Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Convex Duality and Entropy-Based Moment Closures: Characterizing Degenerate Densities

Cory D. Hauck, C. David Levermore, and André L. Tits

Abstract— **A common method for constructing a function from a finite set of moments is to solve a constrained minimization problem. The idea is to find, among all functions with the given moments, that function which minimizes a physically motivated, strictly convex functional. In the kinetic theory of gases, this functional is the kinetic entropy; the given moments are macroscopic densities; and the solution to the constrained minimization problem is used to formally derive a closed system of partial differential equations which describe how the macroscopic densities evolve in time. Moment equations are useful because they simplify the kinetic, phase-space description of a gas, and with entropy-based closures, they retain many of the fundamental properties of kinetic transport.**

Unfortunately, in many situations, macroscopic densities can take on values for which the constrained minimization problem does not have a solution. In this paper, we give a geometric description of these so-called *degenerate densities* in a very general setting. Our key tool is the complementary slackness condition that is derived from a dual formulation of a minimization problem with relaxed constraints. We show that the set of degenerate densities is a union of convex cones defined by the complementary slackness conditions and, under reasonable assumptions, that this set is small in both a topological and a measure-theoretic sense. This result is important for further assessment and implementation of entropy-based moment closures. An expanded version of this work can be found in [Hauck et al., SIAM J. Contr. Optim., Vol. 47, 2008, pp. 1977-2015].

I. Introduction

In gas dynamics, the kinetic description of a gas is often simplified by using moment equations. In this reduced setting, a gas is characterized by a finite-dimensional vector $\boldsymbol{\rho}$ of densities that are moments of the kinetic distribution function F with respect to polynomials of the microscopic velocity. Evolution equations for $\boldsymbol{\rho}$ are derived by taking moments of the Boltzmann equation which governs the evolution of F. The derivation requires that an approximation for F be reconstructed from the densities $\boldsymbol{\rho}$, giving what is called a *closure*.

One well-known method for prescribing a closure is to find a function that minimizes the kinetic entropy subject to the constraint that its moments agree with $\boldsymbol{\rho}$. Such closures are called *entropy-based closures*. In recent years, they have generated substantial interest due to important structural properties which they inherit from the Boltzmann equation. These properties were first brought to light in [2].

In cases where the moments are continuous with respect to the relevant topology, there is always an entropy minimizer [3], [4]. Unfortunately, in classical gas dynamics, this is not usually the case. As a result, there are often physically relevant densities for which the constrained entropy minimization problem does not have a solution. In such cases, entropy-based closures are not well-defined, and these densities are called *degenerate*. In this paper, we provide a geometric description for the set of degenerate densities in a general setting.

This is a condensed version of work that appears in [1]. The full version contains additional details, examples, and complete proofs.

A. Moment systems and entropy-based closures

Consider a gas that is enclosed in a container, represented mathematically by the set $\Omega \subset \mathbb{R}^d$ (typically $d = 3$). The kinetic distribution function $F = F(v, x, t)$ which describes the kinetic state of the gas is a nonnegative function that is defined for positions $x \in \Omega$, velocities $v \in \mathbb{R}^d$, and times $t \geq 0$ so that, for any measurable set $\Lambda \subset \Omega \times \mathbb{R}^d$,

$$\int_\Lambda F(v, x, t)\, dvdx \tag{1}$$

gives the number of particles at time t with positions x and velocities v such that $(v, x) \in \Lambda$. The evolution of F is governed by the Boltzmann transport equation

$$\partial_t F + v \cdot \nabla_x F = \mathcal{C}(F)\,, \tag{2}$$

where $\mathcal{C}$ is an integral operator that describes the collisions between particles which drive the system to local thermal equilibrium.

Solutions of (2) formally satisfy the local balance law [5]

$$\partial_t \mathcal{H}(F) + \nabla_x \cdot \mathcal{J}(F) = \mathcal{S}(F)\,, \tag{3}$$

where the functionals

$$\mathcal{H}(g) \equiv \langle g \log(g) - g \rangle \quad \text{and} \quad \mathcal{J}(g) \equiv \langle v(g \log(g) - g) \rangle \tag{4}$$

An expanded version of this paper, with complete proofs and additional details can be found in [1]. The work of the first author was funded in part by the Department of Mathematics at the University of Maryland, College Park under National Science Foundation VIGRE grant DMS-0240049 and in part by the U.S. Department of Energy at Los Alamos National Laboratory under contract DE-AC52-06NA25396 and the DOE Office of Science Advanced Computing Research (ASCR) Program. The work of the third author was supported in part by the National Science Foundation under Grant DMI-0422931 and by the U.S. Department of Energy under Grant DEFG0204ER25655

C. D. Hauck is with the Computational Physics Group and Center for Nonlinear Studies at the Los Alamos National Laboratory, Los Alamos, NM, 87545, USA cdhauck@lanl.gov

C. D. Levermore is with the Department of Mathematics and Institute for Physical Sciences and Technology, University of Maryland, College Park, MD 20742 lvrmr@math.umd.edu

A. L. Tits is with the Department of Electrical and Computer Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742 andre@umd.edu

are the *kinetic entropy* and *kinetic entropy flux*, respectively, and

$$\mathcal{S}(g) \equiv \langle \log(g)\mathcal{C}(g)\rangle \tag{5}$$

is the *kinetic entropy dissipation*. Here and throughout this paper, $\langle\cdot\rangle$ denotes Lebesgue integration over all $v \in \mathbb{R}^d$, and we assume that all such integrals are well-defined. According to Boltzmann's "H-theorem" [5], $\mathcal{S}(g) \leq 0$, with equality if and only if $\mathcal{C}(g) = 0$. In this case, g is said to be in a state of local thermal equilibrium, and it takes the form of a Maxwellian distribution

$$\mathcal{M}_{\rho,u,\theta}(v) \equiv \frac{\rho}{(2\pi\theta)^{d/2}} \exp\left(-\frac{|v-u|^2}{2\theta}\right), \tag{6}$$

where ρ and θ are positive scalars and $u \in \mathbb{R}^d$. In this way, $\mathcal{H}$ acts as a Lyapunov functional for (2).

