Rigorous derivation of the macroscopic equations for the lattice Boltzmann method via the corresponding Finite Difference scheme

Thomas Bellotti

Rigorous derivation of the macroscopic equations for the lattice Boltzmann method *via* the corresponding Finite Difference scheme

Thomas Bellotti
(thomas.bellotti@polytechnique.edu)
CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris
91128 Palaiseau Cedex, France.

May 4, 2022

Abstract

Lattice Boltzmann schemes are efficient numerical methods to solve a broad range of problems under the form of conservation laws. However, they suffer from a chronic lack of clear theoretical foundations. In particular, the consistency analysis is still an open issue. We propose a rigorous derivation of the macroscopic equations for any lattice Boltzmann scheme under acoustic scaling. This is done by passing from a kinetic (lattice Boltzmann) to a macroscopic (Finite Difference) point of view at a fully discrete level in order to eliminate the non-conserved moments relaxing away from the equilibrium. We rewrite the lattice Boltzmann scheme as a multi-step Finite Difference scheme on the conserved variables, as introduced in our previous contribution. We then perform the usual consistency analysis for Finite Difference by exploiting its precise characterization using matrices of Finite Difference operators. Though we present the derivation until second-order under acoustic scaling, we provide all the elements to extend it to higher orders and to other scalings, since the kinetic-macroscopic connection is conducted at the fully discrete level. Finally, we show that our strategy yields, in a mathematically rigorous setting, the same results as previous works in the literature.

Keywords: Lattice Boltzmann, Finite Difference, macroscopic equations, consistency, Taylor expansions
2000 MSC: 65M75, 65M06

1 Introduction

Lattice Boltzmann methods form a vast category of numerical schemes to address the approximation of the solution of Partial Differential Equations (PDEs) under the form of conservation laws, called macroscopic equations. These numerical schemes act in a kinetic fashion by employing a certain number $q \in \mathbb{N}^\star$ of discrete velocities, larger than the number $N \in \mathbb{N}^\star$ of macroscopic equations to be solved. The scheme proceeds *via* a kinetic-like algorithm made up of two distinct steps. The first one is a local non-linear collision phase on each site of the mesh, followed by a lattice-constrained transport which is inherently linear. The local nature of the collision phase allows for massive parallelization of the method and the fact that the "particles" are constrained to dwell on the lattice allows to implement the stream phase as a pointer shift in memory. This results in a very efficient numerical method capable of reaching problems of important size in terms of computational and memory cost. The historical seminal papers from the end of the eighties are [33] and [21], while for a general modern presentation of the lattice Boltzmann schemes and their extremely broad fields of application, including hyperbolic systems of conservation laws, the quasi-incompressible Navier-Stokes equations, multi-phase systems and porous media, the interested reader can consult [42], [19] and [30]. The presentation of this plethora of interesting applications is however beyond the scope of our contribution.

To our understanding, the highest price to pay for this highly efficient implementation of the method is the lack of pure theoretical understanding on why the overall procedure works well at approximating the solution of the target macroscopic equations. This is essentially due to the fact that – the standpoint of the lattice Boltzmann schemes being kinetic – the number of discrete velocities is larger than the number of target equations.

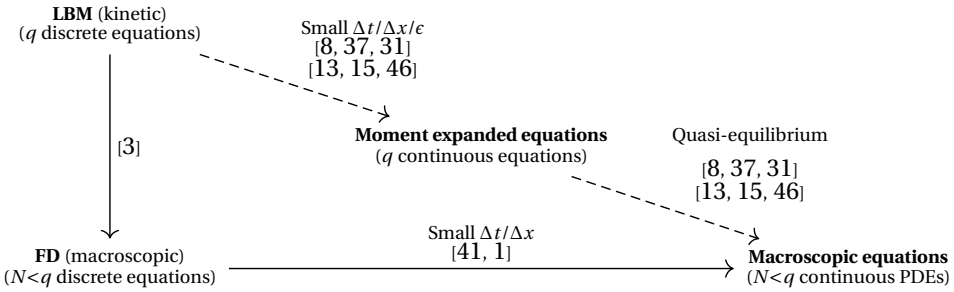

Figure 1: Different paths to recover the macroscopic equations. The formal approaches available in the literature [8, 37, 31, 13, 15, 46] rely on the path marked with dashed arrows. They perform Taylor expansions for small discretization parameters and then utilize the quasi-equilibrium of the non-conserved moments to get rid of them. Our way of proceeding is marked with full arrows: we eliminate exactly the non-conserved moments at the discrete level as in [3] and we perform the usual consistency analysis for Finite Difference schemes as in [41, 1].

Therefore, the formal analyses for lattice Boltzmann schemes available in the literature try to bridge the gap between a kinetic and a macroscopic point of view relying essentially on the quasi-equilibrium of the non-conserved variables. In particular, as far as the consistency with the macroscopic equations is concerned in the limit of small discretization parameters, two main approaches are at our disposal. The first one is based on the Chapman-Enskog expansion [7, 23] from statistical mechanics, shaped to the context of lattice Boltzmann schemes, see for example [8, 37, 31]. The second approach features the so-called equivalent equations introduced by Dubois [13, 15], consisting in performing a Taylor expansion of the scheme both for the conserved and non-conserved moments and progressively re-inject the developments order-by-order. This approach has proved to yield information in accordance with the numerical simulations, see [16, 17, 2]. Despite their proved empirical reliability and the fact that they yield the same results at the dominant orders (see [14] for instance) these two strategies are both formal: the Chapman-Enskog expansion relies on a multi-scale development with no clear mathematical foundation whereas the method of the equivalent equations writes the expansions also on the non-conserved variables, for which no equation under the form of a PDE is known. Other approaches known in the literature are the asymptotic analysis under parabolic scaling deployed in [27, 25, 26] as well as the Maxwell iteration method [46, 47], which shares strong bonds with the equivalent equations method presented before. The previous list of formal analysis techniques does not aim at being exhaustive (the interested reader can refer to [30]) and one should be aware that, despite efforts in this direction [6], there is no consensus on which is the right method to use [30].

A staple of all the previously mentioned approaches is that the expansion for the discretization parameters (time and space steps) tending to zero is performed on the kinetic numerical scheme, where both conserved and non-conserved variables are present. Eventually, the non-conserved variables are formally eliminated from the continuous formulation by scaling arguments, so to speak, using quasi-equilibrium. This corresponds to follow the diagonal path on Figure 1. In this contribution, we develop the other path, namely the top-down movement followed by the left-right one on Figure 1. In particular, in order to fill the hollow between lattice Boltzmann schemes and traditional approaches known to numerical analysts, such as Finite Difference schemes, we recently introduced [3] a formalism to recast any lattice Boltzmann scheme, regardless of its linearity, as a multi-step Finite Difference scheme solely on the conserved moments. It should be stressed that our standpoint, where lattice Boltzmann schemes are studied in terms of their Finite Difference counterpart, must not be seen as the right way of implementing them, because one would lose most of the previously mentioned computational efficiency coming from the kinetic vision. Conversely, our way of writing the scheme should be seen as a sort of one-way mathematical transform to pass from a kinetic standpoint to a macroscopic one in a purely discrete setting. The elimination of the non-conserved moments is carried exactly on the discrete formulation by algebraic devices, thus independently from the time-space scaling. The price to pay for the non-conserved moments relaxing away from the equilibria is the multi-step nature of the Finite Difference scheme. In our previous proposal

[3], it has been crucial to be able to provide, thanks to a systematic mathematical approach, a precise description of the main ingredient needed to reduce the lattice Boltzmann scheme to a Finite Difference scheme, namely the characteristic polynomial of matrices of Finite Difference operators. We are therefore allowed to utilize this characteristic polynomial as a tool satisfying certain properties alone from the particular underlying lattice Boltzmann scheme. Quite the opposite, using the algorithm proposed by [18], one is compelled to explicitly write down the corresponding Finite Difference scheme in order to perform the Taylor expansions to recover the continuous macroscopic equations. In our case, the mathematical understanding that we *a priori* have on the corresponding (macroscopic) Finite Difference schemes, regardless of the (kinetic) lattice Boltzmann scheme it stands for, allows the following theoretical discussion.

The main findings presented in the present paper are the following.

- We propose a procedure to rigorously analyze the consistency of any lattice Boltzmann scheme *via* its corresponding Finite Difference scheme.
- In the case of acoustic scaling between space and time discretization, we rigorously find the expression of the macroscopic PDEs approximated by any scheme until second order.
- Under acoustic scaling, these macroscopic PDEs are the same than the ones obtained by [15] until second order and to the ones from [47] at any order.

The paper is structured as follows. In Section 2, we set notations and assumptions concerning the lattice Boltzmann schemes we shall work with. Section 3 is devoted to recall the main results from our previous work [3] concerning the recast of any lattice Boltzmann scheme as a Finite Difference scheme. These results are then stated in a slightly different manner, facilitating the following analysis. The main result of the work is stated in Section 4 and comes under the form of two theorems, which are eventually proved in Section 5 under the assumption of dealing with one conservation law, for the sake of keeping the presentation and the notations as simple as possible. In Section 6, we indicate how the previous proof is easily extended to several conservation laws, whereas Section 7 is devoted to hint the links with some available approaches to find the macroscopic equations available in the literature. The conclusions and perspectives of this work are drawn in Section 8.

