On the Renormalization Maps for the ϕ-Divergence Moment Closures Applied in Radiative Transfer

Micheal R A Abdelmalik, Zhenning Cai, Teddy Pichard

On the renormalization maps for the φ-divergence moment closures applied in radiative transfer

M.R.A. Abdelmalik*, Z. Cai†, T. Pichard‡

Abstract

The φ-divergence-based moment method was recently introduced Abdelmalik et al. (2023) for the discretization of the radiative transfer equation. At the continuous level, this method is very close to the entropy-based M_N methods and possesses its main properties, i.e. entropy dissipation, rotational invariance and energy conservation. However, the φ-divergence based moment systems are easier to resolve numerically due to the improved conditioning of the discrete equations. Moreover, exact quadrature rules can be used to compute moments of the distribution function, which enables the preservation of energy conservation, entropy dissipation and rotational invariants, discretely. In this paper, we consider different variants of the $\varphi-$divergence closures that are based on different approximations of the exponential function and the Planck function. We compare the approximation properties of the proposed closures in the numerical benchmarks.

Keywords: Method of moments, φ-divergence, Function approximation, Radiative transfer.

1 Introduction

This paper is a follow-up to Abdelmalik et al. (2023) and is concerned about the angular approximation of the solution to the radiative transfer equation (RTE). This kinetic equation consists in a transport at a velocity Ω with a constant norm combined with a collision operator. Among the properties that this model satisfies, we want at least to preserve at the discrete level: the conservation of energy, the dissipation of a convex entropy and the rotational invariance.

The numerical methods commonly used for this model include the Monte-Carlo solvers (DSMC;Carlson (1963); Lewis and Miller (1984); Pomraning (1973); Mihalas and Mihalas (1983)) and the discrete ordinates method (DOM;Carlson (1963); Lewis and Miller (1984); Pomraning (1973); Mihalas and Mihalas (1983)), but those are computationnally very expensive, not rotationally invariant and incapable to capture the equilibria distribution the system converges to. Our work falls within the context of the method of moments which is an efficient alternative. The most popular models in this family are the polynomial P_N models (Chandrasekhar (1950); Pomraning (1973); Hesthaven et al. (2009); Canuto et al. (2006)) and the entropy-based M_N models (Minerbo (1978); Levermore (1996)) which rely on approximating the Ω-dependencies of the solution respectively by polynomials or by a distribution minimizing the entropy under moment constraints. Those are efficients, but the first fails at modelling beams, the second requires high computational costs due to need to solve large numbers of (potentially ill-conditioned) optimization problems. Those were an inspiration for many other techniques developed in this field, including the simplified models (Frank et al. (2007); McClarren (2010)), the flux-limited diffusion (Olson et al. (2000); Humbird and McClarren (2017)),

*Department of Mechanical Engineering, Eindhoven University of Technology, Groene Loper 3, 5612 AE Eindhoven, Netherlands (m.abdel.malik@tue.nl)

†Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road, Singapore 119076 (matcz@nus.edu.sg)

‡CMAP, CNRS, École polytechnique, Institut polytechnique de Paris, 91120, Palaiseau, France (teddy.pichard@polytechnique.edu)

the interpolative methods (Pichard et al. (2017); Li and Li (2020); Sarr and Groth (2020)) and others (see e.g. Schneider (2016); Pichard (2020)).

Recently, the entropy-based method of moment was simplified by exploiting φ-divergence techniques (Csiszár (1972)) while preserving its main properties, originally in the context of rarefied gases (Abdelmalik (2017); Abdelmalik and van Brummelen (2016)) and for the RTE in Abdelmalik et al. (2023). This method consists in a Galerkin approximation where the test function space chosen to be polynomials set and the approximation set is chosen to be a non-linear renormalization applied to the same polynomial set. This renormalization is chosen to be a high degree polynomial approximation of the exponential for the final model to dissipate an approximation of Boltzmann entropy. This model was shown to preserve the three aforementionned properties of conservation of energy, rotational invariance and entropy dissipation. Therefore, it possesses the same properties as M_N models but the optimization problem to solve requires a lower cost and they rely on exact quadrature rules and therefore preserves further in the construction of the discretization the rotational invariance.

The present work exploits the versatility of the method in order to address two issues regarding the angular approximation of a kinetic solution: First, when the degree of the renormalization mapping tends to infinity, our method falls back onto the M_N method based on the Boltzmann entropy, it captures therefore exactly the beam distributions but requires to solve worse-conditioned optimization problems. We would like to adapt our method in order to choose the compromise between the acccuracy in the beam regime and the condition of the optimization problems to solve. Second, Boltzmann entropy is often considered in radiative transfer by analogy with the rarefied gases models, but other entropies are more physically relevant in this context. We would like to adapt our method such that it converges toward other entropy-based models. These two problems are tackled by constructing other polynomial renormalization mappings, and they eventually preserve the aforementionned properties.

The paper is organized as follows. The next section recalls the radiative transfer equation and its properties. The next recalls the construction of our φ-divergence moment closure. The novel construction and study of polynomial renormalization mappings arises in Section 4. Section 5 is devoted to numerical experiments with the present method to approximate functions corresponding to different physical regimes and Section 6 to conclusion.

2 Radiative transfer equation

Consider the radiative transfer equation (RTE)

$$\partial_t I + \Omega \cdot \nabla_x I = LI := \sigma \left(\frac{1}{4\pi} \int_{\mathbb{S}^2} I \, \mathrm{d}\Omega - I \right), \tag{1a}$$

where the unknown $I(t, x, \Omega)$ is the radiative intensity depending on $\Omega \in \mathbb{S}^2$, $x \in \mathbb{R}^3$ and $t \in]0, T[$. This equation is supplemented with initial condition

$$I(t = 0, x, \Omega) = I^0(x, \Omega). \tag{1b}$$

Remark that considering unbounded spatial domain avoids the difficulties emerging with boundary conditions.

This problem is well-posed and preserves the integrability of the initial condition (Dautray and Lions (2000)) in the sense: if $I^0 \in \mathcal{L}^p(\mathbb{R}^d \times \mathbb{S}^2)$, then there exists a unique function $I \in \mathcal{C}\left((0, T); \mathcal{L}^p(\mathbb{R}^d \times \mathbb{S}^2)\right)$ satisfying (1). Furthermore, if $I^0 \geq 0$, then $I \geq 0$. Following this result, the focus is made on $\mathcal{L}^1$ functions in the rest of the paper.

Using the chain rule, this equation is shown to dissipate any entropy, i.e. for all convex scalar functions

η, then

$$\partial_t H(I) + \nabla_x \cdot G(I) = S(I) \leq 0, \tag{2a}$$

$$H(I) = \int_{\mathbb{S}^2} \eta(I(\Omega)) d\Omega, \tag{2b}$$

$$G(I) = \int_{\mathbb{S}^2} \Omega \eta(I(\Omega)) d\Omega, \tag{2c}$$

$$S(I) = \int_{\mathbb{S}^2} \eta'(I(\Omega)) LI(\Omega) d\Omega. \tag{2d}$$

In the present case, the space of collisional invariants

$$\mathtt{C} = \left\{ f \text{ s.t. } \int_{\mathbb{S}^2} f(\Omega) LI(\Omega) d\Omega = 0 \right\}$$

is one-dimensional and composed only of the isotropic functions

$$f \in \mathtt{C} \quad \Leftrightarrow \quad f(\Omega) = \frac{1}{4\pi} \int_{\mathbb{S}^2} f(\Omega') d\Omega'.$$

Therefore, for all convex function η, H is minimum when I is isotropic and the system is attracted toward such equilibria.

For the construction of entropy-based moment closures as below, the Boltzmann-Shannon (afterward denoted with the subscript BS) entropy

$$\eta_{BS}(I) = I \log I$$

is often used as a comparison of the kinetic model (1a) with rarefied gas models. But the Bose-Einstein (afterward denoted with the subscript BE) entropy

$$\eta_{BE}(I) = I \log I - (I+1) \log(I+1)$$

is more meaningful when considering more physically realistic collision models in (1): Interaction of the radiations with matter is often modeled by a coupled system of the form (see e.g. Lowrie et al. (1999); Pomraning (1973); Mihalas and Mihalas (1983) ; some of the physical constants are set to one for simplicity)

$$\partial_t I + \Omega \cdot \nabla_x I = \sigma \left(B_{\nu,T} - I \right), \tag{3a}$$

$$\partial_t T = - \int_0^\infty \int_{\mathbb{S}^2} \sigma (B_{\nu,T} - I) d\nu d\Omega, \tag{3b}$$

where I depends additionally on a frequency variable $\nu \in \mathbb{R}^+$, the matter temperature T depends only on (t,x) and the Planck distribution takes the form $B_{\nu,T} = \left(\exp(\frac{h\nu}{T}) - 1 \right)^{-1}$. This coupled system satisfies the additional inequality (2) where the entropy $H(I,T) = \int_{\mathbb{S}^2} \int_0^\infty \nu^2 \eta_{BE}(I) d\nu d\Omega - \log T$ and its flux $G(I,T) = \int_{\mathbb{S}^2} \int_0^\infty \Omega \nu^2 \eta_{BE}(I) d\nu d\Omega$ yield a negative entropy source $S(I,T) \leq 0$. We refer to Dubroca and Feugeas (1999) for a study of the radiative entropy dynamics.

Finally, Equation (1) was shown to be rotationally invariant, meaning that its solution I satisfies

$$(\partial_t I)(\mathcal{O}\Omega) = \partial_t (I(\mathcal{O}\Omega)), \quad (\Omega \cdot \nabla_x I)(\mathcal{O}\Omega) = (\mathcal{O}\Omega) \cdot \nabla_x (I(\mathcal{O}\Omega)), \quad (LI)(\mathcal{O}\Omega) = L(I(\mathcal{O}\Omega)),$$

for all rotation matrices $\mathcal{O} \in SO(3)$.

Our aim in this paper is to provide and study approximations of a function I with respect to the Ω-variable that preserve the properties presented in this section at the moment level and that are adapted to a selected convex entropy. Dynamical simulations for (1a) with a first version of the present approximation were performed in Abdelmalik et al. (2023). Extensions to (3) when considering the frequency dependency would require further techniques, such as multi-group approximations, and are left as future work.

3 φ-divergence-based moment equations

We recall here the construction of the moment closure from Abdelmalik et al. (2023) and the problems tackled in the present work.

3.1 Construction of the Galerkin framework

The moment system is obtained as a Galerkin approximation of (1a) in the Ω variable. This formulation requires three elements, the choices and properties are recalled here:

- A *finite dimensional* test functions space M, which must be a *subset of the solution set dual* $\mathcal{L}^\infty(\mathbb{S}^2)$. In order to preserve those properties at the numerical level, M must *contain the collision invariants* $1 \in M$; and M must be *rotational invariant*. The natural choice to satisfy both properties is the set of polynomials up to a certain degree N
$$M := \mathbb{P}_N(\mathbb{S}^2).$$

- A renormalization map β to account for non-linearity in the approximation. For a convex function η, choosing $\beta = (\eta')^{-1}$ corresponds to dissipating η at the underlying kinetic level (see e.g. Levermore (1996)). Especially, β must be *monotonically increasing* to match such an entropy. Natural choices include
$$\beta_{BS}(g) = \exp(g) = (\eta'_{BS})^{-1}(g), \qquad \beta_{BE}(g) = \frac{1}{\exp(g)-1} = (\eta'_{BE})^{-1}(g). \tag{4}$$
In our work, we aim at imposing $\beta \in \mathbb{P}_K(\mathbb{R})$ with $K \geq 1$ for numerical quadrature (up to a sufficient order) to be exact, this provided a rotationally invariant algorithm in Abdelmalik et al. (2023). We chose renormalizations of the form
$$\beta_K(g) = \left(1 + \frac{g}{K}\right)^K = (\eta'_K)^{-1}(g) \quad \text{with} \quad \eta_K(I) = KI\left(\frac{K}{K+1} I^{1/K} - 1\right), \tag{5}$$
which are polynomial approximations of exp, and monotonically increasing for odd $K \geq 1$. We exhibit in the next section other choices.

