Uncertainty propagation; intrusive kinetic formulations of scalar conservation laws

Bruno Després, Benoît Perthame

Uncertainty propagation; intrusive kinetic formulations of scalar conservation laws

Bruno Després*† Benoît Perthame*†‡

April 27, 2015

Abstract

We study two intrusive methods for uncertainty propagation in scalar conservation laws based on their kinetic formulations. The first one is based on expansions on an orthogonal family of polynomials.

The first method uses convolutions based on Jackson kernels and we prove that it satisfies BV bounds and converges to the entropy solution but with a spurious damping phenomenon. Therefore we introduce a second method, which is based on projection on layered Maxellians, and which arises as a minimization of entropy. Our construction of layered Maxwellians relies on the Bojavic-Devore theorem about best L^1 polynomial approximation. This new method, denoted below as a kinetic polynomial method, satisfies the maximum principle by construction as well as partial entropy inequalities and thus provides an alternative to the standard method of moments which, in general, does not satisfy the maximum principle.

Simple numerical simulations for the Burgers equation illustrate these theoretical results.

Key words Uncertainty propagation, kinetic formulation of conservation laws, maximum principle, entropy dissipation, chaos polynomial.
Mathematics Subject Classification (2010) 35L65; 35R60; 35A35

1 Introduction

We address the question of constructing intrusive kinetic methods for scalar conservation laws in view of uncertainty quantification (UQ) and propagation. The starting point is a scalar conservation law in dimension d

$$\partial_t u + \nabla \cdot F(u) = 0, \tag{1}$$

where u is the unknown, $x \in \mathbb{R}^d$ is the space variable, $t \geq 0$ is the time variable and $F : \mathbb{R} \to \mathbb{R}^d$ is the flux (we use at some places that its second derivatives is locally bounded). We assume the solution u is a function of an additional variable ω, called the uncertainty variable. For simplicity of notations, we take in this work $\omega \in I \subset \mathbb{R}$, a bounded interval. It can be generalized to $\omega \in I^p \subset \mathbb{R}^p$, $p > 1$ by tensorisation.

*Sorbonne Universités, UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France, Email: benoit.perthame@upmc.fr

†CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France

‡INRIA-Paris-Rocquencourt, EPC MAMBA, Domaine de Voluceau, BP105, 78153 Le Chesnay Cedex, France

The unknown u is therefore a function of (x, ω, t), meaning that we would like to solve an infinite series of conservation laws (1), where the dependency on ω stems from the initial data. Throughout all this paper, we assume it is nonnegative, upper bounded by some $U_M > 0$ and belongs to $L^1(dxd\mu(\omega))$

$$0 \leq u(x, \omega, 0) = u^{\text{init}}(x, \omega) \leq U_M, \qquad \int u^{\text{init}}(x, \omega) dx d\mu(\omega) < \infty. \tag{2}$$

A central hypothesis is that the uncertainty obeys a certain given statistic, with probability density $d\mu(\omega)$, that is $\int d\mu(\omega) = 1$. This information is used to derive reduced models. For this task, one determines the so-called *chaos polynomials* $p_i(\omega)$, with $d^o p_i = i$, which are orthonormal for the measure $d\mu(\omega)$

$$\int p_i(\omega) p_j(\omega) d\mu(\omega) = \delta_{ij}.$$

The terminology chaos polynomials has been coined in the seminal work of Wiener [22]. Such methods are widely used in engineering, see for example [15, 3] and references therein. Since the polynomials determine a Hilbert basis of L^2_μ, and using the vector space P^N_ω of polynomials of degree less or equal to N, any function $v \in L^2_\mu$ is expanded and approximated as

$$v = \sum_{i=0}^{\infty} v_i p_i, \qquad v_i = \int v(\omega) p_i(\omega) d\mu(\omega), \qquad v^{(N)} = \sum_{i=0}^{N} v_i p_i \in P^N_\omega. \tag{3}$$

The approximation $v^{(N)}$ has optimal spectral accuracy properties which have been proved in [7, 1]. We use the notation with parenthesis to distinguish it from other polynomial approximations to be introduced later.

In order to construct a closed system for the evolution of the moments for solutions of (1), one can use the so-called *method of moments*, and consider the model system of conservation laws,

$$\partial_t u^N_i + \nabla \cdot \int F(u^N) p_i(\omega) d\mu(\omega) = 0, \quad 0 \leq i \leq N. \tag{4}$$

Following [8] and the references therein, such a method can be also rewritten as

$$\begin{cases} \partial_t u^N + \nabla \cdot F^{(N)}(u^N) = 0, \qquad F^{(N)}(u^N) = \displaystyle\sum_{i=0}^{N} \left(\int F(u^N(\omega')) p_i(\omega') d\mu(\omega') \right) p_i(\omega), \\ u^N(x, \omega, t) = \displaystyle\sum_{i=0}^{N} u^N_i(x, t) p_i(\omega). \end{cases} \tag{5}$$

In particular, in [8], the authors show how to model uncertainty in the coefficients F of the equations (1)–(2), using an additional variable. Even if the theory of existence, uniqueness and the theory of numerical approximation is well established for scalar conservation laws like (1), there is no such theory for systems of partial differential equations like (4) or for the convergence of the solutions of (4) towards solutions of (1) parametrized by the parameter ω, see [8]. Some results may be found nevertheless in [10] for the advection equation, in [8] where spectral convergence is proved with a weak-strong method but before any shock, and in [20] for Monte-Carlo methods applied to conservation laws. Note also the convergence proof in the recent work [4], for the method of moments in the framework of kinetic equations. We also refer to [13] for conservation laws with random right hand

sides and to [12] for an asymptotic-preserving extension to transport equations with random inputs. One of the reason of these mathematical difficulties lies in the divorce between the standard L^1-based theory for conservation laws, and the L^2-based theory of spectral approximation. Similar concerns may be found in a PDE context in [5, 14], or in [11] for l^1 minimization algorithmic issues related to the use of polynomial chaos expansions.

To circumvent these difficulties, we propose to consider the kinetic formulation of conservation laws which has the advantage of reconciliating L^1 and L^2 theories. It writes as a Boltzmann equation for $t \geq 0$, $x \in \mathbb{R}^d$ and $\xi \geq 0$, in a BGK (relaxation) form,

$$\left\{ \begin{array}{l} \partial_t f_\varepsilon + a(\xi).\nabla f_\varepsilon + \frac{1}{\varepsilon} f_\varepsilon = \frac{1}{\varepsilon} M(u_\varepsilon;\xi), \qquad a = \nabla F, \\ \\ u_\varepsilon(x,t) = \displaystyle\int f_\varepsilon(x,\xi,t) d\xi, \\ \\ f_\varepsilon(t=0) = M\big(u^{\rm init};\xi\big), \end{array} \right. \tag{6}$$

still assuming (2) for $u^{\rm init}$, and

$$M(u;\xi) = \mathbb{1}_{\{0<\xi<u\}} \tag{7}$$

is called a *Maxwellian* in the rest of this work. Notice that the non negativity $u \geq 0$ is needed for this definition to make sense. That is why we assume the initial data is non negative $u^{\rm init} \geq 0$ throughout this work. This assumption simplifies some non essential technicalities and allows us to disregard the negative part of M; the reader can find in [19, 16, 18] the adaptation for general sign of the initial data as well as convergence proofs of (6) to (1). We recall, for later use, that $M(u;\xi)$ is a universal minimizer for a family of *entropy functionals* [5, 19, 16, 18]; for all convex functionals $S(\xi)$,

$$M(u;\cdot) = \underset{u=\int g d\xi,\ 0\leq g\leq 1}{\rm argmin} \int S'(\xi) g d\xi. \tag{8}$$

The idea pursued in this work is firstly to write (6) for all ω, and secondly to modify it in a polynomial manner so as to consider the equation (we call it the *intrusive kinetic formulation*)

$$\left\{ \begin{array}{l} \partial_t f_\varepsilon^N + a(\xi).\nabla f_\varepsilon^N + \frac{1}{\varepsilon} f_\varepsilon^N = \frac{1}{\varepsilon} M^N\big(u_\varepsilon^N;\xi,\omega\big), \\ \\ u_\varepsilon^N(x,\omega,t) = \displaystyle\int f_\varepsilon^N(x,\xi,\omega,t) d\xi, \\ \\ f_\varepsilon^N(t=0) = M^N\big(u^{\rm init};\xi,\omega\big), \end{array} \right. \tag{9}$$

where $0 \leq M^N(u_\varepsilon^N;\xi,\omega) \leq 1$ is a suitable polynomial modification of the Maxwellian M. Notice that $\int f_\varepsilon^N(t=0) d\xi d\omega = \int u^{\rm init} d\omega$ but the initial data needs not be at 'equilibrium' since $u^{\rm init}$ usually does not belong to P_ω^N.

The function f_ε^N is naturally a polynomial in ω of degree less or equal to N and we will show the intrusive kinetic formulation (9) of conservation laws is endowed with some convenient mathematical properties such as, for all $t \geq 0$,

$$0 \leq f_\varepsilon^N \leq 1, \qquad \left\| \int f_\varepsilon^N(x,\xi,\omega,t) d\xi \right\|_{L^\infty_{x\omega}} \leq \left\| \int f_\varepsilon^N(x,\xi,\omega,0) d\xi \right\|_{L^\infty_{x\omega}}.$$

Notice that the solutions of (9) depend now on two parameters ε and N. We describe two families of methods to build the kinetic polynomial M^N.

The first construction is based on convolution kernels such as Fejer or Jackson kernels and reads $M^N(u_\varepsilon^N;\xi,\omega) = G^N *_\omega M(u_\varepsilon^N;\xi)$ for some apropriate kernel G. Since this approach has natural comparison inequalities, it is possible to pass for all time t to the limit $\varepsilon \to 0$, but with $N \to \infty$. However $\varepsilon N^2 \to \infty$ is needed for the limit to satisfy the correct equation (1) for all ω, in the weak sense. Nevertheless, this result is a strong improvement with respect to [8] where the convergence is proved by means of a weak-strong method but only before the time of shock (for example if there is a discontinuity in the initial data, the result [8] simply does not apply). As shown below, a similar property holds for strong convergence, which is a corollary of Proposition 2.10.

Theorem 1.1 (Convolved Maxwellian) *Consider the Jackson kernel. Assume the initial data is BV in all variables. One has convergence in L^1_{loc} of f_ε^N towards f with a rate of strong convergence $O\left(\frac{1}{\varepsilon N}\right)$, and a rate of weak convergence $O\left(\frac{1}{\varepsilon N^2}\right)$.*

This unfortunate consequence -the limit $\varepsilon \to 0$ independently of N is not possible- will be confirmed by considering the formal moment models which will be shown to be non consistent. Consequently this first family of results is probably of minor practical interest for numerical approximation even if endowed with rigorous approximation results.

The second family of construction intends to be less sensitive to N, so that it is possible to pass to the limit $\varepsilon \to 0$ keeping N fixed. This will be performed with an original polynomial approximation of M, that we call a *kinetic polynomial* M^N. It is described thanks to an algorithm based on sharp polynomial properties. The structure of M^N is layered in ξ,

$$\left\{ \begin{array}{l} M^N(u^N;\xi,\omega) = \displaystyle\sum_{l=0}^{L} h_l^N(\omega) \mathbb{1}_{\{\xi_l<\xi<\xi_{l+1}\}}, \quad 0=\xi_0<\xi_1<\cdots<\xi_L<\xi_{L+1}=u_+=\max_I u^N(\omega), \\ h_l^N \in P_\omega^N, \end{array} \right. \tag{10}$$

and is a natural option in the context of kinetic formulations. It defines an original approximation method which has the following form in dimension one for scalar conservation law

$$\partial_t u^N + \partial_x F^N[u^N] = 0, \quad F^N[u^N] = \int a(\xi) M^N(u^N;\xi,\omega) d\xi = \sum_{l\geq 0} h_l^N(\omega) \left(F(\xi_{l+1}) - F(\xi_l)\right). \tag{11}$$

An interesting asset of this *kinetic polynomial method* is the exact formula for $F^N[u^N]$, in this sense no quadrature rule is needed in opposition to the method of moments which requires to evaluate the flux $\int F(u^N) p_i(\omega) d\mu(\omega)$ in (4). Even if the theory of approximation is less advanced than for the convolution case, it seems to be much more adapted for the numerical approximation since this new method (11) is endowed with a natural maximum principle which is not the case for the general moment approximation (4). This property can be considered as one the main results of this work, see Theorem 4.1. For the Burgers flux $f(u) = \frac{u^2}{2}$ simple numerical results which satisfy the maximum principle illustrate the potential interest of this second family of methods.

The organization of this work is as follows. Section 2 is devoted to the convolution approach and the proof of Theorem 1.1. Section 2.3 deals with the design of the layered kinetic polynomial Maxwellian (10) and the proof of Theorem 3.10 about the completion of the algorithm that constructs the layered Maxwellian. Theorem 4.1 about the stability in the maximum norm of the scheme (11) is proved in the last section 4. We complete this work with simple numerical illustrations of the new method (10) for the case $N = 2$.

2 Convolution kernels

A simple idea to construct a polynomial approximation M^N of M, is to use a convolution method under the form

$$M^N(u_\varepsilon^N;\xi,\omega) = G^N *_\omega M(u_\varepsilon^N;\xi) := \int G^N(\omega,\omega')M\big(u_\varepsilon^N(\omega');\xi\big)d\mu(\omega')$$

where the convolution kernel G^N can be decomposed, using the orthonormal basis defined by (3), as

$$G^N(\omega,\omega') = \sum_{i=0}^{N} c_i p_i(\omega)p_i(\omega'), \tag{12}$$

where c_i are appropriate coefficients, and where G^N satisfies

$$G^N \geq 0, \qquad \int G^N(\omega,\omega')d\mu(\omega') = c_0 = 1 = \int G^N(\omega,\omega')d\mu(\omega). \tag{13}$$

The theory of polynomial kernel approximation [9, 21] asserts that convolution kernels exist which satisfy the requirements (12)–(13).

For example, considering the measure $d\mu(\omega) = \frac{d\omega}{\pi\sqrt{1-\omega^2}}$, on the interval $\omega \in I = (-1,1)$, they are built on the Tchebycheff orthonormal polynomials

$$T_i(\omega) = \cos\left(i \arccos \omega\right), \qquad -1 \leq \omega \leq 1.$$

According to formula (3), a function v can be represented as the infinite series

$$v(\omega) = \mu_0 + 2\sum_{i=1}^{\infty} \mu_i T_i(\omega), \qquad \mu_i = \int_I v(\omega')T_i(\omega')d\mu(\omega').$$

The following kernels are suitable modifications of the truncated Dirichlet series

$$v_D^N(\omega) = \mu_0 + 2\sum_{i=1}^{N} \mu_i T_i(\omega) = G *_\omega v$$

for the choice $c_i \equiv 1$, which generates oscillations [21] and thus, is not convenient for our purposes since (13) is not satisfied.