In order to reduce computational cost, the kinetic description of a gas provided by F is often simplified by retaining only a finite number of its velocity averages, or *moments*. Equations which govern the evolution of these moments are derived by integrating (2) with respect to a vector

$$\mathbf{m} = (m_0, \ldots, m_{n-1})^T \tag{7}$$

whose components are (typically) homogeneous polynomials in v. These equations take the form

$$\partial_t \boldsymbol{\rho} + \nabla_x \cdot \langle v\mathbf{m}F\rangle = \langle \mathbf{m}\mathcal{C}(F)\rangle, \tag{8}$$

where the moments

$$\boldsymbol{\rho} = \boldsymbol{\rho}(x,t) \equiv \langle \mathbf{m}F\rangle \tag{9}$$

are the *spatial densities* associated with F.

In general, (8) is not a closed system because there is no way to express the flux terms $\langle v\mathbf{m}F\rangle$ and collision terms $\langle \mathbf{m}\mathcal{C}(F)\rangle$ in terms of $\boldsymbol{\rho}$. Furthermore, in a moment description, an exact expression for F is not available. An alternative is to approximate F by an ansatz of the form

$$\mathcal{F}[\boldsymbol{\rho}] = \mathcal{F}(v, \boldsymbol{\rho}(x,t)). \tag{10}$$

By substituting $\mathcal{F}$ for F in (8), the evolution of $\boldsymbol{\rho}$ can be approximated by the closed system of balance laws

$$\partial_t \boldsymbol{\rho} + \nabla_x \cdot \mathbf{f}(\boldsymbol{\rho}) = \mathbf{c}(\boldsymbol{\rho}), \tag{11}$$

where the flux term $\mathbf{f}$ and collision term $\mathbf{c}$ are given by

$$\mathbf{f}(\boldsymbol{\rho}) = \langle v\mathbf{m}\mathcal{F}[\boldsymbol{\rho}]\rangle \qquad \text{and} \qquad \mathbf{c}(\boldsymbol{\rho}) = \langle \mathbf{m}\mathcal{C}(\mathcal{F}[\boldsymbol{\rho}])\rangle. \tag{12}$$

One way to specify $\mathcal{F}$ is to invoke the principle of entropy minimization (or *maximization* in the physics community, where the term "entropy" refers to $-\mathcal{H}$ and has been widely used for over a century). The probabilistic interpretation of entropy dates back to Boltzmann [6], [7], who argued that the entropy of a system of identical particles depends on the number of microstates (particle arrangements in phase space) that are consistent with the macroscopic state of the system. This dependence is expressed by the famous logarithmic relationship known as *Boltzmann's entropy formula* [8] (and also as Boltzmann's equation, although distinct from (2)) and was first presented in its popular form by Planck [9], [10]. The practical application of entropy as a tool for statistical inference was championed by Jaynes although, in [11], Jaynes himself attributes the original mathematical concepts to Gibbs, who generalized Boltzmann's entropy formula [12]. Jaynes also credits Shannon [13] for illuminating the central role that entropy plays in the theory of information. The relationship between statistics and information theory was further pursued by Kullback [14]. Many of the first rigorous results concerning entropy minimization can be found in the work of Csiszar [15] and references therein.

Closures which are based on the entropy minimization principle use the ansatz

$$\mathcal{F}[\boldsymbol{\rho}] = \arg\min_{g \in \mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m}g\rangle = \boldsymbol{\rho}\} \tag{13}$$

at each x and t to formally close (8). Here

$$\mathbb{F}_{\mathbf{m}} \equiv \left\{ g \in L^1(\mathbb{R}^d) : g \gneqq 0 \text{ and } |\mathbf{m}g| \in L^1(\mathbb{R}^d)\right\}, \tag{14}$$

and $|\cdot|$ is the standard Euclidean norm. The vector $\mathbf{m}$ is not arbitrary; it must form a basis for an admissible polynomials space $\mathbb{M}$ that satisfies certain physically motivated properties [2].

If the minimizer in (13) exists, it is unique and the closure is well-defined. In such cases, (11) is a hyperbolic system of PDEs whose solutions satisfy the local dissipation law

$$\partial_t h(\boldsymbol{\rho}) + \nabla_x \cdot j(\boldsymbol{\rho}) = s(\boldsymbol{\rho}), \tag{15}$$

where

$$h(\boldsymbol{\rho}) \equiv \mathcal{H}(\mathcal{F}[\boldsymbol{\rho}]) \tag{16}$$

is a strictly convex function of $\boldsymbol{\rho}$ and where

$$j(\boldsymbol{\rho}) \equiv \mathcal{J}(\mathcal{F}[\boldsymbol{\rho}]), \qquad s(\boldsymbol{\rho}) \equiv \mathcal{S}(\mathcal{F}[\boldsymbol{\rho}]) \leq 0. \tag{17}$$

Although any choice for the ansatz $\mathcal{F}[\boldsymbol{\rho}]$ will yield a system of the form (11), it is the entropy ansatz that gives (15). This dissipation law implies the existence of a well-posed linear L^2 (Hilbert space) theory for (11) [16]. Furthermore, h acts as a Lyapunov function for (11). To see this, note that (15) is simply (3) evaluated at $F = \mathcal{F}[\boldsymbol{\rho}]$; and like in Boltzmann's H-theorem, $s(\boldsymbol{\rho})$ vanishes if and only if $\mathcal{C}(\mathcal{F}[\boldsymbol{\rho}]) = 0$, in which case $\mathcal{F}[\boldsymbol{\rho}]$ takes the form of a Maxwellian distribution [2].

