A CONVERGENT METHOD FOR LINEAR HALF-SPACE KINETIC EQUATIONS

QIN LI, JIANFENG LU, AND WEIRAN SUN

Abstract. We give a unified proof for the well-posedness of a class of linear half-space equations with general incoming data and construct a Galerkin method to numerically resolve this type of equations in a systematic way. Our main strategy in both analysis and numerics includes three steps: adding damping terms to the original half-space equation, using an inf-sup argument and even-odd decomposition to establish the well-posedness of the damped equation, and then recovering solutions to the original half-space equation. The proposed numerical methods for the damped equation is shown to be quasi-optimal and the numerical error of approximations to the original equation is controlled by that of the damped equation. This efficient solution to the half-space problem is useful for kinetic-fluid coupling simulations.

1. Introduction

In this paper we propose a Galerkin method for computing a class of half-space kinetic equation with given incoming data:

$$\begin{aligned} (v_1+u)\partial_x f + \mathcal{L}f = 0, \qquad & x\in[0,+\infty),\ v\in\mathbb{V}\,, \\ f\big|_{x=0} = \phi(v), \qquad & v_1+u>0\,. \end{aligned} \tag{1.1}$$

Here $u\in\mathbb{R}$ is a given constant and $\mathbb{V}\subseteq\mathbb{R}^\nu$ with $\nu\geq 1$. We are mainly concerned with several cases for $\mathbb{V}$: $\mathbb{V}=[-1,1]$, $\mathbb{V}=\mathbb{S}^{\nu-1}$ or $\mathbb{V}=\mathbb{R}^\nu$. Typical examples are $\mathbb{V}=[-1,1]$ or $\mathbb{V}=\mathbb{S}^2$ (with $u=0$) for neutron transport equations, $\mathbb{V}=\mathbb{R}^3$ (with arbitrary $u\in\mathbb{R}$) for linearized Boltzmann equations, and $\mathbb{V}=\mathbb{R}^1,\mathbb{R}^3$ for linearized BGK equations. If $\mathbb{V}=\mathbb{S}^2$, then in spherical coordinates $v=(\cos\phi,\cos\theta\sin\phi,\sin\theta\sin\phi)\in\mathbb{S}^2$ and $v_1=\cos\phi$. If $\mathbb{V}=\mathbb{R}^3$, then $v=(v_1,v_2,v_3)\in\mathbb{R}^3$ and v_1 is the first component of v. Our framework also covers systems with multiple species, in which case, the ambient space for the density function f is $(L^2(\,\mathrm{d}v))^d$ with $d\geq 1$, where d is the number of species and $\mathrm{d}v$ is the usual Lebesgue measure for $\mathbb{R}^d$ or the rotation-invariant measure on $\mathbb{S}^{d-1}$. The linear operator $\mathcal{L}$ only acts on the velocity variable v and is nonlocal in v in general. The specific structure and main assumptions regarding $\mathcal{L}$ will be given in Section 2.

While the half-space kinetic problem is an important problem on its own, the study of such equations in the form of (1.1) is also our first step in developing systematic algorithms that resolve couplings between kinetic and moment-closure equations. This type of couplings are common in domain-decomposition methods for solving kinetic equations with multi-scales. Multi-scale phenomenon naturally arises in gas dynamics where dense and dilute fluids coexist in a region. It is also common in transport phenomenon where the background medium has different scattering and absorption properties. Taking gas dynamics as an example, in a typical domain-decomposition method, one uses the original kinetic equation for the dilute part and a fluid equation for the dense side. The boundary conditions for the two sides are encoded in the density function near their interface. Due to the sharp variation in scales, the density function in general will have a finite jump across

Date: April 1, 2014.

We would like to express our gratitude to the NSF grant RNMS11-07444 (KI-Net), whose activities initiated our collaboration. J.L. would also like to thank Jian-Guo Liu for helpful discussions. W.S. would like to thank Cory Hauck for pointing out the reference [ES12]. The research of J.L. was supported in part by the Alfred P. Sloan Foundation and the National Science Foundation under award DMS-1312659. The research of W.S. was supported in part by the Simon Fraser University President's Research Start-up Grant PRSG-877723 and NSERC Discovery Individual Grant #611626.

the interface, thus boundary layers will form there. Half-space equations are exactly leading-order boundary-layer equations whose solutions bridge the gap between the fluid and kinetic boundary conditions. Therefore understanding the well-posedness of (1.1) and constructing accurate and efficient numerical schemes to resolve this equation will provide explicit characterization of the couplings. We note that due to the lack of efficient and accurate solvers of the half-space equation, various methods have been developed to bypass solving the half-space equation (see for examples [DJM05, LM12]). Our goal instead is to resolve the kinetic-fluid coupling through efficient and accurate solutions of the half-space equations, which will be the focus of our forthcoming work.

In the literature the well-posedness of equation (1.1) has been referred to as the Milne problem and has long been investigated [BSS84, BY12, CGS88, Gol08, UYY03] for various models. For example, for the conservative neutron transport equation where the null space of $\mathcal{L}$ consists of only constant functions, it has been shown [BSS84, BLP79] that there exists a unique solution $f \in L^\infty(\,\mathrm{d}x\,\mathrm{d}v)$ and the end-state of f at $x = \infty$ is a constant. In general for models where $\dim(\operatorname{Null}\mathcal{L}) > 1$, additional boundary conditions are needed for (1.1) to have a unique solution [CGS88]. A typical example is the linearized Boltzmann equation of one species over $v \in \mathbb{R}^3$. In this case we have $\dim(\operatorname{Null}\mathcal{L}) = 5$. The well-posedness of this equation with hard-sphere collisions is fully resolved in the fundamental work by Coron, Golse, and Sulem [CGS88]. In this work, it is shown that depending on the choices of u, one needs to prescribe various numbers of additional boundary conditions such that (1.1) is well-posed. These numbers of boundary conditions correspond to the counting of the incoming Euler characteristics at $x = \infty$. The proof in [CGS88] relies mainly on the energy method. Subsequently, a different proof using a variational formulation of (1.1) for the linearized Boltzmann equation is given in [UYY03]. The key idea in [UYY03] is to revise (1.1) by adding certain damping terms. The revised collision operator thus obtained is coercive and it enforces the end-state of f at $x = \infty$ to be zero. By the conservation properties of $\mathcal{L}$, the authors then show that (1.1) is well-posed for a large class of incoming data. One restriction in [UYY03] is that u cannot be chosen in the way such that the Mach number of the system is $-1, 1$, or 0. This restriction was later removed in [Gol08].

The variational formulation is also a common tool in proving the well-posedness of the neutron transport equations over general bounded domains Ω in $\mathbb{R}^\nu_x$. There is a vast literature in this direction and we will only review some of the main framework and results in [ES12] which are most relevant to us. Specifically, suppose $\Omega \subseteq \mathbb{R}^\nu_x$ is a bounded domain with a C^1 boundary such that the normal direction on $\partial\Omega$ is well-defined. Consider the neutron transport equation

$$v \cdot \nabla_x f + \mathcal{L} f = 0\,, \tag{1.2}$$

with the incoming boundary

$$f(x, v) = \phi(x, v)\,, \qquad x \in \partial\Omega,\ v \cdot n > 0\,, \tag{1.3}$$

where n is the normal direction on $\partial\Omega$ and $v \in \mathcal{S}^{\nu-1}$ is a unit vector. The scattering operator $\mathcal{L}$ is given by

$$\mathcal{L} f = c_a f + c_s \left(f - \int_{\mathcal{S}^{\nu-1}} \sigma(v \cdot v') f(v')\,\mathrm{d}\sigma \right),$$

where c_a and c_s are the absorption and scattering coefficients and σ is the scattering kernel. In [ES12] it is assumed that $0 < \alpha_0 < c_a < \widetilde{\alpha}_0$ and $0 \le c_s \le c_0$, thus $\mathcal{L}$ is subcritical and $\operatorname{Null}\mathcal{L} = \emptyset$. The main novelty of [ES12] is that one decomposes the solution f into its even and odd parts in v and imposes different regularities for these two parts. Using this mixed regularity, the authors of [ES12] write (1.2) into a variational form and verify that the bilinear operator involved satisfies an inf-sup condition over a properly chosen function space. Moreover, they show that for appropriately constructed Galerkin approximations, the bilinear operator satisfies the inf-sup condition over finite-dimensional approximation spaces as well. This

then shows the Galerkin approximation is quasi-optimal. We also note that the even-odd parity was widely used for transport equations, see for example [JPT01].

There are two main goals in our paper: First, we will generalize the analysis in [ES12, Gol08, UYY03] to obtain a unified proof for the well-posedness of half-space equations in the form of (1.1). Second, we will develop a systematic Galerkin method to numerically resolve (1.1) and obtain accuracy estimates for our scheme.

We now briefly explain our main results and compare them with previous ones in the literature. In terms of analysis, as in [CGS88], we show that with appropriate additional boundary conditions at $x = \infty$, equation (1.1) has a unique solution. The basic framework we use is the even-odd variational formulation in [ES12]. Compared with [ES12], here we allow the linear operator $\mathcal{L}$ to have a nontrivial null space and the background velocity u to be any arbitrary constant for general models. The number of additional boundary conditions will change with u. Due to the loss of coercivity of $\mathcal{L}$, if one directly applies the variational method in [ES12] then the bilinear operator $\mathcal{B}$ ceases to satisfy the inf-sup condition. To overcome this degeneracy, we utilize the ideas in [Gol08, UYY03] by adding damping terms to (1.1) and reconstructing solutions to (1.1) from the damped equation. In the case of linearized Boltzmann equation with a single species, we thus recover the results (in the L^2 spaces) in [Gol08, UYY03]. The main differences between our work and [Gol08, UYY03] are: First, we use a different variational formulation which is convenient for performing numerical analysis. Second, the reconstruction in [Gol08, UYY03] is restricted to a set of incoming data with a finite codimension such that the damping terms are identically zero. Here we use slightly different damping terms and we recover solutions to (1.1) from the damped equation for any incoming data. On the other hand, our main concern is the convergence and accuracy of the numerical scheme and the basic L^2-spaces are sufficient for this purpose. Therefore, except for the hard sphere case, we do not try to achieve decay estimates of the half-space solution to its end-state at $x = \infty$, while in the literature there are a lot of works that show subexponential or superpolynomial decay of the solution to its end-state for hard or soft potentials for the linearized Boltzmann equation(see for example [CLY04, WYY06, WYY07]).

Our analysis also applies to linearized Boltzmann equations with multiple species and linear neutron transport equations with critical or subcritical scatterings, thus providing an alternative proof to the well-posedness result (in the L^2-space) in [BY12].

In parallel with the analysis, numerically we first solve the damped half-space equation and then recover the solution to the original equation. We will use a spectral method and achieve quasi-optimal accuracy (for the damped equation) as in [ES12]. The spectral method dates back to Degond and Mas-Gallic [DMG87] for solving radiative transfer equations, and was later extended by Coron [Cor90] to solving the linearized BGK equation as well. Compared with these works, our approach differs in three ways: First, as a result of using the even-odd formulation, we can derive explicit boundary conditions for the approximate equations. In particular, the number of these boundary conditions is shown to be consistent with the number of the unknowns. Hence our discrete systems are always well-posed. This was not the case in [Cor90] where a least square method was used to solve a potentially overdetermined problem. Second, the method in [Cor90] used Hermite polynomials defined on the whole velocity space as their basis functions. This leads to severe Gibbs phenomenon, since in general the solution to the half-space equation has a finite jump at $x = 0$ and $v = -u$. Here we choose to use basis functions with jumps at $v = -u$ which naturally fit into the even-odd formulation. Third, we will treat the cases with arbitrary bulk velocities u in a uniform way while in [Cor90] different schemes are used for the cases $u = 0$ and $u \neq 0$.

There are also non-spectral methods developed for solving the half-space equations. For example, the work by Golse and Klar [GK95] uses Chapman-Enskog approximation with diffusive closures. The accuracy of these approximations would be hard to analyze: the iterative approach couples the error from the systematic

expansion truncation with the numerical error. Moreover, this work ([GK95]) also treats the cases $u = 0$ and $u \neq 0$ separately. The very recent work by Besse *et al.* [BBG$^+$11] treats the half-space problem as a boundary layer matching kinetics with the limiting fluid equation, where a Marshak type approximation [Mar47] is applied for boundary fluxes. Similar idea was also used in [Del03]. As shown already in [Cor90], in general the Marshak approximation does not yield accurate approximations to the half-space problem.

The layout of this paper as follows: in Section 2, we gather the basic information related to the linear operator $\mathcal{L}$ and set up the variational formulation. In Section 3, we summarize the main results in this paper which include both the analytical results and the numerical scheme. We give proofs in Section 4. Section 5 presents more details of the numerical scheme and numerical results for linearized BGK and linear transport equations.

2. Basic Setting

In this section, we collect the basic information about the collision operator $\mathcal{L}$ and present the variational formulation of a damped version of (1.1).

2.1. **Linear collision operator.**

In order to state the main assumptions imposed on $\mathcal{L}$, we first introduce some notations. Denote $\operatorname{Null}\mathcal{L}$ as the null space of $\mathcal{L}$. Let $\mathcal{P} : (L^2(\,\mathrm{d}v))^d \to \operatorname{Null}\mathcal{L}$ be the projection onto $\operatorname{Null}\mathcal{L}$. Define the weight function

$$a(v) = (1 + |v|)^{\omega_0}\,, \tag{2.1}$$

for some $0 \leq \omega_0 \leq 1$. The main assumptions on $\mathcal{L}$ are as follows:

(A1) $\mathcal{L} : (L^2(\,\mathrm{d}v))^d \to (L^2(\,\mathrm{d}v))^d$ is self-adjoint, nonnegative with its domain given by

$$\mathcal{D}(\mathcal{L}) = \{f \in (L^2(\,\mathrm{d}v))^d \big|\, a(v)f \in (L^2(\,\mathrm{d}v))^d\}\,,$$

where $a(v)$ is defined in (2.1).

(A2) $\mathcal{L} : (L^2(a\,\mathrm{d}v))^d \to (L^2(\frac{1}{a}\,\mathrm{d}v))^d$ is bounded, that is, there exists a constant $C_0 > 0$ such that

$$\|\mathcal{L}f\|_{(L^2(\frac{1}{a}\,\mathrm{d}v))^d} \leq C_0 \|f\|_{(L^2(a\,\mathrm{d}v))^d}\,.$$

(A3) $\operatorname{Null}\mathcal{L}$ is finite dimensional and $\operatorname{Null}\mathcal{L} \subseteq (L^p(\,\mathrm{d}v))^d$ for all $p \in [1, \infty)$.

(A4) $\mathcal{L}$ has a spectral gap: there exists $\sigma_0 > 0$ such that

$$\langle f,\ \mathcal{L}f\rangle \geq \sigma_0 \left\|\mathcal{P}^\perp f\right\|^2_{(L^2(a\,\mathrm{d}v))^d} \qquad \text{for any } f \in (L^2(a\,\mathrm{d}v))^d\,,$$

where $\mathcal{P}^\perp = \mathcal{I} - \mathcal{P}$ is the projection onto the null orthogonal space $(\operatorname{Null}\mathcal{L})^\perp$.

One operator that is of particular importance is $\mathcal{P}_1 : \operatorname{Null}\mathcal{L} \to \operatorname{Null}\mathcal{L}$ which is defined by

$$\mathcal{P}_1(f) = \mathcal{P}((v_1 + u)f) \qquad \text{for any } f \in \operatorname{Null}\mathcal{L}\,.$$

Note that $\mathcal{P}_1$ is a symmetric operator on the finite dimension space $\operatorname{Null}\mathcal{L}$. Therefore, its eigenfunctions form a complete set of basis of $\operatorname{Null}\mathcal{L}$. Denote H^+, H^-, H^0 as the eigenspaces of $\mathcal{P}_1$ corresponding to positive, negative, and zero eigenvalues respectively and denote their dimensions as

$$\dim H^+ = \nu_+, \qquad \dim H^- = \nu_-, \qquad \dim H^0 = \nu_0\,.$$

Let $X_{+,i}, X_{-,j}, X_{0,k}$ be the associated unit eigenfunctions with $1 \leq i \leq \nu_+$, $1 \leq j \leq \nu_-$, and $1 \leq k \leq \nu_0$ for $\nu_\pm, \nu_0 \neq 0$. Note that if any of $\nu_\pm, \nu_0$ is equal to zero, then we simply do not have any eigenfunction associated with the corresponding eigenspace. By their definitions, these eigenfunctions satisfy

$$\langle X_{\alpha,\gamma}, X_{\alpha',\gamma'}\rangle_v = \delta_{\alpha\alpha'}\delta_{\gamma\gamma'}\,, \qquad \langle (v+u)X_{\alpha,\gamma},\ X_{\alpha',\gamma'}\rangle_v = 0 \text{ if } \alpha \neq \alpha' \text{ or } \gamma \neq \gamma'\,,$$
$$\langle (v+u)X_{0,j},\ X_{0,k}\rangle_v = 0\,, \quad \langle (v+u)X_{+,j},\ X_{+,i}\rangle_v > 0\,, \qquad \langle (v+u)X_{-,j},\ X_{-,j}\rangle_v < 0\,,$$

where $\alpha \in \{+,-,0\}$, $\gamma \in \{i,j,k\}$, $1 \le i \le \nu_+$, $1 \le j \le \nu_-$, and $1 \le k \le \nu_0$.