Exemple 2.1 (Fejer Kernel) *The Fejer kernel G_F^N is defined by the coefficients*

$$c_0 = 1 \text{ and } c_i = 2\frac{N+1-i}{N+1}, \quad 1 \leq i \leq N.$$

The truncated Fejer series $v_F^N(\omega) = c_0\mu_0 + 2\sum_{i=1}^{N} c_i\mu_i T_i(\omega)$ is such that $v \geq 0 \Longrightarrow v_N^F \geq 0$. It comes from the integral representation [9]

$$v_F^N(\omega) = \int_0^{2\pi} v(\cos(t-u))K_F^N(u)du = \int_0^{2\pi} v(\cos u)K_F^N(t-u)du, \quad \omega = \cos t, \tag{14}$$

with the kernel $K_F^N(u) = \frac{1}{2\pi(N+1)}\left(\frac{\sin(N+1)\frac{u}{2}}{\sin\frac{u}{2}}\right)^2$. This representation formula shows the property (13). Assuming $v \in L^\infty(I)$, which is relevant in our context, one deduces from [9][Corollary 2.5 page 7] the following approximation property: $\lim_{N\to\infty} v_F^N(\omega) = v(\omega)$ almost everywhere.

Exemple 2.2 (Jackson Kernel) *The Jackson integral representation writes*

$$v_J^N(\omega) = \int_0^{2\pi} v(\cos(t-u))K_J^N(u)du = \int_0^{2\pi} v(\cos u)K_J^N(t-u)du, \quad \omega = \cos t. \tag{15}$$

The Jackson kernel $K_J^N(u) = \lambda_J^N \left(\frac{\sin(N+1)\frac{u}{2}}{\sin\frac{u}{2}}\right)^4$ *is a convenient renormalization of the square of the Fejer kernel [9], where the normalization coefficient* λ_J^N *is defined by* $\int_0^{2\pi} K_J^N(u)du = 1$*. By definition* v_F^N *is a polynomial of degree* $\leq 2N$ *in the variable* ω*. One has better approximation properties for the Jackson kernel than for the Fejer kernel, cf. [9][Theorem 2.2 page 205], since one can prove an optimal error estimate*

$$\left\|v - v_J^N\right\|_{L_\mu^1(I)} \leq C\mathrm{mod}_2(v, \frac{1}{N}) \tag{16}$$

with

$$\mathrm{mod}_2(v,\alpha) := \int_0^\pi |v(\cos(t+\alpha)) - 2v(\cos(t)) + v(\cos(t-\alpha))|\, dt. \tag{17}$$

Notice that $\mathrm{mod}_2(f,\alpha)$ *is an integral with respect to the trigonometric variable* $t \in (0,\pi)$*, while the error is an integral with respect to the original variable* $\omega = \cos t$*. This is the reason of the weighted* L_μ^1 *norm used for* (16)*. One deduces that it also holds:*

$$\left\|v - v_J^N\right\|_{L_\mu^1(I)} \leq C\mathrm{mod}_1(v, \frac{1}{N}) \tag{18}$$

with

$$\mathrm{mod}_1(v,\alpha) = \int_0^{2\pi} |v(\cos(t+\alpha) - v(\cos t)|\, dt \quad \textit{for} \;\; \alpha > 0. \tag{19}$$

Exemple 2.3 (The modified Jackson Kernel) *This kernel can be found in [17] and is studied for modern physical calculations in [21]. The polynomial is defined by*

$$v^N = c_0\mu_0 + 2\sum_{i=1}^N c_i\mu_i T_i(\omega), \quad c_i = \frac{(N+2-i)\cos\frac{\pi i}{N+2} + \sin\frac{\pi i}{N+2}\cot\frac{\pi}{N+2}}{N+2}. \tag{20}$$

This is also a positive kernel [21] and it admits a convolution representation like (15). The approximation properties in L_μ^1 *are similar to those of the Jackson kernel due to the estimates (61)-(69) in [21]. For practical computations, this kernel has two major interests. The coefficients* c_n *are known by a simple analytical formula, and its degree is arbitrary, unlike the Jackson kernel which has even order. The corresponding kernel is denoted as* $K_{mod,J}^N$*.*

2.1 Entropy and a priori bounds

We review some a priori bounds and properties satisfied by the equation (9) recalling that we use well-prepared initial data

$$M^N(u^{\mathrm{init}};\xi,\omega) = G^N *_\omega M(u^{\mathrm{init}};\xi,\omega). \tag{21}$$

The stability estimates are the same for any of the three kernels mentioned above. However, we detail the approximation estimates only for the Jackson kernel, for which the theory of approximation is well established. The generalization to the modified Jackson kernel is left to the reader together with the study of the non optimal approximation properties of the Fejer kernel.

Proposition 2.4 *For any of the three kernels, there is a unique solution of (9) and we have the properties*

1. $$0 \le f_\varepsilon^N \le 1.$$

2. $$0 \le u_\varepsilon^N \le U_M := \sup_{x,\omega} u^{\rm init}(x,\omega), \qquad f_\varepsilon^N(x,\xi,\omega,t) \equiv 0 \text{ for } \xi \ge U_M.$$

3. For all smooth convex functions $S(\cdot)$, we have

$$\partial_t \int S'(\xi) f_\varepsilon^N(x,\xi,\omega,t) d\xi d\mu(\omega) + \mathrm{div} \int a(\xi) S'(\xi) f_\varepsilon^N(x,\xi,\omega,t) d\xi d\mu(\omega) \le 0.$$

4. (Contraction principle) Consider two solutions f_ε^N and g_ε^N of (9), then

$$\int \left| f_\varepsilon^N(t) - g_\varepsilon^N(t) \right| dx d\xi d\mu(\omega) \le \int \left| f_\varepsilon^N(0) - g_\varepsilon^N(0) \right| dx d\xi d\mu(\omega).$$

5. (Comparison principle) Also, if $f_\varepsilon^N(0) \le g_\varepsilon^N(0)$,then $f_\varepsilon^N(t) \le g_\varepsilon^N(t)$ for all $t \ge 0$.

6. Finally, the BV bounds in space are propagated, for all $t \ge 0$,

$$\int |\nabla_x f_\varepsilon^N| dx d\xi d\mu(\omega) \le C^{\rm init}, \qquad \int |\nabla_x u_\varepsilon^N| dx d\mu(\omega) \le C^{\rm init}. \tag{22}$$

Proof. 1. Indeed, from (13) we have $G^N *_\omega M(u_\varepsilon^N;\xi) \ge 0$ and thus $f_\varepsilon^N \ge 0$.

2. Also, we have $G^N *_\omega M(u_\varepsilon^N;\xi) \le \int G^N(\omega,\omega') d\mu(\omega') = 1$ and thus $f_\varepsilon^N \le 1$.

3. Similarly, for $\xi \ge U_M$, we have $M(u^{\rm init};\xi) \equiv 0$ and thus $f_\varepsilon^N(t=0) = 0$. This property is propagated by the equation because $f(\xi) = 0$ for $\xi \ge U_M$ implies $u = \int f d\xi \le U_M$, which itself implies $M = 0$ for $\xi \ge U_M$ and thus $G *_\omega M = 0$ for $\xi \ge U_M$.

4. To prove the entropy inequality, we can always assume $S(0) = 0$. Then, after multiplying the equation by $S'(\xi)$, integrating in ξ and ω, we notice that the right hand side is $\frac{1}{\varepsilon}$ times

$$\begin{aligned}
&\int S'(\xi) G^N *_\omega M(u_\varepsilon^N;\xi) d\xi d\mu((\omega) - \int S'(\xi) f_\varepsilon^N(x,\xi,\omega,t) d\xi d\mu(\omega) \\
&\quad = \int G^N *_\omega S(u_\varepsilon^N) d\mu(\omega) - \int S'(\xi) f_\varepsilon^N(x,\xi,\omega,t) d\xi d\mu(\omega) \\
&\quad = \int S(u_\varepsilon^N) d\mu(\omega) - \int S'(\xi) f_\varepsilon^N(x,\xi,\omega,t) d\xi d\mu(\omega) \le 0
\end{aligned}$$

because this is true ω by ω, thanks to the universal entropy minimisation principle (8). Here we have also used $\int G^N(\omega,\omega') d\mu(\omega) = 1$.

5. One has

$$\begin{aligned}
\partial_t |f_\varepsilon^N - g_\varepsilon^N| + a(\xi).\nabla_x |f_\varepsilon^N - g_\varepsilon^N| + \tfrac{1}{\varepsilon} |f_\varepsilon^N - g_\varepsilon^N| \quad &= \mathrm{sgn}(f_\varepsilon^N - g_\varepsilon^N)\, G *_\omega \left(M(u_\varepsilon^N;\xi) - M(v_\varepsilon^N;\xi) \right) \\
&\le G *_\omega \left| M(u_\varepsilon^N;\xi) - M(v_\varepsilon^N;\xi) \right|.
\end{aligned}$$

So

$$\frac{d}{dt} \int |f_\varepsilon^N - g_\varepsilon^N| dx d\xi d\mu(\omega) + \frac{1}{\varepsilon} \int |f_\varepsilon^N - g_\varepsilon^N| dx d\xi d\mu(\omega) \le \frac{1}{\varepsilon} \int G *_\omega \left| M(u_\varepsilon^N;\xi) - M(v_\varepsilon^N;\xi) \right| dx d\xi d\mu(\omega)$$

$$\leq \frac{1}{\varepsilon}\int \left|M(u_\varepsilon^N;\xi)-M(v_\varepsilon^N;\xi)\right| dxd\xi d\mu(\omega) = \frac{1}{\varepsilon}\int \left|u_\varepsilon^N - v_\varepsilon^N\right| dxd\mu(\omega) = \frac{1}{\varepsilon}\int |f_\varepsilon^N - g_\varepsilon^N| dxd\xi d\mu(\omega).$$

6. The comparison principle is a simple variant using $(...)_+$ in place of the absolute value.
7. The BV bound, is an immediate consequence of the contraction principle combined with translational invariance.

The proof is complete. □

The entropy inequality of Proposition 2.4 is integrated over ω. It is possible to get a sharper estimate by using a test function $\varphi \geq 0$ for the ω variable. The price is to work with weak topology.

Proposition 2.5 (Estimates in weak norms in ω) *Consider solutions of (9) with the Jackson kernel. For all smooth convex functions $\xi \mapsto S(\xi)$ and all non negative function $\omega \mapsto \varphi(\omega)$, there exists a constant C_{S,U_M} such that*

$$\partial_t \int \varphi(\omega) S'(\xi) f_\varepsilon^N(x,\xi,\omega,t) d\xi d\mu(\omega) + \operatorname{div} \int \varphi(\omega) a(\xi) S'(\xi) f_\varepsilon^N(x,\xi,\omega,t) d\xi d\mu(\omega) \leq C_{S,U_M} \frac{\operatorname{mod}_2(\varphi, \frac{1}{N})}{\varepsilon}.$$

Proof. We only have to evaluate the right hand side, that is $1/\varepsilon$ times A with

$$A = \int \varphi(\omega) S'(\xi) G^N *_\omega M(u_\varepsilon^N;\xi) d\xi d\mu(\omega) - \int \varphi(\omega) S'(\xi) f_\varepsilon^N(x,\xi,\omega,t) d\xi d\mu(\omega).$$

The kernel being symmetric one has (normalizing S with $S(0)=0$)

$$\int \varphi(\omega) S'(\xi) G^N *_\omega M(u_\varepsilon^N;\xi) d\xi d\mu(\omega) = \int S'(\xi) M(u_\varepsilon^N;\xi) \left(G^N *_\omega \varphi(\omega)\right) d\xi d\mu(\omega)$$

$$\leq \int S'(\xi) M(u_\varepsilon^N;\xi) \varphi(\omega) d\xi d\mu(\omega) + \int S'(\xi) M(u_\varepsilon^N;\xi) r^N(\omega) d\xi d\mu(\omega)$$

where the convolution error $r^N = G^N *_\omega \varphi - \varphi$ is controled as, recalling (16)–(17),

$$\left\|r^N(\omega)\right\|_{L^1_\mu} \leq \operatorname{mod}_2(\varphi, \frac{1}{N}).$$

Therefore, using the L^∞ bounds of Proposition 2.4, the last term is bounded as

$$\left|\int S'(\xi) M(u_\varepsilon^N;\xi) r^N(\omega) d\xi d\mu(\omega)\right| \leq \sup_\omega |S(u_\varepsilon^N)| \ \left\|r^N(\omega)\right\|_{L^1_\mu} \leq \max_{0\leq\xi\leq U_M} \|S(\xi)\| \operatorname{mod}_2(\varphi, \frac{1}{N}).$$

Finally, using the universal entropy minimisation principle (8), $A \leq C_{S,U_M} \operatorname{mod}_2(\varphi, \frac{1}{N})$ which ends the proof. □

Next question is to obtain a BV bound with respect to ω. In view of the comparison principle of Proposition 2.4, it is appealing to rely on transformations which commute with the operators of the system. Fortunately, the kernels based on Tchebycheff polynomials are endowed with such a natural transformation. This is immediately visible in (15) which is written as convolutions. This is also true for the modified Jackson kernel (20) which can also be written as a convolution [21].