The entropy minimization procedure yields an entire hierarchy of systems with the aforementioned properties whose members are generated by appending an initial choice of $\mathbf{m}$ with additional polynomial components. For this reason, entropy-based closures have been applied to other areas of kinetic theory such as radiation transport [17], [18] and charge transport in semiconductors [19]–[21]. (Additional references for charge transport can be found in [19].) In the case of gas dynamics, the moment hierarchy begins with the canonical choice $\mathbf{m} = (1, v_1, \ldots, v_d, \frac{1}{2}|v|^2)^T$. For this choice, $\mathcal{F}[\boldsymbol{\rho}]$ is always a Maxwellian, and the entropy-based closure generates Euler's equations for a compressible gas.

B. Realizability and degenerate densities

The formal structure of entropy-based closures hinges on the assumption that for each $\boldsymbol{\rho}$ in the class of *realizable densities*

$$\mathcal{R}_{\mathbf{m}} \equiv \{\boldsymbol{\rho} \in \mathbb{R}^n : \boldsymbol{\rho} = \langle \mathbf{m} g \rangle,\, g \in \mathbb{F}_{\mathbf{m}}\}, \tag{18}$$

there is a minimizer for (13). However, for most choices of $\mathbf{m}$, there exist realizable values of $\boldsymbol{\rho}$ for which such a minimizer does not exist. For these densities, which we term *degenerate*, the entropy-based closure is not well-defined and modifications must be made to the entropy-based procedure. There are essentially two approaches:

1) Show that the set of nondegenerate densities is invariant under the dynamics of the balance law (11) with the entropy-based closure (as discussed in [22]) or impose such a condition in a way that is physically reasonable and mathematically justifiable.
2) Develop a modified closure that (i) is well-posed for *all* physically realizable values of $\boldsymbol{\rho}$, (ii) recovers the entropy-based closure whenever the minimizer in (13) exists, and (iii) generates systems of hyperbolic PDEs that dissipate a physically meaningful, convex entropy. This was attempted in [23].

For either approach, it is important to show that the set $\mathcal{D}_{\mathbf{m}}$ of degenerate densities is small in some sense, thereby minimizing the class of physically realizable spatial densities which require special treatment. In the first approach, this means limiting the number of initial conditions which must be discarded; in the second, it means limiting the number of densities in $\mathcal{R}_{\mathbf{m}}$ which require a modified closure.

Another reason to study $\mathcal{D}_{\mathbf{m}}$ is that the equilibrium densities, which are moments of a Maxwellian distribution (6), lie on its boundary [2], [4], [22], [23]. Because the kinetic entropy drives solutions of (3) toward local thermal equilibrium, trajectories defined by solutions to (15) will, at times, come very close to $\mathcal{D}_{\mathbf{m}}$. Thus it is very important to have a detailed understanding of its geometry.

C. Previous work

Previous studies of the set $\mathcal{D}_{\mathbf{m}}$ can be found in [4], [22], [23]. In [22], Junk provides a geometric description for $\mathcal{D}_{\mathbf{m}}$ in a one-dimensional setting ($d = 1$) with $\mathbf{m} = (1, v, v^2, v^3, v^4)^T$. In turns out in this case that $\mathcal{D}_{\mathbf{m}}$ is a codimension one manifold. This result was discovered, in part, by extending the definition of h given by (17) to include cases where the minimizer in (13) does not exist. This is done by replacing the minimum in (13) with an infimum, viz.,

$$h_{\mathrm{J}}(\boldsymbol{\rho}) \equiv \inf_{g \in \mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m} g \rangle = \boldsymbol{\rho}\}. \tag{19}$$

Later, in [4], Junk considers a more general case in which $\mathbf{m}$ consists of a radial component $|v|^N$, for some even integer $N \geq 2$, plus polynomial components of lower degree. For such cases, he provides an integrability condition to determine whether $\mathcal{D}_{\mathbf{m}}$ is nonempty. In practice, this condition is easily checked and extensible to more general choices of $\mathbf{m}$. However, a description of the geometry of $\mathcal{D}_{\mathbf{m}}$, as given in [22], is still lacking for the general setting.

In [23], Schneider introduces a different extension for h by relaxing the constraints in (13). In order to state this problem precisely, we decompose $\mathbf{m}$ into subvectors:

$$\mathbf{m} = (\mathbf{m}_0^T, \mathbf{m}_1^T, \mathbf{m}_2^T, \ldots, \mathbf{m}_N^T)^T, \tag{20}$$

where the n_j components of $\mathbf{m}_j$ are the homogeneous, jth degree polynomial components of $\mathbf{m}$. Thus any polynomial $p \in \mathbb{M}$ can be expressed as the sum of its homogeneous components:

$$p = \boldsymbol{\alpha}^T \mathbf{m} = \sum_{j=1}^{N} \boldsymbol{\alpha}_j^T \mathbf{m}_j, \tag{21}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^n$ is a vector of constant coefficients that decomposes into subvectors

$$\boldsymbol{\alpha} = \left(\boldsymbol{\alpha}_0^T, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_2^T, \ldots, \boldsymbol{\alpha}_N^T\right)^T. \tag{22}$$

With the preceding notation, the relaxed constraint problem is

$$h_{\mathrm{S}}(\boldsymbol{\rho}) \equiv \min_{g \in \mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m} g \rangle \preceq^{\circ} \boldsymbol{\rho}\}, \tag{23}$$

where the relation $\langle \mathbf{m} g \rangle \preceq^{\circ} \boldsymbol{\rho}$ means that

$$\langle \mathbf{m}_j g \rangle = \boldsymbol{\rho}_j,\ 0 \leq j \leq N-1, \tag{24}$$

and

$$\boldsymbol{\alpha}_N^T \langle \mathbf{m}_N g \rangle \leq \boldsymbol{\alpha}_N^T \boldsymbol{\rho}_N \quad \text{whenever} \quad \boldsymbol{\alpha}_N^T \mathbf{m}_N \geq 0. \tag{25}$$

The benefit of using the relaxed constraint set in (23) is that it is closed in the weak-$L^1(\mathbb{R}^d)$ topology, and as a result, the minimizer in (23) always exists. While a minimizing sequence for $\mathcal{H}$ in the constraint set for (13) does converge weakly in $L^1(\mathbb{R}^d)$, the constraint set is not closed with respect to this topology. Thus the infimum might not be attained.

