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. Kinetic and Related Models , 2020, 13 (6), pp.1243-1280. 10.3934/xx.xx.xx.xx . hal-02423328v2

HAL Id: hal-02423328

https://hal.science/hal-02423328v2

Submitted on 5 Aug 2020

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

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 standard techniques providing a realizable closure for the full system. This technique 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 distriution and a chosen regular equilibrium function. The starting point is a kinetic equation of the form

$$\partial_t f + s\partial_x f = C(f), \tag{1}$$

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.*

$$dN = f(t, s, x)dsdx$$

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 $\mathbb{R}$.

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.* [15, 44, 19]). 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 [50, 36, 56, 4, 7, 49, 47]). Those moments, afterward written $\mathbf{f}$ in bold, are weighted integrals of f against a vector $\mathbf{b}(s)$ of polynomial weights. This yields

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

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

$$\partial_t \mathbf{f} + \partial_x \mathbf{F} = \mathbf{C}, \tag{3a}$$

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

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

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

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

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

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 ([44, 45, 38, 42]) or for the atom-based closures K_N (based on the idea of [34] then developped in [46, 47] and exploited in [57]). 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 [51, 50, 40]). 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 overdiffusive 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 [36]). Instead, closures based on a positive reconstruction $\tilde{f}$ capture

these phenomena properly, even with low order moments (see *e.g.* [51, 40, 55]). 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 [61, 37]. Several TMP were solved, mostly when the set of integration is 1D. We may list the truncated moment problems of Hausdorff ($E = [-1, +1]$;[29]), Hamburger ($E = \mathbb{R}$; [26]), Stieltjes ($E = \mathbb{R}^+$; [61]), Toeplitz ($E = S^1$), see *e.g.* [16] for a review on those results.

The objective of this work is to provide a numerically tractable closure (4b) for 1D problems that is realizable and generically applicable to arbitrary high order moment models over 1D sets. It is 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 ([16, 17, 18]) and of a regular integrable function. Such atomic decompositions was also used for the construction of K_N closure ([34, 46, 47, 57]) for moments over $[-1, +1]$, or of quadrature-based closures ([21, 64, 40]) 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 retreives exactly the equilibrium distribution f_{eq} (with K_N only its moment flux $\mathbf{F}(\mathbf{f}_{eq})$ is obtained).
- This projection technique 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. However, the choice of the equilibrium function f_{eq} requires further investigation (discussed in conclusion) for this closure to be applicable when E is unbounded.

In the following, the superscript $\tilde{}$ is droped, 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 analyzed 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

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

The monomial basis is used for its simplicity, though others can equally be used, *e.g.* the Legendre basis are often prefered 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; [39]). This is used *e.g.* to construct the partial moments method ([23, 22, 59, 56]).

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, *i.e.* $f \in L^1(E)^+$ if $f \in L^1(E)$ and

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

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

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

- 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

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

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

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

- Define also the convex cone

$$\mathcal{R}^m_{\mathbf{b}} = \left\{ \sum_{i=1}^{J} \alpha_i \mathbf{b}(s_i), \quad J < \infty, \quad \alpha_i \in \mathbb{R}^{*+}, \quad s_i \in E \right\}. \tag{6d}$$

  Remark that this is the set of moments with respect to positive discrete measures over E since $\mathbf{b}(s_i) = \int_E \mathbf{b}(s)\delta_{s_i}(s)$ where δ_{s_i} is the Dirac measure in $s_i \in E$.
- Finally, we will use extensively the closure set of $\mathcal{R}_{\mathbf{b}}$ in $\mathbb{R}^{\text{Card}(\mathbf{b})}$

$$\mathcal{R}^c_{\mathbf{b}} = \overline{\mathcal{R}}_{\mathbf{b}} \cap \mathbb{R}^{\text{Card}(\mathbf{b})}. \tag{6e}$$

Remark 1. Any Dirac measures on $E \subset \mathbb{R}$ can be interpreted as the limit of a sequence $f^\epsilon ds$ with $f^\epsilon \in L^1_N(E)^+$ when $\epsilon \to 0$. This provides that

$$\mathcal{R}^m_{\mathbf{b}} \subset \mathcal{R}^c_{\mathbf{b}}.$$

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 $p = \boldsymbol{\lambda}\mathbf{b}$ onto

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

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

If the vector $\mathbf{V} = \int_E \mathbf{b}(s)f(s)ds$ is the vector of moments of a function $f \in L^1_{\mathrm{Card}(\mathbf{b})+1}(E)^+$, 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 $\boldsymbol{f} = (f^0, f^1, f^2) \in \mathbb{R}^3$, and the vector of monomials $\mathbf{b}(s) = (1, s, s^2)$. The Riesz function according to the vector $\boldsymbol{f}$ of the polynomial

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

reads

$$\begin{array}{rcccccc} R_{\boldsymbol{f}}(p) & = & R_{\boldsymbol{f}}(1) & + & 3R_{\boldsymbol{f}}(s) & - & R_{\boldsymbol{f}}(s^2) \\ & = & f^0 & + & 3 \quad f^1 & - & f^2. \end{array}$$

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

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

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 exhbibit a neighborhood of $\mathbf{V}$ included in $\mathcal{R}_{\mathbf{b}}$. Write

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

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 $(\mathbf{V}_i^0)_{\{i=1,\dots,\mathrm{Card}(\mathbf{b})\}}$ of its column is a basis of $\mathbb{R}^{\mathrm{Card}(\mathbf{b})}$. Therefore, for all $\alpha > 0$, the set

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

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

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

With this choice of α, one shows that

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

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

Remark 2.
- This also provides that $\mathcal{R}^c_{\mathbf{b}}$ is a (closed) convex cone and that $\mathcal{R}_{\mathbf{b}} = int(\mathcal{R}^c_{\mathbf{b}})$.
- 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.* [3, 52, 53]).

2.3. **Characterizations of the realizability domain.** In this subsection, we recall the well-known Hausdorff, Hamburger and Stieltjes moment problems ([1, 16]) which provides characterizations of the realizability property.

Theorem 2.3 (Hausdorff)**.** *Suppose that $s\in E=[-1,+1]$ and either $\mathbf{b}=\mathbf{b}_{2K}$ (even case) or $\mathbf{b}=\mathbf{b}_{2K+1}$ (odd case). Define the matrices*

$$\textit{Even case}: \quad M^1_{\mathbf{V}}=R_{\mathbf{V}}(\mathbf{b}_K\mathbf{b}_K^T), \qquad M^2_{\mathbf{V}}=R_{\mathbf{V}}\left((1-s^2)\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T\right), \tag{11a}$$

$$\textit{Odd case}: \quad M^1_{\mathbf{V}}=R_{\mathbf{V}}\left((1+s)\mathbf{b}_K\mathbf{b}_K^T\right), \quad M^2_{\mathbf{V}}=R_{\mathbf{V}}\left((1-s)\mathbf{b}_K\mathbf{b}_K^T\right). \tag{11b}$$

Then,

- *A vector $\mathbf{V}\in\mathcal{R}_{\mathbf{b}}$ if and only if $M^1_{\mathbf{V}}$ and $M^2_{\mathbf{V}}$ are positive definite.*
- *A vector $\mathbf{V}\in\mathcal{R}^m_{\mathbf{b}}$ if and only if $M^1_{\mathbf{V}}$ and $M^2_{\mathbf{V}}$ are positive semi-definite.*
- *A vector $\mathbf{V}\in\mathcal{R}^c_{\mathbf{b}}$ if and only if $M^1_{\mathbf{V}}$ and $M^2_{\mathbf{V}}$ are positive semi-definite.*

Theorem 2.4 (Hamburger)**.** *Suppose that $s\in E=\mathbb{R}$ and either $\mathbf{b}=\mathbf{b}_{2K}$ (even case) or $\mathbf{b}=\mathbf{b}_{2K+1}$ (odd case). Define*

$$M^1_{\mathbf{V}} = R_{\mathbf{V}}(\mathbf{b}_K\mathbf{b}_K^T). \tag{12a}$$

Then,

- *A vector $\mathbf{V}\in\mathcal{R}_{\mathbf{b}}$ if and only if $M^1_{\mathbf{V}}$ is positive definite.*

Write furthermore

$$J = Rank(M^1_{\mathbf{V}}), \qquad M^2_{\mathbf{V}} = R_{\mathbf{V}}(\mathbf{b}_{J-1}\mathbf{b}_{J-1}^T), \qquad \mathbf{V}^j = R_{\mathbf{V}}(s^j\mathbf{b}_{J-1}), \tag{12b}$$

- *A vector $\mathbf{V}\in\mathcal{R}^m_{\mathbf{b}}$ if and only if $M^2_{\mathbf{V}}$ is positive definite and*

$$R_{\mathbf{V}}(s^j) = \begin{cases} \mathbf{V}^i(M^2_{\mathbf{V}})^{-1}\mathbf{V}^i & \textit{if} \quad j=2i \textit{ is even}, \\ \mathbf{V}^i(M^2_{\mathbf{V}})^{-1}\mathbf{V}^{i+1} & \textit{if} \quad j=2i+1 \textit{ is odd} \end{cases} \tag{12c}$$

  for all $2J<j\le Card(\mathbf{b})$,
- *A vector $\mathbf{V}\in\mathcal{R}^c_{\mathbf{b}}$ if and only if $M^1_{\mathbf{V}}$ is positive semi-definite.*

Theorem 2.5 (Stieltjes)**.** *Suppose that $s\in E=\mathbb{R}^+$ and either $\mathbf{b}=\mathbf{b}_{2K}$ (even case) or $\mathbf{b}=\mathbf{b}_{2K+1}$ (odd case). Define*

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

$$\textit{Odd case}: \quad M^1_{\mathbf{V}}=R_{\mathbf{V}}\left(\mathbf{b}_K\mathbf{b}_K^T\right) \qquad M^2_{\mathbf{V}}=R_{\mathbf{V}}(s\mathbf{b}_K\mathbf{b}_K^T). \tag{13b}$$

Then,

- *A vector $\mathbf{V}\in\mathcal{R}_{\mathbf{b}}$ if and only if $M^1_{\mathbf{V}}$ and $M^2_{\mathbf{V}}$ are positive definite.*

Write furthermore

$$J_1 = Rank(M^1_{\mathbf{V}}), \quad J_2 = Rank(M^2_{\mathbf{V}}), \tag{13c}$$

$$J = \min(J_1, J_2), \qquad \mathbf{V}^j = R_{\mathbf{V}}(s^j\mathbf{b}_{J-1}), \tag{13d}$$

and, if $J_1 \leq J_2$

$$M^3_{\mathbf{V}} = R_{\mathbf{V}}(\mathbf{b}_{J_1-1}\mathbf{b}^T_{J_1-1}), \qquad M^4_{\mathbf{V}} = R_{\mathbf{V}}(s\mathbf{b}_{J_1-2}\mathbf{b}^T_{J_1-2}), \qquad L = 2J_1. \tag{13e}$$

or, if $J_1 > J_2$

$$M^3_{\mathbf{V}} = R_{\mathbf{V}}(s\mathbf{b}_{J_2-1}\mathbf{b}^T_{J_2-1}), \qquad M^4_{\mathbf{V}} = R_{\mathbf{V}}(\mathbf{b}_{J_2-1}\mathbf{b}^T_{J_2-1}), \qquad L = 2J_2+1. \tag{13f}$$

- *A vector* $\mathbf{V} \in \mathcal{R}^m_{\mathbf{b}}$ *if and only if* $M^3_{\mathbf{V}}$ *and* $M^4_{\mathbf{V}}$ *are positive definite and*

$$R_{\mathbf{V}}(s^j) = \begin{cases} \mathbf{V}^i (M^3_{\mathbf{V}})^{-1}\mathbf{V}^i & \text{if} \quad j = 2i \text{ is even,} \\ \mathbf{V}^i (M^3_{\mathbf{V}})^{-1}\mathbf{V}^{i+1} & \text{if} \quad j = 2i+1 \text{ is odd,} \end{cases} \tag{13g}$$

  for all $L < j \leq Card(\mathbf{b})$.
- *A vector* $\mathbf{V} \in \mathcal{R}^c_{\mathbf{b}}$ *if and only if* $M^1_{\mathbf{V}}$ *and* $M^2_{\mathbf{V}}$ *are positive semi-definite.*

Proofs of the characterizations of $\mathcal{R}_{\mathbf{b}}$ can be found *e.g.* in [1, 2, 35, 16, 37]. The characteriszations of $\mathcal{R}^c_{\mathbf{b}}$ are obtained by taking the closure of $\mathcal{R}_{\mathbf{b}}$. The characterizations of $\mathcal{R}^m_{\mathbf{b}}$ in the case of Hausdorff is also well-established in the literature ([1, 2, 35, 16, 37]). The characterizations of $\mathcal{R}^m_{\mathbf{b}}$ in the cases of Hamburger and Stieltjes are obtained by reformulating results from [16]. These reformulations are provided for completeness.