For any f and $0<\alpha$, we define the linear operator (which mimicks translations)

$$g=Q_\alpha f \iff g(\omega)=\frac{f(\cos(t+\alpha))-f(\cos(t))}{\alpha},\quad \omega=\cos t\in[-1,1]. \tag{23}$$

The aforementioned convolution formulas have the consequence that the operators commute

$$Q_\alpha G^N *_\omega = G^N *_\omega Q_\alpha. \tag{24}$$

Proposition 2.6 (BV bound in ω) *Consider solutions of (9) with any of the three kernels. As a consequence of the commutation property (24), one has that*

$$\int \left|Q_\alpha f_\varepsilon^N(t)\right| dxd\xi d\mu(\omega) \le \int \left|Q_\alpha f_\varepsilon^N(0)\right| dxd\xi d\mu(\omega). \tag{25}$$

Proof. The proof is a slight modification of the comparison principle in Proposition 2.4. We set $g^N_{\varepsilon\alpha}=Q_\alpha f^N_\varepsilon$ which satisfies

$$\partial_t g^N_{\varepsilon\alpha}+a(\xi).\nabla g^N_{\varepsilon\alpha}+\frac{1}{\varepsilon}g^N_{\varepsilon\alpha}=\frac{1}{\varepsilon}Q_\alpha G^N *_\omega M(u^N_\varepsilon;\xi)=\frac{1}{\varepsilon}G^N *_\omega Q_\alpha M(u^N_\varepsilon;\xi).$$

Therefore

$$\begin{aligned}\partial_t|g^N_{\varepsilon\alpha}|+a(\xi).\nabla_x|g^N_{\varepsilon\alpha}|+\tfrac{1}{\varepsilon}|g^N_{\varepsilon\alpha}| &= \operatorname{sgn}(g^N_{\varepsilon\alpha})\, G *_\omega Q_\alpha M(u^N_\varepsilon;\xi)\\ &\le G *_\omega \left|Q_\alpha M(u^N_\varepsilon;\xi)\right|\end{aligned}$$

and

$$\begin{aligned}\tfrac{d}{dt}\textstyle\int|g^N_{\varepsilon\alpha}| \quad dxd\xi d\mu(\omega)+\tfrac{1}{\varepsilon}\int|g^N_{\varepsilon\alpha}|dxd\xi d\mu(\omega) &\le \tfrac{1}{\varepsilon}\int G *_\omega \left|Q_\alpha M(u^N_\varepsilon;\xi)\right| dxd\xi d\mu(\omega)\\ &\le \tfrac{1}{\varepsilon}\int \left|Q_\alpha M(u^N_\varepsilon;\xi)\right| dxd\xi d\mu(\omega) \le \tfrac{1}{\varepsilon}\int \left|Q_\alpha u^N_\varepsilon\right| dxd\mu(\omega)\\ &\le \tfrac{1}{\varepsilon}\int \left|Q_\alpha f^N_\varepsilon\right| dxd\xi d\mu(\omega) = \tfrac{1}{\varepsilon}\int \left|g^N_{\varepsilon\alpha}\right| dxd\xi d\mu(\omega).\end{aligned}$$

So that we conclude $\frac{d}{dt}\int|g^N_{\varepsilon\alpha}|dxd\xi d\mu(\omega)\le 0$, and thus $\int|g^N_{\varepsilon\alpha}(t)|dxd\xi d\mu(\omega)\le\int|g^N_{\varepsilon\alpha}(0)|dxd\xi d\mu(\omega)$. which is the announced result. □

Proposition 2.7 (BV bound in time) *Consider solutions of (9) with any of the three kernels. The time derivative is bounded as follows:*

$$\int \left|\partial_t f^N_\varepsilon(t)\right| dxd\xi d\mu(\omega) \le \int \left|\partial_t f^N_\varepsilon(0)\right| dxd\xi d\mu(\omega) \le C_{U_M}\int|\nabla_x u^{\text{init}}|dxd\mu(\omega).$$

Proof. This is once again a consequence of the comparison principle. Set $g^N_\varepsilon(t)=f^N_\varepsilon(t+\alpha)$, with $\alpha>0$. The comparison principle of Proposition 2.4 yields that

$$\int\left|\frac{f^N_\varepsilon(t+\alpha)-f^N_\varepsilon(t)}{\alpha}\right| dxd\xi d\mu(\omega)\le\int\left|\frac{f^N_\varepsilon(\alpha)-f^N_\varepsilon(0)}{\alpha}\right| dxd\xi d\mu(\omega).$$

We can pass to the limit $\alpha=0^+$ using that $\partial_t f^N_\varepsilon(0)=-a(\xi)\cdot\nabla_x f^N_\varepsilon(0)$ for the well-prepared data under considerations. Because we have

$$\int|\nabla_x M^N\big(u^{\text{init}};\xi,\omega\big)|dxd\xi d\mu(\omega)\le\int|\nabla_x M\big(u^{\text{init}};\xi\big)|dxd\xi d\mu(\omega)\le\int|\nabla_x u^{\text{init}}|dxd\mu(\omega),$$

the proof is complete. □

For our next statement, we use truncation functions in x with the properties (with R a parameter allowing localization on sets as large as we wish)

$$\chi \in \mathcal{C}^2(\mathbb{R}^d), \quad \chi \in L^1(\mathbb{R}^d), \quad \chi > 0, \quad \chi = 1 \text{ in } B_R, \quad |\nabla \chi| \leq \chi. \tag{26}$$

In other words we choose truncation functions which decay as $e^{-|x|}$ at infinity.

Proposition 2.8 (Estimate of derivative in ξ) *Consider any of the three kernels and assume $u^{\text{init}} \in BV_x$ and consider a nonnegative truncation function satisfying* (26)*. Then, solutions of (9) have bounded ξ derivatives*

$$\int \chi(x) \left|\partial_\xi f_\varepsilon^N(t)\right| dxd\xi d\mu(\omega) \leq e^{-\frac{t}{\varepsilon}} \int \chi(x) \left|\partial_\xi f_\varepsilon^N(0)\right| dxd\xi d\mu(\omega) + C_R$$

where C_R depends on the radius in (26) *and on the $x - BV$ estimates (22).*

Proof. Indeed one can differentiate the equation with respect to ξ

$$\partial_t \partial_\xi f_\varepsilon^N + a(\xi).\nabla \partial_\xi f_\varepsilon^N + \frac{1}{\varepsilon} \partial_\xi f_\varepsilon^N = \frac{1}{\varepsilon} G^N *_\omega \partial_\xi M(u_\varepsilon^N; \xi) - a'(\xi).\nabla f_\varepsilon^N$$

where $\partial_\xi M(u_\varepsilon^N; \xi) = \delta(\xi - u_\varepsilon^N)$ is a measure. It yields

$$\partial_t \left|\partial_\xi f_\varepsilon^N\right| + a(\xi).\nabla \left|\partial_\xi f_\varepsilon^N\right| + \frac{1}{\varepsilon} \left|\partial_\xi f_\varepsilon^N\right| = \frac{1}{\varepsilon} \left|G^N *_\omega \partial_\xi M(u_\varepsilon^N; \xi)\right| + \left|a'(\xi).\nabla f_\varepsilon^N\right|,$$

Here the right hand side can be bounded. Firstly one has that

$$\int \left|G^N *_\omega \partial_\xi M(u_\varepsilon^N; \xi)\right| d\xi d\mu(\omega) \leq \int \left|\partial_\xi M(u_\varepsilon^N; \xi)\right| d\xi d\mu(\omega) = \int d\mu(\omega) = C_1,$$

Consequently, we find

$$\begin{aligned} \partial_t \chi(x) \textstyle\int \left|\partial_\xi f_\varepsilon^N\right| d\xi d\mu(\omega) \quad & + \textstyle\int a(\xi).\nabla [\chi(x) \left|\partial_\xi f_\varepsilon^N\right|] \, d\xi d\mu(\omega) + \textstyle\int \chi(x) \frac{1}{\varepsilon} \left|\partial_\xi f_\varepsilon^N\right| d\xi d\mu(\omega) \\ & \leq \chi(x) \tfrac{C_1}{\varepsilon} + \textstyle\int \chi(x) \left|a'(\xi).\nabla f_\varepsilon^N\right| d\xi d\mu(\omega) + \textstyle\int \left|\partial_\xi f_\varepsilon^N\right| a(\xi).\nabla \chi(x) d\xi d\mu(\omega) \end{aligned}$$

Secondly one has

$$\int \left|a'(\xi).\nabla f_\varepsilon^N\right| dxd\xi d\mu(\omega) \leq C_2 \int \left|\nabla f_\varepsilon^N\right| dxd\xi d\mu(\omega) \leq C_3$$

where we used Proposition 2.4, and $|a'| \leq C_3$ because $\xi \leq U_M$ in these integrals. Therefore, using the truncation properties of (26),

$$\frac{d}{dt} \int \chi(x) \left|\partial_\xi f_\varepsilon^N\right| dxd\xi d\mu(\omega) + \frac{1}{\varepsilon} \int \chi(x) \left|\partial_\xi f_\varepsilon^N\right| dxd\xi d\mu(\omega) \leq \frac{C_1}{\varepsilon} + C_3 \int \chi dx + C_4 \int \chi(x) \left|\partial_\xi f_\varepsilon^N\right| dxd\xi d\mu(\omega).$$

This yields the result using Gronwall's lemma. □

2.2 Convergence

We establish some conditions for the convergence of f_ε^N towards a correct limit f as $N \to \infty$ and $\varepsilon \to 0$. To begin with, we notice that the solutions are BV in the domain

$$(x, \xi, \omega, t) \in \mathcal{D}_\alpha = \mathbb{R}^d \times [0, U_M] \times [1-\alpha, 1+\alpha] \times [0, T], \quad 0 < \alpha < 1,$$

due to the previous propositions.

The restriction in direction ω comes from the transformation (23), since we recognize the standard BV criterion for the variable t but mapping it back to the variable ω, we loose the BV bound because of the degeneracy at the end points due to the vanishing derivative of the cosine function. This is why the BV property (in particular for the measure with respect to ω which writes $\frac{d\omega}{\pi\sqrt{1-\omega^2}}$) is obtained readily only in D_α for $\alpha > 0$. Nevertheless the functions f_ε^N also belong to L^∞. Therefore,
1. as $N \to \infty$, $\varepsilon \to 0$, from any sequence $f_\varepsilon^N \in L^\infty \cap \mathcal{D}_\alpha$ for all $0 < \alpha$, one can extract a subsequence that converges strongly in $L^{1,loc}_{x\xi\mu(\omega)}(\mathcal{D}_0)$ to a limit f which satisfies $0 \le f \le 1$ and $\int f dx d\xi d\mu(\omega) \le \int u^{\text{init}}(x,\omega) dx d\mu(\omega)$.
2. for the same subsequence, $u_\varepsilon^N = \int f_\varepsilon^N d\xi$ converges strongly in $L^{1,loc}_{x\mu(\omega)}(\mathcal{D}_0)$ to the limit $u = \int f d\xi$,
3. and $M(u_\varepsilon^N; \xi)$ converges strongly in $L^{1,loc}_{x\xi\mu(\omega)}(\mathcal{D}_0)$ to the limit $M(u;\xi)$,
4. For the Jackson kernel, $M^N(u_\varepsilon^N; \xi, \omega)$ converges strongly in $L^{1,loc}_{x\xi\mu(\omega)}(\mathcal{D}_0)$ to $M(u;\xi)$. This is because

$$M^N(u_\varepsilon^N;\xi,\omega) - M(u_\varepsilon^N;\xi) = (G^N - I) *_\omega [M(u_\varepsilon^N;\xi) - M(u;\xi)] + (G^N - I) *_\omega M(u;\xi).$$

The first term converges to 0 thanks to (13) and 3., while the second term converges to 0 because

$$\int |(G^N - I) *_\omega M(u;\xi)| dx d\xi d\mu(\omega) \le C \int \text{mod}_1\left(M(u;\xi), \frac{1}{N}\right) \le C \int \text{mod}_1\left(M(u^{\text{init}};\xi), \frac{1}{N}\right)$$

where we use the standard L^1_x contraction inequality for solutions of the scalar conservation law (1). A BV bound of the initial data with respect to the variable ω shows the convergence to zero of this term (with the same argument as before for the degeneracy at the endpoints),
5. and, still for the Jackson kernel, $f = M(u;\xi,\omega)$ because the equation (9) gives

$$M^N(u_\varepsilon^N;\xi,\omega) - f_\varepsilon^N = \varepsilon[\partial_t f_\varepsilon^N + a(\xi).\nabla f_\varepsilon^N] \to 0 \qquad \text{in } \mathcal{D}'.$$

Consequently, the questions of interest are to establish some conditions which imply that $f = M(u;\xi)$ and u satisfy the correct limit equations and to obtain error estimates.

Proposition 2.9 (The limit is a weak solution) *Consider the Jackson kernel. Assume that $N^2\varepsilon \to \infty$. Then, the full sequence f_ε^N (see the construction before) converges to $f = M(u;\xi)$ and it is a weak solution of*

$$\partial_t f + a(\xi).\nabla f = \partial_\xi m, \qquad f(t=0) = M(u^{\text{init}};\xi), \tag{27}$$

where m is a non negative measure, and thus $u = \int f d\xi$ is the unique entropy solution of (1), that is for a.e. ω.

Proof. The theory in [19, 16, 18] immediately shows that u is the entropy solution of the conservation law, a.e. with respect to ω, and thus uniqueness as soon as (27) is established (we recall the $L^1 \cap L^\infty$ assumption for u^{init} following (2)).

To prove that (27), holds, we write

$$\partial_t f_\varepsilon^N + a(\xi).\nabla f_\varepsilon^N = \frac{1}{\varepsilon}\left(M(u_\varepsilon^N;\xi) - f_\varepsilon^N\right) + \frac{1}{\varepsilon}\left(G^N *_\omega - I_\omega\right) M(u_\varepsilon^N;\xi). \tag{28}$$

The first three terms pass to the limit from the observations above. It remains to prove that $\frac{1}{\varepsilon}\left(G^N *_\omega - I_\omega\right) M(u_\varepsilon^N;\xi)$ tends to zero in the weak sense. For this purpose, it is sufficient to use a smooth test function $\varphi(x,\xi,\omega,t)$ with compact support $\mathcal{C}$, and observe that

$$\begin{aligned}\tfrac{1}{\varepsilon}\int_{\mathcal{C}} \varphi(x,\xi,\omega,t) \quad & \left((G^N *_\omega - I_\omega)M(u_\varepsilon^N;\xi)\right) dxd\xi d\mu(\omega)dt \\ &= \tfrac{1}{\varepsilon}\int_{\mathcal{C}} \left((G^N *_\omega - I_\omega)\varphi\right)\ M(u_\varepsilon^N;\xi)dxd\xi d\mu(\omega)dt.\end{aligned}$$

To conclude the proof, we use the same argument as in Proposition 2.5, and notice that for the test function we can assume it has compact support in $\omega \in (-1,1)$

$$\|(G^N *_\omega - I_\omega)\varphi\|_{L^1(d\mu)} \leq C\mathrm{mod}_2(\varphi, \frac{1}{N}) \leq \frac{C}{N^2}$$

and the equation is established. □

Our next result quantifies strong convergence, based on the previous comparison estimates. The compatibility relation between ε and N is nevertheless more stringent than for weak convergence since we need $N\varepsilon \to 0$. This is the lowest rate for N that we can use to handle the right hand side of (29).