D. Overview of Main Results

Our main contribution is a geometrical description of the set $\mathcal{D}_{\mathbf{m}}$ in a general setting, based on a dual formulation of (23). Our results, which recover and extend many previous results from [4], [22], [23] are summarized in the following theorems.

- In Theorem 4, we prove strong duality for both the equality constraint problem (19) and the relaxed constraint problem (23). We conclude that $h_{\mathrm{S}} = h_{\mathrm{J}}$, even when the infimum in (23) is not attained. We also prove a *complementary slackness condition* for (23) which serves as the basis of our geometrical description.
- In Theorem 7, we show that the set $\mathcal{D}_{\mathbf{m}}$ is a union of convex cones that are defined by complementary slackness condition from Theorem 4.
- In Theorem 9, we show that, under reasonable assumptions, the set $\mathcal{D}_{\mathbf{m}}$ is a nowhere dense subset of $\mathcal{R}_{\mathbf{m}}$ that has Lebesgue measure zero and is restricted to the boundary of the nondegenerate, realizable densities. The assumptions we employ hold in all known cases. Whether they hold in general is (to our knowledge) an open question in analysis and algebraic geometry.

II. Entropy minimization with relaxed constraints

The main result from [23] for the relaxed constraint problem is the following.

Theorem 1 (Schneider [23]): For any $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, there is a unique minimizer for (23) of the form $G_{\boldsymbol{\alpha}}$, where

$$G_{\boldsymbol{\alpha}} \equiv \exp(\boldsymbol{\alpha}^T \mathbf{m}) \tag{26}$$

and

$$\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}} \equiv \{\boldsymbol{\alpha} \in \mathbb{R}^n : G_{\boldsymbol{\alpha}} \in \mathbb{F}_{\mathbf{m}}\}. \tag{27}$$

Conversely,

$$\begin{aligned} \mathcal{H}(G_{\boldsymbol{\alpha}}) &= \min_{g \in \mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m} g \rangle \preceq^{\circ} \langle \mathbf{m} G_{\boldsymbol{\alpha}} \rangle\} \\ &= \min_{g \in \mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m} g \rangle = \langle \mathbf{m} G_{\boldsymbol{\alpha}} \rangle\}, \end{aligned} \tag{28}$$

for each $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$.

We define $\mathbf{a} : \mathcal{R}_{\mathbf{m}} \to \mathcal{A}_{\mathbf{m}}$ as the mapping which assigns to $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ the vector $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$ such that $G_{\boldsymbol{\alpha}}$ solves (23)—that is,

$$G_{\mathbf{a}(\boldsymbol{\rho})} \equiv \arg \min_{g \in \mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m} g \rangle \preceq^{\circ} \boldsymbol{\rho}\}. \tag{29}$$

We define $\mathbf{r} : \mathcal{A}_{\mathbf{m}} \to \mathbb{R}^n$ as the mapping which generates the moments of $G_{\boldsymbol{\alpha}}$:

$$\mathbf{r}(\boldsymbol{\alpha}) \equiv \langle \mathbf{m} G_{\boldsymbol{\alpha}} \rangle. \tag{30}$$

The image of $\mathcal{A}_{\mathbf{m}}$ under $\mathbf{r}$ is the set of *exponentially realizable densities*:

$$\mathcal{R}_{\mathbf{m}}^{\exp} \equiv \mathbf{r}(\mathcal{A}_{\mathbf{m}}) \subset \mathcal{R}_{\mathbf{m}}. \tag{31}$$

The following theorem relates $\mathbf{r}$ and $\mathbf{a}$.

Theorem 2: The mapping $\mathbf{r}$ is one-to-one from $\mathcal{A}_{\mathbf{m}}$ onto $\mathcal{R}_{\mathbf{m}}^{\exp}$ with inverse $\mathbf{a}$. It is a diffeomorphism between $\operatorname{int} \mathcal{A}_{\mathbf{m}}$ and $\operatorname{int} \mathcal{R}_{\mathbf{m}}^{\exp}$.

Proof: We first identify $\mathbf{a}$ as the inverse of $\mathbf{r}$. Since $\mathbf{r}$ is onto $\mathcal{R}_{\mathbf{m}}^{\exp}$, we need only to show that $\mathbf{a}(\mathbf{r}(\boldsymbol{\alpha})) = \boldsymbol{\alpha}$ for each $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$. According to Theorem 1 and the implicit definition of $\mathbf{a}$ in (29)

$$\mathcal{H}(G_{\boldsymbol{\alpha}}) = \min_{g \in \mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m} g \rangle \preceq^{\circ} \mathbf{r}(\boldsymbol{\alpha})\} = \mathcal{H}(G_{\mathbf{a}(\mathbf{r}(\boldsymbol{\alpha}))}) \tag{32}$$

and, since the minimizer is unique, it follows that $\mathbf{a}(\mathbf{r}(\boldsymbol{\alpha})) = \boldsymbol{\alpha}$. On $\operatorname{int} \mathcal{A}_{\mathbf{m}}$ $\mathbf{r}$ is smooth with Jacobian

$$\frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}) = \frac{\partial^2 h^*}{\partial \boldsymbol{\alpha}^2}(\boldsymbol{\alpha}, \boldsymbol{\rho}) = \langle \mathbf{m}\mathbf{m}^T G_{\boldsymbol{\alpha}} \rangle \tag{33}$$

that is a positive-definite matrix. The inverse function theorem implies then that $\mathbf{r}$ is a diffeomorphism from $\operatorname{int} \mathcal{A}_{\mathbf{m}}$ onto $\operatorname{int} \mathcal{R}_{\mathbf{m}}^{\exp}$. ∎

III. Dual formulation

Because $\mathcal{H}$ is convex on $\mathbb{F}_{\mathbf{m}}$ and the constraints in (23) are linear, a dual treatment to the relaxed-constraint problem, e.g., [24]–[26] is appropriate.