- A finite dimensional trial functions space V such that $\beta(V) \subset \mathcal{L}^1$ is a subset of the solution set. Again, $\beta(V)$ must *contain the equilibrium distributions* $\mathtt{C} \subset \beta(V)$; and V must be *rotational invariant*, and we choose again
$$V := \mathbb{P}_N(\mathbb{S}^2).$$

Eventually, this yields seeking $g \in \mathbb{P}_N(\mathbb{S}^2)$ such that for all $m \in \mathbb{P}_N(\mathbb{S}^2)$

$$\forall (t,x) \in (0,T) \times \mathbb{R}^d, \quad \int_{\mathbb{S}^2} m(\Omega)\left[(\partial_t + \Omega \cdot \nabla_x)\beta(g) - L\beta(g(t,x,\cdot))(\Omega)\right] d\Omega = 0, \tag{6a}$$

$$\forall x \in \mathbb{R}^d, \quad \int_{\mathbb{S}^2} m(\Omega)\left[\beta(g(t=0,x,\Omega)) - I^0(x,\Omega)\right] d\Omega = 0. \tag{6b}$$

Let $\boldsymbol{m}$ denote a vector of all of the basis functions $m(\Omega) \in \mathbb{P}_N(\mathbb{S}^2)$, under the constraints mentioned in the last paragraph. Then (6) can be expressed as a system of moment equations

$$\partial_t \mathbf{U} + \mathrm{div}_x(\mathbf{F}(\mathbf{U})) = \mathbf{L}\mathbf{U}, \tag{7a}$$

$$(\mathbf{U}, \mathbf{F}(\mathbf{U}), \mathbf{L}\mathbf{U}) = \int_{\mathbb{S}^2} \mathbf{m}(\Omega)\left(\beta(\boldsymbol{\lambda}^T\mathbf{m}(\Omega)),\ \Omega^T\beta(\boldsymbol{\lambda}^T\mathbf{m}(\Omega),\ L\beta(\boldsymbol{\lambda}^T\mathbf{m}(\Omega)\right) \mathrm{d}\Omega, \tag{7b}$$

which possesses a symmetrizer constructed from η (the anti-derivative of β^{-1}) and $\boldsymbol{\lambda}$ are the so-called entropic variables (Godlewski and Raviart (1996)) in which (7a) can be written in the so-called symmetric hyperbolic form. Therefore, (7a) possesses a convenient structure for the study of its well-posedness (Kawashima and Yong (2004)).

Remark 1. The entropic variables $\boldsymbol{\lambda}$ introduced in (7b) can also be conceived of as Lagrange multipliers that enforce the moment constraints in the so-called entropy minimization problem (Levermore (1996))

$$\text{Find argmin}\left\{H(h) : \int_{\mathbb{S}^2} \boldsymbol{m} I d\Omega = \int_{\mathbb{S}^2} \boldsymbol{m} h d\Omega\right\}. \tag{8}$$

3.2 Position of the problem

In this paper, we investigate several modifications of the renormalization map (5) and its impact on the closure Abdelmalik et al. (2023). The objective is to alter the convergence of the φ-divergence solution when $K \to +\infty$ either to accelerate the convergence or to modify the limit value:

First, when considering the moments of a Dirac distribution, e.g. $\mathbf{U} = \mathbf{m}(e_1) = \int_{\mathbb{S}^2} \mathbf{m}\delta_{e_1}$, numerical experiments in Abdelmalik et al. (2023) showed that the reconstruction β_K does converge in H^{-2} toward the distribution δ_{e_1} when $K \to +\infty$. However, the rate of convergence is slow. This slow rate of convergence is also illustrated on Fig. 1 where the function $\beta_K(x)$ is plotted for various K odd together with its limits $\lim_{K\to\infty} \beta_K = \exp$.

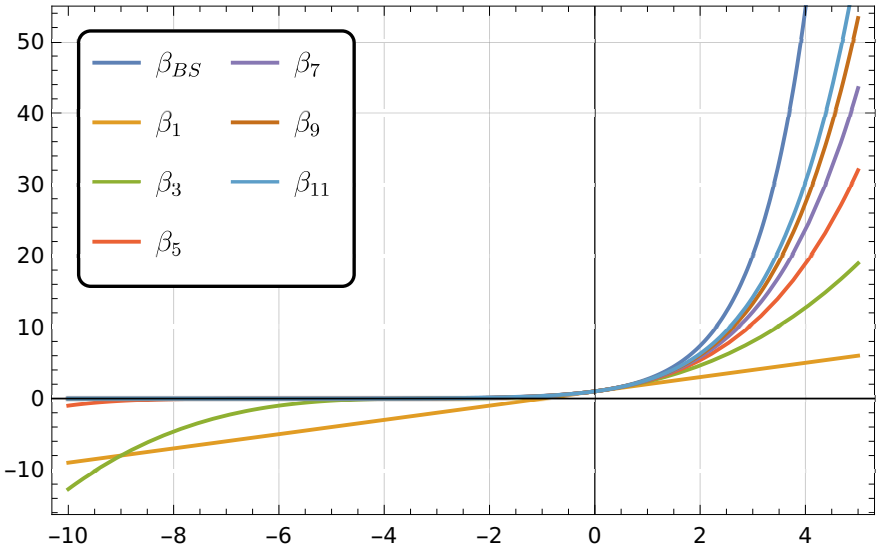


Figure 1: Renormalization mappings β_K for odd K and exponential function.

Formally, the Dirac distribution corresponds to the values $x \to -\infty$ and $x \to \infty$. Indeed, considering a Gaussian mollifier $\frac{1}{\sqrt{\pi\sigma}}\exp(-\frac{y^2}{\sigma})$, then $x = \frac{-y^2}{\sigma}$ takes for values $\pm\infty$ in the limit.

One observes that the sequence $(\beta_K)_{K\in\mathbb{N}}$ indeed converges pointwisely for bounded x toward the exponential function, but this convergence is very slow. Furthermore, the rate of the exponential in the limit $x \to +\infty$ is not accurately captured and one needs high order K, and therefore higher complexity, to capture such large values. Similarly, the zero limit when $x \to -\infty$ can not be reached by any β_K function with finite K as they are polynomials and cannot have bounded value in $-\infty$. Eventually, such Dirac distributions can only be approximated and we only aim at improving the range of accuracy of such approximations.

Second, when $K \to +\infty$, the sequence $(\beta_K)_{K\in\mathbb{N}}$ of approximations may only converge toward the exponential. As mentioned in the previous section, other types of equilibrium can be expected from the solution of the RTE, typically the Planck distribution β_{BE} defined in (4). The β_K approximation does not possess the flexibility to converge toward other $\beta = (\eta')^{-1}$ functions.

Therefore, the objective in the next section is to provide another type of approximations which is flexible enough to control the convergence when $K \to \infty$, i.e. both the convergence rate and the limit function. Applications of the method to the RTE were shown in Abdelmalik et al. (2023) and we expect the present

modifications to have little impact on such academic applications, but could be required for further more realistic ones.

4 Other monotonically increasing polynomial approximations

The solutions considered in this paper consist in *polynomial* approximations. The integral of such polynomial functions can be computed exactly using an appropriate quadrature rule. Therefore, the coefficients $\boldsymbol{\lambda}$ of the approximation in (7) can be obtained using Newton iterations (Abdelmalik et al. (2023)) that can be computed exactly. Remark that rotation invariance is lost when constructing entropy-based closures (see e.g. Hauck (2011)) due to non-exact integration while the present choice of polynomial approximation allows to compute the integrals exactly.

For the model to possess a convex entropy dissipated, we still need this approximation to be *monotonically increasing*. Therefore, we require

$$\beta' > 0.$$

4.1 Taylor expansion

A first idea originated in the observation that the Taylor expansion of the exponential around zero converges faster (empirically) than the sequence $(\beta_K)_{K\in\mathbb{N}}$. This consists in writing in the shifted monomial basis $\mathbf{b} = \mathbf{b}^K(x) := \left(1, \dots, (x-x_0)^K\right)^T$

$$T_K(x) = \sum_{k=0}^{K} \alpha_k (x-x_0)^k = \boldsymbol{\alpha}^T \mathbf{b} \tag{9}$$

where $\alpha_k = \frac{\beta^{(k)}(x_0)}{k!}$ for a generic function β. We provide a simple characterization of monotonically increasing polynomials of this form.

Proposition 4.1. *Suppose that $\beta \in \mathcal{C}^{2K+2}$ is such that $\beta^{(i)} \geq 0$ for all $1 \leq i \leq 2K+2$. Then the polynomial T_{2K+1} is monotonically increasing.*

Proof. This simply follows from the Taylor formula of β' with remainder: there exists $\xi \in [x, x_0]$ such that

$$\beta'(x) - T'_{2K+1}(x) = \frac{\beta^{(2K+2)}(\xi)}{(2K+2)!}(x-x_0)^{2K+1},$$

which is negative for $x < x_0$. Therefore, for $x \leq x_0$

$$T'_{2K+1}(x) \geq \beta'(x) \geq 0.$$

The derivative $T'_{2K+1}(x)$ is also positive for $x > x_0$, then it is monotonically increasing. □

All the derivatives of the exponential $\beta_{BS} = \exp$ are positive, then it satisfies this property and the polynomial

$$T_{2K+1}(x) = \sum_{k=0}^{2K+1} \frac{e^{x_0}}{k!}(x-x_0)^k$$

is monotonically increasing for odd degree $2K+1$. The convergence radius of this sequence is infinite. Therefore, we have convergence of the approximation toward the desired results for all $x \in \mathbb{R}$

$$T_{2K+1}(x) \underset{K\longrightarrow\infty}{\longrightarrow} \exp(x).$$

Concerning the Planck function β_{BE} minimizing the Bose-Einstein entropy, it reads for all strictly negative reals $x \in \mathbb{R}^{*,-}$

$$\beta_{BE}(x) = \frac{1}{e^{-x} - 1} > 0,$$

and one verifies that its derivative satisfies $\beta'_{BE} = (1 + \beta_{BE})\beta_{BE}$ such that all the successive derivatives of β_{BE} are polynomials in β_{BE} with positive coefficients. Especially, those derivatives are all strictly positive for all $x \in \mathbb{R}^{*,-}$ since $\beta_{BE} > 0$. Therefore, the polynomial T_{2K+1} of odd degree with the Planck function are also monotonically increasing. The convergence radius δ of this sequence remains bounded and it depends on the chosen point of expansion x_0. This radius $\delta < |x_0|$ since the function β_{BE} is singular in zero (and undefined after). Therefore, we only have convergence of the approximation toward the desired results for all $x \in (x_0 - \delta, x_0 + \delta) \subsetneq \mathbb{R}^{*,-}$

$$T_{2K+1}(x) \underset{K\to\infty}{\to} \beta_{BE}(x).$$

Especially, we do not have convergence $T_{2K+1}(x) \underset{K\to\infty}{\not\to} \beta_{BE}(x)$ for all points $x < 2x_0$. These are illustrated numerically in Section 4.3 below, which even exhibit divergence for such values of x.

4.2 Optimized parameters

Another idea is to minimize the $\mathcal{L}^2$ difference between the function β to approximate (β_{BS} or β_{BE}) and the polynomials of a given degree, with the constraint that the polynomials must be monotonically increasing. This can be mathematically represented as

$$\begin{aligned} O_{2K+1} = \underset{p\in\mathbb{P}_{2K+1}}{\operatorname{argmin}} \frac{1}{2}\int_c^d |p(x) - \beta(x)|^2 \,\mathrm{d}x, \\ \text{subject to } p'(x) \geq 0, \quad \forall x \in \mathbb{R}. \end{aligned} \tag{10}$$

4.2.1 Reformulation of the approximation

In order to enforce the constraint, we use the fact that all non-negative one-variable polynomials can be represented as the sum of two squares, one of which has a lower degree than the other (Lasserre (2009); Schmuedgen (2017); Szegö (1939)). Thus, the derivative of the polynomial $O_{2K+1}(x)$ has the form:

$$O'_{2K+1}(x) = (a_0 + a_1 x + \cdots + a_K x^K)^2 + (b_0 + b_1 x + \cdots + b_{K-1} x^{K-1})^2.$$

Integrating this equation gives

$$\begin{aligned} O_{2K+1}(x) &= C + \sum_{i=0}^{K}\sum_{j=0}^{K}(a_i a_j + b_i b_j)\frac{x^{i+j+1}}{i+j+1} \\ &= C + \sum_{n=1}^{K+1}\sum_{i=0}^{n-1}(a_i a_{n-1-i} + b_i b_{n-1-i})\frac{x^n}{n} \\ &\quad + \sum_{n=K+2}^{2K+1}\sum_{i=n-1-K}^{K}(a_i a_{n-1-i} + b_i b_{n-1-i})\frac{x^n}{n}, \end{aligned} \tag{11}$$

where we assumed $b_K = 0$. With this form of O_{2K+1}, we can turn the optimization problem into an unconstrained optimization problem. For simplicity, we define intermediate parameters

$$\alpha_0 = C, \qquad \alpha_n = \frac{1}{n}\sum_{i=\max(0,n-1-k)}^{\min(k,n-1)}(a_i a_{n-1-i} + b_i b_{n-1-i}), \quad n = 1, \cdots, 2K+1,$$

so that α_n is the n-th coefficient of the polynomial O_{2K+1}. Let

$$\boldsymbol{w} = (C, a_0, \cdots, a_K, b_0, \cdots, b_{K-1})^T, \qquad \boldsymbol{\alpha}(\boldsymbol{w}) = (\alpha_0, \alpha_1, \cdots, \alpha_{2K+1})^T.$$

Then the Jacobian matrix $J = \partial\boldsymbol{\alpha}/\partial\boldsymbol{w}$ has the following form:

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2A & 2B \end{pmatrix},$$

where

$$A = \begin{pmatrix} a_0 & 0 & \dots & & \dots & 0 \\ \frac{a_1}{2} & \frac{a_0}{2} & 0 & & & \vdots \\ \frac{a_2}{3} & \frac{a_1}{3} & \frac{a_0}{3} & \ddots & & \\ \vdots & \frac{a_2}{4} & \frac{a_1}{4} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \frac{a_2}{5} & \ddots & \ddots & 0 \\ \frac{a_K}{K+1} & \vdots & \vdots & \ddots & \ddots & \frac{a_0}{K+1} \\ 0 & \frac{a_K}{K+2} & \vdots & & \ddots & \frac{a_1}{K+2} \\ \vdots & \ddots & \frac{a_K}{K+3} & & & \frac{a_2}{K+3} \\ & & \ddots & \ddots & & \vdots \\ & & & & \ddots & \vdots \\ & & & & \ddots & \frac{a_K}{2K+1} \\ & & & & & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots & & & \dots & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} b_0 & 0 & \dots & & \dots & 0 \\ \frac{b_1}{2} & \frac{b_0}{2} & 0 & & & \vdots \\ \frac{b_2}{3} & \frac{b_1}{3} & \frac{b_0}{3} & \ddots & & \\ \vdots & \frac{b_2}{4} & \frac{b_1}{4} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \frac{b_2}{5} & \ddots & \ddots & 0 \\ \frac{b_{K-1}}{K} & \vdots & \vdots & \ddots & \ddots & \frac{b_0}{K} \\ 0 & \frac{b_{K-1}}{K+1} & \vdots & & \ddots & \frac{b_1}{K+1} \\ \vdots & \ddots & \frac{b_{K-1}}{K+2} & & & \frac{b_2}{K+2} \\ & & \ddots & \ddots & & \vdots \\ & & & & \ddots & \vdots \\ & & & & \ddots & \frac{b_{K-1}}{2K-1} \\ & & & & & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots & & & \dots & 0 \end{pmatrix}.$$

To solve the optimization problem, we first reformulate the objective function $\boldsymbol{w} \mapsto f(\boldsymbol{\alpha}(\boldsymbol{w}))$ as a function of the intermediate parameters $\boldsymbol{\alpha}$:

$$\begin{aligned} f(\boldsymbol{\alpha}) &= \frac{1}{2}\int_c^d \left| \sum_{n=0}^{2K+1} \alpha_n x^n - \beta(x) \right|^2 \mathrm{d}x \\ &= \frac{1}{2}\boldsymbol{\alpha}^T M \boldsymbol{\alpha} - \boldsymbol{\beta}^T \boldsymbol{\alpha} + \frac{1}{2}\int_c^d \beta(x)^2 \,\mathrm{d}x. \end{aligned}$$

Note that the value of the last integral does not matter in our optimization problem, and the matrix M and the vector $\boldsymbol{\beta}$ are given by

$$M_{i,j} = \frac{d^{i+j+1} - c^{i+j+1}}{i+j+1}, \qquad \beta_j = \int_c^d x^j \beta(x) dx.$$

In the case of the exponential $\beta = \beta_{BS}$, the second coefficient rewrites

$$\beta_j = (-1)^j \left[\Gamma(j+1, -d) - \Gamma(j+1, -c)\right],$$

where the incomplete Γ function is well-implemented in standard numerical libraries. In the case $\beta = \beta_{BE}$, the integral can be represented using the polylogarithmic function $\mathrm{Li}_s(z)$:

$$\beta_j = \sum_{k=0}^{j} \frac{(-1)^{j+k} j!}{k!} \left[d^k \,\mathrm{Li}_{j+1-k}(\mathrm{e}^d) - c^k \,\mathrm{Li}_{j+1-k}(\mathrm{e}^c)\right].$$

4.2.2 Details on the numerical computations

The minimum of $f(\boldsymbol{\alpha}(\boldsymbol{w}))$ is attained where the gradient annihilates. Then we need to solve the nonlinear system $\nabla_{\boldsymbol{w}} f(\boldsymbol{\alpha}(\boldsymbol{w})) = 0$, or

$$[J(\boldsymbol{w})]^T M \boldsymbol{\alpha}(\boldsymbol{w}) = 0. \tag{12}$$

Since $J(\boldsymbol{w})$ is linear and $\boldsymbol{\alpha}(\boldsymbol{w})$ is quadratic, these are actually cubic equations in $\boldsymbol{w}$.

In the following, we compute numerically some parameters $\boldsymbol{w}$ and $\boldsymbol{\alpha}(\boldsymbol{w})$ by using Newton's method for (12). Note that the function $\boldsymbol{w} \mapsto f(\boldsymbol{\alpha}(\boldsymbol{w}))$ is quartic and it is infinity in infinity. However, it might not be convex at all $\boldsymbol{w} \in \mathbb{R}^{2K+2}$. Especially, it may possess several local extrema. Therefore, we cannot guarantee the convergence of Newton's method toward a global minimum. In practice, we choose 500 initial values of $\boldsymbol{w}$ randomly and pick the final solution with the minimum value of the objective function.

Using the density of the (positive) polynomials in $\mathcal{L}^2$, we obtain that this approximation converges in $\mathcal{L}^2(c,d)$ toward the desired β function. However, the convergence is again restricted to the chosen interval (c,d) and we can not guarantee convergence out of it. These are illustrated numerically in the next paragraph.

4.3 Comparing the monotonic polynomial approximations

4.3.1 Approximation for the Boltzmann-Shannon entropy

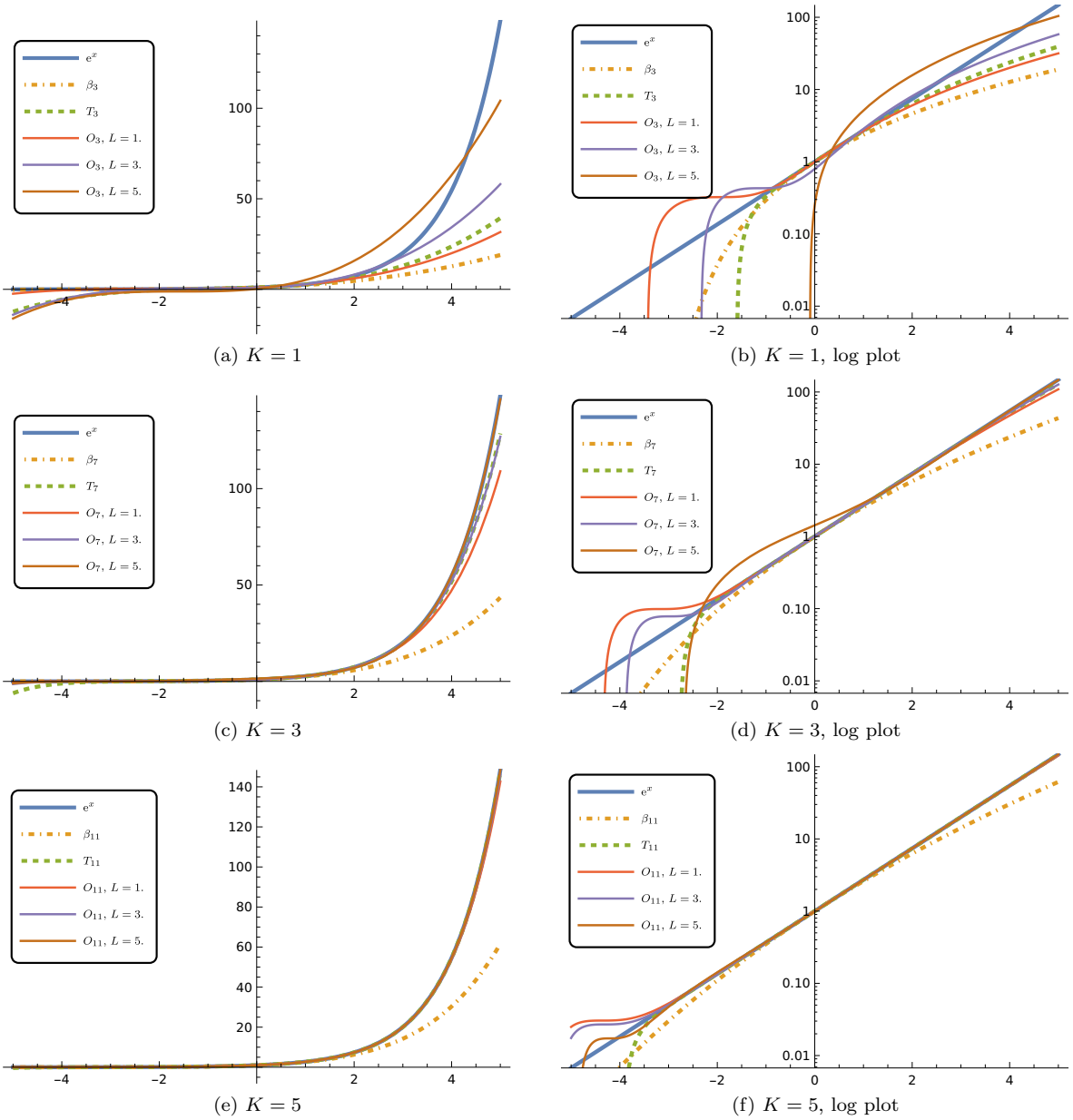
The approximations of the exponential β_{BS} are compared in Fig. 2 for different degrees $2K+1$. The point of expansion is chosen to be $x_0 = 0$ for the Taylor approximation and the range for the optimized approximation is $[-L, L]$ with $L = 1, 3, 5$. In general, by comparing the β_{2K+1} model and the T_{2K+1} model, the β_{2K+1} model gives a better approximation for $x \in \mathbb{R}^{*,-}$, while the T_{2K+1} model does better for $x \in \mathbb{R}^{*,+}$. For the optimized approximation, the function depends on the choice of the range $[-L, L]$. As expected, when K is fixed, the function is approximated within the range $[-L, L]$ for smaller values of L. It balances the quality of the approximation for both the negative and positive parts, and generally looks better than both β_{2K+1} and T_{2K+1} in the plots with linear scales (left panel of Fig. 2). Similar phenomena can be observed in Fig. 3, where all the results for the optimized approximation are given for $L = 5$. The log plots show a clear difference between the O_{2K+1} model and the β_{2K+1} and T_{2K+1} models: for the latter, the entire approximation is below the exact exponential function, whereas the O_{2K+1} approximation oscillates around the exponential, which is a typical behavior in spectral approximations.

Another interesting phenomenon that can be observed from Fig. 2 is that the O_{2K+1} approximation seems to have a good capability in extrapolations. For instance, the red curve in Fig. 2f is computed by minimizing the $\mathcal{L}^2$ distance (10) only on the interval $[-1, 1]$, but this approximation is also quite accurate for $x \in [-3, 5]$. Since the choice of the approximation range usually depends on some preliminary estimation of the problem, this property allows us to use the O_{2K+1} approximations without too much worries about the function values out of the chosen range in actual simulations.

Fig. 4 plots the $\mathcal{L}^2(-L, L)$ difference between the exponential function β_{BS} and the O_{2K+1} approximations. Unsurprisingly, for a fixed polynomial degree K, this error increases as the size of the interval $[-L, L]$ increases, i.e. as L increases. And for a fixed interval $[-L, L]$, it decreases as the approximation polynomial improves, i.e. as K increases. For a fix L, the gaps between lines are nearly the same, showing the spectral convergence rate with respect to K. When L gets larger, the gap becomes narrower, indicating slower convergence. For a fixed K, the figure implies that the error increases in the form of L^α for a certain value of α depending on K.