First, definitions from [16] are recalled.

Definition 2.6.
- The **rank of a moment vector** $\mathbf{V} \in \mathbb{R}^N$ with $N = 2K$ or $N = 2K+1$ is

$$\mathrm{Rank}(\mathbf{V}) = \begin{cases} K+1 & \text{if } R_{\mathbf{V}}(\mathbf{b}_K\mathbf{b}^T_K) \text{ is positive definite,} \\ \min\limits_{\substack{1\leq j\leq K \\ R_{\mathbf{V}}(\mathbf{b}_j\mathbf{b}^T_j) \text{ singular}}} j & \text{otherwise.} \end{cases}$$

- A moment vector $\mathbf{V}$ is **positively recursively generated** if $R_{\mathbf{V}}(\mathbf{b}_{J-1}\mathbf{b}^T_{J-1})$ is positive definite for $J = \mathrm{Rank}(\mathbf{V})$ and there exists $(\beta_i)_{i=1,\dots,J} \in \mathbb{R}^J$ such that

$$R_{\mathbf{V}}(s^j) = \sum_{i=1}^{J} \beta_i R_{\mathbf{V}}(s^{j-i}) \quad \text{for all} \quad J \leq j \leq N.$$

Lemma 2.7. *Consider* $\mathbf{V} \in \mathcal{R}^m_{\mathbf{b}_N}$ *with* $N = 2K$ *or* $= 2K+1$ *and write*

$$J = Rank(\mathbf{V}), \quad M_{\mathbf{V}} = R_{\mathbf{V}}(\mathbf{b}_{J-1}\mathbf{b}^T_{J-1}) \quad \text{and} \quad \mathbf{V}^j = R_{\mathbf{V}}(s^j\mathbf{b}_{J-1}).$$

Then $\mathbf{V}$ *is positively recursively generated if and only if* $M_{\mathbf{V}}$ *is positive definite and for all* $2J-1 \leq j \leq Card(\mathbf{b})$

$$R_{\mathbf{V}}(s^j) = \begin{cases} \mathbf{V}^i M^{-1}_{\mathbf{V}}\mathbf{V}^i & \text{if} \quad j = 2i \text{ is even,} \\ \mathbf{V}^i M^{-1}_{\mathbf{V}}\mathbf{V}^{i+1} & \text{if} \quad j = 2i+1 \text{ is odd} \end{cases}$$

Proof. We only need to show that the coefficients $R_{\mathbf{V}}(s^j)$ are equal in this representation and in the definition.

If $\mathbf{V}$ is positively recursively generated, then $Rank(R_{\mathbf{V}}(\mathbf{b}_K\mathbf{b}^T_K)) = Rank(\mathbf{V})$. Define its submatrices

$$S^{2i} = \left(\begin{array}{c|c} M_{\mathbf{V}} & V^i \\ \hline (V^i)^T & R_{\mathbf{V}}(s^{2i}) \end{array}\right), \qquad S^{2i+1} = \left(\begin{array}{c|c} M_{\mathbf{V}} & V^{i+1} \\ \hline (V^i)^T & R_{\mathbf{V}}(s^{2i+1}) \end{array}\right).$$

By construction, $Rank(S^j) = Rank(\mathbf{V})$ and these submatrices are singular. Their first columns are linearly independent, then the last one is a combination of the others, and inverting this expression provides the result. □

Remark that this also provides $R_{\mathbf{V}}(s^{i+j}) = \mathbf{V}^i M_{\mathbf{V}}^{-1} \mathbf{V}^j$ in a general manner.

Proof in the case of Hamburger. Theorems 3.1 (odd case) and 3.9 (even case) from [16] state the equivalence between

- $\mathbf{V}$ is generated by a finite positive combination of Dirac measures,
- $\mathbf{V}$ is generated by a positive combination of $Rank(\mathbf{V})$ Dirac measures,
- $\mathbf{V}$ is positively recursively generated.

The first assertion and the reformulaion of the last one using Lemma 2.7 provides the result. The second assertion is used in the case of Stieltjes below. □

Proof in the case of Stieltjes: This is obtained by adapting the proofs of Theorems 5.1 (odd case) and 5.3 (even case) from [16].

First, writing $\mathbf{V}$ as a positive combination of $\mathbf{b}(s_i)$ over $s_i \in \mathbb{R}^+$ in the definitions of $M_{\mathbf{V}}^3$, $M_{\mathbf{V}}^4$ and $R_{\mathbf{V}}(s^j)$, one verifies that these matrices are positive definite and that (13g) holds.

In the other way, supose that $J_1 \leq J_2$ such that $M_{\mathbf{V}}^3$ and $M_{\mathbf{V}}^4$ are positive definite and that (13g) holds. Then $\mathbf{V}$ satisfies the conditions of the Hamburger problem and

$$\mathbf{V} = \sum_{i=1}^{J_1} \alpha_i \mathbf{b}(s_i).$$

By contradiction, suppose that one position $s_1 \leq 0$ is non-positive (the first without loss of generalities). Then, define its Lagrange polynomial l_1 and compute the moment with respect to $sl_1(s)^2$

$$l_1(s) = \prod_{i=2}^{J_1} \frac{s - s_i}{s_1 - s_i}, \qquad R_{\mathbf{V}}(sl_1(s)^2) = \sum_{i=1}^{J_1} \alpha_i s_i l_1(s_i)^2 = \alpha_1 s_1 \leq 0.$$

However, rewriting $l_1 = \boldsymbol{\beta}^T \mathbf{b}_{J_1-1} \in \mathbb{R}_{J_1-1}[X]$, this is also

$$R_{\mathbf{V}}(sl_1(s)^2) = \boldsymbol{\beta}^T M_{\mathbf{V}}^4 \boldsymbol{\beta}$$

which contradicts the positive definiteness of $M_{\mathbf{V}}^4$. Then all the $s_i > 0$, and $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m$.

Similarily, in the case $J_1 > J_2$, then the vector $R_{\mathbf{V}}(s\mathbf{b}_{N-1})$ (without the 0-th moment) satisfies the conditions of the Hamburger problem, and one has

$$R_{\mathbf{V}}(s\mathbf{b}_{N-1}) = \sum_{i=1}^{J_2} \alpha_i \mathbf{b}_{N-1}(s_i).$$

By contradiction, suppose that one position $s_1 \leq 0$ is non-positive (the first without loss of generalities). Then, the moment with repect to $(sl^1(s))^2$ reads

$$R_{\mathbf{V}}(s^2 l_1^2(s)) = R_{\mathbf{V}}(s(sl_1(s)^2)) = \sum_{i=1}^{J_2} \alpha_i s_i l_1(s_i)^2 = \alpha_1 s_1 \leq 0.$$

But again, $R_{\mathbf{V}}(s^2 l_1^2(s)) = \boldsymbol{\beta}^T M_{\mathbf{V}}^4 \boldsymbol{\beta}$ which contradicts is positive definiteness. Then all the $s_i > 0$, and

$$R_{\mathbf{V}}(s\mathbf{b}_{N-1}) = \sum_{i=1}^{J_2} \kappa_i s_i \mathbf{b}_{N-1}(s_i), \qquad \kappa_i = \frac{\alpha_i}{s_i}.$$

Therefore $\mathbf{V}$ is represented by

$$\mathbf{V} = \kappa_0 \mathbf{b}(0) + \sum_{i=1}^{J_2} \kappa_i \mathbf{b}(s_i),$$

where $\kappa_0 = R_{\mathbf{V}}(1) - \sum_{i=1}^{J_2} \kappa_i$. As $M^4_{\mathbf{V}}$ is positive definite, then its submatrix

$$R_{\mathbf{V}}(\mathbf{b}_1 \mathbf{b}_1^T) = \begin{pmatrix} R_{\mathbf{V}}(1) & \sum_{i=1}^{J_2} \kappa_i s_i \\ \sum_{i=1}^{J_2} \kappa_i s_i & \sum_{i=1}^{J_2} \kappa_i s_i^2 \end{pmatrix}$$

is positive definite. The coefficient $\sum \alpha_i s_i^2$ on the diagonal is strictly positive. Using a Cauchy-Schwartz inequality, the determinant of this matrix is positive if and only if $\kappa_0 > 0$. Therefore, there exists a positive discrete representing measure for $\mathbf{V}$. □

Remark 3.

- For all considered sets of integration E, one has

$$\mathcal{R}_{\mathbf{b}} \subset \mathcal{R}^m_{\mathbf{b}}.$$

  - Indeed, for Hausdorff TMP, if $M^1_{\mathbf{V}}$ and $M^2_{\mathbf{V}}$ are positive definite, then they are positive semi-definite.
  - For Hamburger TMP, if $M^1_{\mathbf{V}}$ is positive definite, then $J = K + 1$ and one has $M^1_{\mathbf{V}} = M^2_{\mathbf{V}}$ and $2J > Card(\mathbf{b})$.
  - For Stieltjes TMP, if $M^1_{\mathbf{V}}$ and $M^2_{\mathbf{V}}$ are positive definite, then $M^3_{\mathbf{V}}$ and $M^4_{\mathbf{V}}$ are positive definite and $L = Card(\mathbf{b})$.
- This provides especially that $\mathcal{R}^m_{\mathbf{b}}$ contains the open cone $\mathcal{R}_{\mathbf{b}}$ and is contained into its closure $\mathcal{R}^c_{\mathbf{b}}$. Especially, $\mathcal{R}_{\mathbf{b}}$ is dense in $\mathcal{R}^m_{\mathbf{b}}$ which is dense in $\mathcal{R}^c_{\mathbf{b}}$ and

$$\mathcal{R}_{\mathbf{b}} = int\left(\mathcal{R}^m_{\mathbf{b}}\right), \qquad \mathcal{R}^c_{\mathbf{b}} = \overline{\mathcal{R}^m_{\mathbf{b}}}. \tag{14}$$

- In the case of Hausdorff, $\mathcal{R}^m_{\mathbf{b}} = \mathcal{R}^c_{\mathbf{b}}$ is closed. This holds not in the other two cases, this can be exhibited by the following counterexample. Define

$$\forall \epsilon > 0, \qquad \mathbf{V}_\epsilon = \epsilon^2 \mathbf{b}_2(\epsilon^{-1}) \in \mathcal{R}^m_{\mathbf{b}_2}.$$

  Then at the limit

$$\mathbf{V}_0 := \lim_{\epsilon \to 0} \mathbf{V}_\epsilon = (0,0,1)^T \in \mathcal{R}^c_{\mathbf{b}_2} \backslash \mathcal{R}^m_{\mathbf{b}_2}.$$

  Indeed, $\mathbf{V}_0$ provides a positive semi-definite moment matrices

$$R_{\mathbf{V}_0}(\mathbf{b}_1 \mathbf{b}_1^T) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad R_{\mathbf{V}_0}\left(s \mathbf{b}_0 \mathbf{b}_0^T\right) = \begin{pmatrix} 0 \end{pmatrix},$$

  then $\mathbf{V}_0 \in \mathcal{R}^c_{\mathbf{b}_2}$ in both cases of Hamburger and Stieltjes. However, neither (12c) nor (13g) is satisfied ($\mathbf{V}_0$ is not recursively generated), then $\mathbf{V}_0 \notin \mathcal{R}^m_{\mathbf{b}_2}$ is not generated by a sum of Dirac measures in Hamburger and Stieltjes cases.

2.4. **Representation results on the boundary $\partial \mathcal{R}_{\mathbf{b}}$.** For the construction of the closure, we widely exploit the description of the boundary of the realizability domain ([20]) presented through the following propositions.

2.4.1. *Representation of* $\partial\mathcal{R}_{\mathbf{b}} \cap \mathcal{R}^m_{\mathbf{b}}$. First, a representation result for the vectors of $\partial\mathcal{R}_{\mathbf{b}} \cap \mathcal{R}^m_{\mathbf{b}}$ is recalled.

Proposition 2. *For all vector* $\mathbf{V} \in \partial\mathcal{R}_{\mathbf{b}} \cap \mathcal{R}^m_{\mathbf{b}}$*, there exists a unique representing measure for* $\mathbf{V}$*. This measure is given by*

$$\gamma = \sum_{i=1}^{J} \alpha_i \delta_{s_i}, \tag{15}$$

where $(\alpha_i, s_i)_{i=1,\dots,J} \in (\mathbb{R}^{*+} \times E)^J$ *are pairs of strictly positive weights and distinct positions in* E *and the number* J *is*

$$J = \min_j \left(rank(M^j_{\mathbf{V}}) \right),$$

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

This so-called atomic decomposition ([16]), in 1D, can be deduced as a corollary of the truncated Riesz-Haviland theorem, see *e.g.* [54, 30, 31, 16, 18] or surveys in [20, 37], or of Tchakaloff theorem, see *e.g.* [62, 6, 17, 37].