A. The dual function

The Lagrangian function $\mathcal{L} : \mathbb{F}_{\mathbf{m}} \times \mathbb{R}^n \times \mathcal{R}_{\mathbf{m}} \to \mathbb{R} \cup \{\infty\}$ associated to (23) is

$$\mathcal{L}(g, \boldsymbol{\alpha}, \boldsymbol{\rho}) \equiv \mathcal{H}(g) + \boldsymbol{\alpha}^T (\boldsymbol{\rho} - \langle \mathbf{m} g \rangle) \tag{34}$$

and the dual function $\psi : \mathbb{R}^n \times \mathcal{R}_{\mathbf{m}} \to \mathbb{R} \cup \{-\infty\}$ is

$$\psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) \equiv \inf_{g \in \mathbb{F}_{\mathbf{m}}} \mathcal{L}(g, \boldsymbol{\alpha}, \boldsymbol{\rho}). \tag{35}$$

Theorem 3: For all $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$ and $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, the dual function is

$$\psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) = \mathcal{L}(G_{\boldsymbol{\alpha}}, \boldsymbol{\alpha}, \boldsymbol{\rho}) = \boldsymbol{\alpha}^T \boldsymbol{\rho} - \langle G_{\boldsymbol{\alpha}} \rangle. \tag{36}$$

B. Duality theorems

The following strong duality theorem is an application of a general result from [24, Exercise 8.7] and can be proven following the arguments found in [24, Chapter 8].

Theorem 4: Let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, and let h_{J}, h_{S}, and ψ be given by (19), (23), and (35), respectively. Then

$$h_{\mathrm{J}}(\boldsymbol{\rho}) = h_{\mathrm{S}}(\boldsymbol{\rho}) = \max_{\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}} \psi(\boldsymbol{\alpha}, \boldsymbol{\rho}), \tag{37}$$

where the maximum on the right is attained by a unique $\hat{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}}$. Furthermore, $G_{\hat{\boldsymbol{\alpha}}}$ and $\hat{\boldsymbol{\alpha}}$ satisfy the complementary slackness condition

$$\hat{\boldsymbol{\alpha}}^T \boldsymbol{\rho} = \hat{\boldsymbol{\alpha}}^T \langle \mathbf{m} G_{\hat{\boldsymbol{\alpha}}} \rangle, \tag{38}$$

and $G_{\hat{\boldsymbol{\alpha}}}$ minimizes $\mathcal{L}(g, \hat{\boldsymbol{\alpha}}, \boldsymbol{\rho})$ over $\mathbb{F}_{\mathbf{m}}$, i.e., $\psi(\hat{\boldsymbol{\alpha}}, \boldsymbol{\rho}) = \mathcal{L}(G_{\hat{\boldsymbol{\alpha}}}, \hat{\boldsymbol{\alpha}}, \boldsymbol{\rho})$.

In light of (37), the definition of h given in (16), which applies only to $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$, can be extended to all of $\mathcal{R}_{\mathbf{m}}$ by setting

$$h(\boldsymbol{\rho}) \equiv \max_{\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}} \psi(\boldsymbol{\alpha}, \boldsymbol{\rho}). \tag{39}$$

Theorem 4 is used to prove the following result.

Theorem 5: Given $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, the minimization problem with equality constraints (13) has a minimizer if and only if $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$. In other words,

$$\mathcal{D}_{\mathbf{m}} = \mathcal{R}_{\mathbf{m}} \backslash \mathcal{R}_{\mathbf{m}}^{\exp}. \tag{40}$$

Proof: The converse statement of Theorem 1 implies the "if" statement of Theorem 5. To prove the "only if" statement, let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ be such that (13) has a minimizer. According to (37), this minimizer is also the minimizer of (23) and is therefore given by $G_{\mathbf{a}(\boldsymbol{\rho})}$. Hence, the equality constraint conditions in (13) imply that $\boldsymbol{\rho} = \langle \mathbf{m} G_{\mathbf{a}(\boldsymbol{\rho})} \rangle$, which means $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$. ∎

The essential point of Theorem 5 is that when $\mathcal{D}_{\mathbf{m}}$ is nonempty, there are realizable densities $\boldsymbol{\rho}$ that *cannot* be realized by a functions of the form $G_{\boldsymbol{\alpha}}$. In other words, $\boldsymbol{\rho} \notin \mathcal{R}_{\mathbf{m}}^{\exp}$ even though $\mathbf{a}(\boldsymbol{\rho}) \in \mathcal{A}_{\mathbf{m}}$. It is this idea which lays the foundation for the results in [4], [22] and for the new results of this paper. However, we still need the complementary slackness condition in order to find a useful geometric description for $\mathcal{D}_{\mathbf{m}}$.

IV. Geometry of $\mathcal{D}_{\mathbf{m}}$

The complementary slackness condition (38) which relates $\hat{\boldsymbol{\alpha}}_N$ and $\boldsymbol{\rho}_N$ is the key to characterizing $\mathcal{D}_{\mathbf{m}}$. Indeed, this condition is used to define the convex cones from which $\mathcal{D}_{\mathbf{m}}$ is composed.