4.3.2 Approximation for the Bose-Einstein entropy

Similar experiments are done for the Planckian β_{BE}. The results are plotted in Fig. 5 and 6. Since the function is defined only for negative values, we choose the range of approximation to be $[-L, -1/L]$ with $L = 2, 6, 10$. In Fig. 5, three different choices of x_0 are considered for the Taylor expansion, and the choices are made with $x_0 = -(L + 1/L)/2$, which is the center of the interval $[-L, -1/L]$ used in the O_{2K+1} approximation. The general behavior is similar to the case of Boltzmann-Shannon entropy: the optimized



(a) $K = 1$ (b) $K = 1$, log plot

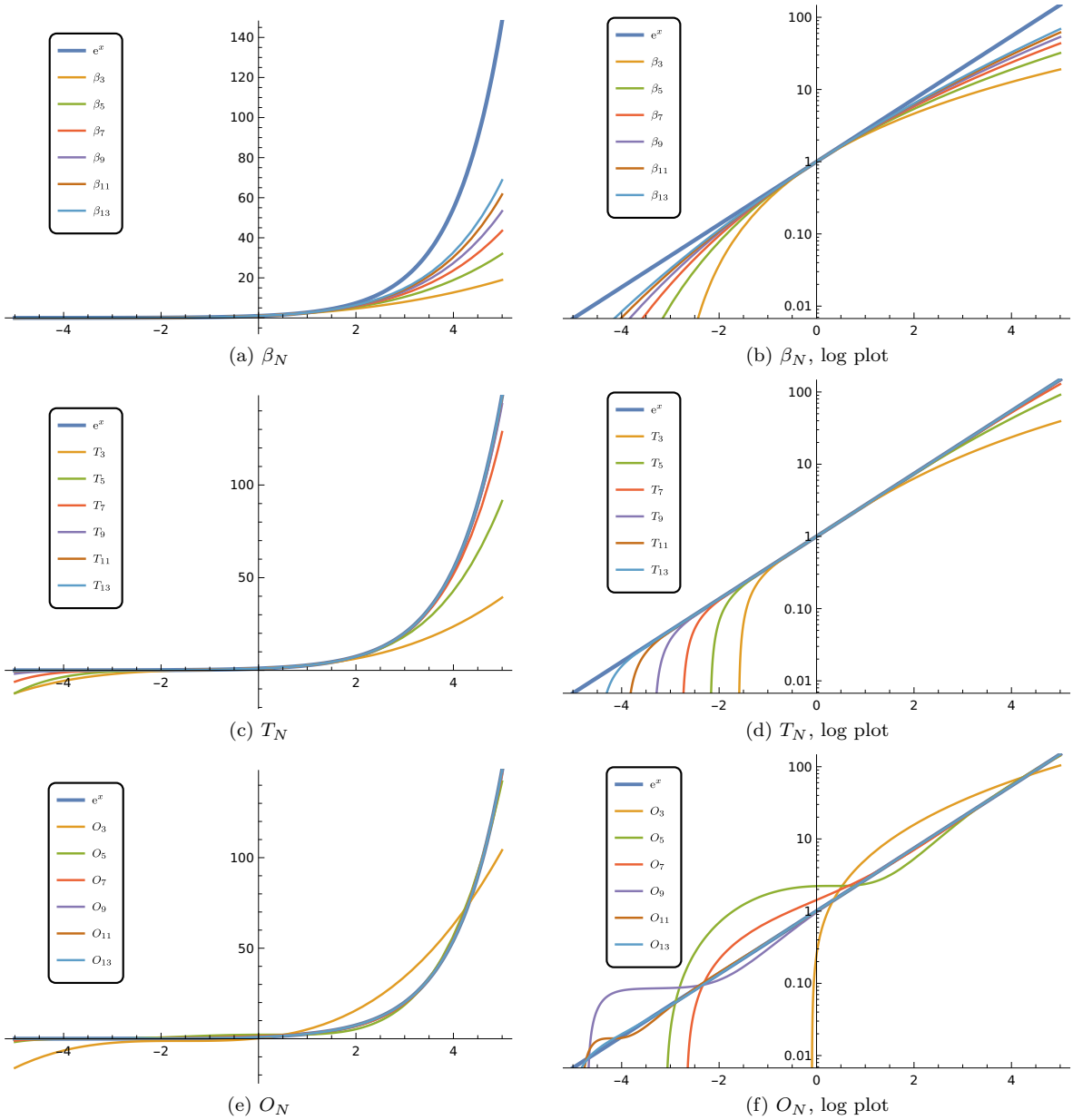
(c) $K = 3$ (d) $K = 3$, log plot

(e) $K = 5$ (f) $K = 5$, log plot

Figure 2: β_{2K+1}, Taylor T_{2K+1} and optimized O_{2K+1} approximations of the exponential function β_{BS}.

approximation fits the Planckian better within the range $[-L, -1/L]$, but it may perform worse than the Taylor approximation out of this interval. The convergence with respect to K can be better observed in Fig. 6, where all Taylor series are expanded about the same point $x_0 = -3.08333$, and the value of L is fixed to be 6 for all optimized approximations. Compared with Taylor approximations, the optimized approach better approximates the part with larger function values, which suppresses the $\mathcal{L}^2$ error more efficiently. Note that both the T_{2K+1} and O_{2K+1} approximations are increasing functions across the entire real axis $\mathbb{R}$, despite the seemingly oscillatory behavior of O_{2K+1} functions.

The $\mathcal{L}^2(-L, -1/L)$ error of the O_{2K+1} approximation for different K and L is given in Fig. 7, where we can again observe the spectral convergence with respect to K for fixed L, and the convergence rates are



(a) β_N (b) β_N, log plot

(c) T_N (d) T_N, log plot

(e) O_N (f) O_N, log plot

Figure 3: β_{2K+1}, Taylor T_{2K+1} and optimized O_{2K+1} approximations of the exponential function β_{BS}.

lower for larger intervals. Comparing Fig. 7 with Fig. 4, we can find that the $\mathcal{L}^2$ error for the Bose-Einstein entropy is significantly larger. This is likely due to the singularity of the Planckian at zero. No polynomial possesses the same property, making the function more difficult to approximate using polynomials. The exponential function, however, tends to infinity only when x tends to infinity, which all polynomials with positive leading coefficients also satisfy.

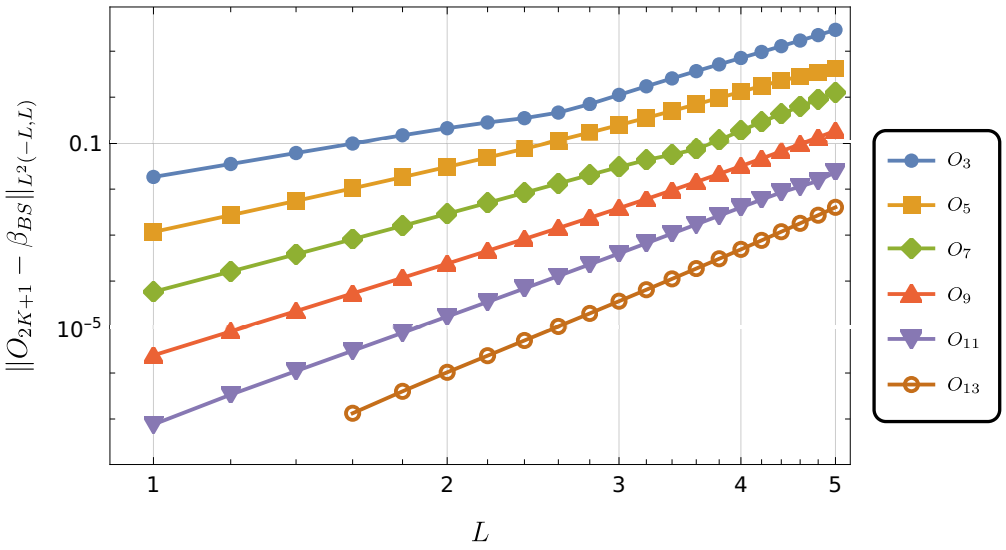


Figure 4: Error plot for the O_{2K+1} approximations of the exponential function β_{BS}

5 Numerical approximation of specific distributions

To complete this study, we reproduce simulations from Abdelmalik et al. (2023) with the different approximations and compare their results. These consist in approximating distributions that correspond to physical regime, namely a near-beam distribution, a distribution corresponding to two beams crossing each others, and smooth distributions. Some applications of the present method to (1a) with the renormalization map β_K shown in (5) were performed in Abdelmalik et al. (2023). The present modifications are expected to have little impact on such academic test cases, but would be valuable when applied to (3) or more realistic radiation models.

Finding the approximation requires solving the moment inversion problem:

$$\int_{\mathbb{S}^2} \mathbf{m}(\Omega)\beta(\boldsymbol{\lambda}^T\mathbf{m}(\Omega))\,\mathrm{d}\Omega = \mathbf{U}, \tag{13}$$

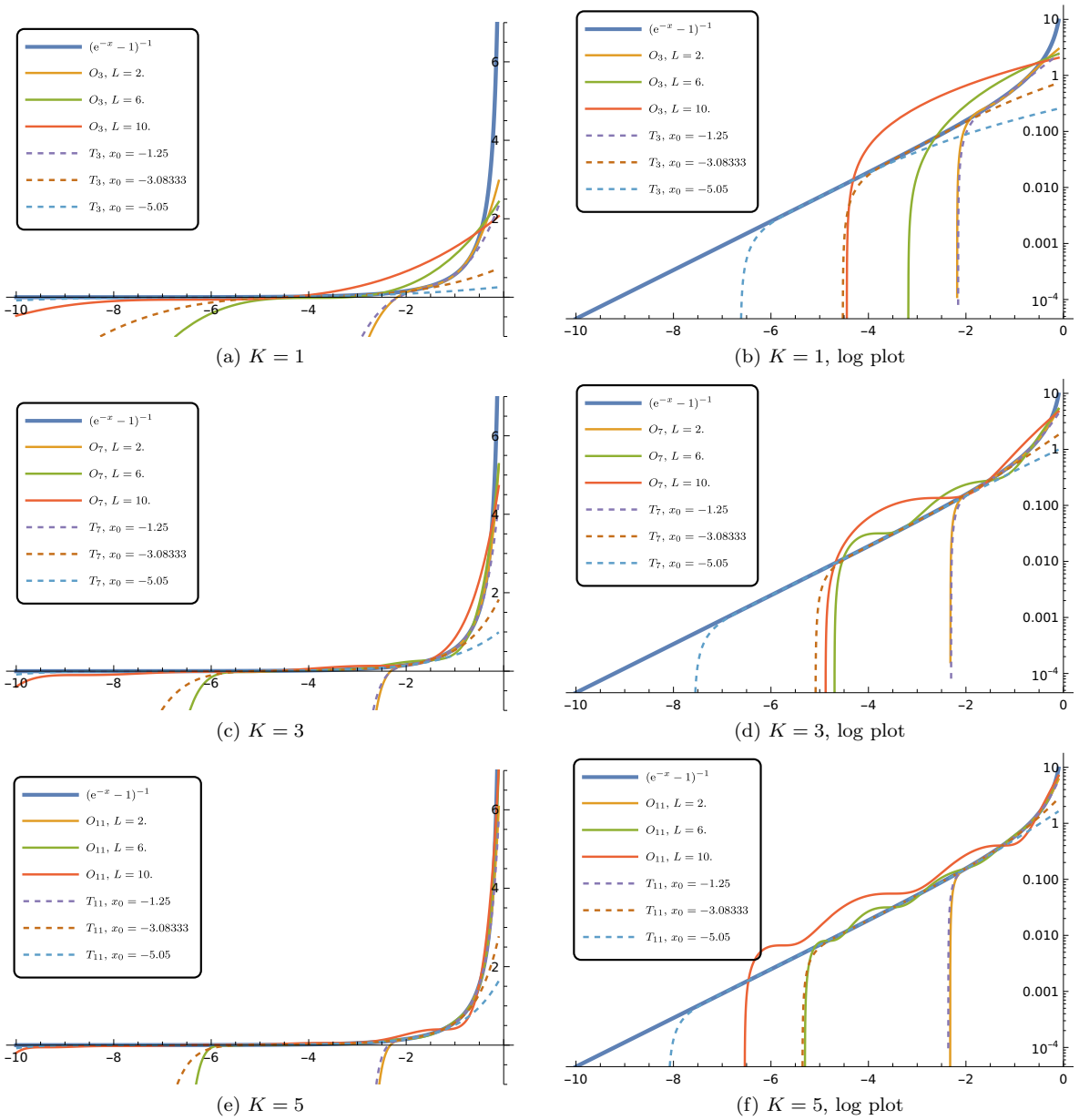
for a given vector of moments $\mathbf{U}$ of specific distributions. The vector function $\mathbf{m}(\Omega)$ is chosen as all real spherical harmonics up to a certain degree N. Here the function $\beta(\cdot)$ is either the β_{2K+1}, T_{2K+1} or O_{2K+1} approximation of the exponential function β_{BS}, and either the T_{2K+1} or O_{2K+1} approximation of the Planckian β_{BE}. The right-hand side $\mathbf{U}$ is given by the moments of a given function, which means we first choose a function $I(\Omega)$, and then set

$$\mathbf{U} = \int_{\mathbb{S}^2} \mathbf{m}(\Omega)I(\Omega)\,\mathrm{d}\Omega.$$

After solving $\boldsymbol{\lambda}$ from (13), the function $\beta(\boldsymbol{\lambda}^T\mathbf{m}(\Omega))$ is regarded as an approximation of $I(\Omega)$. For clarification, we will add the subscript N to the name of the model to denote the moment method, the first subscript N refers to the moment order and the second $2K+1$ to the degree of the polynomial approximation β_{2K+1}, T_{2K+1} or O_{2K+1}. For example, if we use spherical harmonics up to degree N in Ω and choose $\beta(\cdot)$ to be T_{2K+1}, the model is denoted as $T_{N,2K+1}$. Similarly, we will also consider the $\beta_{N,2K+1}$ and $O_{N,2K+1}$ models below.