In the case of Hausdorff, this proposition is sufficient to describe all $\partial\mathcal{R}_{\mathbf{b}}$, since $\partial\mathcal{R}_{\mathbf{b}} \subset \mathcal{R}^m_{\mathbf{b}}$. In the other two cases, this describes only a part of the boundary. This part of the boundary is not exploited in the construction of the closures in the next section, but its description is provided for completeness and for further discussions.

2.4.2. *Representation of* $\partial\mathcal{R}_{\mathbf{b}} \backslash \mathcal{R}^m_{\mathbf{b}}$. In these cases, for the rest $\partial\mathcal{R}_{\mathbf{b}} \backslash \mathcal{R}^m_{\mathbf{b}}$, we exploit the closure relation (6e) to obtain a weaker representation result.

Proposition 3.

- *For all vector* $\mathbf{V} \in \mathcal{R}^c_{\mathbf{b}}$*, there exists a measure* γ^ϵ*, and its moments* $\mathbf{V}^\epsilon$*, of the form*

$$\gamma^\epsilon = \sum_{i=1}^{J} \alpha^\epsilon_i \delta_{s^\epsilon_i}, \qquad \mathbf{V}^\epsilon = \int_E \mathbf{b}(s) d\gamma^\epsilon(s) = \sum_{i=1}^{J} \alpha^\epsilon_i \mathbf{b}(s^\epsilon_i) \tag{16a}$$

where $J = \lfloor \frac{N}{2} \rfloor + 1$ *and* $(\alpha^\epsilon_i, s^\epsilon_i)_{i=1,\dots,J} \in (\mathbb{R}^{*+} \times E)^J$ *are pairs of strictly positive weights and distinct positions in* E*, depending on a parameter* $\epsilon > 0$*, such that*

$$\mathbf{V} = \lim_{\epsilon \to 0} \mathbf{V}^\epsilon. \tag{16b}$$

- *If* $\mathbf{V} \in \partial\mathcal{R}_{\mathbf{b}}$*, then there exists a decomposition of the form* (16) *such that*

$$\lim_{\epsilon \to 0} \alpha^\epsilon_1 = 0. \tag{17}$$

- *If* $\mathbf{V} \in \partial\mathcal{R}_{\mathbf{b}} \backslash \mathcal{R}^m_{\mathbf{b}}$*, then there exists a decomposition of the form* (16) *such that*

$$\lim_{\epsilon \to 0} (\alpha^\epsilon_1, |s^\epsilon_1|) = (0, +\infty), \qquad \lim_{\epsilon \to 0} \alpha^\epsilon_1 \mathbf{b}(s^\epsilon_1) \neq 0_{\mathbb{R}^{Card(\mathbf{b})}}. \tag{18}$$

Proof.

- One remarks that $\mathcal{R}^c_{\mathbf{b}} = \overline{\mathcal{R}^m_{\mathbf{b}}}$, then any vector $\mathbf{V} \in \mathcal{R}^c_{\mathbf{b}}$ is the limit of a certain $\mathbf{V}^\epsilon \in \mathcal{R}^m_{\mathbf{b}}$ when $\epsilon \to 0$.
  The existence of the (non-unique) representation (16a) for any point $\mathbf{V}^\epsilon \in \mathcal{R}^m_{\mathbf{b}}$ is proved in [16].

- Witout loss of generality, let us order $\alpha_1^\epsilon \leq \alpha_2^\epsilon \leq \dots \alpha_J^\epsilon$. By contradiction, suppose that $\lim_{\epsilon\to 0} \alpha_1^\epsilon > 0$.

  First, suppose that all $s_i := \lim_{\epsilon\to 0} s_i^\epsilon$ are different ($s_i \neq s_j$) and bounded ($|s_i| < +\infty$). Using the linearity of the Riesz function

$$R_{\mathbf{V}^\epsilon} = \sum_{i=1}^{J} \alpha_i^\epsilon R_{\mathbf{b}(s_i^\epsilon)},$$

  one easily verifies that all matrices (11-13) are positive definite at the limit $\epsilon \to 0$. Using Theorems 2.3, 2.4 and 2.5, then $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}$, therefore $\mathbf{V} \notin \partial\mathcal{R}_{\mathbf{b}}$.

  Second, suppose that $s_0 := \lim_{\epsilon\to 0} s_1^\epsilon = \lim_{\epsilon\to 0} s_2^\epsilon$ (or any other s_i^ϵ) are equal and bounded. Then

$$\begin{aligned} V &= \lim_{\epsilon\to 0}\left(\alpha_1^\epsilon \mathbf{b}(s_1^\epsilon) + \alpha_2^\epsilon \mathbf{b}(s_2^\epsilon) + \sum_{i=3}^{J} \alpha_i^\epsilon \mathbf{b}(s_i^\epsilon)\right) \\ &= \lim_{\epsilon\to 0}(\alpha_1^\epsilon + \alpha_2^\epsilon)\mathbf{b}(s_0) + \lim_{\epsilon\to 0}\sum_{i=3}^{J} \alpha_i^\epsilon \mathbf{b}(s_i^\epsilon) \\ &= \lim_{\epsilon\to 0}\sum_{i=1}^{J} \beta_i^\epsilon \mathbf{b}(s_i^\epsilon) \end{aligned}$$

  with $\beta_1^\epsilon = 0$, $\beta_2^\epsilon = \alpha_1^\epsilon + \alpha_2^\epsilon$ and $\beta_i^\epsilon = \alpha_i^\epsilon$ for $i > 2$, and we simply exploit this other representation of $\mathbf{V}$.

  Third, suppose that $|s_1^\epsilon| \to +\infty$ (or any other s_i^ϵ) is unbounded. As $\mathbf{b} = \mathbf{b}_N$, if $N > 1$, then $s^{2K} \in Span(\mathbf{b})$ for $K = \lfloor \frac{N}{2} \rfloor$. Then, $\lim_{\epsilon\to 0}(s_i^\epsilon)^{2K} > 0$ for all i and $\lim_{\epsilon\to 0}(s_1^\epsilon)^{2K} = +\infty$, then $R_{\mathbf{V}}(s^{2K}) = \lim_{\epsilon\to 0} R_{\mathbf{V}^\epsilon}(s^{2K}) = +\infty$ is unbounded and can not belong to $\partial\mathcal{R}_{\mathbf{b}}$. If $N = 1$, then the decomposition (16a) possesses a unique atom which is also unbounded.

  This three cases cover all the possibilities. Therefore, this contradicts that $\lim_{\epsilon\to 0} \alpha_1^\epsilon > 0$.
- Consider $\mathbf{V} \in \partial\mathcal{R}_{\mathbf{b}} \backslash \mathcal{R}_{\mathbf{b}}^m$, then it has a representation of the form (16-17). By contradiction, suppose that all such representations have no pair $(\alpha_i^\epsilon, s_i^\epsilon)$ satisfying (18). Then, one verifies again that
  - If all $\lim_{\epsilon\to 0} |s_i^\epsilon| < +\infty$ and $\lim_{\epsilon\to 0} \alpha_i^\epsilon < +\infty$ are bounded, then $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^m$
  - If some α_i^ϵ are unbounded or some $|s_i^\epsilon|$ are unbounded, but not those associated to $\lim_{\epsilon\to 0} \alpha_i^\epsilon = 0$, then $\mathbf{V}$ is unbounded.

  These all contradict $\mathbf{V} \notin \partial\mathcal{R} \backslash \mathcal{R}_{\mathbf{b}}^m$.

  Finally, the case $\lim_{\epsilon\to 0} \alpha_1^\epsilon \mathbf{b}(s_1^\epsilon) = 0_{\mathbb{R}^{Card(\mathbf{b})}}$ is rejected since it has no impact on the moment vector $\mathbf{V}$.

□

2.4.3. *Discussion on the non-uniqueness of the representations in $\mathcal{R}_{\mathbf{b}}^c$.* Even though any vector $\mathbf{V} \in \partial\mathcal{R} \cap \mathcal{R}_{\mathbf{b}}^m$ has a unique representing measure (15) (in $\mathcal{R}_{\mathbf{b}}^m$), the representation (16) (in $\mathcal{R}_{\mathbf{b}}^c$) remains non unique. Furthermore, the representations of the form (16) are not the only ones representing vectors of $\mathcal{R}_{\mathbf{b}}^c$. One could for instance construct ϵ dependent functions (in $L_N^1(E)^+$) which limit captures $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}^c$.

Representation results as those of the last subsections are often exploited to construct closure relations (see next section). Thus, the choice of the set in which such

a representation is chosen is important, especially because it impacts the existence and uniqueness of a representation on the boundary $\partial\mathcal{R}$. Indeed, a discrete measure representation (in $\mathcal{R}^m_{\mathbf{b}}$) exists and is unique over $\partial\mathcal{R}\cap\mathcal{R}^m_{\mathbf{b}}$ but does not exists over $\partial\mathcal{R}\backslash\mathcal{R}^m_{\mathbf{b}}$, while the representation (16) (in $\mathcal{R}^c_{\mathbf{b}}$) exists everywhere but is never unique.

Example. Consider $\mathbf{V}=(1,0,1,0,1)^T$ and $E=\mathbb{R}$. One verifies that

$$\mathbf{V}=\frac{\mathbf{b}_4(+1)+\mathbf{b}_4(-1)}{2}=\int_{\mathbb{R}}\mathbf{b}_4(s)d\gamma(s)\in\partial\mathcal{R}_{\mathbf{b}_4}\cap\mathcal{R}^m_{\mathbf{b}_4}, \tag{19a}$$
$$\gamma=\frac{\delta_1+\delta_{-1}}{2},$$

is the unique representing measure for $\mathbf{V}$. However, $\mathbf{V}$ is also represented *e.g.* by

$$\mathbf{V}=\lim_{\epsilon\to 0}\frac{\mathbf{b}_4(1)+\mathbf{b}_4(-1)+\epsilon^5\mathbf{b}_4(\epsilon^{-1})}{2}=\lim_{\epsilon\to 0}\int_{\mathbb{R}}\mathbf{b}_4(s)d\gamma^\epsilon(s)\in\mathcal{R}^c_{\mathbf{b}}, \tag{19b}$$
$$\gamma^\epsilon=\frac{\delta_1+\delta_{-1}+\epsilon^5\delta_{\epsilon^{-1}}}{2}.$$

These two representations provide different closures, i.e. moment of order $N+1=5$. Indeed, the fifth order moment of the unique representing measure γ from (19a) is 0, while the limit of the representation γ^ϵ from (19b) when $\epsilon\to 0$ would be $\frac{1}{2}$.

In the next section, we always choose the unique measure representation (15) along the boundary $\partial\mathcal{R}_{\mathbf{b}}\cap\mathcal{R}^m_{\mathbf{b}}$ and avoid considering the boundary $\partial\mathcal{R}_{\mathbf{b}}\backslash\mathcal{R}^m_{\mathbf{b}}$.

In the case of Hausdorff, the positions s_i^ϵ are bounded as $E=[-1,1]$. Then, in the representation (16), only the coefficients α_i^ϵ may be unbounded. However, if any $\alpha_i^\epsilon\to+\infty$, then the limit $\mathbf{V}$ is unbounded. This implies that the boundary $\partial\mathcal{R}$ in that case is only represented by a measure, which was shown to be unique in Proposition 2. This means that all representations (16) have the same measure (15) for limit in that case.

Vectors $\mathbf{V}\in\partial\mathcal{R}_{\mathbf{b}}\backslash\mathcal{R}^m_{\mathbf{b}}$ can not be represented by a measure. However, they remain at an arbitrarily small distance from a vector $\mathbf{V}^\epsilon$ represented by a sum of Diracs. One could think of using such an aritrarily close representation. In practice, this remains insufficient to construct closures over $\partial\mathcal{R}_{\mathbf{b}}\backslash\mathcal{R}^m_{\mathbf{b}}$ as some of these vectors can simply not have a bounded closure whatever representation is used. This is illustrated in the next section.

For these reasons, the case of Hausdorff, easier to deal with, is mainly focused on in the next sections, even if part of the present construction is general to all $E\subset\mathbb{R}$. The extension of this construction to such unbounded sets requires further techniques discussed in the conslusive Section 5.3.2 the analysis of which is postponed for future work.

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

$$\int_E \mathbf{b}(s)d\gamma(s) = \mathbf{f}. \tag{20}$$

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 (20).

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. Consider a function $\kappa : \mathbb{R}^{N+1} \to \mathbb{R}$.

- κ is a realizable closure if the function $(Id_{\mathbb{R}^{N+1}}, \kappa)$ sends $\mathcal{R}_{\mathbf{b}_N}$ into $\mathcal{R}_{\mathbf{b}_{N+1}}$.
- κ is a m-realizable closure if it is a realizable closure and $(Id_{\mathbb{R}^{N+1}}, \kappa)$ sends $\mathcal{R}^m_{\mathbf{b}_N}$ into $\mathcal{R}^m_{\mathbf{b}_{N+1}}$.

This definition corresponds to defining κ from a positive integrable function, resp. discrete measure, satisfying the moment constraints (20).