2 Lattice Boltzmann schemes

To start our contribution, we present the classical framework of the multiple-relaxation-times (known as MRT) lattice Boltzmann schemes, as introduced by [11]. For the sake of simplicity, we do not consider source terms which can be effortlessly introduced in the analysis. This fixes the perimeter of the schemes we shall be allowed to treat and study in the sequel.

2.1 Spatial and temporal discretization

We set the problem in spatial dimension $d = 1, 2, 3$ considering the whole space $\mathbb{R}^d$ since this work is not focused on the enforcement of boundary conditions. The space is discretized by a d-dimensional lattice denoted $\mathfrak{L} := \Delta x \mathbb{Z}^d$ with constant step $\Delta x > 0$. The time is uniformly discretized with step $\Delta t > 0$, rendering a time lattice $\mathfrak{Z} := \Delta t \mathbb{N}$. The role of the initial conditions is not investigated and is a subject on its own, see [44, 38]. We introduce the so-called "lattice velocity" $\lambda > 0$ defined by $\lambda := \Delta x / \Delta t$. Observe that in the sequel, namely in Section 4, we shall introduce a particular relation between space step Δx and Δt when $\Delta x \to 0$, in order to provide the main results of the work, as done in [13, 15]. However, until the end of Section 3, the discussion remains valid for any choice of these parameters.

2.2 Discrete velocities

The discrete velocities are an essential ingredient of any lattice Boltzmann scheme. One has to choose $(\boldsymbol{e}_j)_{j=1}^{j=q} \subset \mathbb{R}^d$ with $q \in \mathbb{N}^\star$, discrete velocities, which are multiple of the lattice velocity λ, namely $\boldsymbol{e}_j = \lambda \boldsymbol{c}_j$ for any $j \in [\![1 \,..\, q]\!]$[1] with $(\boldsymbol{c}_j)_{j=1}^{j=q} \subset \mathbb{Z}^d$. Thus, the virtual particles are stuck to the lattice $\mathfrak{L}$ at each time step of the method. We denote the

[1] This shall be a notation to indicate closed intervals of integers, namely for $a, b \in \mathbb{Z}$ with $a \leq b$, then $[\![a \,..\, b]\!] := \{a, a+1, \dots, b\}$.

distribution density of the virtual particles moving with velocity $\boldsymbol{e}_j$ by $f_j = f_j(t,\boldsymbol{x})$ for every $j \in [1\,..\,q]$, depending on the space and time variable.

2.3 Lattice Boltzmann algorithm: collide and stream

As mentioned in the Introduction, any lattice Boltzmann scheme consists in a kinetic algorithm made up of two phases: a local collision phase performed on each site of the lattice $\mathfrak{L}$ and a stream phase where particles are exchanged between different sites of the lattice. Let us independently introduce them.

- *Collision phase.* We adopt the general point of view of the multiple-relaxation-times schemes, where the collision phase is written as a diagonal relaxation towards some equilibria in the moments basis, see [11]. We introduce a change of basis called moment matrix $\boldsymbol{M} \in \mathrm{GL}_q(\mathbb{R})$. Gathering the distributions into $\boldsymbol{f} = (f_1,\dots,f_q)^\mathsf{T}$, the moments are recovered by $\boldsymbol{m} = \boldsymbol{M}\boldsymbol{f}$ and *vice versa*. We also introduce
  - the matrix $\boldsymbol{I} \in \mathrm{GL}_q(\mathbb{R})$, the identity matrix of size q;
  - the matrix $\boldsymbol{S} \in \mathscr{M}_q(\mathbb{R})$, called relaxation matrix. This matrix is diagonal with $N \in [1\,..\,q-1]$ being the number of conserved moments $\boldsymbol{S} = \mathrm{diag}(s_1,\dots,s_N,s_{N+1},\dots,s_q)$, where $s_i \in \mathbb{R}$ for $i \in [1\,..\,N]$ for the conserved moments and $s_i \in]0,2]$ for $i \in [N+1\,..\,q]$, see [13], for the non-conserved ones. Observe that the relaxation parameters corresponding to the conserved moments do not play any role in the lattice Boltzmann algorithm, therefore the matrix $\boldsymbol{S}$ can be singular, without any specific issue. In particular, we shall prove in Section 3.4 that the choice of relaxation parameter for the conserved variables does not have any influence on the outcomes presented in this work. For the sake of presentation, we start numbering the moments by the conserved ones;
  - we employ the notation $\boldsymbol{m}^{\mathrm{eq}}(t,\boldsymbol{x}) = \boldsymbol{m}^{\mathrm{eq}}(m_1(t,\boldsymbol{x}),\dots,m_N(t,\boldsymbol{x}))$ for $t \in \mathfrak{I}$ and $\boldsymbol{x} \in \mathfrak{L}$, where $\boldsymbol{m}^{\mathrm{eq}} : \mathbb{R}^N \to \mathbb{R}^q$ are possibly non-linear functions of the conserved moments. In order to guarantee that the first N moments are conserved through the collision process, the constraints
$$m_i^{\mathrm{eq}}(m_1,\dots,m_N) = m_i, \qquad \forall\, i \in [1\,..\,N], \tag{1}$$
must hold [4].

  Let $t \in \mathfrak{I}$ and $\boldsymbol{x} \in \mathfrak{L}$, the collision phase reads, denoting by $\star$ any post-collision state
$$\boldsymbol{m}^\star(t,\boldsymbol{x}) = (\boldsymbol{I}-\boldsymbol{S})\boldsymbol{m}(t,\boldsymbol{x}) + \boldsymbol{S}\boldsymbol{m}^{\mathrm{eq}}(t,\boldsymbol{x}). \tag{2}$$

- *Stream phase.* The stream phase is diagonal in the space of the distributions and consist in an exact upwind advection of the particle distribution densities. It can be written, for $t \in \mathfrak{I}$ and $\boldsymbol{x} \in \mathfrak{L}$, as
$$f_j(t+\Delta t,\boldsymbol{x}) = f_j^\star(t,\boldsymbol{x}-\boldsymbol{c}_j\Delta x), \tag{3}$$
for any $j \in [1\,..\,q]$.

3 Finite Difference formulation of a lattice Boltzmann scheme

Having defined the lattice Boltzmann schemes, we briefly introduce the setting allowing us to rewrite any lattice Boltzmann scheme (kinetic) as a multi-step Finite Difference scheme (macroscopic) on the N conserved moments of interest. The interested reader can refer to our previous contribution [3] for more details. Then, the formulation of the multi-step Finite Difference scheme is given using a more compact notation which is more suitable to the following discussion. We start with the assumptions needed in the sequel.

Assumptions 3.1 (Finite Difference assumptions)**.** The entries of $\boldsymbol{M}$ and $\boldsymbol{S}$ can depend on Δx and/or on Δt but cannot be a function of the space and time variables.

3.1 Algebraic setting

Let us first introduce the necessary algebraic setting. In particular, we define the shift operators associated with each discrete velocity as well as the derived Finite Difference operators in space. In the following Definition, the time variable does not play any role since kept frozen, thus it is not listed for the sake of readability.

Definition 3.2 (Shift and Finite Difference operators in space)**.** Let $\boldsymbol{z} \in \mathbb{Z}^d$, then the associated shift operator on the lattice $\mathfrak{L}$, denoted $\mathsf{t}_{\boldsymbol{z}}$, is defined in the following way. Take $f : \mathfrak{L} \to \mathbb{R}$ be any function defined on the lattice, then the action of $\mathsf{t}_{\boldsymbol{z}}$ is

$$(\mathsf{t}_{\boldsymbol{z}} f)(\boldsymbol{x}) = f(\boldsymbol{x} - \boldsymbol{z}\Delta x), \qquad \forall \boldsymbol{x} \in \mathfrak{L}.$$

We also introduce $\mathsf{T} := \{\mathsf{t}_{\boldsymbol{z}} \quad \text{with} \quad \boldsymbol{z} \in \mathbb{Z}^d\} \cong \mathbb{Z}^d$. The product $\circ : \mathsf{T} \times \mathsf{T} \to \mathsf{T}$ of two shift operators is defined by

$$\mathsf{t}_{\boldsymbol{z}} \circ \mathsf{t}_{\boldsymbol{w}} := \mathsf{t}_{\boldsymbol{z}+\boldsymbol{w}}, \qquad \forall \boldsymbol{z}, \boldsymbol{w} \in \mathbb{Z}^d.$$