In order to solve the equation (13), we use Newton's method, for which we need to compute the Jacobian

$$\int_{\mathbb{S}^2} \mathbf{m}(\Omega)[\mathbf{m}(\Omega)]^T\beta'(\boldsymbol{\lambda}^T\mathbf{m}(\Omega))\,\mathrm{d}\Omega.$$



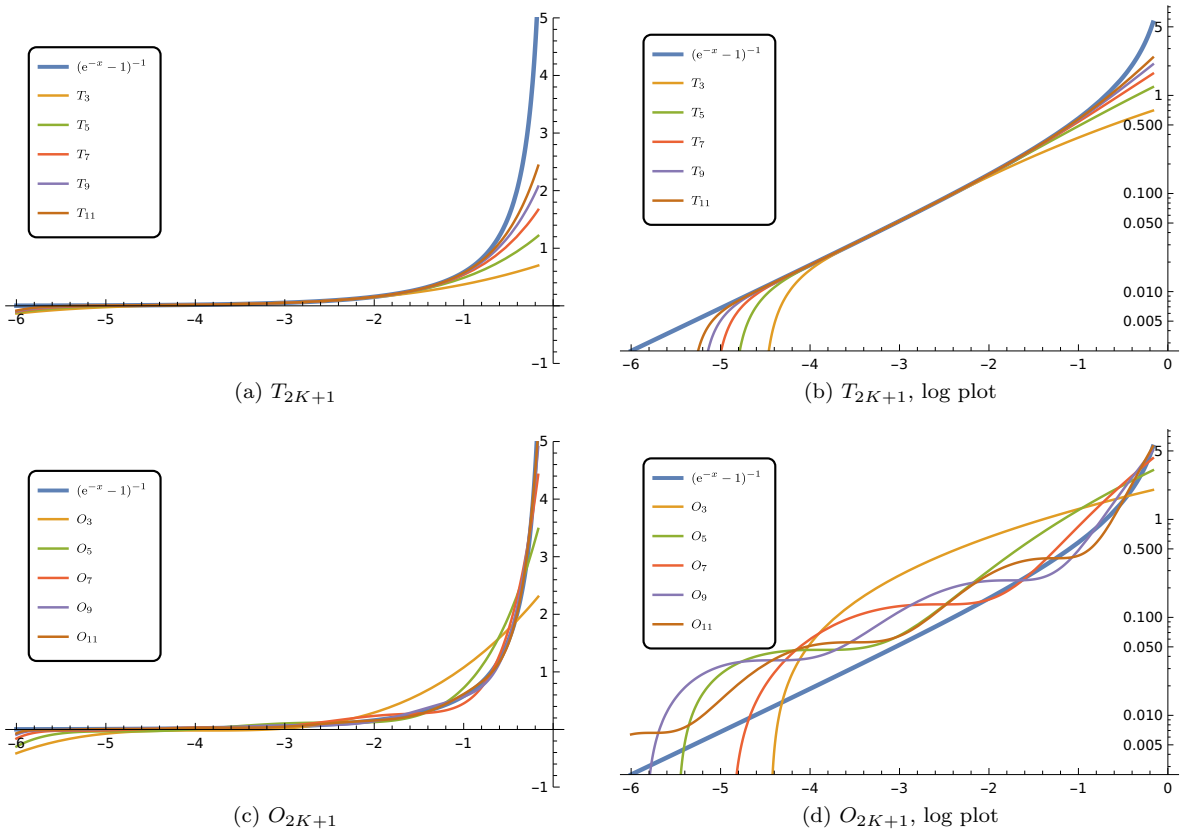
(a) $K=1$ (b) $K=1$, log plot

(c) $K=3$ (d) $K=3$, log plot

(e) $K=5$ (f) $K=5$, log plot

Figure 5: Taylor T_{2K+1} and optimized O_{2K+1} approximations of the Planckian β_{BE}.

Since the integrand is a polynomial of Ω, the integral can be computed exactly using appropriate integration formulas. Here we adopt the Lebedev quadrature (Lebedev (1976); Lebedev and Laikov (1999)). The number of quadrature points is chosen such that the degree of the quadrature is no less than the degree of the polynomial. The iteration stops if the L^2 norm of the residual reaches 10^{-10}. This method to solve (13) was described and studied numerically in Abdelmalik et al. (2023). The present modifications show similar features, especially the Newton method requires a few iterations to reach the desired accuracy.



(a) T_{2K+1} (b) T_{2K+1}, log plot

(c) O_{2K+1} (d) O_{2K+1}, log plot

Figure 6: Taylor T_{2K+1} and optimized O_{2K+1} approximations of the Planckian β_{BE}.

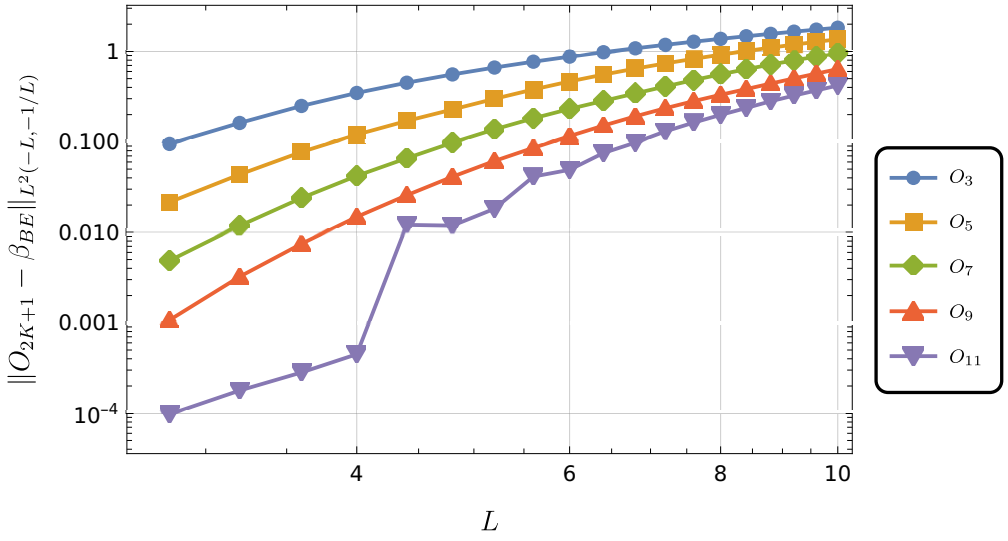


Figure 7: Error plot for the O_{2K+1} approximations of the exponential function β_{BS}

5.1 Single beam approximation

In this section, we apply these approximate entropy models in the approximation of a single beam. Since all the models are rotationally invariant, the direction of the beam does not affect the result. For simplicity, we

consider the approximation of the Dirac-delta function

$$I(\Omega) = \delta(\Omega - \Omega_0),$$

where $\Omega_0 = (0, 0, 1)^T$.

Remark that our method does not guarantee the non-negativity of the distribution. The non-negativity of the approximation is surely beneficial for physical interpretations, and it enforces an exact approximation of the Dirac deltas. However, it is not necessary to obtain a good approximation and relaxing this positivity constraints allows two major advantages for our closure: First, the characterization of the realizability domain, i.e. the set of admissible moment vectors, can be complicated for high order moments, and it is even an open problem in multi-D. Also numerical methods to solve the RTE with a positive approximation need to preserve this realizability domain from one step to another. This can be constraining when considering high order schemes. Relaxing the positivity constraint turns the realizability domain for our approximation into the entire Euclidean space and therefore this makes the simulation a lot easier. Second, if strongly anisotropic distributions are well-captured by positive approximations, the moment inversion problem becomes ill-conditioned for such distributions. It even turns ill-posed in the Dirac delta limit. Relaxing the positivity constraints preserves a finite condition of the optimization problem, other attempts to solve this high-conditioned optimization problem can be found in Hauck (2011); Alldredge et al. (2012, 2014, 2019). In our work, the polynomial degree K is a parameter that allows compromising between the accuracy and the computational cost in this limit. The higher K the better is the approximation of the Diracs, but also the higher is the condition of the optimization problem (8) and the higher the computational cost is needed to solve it. This parameter needs of course to be chosen depending on the considered problem.

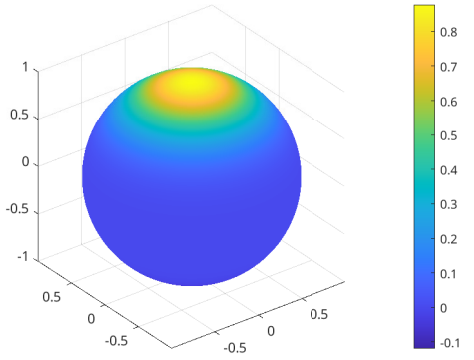
Some approximations for the Boltzmann-Shannon entropy with $N = 1$ and $K = 2$ are given in Fig. 8. The plots show that the $\beta_{1,5}$ model gives a remarkably better result than the $T_{1,5}$. This is not surprising because the value of the approximate function is all below 1.0, which corresponds to $\exp(x)$ with a negative x, where the $\beta_{1,5}$ model can give a better approximation. The quality of the $O_{1,5}$ model shown in Fig. 8c then lies in-between, since it approximates the exponential on the interval $[-5, 5]$, which balances both the positive and the negative parts. In order to improve the result, we can shift the domain to the negative side as in Fig. 8d, so that a result similar to the $\beta_{1,5}$ model can be obtained.

Similar phenomena can be observed when the approximation of the Bose-Einstein entropy is applied. The results are plotted in Fig. 9. For both $T_{1,2K+1}$ and $O_{1,2K+1}$ models, the approximate intensity function is closer to $I(\Omega)$ if the parameters are chosen to fit the range of the function values.

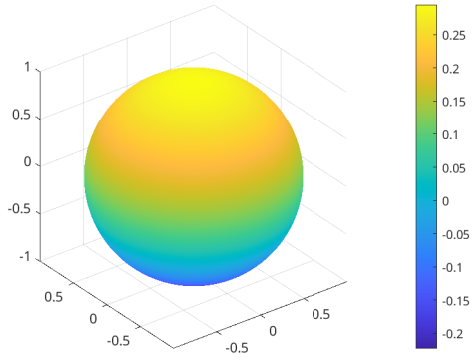
All the results above are only for $N = 1$ and $K = 2$. To improve the results, we can increase either N or K. Fig. 10 includes some results for $K = 6$. Increasing the value of K from 2 to 6 does provide improved result for all other parameters, but the beam is still widely spread for all cases. A more efficient way to get improvement is to increase N from 1 to 3. The results shown in Fig. 11 exhibit much sharper beams compared with all previous results, and the functions are mostly positive except the Taylor model with $x_0 = 0$. In this test case, the optimized model shows the highest peak value for both types of entropy. Meanwhile, all these results show that the $\beta_{N,2K+1}$ model studied in Abdelmalik et al. (2023) is also a good choice for problems involving beams when the Boltzmann-Shannon entropy is considered. However, this model does not have a counterpart for the Bose-Einstein entropy.

When approximating numerically Dirac distributions by means of a Galerkin approach of the form (6), the non-linear renormalization function β needs to accurately capture the zero values. When this β is additionnaly non-increasing, this corresponds to capturing accurately the flat tail $\beta(g) = 0$ when $g < 0$ is large negative. In the previous framework, both with β_{BS} and β_{BE}, this can be improved by choosing $x_0 < 0$ larger in the Taylor approximation or by having the interval $[c, d]$ encompass larger negative values.

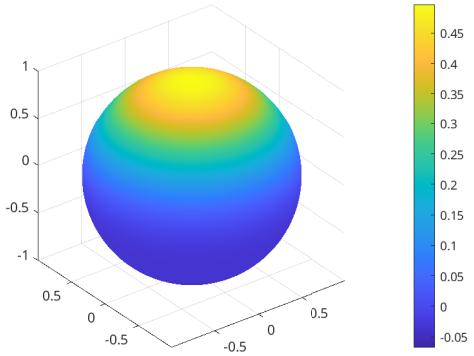
Now, we would like to study the convergence of the Newton method used in our experiments. Since the number of iterations depends on the quality of the initial value, we carry out the test by solving a perturbed moment inversion problem which approximates the Dirac-delta function $I^\epsilon(\Omega) = \delta(\Omega - \Omega_0^\epsilon)$. This moment inversion problem is solved using the function that approximates $I(\Omega)$ as its initial value in the Newton iteration. This mimics the situation in a dynamical problem where a good initial value can usually obtained by taking the solution in the previous time step. In our test, we choose $\Omega_0^\epsilon = (0, \sin(\pi/20), \cos(\pi/20))^T \approx$



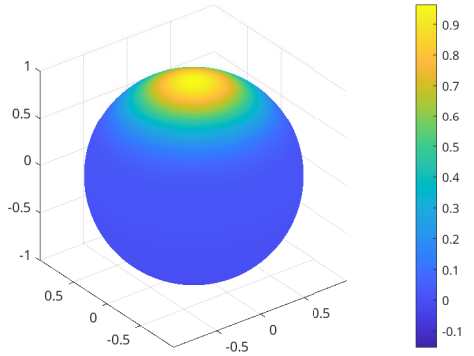
(a) $\beta_{1,5}$



(b) $T_{1,5}$, B-S entropy, $x_0 = 0$



(c) $O_{1,5}$, B-S entropy, $[-5, 5]$



(d) $O_{1,5}$, B-S entropy, $[-10, 0]$

Figure 8: Approximation of a single beam using different models approximating the result of the maximum Boltzmann-Shannon entropy.

