



HAL
open science

A MOMENT CLOSURE BASED ON A PROJECTION ON THE BOUNDARY OF THE REALIZABILITY DOMAIN: 1D CASE

Teddy Pichard

► **To cite this version:**

Teddy Pichard. A MOMENT CLOSURE BASED ON A PROJECTION ON THE BOUNDARY OF THE REALIZABILITY DOMAIN: 1D CASE. 2019. hal-02423328v1

HAL Id: hal-02423328

<https://hal.science/hal-02423328v1>

Preprint submitted on 23 Dec 2019 (v1), last revised 5 Aug 2020 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A MOMENT CLOSURE BASED ON A PROJECTION ON THE BOUNDARY OF THE REALIZABILITY DOMAIN: 1D CASE

TEDDY PICHARD

CMAP, École Polytechnique, CNRS UMR7641,
Institut Polytechnique de Paris, Palaiseau, France

(Communicated by the associate editor name)

ABSTRACT. This work aims to develop and test a projection technique for the construction of closing equations of moment systems. One possibility to define such a closure consists in reconstructing an underlying kinetic distribution from a vector of moments, then expressing the closure based on this reconstructed function.

Exploiting the geometry of the realizability domain, *i.e.* the set of moments of positive distribution function, we decompose any realizable vectors into two parts, one corresponding to the moments of a chosen equilibrium function, and one obtain by a projection onto the boundary of the realizability domain in the direction of equilibrium function. A realizable closure of both of these parts are computed with standart techniques providing a realizable closure for the full system.

Such a technique is available for moments over $[-1, +1]$, over \mathbb{R} or over \mathbb{R}^+ , and is tested for the reduction of a radiative transfer equation in slab geometry.

1. Introduction. This paper aims to develop closure relations for 1D moment models that is based on positive measures and that recovers both purely anisotropic distribution and a chosen regular equilibrium function. The starting point is a kinetic equation of the form

$$\partial_t f + s \partial_x f = C(f), \quad (1)$$

where the unknown f is a distribution function depending on time $t \in \mathbb{R}^+$, position $x \in \Omega \subset \mathbb{R}$ and a state variable $s \in E \subset \mathbb{R}$. In the applications we have in mind, this state variable s corresponds respectively to a cosine direction of flight $s = \mu$ belonging to $E = [-1, +1]$ in radiative transfer, a velocity variable $s = v$ belonging to $E = \mathbb{R}$ in rarefied gas dynamics, or a size of droplets $s \in E = \mathbb{R}^+$ in dispersed flows.

In (1), the unknown f is a density of particles in a phase space $\Omega \times E$, *i.e.*

$$dN = f(t, s, x) ds dx$$

is the quantity of particles in a spatial neighbourhood dx around x and having a state in the neighbourhood ds around s at time t . The operator C models collision effects. Here, we consider only 1D problems such that the space variable and the state variable evolve respectively in subsets of \mathbb{R} .

2010 *Mathematics Subject Classification.* 35L40, 47B35, 35B09.

Key words and phrases. Moment closure, truncated moment problem, realizability property, projection technique.

Due to the high dimensionality of the phase space, equations of the form (1) are often solved numerically using a reduction technique based on a moment extraction (see *e.g.* [14, 40, 18]). Such a technique consists in studying the moments of f according to the s variable instead of f itself. Such moments depend on less variables than f and require therefore less computational efforts to compute (see *e.g.* comparisons in [45, 32, 51, 3, 6, 44, 43]). Those moments, afterward written \mathbf{f} in bold, are weighted integrals of f against a vector $\mathbf{b}(s)$ of polynomial weights. This yields

$$\mathbf{f} = \int_E \mathbf{b}(s)f(s)ds. \quad (2)$$

The moments of the kinetic distribution function f follow an equation of the form

$$\partial_t \mathbf{f} + \partial_x \mathbf{F} = \mathbf{C}, \quad (3a)$$

$$\mathbf{F} = \int_E s\mathbf{b}(s)f(s)ds, \quad \mathbf{C} = \int_E \mathbf{b}(s)C(f)(s)ds. \quad (3b)$$

The system (3a) is underdetermined, because the flux \mathbf{F} and the collision operator \mathbf{C} do not depend on the unknown \mathbf{f} . One common idea to close the system (3a) consists in expressing \mathbf{f} as the moments of an underlying kinetic distribution which is constructed such that it has \mathbf{f} for moments. In practice, one inverts (2), *i.e.* one seeks a function \tilde{f} satisfying

$$\int_E \mathbf{b}(s)\tilde{f}(s)ds = \mathbf{f}, \quad (4a)$$

then, the system (3a) is closed by approaching \mathbf{F} and \mathbf{C} based on \tilde{f} by

$$\mathbf{F} \equiv \tilde{\mathbf{F}}(\mathbf{f}) = \int_E s\mathbf{b}(s)\tilde{f}(s)ds, \quad \mathbf{C} \equiv \tilde{\mathbf{C}}(\mathbf{f}) = \int_E \mathbf{b}(s)C(\tilde{f})(s)ds. \quad (4b)$$

Through this method, one expresses the flux $\tilde{\mathbf{F}}$ and the collision operator $\tilde{\mathbf{C}}$ as a function of the unknown \mathbf{f} in (3b). This corresponds to approximating the s -dependencies of f in (1). Indeed, \tilde{f} can be interpreted as an approximation of f , the moments of which satisfy the same PDE system (3a) as the moments of f .

In this paper, we focus on the construction of the flux function $\tilde{\mathbf{F}}(\mathbf{f})$ and we only consider linear collision operator $\tilde{\mathbf{C}}(\mathbf{f})$. We especially focus on the positivity property of the reconstruction f , referred to as the realizability property at the moment level, for two reasons:

- One major property of the kinetic models of the form (1) is the positivity of the density f of particles in phase space, which is a density function. One commonly expects this property to be preserved through the moment extraction. This is the case *e.g.* for the well-established entropy-based closure M_N ([40, 41, 34, 38]) or for the atom-based closures K_N (based on the idea of [31] then developed in [42, 43] and exploited in [52]). Furthermore, when using such closures, the positivity of the underlying kinetic function needs to be preserved during the computations, for the fluxes $\tilde{\mathbf{F}}$ and the collision operator $\tilde{\mathbf{C}}$ to be well-defined.
- This positivity property is also of major importance for modelling purely anisotropic regimes, *i.e.* when modelling perfect beams of particles. In such a physical limit, the underlying kinetic distribution behaves as a narrow Gaussian or a Dirac peak (see previous work [46, 45, 36]). Non-realizable closures,

such as the polynomial P_N closure, *i.e.* when \tilde{f} is obtained from a spherical harmonics expansion of f , often misbehaves in such a limit. Such closures are generally overdifusive in this case, and one requires a high number of moments for such simulations to be accurate (*e.g.* order up to 21 in 1D and 17 in 3D in [32]). Instead, closures based on a positive reconstruction \tilde{f} capture these phenomena properly, even with low order moments (see *e.g.* [46, 36, 50]). Though, those closures are more difficult to construct and to compute numerically.

At the moment level, a vector \mathbf{V} is said to be realizable if there exists a positive reconstruction \tilde{f} satisfying the moment constraints (4a). The set of realizable vector is called the realizability domain and the problem of characterizing the existence of a positive reconstruction \tilde{f} is called "the Truncated Moment Problem" (TMP) after [56, 33]. Several TMP were solved, mostly when the set of integration is 1D. We may list the truncated moment problems of Hausdorff ($E = [-1, +1]$), Hamburger ($E = \mathbb{R}$; [1]), Stieltjes (\mathbb{R}^+ ; [56]), Toeplitz ($E = S^1$), see *e.g.* [15] for a review on those results.

The objective of this work is to provide a numerically tractable realizable closure (4b) for arbitrary high order moment models in 1D. We present a technique based on a projection on the boundary of the realizability domain in the direction of a chosen equilibrium function f_{eq} which provides such a solution to such 1D moment problems. This solution leads to an atomic closure, *i.e.* the reconstruction \tilde{f} is based on a sum of Dirac peaks called atoms ([15, 16, 17]) and of a regular integrable function. Such atomic decompositions was also used for the construction of K_N closure ([31, 42, 43, 52]) for moments over $[-1, +1]$, or of quadrature-based closures ([20, 59, 36]) for moments over \mathbb{R}^+ , which are also realizable and numerically tractable. However, the present closure differs from those closures as

- the underlying distribution possesses the minimum number of atoms.
- when considering the moments \mathbf{f}_{eq} of an equilibrium distribution f_{eq} for the original problem (1), the underlying distribution of the present closure retrieves exactly the equilibrium distribution f_{eq} (with K_N only its moment flux $\mathbf{F}(\mathbf{f}_{eq})$ is obtained).
- this closure's definition is general to all types of TMP, such as Stieltjes' or Hamburger's, for which the realizability property is enforced by the positivity of moment matrices.

In the following, the superscript $\tilde{\cdot}$ is dropped, and the TMP are understood in terms of integrable functions f as well as in terms of measures γ , *i.e.* replacing $f(s)ds$ by $d\gamma(s)$ in (2-4).

This paper is organized as follows. The next sections recalls definitions and preliminary results exploited in the rest of the paper, especially around the construction and the properties of the realizability domain. Section 3 provides the construction and the numerical computation of realizable closures, namely the Kershaw K_N closure and closure, afterward called Π_N , based on projections on the boundary of the realizability domain. These closures are tested and analyse on practical test cases emerging from the field of radiative transfer. The last section is devoted to conclusive remarks and perspectives of this work.

2. Preliminaries. The present section is devoted to set up the problems considered in the paper and to provide the basics of the theory and state-of-the-art solutions.

2.1. Definitions and notations. In the present work, we focus on polynomial moments and we classically use the monomial basis

$$\mathbf{b}(s) = \mathbf{b}_N(s) = (1, s, \dots, s^N)^T. \quad (5)$$

The monomial basis is used for its simplicity, though others can equally be used, *e.g.* the Legendre basis are often preferred for their orthogonality property. One may also extend the notions presented here, with non-polynomial basis functions as long as it satisfies the pseudo-Haar property (*i.e.* some functional linear independence; [35]). This is used *e.g.* to construct the partial moments method ([22, 21, 54, 51]).

We first give the following definitions.

Definition 2.1. • We denote $L^1(E)^+$ the set of the non-negative integrable functions over E and that are non-zero, in the sense $f \in L^1(E)^+$ iff

$$f \in L^1(E), \quad f \geq 0 \text{ a.e.}, \quad \text{and} \quad \exists(c, d) \in E^2, \quad c < d, \text{ s.t. } \text{essinf}_{s \in [c, d]} f(s) > 0. \quad (6a)$$

- We denote $L_N^1(E)^+$ the set of the non-negative functions which have finite moments over E up to order N , *i.e.* $f \in L_N^1(E)^+$ iff

$$f \in L^1(E)^+, \quad \text{s.t.} \quad s \mapsto s^i f(s) \in L^1(E) \quad \forall i = 0, \dots, N.$$

- A vector $\mathbf{V} \in \mathbb{R}^{N+1}$ is said to be realizable if it is the vector of moments of a positive function, *i.e.* if

$$\exists f \in L_N^1(E)^+, \quad \text{s.t.} \quad \int_E \mathbf{b}_N(s) f(s) ds = \mathbf{V}. \quad (6b)$$

- The set of all realizable vectors is called the realizability domain. It is defined by

$$\mathcal{R}_{\mathbf{b}} = \left\{ \int_E \mathbf{b}(s) f(s) ds, \quad f \in L_{\text{Card}(\mathbf{b})+1}^1(E)^+ \right\}. \quad (6c)$$

- We will also use extensively the closure set of $\mathcal{R}_{\mathbf{b}}$ in $\mathbb{R}^{\text{Card}(\mathbf{b})}$

$$\mathcal{R}_{\mathbf{b}}^m = \overline{\mathcal{R}_{\mathbf{b}}} \cap \mathbb{R}^{\text{Card}(\mathbf{b})}. \quad (6d)$$

The superscript m refers to "measure" as, according to the definitions (6a-6d), any element of this set is the moment vector of a positive measure γ in the sense

$$\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m \quad \Rightarrow \quad \exists \gamma \in \mathcal{M}(E), \quad \mathbf{V} = \int_E \mathbf{b}(s) d\gamma(s). \quad (6e)$$

For notation purposes, we also use extensively the following function.

Definition 2.2 (Riesz functional).

Consider a vector $\mathbf{b} \in (\mathbb{R}[X])^N$ of N polynomials, and a vector $\mathbf{V} \in \mathbb{R}^N$. The Riesz functional $R_{\mathbf{V}}$ associated to \mathbf{V} sends any polynomial $p = \lambda \mathbf{b}$ onto

$$R_{\mathbf{V}}(p) = \lambda \mathbf{V}. \quad (7)$$

Remark that the Riesz functional associated to \mathbf{V} is a linear map from $\text{Span}(\mathbf{b})$ to \mathbb{R} .