Using the notation of $H^\pm, H^0$, we can now state the full equation that we want to study in this paper. Suppose $\mathcal{L}$ is a linear operator in v that satisfies (A1)-(A4). Our goal is to prove the well-posedness of, construct efficient numerical schemes, and obtain accuracy estimates for the following equation:

$$\begin{aligned} (v_1+u)\partial_x f + \mathcal{L}f = 0, \qquad & x \in [0,+\infty),\ v \in \mathbb{V}\,, \\ f\big|_{x=0} = \phi(v), \qquad & v_1+u>0\,, \\ f - f_\infty \in & (L^2(\,\mathrm{d}v\,\mathrm{d}x))^d\,, \end{aligned} \tag{2.2}$$

for some $f_\infty \in H^+ \oplus H^0$. The particular formulation about the end-state f_∞ was given in [CGS88] where the authors proved the well-posedness of the half-space linearized Boltzmann equation:

Theorem 2.1 ([CGS88])**.** *Let $\mathcal{L}$ be the linearized Boltzmann operator with a hard-sphere collision kernel. Then there exists a constant $\beta > 0$ and a unique $f_\infty \in H^+ \oplus H^0$ such that equation* (1.1) *has a unique solution f which satisfies*

$$f - f_\infty \in L^2(e^{2\beta x}\,\mathrm{d}x; L^2(a\,\mathrm{d}v))\,,$$

where $a(v) = 1 + |v|$.

Remark 2.1. The main result in [CGS88] is actually stronger than Theorem 2.1 where $f - f_\infty$ is shown to be in $L^\infty(e^{2\beta x}\,\mathrm{d}x; L^2(\,\mathrm{d}v))$. Here we content ourselves with the L^2-weighted space (in x) since L^2 suffices our needs in proving the quasi-optimal convergence of our numerical scheme.

Many well-known linear or linearized kinetic models satisfy the above assumptions for the collision operators. These include the classical linearized Boltzmann equations for either single-species system or multi-species with hard-sphere collisions and the linear neutron transport equations. The particular equations that we use as numerical examples are the isotropic neutron transport equation (NTE) with slab geometry and the linearized BGK equation. Similar analysis can be carried out to models satisfying (A1)-(A4) without extra difficulties. The main structure of these two equations are as follows. The linear operator of the isotropic NTE is the simplest scattering operator which has the form

$$\mathcal{L}f = f - \int_{-1}^{1} f(v)\,\mathrm{d}v\,. \tag{2.3}$$

In this case, $a(v) = 1 + |v| = \mathcal{O}(1)$ and $(L^2(a\,\mathrm{d}v))^d$ coincides with $(L^2(\,\mathrm{d}v))^d$.

The linearized BGK operator is the linearization of the nonlinear BGK operator, which is introduced as a simplified model that captures some fundamental behavior of the nonlinear Boltzmann equation. The collision operator of the nonlinear BGK is defined as

$$\mathcal{Q}[F] = F - \mathcal{M}[F]\,,$$

where $\mathcal{M}[F]$ is the local Maxwellian associated with F defined by

$$\mathcal{M}[F] = \frac{\rho}{\sqrt{2\pi\theta}} e^{-\frac{|v-u|^2}{2\theta}}\,,$$

where

$$\rho = \int_{\mathbb{R}} F\,\mathrm{d}v\,, \qquad \rho u = \int_{\mathbb{R}} vF\,\mathrm{d}v\,, \qquad \rho u^2 + \rho\theta = \int_{\mathbb{R}} v^2 F\,\mathrm{d}v\,.$$

For a given bulk velocity $u \in \mathbb{R}$, define the global Maxwellian with the steady state $(\rho, u, \theta) = (1, u, 1/2)$ as

$$M_u = \frac{1}{\sqrt{\pi}} e^{-|v-u|^2}\,.$$

Linearizing the operator $\mathcal{Q}$ around M by setting

$$F = M_u + \sqrt{M_u} f\,,$$

we obtain the linearized BGK operator

$$\mathcal{L}_u f = f - m_u,$$

where $m_u(v)$ is f projected onto the kernel space of $\mathcal{L}_u$. In the case of the 1D linearized BGK, one has:

$$\operatorname{Null}\mathcal{L}_u = \operatorname{span}\{\sqrt{M_u},\ v\sqrt{M_u},\ v^2\sqrt{M_u}\}\,.$$

Therefore, $m_u(v)$ is a quadratic function associated with a Maxwellian to 1/2 power:

$$m_u(v) = \big(\widetilde{\rho} + \widetilde{u}(v-u) + \tfrac{\widetilde{\theta}}{2}((v-u)^2-1)\big)\sqrt{M_u}\,,$$

where $(\widetilde{\rho}, \widetilde{u}, \widetilde{\theta})$ are defined in the way such that first three moments of $m(v)$ agree with those of f:

$$\langle f - m_u, v^k\sqrt{M_u}\rangle = \int_{\mathbb{R}} (f-m_u)v^k\sqrt{M_u}\,\mathrm{d}v = 0, \quad k = 0,1,2\,.$$

The half-space equation with the linearized BGK operator that centered at bulk velocity u is:

$$\begin{aligned} &v\partial_x f + \mathcal{L}_u f = 0,\\ &f|_{x=0} = \phi(v)\,, \qquad v > 0\,. \end{aligned} \tag{2.4}$$

Following the classical treatment of the half-space equations, we shift the center of the Maxwellian M_u to the origin by performing the change of variable $v - u \to v$. The half-space equation (2.4) then becomes

$$\begin{aligned} &(v+u)\partial_x f + \mathcal{L} f = 0,\\ &f|_{x=0} = \phi(v)\,, \qquad v + u > 0\,, \end{aligned} \tag{2.5}$$

where

$$\mathcal{L} f = f - m(v)\,, \qquad m(v) = m_0\,, \tag{2.6}$$

and the null space of $\mathcal{L}$ becomes

$$\operatorname{Null}\mathcal{L} = \operatorname{span}\{\sqrt{M},\ v\sqrt{M},\ v^2\sqrt{M}\}$$

with M is the global Maxwellian centered at the origin:

$$M = M_0 = \frac{1}{\sqrt{\pi}} e^{-v^2}\,.$$

It is well known that equation (2.5) is not necessarily well-posed unless appropriate boundary conditions are imposed at $x = \infty$. In this paper we will adopt the framework developed in [CGS88]. To illustrate the boundary conditions used in this setting, we first define a new set of basis functions for $\operatorname{Null}\mathcal{L}$. Let

$$\begin{cases} \chi_0 = \frac{1}{6^{1/2}\pi^{1/4}}\left(2v^2 - 3\right)\exp(-v^2/2),\\ \chi_\pm = \frac{1}{6^{1/2}\pi^{1/4}}\left(\sqrt{6}v \pm 2v^2\right)\exp(-v^2/2)\,, \end{cases} \tag{2.7}$$

such that

$$\begin{cases} \langle \chi_\alpha, \chi_\beta\rangle_v = \int_{\mathbb{R}} \chi_\alpha\chi_\beta\,\mathrm{d}v = \delta_{\alpha\beta}\,,\\ \langle (v+u)\chi_\alpha, \chi_\beta\rangle_v = 0\,, \qquad \alpha \neq \beta\,,\\ \langle (v+u)\chi_0, \chi_0\rangle_v = u_0 = u\,,\\ \langle (v+u)\chi_+, \chi_+\rangle_v = u_+ = u + c\,,\\ \langle (v+u)\chi_-, \chi_-\rangle_v = u_- = u - c\,, \end{cases} \tag{2.8}$$

where $\alpha, \beta \in \{+,-,0\}$, $c = \sqrt{3/2}$, and

$$\langle f, g\rangle_v = \int_{\mathbb{R}} fg\,\mathrm{d}v\,.$$

Using these new basis functions, we can decompose $\operatorname{Null}\mathcal{L}$ into subspaces: $\operatorname{Null}\mathcal{L} = H^+ \oplus H^- \oplus H^0$ with:

$$H^+ = \operatorname{span}\{\chi_\beta |\ u_\beta > 0\}, \quad H^- = \operatorname{span}\{\chi_\beta |\ u_\beta < 0\}, \quad H^0 = \operatorname{span}\{\chi_\beta |\ u_\beta = 0\},$$

where again $\beta \in \{+,-,0\}$. For each fixed $u \in \mathbb{R}$, denote the dimensions of these subspaces as

$$\dim H^+ = \nu_+, \qquad \dim H^- = \nu_-, \qquad \dim H^0 = \nu_0 .$$

Note that $\nu_\pm, \nu_0$ change with u. In particular, we have the following categories:

$$\begin{cases} u < -c: & (\dim H^+, \dim H^-, \dim H^0) = (0,3,0), \\ u = -c: & (\dim H^+, \dim H^-, \dim H^0) = (0,2,1), \\ -c < u < 0: & (\dim H^+, \dim H^-, \dim H^0) = (1,2,0), \\ u = 0: & (\dim H^+, \dim H^-, \dim H^0) = (1,1,1), \\ 0 < u < c: & (\dim H^+, \dim H^-, \dim H^0) = (2,1,0), \\ u = c: & (\dim H^+, \dim H^-, \dim H^0) = (2,0,1), \\ u > c: & (\dim H^+, \dim H^-, \dim H^0) = (3,0,0). \end{cases} \tag{2.9}$$

As mentioned in the introduction, we will present a variational formulation using the mixed regularity idea in [ES12] to prove Theorem 2.1 for general linear or linearized half-space equations when $\mathcal{L}$ satisfies (A1)-(A4). We need to overcome the difficulty due to the non-coercivity of $\mathcal{L}$. Although this degeneracy of $\mathcal{L}$ in some cases can be handled by carefully choosing appropriate function spaces for the variational formulation, we prefer to work with strictly dissipative operators. To this end, we utilize the idea developed in [Gol08, UYY03] to modify the original equation (2.5) by adding in damping terms. The particular damping term are chosen in the way such that we can easily recover the undamped equation (2.5) for any incoming data.

The rest of this section will be devoted to the description of the variational formulation of the damped half-space equation.

2.2. **Variational Formulation.** We will use $\mathbb{V} = \mathbb{R}^3$ as the setting to explain the variational formulation. Other spaces for v will work in a similar way. Let $u \in \mathbb{R}$ be given. We modify (2.2) by adding in damping terms. The modified equation for $f \in (L^2(\,\mathrm{d}x\,\mathrm{d}v))^d$ has the form

$$(v_1 + u)\partial_x f + \mathcal{L}_d f = 0 \tag{2.10}$$

with boundary condition

$$f\big|_{x=0} = \phi(v_1) \quad \text{for } v_1 + u > 0 . \tag{2.11}$$

Here the damped collision operator $\mathcal{L}_d$ is defined by

$$\begin{aligned} \mathcal{L}_d f =& \mathcal{L} f + \alpha \sum_{k=1}^{\nu_+} (v_1+u) X_{+,k} \left\langle (v_1+u) X_{+,k}, f \right\rangle_v \\ &+ \alpha \sum_{k=1}^{\nu_-} (v_1+u) X_{-,k} \left\langle (v_1+u) X_{-,k}, f \right\rangle_v + \alpha \sum_{k=1}^{\nu_0} (v_1+u) X_{0,k} \left\langle (v_1+u) X_0, f \right\rangle_v \\ &+ \alpha \sum_{k=1}^{\nu_0} (v_1+u) \mathcal{L}^{-1}((v_1+u)X_{0,k}) \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), f \right\rangle_v . \end{aligned} \tag{2.12}$$

Here the constant α satisfies that $0 < \alpha \ll 1$. More specific requirement of α will be explained in Section 5. We define the shifted "even" and "odd" parts of a function as

$$f^+(v) = \frac{f(v_1,v_2,v_3) + f(-2u - v_1, v_2, v_3)}{2}, \qquad f^-(v) = \frac{f(v_1,v_2,v_3) - f(-2u - v_1, v_2, v_3)}{2} \tag{2.13}$$

such that $f = f^+ + f^-$ and

$$f^{\pm}(-u+v_1,v_2,v_3) = \pm f^{\pm}(-u-v_1,v_2,v_3)\,.$$

Define the function space

$$\Gamma = \left\{ f \in (L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d \;\middle|\; (v_1+u)\partial_x f^+ \in (L^2(\tfrac{1}{a}\,\mathrm{d}v\,\mathrm{d}x))^d \right\}. \tag{2.14}$$

Then Γ is a Hilbert space with the inner product

$$\langle f, g\rangle_\Gamma = \int_{\mathbb{R}}\int_{\mathbb{R}^3} f\cdot g\, a\,\mathrm{d}v\,\mathrm{d}x + \int_{\mathbb{R}}\int_{\mathbb{R}^3} (v_1+u)\partial_x f^+ \cdot (v_1+u)\partial_x g^+\,\tfrac{1}{a}\,\mathrm{d}v\,\mathrm{d}x\,.$$

Thus the norm of Γ is equivalent to

$$\|f\|_{(L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d} + \|(v_1+u)\partial_x f^+\|_{(L^2(\frac{1}{a}\,\mathrm{d}v\,\mathrm{d}x))^d}\,.$$

Moreover, every element $g \in \Gamma$ has a well-defined trace:

$$\mathcal{T} : \Gamma \to L^2(|v_1+u|\,\mathrm{d}v) \tag{2.15}$$

such that

$$\mathcal{T} g = g^+\big|_{x=0}\,, \qquad \text{for all } g \in C([0,\infty); (L^2(a\,\mathrm{d}v))^d)\,, \tag{2.16}$$

and

$$\int_{\mathbb{R}^3} |v_1+u||g^+|^2\,\mathrm{d}v < \infty\,. \tag{2.17}$$

Now we define a bilinear operator $\mathcal{B} : \Gamma\times\Gamma \to \mathbb{R}$ such that

$$\begin{aligned}
\mathcal{B}(f,\psi) = &-\left\langle f^-, (v_1+u)\partial_x\psi^+\right\rangle_{x,v} + \left\langle (v_1+u)\partial_x f^+, \psi^-\right\rangle_{x,v} + \langle \psi, \mathcal{L} f\rangle_{x,v} \\
&+ \alpha\sum_{k=1}^{\nu_+} \left\langle \langle (v_1+u)X_{+,k}, \psi\rangle_v\,,\ \langle (v_1+u)X_{+,k}, f\rangle_v \right\rangle_x \\
&+ \alpha\sum_{k=1}^{\nu_-} \left\langle \langle (v_1+u)X_{-,k}, \psi\rangle_v\,,\ \langle (v_1+u)X_{-,k}, f\rangle_v \right\rangle_x \\
&+ \alpha\sum_{k=1}^{\nu_0} \left\langle \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), \psi\right\rangle_v\,,\ \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), f\right\rangle_v \right\rangle_x \\
&+ \alpha\sum_{k=1}^{\nu_0} \left\langle \langle (v_1+u)X_{0,k}, \psi\rangle_v\,,\ \langle (v_1+u)X_{0,k}, f\rangle_v \right\rangle_x + \left\langle |v_1+u| f^+, \psi^+\right\rangle_{x=0}\,.
\end{aligned} \tag{2.18}$$

The inner product $\langle\cdot,\cdot\rangle_{x,v}$ in (2.18) is

$$\langle f,\ g\rangle_{x,v} = \int_{\mathbb{R}^3}\int_{\mathbb{R}} f\cdot g\,\mathrm{d}x\,\mathrm{d}v\,.$$

The variational formulation of (2.10) has the form

$$\mathcal{B}(f,\psi) = l(\psi)\,, \qquad \text{for every } \psi \in \Gamma\,. \tag{2.19}$$

Here the linear operator $l(\cdot)$ is given by

$$l(\psi) = 2\int_{v_1+u>0} (v_1+u)\,\phi\,\psi^+\,\mathrm{d}v\,, \tag{2.20}$$

where ϕ is the given incoming data and ψ^+ is the even (with respect to $-u$) part of ψ as defined in (2.13).

3. Main results

In this section we state the main analytical results of this work, and present the ideas of our numerical algorithms. Our strategy is to first solve the damped equation (2.10) and then recover from it the solution to the original undamped equation (2.2). Hence, in the first part we state the well-posedness of the damped equation and the quasi-optimality of the associated Galerkin method for the damped equation. In the second part we explain the details of the recovery procedures and state the convergence result for the original equation.

3.1. **Solution of the damped equation.**

Proposition 3.1 (Well-posedness)**.** *Let $\phi \in (L^2(a(v)\mathbf{1}_{v_1+u>0}\,\mathrm{d}v))^d$ and $0 < \alpha \ll 1$. Then there exists a unique $f \in \Gamma$ such that* (2.19) *holds. Moreover, f satisfies that*

$$(v_1+u)\partial_x f \in (L^2(\tfrac{1}{a}\,\mathrm{d}v\,\mathrm{d}x))^d$$

and it solves the damped half-space equation in the sense of distributions

$$\begin{aligned}
&(v_1+u)\partial_x f + \mathcal{L}_d f = \\
&\quad = (v_1+u)\partial_x f + \mathcal{L} f + \alpha \sum_{k=1}^{\nu_+} (v_1+u)X_{+,k}\left\langle (v_1+u)X_{+,k}, f\right\rangle_v \\
&\quad\quad + \alpha \sum_{k=1}^{\nu_-} (v_1+u)X_{-,k}\left\langle (v_1+u)X_{-,k}, f\right\rangle_v + \alpha \sum_{k=1}^{\nu_0} (v_1+u)X_{0,k}\left\langle (v_1+u)X_{0,k}, f\right\rangle_v \\
&\quad\quad + \alpha \sum_{k=1}^{\nu_0} (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k})\left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), f\right\rangle_v = 0
\end{aligned} \tag{3.1}$$

with the boundary conditions (defined in the trace sense at $x=0$)

$$f|_{x=0} = \phi(v)\,, \qquad v_1+u>0\,. \tag{3.2}$$

Moreover, if $a(v) = 1+|v|$, then there exists $\beta > 0$ such that $(L^2(e^{2\beta x}\,\mathrm{d}x; L^2(a\,\mathrm{d}v)))^d$.