$(0, 0.156, 0.988)^T$, which is a quite significant perturbation. In Tab. 1, we list the number of iterations required for all the methods appeared in the previous tests, which shows that the Newton method is highly efficient for the moment inversion problem. Recall that the stopping criterion is the L^2 norm of the residual less than 10^{-10}, which is quite strict in applications.

5.2 Double beam approximation

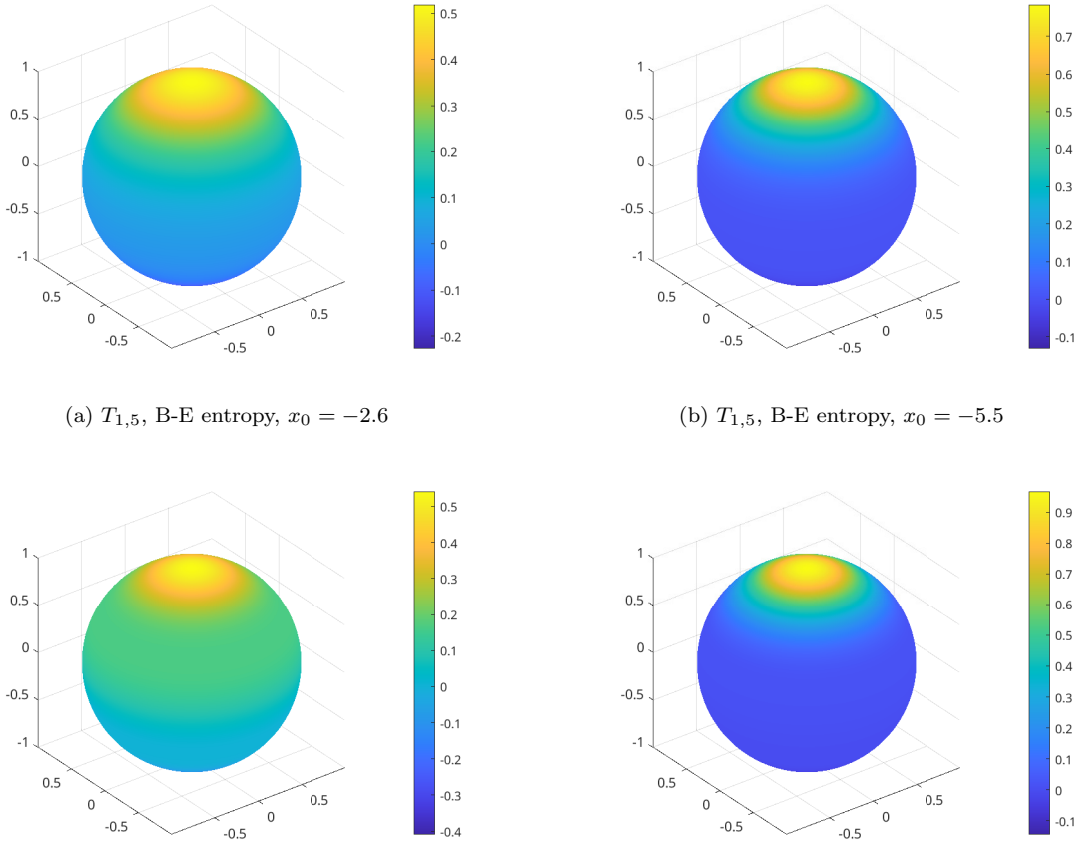
We now considers the approximation of the following function:

$$I(\Omega) = \delta(\Omega - \Omega_1) + \delta(\Omega - \Omega_2),$$

which often occurs when two beams cross each other. Here we focus only on the $T_{N,2K+1}$ and $O_{N,2K+1}$ models, and we refer the readers to Abdelmalik et al. (2023) for results of the $\beta_{N,2K+1}$ models.

Since there are two beams in the intensity function, a model with $N = 1$ cannot give a meaningful approximation. In our experiments, we choose $\Omega_1 = (0, 0, 1)^T$, $\Omega_2 = (1, 0, 0)^T$ and use $N = 3$ and $N = 9$ in the approximation. Other parameters are chosen to be the better combination in the previous subsection. In particular, the value of K is fixed to be 2 in all our examples.

For $N = 3$ (see Fig. 12), the two Taylor models give quite similar results. The two beams are correctly detected with correction locations, and they are both smeared out due to the smooth approximation. The



(a) $T_{1,5}$, B-E entropy, $x_0 = -2.6$

(b) $T_{1,5}$, B-E entropy, $x_0 = -5.5$

(c) $O_{1,5}$, B-E entropy, $[-5, -0.2]$

(d) $O_{1,5}$, B-E entropy, $[-10, -1]$

Figure 9: Approximation of a single beam using different models approximating the result of the maximum Bose-Einstein entropy.

two optimized results provide sharper beams, as the peak value of the distribution is higher. Negative values can still be spotted near the point $(0, -1, 0)^T$, which can be improved by increasing N or K. Here we only perform experiments with $N = 9$, which can be found in Fig. 13. The two bright spots are much more pointy than the results of $N = 3$, and the peak values are now significantly higher. Again, the $O_{N,2K+1}$ models perform slightly better than the $T_{N,2K+1}$ models for both types of entropy.

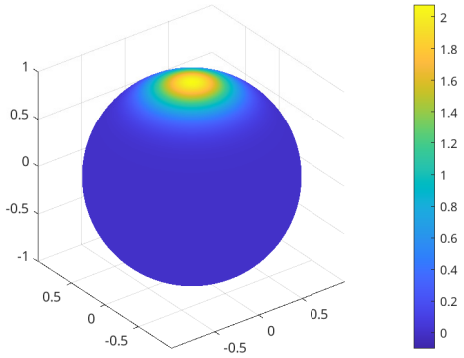
5.3 Approximating smooth functions

We now consider the the approximation of smooth functions and hope to observe spectral convergence. We take the six-Gaussian function considered in Abdelmalik et al. (2023):

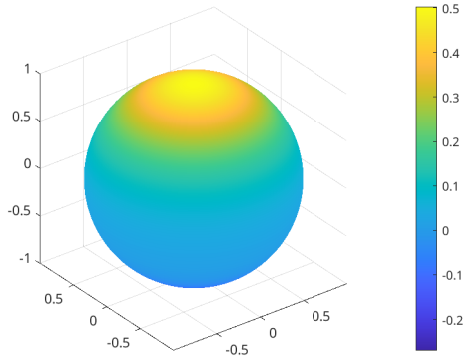
$$I(\Omega) = \sum_{k=1}^{6} \exp\left(-5\|\Omega - \Omega_k\|^2\right),$$

where

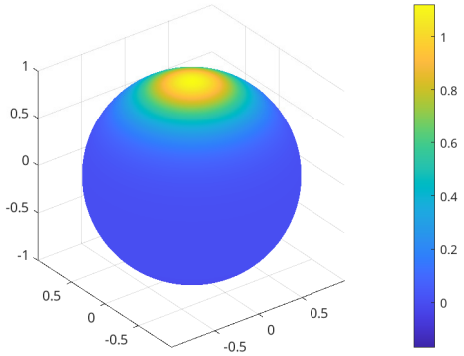
$$\Omega_1 = (1, 0, 0)^T, \quad \Omega_2 = (-1, 0, 0)^T, \quad \Omega_3 = (0, 1, 0)^T,$$
$$\Omega_4 = (0, -1, 0)^T, \quad \Omega_5 = (0, 0, 1)^T, \quad \Omega_6 = (0, 0, -1)^T.$$



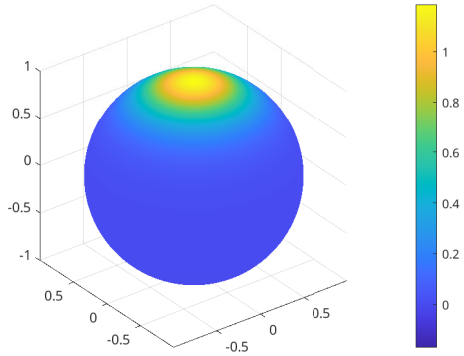
(a) $\beta_{1,13}$



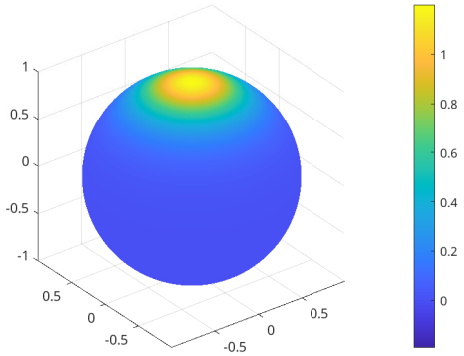
(b) $T_{1,13}$, B-S entropy, $x_0 = 0$



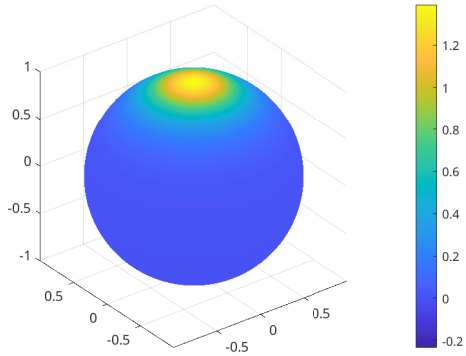
(c) $T_{1,13}$, B-S entropy, $x_0 = -5$



(d) $O_{1,13}$, B-S entropy, $[-10, 0]$



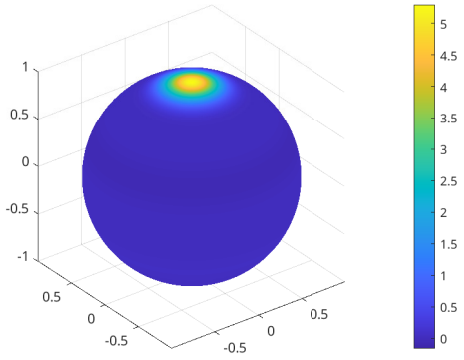
(e) $T_{1,13}$, B-E entropy, $x_0 = -5.5$



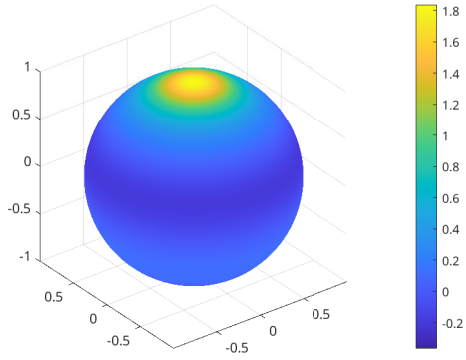
(f) $O_{1,13}$, B-E entropy, $[-10, -1]$

Figure 10: Approximation of a single beam using different models

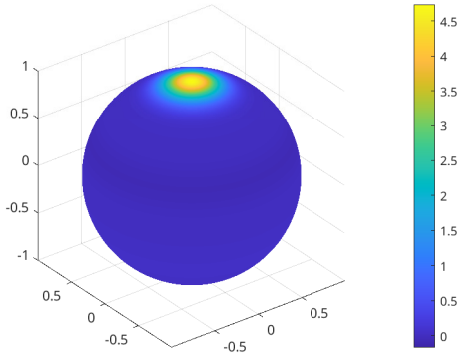
This function and its approximation using the $\beta_{5,5}$ model are plotted in Fig. 14. It can be seen that the $\beta_{5,5}$ approximation overestimates the peak value. Fig. 15 shows some approximations based on the Boltzmann-Shannon entropy. It can be seen by naked eyes that the $O_{5,5}$ model defined by optimization on $[-5, 5]$ gives



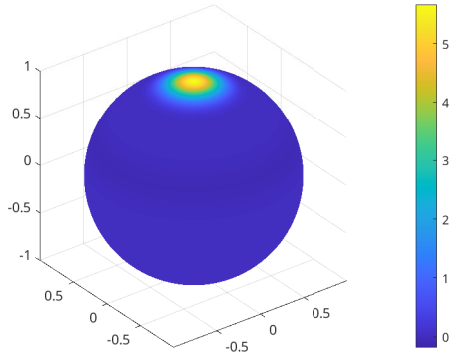
(a) $\beta_{3,5}$



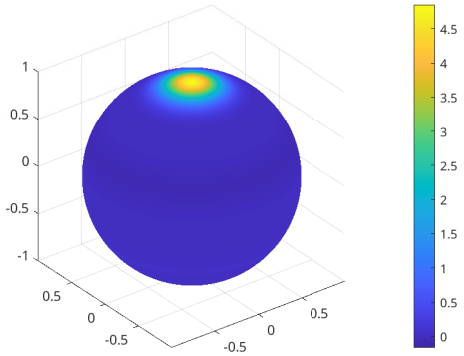
(b) $T_{3,5}$, B-S entropy, $x_0 = 0$



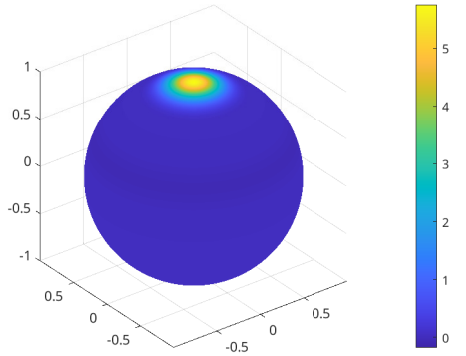
(c) $T_{3,5}$, B-S entropy, $x_0 = -5$



(d) $O_{3,5}$, B-S entropy, $[-10, 0]$



(e) $T_{3,5}$, B-E entropy, $x_0 = -5.5$



(f) $O_{3,5}$, B-E entropy, $[-10, -1]$

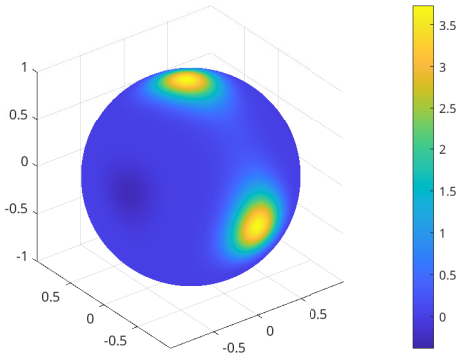
Figure 11: Approximation of a single beam using different models

the best result.