If the vector $\mathbf{V} = \int_E \mathbf{b}(s) f(s) ds$ is the vector of moments of a function $f \in L_{\text{Card}(\mathbf{b})+1}^1(E)^+$, then the Riesz functional of p is the moment of f according to p

$$R_{\mathbf{V}}(p) = \int_E p(s) f(s) ds.$$

In the next sections, the Riesz functional is also applied componentwise to matrices of polynomials

$$R_{\mathbf{V}}(M)_{i,j} = R_{\mathbf{V}}(M_{i,j}).$$

Example: Consider the vector $\mathbf{f} = (f^0, f^1, f^2) \in \mathbb{R}^3$, and the vector of monomials $\mathbf{b}(s) = (1, s, s^2)$. The Riesz function according to the vector \mathbf{f} of the polynomial

$$p(s) = 1 + 3s - s^2$$

reads

$$\begin{aligned} R_{\mathbf{f}}(p) &= R_{\mathbf{f}}(1) + 3R_{\mathbf{f}}(s) - R_{\mathbf{f}}(s^2) \\ &= f^0 + 3f^1 - f^2. \end{aligned}$$

The first problem studied in this paper is the truncated moment problem in 1D

$$\text{Find } \gamma \in \mathcal{M}(E), \quad \text{s.t.} \quad \mathbf{V} = \int_E \mathbf{b}(s) d\gamma(s). \quad (8)$$

2.2. Properties of the realizability domain. In the following sections, we widely exploit the following results.

Proposition 1. *The realizability domain $\mathcal{R}_{\mathbf{b}}$ is an open convex cone.*

Proof. The set $\mathcal{R}_{\mathbf{b}}$ is a convex cone because of the linearity of the integral.

To prove that it is open, for all $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}$, we exhibit a neighborhood of \mathbf{V} included in $\mathcal{R}_{\mathbf{b}}$. Write

$$\mathbf{V} = \int_E \mathbf{b}(s) f(s) ds \in \mathcal{R}_{\mathbf{b}}, \quad \varepsilon := \text{essinf}_{s \in [c,d]} f(s) > 0.$$

Define

$$M_0 = \int_c^d \mathbf{b}\mathbf{b}^T(s) ds.$$

By assumption, for all i , the function $\mathbf{b}_i^2 \neq 0$ on $[a, b]$. Thus, M_0 is symmetric positive definite, and especially non-singular. Then, the family $(\mathbf{V}_i^0)_{\{i=1, \dots, \text{Card}(\mathbf{b})\}}$ of its column is a basis of $\mathbb{R}^{\text{Card}(\mathbf{b})}$. Therefore, for all $\alpha > 0$, the set

$$\mathcal{V} = \left\{ \mathbf{V} + \alpha \sum_i \lambda_i \mathbf{V}_i^0, \quad \lambda_i \in]-1, 1[, \quad i = 1, \dots, \text{Card}(\mathbf{b}) \right\} \quad (9)$$

is a neighborhood of \mathbf{V} in $\mathbb{R}^{\text{Card}(\mathbf{b})}$. Now, chose a coefficient α such that

$$0 < \alpha < \frac{\varepsilon}{\sum_i \|\mathbf{b}_i\|_{\infty, [c,d]}}. \quad (10)$$

With this choice of α , one shows that

$$\left\| \alpha \sum_i \lambda_i \mathbf{b}_i \right\|_{\infty, [c,d]} < \varepsilon, \quad \forall \lambda_i \in]-1, 1[, \quad i = 1, \dots, \text{Card}(\mathbf{b}).$$

Thus, any vector in \mathcal{V} is realized by a function of $f + \alpha \mathbf{1}_{[c,d]} \sum_i \lambda_i \mathbf{b}_i \in L^1_{\text{Card}(\mathbf{b})+1}(E)^+$, thus $\mathcal{V} \subset \mathcal{R}_{\mathbf{b}}$. \square

Remark 1. • This also provides that $\mathcal{R}_{\mathbf{b}}^m$ is a convex cone, and that $\mathcal{R}_{\mathbf{b}} = \text{int}(\mathcal{R}_{\mathbf{b}}^m)$.

- This property is commonly used when constructing numerical schemes for moment equations in order to prove that such schemes preserve the realizability property from one step to another (see *e.g.* [2, 47, 48]).

2.3. Characterizations of the realizability domain. In this subsection, we recall the well-known Hausdorff, Hamburger and Stieltjes moment problems ([1, 15]) which provides characterizations of the realizability property.

Theorem 2.3 (Hausdorff). *Suppose that $\mathbf{b} = \mathbf{b}_{2K}$ (even case) or $\mathbf{b} = \mathbf{b}_{2K+1}$ (odd case). Then a vector $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m$ is realizable iff the matrices*

$$\text{Even case : } \quad M_{\mathbf{V}}^1 = R_{\mathbf{V}}(\mathbf{b}_K \mathbf{b}_K^T), \quad M_{\mathbf{V}}^2 = R_{\mathbf{V}}((1-s^2)\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T), \quad (11a)$$

$$\text{Odd case : } \quad M_{\mathbf{V}}^1 = R_{\mathbf{V}}((1+s)\mathbf{b}_K \mathbf{b}_K^T), \quad M_{\mathbf{V}}^2 = R_{\mathbf{V}}((1-s)\mathbf{b}_K \mathbf{b}_K^T). \quad (11b)$$

are positive semi-definite.

Theorem 2.4 (Hamburger). *Suppose that $\mathbf{b} = \mathbf{b}_{2K}$ (even case) or $\mathbf{b} = \mathbf{b}_{2K+1}$ (odd case). Then a vector $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m$ is realizable iff*

$$\text{Even case : } \quad M_{\mathbf{V}}^1 = R_{\mathbf{V}}(\mathbf{b}_K \mathbf{b}_K^T) \text{ is positive semi-definite,} \quad (12a)$$

$$\text{Odd case : } \quad M_{\mathbf{V}}^1 = R_{\mathbf{V}}(\mathbf{b}_K \mathbf{b}_K^T) \text{ is positive semi-definite} \quad (12b)$$

$$\text{and } \quad R_{\mathbf{V}}(s^{K+1}\mathbf{b}_K^T) \in \text{Im}(M_{\mathbf{V}}^1).$$

Theorem 2.5 (Stieltjes). *Suppose that $\mathbf{b} = \mathbf{b}_{2K}$ (even case) or $\mathbf{b} = \mathbf{b}_{2K+1}$ (odd case). Then a vector $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m$ is realizable iff the matrices*

$$\text{Even case : } \quad M_{\mathbf{V}}^1 = R_{\mathbf{V}}(\mathbf{b}_K \mathbf{b}_K^T), \quad M_{\mathbf{V}}^2 = R_{\mathbf{V}}(s\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T), \quad (13a)$$

$$\text{Odd case : } \quad M_{\mathbf{V}}^1 = R_{\mathbf{V}}(\mathbf{b}_K \mathbf{b}_K^T) \quad M_{\mathbf{V}}^2 = R_{\mathbf{V}}(s\mathbf{b}_K \mathbf{b}_K^T). \quad (13b)$$

are positive semi-definite.

Proofs of these theorems can be found *e.g.* in [1, 15, 33]. For the construction of the closure, we widely exploit the following description of the boundary of the realizability domain ([19]).

Proposition 2. *For all vector $\mathbf{V} \in \partial\mathcal{R}_{\mathbf{b}}^m$ on the boundary of the realizability domain, there exists a unique representing measure for \mathbf{V} . This measure is given by*

$$\gamma = \sum_{i=1}^J \alpha_i \delta_{s_i}, \quad (14)$$

where the coefficients $(\alpha_j)_{j=1,\dots,J} \in (\mathbb{R}^+)^J$ are positive, the positions $(s_j)_{j=1,\dots,J} \in E^J$ and the number of atoms $0 \leq J \leq K$ is

$$J = \min_i (\text{rank}(M_{\mathbf{V}}^i)),$$

where the matrices $M_{\mathbf{V}}^i$ are given either by (11), (12) or (13) depending on the considered problem.

This so-called atomic decomposition ([15]), in 1D, can be deduced as a corollary of the truncated Riesz-Haviland theorem, see *e.g.* [49, 27, 28, 15, 17] or surveys in [19, 33], or of Tchakaloff theorem, see *e.g.* [57, 5, 16, 33].

3. Realizable closures in 1D. We present in this section a strategy to construct a realizable closure for 1D problems. In the following section, we will focus on moments over $s \in [-1, +1]$ and compare such realizable closures to the linear P_N closure.

3.1. Computation of the closure: Decomposition of moment vectors. In the following, we will assume that the collision operator is a simple linear function of the unknown, and we will only focus on the construction of the flux vector \mathbf{F} as a function of a vector of moments \mathbf{f} .

3.1.1. Construction of the closure. The main idea to construct a closure, realizable or not, is to reconstruct from a vector \mathbf{f} a representing measure γ , *i.e.* satisfying

$$\int_E \mathbf{b}(s) d\gamma(s) = \mathbf{f}. \quad (15)$$

Once this reconstruction γ is found, one simply construct the flux vector as the moment flux associated to this representing measure γ , *i.e.* one defines

$$\mathbf{F}(\mathbf{f}) = \int_E s \mathbf{b}(s) d\gamma(s),$$

where γ satisfies (15).

As we only consider monomial basis \mathbf{b}_N defined in (5), one observes that all but the last coefficient of the flux \mathbf{F} already belong to the vector \mathbf{f} . In practice, we denote κ the closure, *i.e.* this only unknown coefficient that can be defined as

$$\kappa(\mathbf{f}) = \int_E s^{N+1} d\gamma(s).$$

Definition 3.1. A function $\kappa : \mathbb{R}^{N+1} \rightarrow \mathbb{R}$ is called a realizable closure iff the vector function $(Id_{\mathbb{R}^{N+1}}, \kappa)$ sends $\mathcal{R}_{\mathbf{b}_N}$ into $\mathcal{R}_{\mathbf{b}_{N+1}}$.

This definition equivals to defining κ from a positive measure γ satisfying the moment constraints (15).

In order to obtain appropriate descriptions of certain physical phenomena, we need to impose the value of the closure and its representing measure when the moment vector corresponds to the moment of specific measures. In practice, we focus on two types of measure that we aim to retrieve when reconstructing the measure γ :

- some equilibrium function f_{eq} for the PDE (1). Then, when

$$\mathbf{f} = \mathbf{V}_{eq} = \int_E \mathbf{b}(s) f_{eq}(s) ds,$$

we aim to construct underlying measure γ and a closure such that

$$d\gamma(s) = f_{eq}(s) ds, \quad \kappa = \int_E s^{N+1} f_{eq}(s) ds.$$

- some purely anisotropic measures, *i.e.* Dirac measures δ_{s_i} in some locations s_i , and potentially sums of such measures. Then, when

$$\mathbf{f} = \mathbf{b}(s_i),$$

we aim to construct underlying measure γ and a closure such that

$$d\gamma(s) = \delta_{s_i}(s), \quad \kappa = s_i^{N+1}.$$

Since δ_{s_i} is the only positive measure representing $\mathbf{b}(s_i)$. This implies that all realizable closures capture these purely anisotropic measure.

3.1.2. Decomposition of the moment vector. The strategy for constructing realizable closures consists in decomposing the known vector of moments $\mathbf{V} \in \mathcal{R}_{\mathbf{b}_N}$ into different realizable parts

$$\mathbf{V} = \sum_i \mathbf{V}_i \quad \text{with} \quad \mathbf{V}_i \in \mathcal{R}_{\mathbf{b}_N},$$

for which a realizable closure $\kappa_i(\mathbf{V}_i)$ is known, or can be computed. Then, one simply define the closure

$$\kappa(\mathbf{V}) = \sum_i \kappa_i(\mathbf{V}_i),$$

which is realizable since the realizability domains are convex cones, *i.e.* we construct $(\mathbf{V}_i, \kappa_i) \in \mathcal{R}_{\mathbf{b}_{N+1}}^m$ then

$$\sum_i (\mathbf{V}_i, \kappa_i(\mathbf{V}_i)) = \left(\sum_i \mathbf{V}_i, \sum_i \kappa_i(\mathbf{V}_i) \right) = (\mathbf{V}, \kappa(\mathbf{V})) \in \mathcal{R}_{\mathbf{b}_{N+1}}^m.$$

For such decompositions, we will exploit two types of vectors $\mathbf{V} \in \mathcal{R}_{\mathbf{b}_N}$ for which we know easily how to construct realizable closures. These are simply the ones described in the previous paragraph that we aim to capture exactly:

- Realizable moments $\mathbf{V}_{eq} \in \mathcal{R}_{\mathbf{b}_N}$ of some **known given** function $f_{eq} \in L_N^1(E)^+$

$$\mathbf{V}_{eq} = \int_E \mathbf{b}_N(s) f_{eq}(s) ds.$$

In practice, we chose a function f_{eq} corresponding to an equilibrium of the PDE (1) we aim to solve. Here the designated closure of such a vector is simply the $N + 1$ -th moment of f_{eq} as this function is data

$$\kappa_{eq} = \int_E s^{N+1} f_{eq}(s) ds.$$

- Moments $\mathbf{V}_s \in \partial \mathcal{R}_{\mathbf{b}_N}^m$ **on the boundary of the realizability domain.** Using Proposition 2, we have the uniqueness of a representing measure γ_s for \mathbf{V}_s which is singular over E , *i.e.*

$$\gamma_s = \sum_{i=1}^J \alpha_i \delta_{s_i} \in \mathcal{M}(E), \quad \mathbf{V}_s = \int_E \mathbf{b}_N(s) d\gamma_s(s) = \sum_{i=1}^J \alpha_i \delta_{s_i}.$$

Thus, there exists a unique realizable closure for such vector

$$\kappa_s(\mathbf{V}_s) = \int_E s^{N+1} d\gamma_s(s) = \sum_i \alpha_i s_i^{N+1}. \quad (16)$$

In the following, we will focus on two realizable closures based on such decomposition methods.