A. Motivation: Behavior of the closure near degeneracy

Even though $\mathcal{D}_{\mathbf{m}}$ is usually nonempty, there is evidence to suggest that if $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$ initially, then densities in $\mathcal{D}_{\mathbf{m}}$ might never be attained during the evolution of the moment system (8). To investigate this possibility, we introduce the function $\chi : \mathcal{R}_{\mathbf{m}} \to \mathbb{R}$, defined by

$$\chi(\boldsymbol{\rho}) \equiv \int_{\mathbb{R}^d} |v\mathbf{m}(v)|\, G_{\mathbf{a}(\boldsymbol{\rho})}(v)\, dv. \tag{41}$$

For the entropy-based closure, χ is closely related to the flux $\mathbf{f}$ in (12), and we show below that χ becomes unbounded as $\boldsymbol{\rho}$ approaches $\mathcal{D}_{\mathbf{m}}$. As pointed out in [22], such divergent behavior raises the possibility that $\mathcal{R}_{\mathbf{m}}^{\exp}$ is invariant under the dynamics of the closure.

Proposition 6: Let $\{\boldsymbol{\rho}_{(j)}\}_{j=1}^{\infty}$ be a sequence in $\mathcal{R}_{\mathbf{m}}^{\exp}$ such that $\boldsymbol{\rho}_{(j)} \to \boldsymbol{\rho}_* \in \mathcal{D}_{\mathbf{m}}$, and for each j, let $\chi_j \equiv \chi(\boldsymbol{\rho}_{(j)})$. Then $\{\chi_j\}_{j=1}^{\infty}$ is unbounded.

Suppose now that it can be proven that $\mathcal{R}_{\mathbf{m}}^{\exp}$ is invariant under the dynamics of the balance law (11) with the entropy-based closure. Then if $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$ initially, the entropy minimization problem with equality constraints (13) will always have a solution, and the formal properties of the closure will be maintained. However, it must be shown—at a minimum—that $\mathcal{D}_{\mathbf{m}}$ is small in some sense, thereby limiting the number of initial conditions in $\mathcal{R}_{\mathbf{m}}$ which must be discarded in order to maintain a well-defined closure. In the following subsections, we show that, under reasonable hypotheses, $\mathcal{D}_{\mathbf{m}}$ is indeed a Lebesgue measure zero set.

B. The complementary slackness condition and normal cones

Because $\boldsymbol{\rho}_j = \langle \mathbf{m}_j G_{\hat{\boldsymbol{\alpha}}} \rangle$ for $j < N$, the only nontrivial part of the complementary slackness condition (38) is

$$\hat{\boldsymbol{\alpha}}_N^T \boldsymbol{\rho}_N = \hat{\boldsymbol{\alpha}}_N^T \langle \mathbf{m}_N G_{\hat{\boldsymbol{\alpha}}} \rangle. \tag{42}$$

This condition and the inequality constraint (25) from the primal problem imply that

$$(\boldsymbol{\alpha}_N - \hat{\boldsymbol{\alpha}}_N)^T (\boldsymbol{\rho}_N - \langle \mathbf{m}_N G_{\hat{\boldsymbol{\alpha}}} \rangle) \le 0 \tag{43}$$

for all $\boldsymbol{\alpha}$ in the cone

$$A_{\mathbf{m}_N} \equiv \{\boldsymbol{\alpha}_N \in \mathbb{R}^{n_N} : \boldsymbol{\alpha}_N^T \mathbf{m}_N \le 0\}. \tag{44}$$

Thus $\boldsymbol{\rho}_N$ is contained in $\mathcal{NC}(A_{\mathbf{m}_N}, \hat{\boldsymbol{\alpha}}_N)$: the normal cone [26] of $A_{\mathbf{m}_N}$ with respect to $\hat{\boldsymbol{\alpha}}_N$ and with vertex $\langle \mathbf{m}_N G_{\hat{\boldsymbol{\alpha}}} \rangle$. The converse statement is also true: Given $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$, every element $\boldsymbol{\rho}_N$ of the normal cone $\langle \mathbf{m}_N G_{\boldsymbol{\alpha}} \rangle + \mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N)$ is the N-th subvector of a density in $\mathcal{R}_{\mathbf{m}}$. To state this more precisely, let $\mathbf{a}_N(\boldsymbol{\rho})$ be the N-th subvector of $\mathbf{a}(\boldsymbol{\rho})$, and for any $y \in \mathbb{R}^n$ and $z \in \mathbb{R}_N^n$ let

$$y +_N z = (y_0^T, y_1^T, \ldots, y_{N-1}^T, y_N^T + z_N^T)^T. \tag{45}$$

Theorem 7: The set $\mathcal{R}_{\mathbf{m}}$ can be expressed as the following union of cones:

$$\mathcal{R}_{\mathbf{m}} = \bigcup_{\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp}} \{\bar{\boldsymbol{\rho}} +_N \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))\}, \tag{46}$$

where the relation '$+_N$' is defined in (45).

For $\bar{\boldsymbol{\rho}} \in \operatorname{int} \mathcal{R}_{\mathbf{m}}^{\exp}$, $\mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))$ is just the origin in $\mathbb{R}^{n_N}$. In such cases, Theorem 7 is trivial, and the construction $\bar{\boldsymbol{\rho}} +_N \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))$ does not generate any new densities. Therefore $\mathcal{D}_{\mathbf{m}}$ is constructed entirely by normal cones attached to $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial \mathcal{R}_{\mathbf{m}}^{\exp}$.