Remark 3.1. Note that $f - f_\infty$ for the neutron transport equations satisfy the exponential decay as $x \to \infty$ by the comment after (2.3).

Using the variational formulation (2.19), we will solve the damped equation (3.1) by a Galerkin method.

Proposition 3.2 (Approximations in $\mathbb{R}^3$)**.** *Suppose $\{\psi_n^{(1)}\}_{n=1}^\infty$ is an orthonormal basis of $(L^2(\,\mathrm{d}v_1))^d$ such that*

- *$\psi_{2n-1}^{(1)}(v_1)$ is odd and $\psi_{2n}^{(1)}(v_1)$ is even in v_1 with respect to $-u$ for any $n \geq 1$;*
- *$(v_1+u)\psi_{2n}^{(1)}(v_1) \in \operatorname{span}\{\psi_1^{(1)}, \cdots, \psi_{2n+1}^{(1)}\}$ for each $n \geq 1$.*

Suppose $\{\psi_n^{(2)}\}, \{\psi_n^{(3)}\}_{n=1}^\infty$ are orthonormal bases for $(L^2(\,\mathrm{d}v_2))^d$ and $(L^2(\,\mathrm{d}v_3))^d$ respectively. Define the closed subspace Γ_{NK} as

$$\Gamma_{NK} = \left\{ g(x,v) \in \Gamma \,\middle|\, g(x,v) = \sum_{m,n=1}^{K} \sum_{k=1}^{2N+1} g_{kmn}(x)\psi_k^{(1)}(v_1)\psi_m^{(2)}(v_2)\psi_n^{(3)}(v_3),\ g_k \in H^1(\,\mathrm{d}x) \right\}.$$

Then there exists a unique $f_{NK} \in \Gamma_{NK}$ such that

$$f_{NK}(x,v) = \sum_{m,n=1}^{K} \sum_{k=1}^{2N+1} a_{kmn}(x)\psi_k^{(1)}(v_1)\psi_m^{(2)}(v_2)\psi_n^{(3)}(v_3)\,, \tag{3.3}$$

which satisfies

$$\mathcal{B}(f_{NK}, g) = l(g)\,, \qquad \text{for every } g \in \Gamma_{NK}\,, \tag{3.4}$$

where $\mathcal{B}$ and l are defined in (2.18) *and* (2.20) *respectively. The coefficients $\{a_{kmn}(x)\}$ satisfy that*

$$a_{kmn}(\cdot) \in C^1[0,\infty) \cap H^1(0,\infty), \qquad 1 \le k \le 2N+1\,,\ 1 \le m, n \le K\,.$$

The following Proposition reformulates (3.4) into an ODE with explicit boundary conditions.

Proposition 3.3. *Let*

$$\mathsf{A} = \left(\left\langle (v_1+u)\psi_k^{(1)},\ \psi_j^{(1)} \right\rangle_{v_1} \right)_{(2N+1)\times(2N+1)}\,.$$

Define two 6-tensors $\mathfrak{A}$ and $\mathfrak{B}$ as

$$\begin{aligned} \mathfrak{A} &= \mathsf{A} \otimes I \otimes I = \left(\mathsf{A}_{ik}\delta_{mj}\delta_{nl}\right)_{(2N+1)^2\times K^2\times K^2}\,, \\ \mathfrak{B}^{ijl}_{kmn} &= -\left\langle \psi_k^{(1)}(v_1)\psi_m^{(2)}(v_2)\psi_n^{(3)}(v_3),\ \mathcal{L}_d\left(\psi_i^{(1)}(v_1)\psi_j^{(2)}(v_2)\psi_l^{(3)}(v_3)\right)\right\rangle_v\,, \end{aligned} \tag{3.5}$$

for $1 \le i,k \le 2N+1$, $1 \le j,m \le K$, and $1 \le l,n \le K$. Then the variational form (3.4) *is equivalent to the following ODE for the coefficients $a_k^{(mn)}(x)$:*

$$\sum_{m,n=1}^{K}\sum_{k=1}^{2N+1} \mathfrak{A}^{ijl}_{kmn}\partial_x a_{kmn}(x) = \sum_{m,n=1}^{K}\sum_{k=1}^{2N+1} \mathfrak{B}^{ijl}_{kmn} a_{kmn}(x), \tag{3.6}$$

together with the boundary conditions at $x = 0$:

$$\begin{aligned} &\sum_{k=1}^{N+1}\left\langle (v_1+u)\psi_{2k-1}^{(1)},\, \psi_{2i}^{(1)}\right\rangle_{v_1} a_{2k-1,jl}(0) + \sum_{k=1}^{N}\left\langle |v_1+u|\psi_{2k}^{(1)}, \psi_{2i}^{(1)}\right\rangle_{v_1} a_{2k,jl}(0) \\ &\qquad = 2\int_{v_1+u>0} (v_1+u)\,\phi\,\psi_{2i}^{(1)}(v_1)\psi_j^{(2)}(v_2)\psi_k^{(3)}(v_3)\,\mathrm{d}v\,, \end{aligned} \tag{3.7}$$

for $i = 1,\cdots,N$ and $j,l = 1,2,\cdots,K$.

Since our numerical examples are both in 1-dimension, we apply Proposition 3.2 to $\mathbb{V} \subseteq \mathbb{R}^1$ and obtain the following Corollary for two special cases:

Corollary 3.4 (Approximations in $\mathbb{R}^1$)**.** *Let $\mathbb{V} = \mathbb{R}^1$ or $\mathbb{V} = [-1,1]$. Let $u \in \mathbb{R}$ be arbitrary if $\mathbb{V} = \mathbb{R}^1$ and $u = 0$ if $\mathbb{V} = [-1,1]$. Suppose $\{\psi_n\}_{n=1}^{\infty}$ is an orthonormal basis of $(L^2(\,\mathrm{d}v))^d$ such that*

- *ψ_{2n-1} is odd and ψ_{2n} is even in v with respect to $-u$ for any $n \ge 1$;*
- *$(v+u)\psi_{2n}(v) \in \operatorname{span}\{\psi_1,\cdots,\psi_{2n+1}\}$ for each $n \ge 1$.*

Define the closed subspace Γ_N as

$$\Gamma_N = \left\{ g(x,v) \in \Gamma \,\middle|\, g(x,v) = \sum_{k=1}^{2N+1} g_k(x)\psi_k(v),\ g_k \in H^1(\,\mathrm{d}x) \right\}\,.$$

Then there exists a unique $f_N \in \Gamma_N$ such that

$$f_N(x,v) = \sum_{k=1}^{2N+1} a_k(x)\psi_k(v)\,, \qquad a_k(x) \in C^1[0,\infty),\ 1 \le k \le 2N+1\,, \tag{3.8}$$

which satisfies

$$\mathcal{B}(f_N, g) = l(g)\,, \qquad \text{for every } g \in \Gamma_N\,, \tag{3.9}$$

where $\mathcal{B}$ and l are defined in (2.18) *and* (2.20) *respectively. Moreover, the variational form* (3.9) *is equivalent to the following ODE for the coefficients $a_k(x)$ together with the boundary conditions at $x = 0$:*

$$\sum_{k=1}^{2N+1} \mathsf{A}_{kl}\partial_x a_k(x) = \sum_{k=1}^{2N+1} \mathsf{B}_{kl} a_k(x)\,, \tag{3.10}$$

$$\sum_{k=1}^{N+1} \langle (v+u)\psi_{2k-1},\, \psi_{2j}\rangle_v\, a_{2k-1}(0) + \sum_{k=1}^{N} \langle |v+u|\psi_{2k}, \psi_{2j}\rangle_v\, a_{2k}(0) = 2\int_{v+u>0} (v_1+u)\,\phi\,\psi_{2j}\,\mathrm{d}v\,, \tag{3.11}$$

where $1 \le j \le N$ and

$$\mathsf{A}_{kl} = \langle (v+u)\psi_k,\ \psi_l\rangle_v\ , \qquad \mathsf{B}_{kl} = -\langle \psi_k,\ \mathcal{L}_d\psi_l\rangle_v\ , \qquad 1 \le i,j \le 2N+1\,. \tag{3.12}$$

We now state the quasi-optimality of the Galerkin approximation.

Proposition 3.5 (Quasi-Optimality)**.** *Suppose f_N is an approximation defined in Proposition 3.4, and f is the unique solution to the damped equation. Then there exits a constant C_0 such that*

$$\|f - f_N\|_\Gamma \le C_0 \inf_{w\in\Gamma_N} \|f - w\|_\Gamma\,,$$

where $\|\cdot\|_\Gamma$ is the norm defined in (2.14)*.*

Numerically, the approximate solutions f_{NK} in (3.6) (or f_N (3.10) in 1D) will be solved using the method of generalized eigenvalues. In particular, we define the generalized eigenvalues and its associated eigen-tensor for $(\mathfrak{A}, \mathfrak{B})$ as $\lambda \in \mathbb{R}$ and $\eta = (\eta_{kmn})_{(2N+1)\times K\times K}$ such that

$$\mathfrak{A}\eta = \sum_{m,n=1}^{K}\sum_{k=1}^{2N+1} \mathfrak{A}^{ijl}_{kmn}\,\eta_{kmn} = \lambda \sum_{m,n=1}^{K}\sum_{k=1}^{2N+1} \mathfrak{B}^{ikl}_{kmn}\,\eta_{kmn} \tag{3.13}$$

for all $1 \le i \le 2N+1$, $1 \le j,l \le K$. When reduced to 1D system, the generalized eigenvalue problem for (A, B) becomes

$$\mathsf{A}\,\eta = \lambda \mathsf{B}\,\eta\,. \tag{3.14}$$

To solve for the coefficient $a(x)$, we take (3.10) as an example. Define $\gamma(x) = \eta^{\mathrm{T}}\mathsf{B}\,a(x)$ and multiply (3.10) by η^T from the left. We then obtain the equation for γ as

$$\eta^{\mathrm{T}}\mathsf{A}\,\partial_x a(x) = \eta^{\mathrm{T}}\mathsf{B}\,a(x) \quad \Rightarrow \quad \lambda\partial_x\gamma(x) = \gamma(x)\,.$$

If $\lambda = 0$, then we immediately get the constraint

$$\gamma(x) = \eta^{\mathrm{T}}\mathsf{B}\,a = 0\,. \tag{3.15}$$

If $\lambda \neq 0$, then we have

$$\gamma(x) = e^{x/\lambda}\gamma(0)\,.$$

Depending on the signs of the eigenvalues, γ either grows exponentially to infinity or decays exponentially to zero; as we look for bounded decaying solutions, this gives us constraints to $\gamma(0)$ for the growing modes: If $\lambda > 0$, then we have the constraints

$$\gamma = \eta^{\mathrm{T}}\mathsf{B}\,a = 0. \tag{3.16}$$

Note that we do not need constraints for modes with negative eigenvalues. The total number of constraints in the form of (3.16) is determined by the number of positive generalized eigenvalues. The following Proposition gives the signature of (A, B):

Proposition 3.6. *Let $\mathfrak{A}, \mathfrak{B}$ be the 6-tensors defined in* (3.5) *with any arbitrary $u \in \mathbb{R}$ and $N, K \geq 1$. Then there are NK^2 positive generalized eigenvalues, NK^2 negative eigenvalues, and K^2 zero eigenvalue for the pair $(\mathfrak{A}, \mathfrak{B})$.*

A direct application of Proposition 3.6 to the 1D case gives:

Corollary 3.7. *Let A, B be the matrices defined in* (3.12) *with any arbitrary $u \in \mathbb{R}$ and $N \geq 1$. Then there are N positive generalized eigenvalues, N negative eigenvalues, and one zero eigenvalue for the pair (A, B).*

Therefore in total we have $N+1$ equations for $a(0)$ given by the constraints (3.15) and (3.16). Combining them with the N equations given by the boundary conditions (3.11) for $a(0)$, we get $2N+1$ equations for $2N+1$ unknowns $\{a_k(0)\}$. The linear system (3.10) for a is then uniquely solvable, which further uniquely determines the approximate solution $f_N(x, v)$ by (3.8).

3.2. **Recovery of the undamped solution.** From the solution of the damped equation (2.12), we can explicitly construct solutions to the original undamped equation (2.2). Specifically, for each $1 \leq i \leq \nu_+$, let $g_{+,i}$ be the solution to (2.12) with boundary conditions given by $X_{+,i}$:

$$g_{+,i}|_{x=0} = X_{+,i}, \quad v_1 + u > 0.$$

Similarly, for each $1 \leq j \leq \nu_0$, denote $g_{0,j}$ as the solution to (3.1) where the incoming boundary data is given by $X_{0,j}$. Let C be the block matrix defined by

$$C = \begin{pmatrix} C_{++} & C_{+0} \\ C_{0+} & C_{00} \end{pmatrix}, \tag{3.17}$$

where

$$\begin{aligned} C_{++,ii'} &= \langle (v_1+u) X_{+,i}, g_{+,i'} \rangle, & C_{+0,ij'} &= \langle (v_1+u) X_{+,i}, g_{0,j'} \rangle, \\ C_{0+,ji'} &= \langle (v_1+u) X_{0,j}, g_{+,i'} \rangle, & C_{00,jj'} &= \langle (v_1+u) X_{0,j}, g_{0,j'} \rangle \end{aligned}$$

for $1 \leq i, i' \leq \nu_+$ and $1 \leq j, j' \leq \nu_0$. Define the coefficient vector $\eta = \left(\eta_{+,1}, \cdots \eta_{+,\nu_+}, \eta_{0,1}, \cdots, \eta_{0,\nu_0}\right)^{\mathrm{T}}$ such that

$$C\eta = U_f, \tag{3.18}$$

where $U_f = \left(u_{+,1}, \cdots, u_{+,\nu_+}, u_{0,1}, \cdots, u_{0,\nu_0}\right)^{\mathrm{T}}$ with $u_{+,i} = \langle (v_1+u) X_{+,i}, f \rangle_{x=0}$, $1 \leq i \leq \nu_+$ and $u_{0,j} = \langle (v_1+u) X_{0,j},\ f \rangle_{x=0}$, $1 \leq j \leq \nu_0$.

Proposition 3.8 (Recovery)**.** *Let $\phi \in (L^2(a(v)\mathbf{1}_{v_1+u>0}\,\mathrm{d}v))^d$. Then C is invertible, hence* (3.18) *is uniquely solvable. Moreover,*

$$f_\phi = f - \sum_{i=1}^{\nu_+} \eta_{+,i}(g_{+,i} - X_{+,i}) - \sum_{j=1}^{\nu_0} \eta_{0,j}(g_{0,j} - X_{0,j})$$

is the unique solution to the half-space equation

$$\begin{aligned} (v_1+u)\partial_x f + \mathcal{L} f &= 0, \\ f|_{x=0} &= \phi(v)\,, \qquad v_1 + u > 0\,, \\ f_\phi - f_{\phi,\infty} &\in L^2(e^{2\beta x}\,\mathrm{d}x; L^2(\,\mathrm{d}v))\,, \end{aligned} \tag{3.19}$$

where $f_{\phi,\infty} \in H^+ \oplus H^0$ is the end-state given by

$$f_{\phi,\infty} = \sum_{j=1}^{\nu_+} \eta_{+,j} X_{+,j} + \sum_{k=1}^{\nu_0} \eta_{0,k} X_{0,k}\,.$$

Applying the recovery construction in the finite dimensional numerical space Γ_N, we obtain the numerical solution of the undamped equation (2.2) in Γ_N. The resulting approximation for the undamped equation is almost quasi-optimal with a correction term.

Proposition 3.9. *Suppose $f_{\phi,N}$ is constructed as in Proposition 3.8 with f, $g_{+,i}$, $g_{0,j}$ being numerical approximations obtained in Proposition 3.4 with appropriate boundary conditions. Suppose f_ϕ is the unique solution to the equation* (2.5)*. Then there exists a constant C_0 such that*

$$\|f_\phi - f_{\phi,N}\|_\Gamma \le C_0 \left(\inf_{w\in\Gamma_N} \|f_\phi - w\|_\Gamma + \inf_{w\in\Gamma_N} \|f - w\|_\Gamma + \delta_N \|f\|_{(L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d} \right),$$

where $\|\cdot\|_\Gamma$ is the norm defined in (2.14) *and*

$$\delta_N := \sum_{i=1}^{\nu_+} \inf_{w\in\Gamma_N} \|g_{+,i} - w\|_\Gamma + \sum_{j=1}^{\nu_0} \inf_{w\in\Gamma_N} \|g_{0,j} - w\|_\Gamma .$$

Remark 3.2. Note that in the above reconstruction scheme, the solutions $g_{+,i}$ for $1 \le i \le \nu_+$ and $g_{0,j}$ for $1 \le j \le \nu_0$ can be precomputed, as they do not depend on the prescribed incoming data ϕ. In particular, we can use a higher order approximation (larger N) for these functions.

4. Convergence Proof

In this section we will show the proofs for the Propositions stated in Section 3.