3.2. Kershaw K_N closure for Hausdorff problem. For completeness, we recall the construction of K_N closure ([31, 42, 52]) for Hausdorff problem $E = [-1, +1]$.

Decomposition of \mathbf{V} : Here, we decompose a vector $\mathbf{V} \in \mathcal{R}_{\mathbf{b}_N}$ into two parts

$$\mathbf{V} = \alpha \mathbf{V}_1 + (1 - \alpha) \mathbf{V}_2, \quad \mathbf{V}_1 = \mathbf{V} = \mathbf{V}_2,$$

the α coefficient will be explicated below. The closure becomes

$$\kappa_{K_N}(\mathbf{V}) = \alpha \kappa_1(\mathbf{V}) + (1 - \alpha) \kappa_2(\mathbf{V}),$$

and we use different closures κ_1 for \mathbf{V}_1 and κ_2 for \mathbf{V}_2 .

Closures for \mathbf{V}_1 and \mathbf{V}_2 : Both closures are constructed such that $(\mathbf{V}_i, \kappa_i) \in \partial \mathcal{R}_{\mathbf{b}_{N+1}}$ are on the boundary of the realizability domain. In practice, using the notations (11-13), we chose these closures $\kappa_i(\mathbf{V}_i)$ such that

$$M_{(\mathbf{V}_i, \kappa_i)}^i \quad \text{is positive semi-definite and singular.}$$

The numerical computation of these closures is detailed in Subsection 3.4 below.

Computation of κ_{K_N} : Finally, the coefficients $\alpha \in [0, 1]$ in the convex combination is chosen in order to recover the closure κ_{eq} of some given equilibrium state \mathbf{V}_{eq} , *i.e.*

$$\mathbf{V}_{eq} = \int_{-1}^{+1} \mathbf{b}_N(s) f_{eq}(s) ds, \quad \kappa_{eq} = \int_{-1}^{+1} s^{N+1} f_{eq}(s) ds,$$

typically for the given function $f_{eq} = 1$. Thus α is defined by

$$\kappa_{K_N}(\mathbf{V}_{eq}) = \kappa_{eq} \quad \Rightarrow \quad \alpha = \frac{\kappa_{eq} - \kappa_2(\mathbf{V}_{eq})}{\kappa_1(\mathbf{V}_{eq}) - \kappa_2(\mathbf{V}_{eq})}.$$

Remark 2. In practice, Kershaw method recover the exact value of the closure $\kappa_{K_N} = \kappa_{eq}$ when $\mathbf{f} = \mathbf{V}_{eq}$, but it does not retrieve the exact value of the representing measure $d\gamma_{K_N}(s) \neq f_{eq}(s) ds$.

Representing measure γ_{K_N} : In practice, with such a K_N model, the kinetic distribution f is approximated by a measure γ_{K_N} of the form

$$\gamma_{K_N} = \gamma_1 \alpha + \gamma_2 (1 - \alpha), \quad \gamma_1 = \sum_{i=1}^J \alpha_i \delta_{s_i}, \quad \gamma_2 = \sum_{i=J+1}^{2J} \alpha_i \delta_{s_i}, \quad (17)$$

where γ_1 and γ_2 are the unique representing measure for $(\mathbf{V}, \kappa_1) \in \partial \mathcal{R}_{\mathbf{b}_{N+1}}^m$ and $(\mathbf{V}, \kappa_2) \in \partial \mathcal{R}_{\mathbf{b}_{N+1}}^m$. Following Proposition 2, the number of Diracs in this K_N representation can reach at most $2J = 2 \min_i \text{rank}(M_{\mathbf{V}}^i)$, *i.e.* $2K = N$ if N is even or $2K = N - 1$ if N is odd. Remark that the number of Diracs in this decomposition, so-called atoms in [15, 17, 16], is not minimized.

A numerical method to compute the coefficient α_i and the positions s_i is given below in Section 3.5.2.

Remark 3. This construction only holds for moments over $[-1, +1]$, because it requires two extensions of a realizable moment vector of size N onto the boundary of the realizability domain for moments of order $N + 1$. And such extensions are unique for Hamburger ($E = \mathbb{R}$) or Stieltjes ($E = \mathbb{R}^+$) problems.

3.3. Projective Π_N closure. The construction of the present projection closure, afterward called Π_N closure for projection (P_N being already taken for polynomial closure), is based on a decomposition of any realizable vector into a regular part \mathbf{V}_{eq} , *i.e.* moments of a given regular function

$$\mathbf{V}_{eq} = \int_E \mathbf{b}_N(s) f_{eq}(s) ds,$$

and a part \mathbf{V}_s on the boundary of the realizability domain, *i.e.* moments of a singular measure $\gamma_s = \sum_i \alpha_i \delta_{s_i}$ according to Proposition 2.

Decomposition of \mathbf{V} : Here, we decompose

$$\mathbf{V} = \bar{x}\mathbf{V}_{eq} + \mathbf{V}_s, \quad (18)$$

and we construct the closure as

$$\kappa_{\Pi_N}(\mathbf{V}) = \bar{x}(\mathbf{V})\kappa_{eq} + \kappa_s(\mathbf{V} - \bar{x}(\mathbf{V})\mathbf{V}_{eq}),$$

where $\kappa_{eq} = \int_E s^{N+1} f_{eq}(s) ds$ is a given value.

Computation of the closures: In practice, we follow two steps the computations of which will be detailed in the next subsections:

1. **Regular part:** The closure of the regular part κ_{eq} is already known, only the multiplicative coefficient \bar{x} needs to be computed. Find a maximum scalar $\bar{x} \geq 0$ s.t.

$$\mathbf{V} - \bar{x}\mathbf{V}_{eq} \in \partial\mathcal{R}_{\mathbf{b}}^m.$$

This part is computed in Subsection 3.4.

2. **Singular part:** Compute the unique realizable closure for the singular part

$$\kappa_s(\mathbf{V} - \bar{x}(\mathbf{V})\mathbf{V}_{eq}).$$

This part is computed in Subsection 3.5.

Remark 4. This construction also generalizes out of the realizability domain. One would only need to seek for a $\bar{x} < 0$. In the result below, we prove that there exists a unique $\bar{x} > 0$ satisfying the second step. This only holds if $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}$ is in the interior of the realizability domain.

Representing measure γ_{Π_N} : In practice, with such a Π_N model, the kinetic distribution f is approximated by a measure γ_{Π_N} of the form

$$\gamma_{\Pi_N}(s) = \bar{x}f_{eq}(s)ds + \gamma_0, \quad \gamma_0 = \sum_{i=1}^J \alpha_i \delta_{s_i}, \quad (19)$$

where γ_0 is the unique representing measure for $\mathbf{V} - \bar{x}\mathbf{V}_{eq} \in \partial\mathcal{R}_{\mathbf{b}_N}^m$. Following Proposition 2, the number of Diracs in this K_N representation can reach at most $J = \min_i \text{rank}(M_{\mathbf{V} - \bar{x}\mathbf{V}_{eq}}^i)$, *i.e.* $K - 1 = \frac{N}{2} - 1$ if N is even or $K - 1 = \frac{N-1}{2} - 1$ if N is odd. This number depends on the choice of f_{eq} but is inferior to the one in (17) for K_N closure.

3.4. Projection on the boundary $\partial\mathcal{R}_{\mathbf{b}}$: Computation of the regular part.

From a chosen regular distribution $f_{eq}(s)$, we construct a projection of any realizable vector $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m$ onto the boundary of the realizability domain. This is simply performed by removing from \mathbf{V} the moments of f_{eq} . For this purpose, we define

$$\mathbf{V}_{eq} := \int_E \mathbf{b}(s)f_{eq}(s)ds.$$

We first exhibit the uniqueness of a decomposition of any realizable vector into the sum of $\mathbf{V}_{eq} \in \mathcal{R}_{\mathbf{b}}$ and a vector on the boundary $\mathbf{V}_s \in \partial\mathcal{R}_{\mathbf{b}}^m$.

Proposition 3. *For all measure realizable vectors $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m$, there exists a unique decomposition of the form*

$$\mathbf{V} = \bar{x}\mathbf{V}_{eq} + \mathbf{V}_s. \quad (20a)$$

where $\mathbf{V}_s \in \partial\mathcal{R}_{\mathbf{b}}^m$ and $\bar{x} \in \mathbb{R}^+$. These parameters are given by

$$\tilde{M}_{\mathbf{V}}^i := (M_{\mathbf{V}_{eq}}^i)^{-1/2} M_{\mathbf{V}}^i \left((M_{\mathbf{V}_{eq}}^i)^{-1/2} \right)^T, \quad (20b)$$

$$\bar{x} = \min_i \min Sp \left(\tilde{M}_{\mathbf{V}}^i \right), \quad \mathbf{V}_s = \mathbf{V} - \bar{x} \mathbf{V}_{eq}, \quad (20c)$$

where $M^{-1/2}$ is positive definite such that $M^{-1/2} M (M^{-1/2})^T = I$ (use e.g. Cholesky decomposition of M), and the matrices $M_{\mathbf{V}}^i$ are defined in (11), (12) or (13) depending on the considered problem.

Proof. We perform the computations in the case of even order moment over $[-1, +1]$ but the method can be generalized to all the other cases. We first prove the existence and uniqueness of the decomposition, then we compute \bar{x} and \mathbf{V}_s .

Existence of the decomposition:

As $\mathbf{V}_{eq} \in \mathcal{R}_{\mathbf{b}}^m$, then its moment matrices $M_{\mathbf{V}}^i$ are symmetric positive definite.

First, we consider the case $\mathbf{V} \in \partial\mathcal{R}_{\mathbf{b}}^m$. Then $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m$. Furthermore, the associated moment matrices (11) are positive semi-definite and one is singular. Write $M_{\mathbf{V}}$ the singular matrix and its eigenvector e_0 associate to the eigenvalue 0.

As the moment matrices of \mathbf{V}_{eq} are positive definite, then $e_0^T M_{\mathbf{V}_{eq}} e_0 > 0$. One obtains

$$\forall y > 0, \quad e_0^T M_{\mathbf{V} - y \mathbf{V}_{eq}} e_0 = -y e_0^T M_{\mathbf{V}_{eq}} e_0 < 0.$$

Thus, for all $y > 0$, we have $\mathbf{V} - y \mathbf{V}_{eq} \notin \mathcal{R}_{\mathbf{b}}^m$, and $\bar{x} = 0$ is the unique scalar of \mathbb{R}^+ such that $\mathbf{V} - \bar{x} \mathbf{V}_{eq} \in \partial\mathcal{R}_{\mathbf{b}}^m$.

Second, we consider the case $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}$. Then its moment matrices (11) are positive definite. Defining

$$\mathbf{W}(y) = \mathbf{V} - y \mathbf{V}_{eq},$$

we have $\mathbf{W}(0) \in \mathcal{R}_{\mathbf{b}}$. One observes that

$$\forall y > z := \frac{\max_i \max(Sp(M_{\mathbf{V}}^i))}{\min_i \min(Sp(M_{\mathbf{V}_{eq}}^i))}, \quad M_{\mathbf{W}(y)}^i \text{ is symmetric negative definite,}$$

then especially $-\mathbf{W}(z) \in \mathcal{R}_{\mathbf{b}}^m$. As $y \mapsto \mathbf{W}(y)$ is a linear (continuous) function of y , then there exists a point $\bar{x} \in]0, z]$ such that $\mathbf{W}(\bar{x}) \in \partial\mathcal{R}_{\mathbf{b}}^m$.

Uniqueness of the decomposition:

Finally, suppose that there exists two points $0 < x_1 < x_2$ such that $\mathbf{W}(x_1) \in \partial\mathcal{R}_{\mathbf{b}}^m$ and $\mathbf{W}(x_2) \in \partial\mathcal{R}_{\mathbf{b}}^m$. By convexity of $\mathcal{R}_{\mathbf{b}}^m$, we have $\mathbf{W}(x) \in \mathcal{R}_{\mathbf{b}}^m$ for all $x \in [x_1, x_2]$. However,

$$\begin{aligned} \mathbf{W}(x) &= \mathbf{W}(x_2)(1 - \alpha) + \mathbf{W}(x_1)\alpha \\ &= \mathbf{W}(x_2) + \alpha(x_2 - x_1) \mathbf{V}_{eq}. \end{aligned}$$

As $\alpha(x_2 - x_1) > 0$, then $\mathbf{W}(x)$ is a positive combination of $\mathbf{W}(x_2) \in \partial\mathcal{R}_{\mathbf{b}}^m$ and of $\mathbf{V}_{eq} \in \mathcal{R}_{\mathbf{b}} = \text{int}(\mathcal{R}_{\mathbf{b}}^m)$, then $\mathbf{W}(x) \in \mathcal{R}_{\mathbf{b}}$ which contradicts with $\mathbf{W}(x_1) \in \partial\mathcal{R}_{\mathbf{b}}^m$. Thus $\bar{x} > 0$ is unique.