The set of Finite Difference operators on the lattice $\mathfrak{L}$ is defined as

$$\mathsf{D} := \mathbb{R}\mathsf{T} = \left\{ \sum_{\mathsf{t}\in\mathsf{T}} \alpha_{\mathsf{t}}\mathsf{t}, \quad \text{where} \quad \alpha_{\mathsf{t}} \in \mathbb{R} \quad \text{and} \quad \alpha_{\mathsf{t}} = 0 \quad \text{almost everywhere} \right\}, \tag{4}$$

the group ring (or group algebra) of T over $\mathbb{R}$. The sum $+ : \mathsf{D} \times \mathsf{D} \to \mathsf{D}$ and the product $\circ : \mathsf{D} \times \mathsf{D} \to \mathsf{D}$ of two elements are defined by

$$\left(\sum_{\mathsf{t}\in\mathsf{T}} \alpha_{\mathsf{t}}\mathsf{t}\right) + \left(\sum_{\mathsf{t}\in\mathsf{T}} \beta_{\mathsf{t}}\mathsf{t}\right) = \sum_{\mathsf{t}\in\mathsf{T}} (\alpha_{\mathsf{t}} + \beta_{\mathsf{t}})\mathsf{t}, \qquad \left(\sum_{\mathsf{t}\in\mathsf{T}} \alpha_{\mathsf{t}}\mathsf{t}\right) \circ \left(\sum_{\mathsf{h}\in\mathsf{T}} \beta_{\mathsf{h}}\mathsf{h}\right) = \sum_{\mathsf{t},\mathsf{h}\in\mathsf{T}} (\alpha_{\mathsf{t}}\beta_{\mathsf{h}})(\mathsf{t}\circ\mathsf{h}).$$

Furthermore, the product of $\sigma \in \mathbb{R}$ with elements of D is given by

$$\sigma \left(\sum_{\mathsf{t}\in\mathsf{T}} \alpha_{\mathsf{t}}\mathsf{t}\right) = \sum_{\mathsf{t}\in\mathsf{T}} (\sigma\alpha_{\mathsf{t}})\mathsf{t}.$$

In the sequel, the products $\circ$ are generally understood.

Remark 3.3. We could achieve exactly the same construction, following Chapter 2 in [9], by considering functions on the lattice $\mathfrak{L}$ as sequences and the Finite Difference operators as sequences with compact support (whence the almost everywhere requirement in Equation (4)). Then, the product $\circ$ can be seen as a convolution (Cauchy) product between compactly supported sequences and the action of a Finite Difference operator on a function as the convolution of a finitely supported sequence with a generic sequence.

Upon introducing the generating displacements along each axis $\mathsf{x}_k := \mathsf{t}_{\boldsymbol{e}_k}$ where $\boldsymbol{e}_k$ is the k-th vector of the canonical basis, for any $k \in [\![1..d]\!]$, we can isomorphically identify $\mathsf{D} \cong \mathbb{R}[\mathsf{x}_1, \mathsf{x}_1^{-1}, \ldots, \mathsf{x}_d, \mathsf{x}_d^{-1}]$, the ring of multivariate Laurent polynomials. On the other hand, the real numbers $\mathbb{R}$ can be viewed as sub-ring of D, being the constant polynomials. This identification can be somehow interpreted as the historical starting point of umbral calculus [39], also known as calculus of Finite Differences [34]: allow to interchange indices in sequences (operators or functions) with exponents (in polynomials). The stream phase Equation (3) can be recast under its non-diagonal form in the space of moments [46, 15] by introducing what we call the moments-stream matrix $\mathbf{T} := \boldsymbol{M}\mathrm{diag}(\mathsf{t}_{\boldsymbol{c}_1}, \ldots, \mathsf{t}_{\boldsymbol{c}_q})\boldsymbol{M}^{-1} \in \mathscr{M}_q(\mathsf{D})$ and merged with the collision phase Equation (2) to obtain the scheme, for any $t \in \mathfrak{Z}$ and for any $\boldsymbol{x} \in \mathfrak{L}$

$$\boldsymbol{m}(t + \Delta t, \boldsymbol{x}) = \mathbf{A}\boldsymbol{m}(t, \boldsymbol{x}) + \mathbf{B}\boldsymbol{m}^{\mathrm{eq}}(t, \boldsymbol{x}), \tag{5}$$

where $\mathbf{A} := \mathbf{T}(\boldsymbol{I} - \boldsymbol{S}) \in \mathscr{M}_q(\mathsf{D})$ and $\mathbf{B} := \mathbf{T}\boldsymbol{S} \in \mathscr{M}_q(\mathsf{D})$.

3.2 Corresponding Finite Difference scheme

With this new compact algebraic form of any lattice Boltzmann scheme, namely Equation (5), we are able to recall the main results proved in [3]. These results encompass the findings from [43], [10] and [18]. The version for one conserved moment can be formulated as follows.

Proposition 3.4 (Corresponding Finite Difference scheme for $N = 1$, [3])**.** *Consider $N = 1$. Then the lattice Boltzmann scheme given by Equation* (5) *corresponds to a multi-step explicit macroscopic Finite Difference scheme on the conserved moment m_1 under the form*

$$m_1(t+\Delta t,\boldsymbol{x}) = -\sum_{k=0}^{q-1} \mathsf{c}_k m_1(t+(1-q+k)\Delta t,\boldsymbol{x}) + \left(\sum_{k=0}^{q-1}\left(\sum_{\ell=0}^{k} \mathsf{c}_{q+\ell-k}\mathbf{A}^\ell\right)\mathbf{B}\boldsymbol{m}^{eq}(t-k\Delta t,\boldsymbol{x})\right)_1, \tag{6}$$

for all $t \in \mathfrak{Z}$ and for all $\boldsymbol{x} \in \mathfrak{L}$, where $(\mathsf{c}_k)_{k=0}^{k=q} \subset \mathsf{D}$ are the coefficients of $\chi_{\mathbf{A}} := det(X\boldsymbol{I} - \mathbf{A}) = \sum_{k=0}^{k=q} \mathsf{c}_k X^k$, the characteristic polynomial of $\mathbf{A}$, with $det(\cdot)$ indicating the determinant of a matrix.

The proof – given in [3] – relies on the fact that D is a commutative ring and that therefore the Cayley-Hamilton theorem [5], stipulating that any square matrix with entries in a commutative ring annihilates its characteristic polynomial, holds.

This result is easily generalized to the case of multiple conservation laws, namely $N > 1$. For this, let us introduce a new notation. For any square matrix $\boldsymbol{C} \in \mathscr{M}_q(\mathfrak{R})$ on a commutative ring $\mathfrak{R}$, consider $\boldsymbol{C}_I := (\sum_{i\in I} \boldsymbol{e}_i \otimes \boldsymbol{e}_i)\boldsymbol{C}(\sum_{i\in I} \boldsymbol{e}_i \otimes \boldsymbol{e}_i) \in \mathscr{M}_q(\mathfrak{R})$ for any $I \subset [1\,..\,q]$, corresponding to the matrix where only the entries with row and column indices in I are kept and the remaining ones are set to zero. Then we have the following statement.

Proposition 3.5 (Corresponding Finite Difference scheme for $N \geq 1$, [3])**.** *Consider $N \geq 1$. Then the lattice Boltzmann scheme given by Equation* (5) *corresponds to a family of multi-step explicit macroscopic Finite Difference schemes on the conserved moments $m_1, \ldots, m_N$. This is, for any $i \in [1\,..\,N]$*

$$\begin{aligned} m_i(t+\Delta t,\boldsymbol{x}) = -\sum_{k=0}^{q-N} \mathsf{c}_{i,k} m_i(t+(k-q+N)\Delta t,\boldsymbol{x}) &+ \left(\sum_{k=0}^{q-N}\left(\sum_{\ell=0}^{k} \mathsf{c}_{i,q+1-N+\ell-k}\mathbf{A}_i^\ell\right)\mathbf{A}_i^\diamond \boldsymbol{m}(t-k\Delta t,\boldsymbol{x})\right)_i \\ &+ \left(\sum_{k=0}^{q-N}\left(\sum_{\ell=0}^{k} \mathsf{c}_{i,q+1-N+\ell-k}\mathbf{A}_i^\ell\right)\mathbf{B}\boldsymbol{m}^{eq}(t-k\Delta t,\boldsymbol{x})\right)_i, \end{aligned} \tag{7}$$

for all $t \in \mathfrak{Z}$ and $\boldsymbol{x} \in \mathfrak{L}$, where $\mathbf{A}_i := \mathbf{A}_{\{i\}\cup[N+1\,..\,q]}$ and $\mathbf{A}_i^\diamond := \mathbf{A} - \mathbf{A}_i$ with $(\mathsf{c}_{i,k})_{k=0}^{k=q+1-N} \subset \mathsf{D}$ which are the coefficients of $\chi_{\mathbf{A}_i} := det(X\boldsymbol{I} - \mathbf{A}_i) = X^{N-1}\sum_{k=0}^{k=q+1-N} \mathsf{c}_{i,k}X^k$, the characteristic polynomial of $\mathbf{A}_i$.

This result is the natural generalization of Proposition 3.4 to the case $N > 1$, in the sense that each sub-problem for any $i \in [1\,..\,N]$ deals with one conserved moment (the i-th) at each time, only trying to eliminate the non-conserved moments while keeping the conserved ones other than the i-th. This is achieved by using a tailored characteristic polynomial for each conserved moment in the problem.

Further comments on Proposition 3.4 and Proposition 3.5 are postponed to the following Section.