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 correpsonds to the moment of specific measures. In practice, we focus on two types of measures that we aim to retreive 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

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

- some purely anistropic 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

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

  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.

These two types of measures are sufficient to represent any vector in $\mathcal{R}^m_{\mathbf{b}}$, and therefore to construct m-closures. However, as discussed in Section 2.4, some vectors of $\mathcal{R}^c_{\mathbf{b}}$ can not be represented by measures and are therefore rejected by the present construction of a closure. One could think of using a representation (16) for vectors in $\partial\mathcal{R}_{\mathbf{b}}\backslash\mathcal{R}^m_{\mathbf{b}}$. This would not necessarily provide a closure, because some vectors of

$\partial\mathcal{R}_{\mathbf{b}}\backslash\mathcal{R}^m_{\mathbf{b}}$ have no representation of the form (16) with a bounded moment of order $N+1$.

Example. Consider $E=\mathbb{R}^+$, $\mathbf{b}(s)=\mathbf{b}_1(s)=(1,s)^T$ and $\mathbf{V}=(0,\ 1)^T$.
The matrices $M^1_{\mathbf{V}}=(0)\in\mathbb{R}^{1\times 1}$ and $M^2_{\mathbf{V}}=(1)\in\mathbb{R}^{1\times 1}$ from (13) are symmetric positive semi-definite, but $\mathbf{V}$ is not recursively generated (the 0-th moment is 0). Therefore $\mathbf{V}\in\partial\mathcal{R}_{\mathbf{b}}\backslash\mathcal{R}^m_{\mathbf{b}}$.
Now, consider a representation of the form (16), *i.e.* here

$$\mathbf{V}=\lim_{\epsilon\to 0}\int_{\mathbb{R}^+}\mathbf{b}(s)\alpha^\epsilon\delta_{s^\epsilon}(s),$$

where α^ϵ and s^ϵ are such that $\lim\limits_{\epsilon\to 0}\alpha^\epsilon=0$ and $\lim\limits_{\epsilon\to 0}\alpha^\epsilon s^\epsilon=1$. The closure would be given by

$$\lim_{\epsilon\to 0}\int_E s^2\alpha^\epsilon\delta_{s^\epsilon}(s)=\lim_{\epsilon\to 0}(\alpha^\epsilon s^\epsilon)s^\epsilon=+\infty.$$

Indeed, $\alpha^\epsilon s^\epsilon$ is bounded by construction, while s^ϵ is not according to Proposition 3.

For this reason, the boundary $\partial\mathcal{R}_{\mathbf{b}}\backslash\mathcal{R}^m_{\mathbf{b}}$ is always rejected for the construction of a closure in the present paper. In the numerical applications below, we only focus on the Hausdorff case for which $\partial\mathcal{R}_{\mathbf{b}}\backslash\mathcal{R}^m_{\mathbf{b}}=\emptyset$. The applications of the present closure to unbounded sets requires further investigation which are discussed in conclusions.

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

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

for which a m-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 $(\mathbf{V}_i,\ \kappa_i)\in\mathcal{R}^m_{\mathbf{b}_{N+1}}$ then

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

For such decompositions, we will exploit two types of vectors $\mathbf{V}\in\mathcal{R}^m_{\mathbf{b}_N}$ for which we know easily how to construct m-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 $f_{eq}\in L^1_N(E)^+$

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

  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 $\mathbf{V}_s \in \partial\mathcal{R}_{\mathbf{b}_N} \cap \mathcal{R}^m_{\mathbf{b}}$ **on the boundary of the realizability domain** represented by discrete measures. Using Proposition 2, there exists a unique representing measure γ_s for $\mathbf{V}_s$ which is singular over E, *i.e.*
$$\gamma_s = \sum_{i=1}^{J} \alpha_i \delta_{s_i} \in \mathcal{M}(E), \qquad \mathbf{V}_s = \int_E \mathbf{b}_N(s) d\gamma_s(s) = \sum_{i=1}^{J} \alpha_i \delta_{s_i}.$$
Thus, any m-realizable closure satisfies for such vector
$$\kappa_s(\mathbf{V}_s) = \int_E s^{N+1} d\gamma_s(s) = \sum_{i=1}^{J} \alpha_i s_i^{N+1}. \tag{21a}$$

In the following, we will focus on two m-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 ([34, 46, 57]) 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
$$\mathbf{V} = \alpha \mathbf{V}_1 + (1-\alpha)\mathbf{V}_2, \qquad \mathbf{V}_1 = \mathbf{V} = \mathbf{V}_2,$$
where the computation of α will be detailed 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 boudary of the realizability domain. In practice, using the notations (11-13), we chose these closures $\kappa_i(\mathbf{V}_i)$ such that
$$M^i_{(\mathbf{V}_i,\kappa_i)} \quad \text{is positive semi-definite and singular.}$$
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.*
$$\mathbf{V}_{eq} = \int_{-1}^{+1} \mathbf{b}_N(s) f_{eq}(s) ds, \qquad \kappa_{eq} = \int_{-1}^{+1} s^{N+1} f_{eq}(s) ds,$$
typically for the given function $f_{eq} = 1$. Thus α is defined by
$$\kappa_{K_N}(V_{eq}) = \kappa_{eq} \quad \Rightarrow \quad \alpha = \frac{\kappa_{eq} - \kappa_2(\mathbf{V}_{eq})}{\kappa_1(\mathbf{V}_{eq}) - \kappa_2(\mathbf{V}_{eq})}.$$

Remark 4. In practice, the Kershaw method recovers the exact value of the closure $\kappa_{K_N} = \kappa_{eq}$ when $\mathbf{f} = \mathbf{V}_{eq}$, but it does not retreive 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
$$\gamma_{K_N} = \gamma_1\alpha + \gamma_2(1-\alpha), \qquad \gamma_1 = \sum_{i=1}^{J} \alpha_i \delta_{s_i}, \qquad \gamma_2 = \sum_{i=J+1}^{2J} \alpha_i \delta_{s_i}, \tag{22}$$
where γ_1 and γ_2 are the unique representing measure for $(\mathbf{V}, \kappa_1) \in \partial\mathcal{R}^m_{\mathbf{b}_{N+1}}$ and $(\mathbf{V}, \kappa_2) \in \partial\mathcal{R}^m_{\mathbf{b}_{N+1}}$. Following Proposition 2, the number of Diracs in this K_N representation can reach at most $2J = 2\min_i rank(M^i_{\mathbf{V}})$, *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 [16, 18, 17], 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 5. 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 $N+1$. However such an extension technique for Hamburger ($E = \mathbb{R}$) or Stieltjes ($E = \mathbb{R}^+$) problems would provide a unique closure κ since this coefficient appears only in one matrix through (12a) or through (13).

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 $\mathbf{V}_s \in \partial\mathcal{R}_{\mathbf{b}}$ on the boundary of the realizability domain.
Decomposition of $\mathbf{V}$: Here, we decompose

$$\mathbf{V} = \bar{x}\mathbf{V}_{eq} + \mathbf{V}_s, \tag{23}$$

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.

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

   This part is computed in Subsection 3.4. In this paper, we will only focus on the cases where this projection $\mathbf{V} - \bar{x}\mathbf{V}_{eq} \in \partial\mathcal{R}_{\mathbf{b}} \cap \mathcal{R}_{\mathbf{b}}^m$ is represented by a discrete measure. Of course, such a requirement only holds under condition over $\mathbf{V}$ and $\mathbf{V}_{eq}$. This is discussed in Section 5.3.2.
2. **Singular part**: Compute a m-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. This construction of κ_s is restricted to moments in $\mathbf{V} - \bar{x}(\mathbf{V})\mathbf{V}_{eq} \in \partial\mathcal{R}_{\mathbf{b}} \cap \mathcal{R}_{\mathbf{b}}^m$.

Remark 6. 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.

Representation of the Π_N closure: Formally, the Π_N model corresponds to approximate the kinetic distribution f by the discrete measure

$$\gamma_{\Pi_N}(s) = \bar{x} f_{eq}(s) ds + \gamma_0, \qquad \gamma_0 = \sum_{i=1}^{J} \alpha_i \delta_{s_i}, \tag{24}$$

where γ_0 is the unique representing measure for $\mathbf{V} - \bar{x}\mathbf{V}_{eq} \in \partial\mathcal{R}_{\mathbf{b}_N} \cap \mathcal{R}^m_{\mathbf{b}}$. Following Proposition 2, the number of Diracs can reach at most $J = \min_i rank(M^i_{\mathbf{V}-\bar{x}\mathbf{V}_{eq}})$, *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 (22) 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 $\mathbf{V} \in \mathcal{R}^m_{\mathbf{b}}$ 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 $\mathbf{V}_s \in \partial\mathcal{R}_{\mathbf{b}}$.

Proposition 4. *For all vectors $\mathbf{V} \in \mathcal{R}^c_{\mathbf{b}}$, there exists a unique decomposition of the form*

$$\mathbf{V} = \bar{x}\mathbf{V}_{eq} + \mathbf{V}_s. \tag{25a}$$

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

$$\tilde{M}^i_{\mathbf{V}} := (M^i_{\mathbf{V}_{eq}})^{-1/2} M^i_{\mathbf{V}} \left((M^i_{\mathbf{V}_{eq}})^{-1/2} \right)^T, \tag{25b}$$

$$\bar{x} = \min_i \min Sp\left(\tilde{M}^i_{\mathbf{V}}\right), \quad \mathbf{V}_s = \mathbf{V} - \bar{x}\mathbf{V}_{eq}, \tag{25c}$$

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 $M^i_{\mathbf{V}}$ are defined in (11), (12a) 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 $\mathbf{V}_{eq} \in \mathcal{R}_{\mathbf{b}}$, then its moment matrices $M^i_{\mathbf{V}_{eq}}$ are symmetric positive definite.

If $\mathbf{V} \in \partial\mathcal{R}_{\mathbf{b}}$, then $\mathbf{V} = \mathbf{V}_s + \bar{x}\mathbf{V}_{eq}$ with $\mathbf{V}_s = \mathbf{V} \in \partial\mathcal{R}_{\mathbf{b}}$ and $\bar{x} = 0 \in \mathbb{R}^+$.

Otherwise, $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}$ and its moment matrices (11) are positive definite. Defining (see Fig. 1)

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

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

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

then especially $-\mathbf{W}(z) \in \mathcal{R}^c_{\mathbf{b}}$. 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 $\mathbf{W}(\bar{x}) \in \partial\mathcal{R}_{\mathbf{b}}$.

Uniqueness of the decomposition:
Suppose that there exists two points $0 \leq x_1 < x_2$ such that $\mathbf{W}(x_1) \in \partial\mathcal{R}_{\mathbf{b}}$ and

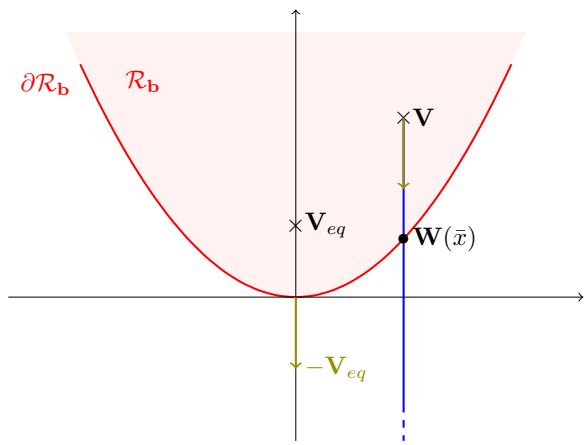

FIGURE 1. Schematic representation of a ray starting at a point $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}$ directed by $-\mathbf{V}_{eq}$ and crossing $\partial\mathcal{R}_{\mathbf{b}}$ in $\mathbf{W}(\bar{x})$.

$\mathbf{W}(x_2) \in \partial\mathcal{R}_{\mathbf{b}}$. By convexity of $\mathcal{R}_{\mathbf{b}}^c$, we have $\mathbf{W}(x) \in \mathcal{R}_{\mathbf{b}}^c$ for all $x \in [x_1, x_2]$. However,

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

with $\alpha \in [0,1]$. As $\alpha(x_2 - x_1) > 0$, then $\mathbf{W}(x)$ is a positive combination of $\mathbf{W}(x_2) \in \partial\mathcal{R}_{\mathbf{b}}$ and of $\mathbf{V}_{eq} \in \mathcal{R}_{\mathbf{b}} = int(\mathcal{R}_{\mathbf{b}}^c)$, then $\mathbf{W}(x) \in \mathcal{R}_{\mathbf{b}}$ which contradicts with $\mathbf{W}(x_1) \in \partial\mathcal{R}_{\mathbf{b}}$. Thus $\bar{x} > 0$ is unique.