Computation of \bar{x} and \mathbf{V}_s :

Remark that $M_{\mathbf{V}_{eq}}^i$ is symmetric positive definite, then invertible and we may define $(M_{\mathbf{V}_{eq}}^i)^{-1/2}$ using its diagonalization. Based on the characterizations of Theorem 2.5, and on the construction (20b), we observe that

$$\begin{aligned} \mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m &\Leftrightarrow M_{\mathbf{V}}^i \text{ symmetric positive semi-definite} \\ &\Leftrightarrow \tilde{M}_{\mathbf{V}}^i \text{ symmetric positive semi-definite.} \end{aligned}$$

Then, exploiting the linearity of the Riesz functional, we obtain

$$\tilde{M}_{\mathbf{W}(y)}^i = \tilde{M}_{\mathbf{V}}^i - yI.$$

Through this formula, considering that $\tilde{M}_{\mathbf{W}(\bar{x})}^i$ is positive semi-definite and singular, we obtain that

$$\bar{x} = \min_i \min Sp(\tilde{M}_{\mathbf{V}}^i).$$

□

In all the following, the coefficient \bar{x} is defined as (20c) where $(M_{\mathbf{V}_{eq}}^i)^{-1/2}$ is the inverse of the Cholesky decomposition of $M_{\mathbf{V}_{eq}}^i$. Remark that since $\mathbf{V}_{eq} \in \mathcal{R}_{\mathbf{b}}$ is in the interior of the realizability domain, then the matrices $M_{\mathbf{V}_{eq}}^i$ are symmetric positive definite and their Cholesky decomposition exist.

3.5. Computation of the closure on the boundary $\partial\mathcal{R}_{\mathbf{b}}^m$: Computation of the singular part. Naively, one could compute all the coefficients α_i and s_i , and then reinject it in (16) to obtain κ_s . However, this would not be very efficient from a numerical point of view.

3.5.1. Computation of the closure κ_s . Instead, as we only look for one higher order moment, it is cheaper to exploit the principle of flat extension proposed by [15, 19], which was also implicitly exploited in the construction of Kershaw K_N closure ([31, 42, 52]). This consists in remarking that the set of moments $\mathbf{f} - \bar{x}\mathbf{f}_{eq}$ is recursively generated, *i.e.* that it is generated by $J < N$ Diracs. If \mathbf{V}_s is generated by J Diracs, so is the extended vector (\mathbf{V}_s, κ_s) and therefore the matrices $M_{(\mathbf{V}_s, \kappa_s)}^i$ are also singular. Remarking that the moment of order $N + 1$, *i.e.* $f^{N+1} = \kappa_s$, appears only once in the matrix in the last diagonal entry. Then, we may decompose: **Hausdorff even case $N = 2K$:**

$$R_{(\mathbf{V}_s, \kappa_s)}((1+s)\mathbf{b}_K \mathbf{b}_K^T) = \begin{pmatrix} R_{\mathbf{V}_s}((1+s)\mathbf{b}_{K-1} \mathbf{b}_{K-1}^T) & R_{\mathbf{V}_s}((1+s)s^K \mathbf{b}_{K-1}) \\ R_{\mathbf{V}_s}((1+s)s^K \mathbf{b}_{K-1})^T & R_{\mathbf{V}_s}(s^{2K}) + \kappa_s \end{pmatrix}, \quad (21a)$$

$$R_{(\mathbf{V}_s, \kappa_s)}((1-s)\mathbf{b}_K \mathbf{b}_K^T) = \begin{pmatrix} R_{\mathbf{V}_s}((1-s)\mathbf{b}_{K-1} \mathbf{b}_{K-1}^T) & R_{\mathbf{V}_s}((1-s)s^K \mathbf{b}_{K-1}) \\ R_{\mathbf{V}_s}((1-s)s^K \mathbf{b}_{K-1})^T & R_{\mathbf{V}_s}(s^{2K}) - \kappa_s \end{pmatrix}. \quad (21b)$$

According to the previous decomposition, these matrices are singular and the last column is in the Span of the others. This provides

$$\kappa_s = -R_{\mathbf{V}_s}(s^{2K}) + R_{\mathbf{V}_s}((1+s)s^K \mathbf{b}_{K-1})^T R_{\mathbf{V}_s}((1+s)\mathbf{b}_{K-1} \mathbf{b}_{K-1}^T)^+ R_{\mathbf{V}_s}((1+s)s^K \mathbf{b}_{K-1}) \quad (22a)$$

$$= R_{\mathbf{V}_s}(s^{2K}) - R_{\mathbf{V}_s}((1-s)s^K \mathbf{b}_{K-1})^T R_{\mathbf{V}_s}((1-s)\mathbf{b}_{K-1} \mathbf{b}_{K-1}^T)^+ R_{\mathbf{V}_s}((1-s)s^K \mathbf{b}_{K-1}), \quad (22b)$$

where the superscript $+$ refers to pseudo-inverse, and it can be computed by standard methods.

Hausdorff odd case $N = 2K + 1$:

$$R_{(\mathbf{V}_s, \kappa_s)}(\mathbf{b}_{K+1} \mathbf{b}_{K+1}^T) = \begin{pmatrix} R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T) & R_{\mathbf{V}_s}(s^{K+1} \mathbf{b}_K) \\ R_{\mathbf{V}_s}(s^{K+1} \mathbf{b}_K)^T & \kappa_s \end{pmatrix},$$

$$R_{(\mathbf{V}_s, \kappa_s)}((1-s^2)\mathbf{b}_K \mathbf{b}_K^T) = \begin{pmatrix} R_{\mathbf{V}_s}((1-s^2)\mathbf{b}_{K-1} \mathbf{b}_{K-1}^T) & R_{\mathbf{V}_s}((1-s^2)s^K \mathbf{b}_{K-1}) \\ R_{\mathbf{V}_s}((1-s^2)s^K \mathbf{b}_{K-1})^T & R_{\mathbf{V}_s}(s^{2K}) - \kappa_s \end{pmatrix}.$$

which leads to

$$\begin{aligned}\kappa_s &= R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K)^T R_{\mathbf{V}_s}(\mathbf{b}_K\mathbf{b}_K^T)^+ R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K) \\ &= R_{\mathbf{V}_s}(s^{2K}) - R_{\mathbf{V}_s}((1-s^2)s^K\mathbf{b}_{K-1})^T R_{\mathbf{V}_s}((1-s^2)\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T)^+ R_{\mathbf{V}_s}((1-s^2)s^K\mathbf{b}_{K-1}),\end{aligned}$$

Hamburger even case $N = 2K$: The vector

$$R_{(\mathbf{V}_s, \kappa_s)}(s^{K+1}\mathbf{b}_K^T) \in \text{Im}(R_{\mathbf{V}_s}(\mathbf{b}_K\mathbf{b}_K^T)),$$

where the matrix on the RHS is singular. This can be rewritten

$$(R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_{K-1})^T, \kappa_s)^T \in \text{Im} \begin{pmatrix} R_{\mathbf{V}_s}(\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T) & R_{\mathbf{V}_s}(s^K\mathbf{b}_{K-1}) \\ R_{\mathbf{V}_s}(s^K\mathbf{b}_{K-1})^T & R_{\mathbf{V}_s}(s^{2K}) \end{pmatrix}.$$

This leads to

$$\kappa_s = R_{\mathbf{V}_s}(s^K\mathbf{b}_{K-1})^T R_{\mathbf{V}_s}(\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T)^+ R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_{K-1})$$

Hamburger odd case $N = 2K + 1$: The matrix

$$R_{(\mathbf{V}_s, \kappa_s)}(\mathbf{b}_{K+1}\mathbf{b}_{K+1}^T) = \begin{pmatrix} R_{\mathbf{V}_s}(\mathbf{b}_K\mathbf{b}_K^T) & R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K) \\ R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K)^T & \kappa_s \end{pmatrix}.$$

is singular. This leads to

$$\kappa_s = R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K)^T R_{\mathbf{V}_s}(\mathbf{b}_K\mathbf{b}_K^T)^+ R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K).$$

Stieltjes even case $N = 2K$: The matrix

$$R_{(\mathbf{V}_s, \kappa_s)}(s\mathbf{b}_K\mathbf{b}_K^T) = \begin{pmatrix} R_{\mathbf{V}_s}(s\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T) & R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_{K-1}) \\ R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_{K-1})^T & \kappa_s \end{pmatrix}.$$

is singular. This leads to

$$\kappa_s = R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_{K-1})^T R_{\mathbf{V}_s}(s\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T)^+ R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_{K-1}).$$

Stieltjes odd case $N = 2K + 1$: The matrix

$$R_{(\mathbf{V}_s, \kappa_s)}(\mathbf{b}_{K+1}\mathbf{b}_{K+1}^T) = \begin{pmatrix} R_{\mathbf{V}_s}(\mathbf{b}_K\mathbf{b}_K^T) & R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K) \\ R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K)^T & \kappa_s \end{pmatrix}.$$

is singular. This leads to

$$\kappa_s = R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K)^T R_{\mathbf{V}_s}(\mathbf{b}_K\mathbf{b}_K^T)^+ R_{\mathbf{V}_s}(s^{K+1}\mathbf{b}_K).$$

Remark 5. These computations can be the source of roundoff errors which can be amplified. In order to smooth them out, we propose the following tricks.

- These computations are all based on the computations of pseudo-inverses M^+ of matrices M which are expected to be singular by construction. In practice, we simply use a basic QR decomposition for such pseudo-inverse. However, as roundoff errors may occur in the construction of those matrices, we filter the lowest eigenvalues of R in the QR decomposition below a certain threshold ($10^{-10} \max_{i,j} R_{i,j}$ in the applications below).
- Similarly, in the Hausdorff case, the closure κ_s is obtained equivalently by two formula (22), in the applications below, again to smooth roundoff errors, we use a convex combinations of the two definitions with a parameter based on the determinant of the non-zero part of the R in the QR decompositions of $R_{\mathbf{V}_s}((1 \pm s)\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T)$.

3.5.2. *Computation of the representing measure γ_s .* Even if its computation is more expensive, it remains possible to compute numerically the coefficients α_i and the positions s_i in the definition of the closure. These computations are also exploited in the numerical experiments in Section 4.

Computation of the positions s_i . One first needs to compute the positions $s_i \in E$. For this purpose, we exploit again the singularity of one of the matrices $M_{\mathbf{V}_s}^i$. We know that these matrices are symmetric positive semi-definite and one is singular. Furthermore, \mathbf{V}_s is a non-negative sum of $\mathbf{b}(s_i)$ for some $s_i \in E$. Let us write \mathbf{X} the eigenvector associated to the eigenvalue zero of $M_{\mathbf{V}_s}^i$, *i.e.*

$$M_{\mathbf{V}_s}^i \mathbf{X} = 0_{\mathbf{b}_K}.$$

Since \mathbf{V}_s is the moment vector of a (singular) measure γ_s , we have

$$\mathbf{X}^T M_{\mathbf{V}_s}^i \mathbf{X} = \int_E p_i(s) (\mathbf{b}(s)^T \mathbf{X})^2 d\gamma(s) = 0,$$

where p_i are the polynomial 1, $(1 \pm s)$, or $(1 - s^2)$ depending on the case considered, associated to the matrix $M_{\mathbf{V}_s}^i$, *i.e.* from formula (11-13). Especially, p_i is non-negative on E . Since γ is non-negative and $p(s) (\mathbf{b}(s)^T \mathbf{X})^2$ also, this implies that the locations s_i of the Diracs composing γ are the roots of $(\mathbf{b}^T \mathbf{X})p$. As this holds for all eigenvectors \mathbf{X} associated to zero, we obtain

$$\text{Supp}(\gamma) = \bigcap_i \bigcap_{M_{\mathbf{V}_s^i} \mathbf{X} = 0} Z(p\mathbf{b}\mathbf{X}), \quad (23)$$

where $\text{Supp}(\gamma)$ is the support of the measure γ , and $Z(p)$ is the zero set of p . This all rewrites:

Hausdorff even case $N = 2K$: • If $R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^N \quad \text{s.t.} \quad R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T) \mathbf{X} = 0_{\mathbb{R}^K} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z(\mathbf{b}_K \mathbf{X}).$$

• If $R_{\mathbf{V}_s}((1 - s^2)\mathbf{b}_{K-1} \mathbf{b}_{K-1}^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^K \quad \text{s.t.} \quad R_{\mathbf{V}_s}((1 - s^2)\mathbf{b}_{K-1} \mathbf{b}_{K-1}^T) \mathbf{X} = 0_{\mathbb{R}^{K-1}} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z((1 - s^2)\mathbf{b}_{K-1} \mathbf{X}).$$

Hausdorff odd case $N = 2K + 1$: • If $R_{\mathbf{V}_s}((1 - s)\mathbf{b}_K \mathbf{b}_K^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^K \quad \text{s.t.} \quad R_{\mathbf{V}_s}((1 - s)\mathbf{b}_K \mathbf{b}_K^T) \mathbf{X} = 0_{\mathbb{R}^K} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z((1 - s)\mathbf{b}_K \mathbf{X}).$$