3.3 A more compact form of corresponding Finite Difference scheme

Although the asymptotic analysis we shall develop in Section 5 can be carried on the formulations from Proposition 3.4 and Proposition 3.5 previously introduced in [3], we propose a different formalism based on shift operators in time. Having utilized both approaches, the advantage of this new standpoint – which shall be adopted in this paper – is to easily deal with the asymptotic analysis of the coefficients of the characteristic polynomial and of the powers of the matrix $\mathbf{A}$ on the right hand side of Equation (6). In particular, this allows for the straightforward generalization of the procedure above second-order. Furthermore, the links with other asymptotic analysis of lattice Boltzmann schemes from the literature – which we shall develop in Section 7 – become noticeably more transparent. To this end, we introduce the following Definition.

Definition 3.6 (Shift operator in time)**.** Let $f : \mathfrak{Z} \to \mathbb{R}$ be any function defined on the time lattice, then the time shift operator z acts as

$$(\mathsf{z}f)(t) = f(t+\Delta t), \qquad \forall\, t \in \mathfrak{Z}.$$

With this, the scheme Equation (5) can be recast under the fully-operatorial form. For any $t \in \mathfrak{Z}$ and for any $\boldsymbol{x} \in \mathfrak{L}$

$$(\mathsf{z}\boldsymbol{I} - \mathbf{A})\boldsymbol{m}(t,\boldsymbol{x}) = \mathbf{B}\boldsymbol{m}^{\mathrm{eq}}(t,\boldsymbol{x}), \tag{8}$$

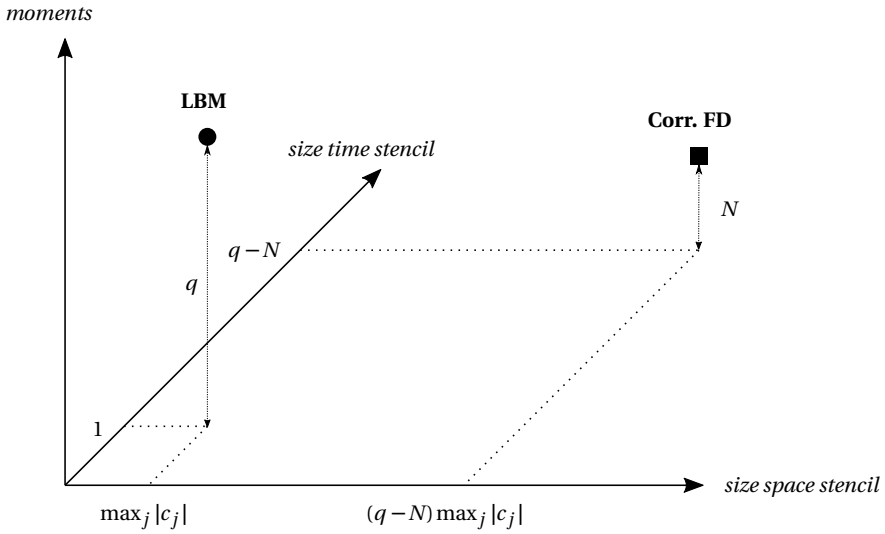

Figure 2: Comparison between lattice Boltzmann scheme (circle) and corresponding Finite Difference schemes (square) in terms of involved moments (respectively q and N), number of time steps (respectively 1 and $q-N$) and size of the maximal spatial stencil (respectively $\max_j |c_j|$ and $(q-N)\max_j |c_j|$).

which corresponds to taking the Z-transform [28] of the scheme in the variable z. Here, the inverse of the resolvent associated with $\mathbf{A}$, namely $\mathsf{z}\boldsymbol{I}-\mathbf{A} \in \mathscr{M}_q(\mathbb{R}[\mathsf{z}] \otimes_{\mathbb{R}} \mathsf{D})$, where $\mathbb{R}[\mathsf{z}] \otimes_{\mathbb{R}} \mathsf{D} \cong \mathbb{R}[\mathsf{z}, \mathsf{x}_1, \mathsf{x}_1^{-1}, \ldots, \mathsf{x}_d, \mathsf{x}_d^{-1}]$, with $\otimes_{\mathbb{R}}$ that indicates the tensor product of $\mathbb{R}$-algebras (see Chapter 16 in [32] or Chapter 2 in [29]), forms a commutative ring. In the sequel, we shall drop the time and the space variables when not strictly needed for the sake of readability, because the system given by Equation (8) is intrinsically time and space invariant thanks to Assumptions 3.1 and since we work on an unbounded domain, without considering the initial conditions.

Proposition 3.4 bis (Corresponding Finite Difference scheme for $N = 1$)**.** *Consider $N = 1$. Then the lattice Boltzmann scheme given by Equation* (5) *or Equation* (8) *corresponds to a multi-step explicit macroscopic Finite Difference scheme on the conserved moment m_1 under the form*

$$det(\mathsf{z}\boldsymbol{I}-\mathbf{A})\, m_1 = \left(adj(\mathsf{z}\boldsymbol{I}-\mathbf{A})\mathbf{B}\boldsymbol{m}^{eq}\right)_1, \tag{9}$$

where adj$(\cdot)$ indicates the adjugate matrix,[2] *also known as classical adjoint, which is the transpose of the cofactor matrix [22].*

Up to a temporal shift of the whole scheme, the corresponding multi-step explicit Finite Difference scheme by Equation (9) *equals the one from Equation* (6)*.*

Proof. The proof can be done starting from Proposition 3.4. Alternatively using the fundamental relation between adjugate and determinant, see Chapter 0 in [22], which is a consequence of the Laplace formula, we have that for any $\boldsymbol{C} \in \mathscr{M}_q(\mathfrak{R})$ where $\mathfrak{R}$ is any commutative ring

$$\boldsymbol{C}\mathrm{adj}(\boldsymbol{C}) = \mathrm{adj}(\boldsymbol{C})\boldsymbol{C} = \det(\boldsymbol{C})\boldsymbol{I}. \tag{10}$$

Hence, multiplying Equation (8) by $\mathrm{adj}(\mathsf{z}\boldsymbol{I}-\mathbf{A})$ yields $\det(\mathsf{z}\boldsymbol{I}-\mathbf{A})\boldsymbol{m} = \mathrm{adj}(\mathsf{z}\boldsymbol{I}-\mathbf{A})\mathbf{B}\boldsymbol{m}^{\mathrm{eq}}$. Selecting the first row gives Equation (9). □

Remark 3.7 (From kinetic to macroscopic)**.** We observe the following facts:

- The procedure can be reversed – when keeping all the lines in $\det(\mathsf{z}\boldsymbol{I}-\mathbf{A})\boldsymbol{m} = \mathrm{adj}(\mathsf{z}\boldsymbol{I}-\mathbf{A})\mathbf{B}\boldsymbol{m}^{\mathrm{eq}}$ – using a multiplication by $\mathsf{z}\boldsymbol{I}-\mathbf{A}$ and then dividing by the polynomial $\chi_{\mathbf{A}}(\mathsf{z}) = \det(\mathsf{z}\boldsymbol{I}-\mathbf{A})$. In this way, one comes back

[2] It is worthwhile observing that the determinant and the adjugate matrix are defined for any square matrix with elements in a commutative ring.

to the lattice Boltzmann scheme by Equation (8). This can be done as long as one does not select and store only the first row as in Equation (9). Contrarily, if this selection is performed, the irreversible passage from the kinetic to the macroscopic formulation is accomplished. The non-conserved moments $m_2, \dots, m_q$ are no longer defined and they cannot be recovered from Equation (9). This fact has been observed by [10]: the same macroscopic Finite Difference scheme can correspond to distinct lattice Boltzmann schemes which can have different evolution equations for the non-conserved moments $m_2, \dots, m_q$. This is not surprising, since for a given monic polynomial, one can find an infinite number of matrices of which is it the characteristic polynomial.

- Though – as previously emphasized – the non-conserved moments are no longer present in the macroscopic Finite Difference scheme by Equation (9), there is a residual shadow of their presence, namely the multi-step nature of the Finite Difference scheme, see Figure 2. In particular, each non-conserved moment m_i relaxing away from the equilibrium, namely with $s_i \neq 1$, for $i \in [\![2 .. q]\!]$, adds a time step to the corresponding Finite Difference scheme solely acting on the conserved moment m_1.

Remark 3.8 (Adjugate and characteristic polynomial)**.** A time shift and a change of variable in Equation (6) allows to express $\mathrm{adj}(\mathsf{z}\boldsymbol{I} - \mathbf{A})$ as a polynomial in z of degree $q-1$ computed from the characteristic polynomial. This relation is indeed classical and reads

$$\mathrm{adj}(\mathsf{z}\boldsymbol{I} - \mathbf{A}) = \sum_{k=0}^{q-1} \left(\sum_{\ell=0}^{q-1-k} \mathsf{c}_{k+\ell+1} \mathbf{A}^\ell \right) \mathsf{z}^k, \qquad \text{where} \qquad \det(\mathsf{z}\boldsymbol{I} - \mathbf{A}) = \sum_{k=0}^{q} \mathsf{c}_k \mathsf{z}^k.$$

In the same way, we can restate Proposition 3.5 using the new formalism.