4.1. Inf-sup condition and well-posedness.

In order to prove the well-posedness of the weak formulation (2.19), we show that the bilinear operator $\mathcal{B}$ satisfies an inf-sup condition. More precisely, we have

Proposition 4.1 (Inf-sup)**.** *Let Γ and $\mathcal{B}$ be the function space and the bilinear operator defined in* (2.14) *and* (2.18) *respectively. Then $\mathcal{B} : \Gamma \times \Gamma \to \mathbb{R}$ satisfies that*

$$\begin{aligned} &\sup_{\|f\|_\Gamma=1} \mathcal{B}(f,\psi) \ge \kappa_0 \|\psi\|_\Gamma , \qquad \text{for any } \psi \in \Gamma , \\ &\sup_{\|\psi\|_\Gamma=1} \mathcal{B}(f,\psi) \ge \kappa_0 \|f\|_\Gamma , \qquad \text{for any } f \in \Gamma \end{aligned} \tag{4.1}$$

for some constant $\kappa_0 > 0$.

Proof. Note that $\mathcal{B}$ is symmetric in its variables. Hence it suffices to show that the second condition in (4.1) holds. To this end, let $f \in \Gamma$ be arbitrary. We only need to find an appropriate ψ such that

$$\mathcal{B}(f,\psi) \ge \kappa_0 \|f\|_\Gamma^2 , \qquad \|\psi\|_\Gamma \le \kappa_1 \|f\|_\Gamma . \tag{4.2}$$

Indeed, if ψ satisfies (4.2), then one can simply let $\Psi = \frac{\psi}{\|\psi\|_\Gamma}$ and obtain the second inequality in (4.1) (with a different constant). The construction of such ψ will be carried out in two steps. First, let $\psi_1 = f$. Then

$$\begin{aligned} \mathcal{B}(f,\psi_1) = \langle f, \mathcal{L}f\rangle_{x,v} &+ \alpha \sum_{k=1}^{\nu_+} \left\langle \langle (v_1+u)X_{+,k}, f\rangle_v^2 \right\rangle_x \\ &+ \alpha \sum_{k=1}^{\nu_-} \left\langle \langle (v_1+u)X_{-,k}, f\rangle_v^2 \right\rangle_x + \alpha \sum_{k=1}^{\nu_0} \left\langle \langle (v_1+u)X_0, f\rangle_v^2 \right\rangle_x \\ &+ \alpha \sum_{k=1}^{\nu_0} \left\langle \langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), f\rangle_v^2 \right\rangle_x + \langle |v_1+u| f^+, f^+\rangle_{x=0} . \end{aligned}$$

Write

$$f = f^\perp + \sum_{i=1}^{\nu_+} f_{+,i} X_{+,i} + \sum_{j=1}^{\nu_-} f_{-,j} X_{+,j} + \sum_{k=1}^{\nu_0} f_{0,k} X_{0,k} ,$$

where $f^\perp = \widetilde{\mathcal{P}} f \in (\mathrm{Null}\,\mathcal{L})^\perp$. By (A3), if we chose $0 < \alpha \ll 1$, then

$$\begin{aligned}
\mathcal{B}(f,\psi_1) \geq{}& \sigma_0 \|f^\perp\|^2_{(L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d} + \frac{\alpha}{4}\sum_{k=1}^{\nu_+} \gamma_{+,k}^2 \left\langle f_{+,k}^2 \right\rangle_x + \frac{\alpha}{4}\sum_{k=1}^{\nu_+} \gamma_{-,k}^2 \left\langle f_{-,k}^2 \right\rangle_x + \left\langle |v_1+u| f^+, f^+ \right\rangle_{x=0} \\
&- \frac{\alpha}{4}\sum_{k=1}^{\nu_+} \left\langle \left\langle (v_1+u) X_{+,k}, f^\perp \right\rangle_v^2 \right\rangle_x - \frac{\alpha}{4}\sum_{k=1}^{\nu_-} \left\langle \left\langle (v_1+u) X_{-,k}, f^\perp \right\rangle_v^2 \right\rangle_x \\
\geq{}& \frac{\sigma_0}{2} \|f^\perp\|^2_{(L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d} + \frac{\alpha}{4}\sum_{k=1}^{\nu_+} \gamma_{+,k}^2 \left\langle f_{+,k}^2 \right\rangle_x + \frac{\alpha}{4}\sum_{k=1}^{\nu_-} \gamma_{-,k}^2 \left\langle f_{-,k}^2 \right\rangle_x + \left\langle |v+u| f^+, f^+ \right\rangle_{x=0} ,
\end{aligned} \tag{4.3}$$

where $\gamma_\pm$'s are defined as

$$\begin{aligned}
&\gamma_{+,i} := \left\langle (v_1+u) X_{+,i}, X_{+,i} \right\rangle_v > 0 , \qquad && 0 \leq i \leq \nu_+ , \\
&\gamma_{-,j} := -\left\langle (v_1+u) X_{-,j}, X_{-,j} \right\rangle_v > 0 , \qquad && 0 \leq j \leq \nu_- .
\end{aligned} \tag{4.4}$$

In addition, if $\nu_0 \neq 0$, then

$$\begin{aligned}
\mathcal{B}(f,\psi_1) \geq{}& \sigma_0 \|f^\perp\|^2_{(L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d} + \alpha \sum_{k=1}^{\nu_0} \left\langle \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), f \right\rangle_v^2 \right\rangle_x \\
\geq{}& \frac{\sigma_0}{2} \|f^\perp\|^2_{(L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d} + \frac{\alpha}{4\nu_0} \left\langle \left(\sum_{k,m=1}^{\nu_0} \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), X_{0,m} \right\rangle_v f_{0,m} \right)^2 \right\rangle_x \\
&- \frac{\alpha}{4}\sum_{k=1}^{\nu_0} \left\langle \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), \sum_{m=1}^{\nu_+} f_{+,k} X_{+,k} \right\rangle_v^2 \right\rangle_x \\
&- \frac{\alpha}{4}\sum_{k=1}^{\nu_0} \left\langle \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), \sum_{m=1}^{\nu_+} f_{-,k} X_{-,k} \right\rangle_v^2 \right\rangle_x \\
&- \frac{\alpha}{4}\sum_{k=1}^{\nu_0} \left\langle \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), f^\perp \right\rangle_v^2 \right\rangle_x .
\end{aligned} \tag{4.5}$$

Since the matrix $\left(\left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), X_{0,m} \right\rangle_v \right)$ is strictly positive, there exists a constant $c_0 > 0$ such that

$$\left\langle \sum_{k=1}^{\nu_0} \left(\sum_{m=1}^{\nu_0} \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), X_{0,m} \right\rangle_v f_{0,m} \right)^2 \right\rangle_x \geq c_0 \sum_{k=1}^{\nu_0} \langle f_{0,m} \rangle_x^2 . \tag{4.6}$$

Hence by multiplying (4.3) by a large enough number and adding it to (4.5), we have

$$\mathcal{B}(f,\psi_1) \geq \kappa_{0,1} \|f\|^2_{(L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d} \qquad \text{for some } \kappa_{0,1} > 0 \ , \tag{4.7}$$

provided $0 < \alpha \ll 1$. Next, let

$$\psi_2 = \frac{1}{(1+|v_1+u|+|v_2|+|v_3|)^{\omega_0}} (v_1+u)\partial_x f^+ .$$

We claim that $\psi_2 \in \Gamma$. Indeed, by the definition of $a(v)$, one can find two constants $c_1, c_2 > 0$ such that

$$\frac{c_1}{a(v)} \leq \frac{1}{(1+|v_1+u|+|v_2|+|v_3|)^{\omega_0}} \leq \frac{c_2}{a(v)} .$$

Here the constants c_1, c_2 depend on u. Thus $\psi_2 \in (L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d$ because

$$\|\psi_2\|_{(L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d} \leq \|(v_1+u)\partial_x f^+\|_{\left(L^2(\frac{1}{a}\,\mathrm{d}v\,\mathrm{d}x)\right)^d} \leq \|f\|_\Gamma .$$

Moreover the definition of ψ_2 implies that

$$\psi_2^+ = 0 \in (L^2(\tfrac{1}{a}\,\mathrm{d}v\,\mathrm{d}x))^d .$$

Hence $\psi_2 \in \Gamma$ and it satisfies

$$\|\psi_2\|_\Gamma \le \|f\|_\Gamma \,. \tag{4.8}$$

Using ψ_2 in $\mathcal{B}$, we have

$$\begin{aligned}
\mathcal{B}(f,\psi_2) =& \left\langle (v_1+u)\partial_x f^+, \psi_2 \right\rangle + \langle \psi_2, \mathcal{L} f\rangle + \alpha \sum_{k=1}^{\nu_+} \left\langle \langle (v_1+u)X_{+,k}, \psi_2\rangle_v \langle (v_1+u)X_{+,k}, f\rangle_v \right\rangle_x \\
&+ \alpha \sum_{k=1}^{\nu_-} \left\langle \langle (v_1+u)X_{-,k}, \psi_2\rangle_v \langle (v_1+u)X_{-,k}, f\rangle_v \right\rangle_x \\
&+ \alpha \sum_{k=1}^{\nu_0} \left\langle \langle (v_1+u)X_{0,k}, \psi_2\rangle_v \langle (v_1+u)X_{0}, f\rangle_v \right\rangle_x \\
&+ \alpha \sum_{k=1}^{\nu_0} \left\langle \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), \psi_2\right\rangle_v \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,k}), f\right\rangle_v \right\rangle_x \\
\ge& \|(v_1+u)\partial_x f^+\|^2_{(L^2(\frac{1}{a}\,\mathrm{d}v\,\mathrm{d}x))^d} - \kappa_2 \|f\|^2_{(L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d} \,,
\end{aligned}$$

for some constant $\kappa_2 > 0$. Hence by taking $\kappa_3 > 0$ large enough, we have that

$$\mathcal{B}(f, \kappa_3\psi_1 + \psi_2) \ge \kappa_0 \|f\|^2_\Gamma \,, \tag{4.9}$$

for some $\kappa_0 > 0$. Recall that by the definition of ψ_1 and (4.8), we also have

$$\|\kappa_3\psi_1 + \psi_2\|_\Gamma \le \sqrt{1+\kappa_3}\, \|f\|_\Gamma \,,$$

which, together with (4.9), shows the inf-sup property of $\mathcal{B}$ on $\Gamma \times \Gamma$.

Next we show that if $a(v) = 1 + |v|$, then there exists $\beta > 0$ such that $f \in (L^2(e^{2\beta x}\,\mathrm{d}x; L^2(a\,\mathrm{d}v)))^d$. The proof will be along the same line for the general case of $a(v)$. We use the standard way to incorporate the exponential into the bilinear form by changing f by $e^{-\beta x} f$. The new bilinear form $\mathcal{B}_\beta$ is

$$\mathcal{B}_\beta(f,\psi) = \mathcal{B}(f,\psi) - \beta \langle (v_1+u) f, \psi\rangle_{x,v} \,,$$

where $\mathcal{B}(f,\psi)$ is defined in (2.18). Note that by Cauchy-Schwartz, if we choose $0 < \beta \ll \alpha$, then by the spectral gap assumption (A4), we have

$$\begin{aligned}
&\left|\beta \langle (v_1+u) f,\ \psi_1\rangle_{x,v}\right| \le \frac{1}{2}\mathcal{B}(f,\psi) \,, \\
&\left|\beta \langle (v_1+u) f,\ \psi_2\rangle_{x,v}\right| \le \frac{1}{2}\mathcal{B}(f,\psi) + \frac{1}{2}\left\langle (v_1+u)\partial_x f^+,\ \psi_2\right\rangle_{x,v} \,.
\end{aligned}$$

Hence, this extra β-term will not affect the inf-sup estimate. □

Given the inf-sup condition of the bilinear form $\mathcal{B}$, we can now prove the well-posedness of the weak formulation (2.19) by using the Babuška-Aziz lemma [BA72]. There are two parts in this lemma and we recall its statement below.

Theorem 4.2 (Babuška-Aziz)**.** *Suppose Γ is a Hilbert space and $\mathcal{B} : \Gamma \times \Gamma \to \mathbb{R}$ is a bilinear operator on Γ. Let $l : \Gamma \to \mathbb{R}$ be a bounded linear functional on Γ.*

(a) If $\mathcal{B}$ satisfies the boundedness and inf-sup conditions on Γ such that

- *there exists a constant $c_0 > 0$ such that $|\mathcal{B}(f,g)| \le c_0 \|f\|_\Gamma \|g\|_\Gamma$ for all $f, g \in \Gamma$;*
- *there exists a constant $\kappa_0 > 0$ such that (4.1) holds,*

then there exists a unique $f \in \Gamma$ which satisfies

$$\mathcal{B}(f, \psi) = l(\psi)\,, \qquad \text{for any } \psi \in \Gamma\,.$$

(b) Suppose Γ_N is a finite-dimensional subspace of Γ. If in addition $\mathcal{B} : \Gamma_N \times \Gamma_N \to \mathbb{R}$ satisfies the inf-sup condition on Γ_N, then there exists a unique solution f_N such that

$$\mathcal{B}(f_N, \psi_N) = l(\psi_N)\,, \qquad \text{for any } \psi_N \in \Gamma_N\,.$$

Moreover, f_N gives a quasi-optimal approximation to the solution f in (a), that is, there exists a constant κ_1 such that

$$\|f - f_N\|_\Gamma \leq \kappa_1 \inf_{w \in \Gamma_N} \|f - w\|_\Gamma\,.$$

Now we proceed to the proof of Proposition 3.1.

Proof of Proposition 3.1. It is straightforward to verify the boundedness of $\mathcal{B}$ and l defined in (2.18) and (2.20). The well-posednes of the variational form is then an immediate consequence of Proposition 4.1 and part (a) of the Babuška-Aziz lemma. In order to show that $(v_1 + u)\partial_x f \in (L^2(\frac{1}{a}\,\mathrm{d}v\,\mathrm{d}x))^d$, we note that the damped equation (3.1) holds in the sense of distributions by choosing the test function $\psi \in C_c^\infty((0,\infty)\times\mathbb{R})$. Thus

$$(v_1 + u)\partial_x f = \beta(v_1 + u)f - \mathcal{L}_d(f) \in (L^2(\tfrac{1}{a}\,\mathrm{d}v\,\mathrm{d}x))^d\,.$$

By the density argument this implies that

$$\left\langle (v_1+u)\partial_x f^-,\ \psi^+ \right\rangle_{x,v} + \left\langle (v_1+u)\partial_x f^+,\ \psi^- \right\rangle_{x,v} - \beta\left\langle (v_1+u)\psi, f\right\rangle_{x,v} + \left\langle \psi, \mathcal{L}_d f\right\rangle_{x,v} = 0\,,$$

for all $\psi \in C^\infty(0,\infty)$. Therefore, if we choose $\phi \in C^\infty[0,\infty)$ and integrate by parts in the variational form (2.19), then boundary terms satisfy

$$\left\langle (v_1+u)f^-,\ \psi^+ \right\rangle_v + \left\langle |v_1+u| f^+,\ \psi^+ \right\rangle_v = 2\int_{v_1+u>0} (v_1+u)\,\phi\,\psi^+\,\mathrm{d}v \qquad \text{at } x = 0\,,$$

which implies,

$$\int_{v_1+u>0} (v_1+u) f\psi^+\,\mathrm{d}v = \int_{v_1+u>0} (v_1+u)\,\phi\,\psi^+\,\mathrm{d}v \qquad \text{at } x = 0\,.$$

Since $\psi^+ \in C^\infty(0,\infty)$ is arbitrary, we have $f = \phi$ at $x = 0$ when $v_1 + u > 0$. In the case of $a(v) = 1 + |v|$, the unique solution that is in the space $(L^2(e^{-2\beta x}\,\mathrm{d}x; L^2(\,\mathrm{d}v)))^d$ is then given by $e^{-\beta x} f$. □

4.2. **The generalized eigenvalue problem.** Let us now prove the properties of the generalized eigenvalue problems (3.13) and (3.14).

Proof of Proposition 3.6 and Corollary 3.7. We first verify that (A, B) has N positive, N negative, and one zero generalized eigenvalues. By the definition of B and the strict coercivity of $\mathcal{L}_d$, the matrix B is symmetric and strictly positive definite. Hence the numbers of positive, negative, and zero generalized eigenvalues are the same with the signature of the matrix $\mathsf{B}^{-1}\mathsf{A}$. Furthermore, by the Sylvestre's Law of Inertia, $\mathsf{B}^{-1}\mathsf{A}$ and A have the same signature. Hence, we only need to count the numbers of positive, negative, and zero eigenvalues of A. Note that by the definition of the basis functions ψ_k in (5.4), A is independent of u since one can perform a change of variable $v + u \to v$ in each entry in A. Thus we only need to study the matrix A_0 with $u = 0$. Change the order of the basis functions such that

$$(\widetilde{\psi}_1, \widetilde{\psi}_2, \cdots, \widetilde{\psi}_{N+1}, \widetilde{\psi}_{N+2}, \cdots \widetilde{\psi}_{2N+1}) = (\psi_1, \psi_3, \cdots, \psi_{2N+1}, \psi_2, \cdots, \psi_{2N}) = P(\psi_1, \psi_2, \cdots, \psi_{2N+1})\,,$$

where P is the similarity matrix. Defined $\widetilde{\mathsf{A}}_0 = P\mathsf{A}_0P^{-1}$. Then $\widetilde{\mathsf{A}}_0$ and A_0 have the same signature. By the even/odd properties of $\widetilde{\psi}_i$, the matrix $\widetilde{\mathsf{A}}_0$ has the form

$$\widetilde{\mathsf{A}}_0 = \begin{pmatrix} 0 & A_1 \\ A_1^{\mathrm{T}} & 0 \end{pmatrix},$$

where $A_1 = \left(\int_{\mathbb{R}} v\psi_{2i}\psi_{2j+1}\right)_{N\times(N+1)}$. Suppose $\eta = (\eta_{1,1}, \cdots, \eta_{1,N}, \eta_{2,1}, \cdots, \eta_{2,N+1})^T = (\eta_1^{\mathrm{T}}, \eta_2^{\mathrm{T}})^{\mathrm{T}}$ is an eigenvector of $\widetilde{\mathsf{A}}_0$ with eigenvalue λ. Then one has

$$A_1\eta_2 = \lambda\eta_1\,, \qquad A_1^{\mathrm{T}}\eta_2 = \lambda\eta_1\,.$$

It is clear that $(\eta_1, -\eta_2)$ is also an eigenvector of $\widetilde{\mathsf{A}}_0$ and the associated eigenvalue is $-\lambda$. This shows the eigenvalues of $\widetilde{\mathsf{A}}_0$ appear in pairs. Since A_1 has a full rank N, we have that $\operatorname{rank}\widetilde{\mathsf{A}}_0 = 2N$. Therefore $\widetilde{\mathsf{A}}_0$, thus A_0 and A, has N positive eigenvalues, N negative eigenvalues, and one zero eigenvalue.