Proposition 2.10 (Strong error bounds) *Consider the Jackson kernel. One has the inequalities*

$$\int \left|f_\varepsilon^N(t) - G^N *_\omega f_\varepsilon(t)\right| dxd\xi d\mu(\omega) \leq C\frac{t}{\varepsilon}\int \mathrm{mod}_1(u^{\mathrm{init}}, \frac{1}{N})dxd\xi, \tag{29}$$

$$\int \left|f_\varepsilon^N(t) - f_\varepsilon(t)\right| dxd\xi d\mu(\omega) \leq C(1+\frac{t}{\varepsilon})\int \mathrm{mod}_1(u^{\mathrm{init}}, \frac{1}{N})dxd\xi, \tag{30}$$

$$\int \left|f_\varepsilon^N(t) - M(u;\xi)\right| dxd\xi d\mu(\omega) \leq c\sqrt{\varepsilon} + C(1+\frac{t}{\varepsilon})\int \mathrm{mod}_1(u^{\mathrm{init}}, \frac{1}{N})dxd\xi. \tag{31}$$

Proof. 1. The convolution of (6) yields the identity

$$\partial_t g_\varepsilon^N + a(\xi).\nabla g_\varepsilon^N + \frac{1}{\varepsilon}g_\varepsilon^N = \frac{1}{\varepsilon}G^N *_\omega M(u_\varepsilon;\xi), \qquad g_\varepsilon^N = G^N *_\omega f_\varepsilon.$$

Set $v_\varepsilon^N = \int g_\varepsilon^N d\xi = G^N *_\omega u_\varepsilon$ and $r_\varepsilon^N = G^N *_\omega \left(M(u_\varepsilon;\xi) - M(v_\varepsilon^N;\xi)\right)$ so that

$$\partial_t g_\varepsilon^N + a(\xi).\nabla g_\varepsilon^N + \frac{1}{\varepsilon}g_\varepsilon^N = \frac{1}{\varepsilon}G^N *_\omega M(v_\varepsilon^N;\xi) + \frac{1}{\varepsilon}r_\varepsilon^N.$$

By definition, see (21), the initial conditions are the same for f_ε^N and g_ε^N:

$$f_\varepsilon^N(0) = g_\varepsilon^N(0) = G^N *_\omega M(u^{\mathrm{init}};\xi) = M^N(u^{\mathrm{init}};\xi,\omega).$$

Since by definition $\partial_t f_\varepsilon^N + a(\xi).\nabla f_\varepsilon^N + \frac{1}{\varepsilon} f_\varepsilon^N = \frac{1}{\varepsilon} G^N *_\omega M(u_\varepsilon^N;\xi)$, one more use of the comparison principle between f_ε^N and g_ε^N yields

$$\int \left| f_\varepsilon^N(t) - g_\varepsilon^N(t) \right| dxd\xi d\mu(\omega) \leq \frac{1}{\varepsilon} \int_0^t \left(\int \left| r_\varepsilon^N(s) \right| dxd\xi d\mu(\omega) \right) ds.$$

Using the definition of r_ε^N and the property (13), one has

$$\begin{aligned}
\int \left| r_\varepsilon^N(x,\xi,\omega,s) \right| dxd\xi d\mu(\omega) &\leq \int \left| M(u_\varepsilon;\xi) - M(v_\varepsilon^N;\xi) \right| dxd\xi d\mu(\omega) \\
&\leq \int \left| u_\varepsilon(x,\omega,s) - v_\varepsilon^N(x,\omega,s) \right| dxd\xi d\mu(\omega) \\
&\leq \int \left| u_\varepsilon(x,\omega,s) - G^N *_\omega u_\varepsilon(x,\omega,s) \right| dxd\mu(\omega) \\
&\leq C \int \mathrm{mod}_1(u_\varepsilon(s), \tfrac{1}{N}) dx.
\end{aligned}$$

The last inequality comes from the approximation property (16) in the L^1_μ norm of the Jackson kernel. It remains to use the standard $L^1_{x\xi}$ contraction estimate for solutions of the kinetic equation (6). One gets

$$\begin{aligned}
\int \mathrm{mod}_1(u_\varepsilon(s), \tfrac{1}{N}) dx &= \int_0^{2\pi} \left(\int \left| u_\varepsilon\left(x, \cos(\tau + \tfrac{1}{N}), s\right) - u_\varepsilon(x, \cos\tau, s) \right| dx \right) d\tau \\
&\leq \int_0^{2\pi} \left(\int \left| u_\varepsilon\left(x, \cos(\tau + \tfrac{1}{N}), 0\right) - u_\varepsilon(x, \cos\tau, 0) \right| dx \right) d\tau \\
&= \int \mathrm{mod}_1(u_\varepsilon(0), \tfrac{1}{N}) dx.
\end{aligned}$$

Since the initial data is independent of ε, that is $u_\varepsilon(0) = u^{\mathrm{init}}$, it ends the proof of the first estimate.

2. We have, using the contraction principle for the BGK equation, cf [19, 16, 18],

$$\begin{aligned}
\int \left| G^N *_\omega f_\varepsilon(t) - f_\varepsilon(t) \right| dxd\xi d\mu(\omega) &\leq \int \mathrm{mod}_1(f_\varepsilon(t), \tfrac{1}{N}) dxd\xi \\
&\leq \int \int_0^{2\pi} |f\varepsilon(x,\xi,\omega + \tfrac{1}{N}, t) - f\varepsilon(x,\xi,\omega,t)| dxd\xi d\omega \\
&\leq \int \int_0^{2\pi} |f\varepsilon(x,\xi,\omega + \tfrac{1}{N}, 0) - f\varepsilon(x,\xi,\omega,0)| dxd\xi d\omega \\
&= \int \mathrm{mod}_1(u^{\mathrm{init}}, \tfrac{1}{N}) dx,
\end{aligned}$$

which proves the second inequality

3. The third estimate is just the consequence of the second and of the standard convergence rate from the BGK model to the scalar conservation law, cf [19, 16, 18]. □

Remark 2.11 *The estimate (29) can be slightly enhanced under the form*

$$\int \left| f_\varepsilon(t) - G^N *_\omega f_\varepsilon(t) \right| dxd\xi d\mu(\omega) \leq \frac{C}{\varepsilon} \int \mathrm{mod}_2(f_\varepsilon(t), \frac{1}{N}) dxd\xi.$$

However, propagation bounds hold true for $\mathrm{mod}_1(f_\varepsilon, \frac{1}{N})$*, but not for* $\mathrm{mod}_2(f_\varepsilon, \frac{1}{N})$.

2.3 Moment equations

Even if the weak or strong convergence estimates do not allow to pass to the limit ε independently of N, it is instructive in view of practical numerical computations to write the formal limit in the regime $\varepsilon N = O(1)$. The unknown of the resulting moment system are the quantities

$$u^N_{\varepsilon,i}(x,t) = \int f^N_{\varepsilon,i}(x,\omega,t)d\xi, \quad f^N_{\varepsilon,i}(x,\omega,t) = \int f^N_\varepsilon(x,\xi,\omega,t)T_i(\omega)d\mu(\omega).$$

We now explain why an artificial damping phenomenon arises.

For convenience we set $N+1 = \frac{1}{\varepsilon}$. The projected equation for the modified Jackson kernel are

$$\partial_t u^N_{\varepsilon,i} + \operatorname{div} \int a(\xi) f^N_{\varepsilon,i} d\xi = \frac{1}{\varepsilon}\left[c_i^{modJ} u^N_{\varepsilon,i} - u^N_{\varepsilon,i}\right]$$

$$= (N+1)\left(\frac{(N+1-i)\cos\frac{\pi i}{N+1} + \sin\frac{\pi i}{N+1}\cot\frac{\pi}{N+1}}{N+1} - 1\right) u^N_{\varepsilon,i}$$

$$= \left((N+1-i)\cos\frac{\pi i}{N+1} + \sin\frac{\pi i}{N+1}\cot\frac{\pi}{N+1} - N - 1\right) u^N_{\varepsilon,i} = -h_N(i) u^N_{\varepsilon,i}.$$

Elementary calculations show that $h_N(0) = 0$, and that $h_N(x) > 0$ for $0 < x < N$ with $h_N(x) \to 0$ for all x as $N \to 0$. One also has that $0 < h_N(i) < i$ for $0 < i \leq N$. It implies after integration in x

$$\partial_t \int u^N_{\varepsilon,i} dx = -h_N(i) \int u^N_{\varepsilon,i} dx \Longrightarrow \int u^N_{\varepsilon,i} dx(t) = e^{-h_N(i)t} \int u^N_{\varepsilon,i} dx(0) \Longrightarrow \lim_{t\to\infty} \int u^N_{\varepsilon,i} dx(t) = 0.$$

This damping phenomenon of the moments $i \neq 0$ also shows up if one uses the Jackson kernel, and is even stronger starting from the Fejer kernel. This is the price to pay for the good theoretical properties. However, with this regard this formulation is less satisfactory than moment methods like (4) which do not damp.

This spurious damping is not satisfactory for practical purposes and motivates the intrinsic method studied in the next section.

3 Kinetic polynomials

In order to define another polynomial system without damping, we introduce directly a different polynomial approximation of the Maxwellian $M(u;\xi) = \mathbb{1}_{\{0<\xi<u\}}$, which we call 'Kinetic Polynomials'. We use a minimization procedure instead of the convolution in ω according to a construction depicted in Figures 1 and 2.

We first present the minimization principles underlying the construction of kinetic polynomials. Then, we present a layered algorithm that allow to build practically and efficiently these polynomials.

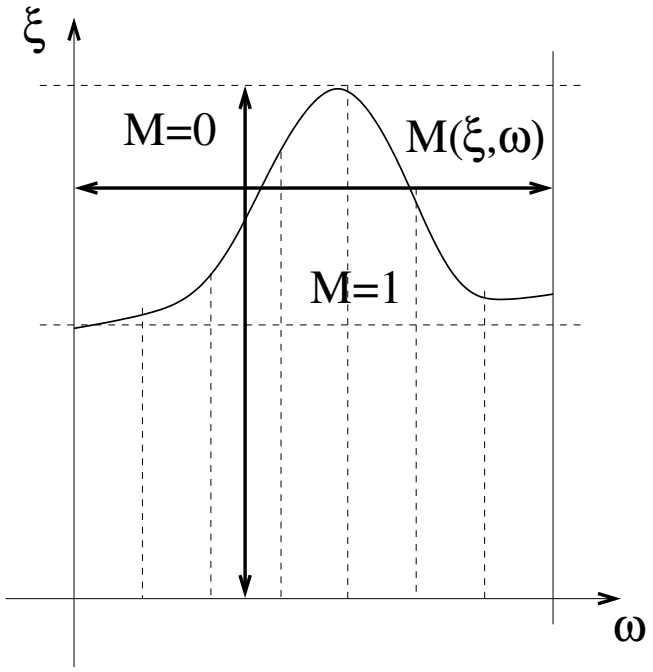


Figure 1: Plot of the upper limit of the support of $M(\omega,\xi)$ in the (ω,ξ) plane, identical to the graph of the function $\omega \mapsto u(\omega)$. The convolution method is based on some averages on horizontal lines, and does not preserve the integral on the vertical lines which is preserved, by construction, for the minimization method of section 3.1.

3.1 Minimization principle for kinetic polynomials

The purpose in this section is to generalize the universal entropy principle (8) and to construct, for any $N \geq 0$, an equilibrium $M^N(u^N;\xi,\omega)$, independent of S, which satisfies

$$M^N(u^N) = \underset{g^N \in K^N(u^N)}{\operatorname{argmin}} \int S'(\xi) g^N d\xi d\mu(\omega), \quad \text{for all admissible functions } S, \tag{32}$$

where $K^N(u^N)$ is the set of states defined by

$$K^N(u^N) = \left\{ g^N(\cdot,\cdot) \in P^N_\omega,\ u^N(\omega) = \int g^N(\xi,\omega) d\xi,\ 0 \leq g^N \leq 1 \right\}, \quad \text{for } u^N(\cdot) \in P^N_\omega.$$

In order to reformulate this problem, we remark that

$$S'(\xi) = \int_0^\infty S''(s) a_s(\xi) ds, \quad a_s(\xi) = \mathbb{I}_{\{0<s<\xi\}} \tag{33}$$

meaning that any function S' such that $S'' \geq 0$ and $S'(0) = 0$ is a non-negative integral of functions $a_s(\xi)$ which also satisfy $a_s' \geq 0$ and $a_s(0) = 0$. So we replace (32) with a family of similar problems

$$M^N(u^N) = \underset{g^N \in K^N(u^N)}{\operatorname{argmin}} \int_\xi^\infty g^N ds d\mu(\omega), \quad \forall \xi. \tag{34}$$

Since the mass must be preserved, that is $\int g^N(s,\omega)dsd\mu(\omega) = \int u^N(\omega)dsd\mu(\omega)$, this problem can be rewritten with the alternative formulation

$$M^N(u^N) = \underset{g^N \in K^N(u^N)}{\operatorname{argmax}} \int_0^\xi g^N dsd\mu(\omega), \quad \forall \xi. \tag{35}$$

Remark 3.1 *Even if the formulation (35) is simpler than (32), it is still involved. Indeed M^N is required to be the single maximizer of an infinite number of maximization problem, i.e. for all ξ. Considering the theory of polynomial approximation, it is possible to imagine that a clever use of (35) combined with the polynomial structure could generate one single maximization problem that encompasses all the properties of M^N. But it is still an open problem to find this unique maximization formulation.*

In the case $N = 0$, the formulation is simply (8). As shown in the literature [5, 19, 16], the degeneracy question is untied by using a strictly convex entropy $S'' > 0$, and noticing that the unique minimum is the same for all strictly convex entropy $S'' > 0$. But in our general case $N > 0$, the formulations (32), (34) and (35) are definitely linearly degenerate in the variable ω. This is the reason of the difficulties encountered for the moment with these maximization formulations.

This is why we try less to analyze the abstract properties of this problem in terms of existence and uniqueness of the solutions, but more to reformulate it in a way adapted to design a constructive algorithm. Hence we will be able to define another formulation, denoted below as the second maximization problem, which reveals to be constructive and is the one that we use in the sequel.

Let $u_+ = \max_{\omega\in I} u^N(\omega)$ and $u_- = \min_{\omega\in I} u^N(\omega)$. We notice that $K^N(u^N)$ is non empty since

$$\widehat{g}^N = \mathbb{1}_{\{0<\xi<u_-\}} + \frac{u^N(\omega) - u_-}{u_+ - u_-}\mathbb{1}_{\{u_-<\xi<u_+\}} \in K^N(u^N).$$

This function vanishes for $u_+ < \xi$ and is identically equal to 1 for $0 < \xi < u_-$. Considering (34) and (35), we deduce that if $M^N(u^N)$ exists, it also vanishes identically for $u_+ < \xi$ and must be identically equal to 1 for $\xi < u_-$.

Therefore we introduce these constraints in the set of possible solutions

$$S^N(u^N) = \left\{ g^N \in K^N(u^N) : \quad g^N = 1 \text{ for } 0 < \xi < u_- \text{ and } g^N = 0 \text{ for } u_+ < \xi \right\}$$

where $u^N(\cdot) \in P^N_\omega$ and $0 \le u^N$. Since $\widehat{g}^N \in S^N(u^N)$, one has $S^N(u^N) \neq \emptyset$. We obtain the constraint maximization problem.