Proposition 3.5 bis (Corresponding Finite Difference scheme for $N \geq 1$)**.** *Consider $N \geq 1$. Then the lattice Boltzmann scheme given by Equation* (5) *or Equation* (8) *corresponds to a family of multi-step explicit macroscopic Finite Difference schemes on the conserved moments $m_1, \dots, m_N$. This is, for any $i \in [\![1 .. N]\!]$*

$$det(\mathsf{z}\boldsymbol{I} - \mathbf{A}_i)\, m_i = \left(adj(\mathsf{z}\boldsymbol{I} - \mathbf{A}_i) \mathbf{A}_i^{\diamond} \boldsymbol{m} \right)_i + \left(adj(\mathsf{z}\boldsymbol{I} - \mathbf{A}_i) \mathbf{B} \boldsymbol{m}^{eq} \right)_i . \tag{11}$$

Up to a temporal shift of the whole scheme, the corresponding multi-step explicit Finite Difference scheme by Equation (11) *equals the one from Equation* (7).

We could call the form of Finite Difference scheme from Proposition 3.5 and Proposition 3.5 bis "canonical" since we shall prove in Section 3.4 that it guarantees that the Finite Difference does not depend on the choice of relaxation parameters for the conserved variables, which do not play any role in the original lattice Boltzmann scheme either, as previously discussed.

Remark 3.9 (Lack of scaling assumption)**.** The results in Proposition 3.4, Proposition 3.5, Proposition 3.4 bis, Proposition 3.5 bis are fully discrete and do not make any assumption on the particular scaling between the time-step Δt and the space-step Δx. In particular, they can be employed both in the case of acoustic scaling, namely $\Delta t \sim \Delta x$, and in the case of diffusive scaling $\Delta t \sim \Delta x^2$.

The previous Remark signifies that the corresponding Finite Difference schemes can be utilized to assess the consistency of the underlying lattice Boltzmann scheme with respect to the macroscopic equations regardless of the particular scaling between time and space discretizations.

3.4 On the choice of relaxation parameters for the conserved moments

In Section 2, we have observed that the choice of relaxation parameters for the conserved moments, namely $s_1, \dots, s_N$, does not change the lattice Boltzmann scheme Equation (2). On the other hand, it could be argued that different choices for $s_1, \dots, s_N$ can affect the formulations of the corresponding Finite Difference schemes resulting from Proposition 3.4 bis and Proposition 3.5 bis. We now show that, as one could hope, this is not the case for the Finite Difference schemes given by Proposition 3.5 bis.

Proposition 3.10. *The multi-step explicit macroscopic Finite Difference schemes given by Equation* (11) *in Proposition 3.5 bis do not depend on the choice of $s_1, \dots, s_N$, the relaxation parameters of the conserved moments.*

Proof. Fix the indices of the conserved moment $i \in [1\,..\,N]$. Let us decompose $\mathbf{B}$, the part of the lattice Boltzmann scheme dealing with the equilibria, as follows: $\mathbf{B} = \mathbf{b}_i \otimes \boldsymbol{e}_i + \mathbf{B}|_{s_i=0}$ where $\mathbf{b}_i = \mathbf{B}_{\cdot,i}$ is the i-th column of $\mathbf{B}$. The dependency of $\mathbf{B}$ on the choice of s_i is now fully contained in $\mathbf{b}_i$. On the other hand $\mathbf{B}|_{s_i=0}$ does not depend on it. The Finite Difference scheme from Proposition 3.5 bis can be therefore recast, upon rearranging and using well-known properties of the external product $\otimes$, as

$$\left(\det(\mathrm{z}\boldsymbol{I} - \mathbf{A}_i) - \boldsymbol{e}_i^{\mathsf{T}} \operatorname{adj}(\mathrm{z}\boldsymbol{I} - \mathbf{A}_i)\mathbf{b}_i\right) m_i = \left(\operatorname{adj}(\mathrm{z}\boldsymbol{I} - \mathbf{A}_i)\mathbf{A}_i^{\diamond}\boldsymbol{m}\right)_i + \left(\operatorname{adj}(\mathrm{z}\boldsymbol{I} - \mathbf{A}_i)\mathbf{B}|_{s_i=0}\boldsymbol{m}^{\mathrm{eq}}\right)_i. \tag{12}$$

The left hand side does not depend on s_j for $j \in [1\,..\,N] \setminus \{i\}$ by construction of $\mathbf{A}_i$ and $\mathbf{b}_i$. On the other hand, the right hand side does not depend on s_j for $j \in [1\,..\,N] \setminus \{i\}$, because $(\mathbf{A}_i^{\diamond})_{\cdot,j} + (\mathbf{B}|_{s_i=0})_{\cdot,j} = (\mathbf{A}_i^{\diamond}|_{s_j=0})_{\cdot,j}$, where we have used Equation (2) and Equation (1). We are left to discuss the possible dependency of Equation (12) on s_i. For the left hand side, we need the following result concerning the determinant of matrices under rank-one updates, whose proof is analogous to that in [12].

Lemma 3.11 (Matrix determinant)**.** *Let $\mathfrak{R}$ be a commutative ring, $\boldsymbol{C} \in \mathscr{M}_q(\mathfrak{R})$ and $\boldsymbol{u}, \boldsymbol{v} \in \mathfrak{R}^q$, then $det(\boldsymbol{C} + \boldsymbol{u} \otimes \boldsymbol{v}) = det(\boldsymbol{C}) + \boldsymbol{v}^{\mathsf{T}} adj(\boldsymbol{C})\boldsymbol{u}$.*

By this Lemma, we deduce that Equation (12) now reads

$$\det(\mathrm{z}\boldsymbol{I} - (\mathbf{A}_i + \mathbf{b}_i \otimes \boldsymbol{e}_i))m_i = \left(\operatorname{adj}(\mathrm{z}\boldsymbol{I} - \mathbf{A}_i)\mathbf{A}_i^{\diamond}\boldsymbol{m}\right)_i + \left(\operatorname{adj}(\mathrm{z}\boldsymbol{I} - \mathbf{A}_i)\mathbf{B}|_{s_i=0}\boldsymbol{m}^{\mathrm{eq}}\right)_i. \tag{13}$$

Observe that $\mathbf{A}_i + \mathbf{b}_i \otimes \boldsymbol{e}_i = \mathbf{A}_i|_{s_i=0}$, thus the left hand side of Equation (13) does not not depend on s_i. The right hand side of Equation (13) is independent of s_i because $\mathbf{A}_i^{\diamond}$ does not depend on it and since the i-th row of $\operatorname{adj}(\mathrm{z}\boldsymbol{I} - \mathbf{A}_i)$ – the transpose of the cofactor matrix of $\mathrm{z}\boldsymbol{I} - \mathbf{A}_i$ – cannot depend on s_i, because only the i-th column of $\mathrm{z}\boldsymbol{I} - \mathbf{A}_i$ depends on s_i. This concludes the proof. □

We have thus shown that the Finite Difference schemes from Proposition 3.5 bis do not depend on the choice of relaxation parameters for the conserved moments and so that we are allowed to take them equal to zero or any other value of specific convenience without loss of generality. In particular, the choice of taking $s_i = 0$ for $i \in [1\,..\,N]$ offers interesting simplifications in the computations to come in Section 5, in a way that shall be clearer by looking at the details. Moreover, this choice has the advantage of showing which moments are conserved at a glance.

4 Main results

Everything is in place to start the standard consistency analysis of Finite Difference schemes [45, 41, 1]. We start from the assumptions allowing us to identify each term once developing in formal power series of Δx, *i.e.* performing Taylor expansions.

Assumptions 4.1 (Scaling assumptions)**.** Assume that:

1. We utilize the acoustic scaling[3] $\lambda > 0$ is a fixed real number as $\Delta x \to 0$, like in [13, 15, 46]. Thus, the only discretization parameter we shall consider in the sequel is Δx.
2. The change of basis $\boldsymbol{M}$, the relaxation matrix $\boldsymbol{S}$ and the moments at equilibrium $\boldsymbol{m}^{\mathrm{eq}}$ are fixed as $\Delta x \to 0$.

We also introduce the spaces of differential operators which shall be obtained by taking the limit $\Delta x \to 0$ as well as other tightly associated concepts.

Definition 4.2 (Time-space differential operators)**.** We define.

- The commutative ring of time-space differential operators $\mathscr{D}$ by

$$\mathscr{D} := \mathbb{R}[\partial_t] \otimes_{\mathbb{R}} \mathbb{R}[\partial_{x_1}, \ldots, \partial_{x_d}] \cong \mathbb{R}[\partial_t, \partial_{x_1}, \ldots, \partial_{x_d}].$$

- Under the acoustic Assumptions 4.1, we consider the commutative ring of formal power series [36, 35]

$$\mathscr{S} := \mathscr{D}[\![\Delta x]\!].$$

[3] Frequently, $\lambda = 1$ is considered in the literature.

