\|_{L^2_k(\mathbb{R}^d)} \le C\,\big(\|f\|_{L^2_{k+1}(\mathbb{R}^d)}\|g\|_{L^1_{k+1}(\mathbb{R}^d)} + \|f\|_{L^1_{k+1}(\mathbb{R}^d)}\|g\|_{L^2_{k+1}(\mathbb{R}^d)}\big), \tag{A.4}$$

where the dependence of the constant is $C := C(d,\|b\|_1)$.

Theorem A.3 and the Leibniz formula

$$\partial^\alpha Q(f,g) = \sum_{|j|\le|\alpha|}\binom{\alpha}{j} Q(\partial^{\alpha-j}f, \partial^j g) \quad \text{for multi-indexes } j\,,\ \alpha\,, \tag{A.5}$$

prove the following theorem; see [27, section 4] for additional discussion.

THEOREM A.4 (Sobolev bound estimate). *Let $\mu > 1+\frac{d}{2\lambda}$. For $f,g\in H^\alpha_{k+\mu}(\mathbb{R}^d)$, the collision operator satisfies*

$$\|Q(f,g)\|^2_{H^\alpha_k(\mathbb{R}^d)} \le C\sum_{j\le\alpha}\binom{\alpha}{j}\Big(\|f\|^2_{H^{\alpha-j}_{k+1}(\mathbb{R}^d)}\|g\|^2_{H^j_{k+\mu}(\mathbb{R}^d)} + \|f\|^2_{H^{\alpha-j}_{k+\mu}(\mathbb{R}^d)}\|g\|^2_{H^j_{k+1}(\mathbb{R}^d)}\Big)\,, \tag{A.6}$$

where the dependence of the constant is $C := C(d,\alpha,\|b\|_1)$.

COROLLARY A.5. *Let $\mu > \frac{d}{2} + \lambda$. For $f \in H^{\alpha}_{k+\mu}(\mathbb{R}^d)$ the collision operator satisfies the estimate*

$$\|Q(f,f)\|_{H^{\alpha}_{k}(\mathbb{R}^d)} \leq C\|f\|^2_{H^{\alpha}_{k+\mu}(\mathbb{R}^d)}. \tag{A.7}$$

The dependence of the constant is given by $C := C(d, \mu, \|b\|_1)$.

In this last section of the appendix we discuss briefly the gain of integrability in the gain collision operator; see [5] for a more detailed discussion.

THEOREM A.6. *The collision operator satisfies the estimate for any $\epsilon > 0$ and $k \geq 0$:*

$$\|Q^{+}_{\lambda}(g,f)\|_{L^2_k(\mathbb{R}^d)} \leq C\|b\|_{\infty}\|g\|_{L^1_k(\Omega_L)}\left(\tfrac{\epsilon^{r'}}{r'}\|f\|_{L^2_k(\mathbb{R}^d)} + \tfrac{1}{r\epsilon^{r}}\|f\|^{1-\theta}_{L^1_k(\Omega_L)}\|f\|^{\theta}_{L^2_k(\mathbb{R}^d)}\right),$$

where $\theta = \frac{1}{d}$, $r = \frac{d-2}{\lambda}$, and C_n is a constant depending only on the dimension.

Acknowledgments. The authors thank and gratefully acknowledged the hospitality and support from the Institute of Computational Engineering and Sciences and the University of Texas Austin.