Corollary 8: The degenerate densities are given by

$$\mathcal{D}_{\mathbf{m}} = \bigcup_{\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial \mathcal{R}_{\mathbf{m}}^{\exp}} \{\bar{\boldsymbol{\rho}} +_N \mathcal{NC}_0(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))\} \tag{47}$$

$$= \bigcup_{\bar{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}} \{\mathbf{r}(\bar{\boldsymbol{\alpha}}) +_N \mathcal{NC}_0(A_{\mathbf{m}_N}, \bar{\boldsymbol{\alpha}}_N)\}, \tag{48}$$

where $\mathcal{NC}_0(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}})) \equiv \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}})) \backslash \{0\}$.

C. Smoothness assumptions on $\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}$

Corollary 8 gives the degenerate densities associated with each $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial \mathcal{R}_{\mathbf{m}}^{\exp}$. However, a clean description of $\mathcal{D}_{\mathbf{m}}$ requires also that $\mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial \mathcal{R}_{\mathbf{m}}^{\exp}$ itself have a nice structure. In particular, we would like to say that $\mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial \mathcal{R}_{\mathbf{m}}^{\exp}$ is a finite union of disjoint manifolds. At this point we are unable to prove such a result in general, in part due to the complicated structure of $\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}$ (the preimage of $\mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial \mathcal{R}_{\mathbf{m}}^{\exp}$ with respect to $\mathbf{r}$). We therefore make two assumptions: first, that $\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}$ is a union of disjoint manifolds with dimensional restrictions that are related to the dimensions of the normal cones in (48) in such a way as to ensure that $\mathcal{D}_{\mathbf{m}}$ is a lower-dimensional subset of $\mathcal{R}_{\mathbf{m}}$; and second, that the mapping $\mathbf{r}$ is diffeomorphic when restricted to each of these manifolds. Thus each dimension k manifold in $\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}$ will be mapped to a dimension k manifold in $\mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial \mathcal{R}_{\mathbf{m}}^{\exp}$. Before stating our assumptions, we define the orthogonal projections $\mathcal{P}_N : \mathbb{R}^n \mapsto \mathbb{R}^{n_N}$ and $\mathcal{P}_{\tilde{N}} : \mathbb{R}^n \mapsto \mathbb{R}^{n-n_N}$ by

$$\mathcal{P}_N(\boldsymbol{\alpha}) \equiv (0, \ldots, 0, 0, \boldsymbol{\alpha}_N^T)^T, \tag{49}$$

$$\mathcal{P}_{\tilde{N}}(\boldsymbol{\alpha}) \equiv \boldsymbol{\alpha} - \mathcal{P}_N(\boldsymbol{\alpha}) = (\boldsymbol{\alpha}_0^T, \boldsymbol{\alpha}_1^T, \ldots, \boldsymbol{\alpha}_{N-1}^T, 0)^T. \tag{50}$$

Assumption I. The set $\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}$ can be decomposed into a finite collection $\mathcal{S}$ of disjoint, smooth (C^∞) manifolds in $\mathbb{R}^n$. Furthermore, if S is one such manifold, then $\mathcal{P}_N$ projects S onto a manifold $S_N \subset \partial A_{\mathbf{m}_N}$ with codimension at least one in $\mathbb{R}^{n_N}$ and $\mathcal{P}_{\tilde{N}}$ projects S onto a manifold $S_{\tilde{N}}$ of codimension at least one in $\mathbb{R}^{n-n_N}$.

Assumption II. If Assumption I holds and if S is an element of the stratification of $\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}$, then for each $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, the restriction of $\mathbf{r}$ to S is infinitely Fréchet differentiable.

When both Assumptions I and II hold, $\mathbf{r}$ is a smooth diffeomorphism with inverse $\mathbf{a}$ when restricted to any manifold in the stratification of $\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}$. It should be noted that Assumptions I and II are known to hold for the examples considered in [4], [22], [23]. Whether or not they hold in general is, to our knowledge, an open question.

D. Smallness of $\mathcal{D}_{\mathbf{m}}$

If Assumptions I and II hold, we can show that $\mathcal{D}_{\mathbf{m}}$ is small in the following sense.

Theorem 9: Suppose that Assumptions I and II hold. Then $\mathcal{D}_{\mathbf{m}}$ has zero Lebesgue measure, $\operatorname{int}\mathcal{R}_{\mathbf{m}}^{\exp}$ is a dense subset of $\mathcal{R}_{\mathbf{m}}$, and $\mathcal{D}_{\mathbf{m}} \subset \partial\mathcal{R}_{\mathbf{m}}^{\exp}$.

Proof: The basic idea of the argument is that the image of a smooth map from a lower-dimensional space to a higher-dimensional space has zero Lebesgue measure. We can construct such a map, whose image cover a portion of $\mathcal{D}_{\mathbf{m}}$, and then cover $\mathcal{D}_{\mathbf{m}}$ with the images from a countable number of similar maps. ■

V. Conclusions and discussion

We have given a description of the set $\mathcal{D}_{\mathbf{m}}$ of degenerate densities based on a geometric interpretation of the complementary slackness condition associated with the dual formulation of (23). Roughly speaking, the set $\mathcal{D}_{\mathbf{m}}$ is constructed by attaching a convex cone to every point in the boundary component $\partial\mathcal{R}_{\mathbf{m}}^{\exp}\cap\mathcal{R}_{\mathbf{m}}^{\exp}$. This description recovers and extends previous results concerning the constrained entropy minimization problem.

Analytically, we see three important open questions that must be solved. First, one must determine if Assumptions I and II hold in a general setting. Concerning Assumption I, this means understanding the structure of the set of polynomials p for which $v \mapsto p(v)e^{p(v)}$ is Lebesgue integrable. For example, do the coefficients of such polynomials form a semialgebraic set? Second, it must be determined whether the sets $\mathcal{R}_{\mathbf{m}}^{\exp}$ and $\mathcal{R}_{\mathbf{m}}$ are invariant under the dynamics of the balance law (11) with the entropy-based closure. (Although not discussed in this paper, such a condition on $\mathcal{R}_{\mathbf{m}}$ is obviously necessary for entropy-based closures to have any practical application.) Finally, it must be determined whether the existence of degenerate densities and the dynamics of (11) near such densities are simply artifacts of the entropy-based closure or if they actually reflect some physically relevant properties of the original Boltzmann equation (2).