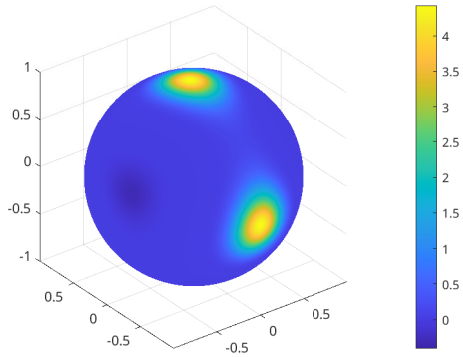
Fig. 16 shows the error decay as N increases. Although all methods provide spectral convergence, the choice of parameters does affect the convergence rate. In this example, the $O_{N,5}$ model optimized on $[-10, 0]$

Table 1: Number of Newton iterations for the perturbed problem

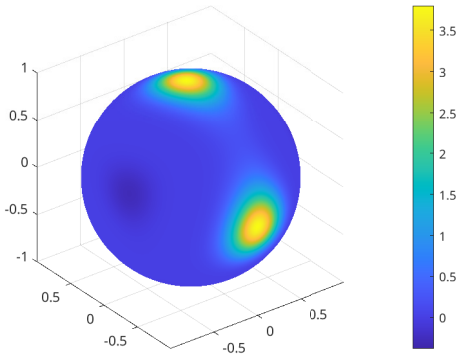
B-S entropy		B-E entropy	
Method	No. of iterations	Method	No. of iterations
$T_{1,5}$, $x_0 = 0$	3	$T_{1,5}$, $x_0 = -2.6$	4
$T_{1,13}$, $x_0 = 0$	4	$T_{1,5}$, $x_0 = -5.5$	5
$T_{1,13}$, $x_0 = -5$	5	$T_{1,13}$, $x_0 = -5.5$	5
$T_{3,5}$, $x_0 = 0$	4	$T_{3,5}$, $x_0 = -5.5$	7
$T_{3,5}$, $x_0 = -5$	7		
$O_{1,5}$, $[-5, 5]$	4	$O_{1,5}$, $[-5, -0.2]$	5
$O_{1,5}$, $[-10, 0]$	5	$O_{1,5}$, $[-10, -1]$	5
$O_{1,13}$, $[-10, 0]$	5	$O_{1,13}$, $[-10, -1]$	5
$O_{3,5}$, $[-10, 0]$	8	$O_{3,5}$, $[-10, -1]$	8



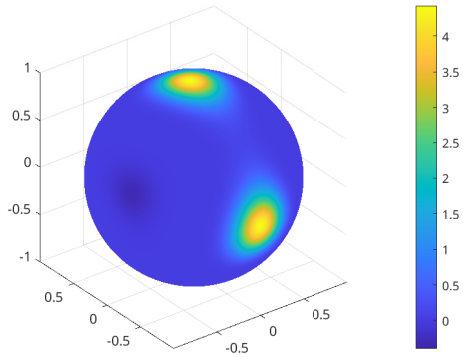
(a) $T_{3,5}$, B-S entropy, $x_0 = -5$



(b) $O_{3,5}$, B-S entropy, $[-10, 0]$



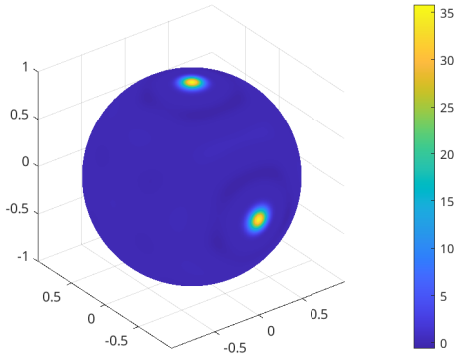
(c) $T_{3,5}$, B-E entropy, $x_0 = -5.5$



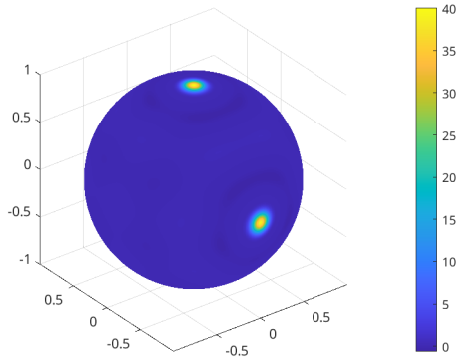
(d) $O_{3,5}$, B-E entropy, $[-10, -1]$

Figure 12: Approximation of a double-beam function using different models.

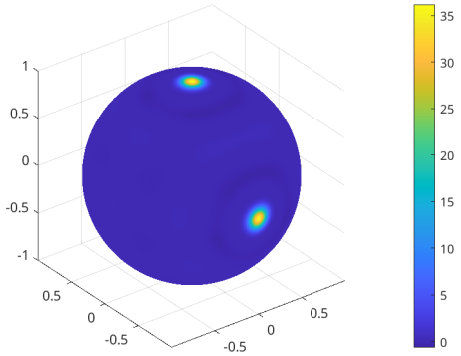
and the $T_{N,5}$ model with $x_0 = -5$ have similar performance, which is worse than the $O_{N,5}$ model optimized on the interval $[-5, 5]$ but better the $T_{N,5}$ model with $x_0 = 0$. Note that the convergence rate of these models is not determined by the quality of approximation to the β function. This is also observed in Abdelmalik et al. (2023), where the P_N (which corresponds to $\beta_{N,1}$) model has the best convergence rate among all



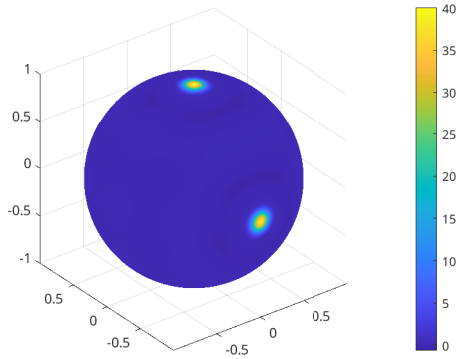
(a) $T_{9,5}$, B-S entropy, $x_0 = -5$



(b) $O_{9,5}$, B-S entropy, $[-10, 0]$

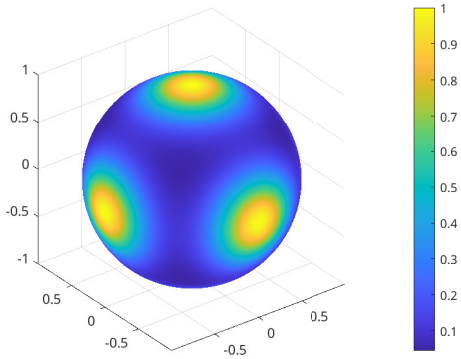


(c) $T_{9,5}$, B-E entropy, $x_0 = -5.5$

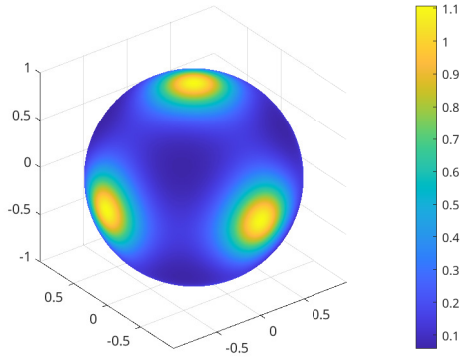


(d) $O_{9,5}$, B-E entropy, $[-10, -1]$

Figure 13: Approximation of a double-beam function using different models.

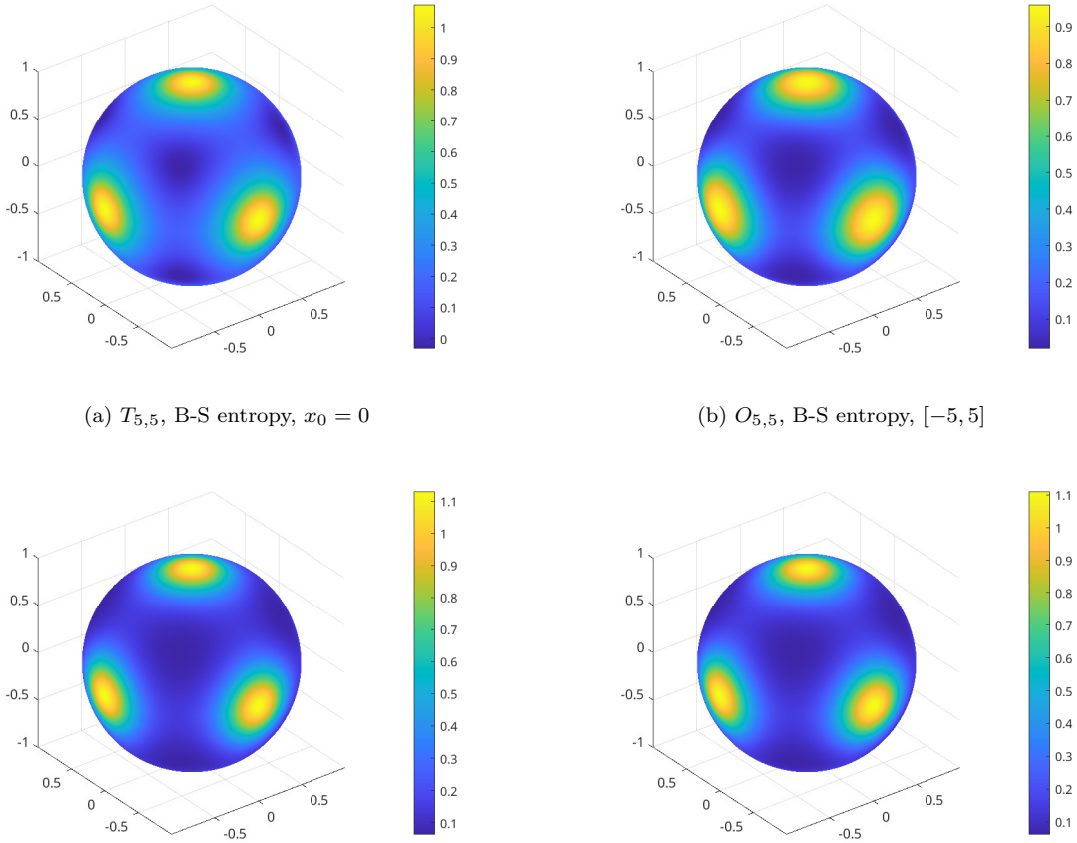


(a) Six-Gaussian function



(b) $\beta_{5,5}$ approximation

Figure 14: The six-Gaussian function and its $\beta_{5,5}$ approximation.