- For any $\delta = \sum_{r=0}^{r=+\infty} \Delta x^r \delta^{(r)} \in \mathscr{S}$, we indicate $\delta = O(\Delta x^{r_\circ})$ for some $r_\circ \in \mathbb{N}$ if $\delta^{(r)} = 0$ for $r \in [0\,..\,r_\circ - 1]$ and $\delta^{(r_\circ)} \neq 0$. The integer $r_\circ$ is called "order" of the formal power series δ, see Chapter 1 in [39].

- Finally, let $\mathsf{d} \in \mathbb{R}[\mathsf{z}] \otimes_{\mathbb{R}} \mathsf{D}$ and $\delta \in \mathscr{S}$, then we indicate $\mathsf{d} \asymp \delta$, called "asymptotic equivalence" of d and δ, if for any smooth function of the time and space variables $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, we have

$$(\mathsf{d} f)(t, \boldsymbol{x}) = \sum_{r=0}^{+\infty} \Delta x^r (\delta^{(r)} f)(t, \boldsymbol{x}), \qquad \forall (t, \boldsymbol{x}) \in \mathbb{R} \times \mathbb{R}^d, \qquad \text{as} \qquad \Delta x \to 0.$$

The previous $O(\cdot)$ notation and the notion of asymptotic equivalence are effortlessly extended to vectors and matrices in an entry-wise fashion. It shall be common and harmless not to distinguish between $\mathscr{M}_q(\mathscr{S})$ and $(\mathscr{M}_q(\mathscr{D}))[\![\Delta x]\!]$.

The momentum-velocity operator matrix $\boldsymbol{\mathcal{G}} \in \mathscr{M}_q(\mathscr{D})$, introduced by [15] with slightly different notations, is defined as follows. It is indeed closely linked to the moment-stream matrix $\mathbf{T} \in \mathscr{M}_q(\mathsf{D})$ that we have previously introduced.

Definition 4.3 (Momentum-velocity operator matrix)**.** The momentum-velocity operator matrix made up of first-order differential operators in space is given by

$$\boldsymbol{\mathcal{G}} := \boldsymbol{M} \left(\sum_{|\boldsymbol{\nu}|=1} \operatorname{diag}\left(\boldsymbol{e}_{\mathbf{1}}^{\boldsymbol{\nu}}, \ldots, \boldsymbol{e}_{\boldsymbol{q}}^{\boldsymbol{\nu}} \right) \partial^{\boldsymbol{\nu}} \right) \boldsymbol{M}^{-1} \in \mathscr{M}_q(\mathscr{D}),$$

where the multi-index notation is employed.

This momentum-velocity operator matrix can be partitioned in four blocks with different meanings according to the different nature (conserved or not) of the corresponding moments, as for Equation (8) in [15]. We are now ready to state and then prove the main results of the present contribution. The Taylor expansions are applied to the solution of the corresponding Finite Difference schemes given by Proposition 3.5 or Proposition 3.5 bis, where non-conserved moments have been removed yielding purely macroscopic discrete equations.

Theorem 4.4 (First order expansion)**.** *Under Assumptions 4.1 and in the limit $\Delta x \to 0$, the conserved moments $m_1, \ldots, m_N$, solution of the corresponding macroscopic Finite Difference schemes given by Proposition 3.5 or Proposition 3.5 bis, asymptotically satisfy the following system of macroscopic PDEs*

$$\partial_t m_i + \gamma_{1,i} = O(\Delta x), \qquad \textit{with} \qquad \gamma_{1,i} = \gamma_{1,i}(m_1, \ldots, m_N) := \sum_{j=1}^{N} \boldsymbol{\mathcal{G}}_{ij} m_j + \sum_{j=N+1}^{q} \boldsymbol{\mathcal{G}}_{ij} m_j^{eq}(m_1, \ldots, m_N),$$

for $i \in [1\,..\,N]$ and for $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times \mathbb{R}$, where $m_1, \ldots, m_N : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ are smooth functions.

The first term in $\gamma_{1,i}$ represents the derivatives of fluxes of the conserved variables, which are necessarily linear, while the second one represents the derivatives of the fluxes given by the equilibria of the non-conserved moments, which can be non-linear. These limit equations can be refined with a second-order expansion yielding diffusion terms, where the so-called Hénon's parameters [20] of type $1/s_j - 1/2$ appear.

Theorem 4.5 (Second order expansion)**.** *Under Assumptions 4.1 and in the limit $\Delta x \to 0$, the conserved moments $m_1, \ldots, m_N$, solution of the corresponding macroscopic Finite Difference schemes given by Proposition 3.5 or Proposition 3.5 bis, asymptotically satisfy the following system of macroscopic PDEs*

$$\partial_t m_i + \gamma_{1,i} + \frac{\Delta x}{\lambda} \sum_{j=N+1}^{q} \left(\frac{1}{s_j} - \frac{1}{2} \right) \boldsymbol{\mathcal{G}}_{ij} \left(\sum_{\ell=1}^{N} \frac{dm_j^{eq}}{dm_\ell} \gamma_{1,\ell} - \sum_{\ell=1}^{N} \boldsymbol{\mathcal{G}}_{j\ell} m_\ell - \sum_{\ell=N+1}^{q} \boldsymbol{\mathcal{G}}_{j\ell} m_\ell^{eq} \right) = O(\Delta x^2),$$

for $i \in [1\,..\,N]$ and for $(t, \boldsymbol{x}) \in \mathbb{R}^+ \times \mathbb{R}$, where $m_1, \ldots, m_N : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ are smooth functions.

One can notice that the diffusive terms appear to be proportional to Δx. This is not surprising, since the only way of having a stable explicit Finite Difference scheme to simulate the heat equation under the acoustic scaling is to consider a diffusion coefficient proportional to Δx, in order to constrain the speed of propagation of information to remain finite in the limit $\Delta x \to 0$, see for instance Theorem 6.3.1 in [41].

In this contribution, we have deliberately neglected the behavior of the schemes close to the initial time $t = 0$. It is dictated by the choice of initial datum for the non conserved moments, which is not unique for lattice Boltzmann schemes since $q > N$ but one only knows the N conserved moments at $t = 0$, being the initial datum of the macroscopic PDEs to be solved. The interested reader can consult [44, 38] for more information on this topic.

Let us sketch the main ideas of the proofs of Theorem 4.4 and Theorem 4.5:

- The result of Proposition 3.5 has allowed to eliminate the non-conserved moments from the discrete scheme, thus has completed the step represented by a vertical arrow in Figure 1. Contrarily to the existing approaches, we do not need (and we cannot, see Remark 3.7) to estimate the Taylor expansions of the non-conserved moments.

- We benefit from the clever formulation from Proposition 3.5 bis instead of that of Proposition 3.5. Indeed, considering $\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i \asymp \mathsf{z}\boldsymbol{I} - \mathbf{A}_i$, we are allowed to write, for every $i \in [1\,..\,N]$

$$\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)\, m_i = \left(\operatorname{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)\boldsymbol{\mathcal{A}}_i^{\diamond}\boldsymbol{m}\right)_i + \left(\operatorname{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)\boldsymbol{\mathcal{B}}\boldsymbol{m}^{\mathrm{eq}}\right)_i,$$

  obtained by replacing matrices with entries in the ring $\mathbb{R}[\mathsf{z}] \otimes_{\mathbb{R}} \mathsf{D}$ of discrete operators by their asymptotic equivalents in the ring $\mathscr{S}$. Here, for example, $\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i) \in \mathscr{S}$, and the expression perfectly makes sense because the determinant and the adjugate are well-defined polynomial functions of any square matrix on a commutative ring, like $\mathscr{S}$. Since the determinant and the adjugate are non-linear functions and thus mix different orders in the expansion $\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i$, if we want to recover a closed-form result at a given order of accuracy, we are compelled to utilize the Taylor expansions of the determinant and the adjugate. However, these expansions are well-known and can be easily computed at any order of accuracy.

  Notice that, on the other hand, if we want to exploit the formulation of [3] stated in Proposition 3.5, we should characterize the asymptotic equivalents of any coefficient of the characteristic polynomial of $\mathbf{A}_i$ and then combine them with the asymptotic equivalents of the time shifts z alone and the terms on the right hand side of Equation (7). Though this is actually feasible and we firstly did it, the computations are extremely involved[4] and very hard to generalize above second-order.

  This justifies the use of the formulation from Proposition 3.5 bis to achieve the step denoted by an horizontal arrow in Figure 1.

5 Detailed proofs

The vast majority of rest of this work is devoted to the detailed proof of Theorem 4.4 and Theorem 4.5 for the scalar case $N = 1$. This choice has been adopted to keep the presentation and the involved notations as simple as possible. The idea behind the generalization to $N > 1$ is eventually given in Section 6 and is straightforward except for the more involved notations.

Let us start by finding, for each shift operator from Definition 3.2, its asymptotically equivalent formal power series in Δx, see for instance [46, 15]. This is formalized by the following Lemma.