Numerically speaking, a full implementation of entropy-based closures for gas dynamics faces many challenges. (An implementation has been attempted in [27], but the issue of degenerate densities was not addressed.) Clearly a discretization of (11) must preserve any invariant properties of $\mathcal{R}_{\mathbf{m}}$ and $\mathcal{R}_{\mathbf{m}}^{\exp}$ with respect to the balance law (11). As pointed out in [22], even if $\mathcal{R}_{\mathbf{m}}^{\exp}$ *is* invariant under (11), solving the dual optimization problem (34) becomes extremely difficult for $\boldsymbol{\rho}$ *near* $\mathcal{D}_{\mathbf{m}}$ because the function h^* is very hard to evaluate. The reason for this is that, as $\boldsymbol{\alpha}$ approaches $\partial\mathcal{A}_{\mathbf{m}}$, the function $G_{\boldsymbol{\alpha}}$ can develop isolated modes that are often overlooked in a numerical quadrature. The result is a regularization effect in which accuracy is lost. In addition, the matrix $\langle \mathbf{m}\mathbf{m}^T G_{\boldsymbol{\alpha}}\rangle$ becomes poorly conditioned near the boundary of $\mathcal{A}_{\mathbf{m}}$. Any minimization algorithm for (34) must be carefully formulated in order to overcome these challenges. Furthermore, as with the degenerate densities themselves, one must determine if these difficulties are by-products of the closure or related in some way to the dynamics of the Boltzmann equation.

References

[1] C. D. Hauck, C. D. Levermore, and A. L. Tits, "Convex duality and entropy-based moment closures: Characterizing degenerate densities," *SIAM J. Control and Optim.*, vol. 47, pp. 1977–2015, 2008.

[2] C. D. Levermore, "Moment closure hierarchies for kinetic theory," *J. Stat. Phys.*, vol. 83, pp. 1021–1065, 1996.

[3] J. M. Borwein and A. S. Lewis, "Duality relationships for entropy-like minimization problems," *SIAM J. Control Optim.*, vol. 1, pp. 191–205, 1991.

[4] M. Junk, "Maximum entropy for reduced moment problems," *Math. Mod. Meth. Appl. S.*, vol. 10, no. 7, pp. 1001–1025, 2000.

[5] C. Cercignani, R. Ilner, and M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, ser. Applied Mathematical Sciences 106. New York: Springer-Verlag, 1994.

[6] L. Boltzmann, "Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten," *Wien. Ber.*, vol. 58, pp. 517–560, 1868.

[7] ——, "Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen," *Wien. Ber.*, vol. 76, pp. 373–435, 1877.

[8] H. B. Callen, *Thermodynamics and an Introduction to Thermostatstics*, 2nd ed. New York: John Wiley and Sons, Inc., 1985.

[9] M. Planck, "Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum," *Verhandlungen der Deutschen Physikalischen Gesellschaft*, vol. 2, pp. 237–245, 1900.

[10] ——, "Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum," *Annalen der Physik*, vol. 4, pp. 553–563, 1901.

[11] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.*, vol. 106, no. 4, pp. 620–630, May 1957.

[12] J. W. Gibbs, *Elementary Principles in Statistical Mechanics*. New York: Charles Scribner's Son's, 1902.

[13] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, vol. 27, pp. 379–423 and 623–656, July and October 1948.

[14] S. Kullback, *Information Theory and Statistics*. New York: Wiley, 1959.

[15] I. Csiszar, "I-divergence geometry of probability distributions and minimization problems," *Ann. Probab.*, vol. 3, no. 1, pp. 146–158, 1975.

[16] J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, 2nd ed. Philadelphia: SIAM, 2004.

[17] B. Dubroca and J.-L. Fuegas, "Étude théorique et numérique d'une hiérarchie de modèles aus moments pour le transfert radiatif," *C.R. Acad. Sci. Paris*, vol. I. 329, pp. 915–920, 1999.

[18] B. Dubroca and A. Klar, "Half-moment closure for radiative transfer equations," *J. Comput. Phys.*, vol. 180, pp. 584–596, 2002.

[19] A. M. Anile, G. Mascali, and V. Romano, *Recent Developments in Hydrodynamical Modeling of Semiconductors*, ser. Lecture Notes in Mathematics. Berlin: Springer-Verlag, 2003, vol. 1823, pp. 1–56, lectures given at the C.I.M.E. Summer School held in Cetraro, Italy on July 15-22, 1998.

[20] P. Degond and C. Ringhofer, "Quantum moment hydrodynamics and the entropy principle," *J. Stat. Phys.*, vol. 112, pp. 587–627, 2003.

[21] M. Junk and V. Romano, "Maximum entropy systems of the semiconductor Boltzmann equation using Kane's dispersion relation." *Continuum Mech. Therm.*, vol. 17, pp. 247–267, 2004.

[22] M. Junk, "Domain of definition of Levermore's five moment system," *J. Stat. Phys.*, vol. 93, no. 5-6, pp. 1143–1167, 1998.

[23] J. Schneider, "Entropic approximation in kinetic theory," *Math. Model. Numer. Anal.*, vol. 38, pp. 541–561, 2004.

[24] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: John Wiley and Sons, Inc., 1969.

[25] S. Boyd and L. Vandenderghe, *Convex Optimization*. New York: Cambridge University Press, 2004.

[26] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont, Massachusetts: Athena Scientific, 2003.

[27] P. L. Tallec and J. P. Perlat, "Numerical analysis of Levermore's moment system," *INRIA preprint*, vol. 3124, pp. 1–36, 1997.