(a) $T_{5,5}$, B-S entropy, $x_0 = 0$

(b) $O_{5,5}$, B-S entropy, $[-5, 5]$

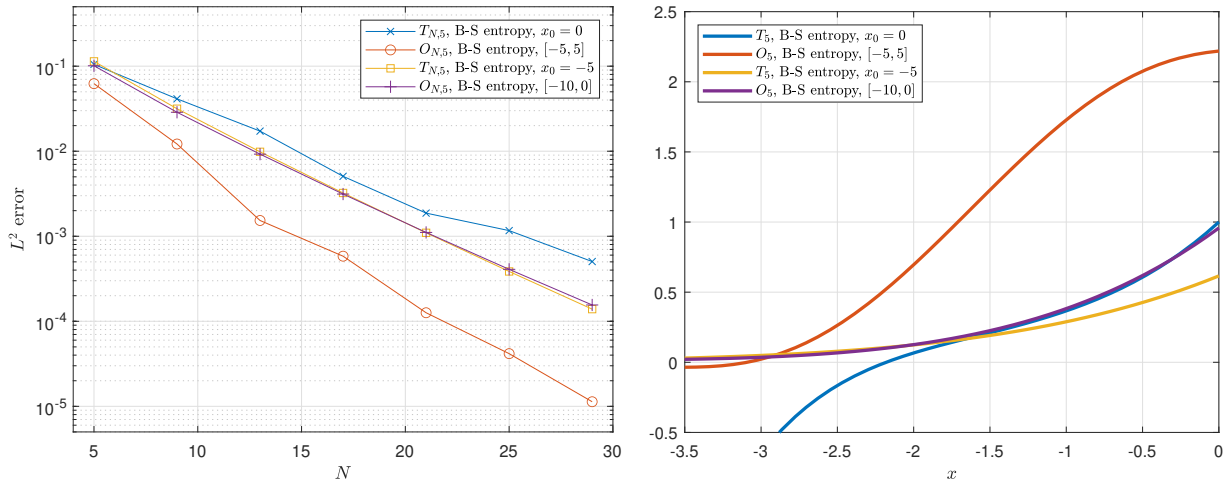
(c) $T_{5,5}$, B-S entropy, $x_0 = -5$

(d) $O_{5,5}$, B-S entropy, $[-10, 0]$

Figure 15: Approximations of the six-Gaussian function based on the Boltzmann-Shannon entropy.

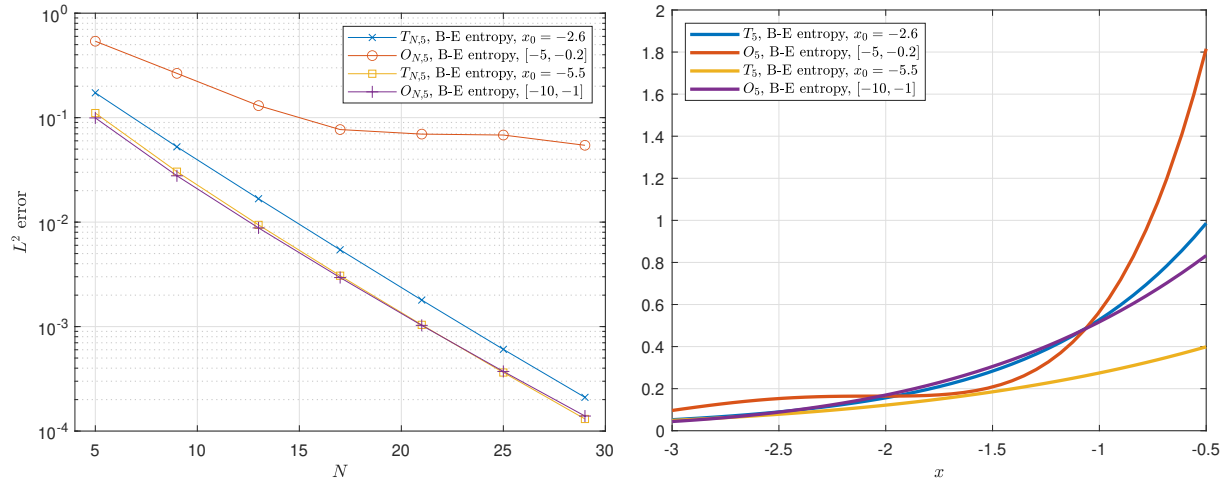
$\beta_{N,K}$ models, although the $\beta_{N,1}$ model only approximates the exponential function by $\mathrm{e}^x \approx 1 + x$, which is a poor fit. Here we conjecture that the $O_{N,5}$ model optimized on $[-5, 5]$ has a better convergence rate because the function $O_5(x)$ is relatively closer to a linear function with slope 1 for a certain range on the negative part of the domain (see Fig. 16b). We focus on the negative part because the value of $I(\Omega)$ is mostly between 0 and 1. This means the $O_{N,5}$ model optimized on $[-5, 5]$ is likely to be closer to the P_N model when approximating this function.

Now we test the performance of the methods based on the Bose-Einstein entropy. The results are shown in Fig. 17. It can be seen that the spectral convergence is observed for three models except the $O_{N,5}$ method optimized on the interval $[-5, -0.2]$. This is likely due to the flatness of the O_5 function on the interval from $[-2.5, -1.5]$ caused by enforcing good approximations in the region closer to zero where the Planckian has larger values. This can be improved by slightly shifting the upper bound of the domain. Fig. 17a gives the result for $O_{N,5}$ method optimized on $[-5, -0.5]$, where a much better convergence rate is obtained. Nevertheless, since the choice of the domain is not optimized, the numerical error is still significantly larger than other lines in Figure 17a.



(a) Error decay (b) Approximation of the exponential

Figure 16: Error decay for approximations of the six-Gaussian function based on the Boltzmann-Shannon entropy and the corresponding approximations of the exponential function



(a) Error decay (b) Approximation of the exponential

Figure 17: Error decay for approximations of the six-Gaussian function based on the Bose-Einstein entropy and the corresponding approximations of the exponential function

6 Conclusion and discussion

In this work, we have considered the approximation of intensity functions based on the maximum entropy method, where the entropy can be chosen according to the underlying physics such as the Boltzmann-Shannon's or Bose-Einstein's entropy. The approximation can be applied in the construction of moment equations that dissipate the selected entropy. Polynomials are adopted in the approximation so that when performing the moment closure, integration can be carried out exactly in the moment inversion problem. Such a method can be considered as an extension of the previous work Abdelmalik et al. (2023), which includes only a specific approximation based on the B-S entropy. Our method works for a wider range of entropy functions, and some additional parameters such as the reference point for the Taylor expansion and the interval for the optimization problem are introduced. It is demonstrated by numerical tests that the

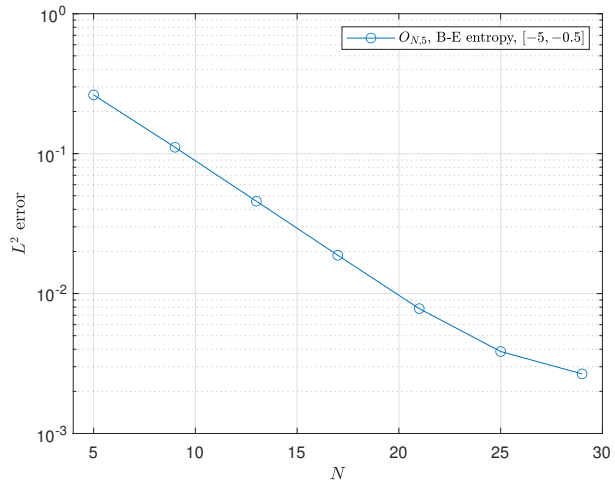


Figure 18: Error decay for approximations of the six-Gaussian function based on the Bose-Einstein entropy

quality of the moment closure depends on the choice of these parameters, but in most cases, decent results can be obtained by optimizing the distance between the polynomial and the physical entropy.

In our future work, we are going to further study their performance by applying these moment models to the radiative transfer equation with interaction with matter. In Abdelmalik et al. (2023), experiments have been done for the $\beta_{N,K}$ model applied to the monochromatic radiative transfer equation without interaction with matter, showing some promising results. For the monochromatic model, we do not expect that using the B-E entropy gives better results, since this model is not clearly connected to the B-E entropy. A more meaningful simulation with the B-E entropy should involve the frequency variable and the temperature of matter as in (3), requiring more sophisticated numerical techniques. Since our models include more parameters, there are more possibilities in the simulation of the radiative transfer equation. For instance, different moment closures based on optimization on different intervals can be chosen for different spatial locations and differential frequencies, allowing better approximations to the original maximum-entropy methods.

Acknowledgements

Zhenning Cai was supported by the Academic Research Fund of the Ministry of Education of Singapore under grant No. A-0004592-00-00.

References

Abdelmalik, M. (2017). *Adaptive algorithms for optimal multiscale model hierarchies of the Boltzmann Equation: Galerkin methods for kinetic theory.* Ph. D. thesis, TU Eindhoven.

Abdelmalik, M., Z. Cai, and T. Pichard (2023). Moment methods for the radiative transfer equation based on φ-divergences. *Comput. Meth. Appl. 417*, 1–29.

Abdelmalik, M. and H. van Brummelen (2016). Moment closure approximations of the boltzmann equation based on φ-divergences. *J. Stat. Phys.* (164), 77–104.

Alldredge, G. W., M. Frank, and C. D. Hauck (2019). A regularized entropy-based moment method for kinetic equations. *SIAM J. Appl. Math. 79*(5).

Alldredge, G. W., C. D. Hauck, D. P. O'Leary, and A. L. Tits (2014, february). Adaptive change of basis in entropy-based moment closures for linear kinetic equations. *J. Comp. Phys. 74*(4), 489–508.

Alldredge, G. W., C. D. Hauck, and A. L. Tits (2012). High-order entropy-based closures for linear transport in slab geometry II: A comutational study of the optimization problem. *SIAM J. Sci. Comput. 34*(4), 361–391.

Canuto, C., M. Y. Hussaini, A. Quarteroni, and T. A. Zang (2006). *Spectral Methods: Fundamentals in Single Domains*. Springer-Verlag.

Carlson, B. (1963). *Methods in computational physics*. ACADEMIC PRESS.

Chandrasekhar, S. (1950). *Radiative transfer*. Oxford.

Csiszár, I. (1972). A class of measures of informativity of observation channels. *Period. Math. Hung. 2*, 191–213.

Dautray, R. and J.-L. Lions (2000). *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 6, Evolution Problems II*. Springer.

Dubroca, B. and J.-L. Feugeas (1999). Entropic moment closure hierarchy for the radiative transfer equation. *C. R. Acad. Sci. Paris Ser. I 329*, 915–920.

Frank, M., A. Klar, E. Larsen, and S. Yasuda (2007). Time-dependent simplified pn approximation to the equations of radiative transfer. *J. Comput. Phys. 226*(2), 2289–2305.

Godlewski, E. and P.-A. Raviart (1996). *Numerical approximation of hyperbolic systems of conservation laws*. Springer.

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

Hesthaven, J. S., S. Gottlieb, and D. Gottlieb (2009). *Spectral Methods for Time-Dependent Problems*. Cambridge.

Humbird, K. D. and R. McClarren (2017). Adjoint-based sensitivity analysis for high-energy density radiative transfer using flux-limited diffusion. *High Energy Density Physics 22*, 12–16.

Kawashima, S. and W.-A. Yong (2004). Dissipative structure and entropy for hyperbolic systems of balance laws. *Arch. Rational Mech. Anal. 174*, 345–364.

Lasserre, J.-B. (2009). *Moment, positive polynomials, and their applications*. Imperial college press.

Lebedev, V. (1976). Quadratures on a sphere. *USSR Comput. Math. Math. Phys. 16*(2), 10–24.

Lebedev, V. and D. Laikov (1999). A quadrature formula for the sphere of the 131st algebraic order of accuracy. *Doklady Mathematics 59*(3), 477–481.

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

Lewis, E. E. and W. F. Miller (1984). *Computational Methods of Neutron Transport*. John Wiley & Sons Inc.

Li, R. and W. Li (2020). 3D B_2 model for radiative transfer equation. *Int. J. Numer. Anal. Modeling 17*(1), 118–150.

Lowrie, R. B., J. E. Morel, and J. A. Hittinger (1999). The coupling of radiation and hydrodynamics. *The Astro. J. 521*, 432–450.

McClarren, R. (2010). Theoretical aspects of the simplified pn equations. *Transport Theory Stat. Phys. 39*(2-4), 73–109.

Mihalas, D. and B. R. W. Mihalas (1983). *Foundations of Radiation Hydrodynamics*. Oxford.

Minerbo, G. N. (1978). Maximum entropy Eddington factors. *J. Quant. Spectros. Radiat. Transfer 20*, 541–545.

Olson, G. L., L. H. Auer, and M. L. Hall (2000). Diffusion, p1, and other approximate forms of radiation transport. *J. Quant. Spec. Rad. Trans. 64* (6), 619–634.

Pichard, T. (2020). A moment closure based on a projection on the boundary of the realizability domain: 1d case. *Kin. rel. models* (13), 1243–1280.

Pichard, T., G. W. Alldredge, S. Brull, B. Dubroca, and M. Frank (2017). An approximation of the M_2 closure: application to radiotherapy dose simulation. *J. Sci. Comput.* (71), 71–108.

Pomraning, G. C. (1973). *Equations of Radiation Hydrodynamics*. Pergamon.

Sarr, J. A. R. and C. P. T. Groth (2020). A second-order maximum-entropy inspired interpolative closure for radiative heat transfer in gray participating media. *J. Quant. Spectros. Radiat. Transfer*, 107238.

Schmuedgen, K. (2017). *The moment problem*. Springer.

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

Szegö, G. (1939). *Orthogonal polynomials*. AMS.