Computation of $\bar{x}$ and $\mathbf{V}_s$:
Remark that $M^i_{\mathbf{V}_{eq}}$ is symmetric positive definite, then invertible and we may define $(M^i_{\mathbf{V}_{eq}})^{-1/2}$ using its diagonalization. Based on the characterizations of Theorems 2.3, 2.4 and 2.5, and on the construction (25b), we observe that

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

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

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

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

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

□

In all the following, the coefficient $\bar{x}$ is defined as (25c) where $(M^i_{\mathbf{V}_{eq}})^{-1/2}$ is the inverse of the Cholesky decomposition of $M^i_{\mathbf{V}_{eq}}$. Remark that since $\mathbf{V}_{eq} \in \mathcal{R}_{\mathbf{b}}$ is in the interior of the realizability domain, then the matrices $M^i_{\mathbf{V}_{eq}}$ are symmetric positive definite and their Cholesky decomposition exist.

3.5. **Computation of the closure on the boundary $\partial\mathcal{R}_{\mathbf{b}}\cap\mathcal{R}_{\mathbf{b}}^m$: Computation of the singular part.** Naively, one could compute all the coefficients α_i and s_i, and then reinject it in (21) to obtain κ_s. However, this would not be very efficient from a numerical point of view. We focus here on the computation in $\partial\mathcal{R}_{\mathbf{b}}\cap\mathcal{R}_{\mathbf{b}}^m$ and its representing measure.

3.5.1. *Computation of the closure κ_s in $\partial\mathcal{R}_{\mathbf{b}}\cap\mathcal{R}_{\mathbf{b}}^m$.* As we only look for one higher order moment, it is cheaper to exploit the principle of flat extension proposed by [16, 20] (available in $\partial\mathcal{R}_{\mathbf{b}}\cap\mathcal{R}_{\mathbf{b}}^m$), which was also implicitly exploited in the construction of Kershaw K_N closure ([34, 46, 57]). 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 $(\mathbf{V}_s,\kappa_s)$ and therefore the matrices $M^i_{(\mathbf{V}_s,\kappa_s)}$ 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 $N=2K$:

$$R_{(\mathbf{V}_s,\kappa_s)}((1+s)\mathbf{b}_K\mathbf{b}_K^T)=\begin{pmatrix} R_{\mathbf{V}_s}((1+s)\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T) & R_{\mathbf{V}_s}((1+s)s^K\mathbf{b}_{K-1}) \\ R_{\mathbf{V}_s}((1+s)s^K\mathbf{b}_{K-1})^T & R_{\mathbf{V}_s}(s^{2K})+\kappa_s \end{pmatrix}, \tag{26a}$$

$$R_{(\mathbf{V}_s,\kappa_s)}((1-s)\mathbf{b}_K\mathbf{b}_K^T)=\begin{pmatrix} R_{\mathbf{V}_s}((1-s)\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T) & R_{\mathbf{V}_s}((1-s)s^K\mathbf{b}_{K-1}) \\ R_{\mathbf{V}_s}((1-s)s^K\mathbf{b}_{K-1})^T & R_{\mathbf{V}_s}(s^{2K})-\kappa_s \end{pmatrix}. \tag{26b}$$

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

$$\kappa_s = -R_{\mathbf{V}_s}(s^{2K})+R_{\mathbf{V}_s}((1+s)s^K\mathbf{b}_{K-1})^T R_{\mathbf{V}_s}((1+s)\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T)^+ R_{\mathbf{V}_s}((1+s)s^K\mathbf{b}_{K-1}) \tag{27a}$$

$$= R_{\mathbf{V}_s}(s^{2K})-R_{\mathbf{V}_s}((1-s)s^K\mathbf{b}_{K-1})^T R_{\mathbf{V}_s}((1-s)\mathbf{b}_{K-1}\mathbf{b}_{K-1}^T)^+ R_{\mathbf{V}_s}((1-s)s^K\mathbf{b}_{K-1}), \tag{27b}$$

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

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

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

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

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 7. These computations can be the source of roudoff 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 occure 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).
- Similarily, in the Hausdorff case, the closure κ_s is obtained equivalently by two formula (27), in the applications below, again to smooth roudoff 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 in $\partial\mathcal{R}_{\mathbf{b}}\cap\mathcal{R}_{\mathbf{b}}^m$.* 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 closue. 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 $M^i_{\mathbf{V}_s}$. 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 $M^i_{\mathbf{V}_s}$, *i.e.*

$$M^i_{\mathbf{V}_s}\mathbf{X} = 0_{\mathbf{b}_K}.$$

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

$$\mathbf{X}^T M^i_{\mathbf{V}_s}\mathbf{X} = \int_E p_i(s)\left(\mathbf{b}(s)^T\mathbf{X}\right)^2 d\gamma(s) = 0,$$

where p_i are the polynomial 1, $(1 \pm s)$, or $(1-s^2)$ depending on the case considered, associated to the matrix $M^i_{\mathbf{V}}$, *i.e.* from formula (11-13). Especially, p_i is non-negative on E. Since γ is non-negative and $p(s)\left(\mathbf{b}(s)^T\mathbf{X}\right)^2$ 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

$$Supp(\gamma) = \bigcap_i \bigcap_{M_{\mathbf{V}^i_s}\mathbf{X}=0} Z\left(p\mathbf{b}\mathbf{X}\right), \tag{28}$$

where $Supp(\gamma)$ 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,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Computation of the coefficients α_i. Finally, in order to retreive 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

$$\frac{\mathbf{f}_i^{n+1} - \mathbf{f}_i^n}{\Delta t} + \frac{\mathcal{F}_{i+\frac{1}{2}}^n - \mathcal{F}_{i-\frac{1}{2}}^n}{\Delta x} = K\left(\frac{R_{\mathbf{f}_i^n}(1)}{2}\mathbf{V}_{iso} - \mathbf{f}_i^{n+1}\right), \tag{29a}$$

$$\mathcal{F}_{i+\frac{1}{2}} = \frac{1}{2}\left[\mathbf{F}(\mathbf{f}_{i+1}^n) + \mathbf{F}(\mathbf{f}_i^n) - (\mathbf{f}_{i+1}^n - \mathbf{f}_i^n)\right]. \tag{29b}$$

One easily proves using standard techniques (see *e.g.* [47, 4, 56, 50, 36]) 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 [34, 46, 57]) on four test cases. The first one is simply meant to study quantitatively the accuracy of the methods. The next two 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. The last test case is a rather elementary test that involving a small perturbation of an equilibrium function. Such a distribution is well-captured only by the Π_N models.

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

$$\partial_t f_n + s\partial_x f_n = -Kf_n, \tag{30a}$$

$$f_n(x) = f(x) \quad \text{on the boundaries } x = 0 \text{ and } x = L, \tag{30b}$$

$$\partial_t f_s + s\partial_x f_s = K\left(\frac{1}{2}\int_{-1}^{+1} f_s(s) + f_n(s)ds - f_s\right), \tag{30c}$$

$$f_s(x) = 0 \quad \text{on the boundaries } x = 0 \text{ and } x = L. \tag{30d}$$

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 (29). 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 [41, 63, 32, 8, 7, 14, 13, 49].

For the third 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 treatement 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.

The last test case corresponds to a kinetic which has an analytical solution, and therefore needs no additional numerical treatement.

The simulations are performed in `Python` using standart `numpy.linalg` functions ([48]), i.e. qr_multiply for the pseudo-inverse and eigh for the minimum eigvenvalue. The advantages of the present construction relies not in the numerical efficiency, but in the generality of models it can be applied to. Here, no particular code optimization was performed for any of the methods to accelerate the simulations. Especially, the K_N and Π_N closures are computed by passing from vector storage of moments to matrices (26) storage, instead of computing compressed closure formulae at fixed N as in [46, 57]. Such a implementation makes those simulations much longer and the K_N, resp. Π_N, simulations below are around nine times, resp. thirteen times, longer the P_N simulations for the same moment order N.

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 $\partial\mathcal{R}_{\mathbf{b}}$.

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 (30) 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, P_7 and the reference solution of (30) (first equation solved analytically, second one solved with P_{24}) are plotted on Fig. 2 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. 3.

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 (30) is the sum of the propagation of a Dirac distribution f_n with a regular one f_s.

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

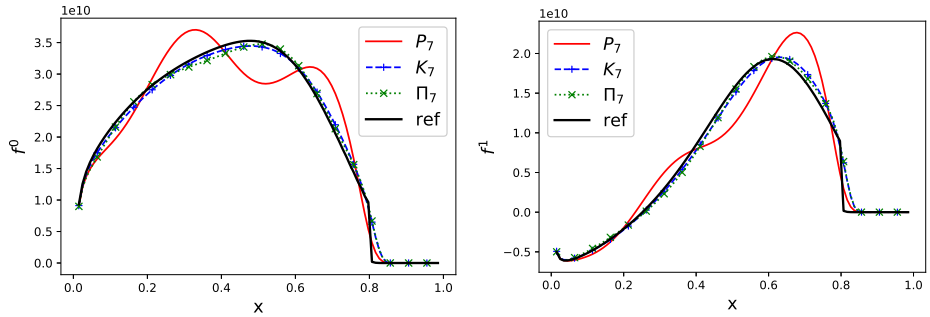

FIGURE 2. Moments of order 0 (left) and 1 (right) obtained with P_7, K_7, Π_7 and reference solution for the simple beam test case.

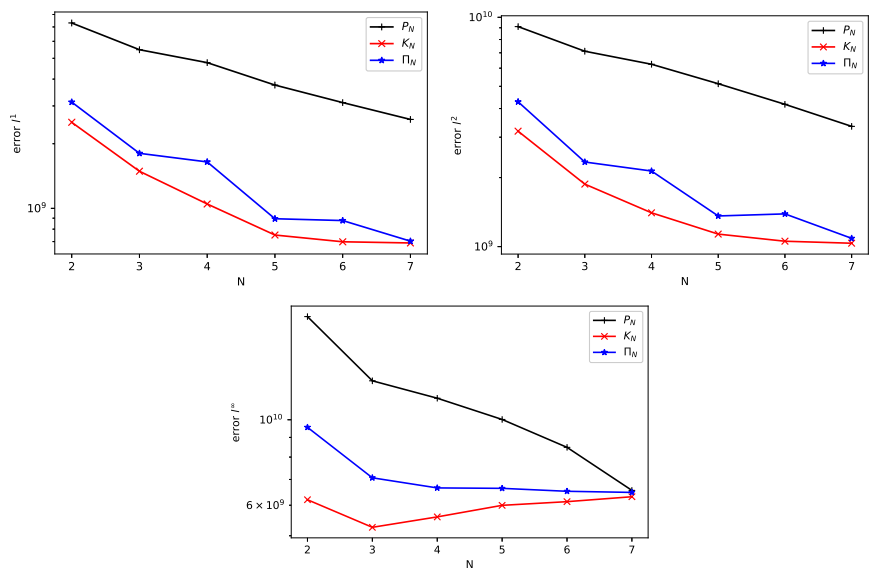

FIGURE 3. Discrete l^1 (top left), l^2 (top right) and l^∞ (bottom) errors on the moment of order 0 compared to a reference solution for the P_N, K_N and Π_N as a function of N for the simple beam test case.

and l^2 and l^∞ norms (the solution being smooth enough for that), but the odd order ones are more accurate. These discrepencies are due to the numerical methods used for the computation of the closure for several reasons which are mainly related to the moment matrices $R_{\mathbf{V}}((1 \pm s)\mathbf{b} \otimes \mathbf{b})$, $R_{\mathbf{V}}((1 - s^2)\mathbf{b} \otimes \mathbf{b})$ 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 $R_{\mathbf{V}}((1 \pm s)\mathbf{b} \otimes \mathbf{b})$, $R_{\mathbf{V}}((1 - s^2)\mathbf{b} \otimes \mathbf{b})$ 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, $R_{\mathbf{V}}((1-s)\mathbf{b}\otimes\mathbf{b})$ should have a smaller dimension than the other matrices (due to the position of the Dirac). Thus, the numerical errors producted during the computations of the different moment matrices may affect differently the QR algorithm used.

- Similarily, the formula (26) and (27) are simply different between odd and even orders and this may affect differently the numerical accuracy.

4.2. **Double beam test.** This test case ([28, 51, 50]) consists in having two beams of particles cross each other. As beams are used, purely anisotrpic 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

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

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. 4. The results with the reference solution of (30) (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. 5.

On this test case, the realizable closures are again more accurate than the polynomial 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.