Problem 3.2 (First maximization problem) *Find a unique* Maxwellian polynomial $M^N(u^N) \in S^N(u^N)$ *such that*

$$\int_0^\xi M^N dsd\mu(\omega) \ge \int_0^\xi \widehat{g}^N dsd\mu(\omega), \quad \forall \xi, \quad \forall \widehat{g}^N \in S^N(u^N).$$

This is a family of maximization problems of a linear functional, over a convex set $S^N(u^N)$. An interpretation is that if a solution exists for (34) or (35), then the mass at each level $\xi \mapsto \int M^N(u^N;\xi,\omega)d\mu(\omega)$ is maximized for smaller ξ: it is already a property of the solution of the initial problem (8). The whole problem is to construct the maximizer (if it exists) for values of ξ between u_- and u_+. As stressed in remark 3.1 the existence and uniqueness of a solution to this

problem is nevertheless an open problem. We will build on the idea that the problem must be analyzed level after level (i.e. layer after layer) to construct layered feasible solutions under the form: $M^N(u^N;\xi,\omega) = \sum_{l=0}^{L} h_l^N(\omega)\mathbb{1}_{\{\xi_l<\xi<\xi_{l+1}\}}$ with $0 = \xi_0 < \cdots < \xi_{L+1} = u_+$ and $h_l^N \in P_\omega^N$ polynomial inside each layer. See Figure 2.

In what follows, we assume that the problem 3.2 admits a unique solution, and detail some immediate consequences which explains the interest of this problem.

Proposition 3.3 *Under the assumption that a solution exists to the maximization problem 3.2, then it is a minimizer of (32) (that is for any S convex).*

Proof. This is a consequence of the identity (33). □

Proposition 3.4 *Cconsider a nonnegative polynomial approximation of the initial data $u^{\mathrm{init,N}} \geq 0$. Under the assumption that a solution exists to the maximization problem 3.2, then the solution of the kinetic equation*

$$\begin{cases} \partial_t f_\varepsilon^N + a(\xi).\nabla f_\varepsilon^N + \frac{1}{\varepsilon} f_\varepsilon^N = \frac{1}{\varepsilon} M^N(u_\varepsilon^N;\xi,\omega), \\ u_\varepsilon^N(x,\omega,t) = \displaystyle\int f_\varepsilon^N(x,\xi,\omega,t)d\xi, \\ f_\varepsilon^N(t=0) = M^N(u^{\mathrm{init,N}};\xi), \end{cases} \tag{36}$$

satisfies the maximum principle. For all smooth convex functions $S(\cdot)$, we have the entropy inequality

$$\partial_t \int S'(\xi) f_\varepsilon^N(x,\xi,\omega,t)d\xi d\mu(\omega) + \mathrm{div} \int a(\xi)S'(\xi) f_\varepsilon^N(x,\xi,\omega,t)d\xi d\mu(\omega) \leq 0.$$

Proof. Immediate. Notice the assumption on the non negativity of the initial data can easily be realized by using the previous method with convolution kernels: $u^{\mathrm{init,N}} = G^N *_\omega u^{\mathrm{init}}$. □

Proposition 3.5 (Derivation of the polynomial system) *Under the assumption that a solution exists to the maximization problem 3.2, and if u_ε^N converges strongly to some u^N, then we can pass to the limit $\varepsilon \to 0$ in (36) and obtain the system of conservation laws*

$$\partial_t u_i^N + \mathrm{div}\ \mathcal{F}_i^N[u^N] = 0, \quad 0 \leq i \leq N, \qquad \mathcal{F}_i^N[u^N] := \int a(\xi) M^N(u^N;\xi,\omega) T_i(\omega) d\xi d\mu(\omega), \tag{37}$$

with the entropy inequalities, for all smooth convex function $S(\cdot)$,

$$\partial_t \mathcal{S}^N[u^N] + \mathrm{div}\ \mathcal{G}^N[u^N] \leq 0,$$

where the entropy and entropy fluxes are defined by

$$\mathcal{S}^N[u^N] := \int S'(\xi) M^N(u^N;\xi,\omega) d\xi d\mu(\omega), \quad \mathcal{G}^N[u^N] := \int S'(\xi) a(\xi) M^N(u^N;\xi,\omega) d\xi d\mu(\omega).$$

Proof. Because of the bounds $0 \leq f_\varepsilon^N(x,\xi,\omega,t) \leq 1$, we may extract subsequences such that $f_\varepsilon^N \rightharpoonup f^N \in P_\omega^N$ in $L^\infty - w*$. Because, u_ε^N converges strongly, $M^N(u_\varepsilon^N)$ also converges strongly to $M^N(u^N)$. From the equation (36) (once multiplied by ε), we conclude that $f^N = M^N(u^N)$. To find the system of conservation laws, it remains to integrate in ξ the equation (36) and pass to the limit. Similarly, we pass to the limit in the entropy inequality. $\square$

This system, which is equivalent to (11) once projected on the orthonormal basis, has a much better structure than those in section 2.3 because they do not contain relaxation terms in the right hand side. Also, the family of entropy inequalities explains that it satisfies the maximum principle.

The rest of this work is devoted to construct a reasonable, or feasible, solution to the maximization problem. It is possible to prove the following results. For $N = 0$, the solution of the first maximization problem 3.2 exists, is unique and is of course equal to $M(u)$ (single layer). For $N = 1$ the solution of the first maximization problem 3.2 exists, is simple to construct in two layers and is also unique. This is explained in section 3.2.

For $N > 1$, we reformulate the problem. The idea is that if one determines the function $M^N(u^N;\xi,\omega)$ under a certain threshold $\xi < \xi_*$, then the maximization formulation 3.2 implies that

$$\int_0^{\xi_*+\alpha} M^N(u^N;\xi,\omega)d\xi d\mu(\omega) \geq \int_0^{\xi_*} M^N(u^N;\xi,\omega)d\xi d\mu(\omega) + \int_{\xi_*}^{\xi_*+\alpha} \widehat{g}^N(\xi,\omega)d\xi d\mu(\omega), \quad \alpha > 0,$$

that is

$$\int_{xi_*}^{\xi_*+\alpha} M^N(u^N;\xi,\omega)d\xi d\mu(\omega) \geq \int_{\xi_*}^{\xi_*+\alpha} \widehat{g}^N(\xi,\omega)d\xi d\mu(\omega), \quad \alpha > 0.$$

The limit $\alpha \to 0^+$ gives another characterization as a maximization problem one layer after the other:

Problem 3.6 (Second maximization problem) *Assume $M^N(u^N) \in S^N(u^N)$ is a solution of the first maximization problem 3.2. Then it is solution of another maximization problem*

$$\lim_{\alpha\to 0^+} \frac{1}{\alpha}\int_\xi^{\xi+\alpha} M^N(\xi,\omega)d\xi d\mu(\omega) \geq \lim_{\alpha\to 0^+} \frac{1}{\alpha}\int_\xi^{\xi+\alpha} \widetilde{g}^N(\xi,\omega)d\mu(\omega), \quad \forall\xi \tag{38}$$

and for all $\widetilde{g}^N$ such that $\widehat{g}^N(\xi',\omega) = M^N(\xi',\omega)\mathbb{1}_{\{0<\xi'<\xi\}} + \widetilde{g}^N(\xi',\omega)\mathbb{1}_{\{\xi<\xi'<\xi+\alpha\}} \in S^N(u^N)$.

Remark 3.7 *Like the first maximization problem, this is an infinite family (that is for all ξ) of maximization problems of a linear functional, over a convex set $S^N(u^N)$. But the difference if that it is now ordered, in the sense that we can try to maximize for small ξ, and after that to maximize on a small layer $\xi + \varepsilon$, and so on.*

We will use this principle to construct the unique solution of the second maximization problem 3.6 with a layered technique: that is we will construct $M^N(u^N)$ step by step (layer by layer).

A natural interpretation is that $M^N(u^N)$ is a function that tries to have as much mass as possible for small ξ. If we see $\xi = 0$ as where a ground state $M^N(0,\omega) = 1$ is, and the function $M^N(u^N)$ as some generalized "particles" distribution like in statistical physics, then the algorithm just tries to pill up "particles" above the ground state.

3.2 $N = 1$

A preliminary result is the following, which is the consequence that polynomials of degree 1 can be characterized by their values at the two end points.

Proposition 3.8 *Consider the case $N = 1$ and $I = [-1, 1]$. Denote the constraint as $u^1(\omega) = \alpha\omega + \beta \geq 0$ and assume for instance that $\alpha \geq 0$. Then, the solution of the first maximization problem 3.2 exists, is unique with the the layered form*

$$M^1(u^1; \xi, \omega) = \mathbb{1}_{\{0<\xi<\min(u(-1),(u(1))\}} + \frac{1+\omega}{2}\mathbb{1}_{\{u(-1)<\xi<u(1)\}} + \frac{1-\omega}{2}\mathbb{1}_{\{u(1)<\xi<u(-1)\}}$$

where one of the two last terms vanishes.

Proof. For a first order polynomial in ω, $g^1(\xi, \omega)$, we set $c(\xi) = g^1(\xi, -1)$ and $d(\xi) = g^1(\xi, 1)$, so that

$$\begin{cases} g^1(\xi, \omega) = c(\xi)\frac{1-\omega}{2} + d(\xi)\frac{1+\omega}{2}, \\ g^1 \in S^1(u^1) \Longleftrightarrow \int c(\xi)d\xi = u(-1),\ \int d(\xi)d\xi = u(1),\ 0 \leq c(\xi) \leq 1 \text{ and } 0 \leq d(\xi) \leq 1. \end{cases}$$

The bounds $0 \leq c, d \leq 1$ guarantee that the linear polynomial is in bounds $0 \leq g^1 \leq 1$. We define the positive coefficients $\gamma = \int \frac{1-\omega}{2}d\mu(\omega)$ and $\delta = \int \frac{1+\omega}{2}d\mu(\omega)$, with

$$\int S'(\xi)g^1(\xi, \omega)d\omega = \gamma \int S'(\xi)c(\xi)d\xi + \delta \int S'(\xi)d(\xi)d\xi.$$

Therefore the maximization problem $N = 1$ is equivalent to two separate maximization problems $N = 0$ like (8). The unknowns are c and d with the constraints $0 \leq c, d \leq 1$. The solution is given by $c(\xi) = \mathbb{1}_{\{0<\xi<u(-1)\}}$ and $d(\xi) = \mathbb{1}_{\{0<\xi<u(1)\}}$ which ends the proof. □

3.3 The general case $N > 1$

We design the solution of the second maximization problem 3.6 hereafter by a constructive method (an algorithm) under the form

$$M^N(u^N; \xi, \omega) = \sum_{l\geq 0} h_l^N(\omega)\mathbb{1}_{\{\xi_l<\xi<\xi_{l+1}\}}, \quad 0 = \xi_0 < \xi_1 < \cdots < \xi_L < \xi_{L+1} = u_+ = \max_I u^N(\omega). \quad (39)$$

The construction shows the uniqueness of the solution. The layer structure of this function is illustrated in Figure2. The integral identity $\int_0^{u_+} M^N(u^N; \xi, \omega)d\xi = u^N(\omega)$ writes

$$\sum_{l\geq 0}(\xi_{l+1} - \xi_l)h_l^N(\omega) = u^N(\omega), \quad \omega \in I. \quad (40)$$

This function is constructed step by step departing from the bottom, the first step being trivial. The second step is the critical one where all the ideas of the method are carefully explained, in particular the role of the Bojavic-Devore theorem for one sided approximation. The other steps are designed with the same method, and Theorem 3.10 guarantees the completion of the algorithm.

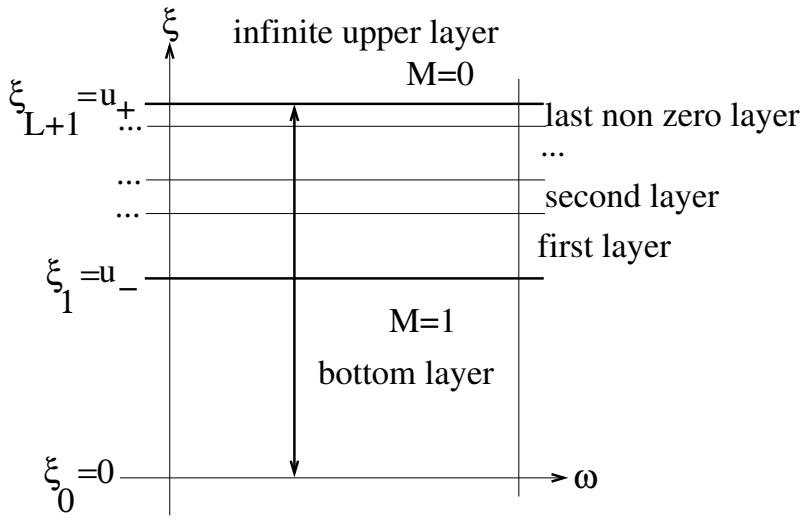


Figure 2: Layered structure of the kinetic polynomial M^N, a polynomial of degree less or equal to N in ω which is constant in each layer (ξ_i, ξ_{i+1}). The vertical line indicates that the integral on the line is equal to the integral on the same line in Figure 1.

3.3.1 Construction of h_0^N in the lower layer

The solution in the lower layer is by definition of S^N

$$h_0^N(\omega) = 1 \quad \forall \omega, \tag{41}$$

with the layer size

$$\xi_1 = u_-. \tag{42}$$

After this stage, it remains to construct the next layers. The integral relation (40) written for the next layers $l = 1, 2, \ldots$ becomes

$$\sum_{l \geq 1} (\xi_{l+1} - \xi_l) h_l^N(\omega) = v^N(\omega) := u^N(\omega) - u_-, \quad \omega \in I. \tag{43}$$

By construction v^N reaches its minimum 0 at some points denoted as ω_i, $1 \leq i \leq p$. It also reaches its maximum $D = u_+ - \xi_1$ at some points denoted as μ_j, $1 \leq j \leq q$, and one has that

$$0 \leq \frac{1}{D} v^N \leq 1. \tag{44}$$

In the sequel the points ω_i and μ_j are called **points of contact**, as illustrated in Figure 3. We naturally define an integer referred to as **the local order of the contact**. It is an even number if ω_j (resp. μ_i) is inside the interval I. It can be any non zero natural number if $\omega_j = \pm 1$ (resp. $\mu_i = \pm 1$) is on the extremities. The **local order of contact** at ω_i (rest. μ_j) will be denoted as $r_i + 1$ (resp. $s_j + 1$). The **total order of contact** of the polynomial $\frac{1}{D} v^N$ is defined by $\sum_i (r_i + 1) + \sum_j (s_j + 1)$. These notions play critical role in the construction of h_1^N which is detailed below.