REFERENCES

[1] R. ALONSO, J. CANIZO, I. GAMBA, AND C. MOUHOT, *A new approach to the creation and propagation of exponential moments in the Boltzmann equation*, Comm. Partial Differential Equations., 38 (2013), pp. 155–169.
[2] R. ALONSO AND E. CARNEIRO, *Estimates for the Boltzmann collision operator via radial symmetry and Fourier transform*, Adv. Math., 223 (2010), pp. 511–528.
[3] R. ALONSO, E. CARNEIRO, AND I. M. GAMBA, *Convolution inequalities for the Boltzmann collision operator*, Comm. Math. Physics., 298 (2010), pp. 293–322.
[4] R. ALONSO AND I. M. GAMBA, *$L^1 - L^\infty$ Maxwellian bounds for the derivatives of the solution of the homogeneous Boltzmann equation*, J. Math. Pures Appl. (9), 89 (2008), pp. 575–595.
[5] R. ALONSO AND I. M. GAMBA, *Gain of integrability for the Boltzmann collisional operator*, Kinet. Relat. Models, 4 (2011), pp. 41–51.
[6] R. ALONSO AND B. LODS, *Free cooling and high-energy tails of granular gases with variable restitution coefficient*, Commun. Math. Sci., 11 (2013), pp. 807–862.
[7] R. ALONSO AND B. LODS, *Boltzmann model for viscoelastic particles: Asymptotic behavior, pointwise lower bounds and regularity*, Commun. Math. Phys., 331 (2014), pp. 545–591.
[8] G. A. BIRD, *Molecular Gas Dynamics*, Clarendon Press, Oxford, 1994.
[9] A. V. BOBYLEV, *Exact solutions of the nonlinear Boltzmann equation and the theory of relaxation of a Maxwellian gas*, Teoret. Mat. Fiz., 60 (1984), pp. 280–310.
[10] A. V. BOBYLEV, *Moment inequalities for the Boltzmann equation and applications to spatially homogeneous problems*, J. Stat. Phys., 88 (1997), pp. 1183–1214.
[11] A. V. BOBYLEV, J. A. CARRILLO, AND I. M. GAMBA, *On some properties of kinetic and hydrodynamic equations for inelastic interactions*, J. Stat. Phys., 98 (2000), pp. 743–773.
[12] A. V. BOBYLEV AND C. CERCIGNANI, *Discrete velocity models without nonphysical invariants*, J. Stat. Phys., 97 (1999), pp. 677–686.
[13] A. V. BOBYLEV, I. M. GAMBA, AND V. A. PANFEROV, *Moment inequalities and high-energy tails for Boltzmann equations with inelastic interactions*, J. Stat. Phys., 116 (2004), pp. 1651–1682.
[14] A. V. BOBYLEV AND S. RJASANOW, *Difference scheme for the Boltzmann equation based on the Fast Fourier Transform*, Eur. J. Mech. B Fluids, 16 (1997), pp. 293–306.
[15] A. V. BOBYLEV AND S. RJASANOW, *Fast deterministic method of solving the Boltzmann equation for hard spheres*, Eur. J. Mech. B Fluids, 18 (1999), pp. 869–887.
[16] A. V. BOBYLEV AND S. RJASANOW, *Numerical solution of the Boltzmann equation using fully conservative difference scheme based on the Fast Fourier Transform*, Transp. Theory Stat. Phys., 29 (2000), pp. 289–310.
[17] F. BOUCHUT AND L. DESVILLETTES, *A proof of the smoothing properties of the positive part of Boltzmann's kernel*, Rev. Mat. Iberoam., 14 (1998), pp. 47–61.