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. 6. These measures are given by (22) for K_N, (24) for Π_N and by a basic polynomial reconstruction for P_N. These are compared to reconstruction obtained with the reference solution given by the sum of the analytical solution of (30a) (here composed of two Diracs of directions $s = \pm 1$) and of the P_{24} solution of (30c). For K_N and Π_N, these representing measures are composed of a regular

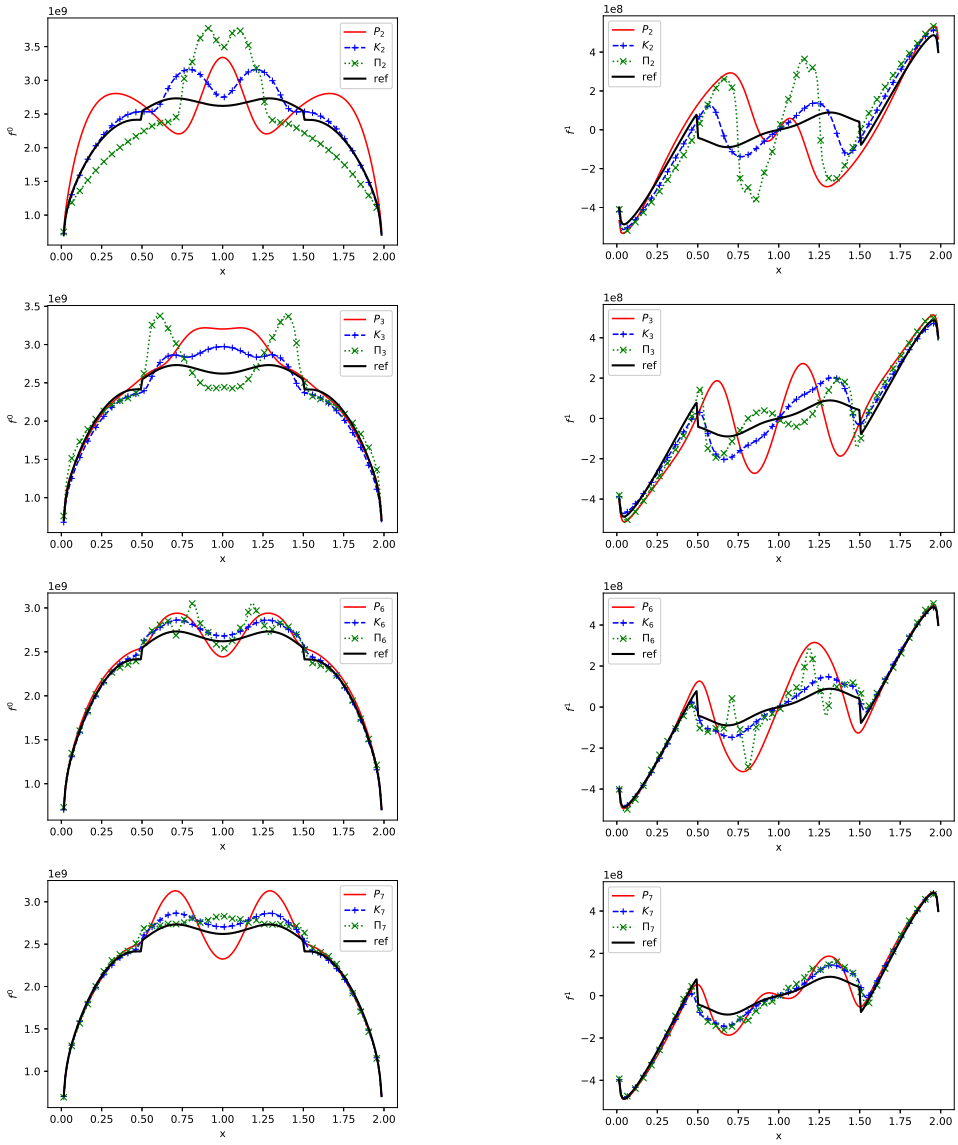

FIGURE 4. Moments of order 0 (left) and 1 (right) obtained with P_N, K_N, Π_N for $N = 2$ (first line), $N = 3$ (second line), $N = 6$ (third line), $N = 7$ (fourth line) and reference solution for the double beam test case.

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.

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.

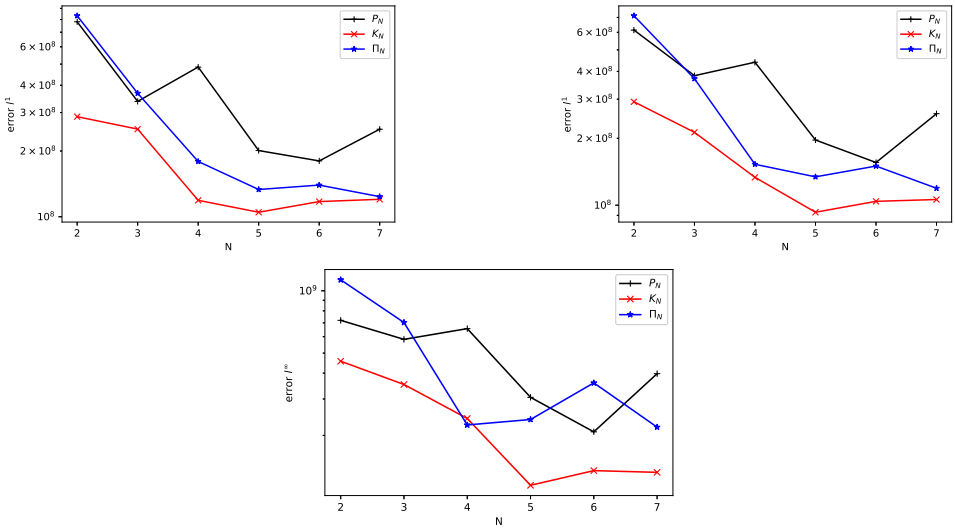

FIGURE 5. Discrete l^1 (top left), l^2 (top right) and l^∞ (bottom) errors on the moment of order 0 compared to a reference solution for the P_N, K_N and Π_N as a function of N for the double beam test case.

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. However, their positions and amplitude seem to oscillate around a stable value near the position of the maximum of the reference distribution. Similarily, the constant value $\bar{x}$ seems to stabilize below the mean value of the reference such that the mass of the two distributions, i.e. f^0, are close.

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 oversestimated 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.** This test case ([12, 25, 28]) 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 collision parameter $K = 1$, the distribution tends to an anisotropic distribution pointed away from the center.

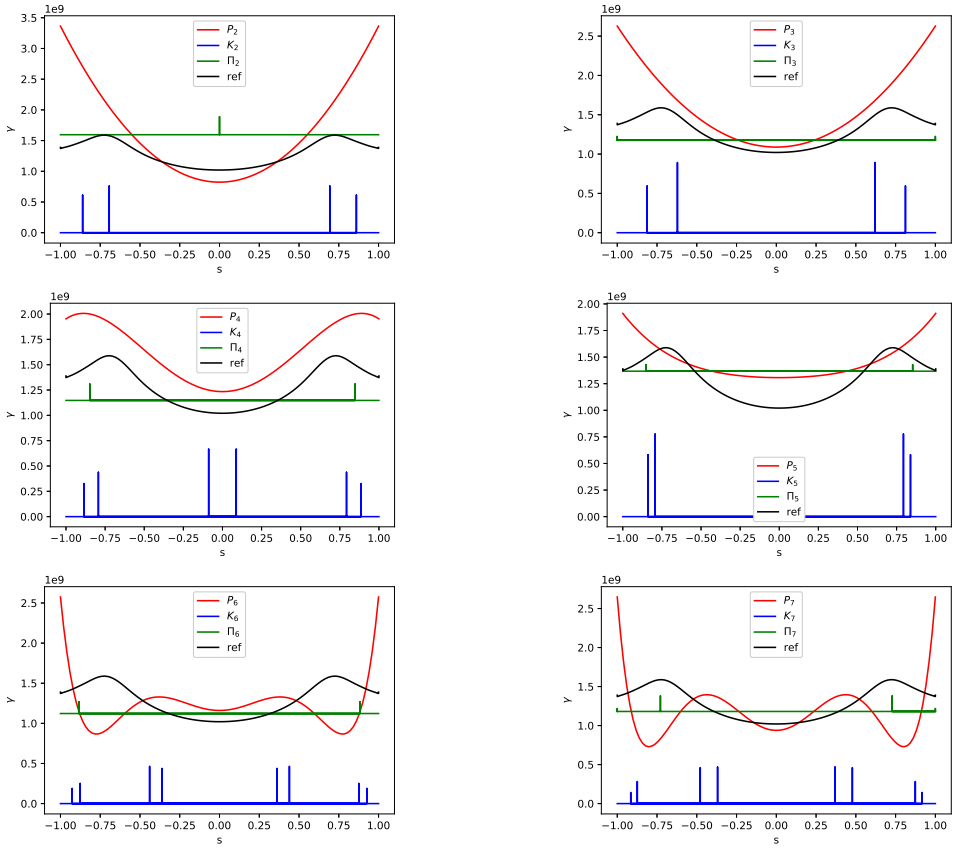

FIGURE 6. Representation of the measures representing the vector $\mathbf{f}(x = L/2)$ with P_N, K_N and Π_N models with $N = 2$ (top left), 3 (top right), 4 (middle left), 5 (middle right), 6 (bottom left), 7 (bottom right) for the double beam test case.

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

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

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. 7. 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. 8.

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

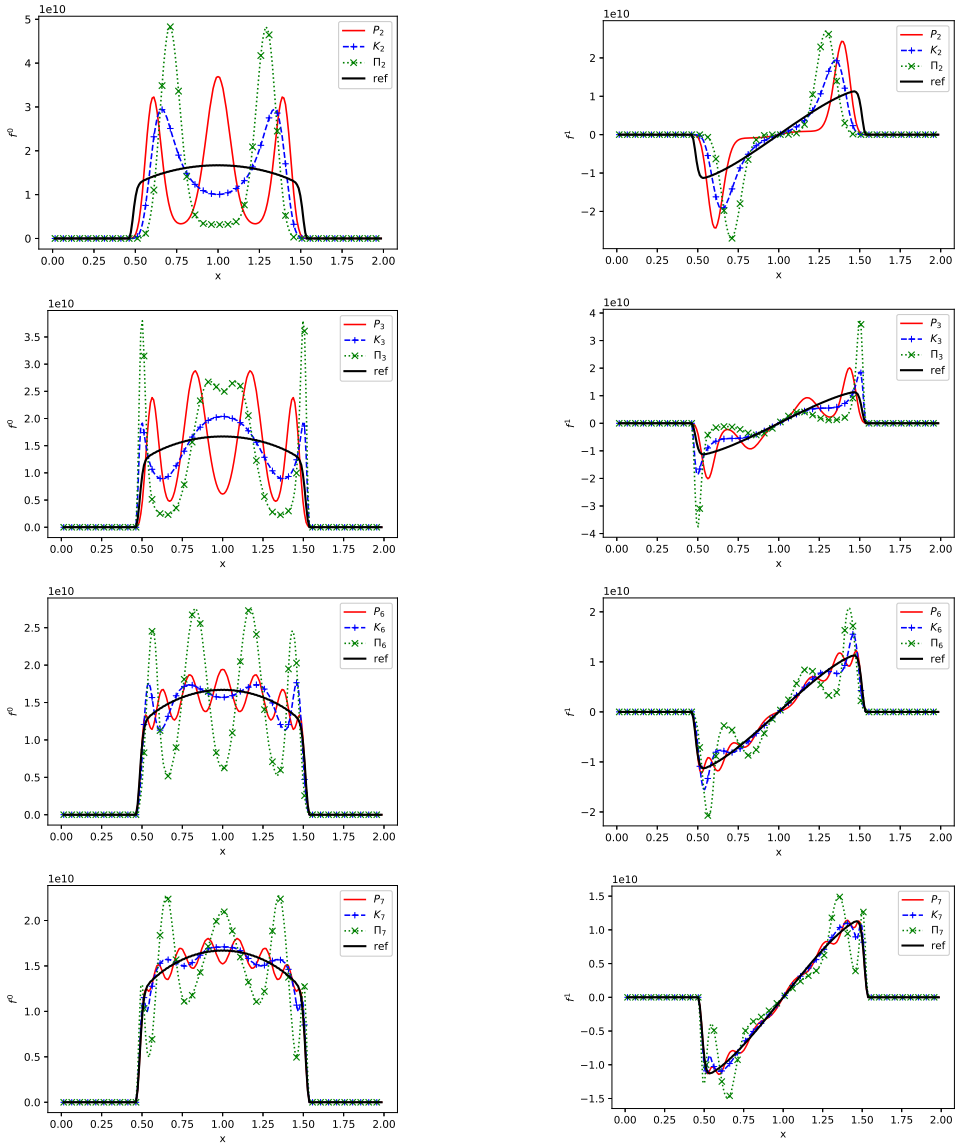

FIGURE 7. Moments of order 0 (left) and 1 (right) obtained with P_N, K_N, Π_N for $N = 2$ (first line), $N = 3$ (second line), $N = 6$ (third line), $N = 7$ (fourth line) and a reference P_{24} solution for the point source test case.

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 dicrease in amplitude. Their amplitudes are higher with Π_N approximation.