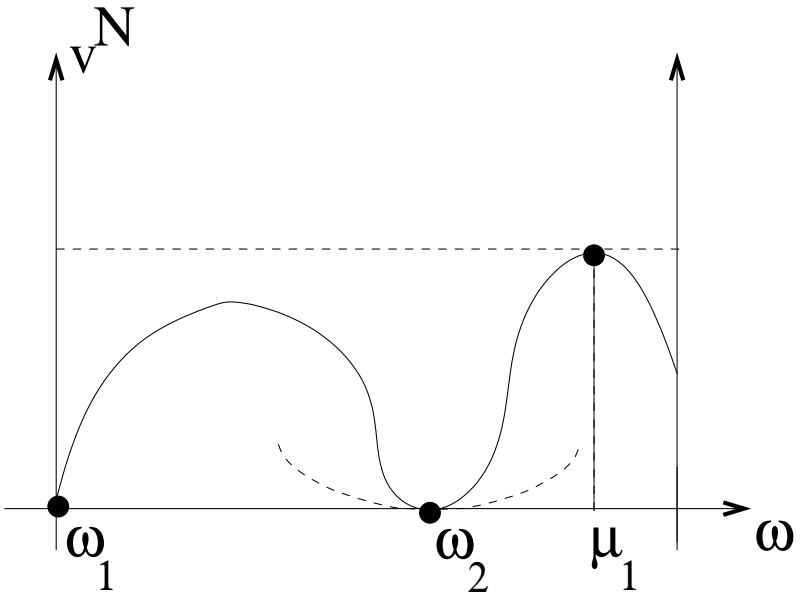


Figure 3: Example of contact points. Here ω_1 has order 1, ω_2 has order 2 and μ_1 has order 2. The total order is $5 = 1 + 2 + 2$. Also is represented at ω_2 a function which has locally a contact order greater or equal to the one of the main function.

3.3.2 Construction of h_1^N in the first layer

The two unknowns that will be defined at the end of this section are the polynomial $h_1^N(\omega)$ and ξ_2 which gives the length $\xi_2 - \xi_1$ of the layer. Due to the points of contact, this construction, detailed in 6 steps, is largely constrained.

Step 1-Expression of the constraints: Let us examine the function

$$g^N(\xi, \omega) = \sum_{l \geq 1} h_l^N(\omega) \mathbb{1}_{\{\xi_l < \xi < \xi_{l+1}\}}.$$

Since we desire to impose

$$\int_{\xi_1}^{D} g^N(\xi, \omega) d\xi = \sum_{l \geq 1} (\xi_{l+1} - \xi_l) h_l^N(\omega) = v^N(\omega) \text{ for all } \omega \in I, \quad \text{and } 0 \leq h_l^N \leq 1 \text{ for all } l, \tag{45}$$

it means that at all points of contact and for all $l \geq 1$

$$h_l^N(\xi, \omega_i) = 0 \text{ for } 1 \leq i \leq p, \text{ and } h_l^N(\xi, \mu_j) = 1 \text{ for } 1 \leq j \leq q. \tag{46}$$

The ω_j and μ_i, defined at the end of the previous step, are the points of contact of the graph of $\frac{1}{D} v^N$ at values 0 and 1. In particular the points of contact of $\frac{1}{D} v^N$ are also points of contact for h_1^N.

Actually it can be proved that the local contact orders for h_1^N are greater or equal to the contact orders for $\frac{1}{D} v^N$. The example depicted in figure 3 illustrates that property: Let us assume that $\omega_2 \in (-1, 1)$ is a contact order $r_1 = 4$ for v^N, and the contact order of h_1^N is less than 4, let us take 2. In this case we can write $h_1^N(\omega) = c(\omega - \omega_2)^2 + O(\omega - \omega_2)^3$, $c > 0$. On the other hand $v^N(\omega) = \widehat{c}(\omega - \omega_2)^4 + O(\omega - \omega_2)^5$, $\widehat{c} > 0$. In this case it is sure that

$$\forall \tau > 0, \quad \exists\, \overline{\omega} \text{ close to } \omega_2 \text{ with } v^N(\overline{\omega}) < \tau h_1^N(\overline{\omega}).$$

So it will not be possible to enforce (45). This is a contradiction. It means the local expansion of h_1^N at the point of contact ω_2 is of the form

$$h_1^N(\omega) = \widetilde{c}(\omega - \omega_2)^{2p} + O(\omega - \omega_2)^{2p+1}, \quad \widetilde{c} > 0, \quad p \geq 2.$$

This situation, when the contact order of h_1^N is greater or equal to the one of v^N, is represented in dashed in the figure 3. This example is easily generalized to all other points of contact: we obtain that all points of contact of v^N are also points of contact for h_1^N with a contact order greater or equal to r_i (resp. s_j).

Let us now consider the polynomial $H = \frac{1}{D}v^N$. By construction one has that

$$(h_1^N - H)^{(r)}(\omega_i) = 0 \quad 0 \leq r \leq r_i \quad 1 \leq i \leq q, \text{ and } (h_1^N - H)^{(s)}(\mu_j) = 0 \quad 0 \leq s \leq s_j \quad 1 \leq j \leq q.$$

It implies that h_1^N can be decomposed as

$$h_1^N(\omega) = H(\omega) + W(\omega)r(\omega) \tag{47}$$

where W is a polynomial which can be written as

$$W(\omega) = \Pi_1^p|\omega - \omega_i|^{r_i}\Pi_1^q|\omega - \mu_j|^{s_j}, \quad \deg(W) \leq N. \tag{48}$$

W is indeed a polynomial over I due since r_i and s_j are even for interior points of contact.

Step 2-Maximization under constraints: We define the admissible set

$$S_1 = \left\{ r \in P^{N-\deg(W)},\ 0 \leq H + Wr \leq 1 \right\}.$$

Since $0 \leq H$ (see 44), one has that $0 \in S_1$ which is non empty. Therefore the maximization problem

$$r = \text{argmax}_{r \in S_1} \int \left(H(\omega) + W(\omega)r(\omega)\right) d\mu(\omega) \tag{49}$$

makes sense. It can be recast also as

$$r = \text{argmax}_{r \in S_1} \int W(\omega)r(\omega)d\mu(\omega). \tag{50}$$

Step 3-Design of h_1^N: The integral (49) defines a linear functional. The set S_1 is non empty, compact and convex. So there exists at least one solution to the maximization problem (49). The Bojavic-Devore theorem [2][Theorem 3] yields the uniqueness of the solution. Actually the original Bojavic-Devore theorem considers only one bound, but it is immediate to extend[1] the technical part [2][Lemma 3] of its proof to the case with two bounds, a lower bound and an upper bound, as in $g \leq P \leq f$ assuming $g < 0 < f$.

[1] A sketch of the proof is as follows, using the notation of the seminal Bojavic-Devore paper. The key part concerns the polynomial $P + \eta Q_\epsilon$ defined bottom of page 146, where by definition $g \leq P \leq f$. To take into account the lower bound which is needed in our formulation, it is needed to check that $g \leq P + \eta Q_\epsilon$ for all x. The verification goes as follows. Case a) If $x \in [a, b]/\mathcal{I}$, then $Q_\varepsilon(x) \geq 0$ by construction. So $g \leq P(x) + \eta Q_\epsilon(x)$. Case b). If $x \in \mathcal{I}$, then $Q_\varepsilon(x) \leq 0$ by construction so it is the dangerous case. But at the same time $P(x_\nu) = f(x_\nu)$ where x_ν the point of contacts of P on f. Since $f \geq \alpha > 0$ and for all $x \in \mathcal{I}$, there exists $x_\nu \in \mathcal{I}$ such that $|x - x_\nu| \leq \varepsilon$, it is sufficient to take $\varepsilon > 0$ as small as needed to guarantee that $P(x) \geq \beta > g$ for all $x \in \mathcal{I}$. Taking $\eta > 0$ small enough ends the proof of the technical Lemma 3. The same method works for the Lemma 4. The rest of the proof is the same.

Theorem 3.9 (Bojavic-Devore [2]) *Let w be a non negative Lebesgue integrable weight function: $\int w(\omega)d\omega > 0$. Let $g < 0 < f$ be two functions both differentiable on (a,b), except possibly at a finite number of points τ_c $(c = 1,\dots,C)$ where $f(\tau_c) = +\infty$ or $g(\tau_c) = -\infty$. Let $n \geq 0$ and $Q^n(z) = \{r^n \in P^n, 0 \leq r^n \leq f\}$.*
Then, there exists a unique maximizer $r^n = argmax_{s^n \in Q^n(z)} \int_a^b w(\omega)s^n(\omega)d\omega$.

Considering the maximization problem (50) with the definition (48) of W, we can apply the Bojanic-Devore theorem with the weight $w(\omega) = W(\omega)\frac{d\mu(\omega)}{d\omega}$, the function $f = \frac{1-H}{W} > 0$, the function $g = \frac{-H}{W} < 0$ and the degree is $n = N - \deg W$. Notice that since W has all the points of contact of H with orders, then $g < 0 < f$. The value of the function f is infinite at a root of W which is not a root of $1 - H$ (with multiplicity). Similarly the function $-g$ is infinite at a root of W which is not a root of $-H$ (with multiplicity). So that all the assumptions of the theorem are fulfilled.

It defines a unique r solution of (49)-(50) and we set $h_1^N = H + Wr$. By definition, the solution satisfies the comparison inequality

$$\int h_1^N(\omega)d\mu(\omega) \geq \int H(\omega)d\mu(\omega), \quad H = \frac{1}{D}v^N(\omega). \tag{51}$$

Step 4-Design of ξ_2: Now that h_1^N has been determined, it remains to decide of a value for ξ_2 such that

$$g^N(\xi,\omega) = h_1^N(\omega) \text{ for } \xi_1 < \xi < \xi_2.$$

Whatever the value of ξ_2, we will have to design the next layers so that

$$\sum_{l\geq 2}(\xi_{l+1} - \xi_l)h_l^N(\omega) = v^N(\omega) - (\xi_2 - \xi_1)h_1^N(\omega), \quad \omega \in I. \tag{52}$$

This problem has the same structure than the original problem over a smaller length $\widetilde{D} = D - (\xi_2 - \xi_1)$ and for a function $\widetilde{v}^N = v^N(\omega) - (\xi_2 - \xi_1)h_1^N(\omega)$. It can be solved with the method we used to construct h_1^N if

$$0 \leq \frac{1}{\widetilde{D}}\widetilde{v}^N \leq 1. \tag{53}$$

Indeed this inequality is the same as (44) which is at the starting point of the construction of the solution in the this layer. So if we manage to guarantee (53) it will be able to continue the construction. Therefore we define ξ_2 as the largest value so that (53) is true. It writes

$$0 \leq \frac{v^N - (\xi_2 - \xi_1)h_1^N}{D - (\xi_2 - \xi_1)} \leq 1 \Longleftrightarrow \left\{ \begin{array}{l} (\xi_2 - \xi_1)h_1^N \leq v^N, \\ (\xi_2 - \xi_1)(1 - h_1^N) \leq D - v^N. \end{array} \right. \tag{54}$$

By construction the polynomial h_1^N divides the polynomial v^N, and has the same points of contact (with equal of greater order of multiplicity) of v^N. We define two rational fractions

$$w^N = \frac{v^N(\omega)}{h_1^N(\omega)} \geq \alpha > 0 \text{ and } z^N = \frac{1 - v^N(\omega)}{1 - h_1^N(\omega)} \geq \beta > 0. \tag{55}$$

In general, w^N and z^N are polynomials, but they may be rational fractions if the contact order of h_1^N at a given point is strictly greater than the contact order of v^N at the same point. We take the largest as possible ξ_2 which satisfies (54), that is

$$\xi_1 < \min\left(\xi_1 + \min_{\omega\in I}\left(w^N(\omega), z^N(\omega)\right), u_+\right) = \xi_2 \leq u_+. \tag{56}$$

Step 5-Proof that $\xi_2 \le u_+$: It stems directly from (56). Indeed (54) yields

$$(\xi_2 - \xi_1) - D + v^N \le (\xi_2 - \xi_1)h_1^N \le v^N \Longrightarrow \xi_2 \le D + \xi_1 = u_+.$$

In other words if $\xi = u_+$ the algorithm just completed and there is nothing more to say. In the other case $\xi_2 < \xi_1 + D = u_+$. It also means that one can forget the u_+ in the definition (56) of ξ_2.

Step 6-Increase of the contact order. We want to show that the total contact order of $\widetilde{v}^N$ is strictly greater than the total contact order of v^N. Since $\widetilde{v}^N = v^N - (\xi_2 - \xi_1)h_1^N$ and h_1^N has all the points of contact of v^N (with greater or equal multiplicity order), then the total contact order cannot decrease. It remains to show that is increases strictly.

We analyze this property for the case $\xi_2 < D$, that is

$$\text{either } \min_{\omega\in I} w^N(\omega) < D \quad \text{ or } \quad \min_{\omega\in I} z^N(\omega) < D.$$

So let us assume $\min_{\omega\in I} w^N(\omega) \le \min_{\omega\in I} z^N(\omega)$ and that $\min_{\omega\in I} w^N(\omega) < D$. The other case is symmetric. Set $d = \xi_2 - \xi_1 < D$ so that $0 \le v^N(\omega) - dh_1^N(\omega)\ \forall \omega \in I$. One also has

$$\text{for all sufficiently small } \varepsilon > 0, \quad \exists \omega_\varepsilon \in I, \quad v^N(\omega_\varepsilon) - (d+\varepsilon)h_1^N(\omega_\varepsilon) < 0. \tag{57}$$

Notice that $\omega_\varepsilon \ne \omega_i$, if not it is in contradiction with (57) since both polynomials vanish at ω_i. Passing to the limit after extraction of a subsequence in I which is a closed bounded interval, there exists $\omega_* \in I$ such that

$$\omega_\varepsilon \to \omega_* \quad \text{ and } \quad v^N(\omega_*) - dh_1^N(\omega_*) = 0. \tag{58}$$

The discussion considers two cases.

• If $\omega_* \ne \omega_i$ for all i, where ω_i is a point of contact of v^N, then it adds one more point of contact in the list of the ω_i's. In this case the property is proved.