• If $R_{\mathbf{V}_s}((1 + s)\mathbf{b}_K \mathbf{b}_K^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^K \quad \text{s.t.} \quad R_{\mathbf{V}_s}((1 + s)\mathbf{b}_K \mathbf{b}_K^T) \mathbf{X} = 0_{\mathbb{R}^K} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z((1 + s)\mathbf{b}_K \mathbf{X}).$$

Hamburger even case $N = 2K$: • If $R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^K \quad \text{s.t.} \quad R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T) \mathbf{X} = 0_{\mathbb{R}^K} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z(\mathbf{b}_K \mathbf{X}).$$

Hamburger odd case $N = 2K + 1$: • If $R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^K \quad \text{s.t.} \quad R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T) \mathbf{X} = 0_{\mathbb{R}^K} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z(\mathbf{b}_K \mathbf{X}).$$

Stieltjes even case $N = 2K$: • If $R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^K \quad \text{s.t.} \quad R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T) \mathbf{X} = 0_{\mathbb{R}^K} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z(\mathbf{b}_K \mathbf{X}).$$

• If $R_{\mathbf{V}_s}(s\mathbf{b}_{K-1} \mathbf{b}_{K-1}^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^{K-1} \quad \text{s.t.} \quad R_{\mathbf{V}_s}(s\mathbf{b}_{K-1} \mathbf{b}_{K-1}^T) \mathbf{X} = 0_{\mathbb{R}^{K-1}} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z(s\mathbf{b}_{K-1} \mathbf{X}).$$

Stieltjes odd case $N = 2K + 1$: • If $R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^K \quad \text{s.t.} \quad R_{\mathbf{V}_s}(\mathbf{b}_K \mathbf{b}_K^T) \mathbf{X} = 0_{\mathbb{R}^K} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z(\mathbf{b}_K \mathbf{X}).$$

• If $R_{\mathbf{V}_s}(\mathbf{s} \mathbf{b}_K \mathbf{b}_K^T)$ is singular,

$$\forall \mathbf{X} \in \mathbb{R}^K \quad \text{s.t.} \quad R_{\mathbf{V}_s}(\mathbf{s} \mathbf{b}_K \mathbf{b}_K^T) \mathbf{X} = 0_{\mathbb{R}^K} \quad \text{then} \quad \text{Supp}(\gamma) \subset Z(\mathbf{s} \mathbf{b}_K \mathbf{X}).$$

Both the eigenvectors \mathbf{X} and the roots of these polynomials can be computed with standart techniques.

Computation of the coefficients α_i . Finally, in order to retrieve the coefficients α_i in front of the Diracs, we use the fact that $\gamma_s = \sum_{i=1}^J \alpha_i \delta_{s_i}$ realize the moments \mathbf{V}_s . Knowing the locations s_i , we simply compute the first J moments of such distribution to obtain

$$R_{\mathbf{V}_s}(\mathbf{b}_J) = \sum_{i=1}^J \alpha_i \mathbf{b}_J(s_i).$$

This is a simple linear system that we can invert to obtain the coefficients α_i .

4. Numerical test for the radiative transfer equation. We consider here the equation (1) with $s \in [-1, +1]$ and its moment system (3a) and a collision operator of the form

$$C(f) = K \left(\frac{1}{2} \int_{-1}^{+1} f(s) ds - f \right), \quad \mathbf{C}(\mathbf{f}) = K \left(\frac{R_{\mathbf{f}}(1)}{2} \mathbf{V}_{iso} - \mathbf{f} \right), \quad \mathbf{V}_{iso}^i = \int_{-1}^{+1} s^i ds.$$

The equation is discretized with the following scheme

$$\frac{\mathbf{f}_i^{n+1} - \mathbf{f}_i^n}{\Delta t} + \frac{\mathcal{F}_{i+\frac{1}{2}}^n - \mathcal{F}_{i-\frac{1}{2}}^n}{\Delta x} = K \left(\frac{R_{\mathbf{f}_i^n}(1)}{2} \mathbf{V}_{iso} - \mathbf{f}_i^{n+1} \right), \quad (24a)$$

$$\mathcal{F}_{i+\frac{1}{2}} = \frac{1}{2} [\mathbf{F}(\mathbf{f}_{i+1}^n) + \mathbf{F}(\mathbf{f}_i^n) + (\mathbf{f}_{i+1}^n - \mathbf{f}_i^n)]. \quad (24b)$$

One easily proves using standart techniques (see *e.g.* [43, 3, 51, 45, 32]) that this schemes preserves realizability as the solution \mathbf{f}_i^{n+1} is a linear combination with positive coefficients of realizable vectors of the form \mathbf{f} or $\mathbf{f} \pm \mathbf{F}(\mathbf{f})$. The consistence with (3a) is classical. The stability is however more complicated to study as it relies on the fact that the eigevalues of the Jacobian of the fluxes are smaller than one, *i.e.* that $Sp(\nabla_{\mathbf{f}} \mathbf{F}(\mathbf{f})) \subset [-1, 1]$, which is commonly expected in radiative transfer, though this remains to verify with the present Π_N closure. This analysis is postponed to futur work.

We test and compare the closure presented in the previous section with the P_N polynomial approximation and the K_N closure (see Section 3.2 and [31, 42, 52]) on three test cases. The first one is simply meant to study quantitatively the accuracy of the methods. The other two cases are known to be difficult to model with moment approaches as they require good approximation of both isotropic and purely anisotropic regimes. The methods are compared qualitatively on these two cases.

For the first two test cases, the obtained solution is compared to a reference which is computed using the following method. Remarking that the kinetic original equation (1) is linear, we decompose its solution f into two parts f_n and f_s

respectively solution of

$$\partial_t f_n + s \partial_x f_n = -K f_n, \quad (25a)$$

$$f_n(x) = f(x) \quad \text{on the boundaries } x = 0 \text{ and } x = L, \quad (25b)$$

$$\partial_t f_s + s \partial_x f_s = K \left(\frac{1}{2} \int_{-1}^{+1} f_s(s) + f_n(s) ds - f_s \right), \quad (25c)$$

$$f_s(x) = 0 \quad \text{on the boundaries } x = 0 \text{ and } x = L. \quad (25d)$$

Physically, f_n and f_s correspond to the distribution of particles that have never scattered in the domain and to the ones that have scattered at least once. The first equation is solved analytically, while the second is solved using a P_{24} moment method and the scheme (24). This decomposition is commonly more accurate than discretizing directly (1) because f_n is analytical and f_s is smooth enough with respect to s such that basic moment approximation is accurate enough. This technique was exploited *e.g.* for the developpement of the codes [37, 58, 29, 7, 6, 13, 12, 44].

For the last test case, the distribution f_n of non-scattered particles is enhanced by the initial condition which is thus a Dirac in time and position. Even if this distribution can be computed, it requires a special treatment when introduced in the second equation. Instead, we compare to the P_{24} solution applied directly to the full equation (1). We found experimentally this solution smooth and accurate enough for the present problem.

4.1. Simple beam test. The beam is modeled by giving as boundary conditions the moments of

$$\mathbf{f}(t, x = 0) = 10^{12} \mathbf{b}(1) \exp(-10t).$$

The initial condition is set to

$$\mathbf{f}(t = 0, x) = 0_{\mathbb{R}^{Card(\mathbf{b})}}.$$

Remark that numerical simulations using realizable closures, typically M_N , commonly require realizable initial and boundary conditions in the sense of $L^1(E)^+$ functions, *i.e.* in $\mathcal{R}_{\mathbf{b}}$. Here, this is not necessary and those are chosen on the boundary $\partial \mathcal{R}_{\mathbf{b}}^m$.

We fix a spatial domain $[0, L = 1]$ meshed with 101 cells, the collision parameter is fixed at $K = 6$ low enough to preserve an anisotropic distribution in the domain, but sufficiently large such that the discontinuity due to the propagation of non-scattered particles f_n in (25) is smoothed down. Such discontinuities may affect the convergence of the method with respect to the number N of moments. The final time $T_{\max} = 0.8$ is chosen such that the beam has not yet reach the other end at final time. The time step is computed from the fixed Courant number at $\frac{\Delta t}{\Delta x} = 0.95$.

The moments of order 0 and 1 obtained at final time with Π_7 , K_7 , \bar{P}_7 and the reference solution of (25) (first equation solved analytically, second one solved with P_{24}) are plotted on Fig. 1 as an indication of the expected solution. The discrete l^1 , l^2 and l^∞ errors of the f^0 component compared to the reference solution are plotted as a function of N on Fig. 2.

On this simple test case, the realizable K_N and Π_N methods provides much more accurate results. This is expected since realizable closures captures perfectly Dirac distributions, while polynomial approximations cannot. And the solution of the present problem (1) or equivalently of (25) is the sum of the propagation of a Dirac distribution f_n with a regular one f_s .

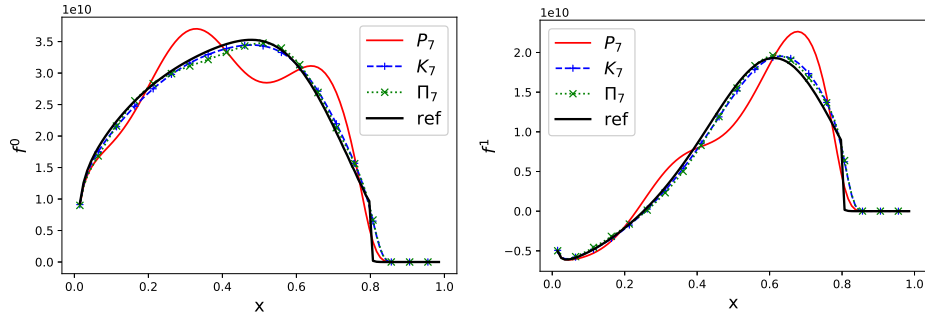


FIGURE 1. Moments of order 0 (left) and 1 (right) obtained with P_7 , K_7 , Π_7 and reference solution for the simple beam test case.

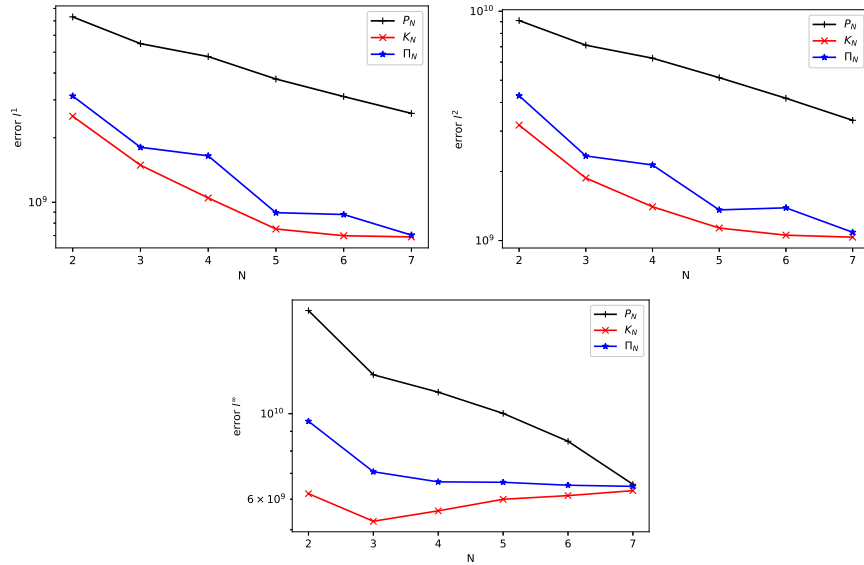


FIGURE 2. Discrete l^1 (top left), l^2 (top right) and l^∞ (bottom) errors on the moment of order 0 compared to a reference solution for the P_N , K_N and Π_N as a function of N for the simple beam test case.

One observes convergence with respect to N in discrete l^1 and l^2 norm for the K_N and Π_N models. This convergence is faster than for the P_N method which is already exponential. The discrete l^∞ error cannot tend to zero on this test because of the discontinuity due to the propagation of the front of f_n .

Qualitatively, the moment results are similar for both odd and even order moment methods. Though, quantitatively, one observes a small difference of precision between the odd and even order Π_N methods. Both methods converge in discrete l^1 and l^2 and l^∞ norms (the solution being smooth enough for that), but the odd order ones are more accurate. These discrepancies are due to the numerical methods used for the computation of the closure for several reasons which are mainly related to

the moment matrices $R_{\mathbf{V}}((1 \pm \mu)\mathbf{b} \otimes \mathbf{b})$, $R_{\mathbf{V}}((1 - \mu^2)\mathbf{b} \otimes \mathbf{b})$ and $R_{\mathbf{V}}(\mathbf{b} \otimes \mathbf{b})$ that we exploit for the construction of Π_N closure:

- The matrices to pseudo-invert have a different size between odd and even orders.
- Some moment matrices $R_{\mathbf{V}}((1 \pm \mu)\mathbf{b} \otimes \mathbf{b})$, $R_{\mathbf{V}}((1 - \mu^2)\mathbf{b} \otimes \mathbf{b})$ and $R_{\mathbf{V}}(\mathbf{b} \otimes \mathbf{b})$ are expected to be singular. We may expect for this test case that the moment solution in most of the cells to be close to the moment of a Dirac peak in $s = +1$. For this reason, $R_{\mathbf{V}}((1 - \mu)\mathbf{b} \otimes \mathbf{b})$ should have a smaller dimension than the other matrices (due to the position of the Dirac). Thus, the numerical errors produced during the computations of the different moment matrices may affect differently the QR algorithm used.
- Similarly, the formula (21) and (22) are simply different between odd and even orders and this may affect differently the numerical accuracy.