Now we claim that each generalized eigenpair (λ, v) of A gives rise to K^2 eigenpairs of $\mathfrak{A}$. Indeed, let $\{w^{(m)}\}_{m=1}^K$ be a set of basis vectors of $\mathbb{R}^K$. Choose the 3-tensor $\eta^{(mn)} = v \otimes w^{(m)} \otimes w^{(n)}$. Then

$$\mathfrak{A}\eta = (\mathsf{A}\otimes I\otimes I)(v\otimes w^{(m)}\otimes w^{(n)}) = (\mathsf{A}v)\otimes w^{(m)}\otimes w^{(n)} = \lambda\eta^{(mn)}\,,$$

for any $1 \le m, n \le K$. Thus each $(\lambda, v\otimes w^m \otimes w^{(n)})$ is an eigenpair of $\mathfrak{A}$.

Note that we can also view $\mathfrak{A}$ and $\mathfrak{B}$ as two matrices of size $((2N+1)K^2)\times((2N+1)K^2)$ by defining a bijection between the indices

$$\Upsilon : \{(i,j,l)|\, i = 1, \cdots, 2N+1, j, l = 1, \cdots, K\} \to \{1, \cdots, (2N+1)K^2\}\,.$$

Then $\mathfrak{B}$ is symmetric and positive definite and $\mathfrak{A}$ is symmetric. Therefore, by a similar argument as for (A, B) using Sylvestre's Law of Inertia, the number of positive, negative, and zero generalized eigenvalues agree with the signature of $\mathfrak{A}$. This shows there are NK^2 positive, NK^2 negative, and K^2 zero generalized eigenvalues for $(\mathfrak{A}, \mathfrak{B})$. □

4.3. **Recovering the Undamped Equation.** In this part we explain the procedures to recover the undamped equation from the damped one, thus proving Proposition 3.8.

Proof of Proposition 3.8. First we recall [CGS88] the uniqueness property of the solution to (3.19): if f is a solution to (3.19) which satisfies $f \in (L^2(a\,\mathrm{d}v\,\mathrm{d}x))^d$ and $(v_1+u)\partial_x f \in (L^2(\frac{1}{a}\,\mathrm{d}v\,\mathrm{d}x))^d$, then f must be unique. For the convenience of the reader, we brief explain its proof: Suppose h is a solution to the half-space equation (3.19) with incoming data $\phi = 0$. Then $\int_{\mathbb{R}^3}(v_1+u)h^2\,\mathrm{d}v$ is decreasing in x. Since there exists $h_\infty \in H^+ \oplus H^0$ such that $h - h_\infty \in (L^2(\,\mathrm{d}v\,\mathrm{d}x))^d$, we can find a sequence x_k such that

$$\int_{\mathbb{R}^3}(v_1+u)h^2(x_k, v)\,\mathrm{d}v \to \int_{\mathbb{R}^3}(v_1+u)h_\infty^2(x_k, v)\,\mathrm{d}v \ge 0\,.$$

Hence $\int_{\mathbb{R}^3}(v_1+u)h^2(x,v)\,\mathrm{d}v \ge 0$ for all $x \ge 0$. This holds in particular at $x = 0$. Since the incoming data is zero at $x = 0$, the outgoing data at $x = 0$ must also be zero and $\int_{\mathbb{R}^3}(v_1+u)h^2(x_k,v)\,\mathrm{d}v = 0$ for all $x \ge 0$. The conservation property of the half-space equation then implies that $h(x,\cdot) \in (\operatorname{Null}\mathcal{L})^\perp$. By multiplying the equation by h and integrate over v, we have $\langle h, Lh\rangle = 0$ for all $x \ge 0$. Hence $\widetilde{\mathcal{P}}h = 0$ for all $x \ge 0$ by the spectral gap of $\mathcal{L}$ in (A4). Therefore $h \equiv 0$ and the solution to the half-space equation is unique.

Denote

$$\begin{aligned}
\vec{U}_+ &= \left(\langle (v_1+u)X_{+,1}, f\rangle_v\,,\ \cdots\,,\ \langle (v_1+u)X_{+,\nu_+}, f\rangle_v\right)^{\mathrm{T}}\,,\\
\vec{U}_- &= \left(\langle (v_1+u)X_{-,1}, f\rangle_v\,,\ \cdots\,,\ \langle (v_1+u)X_{-,\nu_-}, f\rangle_v\right)^{\mathrm{T}}\,,\\
\vec{U}_0 &= \left(\langle (v_1+u)X_{0,1}, f\rangle_v\,,\ \cdots\,,\ \langle (v_1+u)X_{0,\nu_0}, f\rangle_v\right)^{\mathrm{T}}\,,\\
\vec{U}_{\mathcal{L},0} &= \left(\left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,1}), f\right\rangle_v\,,\ \cdots\,,\ \left\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,1})X_{0,\nu_0}, f\right\rangle_v\right)^{\mathrm{T}}\,,
\end{aligned}$$

and

$$\vec{U} = \left(\vec{U}_+^{\mathrm{T}},\ \vec{U}_-^{\mathrm{T}},\ \vec{U}_0^{\mathrm{T}},\ \vec{U}_{\mathcal{L},0}^{\mathrm{T}}\right)^{\mathrm{T}}\,. \tag{4.10}$$

We will separate two different cases according to $\dim H^0$.

Case 1: $\dim(H^0) = 0$. In this case, there are only various numbers of X_+'s and X_-'s. Hence (3.1) reduces to

$$\begin{aligned}
(v_1+u)\partial_x f + \mathcal{L}f + \alpha\sum_{k=1}^{\nu_+}(v_1+u)X_{+,k}\langle (v_1+u)X_{+,k}, f\rangle_v\\
+\alpha\sum_{k=1}^{\nu_-}(v_1+u)X_{-,k}\langle (v_1+u)X_{-,k}, f\rangle_v = 0\,,
\end{aligned} \tag{4.11}$$

and $\vec{U}$ reduces to

$$\vec{U} = \left(\vec{U}_+^{\mathrm{T}},\ \vec{U}_-^{\mathrm{T}}\right)^{\mathrm{T}}\,.$$

Multiplying (4.11) by $X_{+,k}, X_{-,j}$ and integrating over $v\in\mathbb{R}$, we obtain a linear system for $\vec{U}$:

$$\partial_t\vec{U} + A_1\vec{U} = 0\,, \tag{4.12}$$

where the coefficient matrix is diagonal:

$$A_1 = \left(\begin{array}{c|c} \alpha D_+ & 0\\ \hline 0 & -\alpha D_-\end{array}\right)\,, \tag{4.13}$$

where D_+, D_- are positive definite if they exist and

$$D_+ = \mathrm{diag}(\gamma_{+,1},\cdots,\gamma_{+,\nu_+})\,,\qquad D_- = \mathrm{diag}(-\gamma_{-,1},\cdots,-\gamma_{-,\nu_-})\,,$$

where $\gamma_{\pm,k}$ are defined as in (4.4). Since solutions to (4.11) are in $(L^2(\,\mathrm{d}v\,\mathrm{d}x))^d$, it is clear that

$$\langle (v_1+u)X_{-,j}, f(0,\cdot)\rangle_v = \langle (v_1+u)X_{-,j}, f(x,\cdot)\rangle_v = 0\,,\qquad \text{for all } 1\le j\le\nu_- \text{ and } x\ge 0\,.$$

Hence (4.11) further reduces to

$$(v_1+u)\partial_x f + \mathcal{L}f + \alpha\sum_{k=1}^{\nu_+}(v_1+u)X_{+,k}\langle (v_1+u)X_{+,k}, f\rangle_v = 0\,. \tag{4.14}$$

Given any $\phi\in L^2$, suppose f is the solution to (4.14) (or (3.1)) with ϕ as the incoming data. Let g_k be the solution to (4.14) (or (3.1))with the incoming data

$$\phi_k = X_{+,k}\,.$$

We claim that

$$\langle (v_1+u)X_{+,k},\ g_k\rangle_v\Big|_{x=0} \neq 0\,,\qquad \text{for each } k\ge 1\,. \tag{4.15}$$

Suppose not. Then there exists $k\ge 1$ such that

$$\langle (v_1+u)X_{+,k},\ g_k\rangle_v = 0\,,\qquad \text{at } x = 0\,.$$

Since the coefficient matrix A_1 in the ODE (4.12) is diagonal, we have

$$\langle (v_1+u)X_{+,k},\ g_k\rangle_v = 0\,, \qquad \text{for all } x\geq 0\,. \tag{4.16}$$

This further implies that g_k satisfies (3.1) with the incoming data $\phi_k = X_{+,k}$. By the uniqueness of the solution to (3.1), $g_k = X_{+,k}$, which contradicts (4.16). Hence (4.15) holds. Now define

$$\eta_{+,k} = \left.\frac{\langle (v_1+u)X_{+,k}, f\rangle_v}{\langle (v_1+u)X_{+,k}, g_k\rangle_v}\right|_{x=0}, \qquad \text{for each } 1\leq k\leq \nu_+\,, \tag{4.17}$$

and

$$\widetilde{g} = \sum_{k=1}^{\nu_+}\eta_{+,k}g_k\,, \qquad \widetilde{\phi} = \sum_{k=1}^{\nu_+}\eta_{+,k}X_{+,k}\,. \tag{4.18}$$

Then $f-\widetilde{g}$ satisfies (4.14) and

$$\langle (v_1+u)X_{+,k},\ f-\widetilde{g}\rangle_v\big|_{x=0} = \langle (v_1+u)X_{+,k},\ f-\widetilde{g}\rangle_v(x) = 0\,, \qquad \text{for all } x\geq 0\,.$$

Therefore we have

$$\begin{aligned}
&(v_1+u)\partial_x(f-\widetilde{g}) + \mathcal{L}(f-\widetilde{g}) = 0\,,\\
&f-\widetilde{g}|_{x=0} = \phi(v) - \sum_{k=1}^{\nu_+}\eta_{+,k}X_{+,k}\,, \qquad v_1+u>0\,,\\
&f-\widetilde{g}\in L^2(e^{2\beta x}\,\mathrm{d}x; L^2(\,\mathrm{d}v))\,,
\end{aligned}$$

for some $\beta\geq 0$. Thus, $f_\phi = f-g+\widetilde{\phi}$ satisfies (2.2) with

$$f_{\phi,\infty} = \sum_{k=1}^{\nu_+}\eta_{+,k}X_{+,k}\,.$$

Case 2: $\dim H^0\neq 0$. By multiplying $X_{+,j}, X_{-,i}, X_{0,k}, \mathcal{L}^{-1}(v_1\chi_{0,m})$ to (3.1) and integrating over $v\in\mathbb{R}^3$, we have

$$\partial_x\vec{U} + A_2\vec{U} = 0\,, \tag{4.19}$$

where the coefficient matrix A_2 is

$$A_2 = \left(\begin{array}{cc|c|c}
\alpha D_+ & & 0 & \alpha A_{21}\\
& -\alpha D_- & & \alpha A_{22}\\
\hline
\multicolumn{2}{c|}{0} & 0 & \alpha B\\
\hline
\alpha A_{21}^{\mathrm{T}} & \alpha A_{22}^{\mathrm{T}} & I+\alpha B & \alpha D
\end{array}\right), \tag{4.20}$$

where $D_\pm$ are positive diagonal matrices

$$D_+ = \mathrm{diag}(\gamma_{+,1},\cdots,\gamma_{+,\nu_+})_{\nu_+\times\nu_+}\,, \qquad D_- = \mathrm{diag}(-\gamma_{-,1},\cdots,-\gamma_{-,\nu_-})_{\nu_-\times\nu_-}\,,$$

and

$$\begin{aligned}
A_{21,ik} &= \big(\big\langle (v_1+u)X_{+,i},\ \mathcal{L}^{-1}((v_1+u)X_{0,k})\big\rangle_v\big)_{\nu_+\times\nu_0}\,,\\
A_{22,jk} &= \big(\big\langle (v_1+u)X_{-,j},\ \mathcal{L}^{-1}((v_1+u)X_{0,k})\big\rangle_v\big)_{\nu_-\times\nu_0}\,,\\
B_{ij} &= \big\langle (v_1+u)X_{0,i},\ \mathcal{L}^{-1}((v_1+u)X_{0,j})\big\rangle_{v,\nu_0\times\nu_0}\,,\\
D_{ij} &= \big\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,i}),\ \mathcal{L}^{-1}((v_1+u)X_{0,j})\big\rangle_{v,\nu_0\times\nu_0}\,,
\end{aligned}$$

where B is symmetric positive definite and D is symmetric. Note that if we define

$$Q = \left(\begin{array}{cc|c|c} I & & 0 & 0 \\ & I & & \\ \hline 0 & & (\alpha B)^{1/2}(I+\alpha B)^{-1/2} & 0 \\ \hline 0 & & 0 & I \end{array}\right),$$

and

$$\widetilde{A}_2 = \left(\begin{array}{cc|c|c} \alpha D_+ & & 0 & \alpha A_{21} \\ & -\alpha D_- & & \alpha A_{22} \\ \hline 0 & & 0 & (I+\alpha B)^{1/2}(\alpha B)^{1/2} \\ \hline \alpha A_{21}^{\mathrm{T}} & \alpha A_{22}^{\mathrm{T}} & (I+\alpha B)^{1/2}(\alpha B)^{1/2} & \alpha D \end{array}\right).$$

Then

$$A_2 = Q^{-1}\widetilde{A}_2 Q\,. \tag{4.21}$$

Thus A_2 and $\widetilde{A}_2$ have the same signature. In particular, they have the same number of negative eigenvalues. Now we count the number of negative eigenvalues of $\widetilde{A}_2$. Let

$$P = \left(\begin{array}{cc|c|c} I & & 0 & 0 \\ & I & & \\ \hline 0 & & I & 0 \\ \hline -A_{21}^{\mathrm{T}}D_+^{-1} & A_{22}^{\mathrm{T}}D_-^{-1} & 0 & I \end{array}\right).$$

Then P is non-singular and

$$A_3 = P\widetilde{A}_2 P^{\mathrm{T}} = \left(\begin{array}{cc|c|c} \alpha D_+ & & 0 & 0 \\ & -\alpha D_- & & \\ \hline 0 & & 0 & (I+\alpha B)^{1/2}(\alpha B)^{1/2} \\ \hline 0 & & (I+\alpha B)^{1/2}(\alpha B)^{1/2} & \alpha D_1 \end{array}\right),$$

where D_1 is symmetric and

$$D_1 = D - A_{21}^{\mathrm{T}}D_+^{-1}A_{21} + A_{22}^{\mathrm{T}}D_-^{-1}A_{22}\,.$$

By Sylvester's law of inertia, the matrices $\widetilde{A}_2$ and A_3, thus A_2 and A_3, have the same number of negative eigenvalues. The total number of negative eigenvalues of A_3 is determined by that of the submatrix $\begin{pmatrix} 0 & (I+\alpha B)^{1/2}(\alpha B)^{1/2} \\ (I+\alpha B)^{1/2}(\alpha B)^{1/2} & \alpha D_1 \end{pmatrix}$. Define

$$P_1 = \begin{pmatrix} (I+\alpha B)^{-1/4}(\alpha B)^{-1/4} & 0 \\ 0 & (I+\alpha B)^{-1/4}(\alpha B)^{-1/4} \end{pmatrix}.$$

Then

$$A_4 = P_1 \begin{pmatrix} 0 & (I+\alpha B)^{1/2}(\alpha B)^{1/2} \\ (I+\alpha B)^{1/2}(\alpha B)^{1/2} & \alpha D_1 \end{pmatrix} P_1^{\mathrm{T}} = \begin{pmatrix} 0 & I \\ I & \alpha D_2 \end{pmatrix},$$

where

$$D_2 = (I+\alpha B)^{-1/4}(\alpha B)^{-1/4} D_1 (I+\alpha B)^{-1/4}(\alpha B)^{-1/4}\,.$$