As for the last test case, in order to study this phenomenum, we plot on Fig. 9 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. For this test case, even though the initial condition isotropic in $x = L/2$, the analytical solution of (30a) turns into a beam

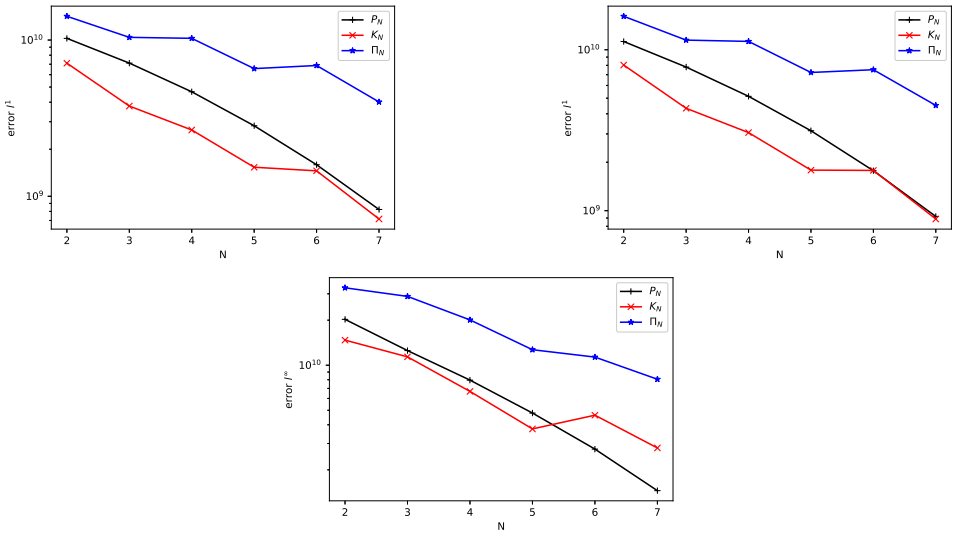

FIGURE 8. Discrete l^1 (top left), l^2 (top right) and l^∞ (bottom) errors on the moment of order 0 compared to a reference P_{24} simulation for the P_N, K_N and Π_N as a function of N for the point source test case.

distribution of direction $(L/2 - x)/t$ at positive times. Thus, in $x = L/2$, this is a Dirac in $s = 0$.

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 $s = 0$ (as for the reference distribution) than with K_N. The valus of $\bar{x}$ again stabilizes below the mean value of the reference distribution. At this local level, the Π_N distribution seems closer to the reference solution, while it is not at the moment level on Fig. 7.

Again, we may interprete 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 classical 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 $\partial\mathcal{R}_{\mathbf{b}}$. In practice, the moment matrix associated to K_7 closure should be singular (*i.e.* on the boundary of $\partial\mathcal{R}_{\mathbf{b}}$), 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

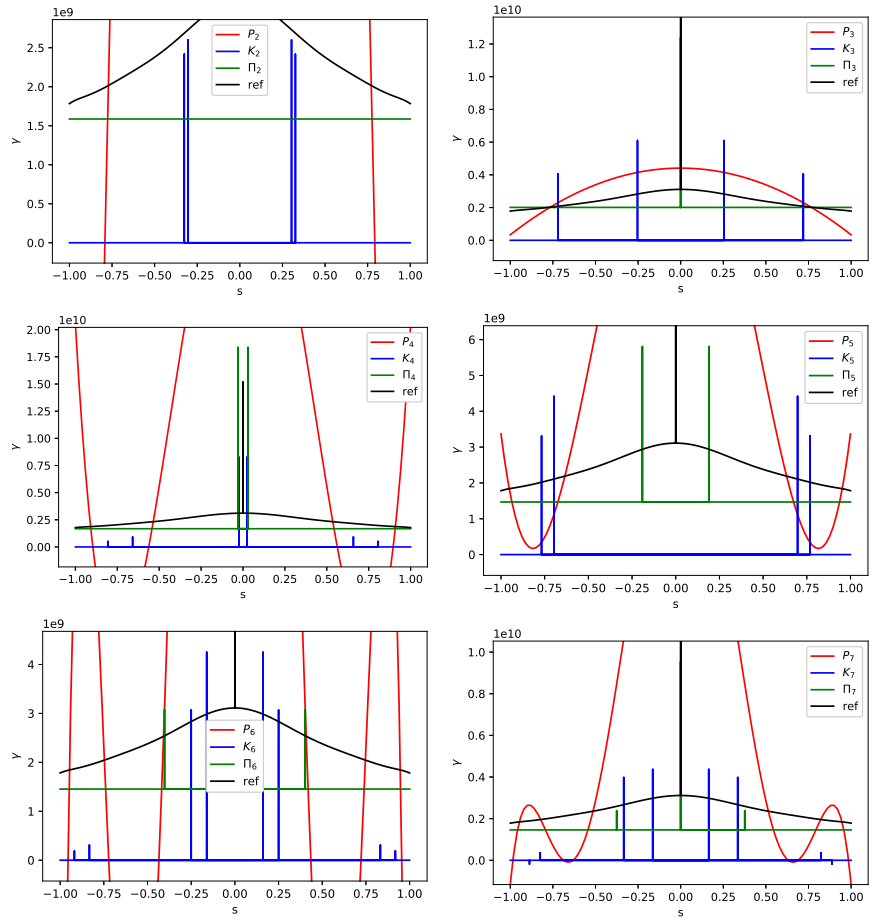

FIGURE 9. Representation of the measures representing the vector $\mathbf{f}(x = L/2)$ with P_N, K_N and Π_N models for $N = 2$ (top left), 3 (top right), 4 (middle left), 5 (middle right), 6 (bottom left), 7 (bottom right) for the point source test case.

of the method. Remark although that the closure remains well-defined and it affects not the accuracy of the simulation.

4.4. **Riemann problem with a small perturbation of the equilibrium.** While the previous tests are commonly performed to study various moment models, this last one is meant to exhibit the advantages of the Π_N models. It consists of a Riemann problem given by the initial conditions by

$$\mathbf{f}(t = 0, x) = 10^{12} \left[20\mathbf{V}_{iso} + 0.5\mathbf{b}(1)\mathbf{1}_{\mathbf{R}^-}(x - L/2)\right],$$

and boundary conditions

$$\mathbf{f}(t, x = 0) = 10^{12} \left[20\mathbf{V}_{iso} + 0.5\mathbf{b}(+1)\right], \qquad \mathbf{f}(t, x = L) = 10^{12} \times 20\mathbf{V}_{iso}.$$

One remarks that the initial condition corresponds to the moments $20\mathbf{V}_{iso}$ of the equilibrium function $20f_{eq}$ slightly perturbed by a beam distribution for $0.5\delta_1(s)$ for all $x < L/2$.

The spatial domain is of size $L = 2$ meshed with 201 cells. The final time is chosen to be $T_{\max} = 0.8$ such that the beam $\mathbf{b}(+1)$ on the left side of the interface has not yet reach the boundary at the end of the simulation. This test case is collisionless $K = 0$ such that the solution to (1) is analytical, it yields

$$f(t, x, s) = 10^{12} \left[20 + 0.5\delta_1(s)\mathbf{1}_{\mathbb{R}^-}(x - t)\right].$$

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. 10. The moments of the analytical solution are also plotted. The exact solution of this PDE is a wave propargating at velocity 1. The distribution at $x = L/2$ is represented on Fig. 11. The solution obtained with the Π_N models even at vey low order $N = 2$ are on top of the analytical solution. The amplitude and the velocity of the discontinuity is perfectly modeled with Π_N approximations. This can be expected for this test case as this solution is one of those Π_N distributions. It is indeed a positive combination of the equilibrium function and a Dirac.

The solutions obtained with P_N and K_N models are not as accurate. They do not capture the right location and amplitude of the analytical wave, and furthermore create artificial ones. Indeed, at each N, one solves an hyperbolic system of $N + 1$ equations. The initial discontinuity creates $N + 1$ waves with different amplitudes and velocities and none of them corresponds to the exact one. The wave structure of the Π_N solution is briefly discussed in Section 5.2 below, but its complete analysis is postponed to futur study. Note that the fluctuations in those plots actually have a small amplitude, it is around 1% of the mean f^0, i.e. $\max f^0 \approx 4.07 \times 10^{13}$ and $\min f^0 \approx 4.03 \times 10^{13}$. Quantitatively, the relative error is not so large, but qualitatively, P_N and K_N models are not able to capture the propagation of such a small perturbation while Π_N models are.

5. Conclusing 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 basic 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 [51, 40, 55] over S^2. For moments up to order 2, the associated realizability domain has been characterized ([34]). 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 $\mathcal{R}^c_{\mathbf{b}_2}$. This non-linear realizability constraint on the numerical solution also affects the robustness of the numerical method. This holds when using

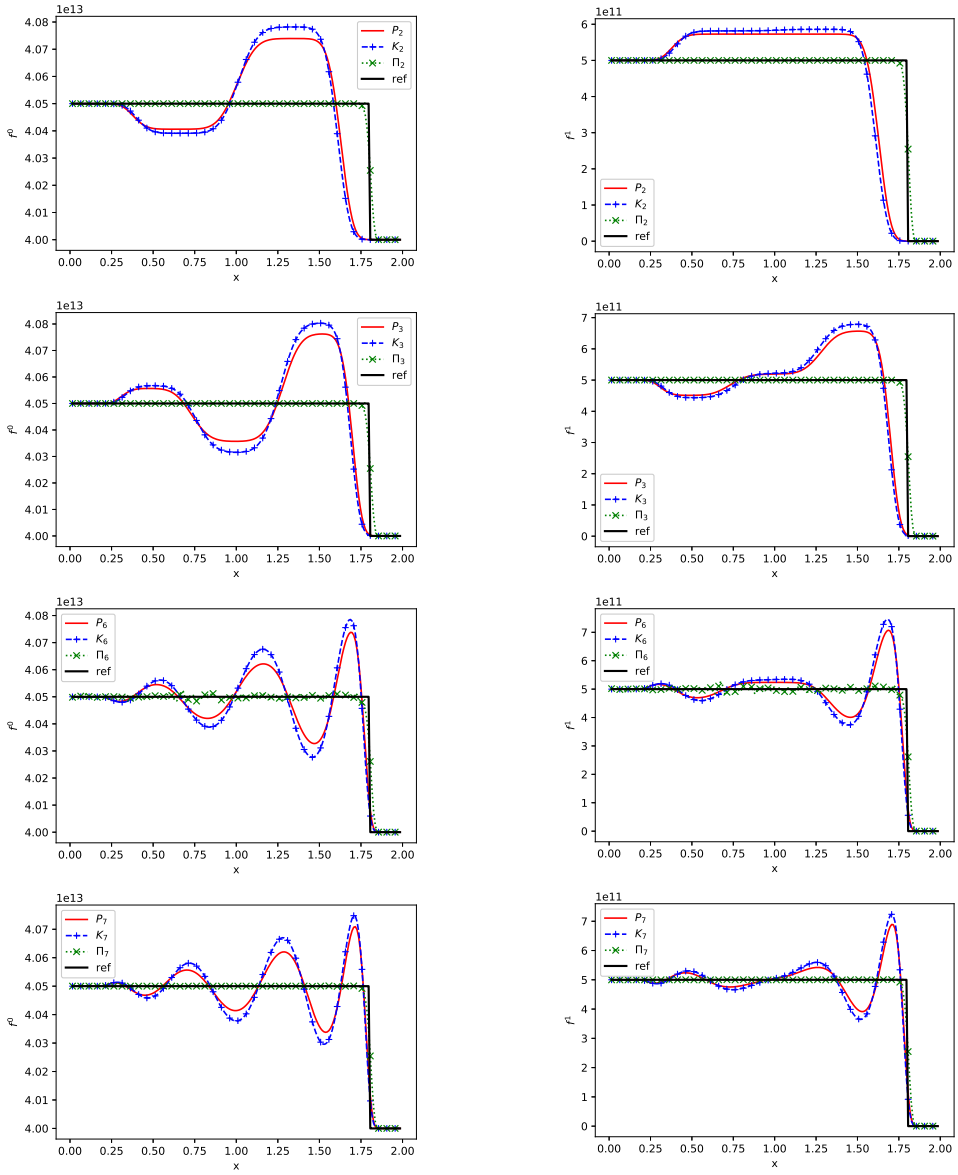

FIGURE 10. Moments of order 0 (left) and 1 (right) obtained with P_N, K_N, Π_N for $N = 2$ (first line), $N = 3$ (second line), $N = 6$ (third line), $N = 7$ (fourth line) and those of the analytical solution for the Riemann problem.

basic first order schemes, such as (29), to the moment system. It affects even more the robustness of a code when using high order numerical schemes, which generally require a particular treatement to preserve realizability (see *e.g.* [28, 5, 58]).

The present construction remains valid out of the realizability domain and can therefore be exploited to extend the 3D second order closures [51, 40, 55] out of $\mathcal{R}_{\mathbf{b}}^{c}$.