4.2. Double beam test. This test case ([26, 46, 45]) consists in having two beams of particles cross each other. As beams are used, purely anisotropic distribution need to be well-modeled by the approach. Furthermore, in the mixing region, the distribution is the sum of two anisotropic distribution, and low order approaches (first order) are insufficient to model such distributions.

The two beams are modeled by giving as boundary conditions the moments of Diracs

$$\begin{aligned} \mathbf{f}(t, x = 0) &= 10^{12}\mathbf{b}(+1) \exp(-20t), \\ \mathbf{f}(t, x = L) &= 10^{12}\mathbf{b}(-1) \exp(-20t), \end{aligned}$$

Again the initial condition is fixed at

$$\mathbf{f}(t = 0, x) = 0_{\mathbb{R}^{Card(\mathbf{b})}}.$$

We fix a spatial domain $[0, L = 2]$ meshed with 201 cells, the collision parameter is again fixed at $K = 6$. The final time $T_{\max} = 1.5$ is chosen such that the beam emerging from one end has not reached the other end at final time but they have crossed each others in the center region. The time step is computed from the fixed Courant number at $\frac{\Delta t}{\Delta x} = 0.95$.

The moments of order 0 and 1 obtained at final time with Π_N , K_N and P_N for $N = 2, 3, 6, 7$ (even and odd together) are plotted on Fig. 3. The results with the reference solution of (25) (sum of resp. analytical and P_{24} solutions) are also given as reference. The discrete l^1 , l^2 and l^∞ errors on the 0-th moment compared to this reference of P_N , Π_N and K_N models are plotted as a function of N on Fig. 4.

On this test case, the realizable closures are again more accurate than the standard P_N closure. This is expected for the same reason as in the previous case. Especially the kinetic distribution in the middle region is composed of two beams of opposite directions. Such distributions are poorly approximated by polynomials, which oscillate around the solution, or by low order moment methods, while high order realizable closures capture them perfectly. There remains some differences between Π_N and K_N . Mainly, the Π_N models present a small bump, alternatively positive and negative, in the middle of the domain in the 0-th moment plot. This bump is characteristic of the moment approximation of double-beam distributions.

As in the previous case, all the models converge with N to the desired solution. Though, the final time T_{\max} is higher than in the previous case, and the two beams cross each others. Such problems are known to be difficult to model with moment approaches and we observe less regular convergence results with respect to N .

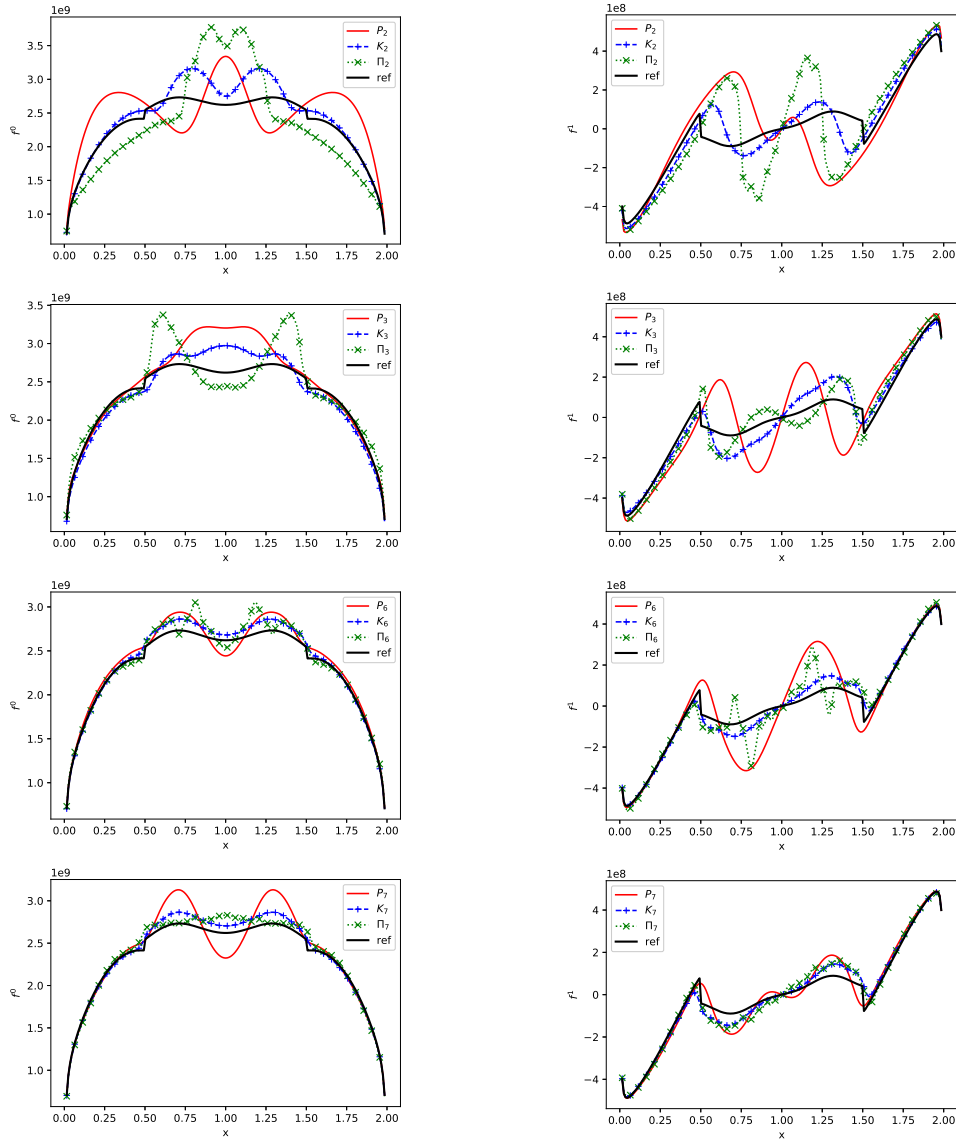


FIGURE 3. Moments of order 0 (left) and 1 (right) obtained with P_N , K_N , Π_N for $N = 2$ (first line), $N = 3$ (second line), $N = 6$ (third line), $N = 7$ (fourth line) and reference solution for the double beam test case.

In order to study the mix of the two beams in the middle region, the measures γ representing the moments $\mathbf{f}(T_{\max}, x = L/2)$ with the different models are represented on Fig. 5. These measures are given by (17) for K_N , (19) for Π_N and by a basic polynomial reconstruction for P_N . For K_N and Π_N , these representing measures are composed of a regular part and of a sum of Diracs. The Diracs are represented by a vertical segment of length α , the coefficient in front of the Diracs, and located in s , the position of the Dirac.

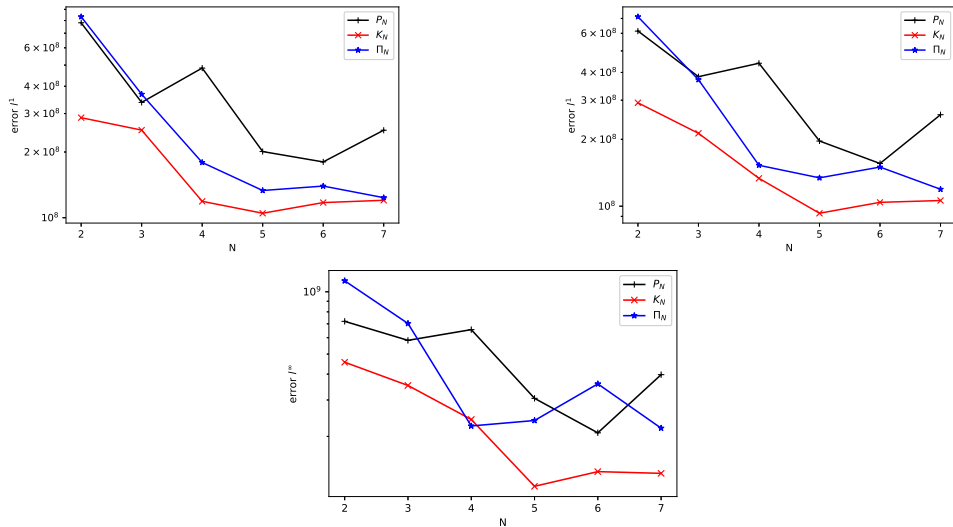


FIGURE 4. Discrete l^1 (top left), l^2 (top right) and l^∞ (bottom) errors on the moment of order 0 compared to a reference solution for the P_N , K_N and Π_N as a function of N for the double beam test case.

Every measure γ is symmetric with respect to $s = 0$ which is expected from the x -symmetry of the problem. For this particular problem, at the very location $x = L/2$, the P_N polynomial approximation is positive. As expected, the representing measures with Π_N models is composed of less Diracs than for the K_N model. Furthermore, these Dirac peaks have a lower amplitude than with K_N . The number of Dirac peaks raises with N for the K_N models, while they seem to remain two for the Π_N models. The positions and the amplitude of those two Diracs and the constant value \bar{x} seem however to oscillate around a stable value.

Comparing the results with the two realizable closures K_N and Π_N , we have a better qualitative and quantitative accuracy with K_N models. This can also be interpreted from the representing measures. On this test case, the expected solution is supposed to be the superposition of two beams located around $s = \pm 1$. The K_N representing measures naturally captures purely anisotropic measures, and even enhanced some, as it is constructed from the sum of (*a priori* an overestimated number of) Dirac peaks. On the contrary, the Π_N construction tends to overestimate the isotropic part of the underlying representing measure as \bar{x} is maximized. This results in an overestimation of the diffusion effects in s , and thus the appearance of this bump in the center region. However, there remain some flexibility in these model, especially in the choice of the equilibrium function f_{eq} the models aim to capture. Studying alternative choices of such equilibrium functions is part of outlooks.

4.3. Point source test. The last test case ([11, 24, 26]) is a 1D version of the line source problem. It consists in defining an isotropic source of particles in the middle of a domain at initial time and letting it spread in all directions. In this test, anisotropic distributions are enhanced from an isotropic one because, using a small

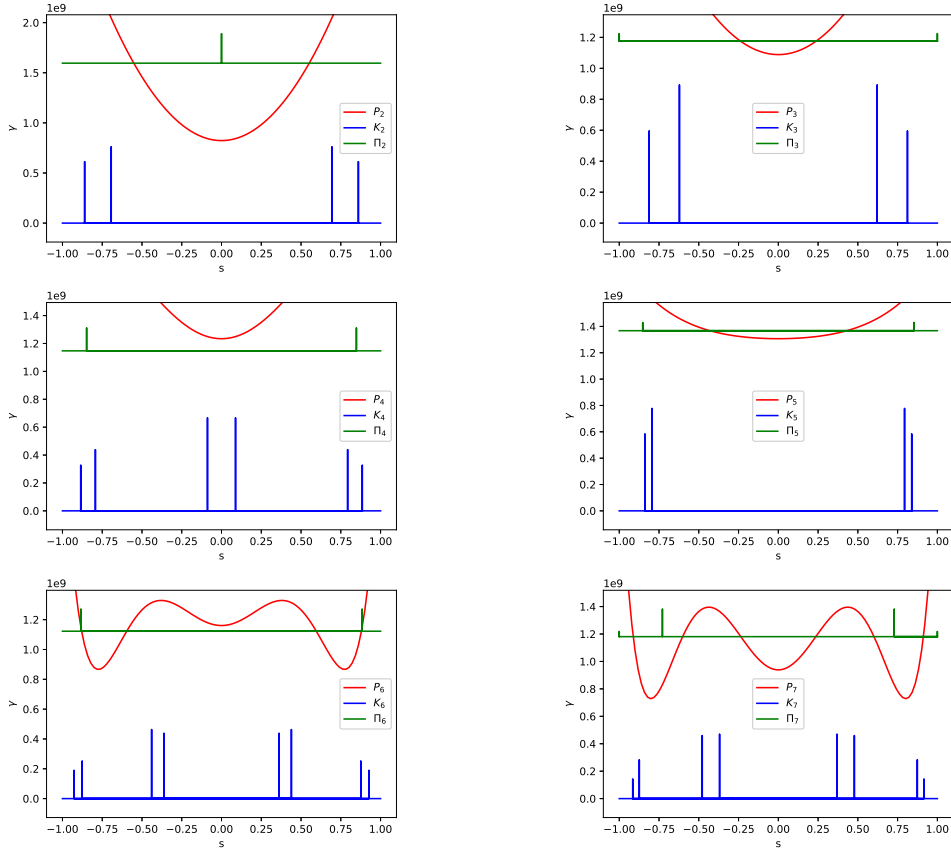


FIGURE 5. Representation of the measures representing the vector $\mathbf{f}(x = L/2)$ with P_N , K_N and Π_N models with $N = 2$ (top left), 3 (top right), 4 (middle left), 5 (middle right), 6 (bottom left), 7 (bottom right) for the double beam test case.

collision parameter K , the distribution tends to a anisotropic distribution pointed away from the center.