• The other case when $\omega_* = \omega_i$ for a given i needs a little more work. Assume the local expansion of v^N can be written

$$v^N(\omega) = c(\omega - \omega_i)^{r_i+2} + O(\omega - \omega_i)^{r_i+3}, \quad 0 < c,\ 0 < r_i,$$

where $-1 < \omega_i < 1$ for the simplicity of notations, so that r_i is necessarily even. The local expansion of h_1^N can be written under a similar form

$$h_1^N(\omega) = \widetilde{c}(\omega - \omega_i)^{\widetilde{r}_i+2} + O(\omega - \omega_i)^{\widetilde{r}_i+3}, \quad \widetilde{0} < \widetilde{c},\ 0 < r_i \le \widetilde{r}_i.$$

But in view of (57-58), it is necessary that $\widetilde{r}_i = r_i$. Then

$$v^N(\omega_\varepsilon) - (d+\varepsilon)h_1^N(\omega_\varepsilon) = (c - (d+\varepsilon)\widetilde{c})(\omega_\varepsilon - \omega_i)^{r_i+2} + O(\omega_\varepsilon - \omega_i)^{r_i+3}$$

Using (57) and $\omega_\varepsilon - \omega_i \ne 0$, we see that

$$(c - (d+\varepsilon)\widetilde{c}) + O(\omega_\varepsilon - \omega_i) < 0.$$

Passing to the limit we obtain that $c - d\widetilde{c} \le 0$. But at the same time $0 \le v^N - dh_1^N$. In view of the local expansion

$$v^N(\omega_\varepsilon) - dh_1^N(\omega_\varepsilon) = (c - d\widetilde{c})(\omega_\varepsilon - \omega_i)^{r_i} + O(\omega_\varepsilon - \omega_i)^{r_i+1}$$

it implies $0 \le c - d\widetilde{c}$. Therefore $c - d\widetilde{c} = 0$, meaning that

$$v^N(\omega_\varepsilon) - dh_1^N(\omega_\varepsilon) = O(\omega_\varepsilon - \omega_i)^{r_i+3}.$$

It shows that the local contact number is increased at least by 1.

In all cases there is strict increase of the total contact number.

3.3.3 Next steps

The problems in the second layer, and in the third layer, fourth layer, ..., have exactly the same structure as in the first layer. It constructs by iteration a sequence (v_l^N, ξ_l, h_l^N). The first steps are $(v_0^N, \xi_0, h_0^N) = (u^N, 0, 1)$ and $(v_l^N, \xi_l, h_l^N) = (v^N, \xi_1, h_1^N)$. The next steps are constructed with the method used to construct (v_l^N, ξ_l, h_l^N) from (v_0^N, ξ_0, h_0^N).

Some properties are guaranteed by construction

- i) one has the iterations $v_{l+1}^N(\omega) = v_l^N(\omega) - (\xi_{l+1} - \xi_l)h_l^N(\omega)$,
- ii) one has the bound $0 \leq \frac{1}{u_+ - \xi_l} v_l^N(\omega) \leq 1$,
- iii) the contact number of h_l^N (and v_l^N) increases strictly at each step,
- iv) the generalization of (51) holds

$$\int h_l^N(\omega) d\mu(\omega) \geq \frac{1}{D - \xi_l} \int v_l^N(\omega) d\mu(\omega).$$

□

Theorem 3.10 (Completion of the algorithm) *The method ends in a finite number of steps $L \leq N$, and the layers fill the interval $[0, u^+]$*

$$0 = \xi_0 < \xi_1 < \cdots < \xi_L < \xi_{L+1} = u_+.$$

By construction the function $M^N(u^N; \xi, \omega) = \sum_{l \geq 0} h_l^N(\omega) \mathbb{1}_{\{\xi_l < \xi < \xi_{l+1}\}}$ is a solution of the second maximization problem 3.6.

Proof. The proof is by contradiction. Assume the algorithm never completes. It construct a infinite series of layers, all of them $[\xi_l, \xi_{l+1}] \subset [0, u_+]$ with $\xi_l < \xi_{l+1}$. Since the total contact number increases at least by 1 at each level, it reaches $N+1$ or more at a certain step L of the algorithm. $N+1$ contacts imply $N+1$ equality constraints for h_L^N. There is only one possibility $h_L^N = \frac{1}{u_+ - \xi_L} v_L^N$. It implies that $\xi_{L+1} = u_+$ and $v_{L+1=0}^N$. The rest of the proof is evident. □

The completion property of the kinetic polynomials is also related to the identity $u^N(\omega) = \int M_{u^N}^N(\xi, \omega) d\xi$. Applying this identity for any μ_j which is an upper point of contact at every layer, one obtains

$$u_+ = u^N(\mu_j) = \sum_{l=0}^{L} (\xi_{l+1} - \xi_l) h_l^N(\mu_j) = \sum_{l=0}^{L} (\xi_{l+1} - \xi_l) = \xi_{L+1}.$$

4 A simple numerical scheme for the projected equations (11)

We discretize in time and space and implement the method under the form

$$\frac{\overline{u}_j^N - u_j^N}{\Delta t} + \frac{F^N[u_j^N] - F^N[u_{j-1}^N]}{\Delta x} = 0 \tag{59}$$

where $u_j^N \in P^N(\omega)$ is a polynomial in ω of degree N (fixed), in cell j and at the current time step $t_n = n\Delta t$ and the generic flux $F^N[u_j^N]$ is constructed with the kinetic polynomial formula (11), a more

detailed formula is (37), accordingly the construction presented in the previous section. The value at next time step $t_{n+1} = (n+1)\Delta t$ in cell j is denoted with a bar $\overline{u}_j^N \in P^N(\omega)$.

We assume the initial data is a positive and bounded polynomial. With our notations it can be written as

$$0 \leq U_m \leq u_j^N(\omega) \leq U_M < \infty, \quad \forall j \text{ and } \forall \omega \in I. \tag{60}$$

We consider the archetype of a convex flux which is the Burgers flux $F(\xi) = \frac{\xi^2}{2}$. This is compatible with the upwinding of the discrete spatial derivative visible in (59) which corresponds to the characteristic lines ω per ω of a non negative initial data.

The following result states that the explicit Euler scheme satisfies the maximum principle (this is a minimal stability requirement) under a CFL condition which is independent of N. The property is here checked directly on the scheme (59) but can also be derived as a consequence of the underlying kinetic formulation.

Theorem 4.1 *Assume the CFL condition $U_M \Delta t \leq \Delta x$. Then*

$$U_m \leq \overline{u}_j^N(\omega) \leq U_M, \quad \forall j \text{ and } \forall \omega \in I. \tag{61}$$

Proof. Using $v_j^N = F^N[u_j^N]$ to simplify the notations, one notices that by construction $\frac{u_j^N(\omega)^2}{2} \leq v_j^N(\omega) \leq \frac{U_M^2}{2}$, where the lower bound is, by the standard Brenier inequality, a consequence of (8). One rewrites the Euler scheme as

$$\overline{u}_j^N(\omega) = u_j^N(\omega) - \frac{\Delta t}{\Delta x} v_j^N(\omega) + \frac{\Delta t}{\Delta x} v_{j-1}^N(\omega). \tag{62}$$

It yields

$$\overline{u}_j^N(\omega) \leq u_j^N(\omega) - \frac{\Delta t}{\Delta x}\frac{u_j^N(\omega)^2}{2} + \frac{\Delta t}{\Delta x}\frac{U_M^2}{2} = (1-\alpha_j^N(\omega))u_j^N(\omega) + \alpha_j^N(\omega) U_M, \quad \alpha_j^N(\omega) = \frac{\Delta t}{\Delta x} \times \frac{u_j^N(\omega) + U_M}{2}.$$

The condition $0 \leq \alpha_j^N(\omega) \leq 1$ yields the propagation of the propagation of the upper bound $\overline{u}_j^N \leq U_M$. This inequality is guaranteed under the CFL condition.

To prove the lower bound in (61), we need a sharper upper bound on $v_j(\omega)$. Let us start with

$$M_j^N(u^N; \xi, \omega) = 1 \text{ for } \xi < U_m \text{ and } \int_{\xi = U_m}^{U_M} M_j^N(u^N; s, \omega) ds = u_j^N(\omega) - U_m.$$

Therefore

$$v_j^N(\omega) = \int_0^{U_M} \xi M_j^N(u^N; s, \omega) ds \leq \frac{U_m^2}{2} + \int_{U_m}^{U_M} \xi M_j^N(u^N; s, \omega) ds.$$

A Brenier type inequality [5] yields the needed upper bound

$$v_j(\omega) \leq \frac{U_m^2}{2} + \int_{U_M - (u^N(\omega) - U_m)}^{U_M} \xi d\xi = \frac{U_m^2}{2} + \frac{U_M^2}{2} - \frac{\left(U_M - \left(u^N(\omega) - U_m\right)\right)^2}{2} \leq \frac{U_m^2}{2} + U_M(u^N(\omega) - U_m).$$

Plugging in (62), we get

$$\overline{u}_j^N(\omega) \geq u_j^N(\omega) - \frac{\Delta t}{\Delta x}\frac{U_m^2}{2} - \frac{\Delta t}{\Delta x} U_M(u_j^N(\omega) - U_m) + \frac{\Delta t}{\Delta x} v_{j-1}^N(\omega).$$

It yields

$$\overline{u}_j^N(\omega) \geq u_j^N(\omega) - \frac{\Delta t}{\Delta x} U_M (u^N(\omega) - U_m) = U_m + \left(1 - U_M \frac{\Delta t}{\Delta x}\right) \left(u_j^N(\omega) - u_-\right) \geq U_m$$

where the last inequality is a consequence of the CFL condition.

By iteration, the maximum principle (61) holds true one step after the other and the proof is complete. □

We now show an explicit construction of the solution of the Theorem 3.10 for the case $N = 2$, and we use it to obtain preliminary numerical results for the corresponding projected equations.

4.1 Construction of the kinetic polynomial for $N = 2$

Let $u^2 \in P^2$ be a non negative polynomial $u^2(\omega) = a\omega^2 + b\omega + c \geq 0$. We assume $a \neq 0$ since the solution for $a = 0$ can be determined by the method for $N = 1$. We first determine u_+ and u_-. Let $\omega_0 = -\frac{b}{2a}$ be the solution of $u'(\omega_0) = 0$.

- If $\omega_0 \geq 1$ and $a > 0$: $u_+ = u^2(-1) = a - b + c$ and $u_- = u^2(1) = a + b + c$.
- If $\omega_0 \geq 1$ and $a < 0$: $u_+ = u^2(1) = a + b + c$ and $u_- = u^2(-1) = a - b + c$.
- If $\omega_0 \leq 1$ and $a > 0$: $u_+ = u^2(1) = a + b + c$ and $u_- = u^2(-1) = a - b + c$.
- If $\omega_0 \leq 1$ and $a < 0$: $u_+ = u^2(-1) = a - b + c$ and $u_- = u^2(1) = a + b + c$.
- If $-1 < \omega_0 < 1$ and $a > 0$: $u_+ = \max(u^2(-1), u^2(1)) = a + |b| + c$ and and $u_- = u^2(\omega_0) = c^2 - \frac{b^2}{4a}$.
- If $-1 < \omega_0 < 1$ and $a < 0$: $u_+ = u^2(\omega_0) = c^2 - \frac{b^2}{4a}$ and and $u_- = \min(u_2(-1), u_2(1)) = a - |b| + c$.

We introduce the notations: $D = u_+ - u_-$, $v_1^2(\omega) = u^2(\omega) - u_-$ and $w^2(\omega) = \frac{v_1^2(\omega)}{D}$.

In case the absolute extrema of u^2 strictly belongs to I, it is the same for w^2. The only possibility to reconstruct a second order polynomial when prescribing one value at ± 1, one value at ω_0 and the zero derivative at ω_0 is to keep the same polynomial. In this case we set $h_1^2 = w^2$ and

$$M_{u^2}^2(\xi, \omega) = \mathbb{1}_{\{0<\xi<u_-\}} + h_1^2(\omega) \mathbb{1}_{\{u_-<\xi<u_+\}} = M_{u^2}^2(\xi, \omega) = \mathbb{1}_{\{0<\xi<u_-\}} + \frac{u^2(\omega) - u_-}{u_+ - u_-} \mathbb{1}_{\{u_-<\xi<u_+\}}. \quad (63)$$

In the other case, w^2 has necessarily its extrema at the boundary. Assume $w^2(-1) = 0$ and $w^2(1) = 1$. We factorize $w^2(\omega) = (1 + \omega)\left(H(\Omega) + W(\omega) r(\omega)\right)$. Necessarily $W(\omega) = 1 - \omega$ so $1 = w^2(1) = 2H(1)$. So more precisely $w^2(\omega) = (1 + \omega)\left(\frac{1}{2} + (1 - \omega)r\right)$, $r \in \mathbb{R}$. The condition $0 \leq w^2 \leq 1$ on the interval reads $-\frac{1}{4} \leq r \leq \frac{1}{4}$.

In view of the representation formula, the maximum of the integral is for $r = \frac{1}{4}$, and the minimum is for $r = -\frac{1}{4}$. Let ξ_1 be the solution of $Dr = \xi_1 \frac{1}{4} + (D - \xi_1)(-\frac{1}{4})$, which is given by $\xi_1 = 2D(r + \frac{1}{4})$ where $D = u_+ - u_-$. One obtains

$$M^2(u^2; \xi, \omega) = \mathbb{1}_{\{0<\xi<u_-\}} + (1 + \omega)(\frac{3}{4} - \frac{1}{4}\omega) \mathbb{1}_{\{u_-<\xi<u_- + \xi_1\}} + \frac{1}{4}(1 + \omega)^2 \mathbb{1}_{\{u_- + \xi_1 < \xi < u_+\}}. \quad (64)$$

The case where $w^2(-1) = 1$ and $w^2(1) = 0$ is deduced by symmetry

$$M^2(u^2; \xi, \omega) = \mathbb{1}_{\{0<\xi<u_-\}} + (1 - \omega)(\frac{3}{4} + \frac{1}{4}\omega) \mathbb{1}_{\{u_-<\xi<u_- + \xi_1\}} + \frac{1}{4}(1 - \omega)^2 \mathbb{1}_{\{u_- + \xi_1 < \xi < u_+\}}. \quad (65)$$

With these formulas one can compute $\int_{\mathbb{R}^+} \xi M_2(\xi,\omega)d\xi$. One finds, in the first case (63),

$$\begin{aligned} F^2[u^2] &= \int_{\mathbb{R}^+} \xi M^2(u^2;\xi,\omega)d\xi = \tfrac{1}{2}u_-^2 + w_2(\omega)\left(\tfrac{1}{2}u_+^2 - \tfrac{1}{2}u_-^2\right) \\ &= \tfrac{1}{2}u_-^2 + (u_2(\omega)-u_-)\tfrac{u_++u_-}{2} = u_2(\omega)\tfrac{u_++u_-}{2} - \tfrac{u_+u_-}{2}. \end{aligned} \tag{66}$$

In the second case (64) the flux is

$$F^2[u^2] = \int_{\mathbb{R}^+} \xi M^2(u^2;\xi,\omega)d\xi = \frac{1}{2}u_-^2 + (1+\omega)(\frac{3}{4} - \frac{1}{4}\omega)\xi_1 + \frac{1}{4}(1+\omega)^2(D-\xi_1). \tag{67}$$

The third case (65) is symmetric and yields

$$F^2[u^2] = \int_{\mathbb{R}^+} \xi M^2(u^2;\xi,\omega)d\xi = \frac{1}{2}u_-^2 + (1-\omega)(\frac{3}{4} + \frac{1}{4}\omega)\xi_1 + \frac{1}{4}(1-\omega)^2(D-\xi_1). \tag{68}$$

4.2 Numerical examples

We compare numerical results obtained for $N = 2$ with the usual moment method, with the new method based on kinetic polynomials and also with a non intrusive method. We use the measure $d\mu(\omega) = \frac{d}{\pi\sqrt{1-\omega^2}}$ and consider first the Burgers equation. Also we restrict the numerics to $N = 2$.