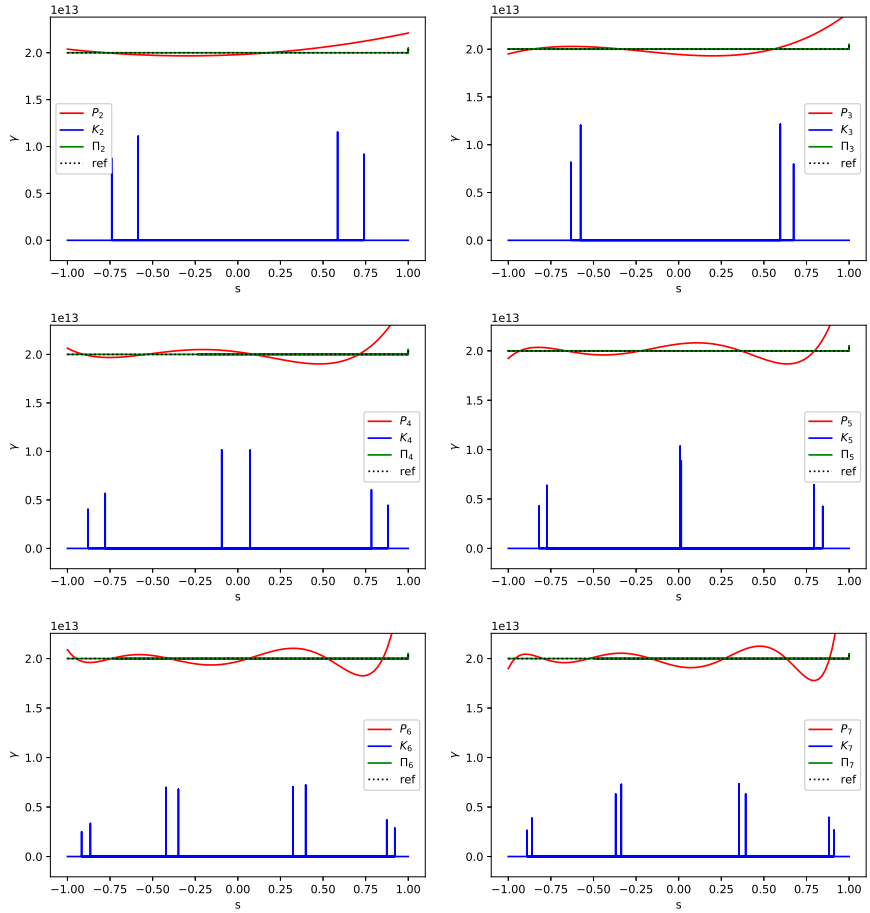

FIGURE 11. Representation of the measures representing the vector $\mathbf{f}(x = L/2)$ with P_N, K_N and Π_N models for $N = 2$ (top left), 3 (top right), 4 (middle left), 5 (middle right), 6 (bottom left), 7 (bottom right) for the Riemann problem.

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 ([43, 9, 10, 11, 60, 33, 38]) under the constraint of satisfying the moments (and potentially of being positive [27]). This reconstruction is known to be of the form $\eta^{*\prime}(\boldsymbol{\lambda}.\mathbf{b}(s))ds$ where η^* is the Legendre dual of η and $\boldsymbol{\lambda}$ are Lagrange multipliers. The resulting system is symmetric hyperbolic ([24]) and dissipates the entropy η.

The K_N closure was shown to be hyperbolic for low order N in [46, 47, 57] 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 (23).

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.**

5.3.1. *Physical equilibrium.* 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.

5.3.2. *Extension to unbounded and multi-D sets of integration.* Having the equlibrium $\mathbf{V}_{eq}$ depend on the unknown $\mathbf{V}$ could also be used to extend the present construction to unbounded sets $E = \mathbb{R}$ or $\mathbb{R}^+$ and to multi-D problems $E = S^2$. Indeed, the main difficulty in these problems relies inthe lack of representation along the boundary $\partial\mathcal{R}_{\mathbf{b}}$ in the general case or just in the construction of such representation. Such construction is straightforward along $\partial\mathcal{R}_{\mathbf{b}} \cap \mathcal{R}_{\mathbf{b}}^m$ as such vectors are uniquely represented by a discrete measure which can be fully computed (or partially to obtain only the closure) using basic linear algebra techniques. Projections of any $\mathbf{V} \in \mathcal{R}_{\mathbf{b}}$ toward a fixed $\mathbf{V}_{eq}$ can not always fall onto $\partial\mathcal{R}_{\mathbf{b}} \cap \mathcal{R}_{\mathbf{b}}^m$. However, one could find an appropriate function $\mathbf{V}_{eq}$ of $\mathbf{V}$ such that the projection to always fall onto such a desirable part of the boundary. Such objective equilibrium function should then capture the physical equilibrium and be such that the projection always point toward $\partial\mathcal{R}_{\mathbf{b}} \cap \mathcal{R}_{\mathbf{b}}^m$.

REFERENCES

[1] N. I. Akhiezer. *The classical moment problem.* Edinburgh : Oliver & Boyd, 1965.

[2] N. I. Akhiezer and M. G. Krein *Some questions in the theory of moments.* AMS Trans. Math. Monographs : Vol. 2, 1962.

[3] 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.

[4] G. W. Alldredge. *Optimization Techniques for Entropy-Based Moment Models of Linear Transport.* PhD thesis, University of Maryland, 2012.

[5] 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.

[6] C. Bayer and J. Teichmann. The proof of Tchakaloff's theorem. *arXiv:0502473*, pages 1–6, 2005.

[7] G. Birindelli. *Modèle entropique pour le calcul de dose en radiothérapie externe et curiethérapie*. PhD thesis, Université de Bordeaux, 2018.

[8] 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.

[9] J. Borwein and A. Lewis. Duality relationships for entropy-like minimization problems. *SIAM J. Control Optim.*, 29(2):325–338, 1991.

[10] J. Borwein and A. Lewis. Partially finite convex programming, part I: Quasi relative interiors and duality theory. *Math. Program.*, 57:15–48, 1992.

[11] J. Borwein and A. Lewis. Partially finite convex programming: Part II. *Math. Program.*, 57:49–83, 1992.

[12] 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.

[13] J. Caron. *Étude et validation clinique d'un modèle aux moments entropique pour le transport de particules énergétiques : application aux faisceaux d'électrons pour la radiothérapie externe*. PhD thesis, Univ. Bordeaux, 2016.

[14] 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.

[15] S. Chandrasekhar. *Radiative transfer*. Dover publications, 1960.

[16] R. Curto and L. Fialkow. Recursiveness, positivity, and truncated moment problem. *Houston j. Math.*, 17(4):603–635, 1991.

[17] R. Curto and L. Fialkow. A duality proof of Tchakaloff's theorem. *J. Math. Anal. Appl.*, 269:519–532, 2002.

[18] R. Curto and L. Fialkow. An analogue of the Riesz-Haviland theorem for the truncated moment problem. *J. Functional Analysis*, 255:2709–2731, 2008.

[19] B. Dubroca and J.-L. Feugeas. Hiérarchie des modèles aux moments pour le transfert radiatif. *C. R. Acad. Sci. Paris*, 329:915–920, 1999.

[20] L. Fialkow. The truncated K-moment problem: a survey. *Theta Ser. Adv. Math.*, 18:25–51, 2016.

[21] R. O. Fox. A quadrature-based third-order moment method for dilute gas-particle flows. *J. Comput. Phys.*, 227:63136350, 2008.

[22] M. Frank. *Partial moment models for radiative transfer*. PhD thesis, T.U. Kaiserslautern, 2005.

[23] M. Frank, B. Dubroca, and A. Klar. Partial moment entropy approximation to radiative heat transfer. *J. Comput. Phys.*, 218:1–18, 2006.

[24] K. O. Friedrichs and P. D. Lax. Systems of conservation equations with a convex extension. *Proc. Nat. Acad. Sci.*, 68(8):1686–1688, 1971.

[25] 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.

[26] H. Hamburger. Über eine Erweiterung des Stieltjesschen Momentenproblems. *Math. Ann.*, 82:120–164, 1921.

[27] C. Hauck and R. McClarren. Positive P_N closures. *SIAM J. Sci. Comput.*, 32(5):2603–2626, 2010.

[28] C. D. Hauck. High-order entropy-based closures for linear transport in slab geometry. *Commun. Math. Sci.*, 9(1):187–205, 2011.

[29] F. Hausdorff. Summationmethoden und Momentfolgen. *Math. Z.*, 9:74–109, 1921.

[30] E. K. Haviland. On the momentum problem for distributions in more than one dimension ii. *Amer. J. Math.*, 57(3):562–568, 1935.

[31] E. K. Haviland. On the momentum problem for distribution functions in more than one dimension. ii. *Amer. J. Math.*, 58(1):164–168, 1936.

[32] H. Hensel, R. Iza-Teran, and N. Siedow. Deterministic model for dose calculation in photon radiotherapy. *Phys. Med. Biol.*, 51:675–693, 2006.

[33] M. Junk. Maximum entropy for reduced moment problems. *Math. Mod. Meth. Appl. S.*, 10(1001–1028):2000, 1998.

[34] D. Kershaw. Flux limiting nature's own way. Technical report, Lawrence Livermore Laboratory, 1976.

[35] M. G. Krein and A. A. Nudel'man. *The Markov moment problem and extremals problems.* AMS Trans. Math. Monographs : Vol. 50, 1977.

[36] K. Kuepper. *Models, numerical methods, and uncertainty quantification for radiation therapy.* PhD thesis, RWTH Aachen University, 2016.

[37] J.-B. Lasserre. *Moment, positive polynomials, and their applications*, volume 1. Imperial college press, 2009.

[38] C. D. Levermore. Moment closure hierarchies for kinetic theories. *J. Stat. Phys.*, 83(5–6):1021–1065, 1996.

[39] A. S. Lewis. Consistency of moment systems. *Can. J. Math.*, 47:995–1006, 1995.

[40] R. Li and W. Li. 3D B_2 model for radiative transfer equation, Part I: Modelling. *arxiv*, 2017.

[41] 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.

[42] J. McDonald and M. Torrilhon. Affordable robust moment closures for cfd based on the maximum-entropy hierarchy. *J. Comput. Phys.*, 251:500–523, 2013.

[43] L. R. Mead and N. Papanicolaou. Maximum entropy in the problem of moments. *J. Math. Phys.*, 25(8):2404–2417, 1984.

[44] G. N. Minerbo. Maximum entropy Eddington factors. *J. Quant. Spectros. Radiat. Transfer*, 20:541–545, 1978.

[45] G. N. Minerbo. Maximum entropy reconstruction from cone-beam projection data. *Comput. Biol. Med.*, 9(1):29–37, 1979.

[46] P. Monreal. Higher order minimum entropy approximations in radiative transfer. *arXiv:0812.3063*, pages 1–18, 2008.

[47] P. Monreal. *Moment realizability and Kershaw closures in radiative transfer.* PhD thesis, RWTH Aachen University, 2012.

[48] T. E. Oliphant *Guide to NumPy* Trelgol Publishing USA, 2006.

[49] J. Page. *Développement et validation de l'application de la force de Lorentz dans le modèle aux moments entropiques M_1. Étude de l'effet du champ magnétique sur le dépôt de dose en radiothérapie externe.* PhD thesis, Univ. Bordeaux, 2018.

[50] T. Pichard. *Mathematical modelling for dose deposition in photontherapy.* PhD thesis, Université de Bordeaux & RWTH Aachen University, 2016.

[51] 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.

[52] T. Pichard, D. Aregba-Driollet, S. Brull, B. Dubroca, and M. Frank. Relaxation schemes for the M_1 model with space-dependent flux: Application to radiotherapy dose calculation. *Commun. Comput. Phys.*, 19(01):168–191, 2016.

[53] T. Pichard, S. Brull, and B. Dubroca. A numerical approach for a system of transport equations in the field of radiotherapy. *Commun. Comput. Phys.*, 25(4):1097–1126, 2019.

[54] M. Riesz. Sur le problème des moments, troisième note. *Ark. Math. Astr. Fys.*, 17(16):1–52, 1923.

[55] J. A. R. Sarr and C. P. T. Groth. A Second-Order Maximum-Entropy Inspired Interpolative Closure for Radiative Heat Transfer in Gray Participating Media. *J. Quant. Spectros. Radiat. Transfer*, accepted for publication, 2020.

[56] F. Schneider. *Moment models in radiation transport equations.* PhD thesis, T.U. Kaiserslautern, 2015.

[57] F. Schneider. Kershaw closures for linear transport equations in slab geometry i: model derivation. *J. Comput. Phys.*, 322:905–919, 2016.

[58] 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.

[59] 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.

[60] J. Schneider. Entropic approximation in kinetic theory. *ESAIM-Math. Model. Num.*, 38(3):541–561, 2004.

[61] T.-J. Stieltjes. Recherches sur les fractions continues. *Anns Fac. Sci. Toulouse : Mathématiques*, 8(4):J1–J122, 1894.

[62] V. Tchakaloff. Formules de cubature mécanique à coefficients non négatifs. *Bull. Sci. Math.*, 81:123–134, 1957.

[63] T.A. Wareing, J.M. McGhee, Y. Archambault, and S. Thompson. Acuros XB ® advanced dose calculation for the Eclipse ™ treatement planning system. *Clinical perspectives*, 2010.

[64] 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