[18] J. E. Broadwell, *Study of rarefied shear flow by the discrete velocity method*, J. Fluid Mech., 19 (1964), pp. 401–414.
[19] C. Cercignani, R. Illner, and M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, Appl. Math. Sci. 106, Springer, New York, 1994.
[20] L. Desvillettes and C. Mouhot, *About L^p estimates for the spatially homogeneous Boltzmann equation*, Ann. Inst. H. Poincaré Anal. Non. Linéaire, 22 (2005), pp. 127–142.
[21] L. Desvillettes and C. Villani, *On the trend to global equilibrium for spatially inhomogeneous kinetic systems: The Boltzmann equation*, Invent. Math., 159 (2005), pp. 245–316, https://doi.org/10.1007/s00222-004-0389-9,.
[22] F. Filbet and C. Mouhot, *Analysis of spectral methods for the homogeneous Boltzmann equation*, Trans. Amer. Math. Soc., 363 (2010), pp. 1947–1980.
[23] F. Filbet, C. Mouhot, and L. Pareschi, *Solving the Boltzmann equation in $Nlog_2N$*, SIAM J. Sci. Comput., 28 (2006), pp. 1029–1053.
[24] F. Filbet and G. Russo, *High order numerical methods for the space non homogeneous Boltzmann equation*, J. Comput. Phys., 186 (2003), pp. 457–480.
[25] E. Gabetta, L. Pareschi, and G. Toscani, *Relaxation schemes for nonlinear kinetic equations*, SIAM J. Numer. Anal., 34 (1997), pp. 2168–2194.
[26] I. M. Gamba and J. R. Haack, *A conservative spectral method for the Boltzmann equation with anisotropic scattering and the grazing collisions limit*, J. Comput. Phys., 270 (2014), pp. 40–57.
[27] I. M. Gamba, V. Panferov, and C. Villani, *On the Boltzmann equation for diffusively excited granular media*, Comm. Math. Phys., 246 (2004), pp. 503–541.
[28] I. M. Gamba, V. Panferov, and C. Villani, *Upper Maxwellian bounds for the spatially homogeneous Boltzmann equation*, Arch. Ration. Mech. Anal., 194 (2009), pp. 253–282.
[29] I. M. Gamba, S. Rjasanow, and W. Wagner, *Direct simulation of the uniformly heated granular Boltzmann equation*, Math. Comput. Model., 42 (2005), pp. 683–700.
[30] I. M. Gamba and S. H. Tharkabhushanam, *Spectral-Lagrangian methods for collisional models of non-equilibrium statistical states*, J. Comput. Phys., 228 (2009), pp. 2012–2036.
[31] I. M. Gamba and S. H. Tharkabhushanam, *Shock and boundary structure formation by spectral-Lagrangian methods for the inhomogeneous Boltzmann transport equation*, J. Comput. Math, 28 (2010), pp. 430–460.
[32] H. Grad, *Singular and nonuniform limits of solutions of the Boltzmann equation*, in Transport Theory SIAM-AMS Proc., Vol. I, Amer. Math. Soc., Providence, RI, 1969, pp. 269–308.
[33] M. Herty, L. Pareschi, and M. Seaid, *Discrete-velocity models and relaxation schemes for traffic flows*, SIAM J. Sci. Comput., 28 (2006), pp. 1582–1596.
[34] I. Ibragimov and S. Rjasanow, *Numerical solution of the Boltzmann equation on the uniform grid*, Computing, 69 (2002), pp. 163–186.
[35] S. Kawashima, *Global solution of the initial value problem for a discrete velocity model of the Boltzmann equation*, Proc. Japan Acad. Ser. A Math. Sci., 57 (1981), pp. 19–24.
[36] D. L., *Some applications of the method of moments for the homogeneous Boltzmann and KAC equations*, Arch. Rational Mech. Anal., 123 (1993), pp. 387–404.
[37] L. D. Landau, *Kinetic equation for the case of Coulomb interaction*, Zh. Eks. Teor. Phys., 7 (1937), pp. 203–209.
[38] L. D. Landau and E. M. Lifschitz, *Statistical Physics*, 3rd ed., Butterworth-Heinemann, Oxford, 1980.
[39] L. Mieussens, *Discrete-velocity models and numerical schemes for the Boltzmann-BGK equation in plane and axisymmetric geometries*, J. Comput. Phys., 162 (2000), pp. 429–466.
[40] S. Mischler, C. Mouhot, and M. Rodriguez Ricard, *Cooling process for inelastic Boltzmann equations for hard spheres.* I. *The Cauchy problem*, J. Stat. Phys., 124 (2006), pp. 655–702.
[41] C. Mouhot, *Rate of convergence to equilibrium for the spatially homogeneous Boltzmann equation with hard potentials*, Comm. Math. Phys., 261 (2006), pp. 629–672.
[42] C. Mouhot and L. Pareschi, *Fast algorithms for computing the Boltzmann collision operator*, Math. Comp., 75 (2006), pp. 1833–1852.
[43] C. Mouhot and C. Villani, *Regularity theory for the spatially homogeneous Boltzmann equation with cut-off*, Arch. Ration. Mech. Anal., 173 (2004), pp. 169–212.
[44] A. Munafo, J. R. Haack, I. M. Gamba, and T. E. Magin, *A spectral-Lagrangian Boltzmann solver for a multi-energy level gas*, J. Comput. Phys., 264 (2014), pp. 152–176.
[45] L. Pareschi and B. Perthame, *A Fourier spectral method for homogenous Boltzmann equations*, Transp. Theory Stat. Phys., 25 (2002), pp. 369–382.

[46] L. Pareschi and G. Russo, *Numerical solution of the Boltzmann equation.* I. *Spectrally accurate approximation of the collision operator*, SIAM J. Numerical Anal., 37 (2000), pp. 1217–1245.
[47] S. Rjasanow and W. Wagner, *Stochastic Numerics for the Boltzmann Equation*, Springer, Berlin, 2005.
[48] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, NJ, 1970.
[49] W. Wagner, *A convergence proof for Bird's direct simulation Monte Carlo method for the Boltzmann equation*, J. Stat. Phys., 66 (1992), pp. 1011–1044.
[50] B. Wennberg, *Entropy dissipation and moment production for the Boltzmann equation*, J. Stat. Phys., 86 (1997), pp. 1053–1066.
[51] C. Zhang and I. M. Gamba, *A conservative discontinuous Galerkin solver for homogeneous Boltzmann equation*, SIAM J. Numer. Anal., 56 (2018), pp. 3040–3070.
[52] C. Zhang and I. M. Gamba, *A conservative scheme for Vlasov Poisson Landau modeling collisional plasmas*, J. Comput. Phys., 340 (2017), pp. 470–497.