Note that D_2 is symmetric. Hence, D_2 has a complete set of eigenvectors. Let $(\lambda, \boldsymbol{E}) = (\lambda, (\boldsymbol{e}_1, \boldsymbol{e}_2)^{\mathrm{T}})$ be an eigenpair of A_4. Then

$$\begin{pmatrix} 0 & I \\ I & \alpha D_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_1^{\mathrm{T}} \\ \boldsymbol{e}_2^{\mathrm{T}} \end{pmatrix} = \lambda \begin{pmatrix} \boldsymbol{e}_1^{\mathrm{T}} \\ \boldsymbol{e}_2^{\mathrm{T}} \end{pmatrix}, \tag{4.22}$$

which is equivalent to

$$\boldsymbol{e}_2 = \lambda \boldsymbol{e}_1\,, \qquad {\boldsymbol{e}_1}^{\mathrm{T}} + \alpha D_2 \boldsymbol{e}_2^{\mathrm{T}} = \lambda \boldsymbol{e}_2^{\mathrm{T}}\,.$$

Note that $\lambda \neq 0$. Since D_2 is symmetric, it has a complete set of orthogonal eigenvectors. Let $\boldsymbol{e}$ be an arbitrary eigenvector of D with eigenvalue $\lambda_{\boldsymbol{e}}$ and take $\boldsymbol{e}_2 = \boldsymbol{e}$. Then

$$\boldsymbol{e}_1 = \frac{1}{\lambda}\boldsymbol{e}\,, \qquad \frac{1}{\lambda}\boldsymbol{e}^{\mathrm{T}} + \alpha D_2 \boldsymbol{e}^{\mathrm{T}} = \lambda \boldsymbol{e}^{\mathrm{T}}\,. \tag{4.23}$$

Thus

$$\frac{1}{\lambda} + \alpha\lambda_{\boldsymbol{e}} - \lambda = 0\,, \tag{4.24}$$

which has exactly one negative solution for λ. Since the set of eigenvectors of D_2 is complete, the matrix A_4 has exactly ν_0 negative eigenvalues. Together with D_-, we have that A_3, thus A_2, has exactly $\nu_- + \nu_0$ negative eigenvalues, which prescribes $\nu_- + \nu_0$ conditions on $\vec{U}$ such that

$$\boldsymbol{E}_k \cdot \vec{U}(x) = 0\,, \qquad 1 \leq k \leq \nu_- + \nu_0\,, \quad x \geq 0\,, \tag{4.25}$$

where $\boldsymbol{E}_k$ are the eigenvectors associated with negative eigenvalues. Write each $\boldsymbol{E}_k$ as

$$\boldsymbol{E}_k = (\boldsymbol{e}_{k,+},\ \boldsymbol{e}_{k,-},\ \boldsymbol{e}_{k,0},\ \boldsymbol{e}_{k,\mathcal{L},0})^{\mathrm{T}}\,.$$

Now we make the following claim:

C1. The matrix given by

$$\boldsymbol{E} = \begin{pmatrix} \boldsymbol{e}_{1,-} & & \boldsymbol{e}_{1,\mathcal{L},0} \\ & \cdots & \\ \boldsymbol{e}_{\nu_-+\nu_0,-} & & \boldsymbol{e}_{\nu_-+\nu_0,\mathcal{L},0} \end{pmatrix}_{(\nu_-+\nu_0)\times(\nu_-+\nu_0)}$$

is nonsingular.

Proof of Claim **C***1.* Suppose not. Let $\mathcal{N}_2$ be the space spanned by the eigenvectors of A_2 with negative eigenvalues. Then there exists a nontrivial vector in $\mathcal{N}_2$ which takes the form

$$\widehat{\boldsymbol{E}} = (\widehat{\boldsymbol{e}}_+, 0, \widehat{\boldsymbol{e}}_0, 0)^{\mathrm{T}}\,.$$

By (4.21), if $\boldsymbol{E} = (\boldsymbol{e}_+, \boldsymbol{e}_-, \boldsymbol{e}_0, \boldsymbol{e}_{\mathcal{L},0})^{\mathrm{T}}$ is an eigenvector of A_2 with eigenvalue λ, then $\boldsymbol{F} = Q(\boldsymbol{e}_+, \boldsymbol{e}_-, \boldsymbol{e}_0, \boldsymbol{e}_{\mathcal{L},0})^{\mathrm{T}}$ is an eigenvector of $\widetilde{A}_2$ with the same eigenvalue. By the definition of Q, if we denote

$$\boldsymbol{F} = (\boldsymbol{f}_+, \boldsymbol{f}_-, \boldsymbol{f}_0, \boldsymbol{f}_{\mathcal{L},0})^{\mathrm{T}}\,,$$

then

$$\boldsymbol{e}_- = \boldsymbol{f}_-, \qquad \boldsymbol{e}_{\mathcal{L},0} = \boldsymbol{f}_{\mathcal{L},0}\,.$$

Let $\mathcal{QN}_2$ as the space spanned by the eigenvectors of $\widetilde{A}_2$ with negative eigenvalues. Then there exists a nontrivial $\widehat{\boldsymbol{F}} \in \mathcal{QN}_2$ such that

$$\widehat{\boldsymbol{F}} = (\widehat{\boldsymbol{f}}_+, 0, \widehat{\boldsymbol{f}}_0, 0)^{\mathrm{T}}\,.$$

Since $\mathcal{QN}_2$ is an invariant subspace of $\widetilde{A}_2$, we have that

$$\widetilde{A}_2\widehat{\boldsymbol{F}} = (\alpha D_+ \widehat{\boldsymbol{f}}_+^{\mathrm{T}}, 0, 0, \widehat{\boldsymbol{f}}_2) \in \mathcal{QN}_2$$

where $\widehat{\boldsymbol{f}}_2 = \alpha A_{21}^{\mathrm{T}} \widehat{\boldsymbol{f}}_+^{\mathrm{T}} + (I + \alpha B)^{1/2} (\alpha B)^{1/2})^{\mathrm{T}} \widehat{\boldsymbol{f}}_0^{\mathrm{T}}$. By the symmetry and non-degeneracy of $\widetilde{A}_2$, the quadratic form given by $\widetilde{A}_2$ on $\mathcal{QN}_2$ is strictly negative. Therefore,

$$\widehat{\boldsymbol{F}}^{\mathrm{T}} \widetilde{A}_2 \widehat{\boldsymbol{F}} = \alpha \widehat{\boldsymbol{f}}_+^{\mathrm{T}} D_+ \widehat{\boldsymbol{f}}_+ \leq 0\,.$$

Since D_+ is strictly positive definite, we have that $\widehat{\boldsymbol{f}}_+ = 0$ and

$$\widehat{\boldsymbol{F}}^{\mathrm{T}} \widetilde{A}_2 \widehat{\boldsymbol{F}} = 0\,,$$

which implies that $\widehat{\boldsymbol{F}} = 0$. This contradicts the assumption that $\widehat{\boldsymbol{F}}$ is non-trivial. Hence the matrix $\boldsymbol{E}$ is non-singular. □

Now we are ready to show the recovery of solutions to the undamped equation from (3.1). Let $g_{+,j}, g_{0,k}$ be the solutions to (3.1) with the incoming data $X_{+,j}, X_{0,k}$ with $1 \le j \le \nu_+$ and $1 \le k \le \nu_0$ respectively. Let C be the matrix defined in (3.17).

We show that C is non-singular. Suppose not. Then there exist constants

$$(\eta_{+,1}, \cdots, \eta_{+,\nu_+}, \eta_{0,1}, \cdots, \eta_{0,\nu_0}) \neq 0$$

such that we can find incoming data

$$\phi_g = \sum_{j=1}^{\nu_+} \eta_{+,j} X_{+,j} + \sum_{k=1}^{\nu_0} \eta_{0,k} X_{0,k}\,,$$

which gives rise to a solution g satisfying that

$$\begin{aligned} &\langle (v_1+u)X_{+,1},\ g\rangle_v = \cdots = \big\langle (v_1+u)X_{+,\nu_+},\ g\big\rangle_v = 0\,, \\ &\langle (v_1+u)X_{0,1},\ g\rangle_v = \cdots = \langle (v_1+u)X_{0,\nu_0},\ g\rangle_v = 0\,, \end{aligned} \qquad \text{at } x = 0. \tag{4.26}$$

By (4.25), this implies that we have

$$\begin{pmatrix} \boldsymbol{e}_{1,-} & & \boldsymbol{e}_{1,\mathcal{L},0} \\ & \cdots & \\ \boldsymbol{e}_{\nu_-+\nu_0,-} & & \boldsymbol{e}_{\nu_-+\nu_0,\mathcal{L},0} \end{pmatrix} \begin{pmatrix} \vec{U}_- \\ \\ \vec{U}_{\mathcal{L},0} \end{pmatrix} = 0 \qquad \text{at } x = 0,$$

where

$$\begin{aligned} \vec{U}_- &= \big(\langle (v_1+u)X_{-,1},\ g\rangle_v, \cdots, \big\langle (v_1+u)X_{-,\nu_-},\ g\big\rangle_v\big)^{\mathrm{T}}\,, \\ \vec{U}_{\mathcal{L},0} &= \big(\big\langle (v_1+u)\mathcal{L}^{-1}((v+u)X_{0,1}),\ g\big\rangle_v, \cdots, \big\langle (v_1+u)\mathcal{L}^{-1}((v+u)X_{0,\nu_0}),\ g\big\rangle_v\big)^{\mathrm{T}}\,. \end{aligned}$$

By Claim (**C**1), this simply implies that at $x = 0$,

$$\begin{aligned} &\langle (v_1+u)X_{-,1},\ g\rangle_v = \cdots = \big\langle (v_1+u)X_{-,\nu_-},\ g\big\rangle_v = 0\,, \\ &\big\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,1}),\ g\big\rangle_v = \cdots = \big\langle (v_1+u)\mathcal{L}^{-1}((v_1+u)X_{0,\nu_0}),\ g\big\rangle_v = 0\,. \end{aligned} \tag{4.27}$$

Hence (4.26) and (4.27) hold for any $x \ge 0$. Thus the solution g satisfies (1.1). Again by the uniqueness of (1.1), we have $\eta_{+,1} = \cdots = \eta_{+,\nu_+} = \eta_{0,1} = \cdots = \eta_{0,\nu_0} = 0$ which is a contraction. Thus C must be non-singular. Suppose f is the solution to (3.1) with the incoming data ϕ. There exists $(\widetilde{\eta}_{+,1}, \cdots, \widetilde{\eta}_{+,\nu_+}, \widetilde{\eta}_{0,1}, \cdots \widetilde{\eta}_{0,\nu_0})$ such that if we define

$$\widetilde{\Phi} = \sum_{j=1}^{\nu_+} \widetilde{\eta}_{+,j} X_{+,j} + \sum_{k=1}^{\nu_0} \widetilde{\eta}_{0,k} X_{0,k}\,, \tag{4.28}$$

then

$$\begin{pmatrix} \widetilde{\eta}_{+,1} \\ \vdots \\ \widetilde{\eta}_{+,\nu_+} \\ \widetilde{\eta}_{0,1} \\ \vdots \\ \widetilde{\eta}_{0,\nu_0} \end{pmatrix} = C^{-1} \begin{pmatrix} \langle (v_1+u)X_{+,1},\ f\rangle_v \\ \vdots \\ \big\langle (v_1+u)X_{+,\nu_+},\ f\big\rangle_v \\ \langle (v_1+u)X_{0,1},\ f\rangle_v \\ \vdots \\ \langle (v_1+u)X_{0,\nu_0},\ f\rangle_v \end{pmatrix}. \tag{4.29}$$

Let $\widetilde{g}$ be the solution to (3.1) with incoming data $\widetilde{\Phi}$. We claim that

$$\vec{U}_{\widetilde{g}} = \vec{U}_f, \qquad \text{at } x = 0\,, \tag{4.30}$$

where $\vec{U}_{\widetilde{g}}$ and $\vec{U}_f$ are defined in (4.10) with incoming data $\widetilde{\Phi}$ and ϕ respectively. Indeed, by (4.29), we have that

$$\vec{U}_{+,\widetilde{g}} = \vec{U}_{+,f}\,, \qquad \vec{U}_{0,\widetilde{g}} = \vec{U}_{0,f}\,, \qquad \text{at } x = 0\,.$$

Thus, if we denote $\vec{U}_{f-\widetilde{g}}$ as the moments defined in (4.10) with incoming data $\phi - \widetilde{\Phi}$, then $\vec{U}_{f-\widetilde{g}}$ has the form

$$\vec{U}_{f-\widetilde{g}} = \begin{pmatrix} 0 \\ U_{-,f-\widetilde{g}} \\ 0 \\ U_{\mathcal{L},0,f-\widetilde{g}} \end{pmatrix}, \qquad \text{at } x = 0\,. \tag{4.31}$$

Together with (4.25), we have

$$\boldsymbol{E} \begin{pmatrix} U_{-,f-\widetilde{g}} \\ U_{\mathcal{L},0,f-\widetilde{g}} \end{pmatrix} = 0\,.$$

Equation (4.30) then follows from the non-degeneracy of $\boldsymbol{E}$ as shown in (**C**1). Given (4.30), we derive that

$$\vec{U}_{\widetilde{g}} = \vec{U}_f, \qquad \text{for all } x \geq 0\,,$$

since $\vec{U}_{f-\widetilde{g}}$ satisfies (4.19). Hence, $f - \widetilde{g}$ satisfies the undamped equation (1.1). Note that $\widetilde{\phi}$ is a solution to (1.1). Therefore the unique solution to (1.1) with incoming data ϕ is given by

$$f_\phi = f - \widetilde{g} + \widetilde{\Phi}\,,$$

where f is the solution to the damped equation with incoming data ϕ. In this case, the end-state of f_ϕ is given by

$$f_{\phi,\infty} = \widetilde{\Phi}\,.$$

We thereby recover the undamped equation from the damped equation in a constructive way. □

4.4. **Quasi-Optimality.** We now prove Proposition 3.2, 3.5, and 3.9. The key observation in [ES12] for the neutron transport equation is that the inf-sup condition holds on the finite-dimensional space as long as $(v_1 + u)\partial_x f^+$ stays in this subspace. In this case, the proof for the well-posedness of the finite-dimension system will be almost a verbatim as for the original equation. Here we have a similar situation for the damped equation (3.1) with a slight modification.

Proof of Proposition 3.2 and 3.5. For the well-posedness part, we only explain the modification in choosing the test functions ψ_1 and ψ_2. For any $f \in \Gamma_N$, we choose

$$\psi_1 = f\,, \qquad \psi_2 = \mathcal{P}_N \left(\frac{1}{(1 + |v_1 + u| + |v_2| + |v_3|)^{\omega_0}} (v_1 + u)\partial_x f^+ \right),$$

where $\mathcal{P}_N : (L^2(\,\mathrm{d}v))^d \to \Gamma_N$ is the projection onto Γ_N. The rest of the estimates are similar to the proof in Section 5, thus omitted.

Now we show that the boundary conditions for the solution to (3.9) are given by (3.11). To this end, we first choose test functions $G_{2j}(x, v) = g(x)\psi_{2j}(v)$ where $g(x) \in C_c^\infty([0, \infty))$ and $1 \leq j \leq N$. Applying G_j in (3.9), we get

$$-\left\langle f_N^-,\ (v_1 + u)\psi_{2j}(v)\partial_x g(x) \right\rangle_{x,v} + \left\langle (\mathcal{L}\psi_{2j})g(x),\ f_N \right\rangle_{x,v} = 0\,, \tag{4.32}$$

where f_N is defined in (3.8) and $f_N = f_N^- + f_N^+$ with

$$f_N^- = \sum_{k=1}^{N+1} a_{2k-1}(x)\psi_{2k-1}\,, \qquad f_N^+ = \sum_{k=1}^{N} a_{2k}(x)\psi_{2k}\,.$$

By integration by parts in (4.32) we obtain

$$\left\langle \sum_{k=1}^{N+1} \psi_{2k-1}\partial_x a_{2k-1}(x),\ (v_1+u)\psi_{2j}(v)g(x) \right\rangle_{x,v} + \langle (\mathcal{L}\psi_{2j})g(x),\ f_N\rangle_{x,v} = 0\,.$$

Since $g \in C_c^\infty([0,\infty))$ is arbitrary, we have

$$\sum_{k=1}^{N+1} \langle \psi_{2k-1},\ (v_1+u)\psi_{2j}(v)\rangle_v\, \partial_x a_{2k-1}(x) + \langle (\mathcal{L}\psi_{2j}),\ f_N\rangle_v = 0\,, \tag{4.33}$$

for each $1 \le j \le N$ and $x \in [0,\infty)$. Note we choose $\widetilde{G}_{2j} = \widetilde{g}(x)\psi_{2j}(x)$ where $\widetilde{g} \in C^\infty([0,\infty))$. Then equation (3.9) becomes

$$\begin{aligned} -\left\langle f_N^-,\ (v+u)\psi_{2j}(v)\partial_x g(x)\right\rangle_{x,v} + \langle (\mathcal{L}\psi_{2j})g(x),\ f_N\rangle_{x,v} + \left\langle (v_1+u)f_N^+,\ \psi_{2j}\widetilde{g}(0)\right\rangle_{x=0} \\ = 2\int_{v_1+u>0} (v_1+u)\phi(v)\psi_{2j}(v)g(0)\,\mathrm{d}v\,, \end{aligned} \tag{4.34}$$

for each $1 \le j \le N$. The set of N boundary conditions (3.11) then follows from integrating by parts in (4.34) and applying (4.33). □

The error estimate of the undamped equation follows from Proposition 3.5.