The final time $T_{\max} = 0.5$ is again chosen such that no particles has reach a boundary. In practice, we use as boundary conditions the moments of a realizable very small distribution

$$\mathbf{f}(t, x = 0) = 0_{\mathbb{R}^{Card(\mathbf{b})}}, \quad \mathbf{f}(t, x = L) = 0_{\mathbb{R}^{Card(\mathbf{b})}},$$

while the initial condition is fixed as

$$\mathbf{f}(t = 0, x) = 10^{12} \delta_{L/2}(x) \mathbf{V}_{iso},$$

i.e. a large isotropic distribution in the center of the domain.

The moments of order 0 and 1 obtained at final time with Π_N , K_N and P_N for $N = 2, 3, 6, 7$ (even and odd together) are plotted on Fig. 6. The results with a P_{24} model are also given as a high order reference. The discrete l^1 , l^2 and l^∞ errors on the 0-th moment compared to the most refined solution P_{24} of P_N , Π_N and K_N models are plotted as a function of N on Fig. 7.

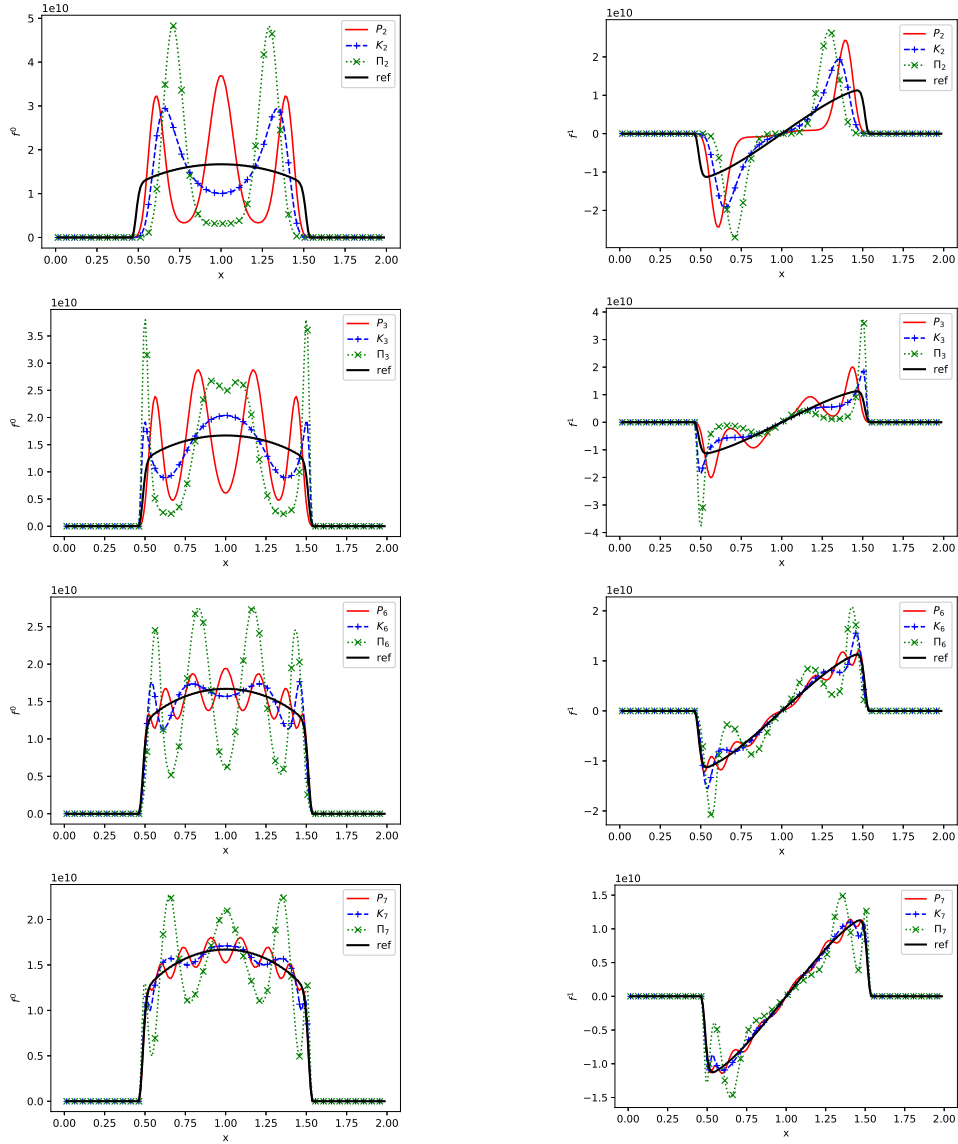


FIGURE 6. Moments of order 0 (left) and 1 (right) obtained with P_N , K_N , Π_N for $N = 2$ (first line), $N = 3$ (second line), $N = 6$ (third line), $N = 7$ (fourth line) and a reference P_{24} solution for the point source test case.

Again, we observe convergence with respect to N toward the desired solution. However, we see rather different phenomena compared to the last test case. Here, the accuracy of Π_N method is lower than the others, even than the polynomial P_N one. Both realizable closures K_N and Π_N , as well P_N closure, present oscillations around the reference solution. The quantity of these oscillations raise with N but decrease in amplitude. Their amplitudes are higher with Π_N approximation.

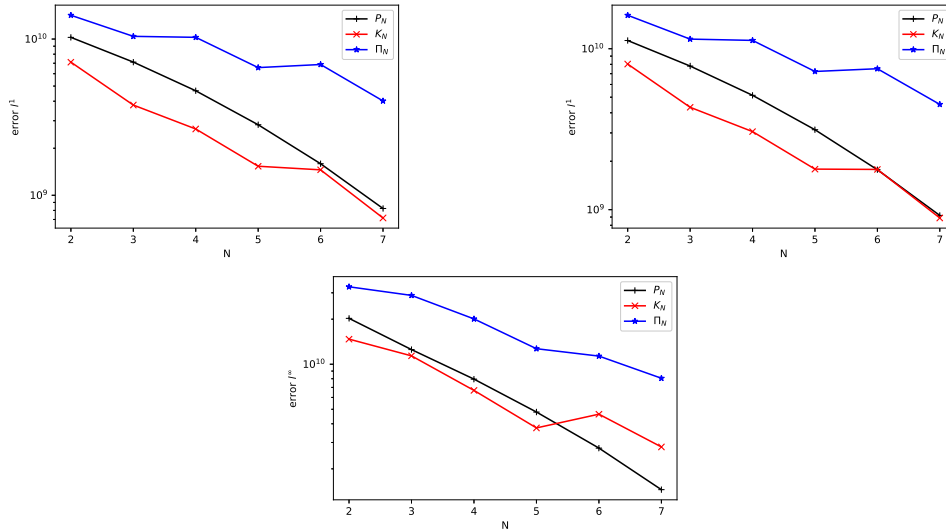


FIGURE 7. Discrete l^1 (top left), l^2 (top right) and l^∞ (bottom) errors on the moment of order 0 compared to a reference P_{24} simulation for the P_N , K_N and Π_N as a function of N for the point source test case.

As for the last test case, in order to study this phenomenon, we plot on Fig. 8 the measures representing the moments $\mathbf{f}(T_{\max}, x = L/2)$ in the middle of the domain, where the source was at initial time.

Again, the measure are all symmetric with respect to $s = 0$. However, for this test case, the P_N approximation is no more positive and have large amplitudes. Again the Π_N measures possess less Diracs than K_N ones and those are located closer to the center $s = 0$ than with K_N .

Again, we may interpret the difference of accuracy of the models from their representing measures. On this test case as in the previous one, we expect the solution at one location to be an accumulation of beams, *i.e.* the initial one emerging from the middle of the domain, and secondary ones created by collisional effects. Such distribution are better captured by an accumulation of Diracs such as the one used with K_N , and less by isotropic distribution such as the one maximized with Π_N model. Here, we exploited the most standart equilibrium function f_{eq} , though a better understanding of the expected solution could lead to use a better adapted function in the definition of K_N and Π_N models.

One may observe also a very small negative amplitude of two Diracs (symmetric) in the K_7 representation. These negative values are due to the numerical approximation of the position of $\partial\mathcal{R}_{\mathbf{b}}^m$. In practice, the moment matrix associated to K_7 closure should be singular (*i.e.* on the boundary of $\partial\mathcal{R}_{\mathbf{b}}^m$), while round-off errors may prevent such matrices to have exactly 0 as eigenvalues. In our computations, we chose to filter away the smallest eigenvalues of those moment matrices, *i.e.* the ones below a threshold of $10^{-4}R_{\mathbf{f}}(1)$. This seems sufficient to have decent simulation results, though some small negative value may appear in the representing measures of the method. Remark although that the closure remains well-defined and it affects not the accuracy of the simulation.

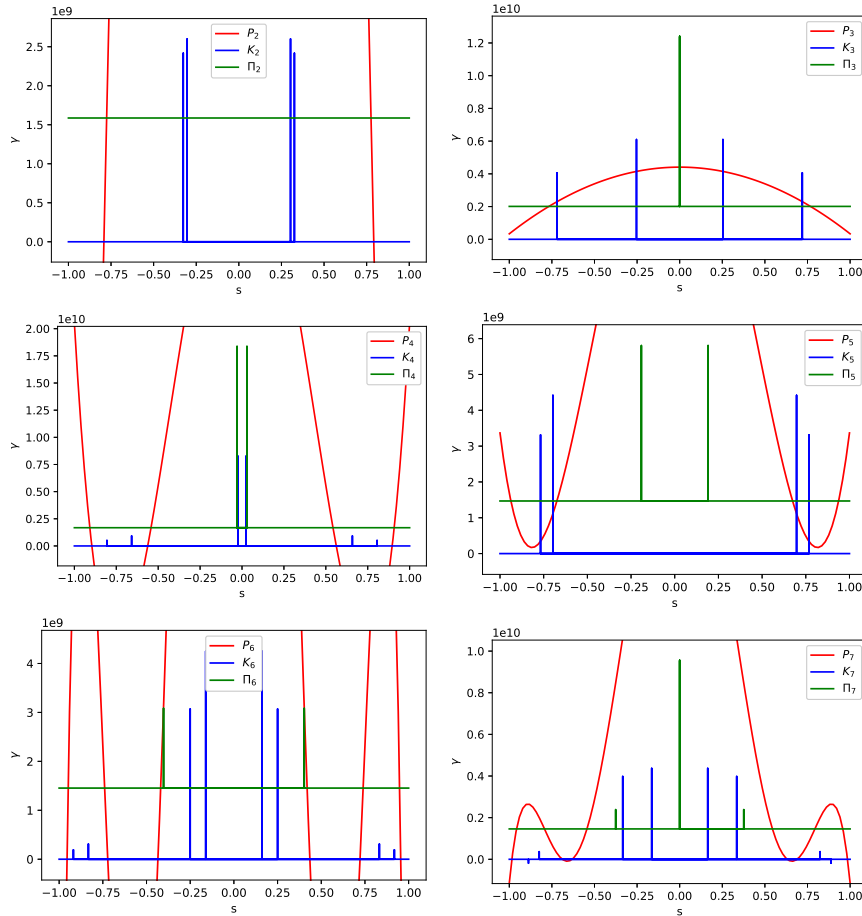


FIGURE 8. Representation of the measures representing the vector $\mathbf{f}(x = L/2)$ with P_N , K_N and Π_N models for $N = 2$ (top left), 3 (top right), 4 (middle left), 5 (middle right), 6 (bottom left), 7 (bottom right) for the point source test case.

5. Concluding discussions and perspectives. We have presented a method based on the study of the realizability domain to construct realizable closures for moment models over 1D domains. This method consists in projecting a realizable vector on the boundary of the realizability domain along the direction of a given realizable vector. The numerical computation of the closure relies on standard numerical techniques: one Cholesky decomposition, the computations of eigenvalues of symmetric positive semi-definite matrices and one pseudo-inverse. However, several aspects of the construction and of the analysis of method are still missing. We list here some that are left as perspectives.

5.1. Closure for multi-D integration domain. The projection techniques presented in this paper entirely depend on the knowledge of the position of the boundary of the realizability domain. If this boundary is well known and characterized for

moments in 1D, this remains a difficult question for moment over multi-D domains, typically on the unit sphere S^2 or on \mathbb{R}^3 for the present applications.

Similar projection techniques were already exploited for the construction of approximations of the second order entropy-based M_2 closure in 3D in [46, 36, 50] over S^2 . For moments up to order 2, the associated realizability domain has been characterized ([31]). However, those constructions are only valid inside the realizability domain and on its boundary. Especially, they are based on the computation of some parameters characterizing the distance to the boundary, and which can not be generalized out of $\mathcal{R}_{\mathbf{b}_2}^m$. This non-linear realizability constraint on the numerical solution also affects the robustness of the numerical method. This holds when using basic first order schemes, such as (24), to the moment system. It affects even more the robustness of a code when using high order numerical schemes, which generally require a particular treatment to preserve realizability (see *e.g.* [26, 4, 53]).