Lemma 5.1 (Series expansion of a shift operator in space)**.** *Let $\boldsymbol{z} \in \mathbb{Z}^d$, then the associated shift operator in space $\mathsf{t}_{\boldsymbol{z}} \in \mathbb{R}[\mathsf{z}] \otimes_{\mathbb{R}} \mathsf{D}$ is asymptotically equivalent, in the limit of $\Delta x \to 0$, to the formal power series of differential operators of the form*

$$\mathsf{t}_{\boldsymbol{z}} \asymp \sum_{|\boldsymbol{\nu}| \geq 0} \frac{(-\Delta x)^{|\boldsymbol{\nu}|}\boldsymbol{z}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\partial^{\boldsymbol{\nu}} \in \mathscr{S}.$$

Proof. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a smooth function of the spatial variable. Then performing a Taylor expansion for $\Delta x \to 0$ yields

$$(\mathsf{t}_{\boldsymbol{z}} f)(\boldsymbol{x}) = f(\boldsymbol{x} - \boldsymbol{z}\Delta x) = \sum_{|\boldsymbol{\nu}| \geq 0} \frac{(-\Delta x)^{|\boldsymbol{\nu}|}\boldsymbol{z}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\partial^{\boldsymbol{\nu}} f(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

□

[4] Probably, a deeper mastery of the elementary symmetric polynomials, the Newton's identities, the Bell polynomials and the Feddeev-Leverrier algorithm could simplify many reasonings.

The extension of Lemma 5.1 to any Finite Difference operator in D according to Definition 3.2 is done by linearity. With this in mind, recalling the definition of $\mathbf{T} \in \mathscr{M}_q(\mathsf{D})$, the moments-stream matrix and using Assumptions 4.1, we have that

$$\mathbf{T} := \boldsymbol{M}\mathrm{diag}(\mathsf{t}_{c_1},\dots,\mathsf{t}_{c_q})\boldsymbol{M}^{-1} \asymp \boldsymbol{M}\Big(\sum_{|\boldsymbol{\nu}|\geq 0}\frac{(-\Delta x)^{|\boldsymbol{\nu}|}}{\boldsymbol{\nu}!}\mathrm{diag}\Big(\boldsymbol{c}_1^{\boldsymbol{\nu}},\dots,\boldsymbol{c}_q^{\boldsymbol{\nu}}\Big)\partial^{\boldsymbol{\nu}}\Big)\boldsymbol{M}^{-1} =: \boldsymbol{\mathcal{T}} \in \mathscr{M}_q(\mathscr{S}). \tag{14}$$

Accordingly, we introduce $\boldsymbol{\mathcal{A}} := \boldsymbol{\mathcal{T}}(\boldsymbol{I}-\boldsymbol{S}) \in \mathscr{M}_q(\mathscr{S})$ and $\boldsymbol{\mathcal{B}} := \boldsymbol{\mathcal{T}}\boldsymbol{S} \in \mathscr{M}_q(\mathscr{S})$ such that $\boldsymbol{\mathcal{A}} \asymp \mathbf{A}$ and $\boldsymbol{\mathcal{B}} \asymp \mathbf{B}$. The tight bond between the momentum-velocity operator matrix $\boldsymbol{\mathcal{G}} \in \mathscr{M}_q(\mathscr{D})$ from [15] and our moments-stream matrix $\mathbf{T} \in \mathscr{M}_q(\mathsf{D})$ and its asymptotic equivalent matrix $\boldsymbol{\mathcal{T}} \in \mathscr{M}_q(\mathscr{S})$ is given by the following Lemma.

Lemma 5.2 (Link between $\boldsymbol{\mathcal{G}}$ and $\boldsymbol{\mathcal{T}}^{(r)}$)**.** *For any order $r \in \mathbb{N}$, the matrix $\boldsymbol{\mathcal{T}}^{(r)} \in \mathscr{M}_q(\mathscr{D})$ is linked to $\boldsymbol{\mathcal{G}} \in \mathscr{M}_q(\mathscr{D})$ by*

$$\boldsymbol{\mathcal{T}}^{(r)} = \frac{(-1)^r}{\lambda^r r!}\boldsymbol{\mathcal{G}}^r.$$

Moreover, using the Assumptions 4.1, we also have

$$\boldsymbol{\mathcal{A}}^{(r)} = \frac{(-1)^r}{\lambda^r r!}\boldsymbol{\mathcal{G}}^r(\boldsymbol{I}-\boldsymbol{S}), \qquad \boldsymbol{\mathcal{B}}^{(r)} = \frac{(-1)^r}{\lambda^r r!}\boldsymbol{\mathcal{G}}^r\boldsymbol{S}.$$

Proof. By Equation (21) in [15], we have that $\mathbf{T} \asymp \boldsymbol{\mathcal{T}} = \exp\big(-\frac{\Delta x}{\lambda}\boldsymbol{\mathcal{G}}\big)$. A Taylor expansion of the exponential function yields the result. Using Assumptions 4.1, one obtains that $(\boldsymbol{I}-\boldsymbol{S})$ and $\boldsymbol{S}$ do not perturb the orders of the expansion. □

As far as the time variable is concerned, we can complete by the development of the time shift operator z in order to provide the overall expansion of the inverse of the resolvent $\mathsf{z}\boldsymbol{I}-\mathbf{A} \in \mathscr{M}_q(\mathbb{R}[\mathsf{z}]\otimes_{\mathbb{R}}\mathsf{D})$.

Lemma 5.3 (Expansion of the inverse of the resolvent)**.** *Under Assumptions 4.1 and in the limit of $\Delta x \to 0$, the inverse of the resolvent $\mathsf{z}\boldsymbol{I}-\mathbf{A} \in \mathscr{M}_q(\mathbb{R}[\mathsf{z}]\otimes_{\mathbb{R}}\mathsf{D})$ is asymptotically equivalent to $\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}} \in \mathscr{M}_q(\mathscr{S})$, where*

$$\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}} = \sum_{r=0}^{+\infty}\frac{\Delta x^r}{\lambda^r r!}\big(\partial_t^r\boldsymbol{I}-(-1)^r\boldsymbol{\mathcal{G}}^r(\boldsymbol{I}-\boldsymbol{S})\big) = \boldsymbol{S}+\frac{\Delta x}{\lambda}(\partial_t\boldsymbol{I}+\boldsymbol{\mathcal{G}}(\boldsymbol{I}-\boldsymbol{S}))+\frac{\Delta x^2}{2\lambda^2}\big(\partial_{tt}\boldsymbol{I}-\boldsymbol{\mathcal{G}}^2(\boldsymbol{I}-\boldsymbol{S})\big)+O(\Delta x^3). \tag{15}$$

Proof. The standard Taylor expansion of z, using the assumption on the acoustic scaling, gives the claim. □

The consistency analysis of the Finite Difference schemes from Proposition 3.5 bis could be carried on infinite formal power series of differential operators $\mathscr{S}$ on the formulation

$$\det(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}}_i)m_i = \big(\mathrm{adj}(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}}_i)\boldsymbol{\mathcal{A}}_i^{\diamond}\boldsymbol{m}\big)_i + \big(\mathrm{adj}(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}}_i)\boldsymbol{\mathcal{B}}\boldsymbol{m}^{\mathrm{eq}}\big)_i, \tag{16}$$

for each $i \in [1\,..\,N]$, because the determinant and the adjugate perfectly make sense for any square matrix on a commutative ring, like $\mathscr{S}$. However, in order to prove Theorem 4.4 and Theorem 4.5, where formal power series are truncated at a certain order, we shall need Equation (15) from Lemma 5.3 as well as the Taylor expansions of the determinant and the adjugate matrix around a given matrix. Indeed, these are non-linear functions and thus mix different orders in the expansions $\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}} \in \mathscr{M}_q(\mathscr{S})$. Since the product of the relaxation parameters for the non-conserved moments is a quantity which shall frequently appear in the computations to come, we fix a special notation for it, namely setting $\Pi := \prod_{i=2}^{i=q} s_i \neq 0$.

5.1 Determinant

We start by studying the expansion of the determinant up to second-order in the perturbation. For this, we need to characterize its derivatives. The expansion can be carried at higher order by employing the very same strategy.

Lemma 5.4 (Derivatives and expansion of the determinant function)**.** *Let $\boldsymbol{C} \in GL_q(\mathfrak{R})$ and $\boldsymbol{D},\boldsymbol{E} \in \mathscr{M}_q(\mathfrak{R})$, where $\mathfrak{R}$ is a commutative ring. Then the determinant function*

$$\begin{aligned} det\colon \mathscr{M}_q(\mathfrak{R}) &\to \mathfrak{R} \\ \boldsymbol{C} &\mapsto det(\boldsymbol{C}), \end{aligned}$$

has the following derivatives.

$$D_{\boldsymbol{C}}\left(det(\boldsymbol{C})\right)(\boldsymbol{D}) = det(\boldsymbol{C})\, tr(\boldsymbol{C}^{-1}\boldsymbol{D}), \tag{17}$$

$$D_{\boldsymbol{CC}}\left(det(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{E}) = det(\boldsymbol{C})\left(tr(\boldsymbol{C}^{-1}\boldsymbol{E})\, tr(\boldsymbol{C}^{-1}\boldsymbol{D}) - tr(\boldsymbol{C}^{-1}\boldsymbol{E}\boldsymbol{C}^{-1}\boldsymbol{D})\right), \tag{18}$$

where $tr(\cdot)$ *indicates the trace,* i.e. *the sum of the diagonal entries. Equation* (17) *is known as Jacobi formula. Moreover, the second-order Taylor expansion of the determinant function reads*

$$det(\boldsymbol{C}+\boldsymbol{D}) = det(\boldsymbol{C}) + D_{\boldsymbol{C}}\left(det(\boldsymbol{C})\right)(\boldsymbol{D}) + \frac{1}{2} D_{\boldsymbol{CC}}\left(det(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{D}) + O(\|\boldsymbol{D}\|^3),$$

where the derivatives are given by Equation (17) *and Equation* (18).