Proof of Proposition 3.9. First we note that there exist constants $\kappa_4, \widetilde{\kappa}_4 > 0$ such that

$$\begin{aligned} \|\widetilde{\Phi} - \widetilde{\Phi}_N\|_\Gamma &= \|\widetilde{\Phi} - \widetilde{\Phi}_N\|_{(L^2(\,\mathrm{d}v))^d} \\ &\le \kappa_4 \|f - f_N\|_\Gamma + \widetilde{\kappa}_4 \|f\|_{(L^2(\,\mathrm{d}v\,\mathrm{d}x))^d} \left(\sum_{i=1}^{\nu_+} \|g_{+,i} - g_{+,i,N}\|_\Gamma + \sum_{j=1}^{\nu_0} \|g_{0,j} - g_{0,j,N}\|_\Gamma \right)\,, \end{aligned}$$

where $\widetilde{\Phi}$ is defined in (4.18) or (4.28), $\widetilde{\Phi}_N$ is given by the same formula for $\widetilde{\Phi}$ with f changed to f_N, and $g_{+,i,N}, g_{0,j,N}$ are the approximations in Γ_N to $g_{+,i}, g_{0,j}$. Second, since $f - \widetilde{g}$ is a solution to the damped equation, we have

$$\|(f-\widetilde{g}) - (f_N - \widetilde{g}_N)\|_\Gamma \le \kappa_5 \inf_{w\in\Gamma_N} \|w - (f-\widetilde{g})\|_\Gamma\,,$$

where $f_N, \widetilde{g}_N \in \Gamma_N$. Therefore,

$$\begin{aligned} \|(f-\widetilde{g}+\widetilde{\Phi}) - (f_N - \widetilde{g}_N + \widetilde{\Phi}_N)\|_\Gamma &\le \kappa_5 \inf_{w\in\Gamma_N} \|w - (f-\widetilde{g})\|_\Gamma + \|\widetilde{\Phi} - \widetilde{\Phi}_N\|_\Gamma \\ &\le \kappa_5 \inf_{w\in\Gamma_N} \|w - (f - \widetilde{g} + \widetilde{\Phi})\|_\Gamma + \kappa_4 \|f - f_N\|_\Gamma \\ &\le \kappa_6 \left(\inf_{w\in\Gamma_N} \|f_\phi - w\|_\Gamma + \inf_{w\in\Gamma_N} \|f - w\|_\Gamma + \delta_N \|f\|_{(L^2(\,\mathrm{d}v\,\mathrm{d}x))^d} \right)\,, \end{aligned}$$

where

$$\delta_N = \sum_{i=1}^{\nu_+} \inf_{w\in\Gamma_N} \|g_{+,i} - w\|_\Gamma + \sum_{j=1}^{\nu_0} \inf_{w\in\Gamma_N} \|g_{0,j} - w\|_\Gamma\,.$$

Note that the second inequality holds because $\widetilde{\Phi} \in H^+ \oplus H^0 \subseteq \Gamma_N$. □

5. Numerical Examples

As explained in Section 3, the overall strategy to solve the half-space equation consists of two steps: First, we solve for numerical solution to the half-space damped equation (3.1) using the Galerkin approximation; Second, we recover the undamped solution by Proposition 3.8, which involves the solution of the damped equation with various boundary conditions to obtain the matrix C in the linear system (3.18).

We will consider the linearized BGK equation and the linear transport equation below. We will restrict ourselves here to one dimensional examples. As in Proposition 3.4, for the Galerkin approximation, we specify a set of even and odd functions to form the approximation space Γ_N, whose choice depends on the particular equation under study. Using Corollary 3.7, the solution of the system under Galerkin approximation (3.10)–(3.11) is reduced to solving the generalized eigenvalue problem (3.14), for which we need to assemble the matrices A and B using Gaussian quadrature. This will be discussed in more details below.

Our algorithm is implemented in MATLAB. The Gaussian quadrature abscissas and weights are obtained using symbolic calculations to guarantee precision.

5.1. **Linearized BGK equation.** We first consider the case of one dimension linearized BGK equation. For this equation, our construction of basis functions is based on the half-space Hermite polynomials. Those are orthogonal polynomials defined on the positive half v-axis with weight functions given by $\exp(-v^2)$: $\{B_n(v), v > 0\}$ such that $B_n(v)$ is a polynomial of order n and satisfies

$$\int_0^\infty B_m(v)B_n(v)e^{-v^2}\,\mathrm{d}v = \delta_{nm}. \tag{5.1}$$

The orthogonal polynomials can be constructed using three term recursion formula (see for example [Shi81]), and the details are recalled in Appendix A for completeness.

The basis functions ψ we need are either odd or even with respect to $v = -u$, hence we shift the functions B_n's by $-u$ and make even / odd extensions:

$$B_n^E(v) = \begin{cases} B_n(v+u)/\sqrt{2}, & v > -u; \\ B_n(-v-u)/\sqrt{2}, & v < -u. \end{cases} \tag{5.2}$$

$$B_n^O(v) = \begin{cases} B_n(v+u)/\sqrt{2}, & v > -u; \\ -B_n(-v-u)/\sqrt{2}, & v < -u. \end{cases} \tag{5.3}$$

Finally, the basis functions ψ_k's are obtained by multiplying these functions by the square root of the Maxwellian: for $n \geq 1$

$$\begin{aligned} \psi_{2n-1} &= B_{n-1}^O e^{-(v+u)^2/2}, \\ \psi_{2n} &= B_{n-1}^E e^{-(v+u)^2/2}. \end{aligned} \tag{5.4}$$

By definition, ψ_{2n-1} is odd, ψ_{2n} is even, and they form a orthonormal basis of $L^2(\mathrm{d}v)$. For a fixed n, $(v+u)\psi_{2n}(v)$ is a odd function with respect to $v = -u$. For $v > -u$,

$$(v+u)\psi_{2n}(v) = (v+u)B_{n-1}(v+u)e^{-(v+u)^2/2}/\sqrt{2}.$$

Since $(v+u)B_n(v+u)$ is a n-th order polynomial in $v+u$, there exists an expansion

$$(v+u)B_{n-1}(v+u) = \sum_{i=0}^{n} \alpha_i B_i(v+u). \tag{5.5}$$

This yields that

$$(v+u)\psi_{2n}(v) = \sum_{i=0}^{n} \alpha_i \psi_{2i+1} \in \mathrm{span}\{\psi_1, \cdots, \psi_{2n+1}\}. \tag{5.6}$$

Therefore, $\Gamma_N = \mathrm{span}\{\psi_1, \cdots, \psi_{2N+1}\}$ satisfies the condition of Proposition 3.4 and the variational formulation (3.9)–(3.11) is well-posed. The $(2N+1)\times(2N+1)$ matrices A and B are then given by

$$\mathsf{A}_{ij} = \int_{\mathbb{R}} (v+u)\psi_i\psi_j \,\mathrm{d}v \quad \text{and} \quad \mathsf{B}_{ij} = -\int_{\mathbb{R}} \psi_i \mathcal{L}_d \psi_j \,\mathrm{d}v.$$

Note that both matrices are symmetric. The matrix A can be obtained by using the recurrence relation of the orthogonal polynomials. For the matrix B, recall that

$$\begin{aligned}
\mathcal{L}\psi_i &= \psi_i - m_i = \psi_i - \chi_0 \int_{\mathbb{R}} \psi_i \chi_0 \,\mathrm{d}v - \chi_+ \int_{\mathbb{R}} \psi_i \chi_+ \,\mathrm{d}v - \chi_- \int_{\mathbb{R}} \psi_i \chi_- \,\mathrm{d}v. \\
\mathcal{L}_d\psi_i &= \mathcal{L}\psi_i + \alpha \sum_{k=1}^{\nu_+} (v+u)X_{+,k} \int_{\mathbb{R}} (v+u)X_{+,k}\psi_i \,\mathrm{d}v \\
&\quad + \alpha \sum_{k=1}^{\nu_-} (v+u)X_{-,k} \int_{\mathbb{R}} (v+u)X_{-,k}\psi_i \,\mathrm{d}v + \alpha \sum_{k=1}^{\nu_0} (v+u)X_{0,k} \int_{\mathbb{R}} (v+u)X_0\psi_i \,\mathrm{d}v \\
&\quad + \alpha \sum_{k=1}^{\nu_0} (v+u)\mathcal{L}^{-1}((v+u)X_{0,k}) \int_{\mathbb{R}} (v+u)\mathcal{L}^{-1}((v+u)X_{0,k})\psi_i \,\mathrm{d}v.
\end{aligned}$$

All the integrals involved in calculating B can be easily made exact up to machine precision by using Gaussian quadrature. For simplicity, let us just focus on

$$\int \psi_{2j}\chi_0 \,\mathrm{d}v$$

and note that the other integrals share the same structure, that is, the integrand is a product of two polynomials and two Gaussians $e^{-v^2/2}$ and $e^{-(v+u)^2/2}$. To evaluate this type of integral using Gaussian quadrature, we first split the integral into two parts:

$$\int_{\mathbb{R}} \psi_{2j}\chi_0 \,\mathrm{d}v = \int_{-u}^{\infty} \psi_{2j}\chi_0 \,\mathrm{d}v + \int_{-\infty}^{-u} \psi_{2j}\chi_0 \,\mathrm{d}v.$$

Note that ψ_{2j}, on either side of $-u$, is a $(j-1)$-th order polynomial multiplied by $\exp(-(v+u)^2/2)$, while χ_0 is a quadratic function multiplied with a different weight function $\exp(-v^2/2)$. The product of two Gaussians centered at different locations could be combined into a single Gaussian:

$$\begin{aligned}
\int_{-u}^{\infty} \psi_{2j}\chi_0 \,\mathrm{d}v &= \frac{\sqrt{2}}{2} \int_{-u}^{\infty} B_{j-1}(v+u) \frac{\chi_0(v)}{e^{-v^2/2}} e^{-\frac{(v+u)^2+v^2}{2}} \,\mathrm{d}v \\
&= \frac{\sqrt{2}}{2} e^{-u^2/4} \int_0^{\infty} B_{j-1}(v) \frac{\chi_0(v-u)}{e^{-(v-u)^2/2}} e^{-(v-u/2)^2} \,\mathrm{d}v.
\end{aligned} \tag{5.7}$$

Similarly, we have the integration for $v < -u$:

$$\begin{aligned}
\int_{-\infty}^{-u} \psi_{2j}\chi_0 \,\mathrm{d}v &= \frac{\sqrt{2}}{2} \int_{-\infty}^{-u} B_{j-1}(-v-u) \frac{\chi_0(v)}{e^{-v^2/2}} e^{-\frac{(v+u)^2+v^2}{2}} \,\mathrm{d}v \\
&= \frac{\sqrt{2}}{2} e^{-u^2/4} \int_0^{\infty} B_{j-1}(v) \frac{\chi_0(-v-u)}{e^{-(v+u)^2/2}} e^{-(v+u/2)^2} \,\mathrm{d}v.
\end{aligned} \tag{5.8}$$

The integrals (5.7) and (5.8) can be evaluated up to machine precision by Gaussian quadrature based on weight $e^{-(v-u/2)^2}$ and $e^{-(v+u/2)^2}$ respectively, as $B_{j-1}\chi_0 e^{v^2/2}$ is a polynomial with its degree up to $N+3$. The boundary condition (3.11) requires the numerical evaluation of the integral

$$\int_{v+u>0} (v+u)\phi\psi_{2j} \,\mathrm{d}v.$$

We calculate this using Gaussian quadrature with the weight $e^{-(v+u)^2}$. The error of the quadrature depends on the number of quadrature points and the regularity of the incoming data ϕ.

We now present some numerical results for the linearized BGK equation. In the first set of examples, we compare our numerical results with analytical solutions, when the specified boundary data ϕ is given by the restriction of some $f \in H^0 \oplus H^+$ on $v > -u$. In this case, the solution to the undamped equation (1.1) is simply f on the whole velocity space. As discussed in (2.9), the dimension of the space $H^0 \oplus H^+$ depends on the bulk velocity u and the sound speed, which is $c = \sqrt{3/2}$ in our case as $T = 1/2$. We will choose $\chi_{+/-/0}$ defined in (2.7) as the incoming data. By the uniqueness of the half-space equation, the solution will simply be $\chi_{+/0}$ when the incoming data is chosen as $\chi_{+/0}$. We take six choices of u corresponding to the six cases listed in (2.9) (the case $u < -c$ gives an empty $H^0 \oplus H^+$ hence not included). The results are shown in Figures 1–6 below. In all these figures, the blue squared line is the incoming data, given by χ_-, χ_0 and χ_+ respectively. The green triangle line is the solution at $x = \infty$, and the red dotted line is the solution at $x = 0$.

Several remarks are in order: First, when the χ modes lie in $H^0 \oplus H^+$ for the given bulk background velocity u, we observe in Figure 1-6 that the solution at $x = 0$ gives a perfect match. We thus recover the exact solution from the numerical scheme. Second, we note that in general, the solution exhibits a jump at $v = -u$, as clearly seen for example in Figure 1(left). This justifies our choice of the even-odd formulation and basis functions from the half-space Hermite polynomials. Finally, we remark that we have used a filtering (with 2nd order cosine filter) to reduce the Gibbs oscillations caused by the large derivatives in some cases (for instance Figure 2(left)).

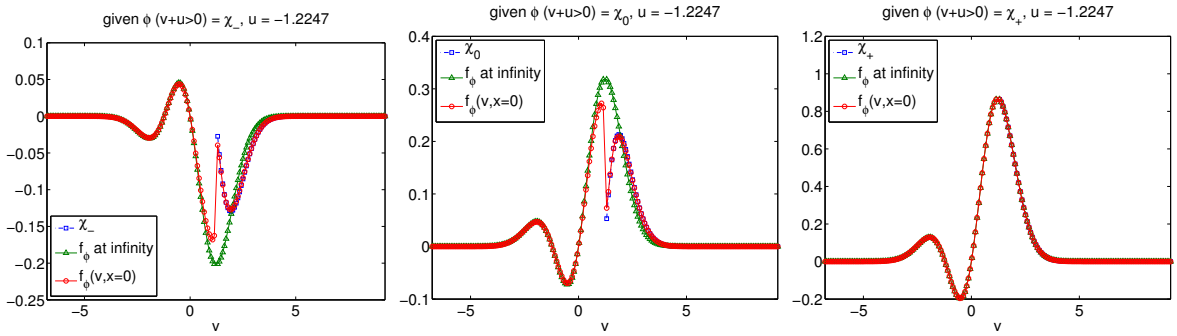

FIGURE 1. $u = -\sqrt{1.5} = -c$. In this case $\chi_+ \in H^0$, and χ_- and χ_0 are in H^-.

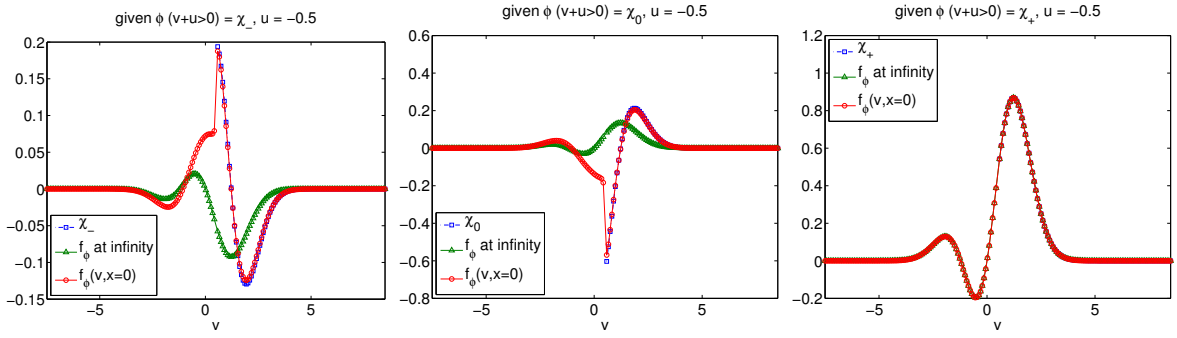

FIGURE 2. $-c < u = -0.5 < 0$. In this case $\chi_+ \in H^+$, and χ_- and χ_0 are in H^-.

Next, we consider an example where the exact solution is not known. We solve the equation (1.1) for $u = 0$ with boundary data $\phi = v^3, v > 0$. The numerical solution is shown in Figure 7.

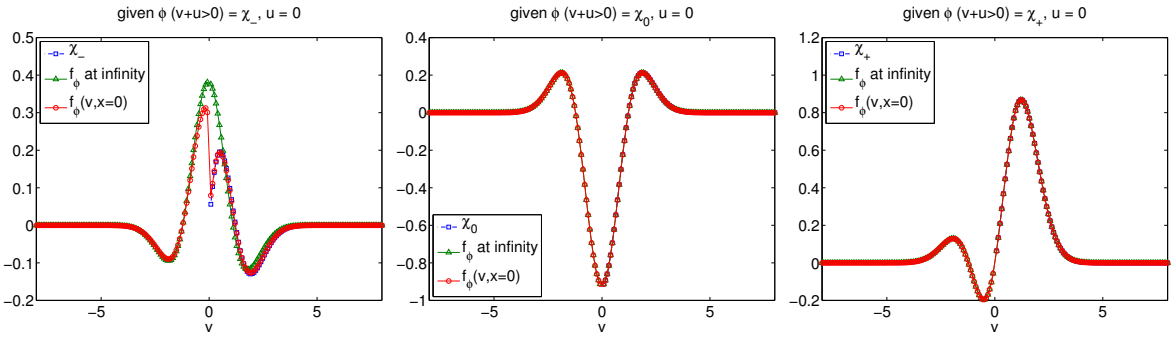

FIGURE 3. $u = 0$. In this case $\chi_+ \in H^+$, $\chi_0 \in H^0$ and $\chi_- \in H^-$.