The present construction remains valid out of the realizability domain and can therefore be exploited to extend the 3D second order closures [46, 36, 50] out of $\mathcal{R}_{\mathbf{b}}^m$.

5.2. Hyperbolicity and entropy decay. Only the realizability of the present closure was studied here. However, among the main desirable properties expected from a moment closure, one should list hyperbolicity of the resulting moment system of equation, *i.e.* that the Jacobian of the flux \mathbf{F} needs to be diagonalizable in \mathbb{R} ; and entropy decay through this model, *i.e.* the existence of a convex function h that is dissipated through the the moment system (3a).

From the present construction, it is unclear if both of these properties are satisfied by our approach. These two properties are obtained jointly in the construction of the entropy-based M_N models. For such a closure, the representing measure γ is constructed by minimizing an entropy η ([39, 8, 9, 10, 55, 30, 34]) under the constraint of satisfying the moments (and potentially of being positive [25]). This reconstruction is known to be of the form $\eta^*(\boldsymbol{\lambda} \cdot \mathbf{b}(s)) ds$ where η^* is the Legendre dual of η and $\boldsymbol{\lambda}$ are Lagrange multipliers. The resulting system is symmetric hyperbolic ([23]) and dissipates the entropy η .

The K_N closure was shown to be hyperbolic for low order N in [42, 43, 52] by straight computations of the Jacobian of the flux. The present Π_N construction is *a priori* not based on an entropy minimization. However, the representing measure γ behind Π_N closure is constructed to minimize the distance, in a certain measure sense, to the chosen equilibrium function $f_{eq}(s) ds$. In our computations, this results in maximizing \bar{x} in (18).

Furthermore, the present approach was only tested on a simple radiative transfer equation. Experimentally, such problems over $s \in [-1, +1]$ present less stiffness than problems over non-bounded domains $s \in \mathbb{R}$ or $s \in \mathbb{R}^+$. For instance, one typically requires the bounds on $Sp(\nabla_{\mathbf{f}}(\mathbf{F}(\mathbf{f})))$ to construct stable numerical scheme for (3a). If those bounds are highly expected to be ± 1 for moments over $s \in [-1, +1]$ (this holds for most of the reasonable closures), they are unknown for moments on unbounded domains.

5.3. Solution-dependent equilibrium functions. For application in radiative transfer, one expects the solution to relax toward an equilibrium represented by an isotropic distribution. This choice is neither realistic for moments on unbounded domains, nor possible since such a constant function is not integrable. One more realistic choice would be to define the equilibrium function from a minimum-entropy

solution, typically a Maxwellian. However, such a Maxwellian is defined from the first moments. Such a choice of solution-dependent equilibrium function would not affect the present construction of the closure. It could be interpreted as a realizable correction of low order moment method exploiting higher order moments.

REFERENCES

- [1] N. I. Akhiezer. *The classical moment problem*. Edinburgh : Oliver & Boyd, 1965.
- [2] G. Alldredge and F. Schneider. A realizability-preserving discontinuous galerkin scheme for entropy-based moment closures for linear kinetic equations in one space dimension. *J. Comput. Phys.*, 295:665–684, 2015.
- [3] G. W. Alldredge. *Optimization Techniques for Entropy-Based Moment Models of Linear Transport*. PhD thesis, University of Maryland, 2012.
- [4] G. W. Alldredge, C. D. Hauck, and A. L. Tits. High-order entropy-based closures for linear transport in slab geometry II: A computational study of the optimization problem. *SIAM J. Sci. Comput.*, 34(4):361–391, 2012.
- [5] C. Bayer and J. Teichmann. The proof of Tchakaloff’s theorem. *arXiv:0502473*, pages 1–6, 2005.
- [6] G. Birindelli. *Modèle entropique pour le calcul de dose en radiothérapie externe et curi-thérapie*. PhD thesis, Université de Bordeaux, 2018.
- [7] G. Birindelli, J.-L. Feugeas, J. Caron, B. Dubroca, G. Kantor, J. Page, T. Pichard, V.T. Tikhonchuk, and Ph. Nicolaï. High performance modelling of the transport of energetic particles for photon radiotherapy. *Phys. Medica*, 42(8):305–312, 2017.
- [8] J. Borwein and A. Lewis. Duality relationships for entropy-like minimization problems. *SIAM J. Control Optim.*, 29(2):325–338, 1991.
- [9] J. Borwein and A. Lewis. Partially finite convex programming, part I: Quasi relative interiors and duality theory. *Math. Program.*, 57:15–48, 1992.
- [10] J. Borwein and A. Lewis. Partially finite convex programming: Part II. *Math. Program.*, 57:49–83, 1992.
- [11] T. A. Brunner and J. P. Holloway. One-dimensional Riemann solvers and the maximum entropy closure. *J. Quant. Spectros. Radiat. Transfer*, 69:543–566, 2000.
- [12] J. Caron. *Étude et validation clinique d’un modle aux moments entropique pour le transport de particules nergtiques : application aux faisceaux d’lectrons pour la radiothérapie externe*. PhD thesis, Univ. Bordeaux, 2016.
- [13] J. Caron, J.-L. Feugeas, B. Dubroca, G. Kantor, C. Dejean, G. Birindelli, T. Pichard, Ph. Nicolaï, E. d’Humières, M. Frank, and V. Tikhonchuk. Deterministic model for the transport of energetic particles. application in the electron radiotherapy. *Phys. Medica*, 31(8):912–921, 2016.
- [14] S. Chandrasekhar. *Radiative transfer*. Dover publications, 1960.
- [15] R. Curto and L. Fialkow. Recursiveness, positivity, and truncated moment problem. *Houston j. Math.*, 17(4):603–635, 1991.
- [16] R. Curto and L. Fialkow. A duality proof of Tchakaloff’s theorem. *J. Math. Anal. Appl.*, 269:519–532, 2002.
- [17] R. Curto and L. Fialkow. An analogue of the Riesz-Haviland theorem for the truncated moment problem. *J. Functional Analysis*, 255:2709–2731, 2008.
- [18] B. Dubroca and J.-L. Feugeas. Hirarchie des modles aux moments pour le transfert radiatif. *C. R. Acad. Sci. Paris*, 329:915–920, 1999.
- [19] L. Fialkow. The truncated K-moment problem: a survey. *Theta Ser. Adv. Math.*, 18:25–51, 2016.
- [20] R. O. Fox. A quadrature-based third-order moment method for dilute gas-particle flows. *J. Comput. Phys.*, 227:63136350, 2008.
- [21] M. Frank. *Partial moment models for radiative transfer*. PhD thesis, T.U. Kaiserslautern, 2005.
- [22] M. Frank, B. Dubroca, and A. Klar. Partial moment entropy approximation to radiative heat transfer. *J. Comput. Phys.*, 218:1–18, 2006.
- [23] K. O. Friedrichs and P. D. Lax. Systems of conservation equations with a convex extension. *Proc. Nat. Acad. Sci.*, 68(8):1686–1688, 1971.
- [24] C. K. Garrett and C. D. Hauck. A comparison of moment closures for linear kinetic transport equations: The line source benchmark. *Transport theory and Stat. Phys.*, 42:203–235, 2013.

- [25] C. Hauck and R. McClarren. Positive P_N closures. *SIAM J. Sci. Comput.*, 32(5):2603–2626, 2010.
- [26] C. D. Hauck. High-order entropy-based closures for linear transport in slab geometry. *Commun. Math. Sci.*, 9(1):187–205, 2011.
- [27] E. K. Haviland. On the momentum problem for distributions in more than one dimension ii. *Amer. J. Math.*, 57(3):562–568, 1935.
- [28] E. K. Haviland. On the momentum problem for distribution functions in more than one dimension. ii. *Amer. J. Math.*, 58(1):164–168, 1936.
- [29] H. Hensel, R. Iza-Teran, and N. Siedow. Deterministic model for dose calculation in photon radiotherapy. *Phys. Med. Biol.*, 51:675–693, 2006.
- [30] M. Junk. Maximum entropy for reduced moment problems. *Math. Mod. Meth. Appl. S.*, 10(1001–1028):2000, 1998.
- [31] D. Kershaw. Flux limiting nature’s own way. Technical report, Lawrence Livermore Laboratory, 1976.
- [32] K. Kuepper. *Models, numerical methods, and uncertainty quantification for radiation therapy*. PhD thesis, RWTH Aachen University, 2016.
- [33] J.-B. Lasserre. *Moment, positive polynomials, and their applications*, volume 1. Imperial college press, 2009.
- [34] C. D. Levermore. Moment closure hierarchies for kinetic theories. *J. Stat. Phys.*, 83(5–6):1021–1065, 1996.
- [35] A. S. Lewis. Consistency of moment systems. *Can. J. Math.*, 47:995–1006, 1995.
- [36] R. Li and W. Li. 3D B_2 model for radiative transfer equation, Part I: Modelling. *arxiv*, 2017.
- [37] D. S. Lucas, H. D. Gougar, T. Wareing, G. Failla, J. McGhee, D. A. Barnett, and I. Davis. Comparison of the 3-D deterministic neutron transport code Attila® to measure data, MCNP and MCNPX for the advanced test reactor. Technical report, Idaho National Laboratory, 2005.
- [38] J. McDonald and M. Torrilhon. Affordable robust moment closures for cfd based on the maximum-entropy hierarchy. *J. Comput. Phys.*, 251:500–523, 2013.
- [39] L. R. Mead and N. Papanicolaou. Maximum entropy in the problem of moments. *J. Math. Phys.*, 25(8):2404–2417, 1984.
- [40] G. N. Minerbo. Maximum entropy Eddington factors. *J. Quant. Spectros. Radiat. Transfer*, 20:541–545, 1978.
- [41] G. N. Minerbo. Maximum entropy reconstruction from cone-beam projection data. *Comput. Biol. Med.*, 9(1):29–37, 1979.
- [42] P. Monreal. Higher order minimum entropy approximations in radiative transfer. *arXiv:0812.3063*, pages 1–18, 2008.
- [43] P. Monreal. *Moment realizability and Kershaw closures in radiative transfer*. PhD thesis, RWTH Aachen University, 2012.
- [44] J. Page. *Dveloppement et validation de l’application de la force de Lorentz dans le modle aux moments entropiques M_1 . Etude de l’effet du champ magnétique sur le dpôt de dose en radiothrapie externe*. PhD thesis, Univ. Bordeaux, 2018.
- [45] T. Pichard. *Mathematical modelling for dose deposition in photontherapy*. PhD thesis, Universit de Bordeaux & RWTH Aachen University, 2016.
- [46] T. Pichard, G. W. Alldredge, S. Brull, B. Dubroca, and M. Frank. An approximation of the M_2 closure: application to radiotherapy dose simulation. *J. Sci. Comput.*, 71:71–108, 2017.
- [47] T. Pichard, D. Aregba-Driollet, S. Brull, B. Dubroca, and M. Frank. Relaxation schemes for the m1 model with space-dependent flux: Application to radiotherapy dose calculation. *Commun. Comput. Phys.*, 19(01):168–191, 2016.
- [48] T. Pichard, S. Brull, and B. Dubroca. A numerical approach for a system of transport equations in the field of radiotherapy. *Commun. Comput. Phys.*, 2018 (accepted).
- [49] M. Riesz. Sur le problme des moments, troisieme note. *Ark. Math. Astr. Fys.*, 17(16):1–52, 1923.
- [50] J. A. R. Sarr and C. P. T. Groth. Maximum-entropy inspired interpolative closure for radiative heat transfer in gray participating media. *9th International Symposium on Radiative Transfer (RAD-19)*, 2019.
- [51] F. Schneider. *Moment models in radiation transport equations*. PhD thesis, T.U. Kaiserslautern, 2015.
- [52] F. Schneider. Kershaw closures for linear transport equations in slab geometry i: model derivation. *J. Comput. Phys.*, 322:905–919, 2016.

- [53] F. Schneider, G. W. Alldredge, and J. Kall. A realizability-preserving high-order kinetic scheme using weno reconstruction for entropy-based moment closures of linear kinetic equations in slab geometry. *Kinetic and related models*, 9(1):193–215, 2016.
- [54] F. Schneider, J. Kall, and A. Roth. First-order quarter- and mixed-moment realizability theory and Kershaw closures for a Fokker-Planck equation in two space dimensions. *Kinetic and related models*, 10(04):1127–1161, 2017.
- [55] J. Schneider. Entropic approximation in kinetic theory. *ESAIM-Math. Model. Num.*, 38(3):541–561, 2004.
- [56] T.-J. Stieltjes. Recherches sur les fractions continues. *Ann. Fac. Sci. Toulouse : Mathématiques*, 8(4):J1–J122, 1894.
- [57] V. Tchakaloff. Formules de cubature mécanique à coefficients non négatifs. *Bull. Sci. Math.*, 81:123–134, 1957.
- [58] T.A. Wareing, J.M. McGhee, Y. Archambault, and S. Thompson. Acuros XB [®] advanced dose calculation for the Eclipse [™] treatment planning system. *Clinical perspectives*, 2010.
- [59] C. Yuan, F. Laurent, and R. O. Fox. An extended quadrature method of moments for population balance equations. *J. Aerosol Sci.*, 51:123, 2012.

E-mail address: teddy.pichard@polytechnique.edu