The moment method, that is solving (4), is easy to compute since the orthonormal basis is formed of the Tchebycheff polynomials and the Burgers flux is also a polynomial. It writes

$$\partial_t \begin{pmatrix} a \\ b \\ c \end{pmatrix} + \partial_x \begin{pmatrix} \frac{a^2+b^2+c^2}{2} \\ ab + \frac{bc}{\sqrt{2}} \\ ac + \frac{b^2}{2\sqrt{2}} \end{pmatrix} = 0.$$

The kinetic polynomials method use the fluxes defined in (66-68).

We consider the initial data

$$u^{\text{ini}}(x,\omega) = \begin{cases} 3 & \text{for } x < 1/2 \text{ and } -1 < \omega < 0, \\ 5 & \text{for } x < 1/2 \text{ and } 0 < \omega < 1, \\ 1 & \text{for } 1/2 < x \text{ and } -1 < \omega < 1. \end{cases} \tag{69}$$

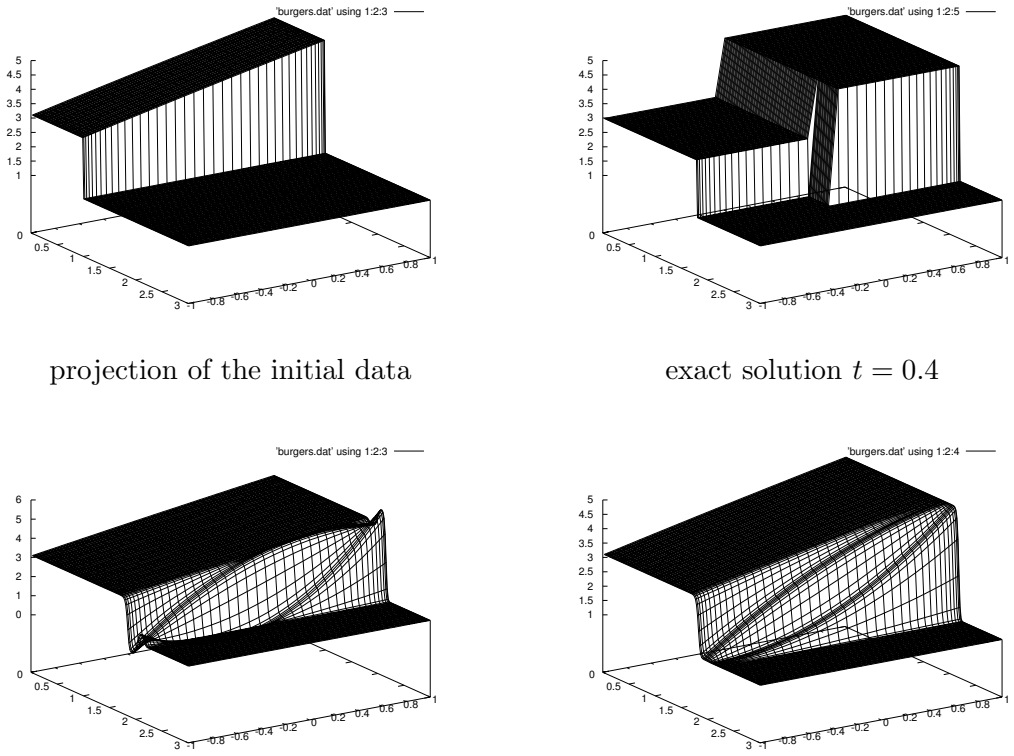
The exact solution is a shock at velocity 2 for $\omega < 0$, and another shock at velocity 3 for $0 < \omega$. The results plotted in Figure 4 show the gain in term of stability of the new method with respect to the moment method.

We consider still the Burgers equation, but with another initial data, which is continuous,

$$u^{\text{ini}}(x,\omega) = \begin{cases} 12 & \text{for } x - \omega/5 < 1/2, \\ 1 & \text{for } x - \omega/5 < 3/2, \\ 12 - 11\,(x - \omega/5 - 1/2) & \text{in between.} \end{cases} \tag{70}$$

The exact solution is a compressive ramp on all lines, and a shock at time $T = \frac{1}{11}$. Therefore the exact solution is continuous in x and ω directions for $t < T$, and is discontinuous in the ω direction for $T < t$.

Notice that we preprocess the initial data with the modified Jackson kernel in order to make sure the initial data is in bounds. As one can see it in Figure 5, the compressive nature of the solution is



projection of the initial data | exact solution $t = 0.4$

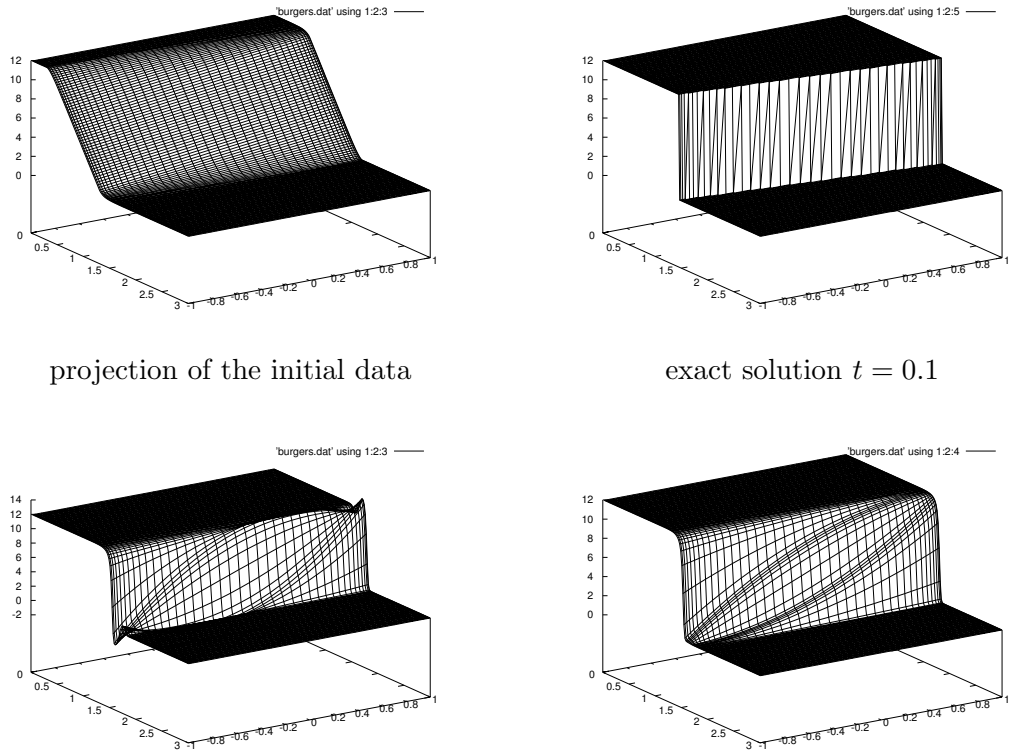
moment solution $t = 0.4$ | new method $t = 0.4$

Figure 4: Burgers flux. Illustration of the maximum principle satisfied by the method based on kinetic polynomials for the initial data (69). The usual moment solution does not satisfy the maximum principle which is a salient property of the exact solution.

evident at time $t = 0.1 > T$. The moment method generates oscillations. On the contrary, the new method based on kinetic polynomials satisfies the maximum principle at all times.

It is interesting also to compare with the non intrusive method which is fierce competitor due to its simplicity and versatility in engineering. This last method amounts to decide quadrature points denoted as ω_i with $1 \leq N+1$, to solve in parallel $N+1$ standard conservation laws, and after that to reconstruct the function in the ω direction in order to get an approximation for all ω. This method also has its difficulties. The first one is the approximation of the initial data if it is discontinuous in ω. So if a quadrature point if at the discontinuity, the initialization is ambiguous. The second one is the oscillations may show up after reconstruction with Lagrange interpolation techniques.

In the test, we use the general prescription for quadratures. We take the roots of the Tchebycheff polynomial $T_3(\omega) = 4\omega^4 - 3\omega$: $\omega_1 = -\sqrt{\frac{3}{4}}$, $\omega_2 = 0$ and $\omega_3 = \sqrt{\frac{3}{4}}$. We do not perform a test for the first initial data (69) due to the ambiguity of the initialization at ω_2. The results for the second problem (70) are displayed in figure 6. Without any surprise, we observe that the maximum principle is not satisfied, an artefact also observed with the moment method.



projection of the initial data | exact solution $t = 0.1$

moment solution $t = 0.1$ | new method $t = 0.1$

Figure 5: Burgers flux. Illustration of the maximum principle satisfied by the new method, for the initial data (70). The usual moment solution does not satisfy the maximum principle which is a salient property of the exact solution.

References

[1] C. Bernardi and Y. Maday, Polynomial interpolation results in Sobolev space, Journal of Comp. and Applied Math., 43, 53-80, 1992.

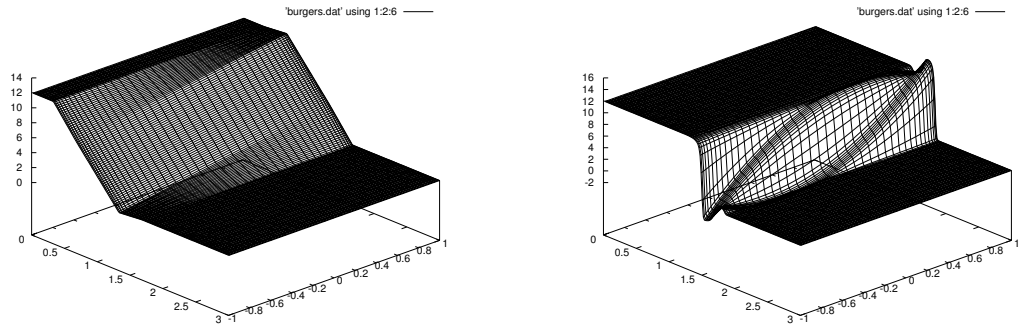
[2] R. Bojanovic and R.A. Devore, On polynomials of best one side approximation, L'enseignement mathématique, 12, 1966.

[3] H. Bijl, D. Lucor, S. Mishra and C. Schwab Editors, Uncertainty Quantification in Computational Fluid Dynamics, Lecture Notes in Computational Science and Engineering, 92, 2010.

[4] Y. Bourgault, D. Broizat and P.-E. Jabin, Convergence rate for the method of moments with linear closure relations, 1-27, Kinetic and Related Models (KRM) Vol. 8, No 1, March 2015.

[5] Y. Brenier, Résolution d'équations d'évolution quasi-linéaires en dimension N d'espace à l'aide d'équations linéaires en dimension $N + 1$, Journal of Differential Equations, 50, 375-390, 1983.

[6] Y. Brenier, L^2 formulation of multidimensional scalar conservation laws. Arch. Ration. Mech. Anal. 193 (2009), no. 1, 1-19.



reconstructed initial data reconstructed solution at $t = 0.1$

Figure 6: Burgers flux. Standard non intrusive method with the Gauss quadrature points $\omega_1 = -\sqrt{\frac{3}{4}}$, $\omega_2 = 0$ and $\omega_3 = \sqrt{\frac{3}{4}}$. This figure illustrates that the maximum principle is not satisfied for the non intrusive method, even for the initial data (70).

[7] C. Canuto and A. Quarteroni, Approximation Results for Orthogonal Polynomials in Sobolev Spaces Math. of Comp., 38, 157, 67-86, 1982.

[8] B. Despres, D. Lucor and G. Poette, Robust Uncertainty Propagation in Systems of Conservation Laws with the Entropy Closure Method, chapter in Uncertainty Quantification in Computational Fluid Dynamics, LNCSE 92, Springer series 2010.

[9] R.A. Devore and G.G. Lorenz, Constructive approximation, Springer, 1981.

[10] D. Gottlieb and D. Xiu, Galerkin Method for Wave Equations with Uncertain Coefficients. Commun. Comp. Phys., 3:505-518, 2008.

[11] J.D. Jakeman, M.S. Eldred and K. Sargsyan, Enhancing l^1 minimization estimates of polynomial chaos expansions using basis selection, Journal of Computational Physics, Volume 289, 15 May 2015, Pages 18-34.

[12] S. Jin, D. Xiu and X. Zhu, Asymptotic-preserving methods for hyperbolic and transport equations with random inputs and diffusive scalings, Journal of Computational Physics 289 (2015) 35-52.

[13] S. Jin, D. Xiu and X. Zhu, A well balanced stochastic galerkin method for scalar conservation laws with random inputs, preprint 2015, online http://www.math.wisc.edu/jin/PS/WBUQ.pdf.

[14] M.-J. Kand and A. Vasseur, Criteria on contraction for entropic discontinuities of the system of conservation laws via weighted relative entropy. Preprint 2015.

[15] G. Lin, C.-H. Su, and G. E. Karniadakis, The Stochastic Piston Problem. PNAS, 101(45):15840-15845, 2004.

[16] P.-L. Lions, B. Perthame and E. Tadmor, A kinetic formulation of multidimensional scalar conservation laws and related equations, J. AMS, 7-1, 169-191, 1994.

[17] G. Meinardus, Approximation of functions: theory and numerical methods, Springer-Verlag, 1967.

[18] B. Perthame, *Kinetic formulation of conservation laws*, volume 21 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[19] B. Perthame and E. Tadmor, A kinetic equation with kinetic entropy functions for scalar conservation laws, Comm. Math. Phys. 136, 501-517, 1991.

[20] Ch. Schwab, S. Mishra and J. Sukys, Multi-level Monte Carlo finite volume methods for non- linear systems of conservation laws in multi-dimensions. Journal of Computational Physics Volume 231, Issue 8, 2012, Pages 3365-3388.

[21] A. Weisse, G. Wellein, A. Alvermann and H. Fehske, The kernel polynomial method, Reviews of Modern Physics, Vol. 78, 2006.

[22] N. Wiener, The Homogeneous Chaos. Amer. J. Math., 60:897-936, 1938.