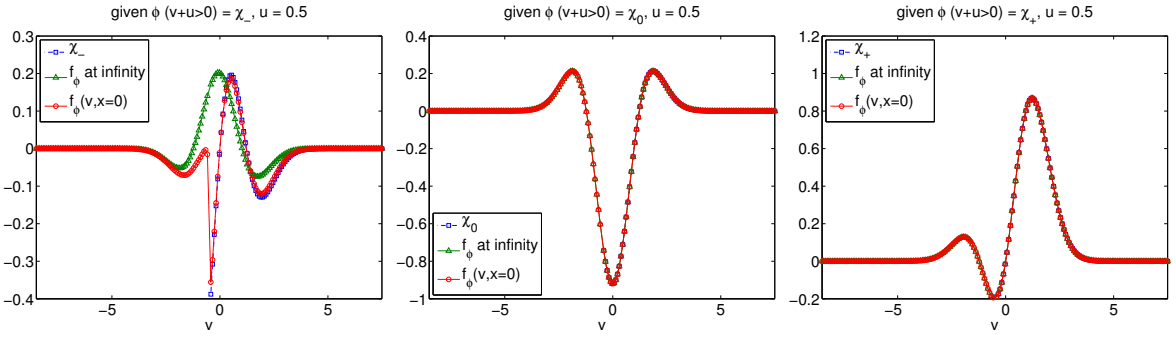

FIGURE 4. $0 < u = 0.5 < c$. In this case χ_+ and χ_0 are in H^+, and $\chi_- \in H^-$.

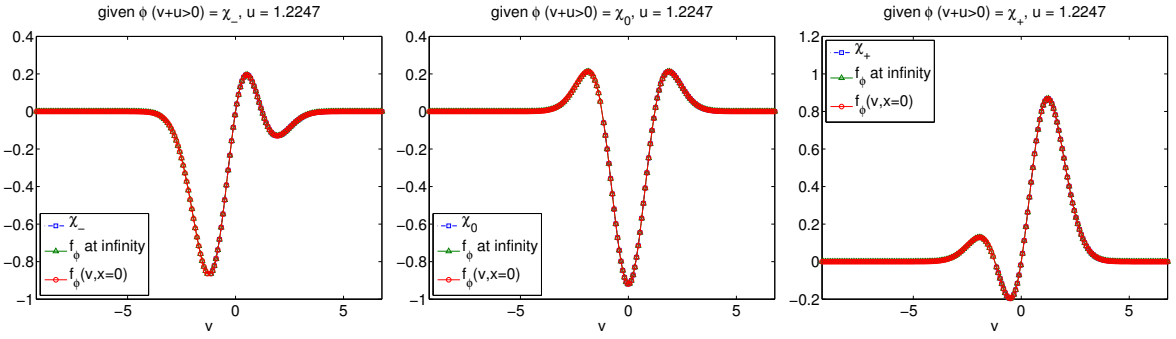

FIGURE 5. $u = \sqrt{1.5} = c$. In this case χ_+ and χ_0 are in H^+, and $\chi_- \in H^0$.

5.2. **Isotropic neutron transport equation.** We further consider the isotropic neutron transport equation. The construction of the basis functions is similar to the linearized BGK case. However, instead of using half-space Hermite polynomials, we start with Legendre polynomials on the interval $[0, 1]$ and carry out even / odd extensions. The Legendre polynomials, which are orthogonal polynomials for constant weight function, are used as the equilibrium states for the neutron transport equation are simply constants. We then apply Gauss-Legendre quadrature to assemble A and B for the generalized eigenvalue problem. We will skip further details here, as the construction is relatively straightforward compared with the linearized BGK case.

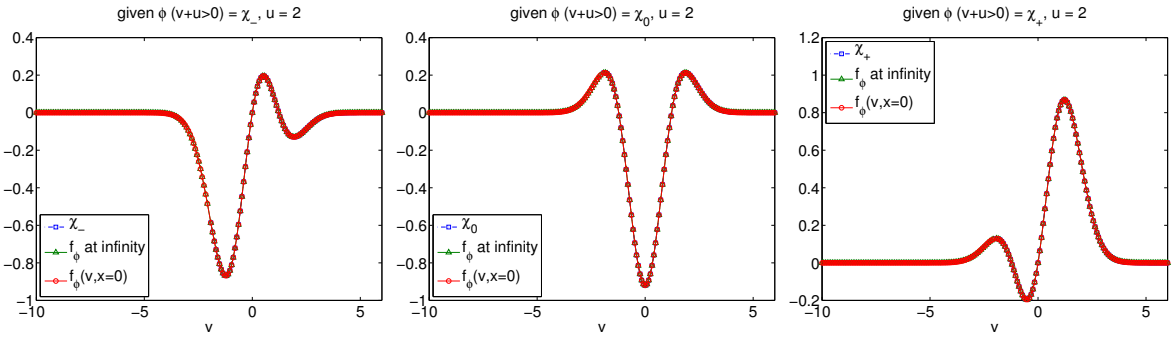

FIGURE 6. $u = 2 > c$. In this case all χ are in H^+.

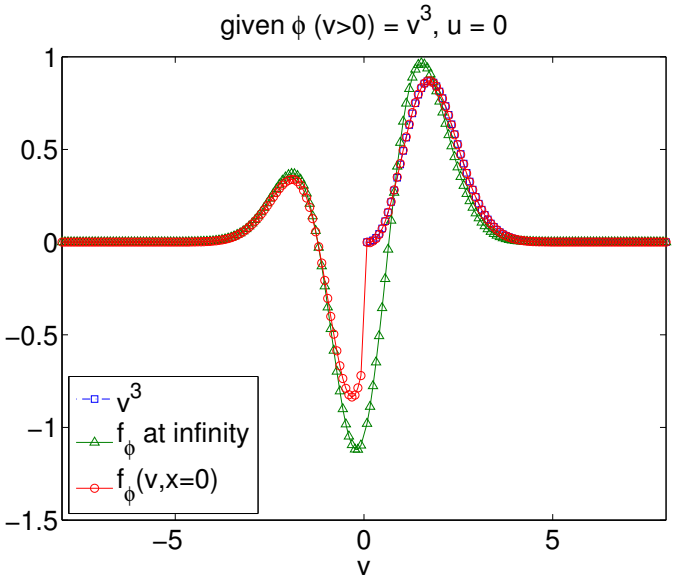

FIGURE 7. Blue boxed line is the input data $\phi = v^3(v > 0)$. Green triangle line is the solution at infinity and the red circled line is the solution at the boundary. N = 36 here.

To validate our methods in this case, we compare numerical solution with analytical solution with boundary data given by $\phi = v$ for $v \in [0, 1]$. The analytical solution is known in case as

$$f_\phi(-v) = \frac{1}{\sqrt{3}} H(v) - v, \qquad v > 0\,, \tag{5.9}$$

where H is the Chandrasekhar H-function. In Figure 8 we plot both analytical and numerical solutions. The plot shows good agreement of the numerical solution with the exact one. Furthermore, the limit at $x = \infty$ of the solution to the half-space isotropic NTE is a constant, whose amplitude is known as the extrapolation length. In Table 1 we compare our numerical approximation of extrapolation length with the exact result, which is again in good agreement. In comparison, we note that the approximate value for the extrapolation length obtained in [Cor90] is 0.71040377 with 70 modes, while we achieve better results with piecewise polynomial of only 12-th order.

TABLE 1. Numerical approximations of the extrapolation length.

4	0.709324539775964	24	0.710445373807707	44	0.710446026371328	64	0.710446075479882
8	0.710386430787361	28	0.710445703544666	48	0.710446044962143	68	0.710446078520678
12	0.710434523809144	32	0.710445863417934	52	0.710446057194912	72	0.710446080785171
16	0.710442451548528	36	0.710445948444682	56	0.710446065509628	76	0.710446082499459
20	0.710444603305304	40	0.710445997010591	60	0.710446071320336	exact	0.710446089598763

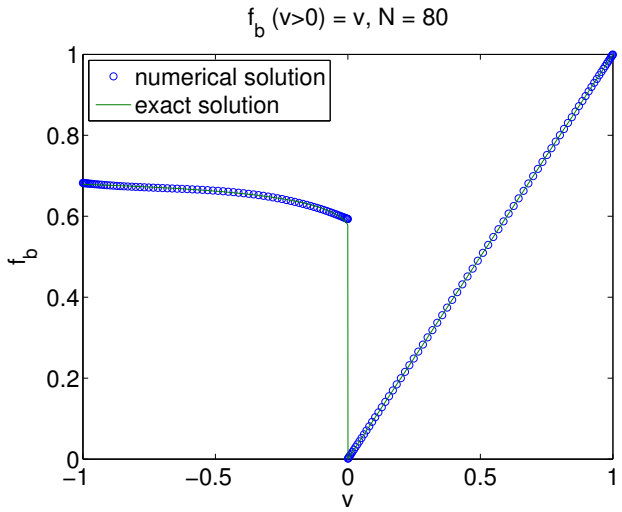

FIGURE 8. Analytical solution and numerical solution to the isotropic neutron transport equation at $x = 0$.

APPENDIX A. HALF-HERMITE POLYNOMIAL

Here we derive the half-space orthogonal polynomial with weight $\exp(-(v-u)^2)$ with u a real number. The zeroth order half space Hermite polynomial is:

$$B_0 = \frac{1}{\sqrt{m_0}} \quad \text{with} \quad m_0 = \frac{\sqrt{\pi}}{2}\left(1 + \operatorname{erf}(u)\right). \tag{A.1}$$

The higher order polynomials are defined through recurrence relation:

$$\sqrt{\beta_{n+1}} B_{n+1} = (v - \alpha_n) B_n - \sqrt{\beta_n} B_{n-1}, \tag{A.2}$$

where α and β are defined by

$$\begin{cases} \beta_{n+1} = n + \dfrac{1}{2} + u\alpha_n - \alpha_n^2 - \beta_n; \\ \alpha_{n+1} = u - \alpha_n + \dfrac{1}{2\beta_{n+1}} \displaystyle\sum_{k=0}^{n} \alpha_k \end{cases} \tag{A.3}$$

with $\alpha_0 = m_1/m_0$ and $\beta_1 = \sqrt{m_0 m_2 - m_1^2}/m_0$, where m_i, $i = 0, 1, 2$ are moments of the Gaussian:

$$m_i = \int_0^\infty v^i e^{-(v-u)^2} \, \mathrm{d}v, \qquad i = 0, 1, 2. \tag{A.4}$$

The deduction formula are derived from the Christoffel-Darboux identity

$$\sum_{k=0}^{n} B_k^2 = \sqrt{\beta_{n+1}} \left(B'_{n+1} B_n - B_{n+1} B'_n \right) \tag{A.5}$$

as follows. By orthogonality of $\{B_n\}$, we get

$$\alpha_n = \int_0^\infty v B_n^2 e^{-(v-u)^2}\,\mathrm{d}v, \quad \text{and} \quad \sqrt{\beta_{n+1}} = \int_0^\infty v B_n B_{n+1} e^{-(v-u)^2}\,\mathrm{d}v.$$

Integrate the identity (A.5) over v with the weight, we get

$$\begin{aligned} n+1 &= \sqrt{\beta_{n+1}} \int_0^\infty B'_{n+1} B_n e^{-(v-u)^2}\,\mathrm{d}v = \int_0^\infty v B_{n+1} B'_{n+1} e^{-(v-u)^2}\,\mathrm{d}v \\ &= -\frac{1}{2} + \int_0^\infty v^2 B_{n+1}^2 e^{-(v-u)^2}\,\mathrm{d}v - u\alpha_n, \end{aligned}$$

where the second equality is obtained by taking the inner product with B'_{n+1} of recursion equation (A.2), and the third comes from integration by parts. From this we get the first deduction relation in (A.3). Next multiply (A.5) with v and then integrate, we obtain

$$\begin{aligned} \sum_{k=0}^{n} \alpha_k &= \sqrt{\beta_{n+1}} \int_0^\infty v B'_{n+1} B_n e^{-(v-u)^2}\,\mathrm{d}v \\ &= \sqrt{\beta_{n+1}} \left(2\int_0^\infty v^2 B_{n+1} B_n e^{-(v-u)^2}\,\mathrm{d}v - 2u \int_0^\infty v B_{n+1} B_n e^{-(v-u)^2}\,\mathrm{d}v \right) \\ &= 2\beta_{n+1} \left(\alpha_n + \alpha_{n+1} - u \right), \end{aligned}$$

where the first equality comes from the fact that $\int_0^\infty v B_{n+1} B'_n e^{-(v-u)^2}\,\mathrm{d}v = 0$, the second is due to integration by parts, and the third comes from integrating the recursion equation (A.2) multiplied by vB_{n+1}. This gives the other deduction relation in (A.3).

References

[BA72] I. Babuška and A.K. Aziz, *Survey lectures on the mathematical foundations of the finite element method*, The Mathematical foundation of the Finite Element method with Applications to Partial Differential Equations (A.K. Aziz (ed.)), 1972, pp. 1–359.

[BBG+11] C. Besse, S. Borghol, T. Goudon, I. Lacroix-Violet, and J.-P. Dudon, *Hydrodynamic regimes, Knudsen layer, numerical schemes: definition of boundary fluxes*, Adv. Appl. Math. Mech. **3** (2011), no. 5, 519–561.

[BLP79] A. Bensoussan, J.L. Lions, and G.C. Papanicolaou, *Boundary-layers and homogenization of transport processes*, J. Publ. RIMS Kyoto Univ. **15** (1979), 53–157.

[BSS84] C. Bardos, R. Santos, and R Sentis, *Diffusion approximation and computation of the critical size*, Trans. Amer. Math. Soc. **284** (1984), no. 2, 617–649.

[BY12] C. Bardos and X. Yang, *The classification of well-posed kinetic boundary layer for hard sphere gas mixtures*, Comm. Partial Differential Equations **37** (2012), no. 7, 1286–1314.

[CGS88] F. Coron, F. Golse, and C. Sulem, *A classification of well-posed kinetic layer problems*, Comm. Pure Appl. Math. **41** (1988), 409–435.

[CLY04] C.-C. Chen, T.-P. Liu, and T. Yang, *Existence of boundary layer solutions to the Boltzmann equation*, Anal. Appl. **2** (2004), no. 4, 337–363.

[Cor90] F. Coron, *Computation of the asymptotic states for linear half space kinetic problems*, Transport Theory Statist. Phys. **19** (1990), no. 2, 89–114.

[Del03] S. Dellacherie, *Coupling of the Wang Chang-Uhlenbeck equations with the multispecies Euler system*, J. Comput. Phys. **189** (2003).

[DJM05] P. Degond, S. Jin, and L. Mieussens, *A smooth transition model between kinetic and hydrodynamic equations*, J. Comp. Phys **209** (2005), 665–694.

[DMG87] P. Degond and S. Mas-Gallic, *Existence of solutions and diffusion approximation for a model Fokker-Planck equation*, Transport Theory Statist. Phys. **16** (1987), no. 4–6, 589–636.

[ES12] H. Egger and M. Schlottbom, *A mixed variational framework for the radiative transfer equation*, Math. Models Methods Appl. Sci. **22** (2012), 1150014.

[GK95] F. Golse and A. Klar, *A numerical method for computing asymptotic states and outgoing distributions for kinetic linear half-space problems*, J. Stat. Phys. **80** (1995), no. 5–6, 1033–1061.

[Gol08] F. Golse, *Analysis of the boundary layer equation in the kinetic theory of gases*, Bull. Inst. Math. Acad. Sin. (N.S.) **3** (2008), no. 1, 211–242.

[JPT01] S. Jin, L. Pareschi, and G. Toscani, *Uniformly accurate diffusive relaxation schemes for multiscale transport equations*, SIAM J. Numer. Anal. **38** (2001), 913–936.

[LM12] M. Lemou and F. Méhats, *Micro-macro schemes for kinetic equations including boundary layers*, SIAM J. Sci. Comput. **34** (2012), no. 6, B734–B760.

[Mar47] R. E. Marshak, *Note on the spherical harmonic method as applied to the Milne problem for a sphere*, Phys. Rev. **71** (1947), 443–446.

[Shi81] B. Shizgal, *A Gaussian quadrature procedure for use in the solution of the Boltzmann equation and related problems*, J. Comput. Phys. **41** (1981), 309–328.

[UYY03] S. Ukai, T. Yang, and S.-H. Yu, *Nonlinear boundary layers of the Boltzmann equation. I. Existence*, Comm. Math. Phys. **236** (2003), no. 3, 373–393.

[WYY06] W. Wang, T. Yang, and X. Yang, *Nonlinear stability of boundary layers of the Boltzmann equation for cutoff hard potentials*, J. Math. Phys **47** (2006), no. 8, 083301.

[WYY07] W. Wang, T. Yang, and X. Yang, *Existence of boundary layers to the Boltzmann equation with cutoff soft potentials*, J. Math. Phys **48** (2007), no. 7, 073304.

Computing and Mathematical Sciences, California Institute of Technology, 1200 E California Blvd. MC 305-16, Pasadena, CA 91125 USA.

E-mail address: qinli@caltech.edu

Department of Mathematics, Physics, and Chemistry, Duke University, Box 90320, Durham, NC 27708 USA.

E-mail address: jianfeng@math.duke.edu

Department of Mathematics, Simon Fraser University, 8888 University Dr., Burnaby, BC V5A 1S6, Canada

E-mail address: weirans@sfu.ca