Proof. The Jacobi formula Equation (17) is a standard result, see Chapter 0 in [22] or Chapter 5 in [48]. Let us prove Equation (18).

$$\begin{aligned}
\mathrm{D}_{\boldsymbol{CC}}\left(\det(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{E}) &:= \mathrm{D}_{\boldsymbol{C}}\left(\mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\right)(\boldsymbol{D})\right)(\boldsymbol{E}) = \mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\right)(\boldsymbol{E}),\\
&= \mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\right)(\boldsymbol{E})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D}) + \det(\boldsymbol{C})\mathrm{D}_{\boldsymbol{C}}\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\right)(\boldsymbol{E}),\\
&= \det(\boldsymbol{C})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{E})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D}) + \det(\boldsymbol{C})\mathrm{tr}(\mathrm{D}_{\boldsymbol{C}}\left(\boldsymbol{C}^{-1}\boldsymbol{D}\right)(\boldsymbol{E})),\\
&= \det(\boldsymbol{C})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{E})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D}) - \det(\boldsymbol{C})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{E}\boldsymbol{C}^{-1}\boldsymbol{D}),
\end{aligned}$$

where we have used, in this order, the product rule for derivatives, the Jacobi formula Equation (17), the linearity of the trace and the fact that $\mathrm{D}_{\boldsymbol{C}}\left(\boldsymbol{C}^{-1}\right)(\boldsymbol{D}) = -\boldsymbol{C}^{-1}\boldsymbol{D}\boldsymbol{C}^{-1}$, see Chapter 5 in [48]. □

Remark 5.5 (On the invertibility assumption)**.** There exists a form of the Jacobi formula Equation (17) for general $\boldsymbol{C} \in \mathscr{M}_q(\mathfrak{R})$ without assuming invertibility, under the form $\mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\right)(\boldsymbol{D}) = \mathrm{tr}(\mathrm{adj}(\boldsymbol{C})\boldsymbol{D})$, which is equivalent to Equation (17) since Equation (10) holds. Nevertheless, we decided to state Lemma 5.4 using the invertibility assumption. This is done, as we shall see, without loss of generality by taking advantage of some invertible approximation of real matrices and allows to easily find the formulæ for higher order derivatives and expansions *via* basic differential calculus as illustrated in the previous proof.

In the sequel, we shall take $\mathfrak{R} = \mathscr{S}$ and $\boldsymbol{C} = \boldsymbol{S} \in \mathrm{GL}_q(\mathbb{R}) \subset \mathrm{GL}_q(\mathscr{S})$ and $\boldsymbol{D} = O(\Delta x) \in \mathscr{M}_q(\mathscr{S})$. To simplify the computations and relying on the findings of Section 3.4, we can consider $\boldsymbol{S}$ singular by having $s_1 = 0$. To avoid the difficulties linked with singular matrices, in the spirit of Remark 5.5, we take advantage of the fact that the derivatives of the determinant (and the determinant itself) around $\boldsymbol{C}$ are smooth (indeed, polynomial) functions of $\boldsymbol{C}$. Thus, we introduce the non-singular approximation $\boldsymbol{S}$ where $s_1 \neq 0$, which is such that $\boldsymbol{S} \to \boldsymbol{S}|_{s_1=0}$ as $s_1 \to 0$ for any matricial topology.

We are now ready to use the expansion given by Lemma 5.3 into the terms stemming from Lemma 5.4 to find the leading order terms of the left hand side of Equation (9), namely of $\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) \in \mathscr{S}$. This is nothing but computing the Taylor series of composite functions (see the Faà di Bruno's formulæ[24]) or the composition of formal series

$$\begin{aligned}
\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) = \det(\boldsymbol{S}) &+ \Delta x \mathrm{D}_{\boldsymbol{S}}\left(\det(\boldsymbol{S})\right)\left((\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(1)}\right)\\
&+ \Delta x^2 \left(\mathrm{D}_{\boldsymbol{S}}\left(\det(\boldsymbol{S})\right)\left((\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(2)}\right) + \mathrm{D}_{\boldsymbol{SS}}\left(\det(\boldsymbol{S})\right)\left((\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(1)}\right)\left((\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(1)}\right)\right) + O(\Delta x^3).
\end{aligned}$$

- One clearly has $\det(\boldsymbol{S}) = s_1 \Pi$, because the matrix $\boldsymbol{S}$ is diagonal. Thus, the Taylor expansion of $\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})$ does not contain zero-order terms if $s_1 = 0$.
- Let $\boldsymbol{C} = \boldsymbol{S} \in \mathrm{GL}_q(\mathbb{R}) \subset \mathrm{GL}_q(\mathscr{S})$ and $\boldsymbol{D} = \frac{\Delta x}{\lambda}\left(\partial_t \boldsymbol{I} + \boldsymbol{\mathcal{G}}(\boldsymbol{I} - \boldsymbol{S})\right) + \frac{\Delta x^2}{2\lambda^2}\left(\partial_{tt}\boldsymbol{I} - \boldsymbol{\mathcal{G}}^2(\boldsymbol{I} - \boldsymbol{S})\right) + O(\Delta x^3) \in \mathscr{M}_q(\mathscr{S})$ from Lemma 5.3. Using Equation (17) from Lemma 5.4 and performing elementary computations, we have

$$\begin{aligned}
\mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\right)(\boldsymbol{D}) = s_1 \Pi \Bigg(&\frac{\Delta x}{\lambda}\left(\frac{1}{s_1}\left(\partial_t + \boldsymbol{\mathcal{G}}_{11}\right) + \sum_{i=2}^{q} \frac{1}{s_i}\left(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii}\right)\right)\\
&+ \frac{\Delta x^2}{2\lambda^2}\left(\frac{1}{s_1}\left(\partial_{tt} - \boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{11} - \sum_{\ell=2}^{q} \boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell 1}\right) + \sum_{i=2}^{q} \frac{1}{s_i}\left(\partial_{tt} - (1-s_i)\sum_{\ell=1}^{q} \boldsymbol{\mathcal{G}}_{i\ell}\boldsymbol{\mathcal{G}}_{\ell i}\right)\right)\Bigg) + O(\Delta x^3).
\end{aligned} \tag{19}$$

We keep this expression without taking the limit in s_1 for future use. On the other hand, taking the limit for $s_1 \to 0$ yields the derivative around the singular matrix $\boldsymbol{S}|_{s_1=0}$ instead of $\boldsymbol{S} \in \mathrm{GL}_q(\mathbb{R})$ for $s_1 \neq 0$.

$$\lim_{s_1 \to 0} \mathrm{D}_{\boldsymbol{C}} \left(\det(\boldsymbol{C})\right)(\boldsymbol{D}) = \Pi \frac{\Delta x}{\lambda} \left(\partial_t + \boldsymbol{\mathcal{G}}_{11} + \frac{\Delta x}{2\lambda} \left(\partial_{tt} - \boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{11} - \sum_{\ell=2}^{q} \boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell 1} \right) \right) + O(\Delta x^3). \tag{20}$$

This gives all the first-order term and part of the second-order term in the series $\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})$.

- Let $\boldsymbol{C} = \boldsymbol{S} \in \mathrm{GL}_q(\mathbb{R}) \subset \mathrm{GL}_q(\mathscr{S})$ and $\boldsymbol{D} = \frac{\Delta x}{\lambda}\left(\partial_t \boldsymbol{I} + \boldsymbol{\mathcal{G}}(\boldsymbol{I} - \boldsymbol{S})\right) + O(\Delta x^2) \in \mathscr{M}_q(\mathscr{S})$ from Lemma 5.3. Using Equation (18) from Lemma 5.4, we have, after some algebra

$$\begin{aligned} \mathrm{D}_{\boldsymbol{CC}} \left(\det(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{D}) = s_1 \Pi \frac{\Delta x^2}{\lambda^2} \Bigg(& \frac{2}{s_1}(\partial_t + \boldsymbol{\mathcal{G}}_{11}) \sum_{i=2}^{q} \frac{1}{s_i}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii}) + \left(\sum_{i=2}^{q} \frac{1}{s_i}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii}) \right)^2 \\ & - \frac{2}{s_1} \sum_{\ell=2}^{q} \left(\frac{1}{s_\ell} - 1 \right) \boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell 1} - \sum_{i=2}^{q} \frac{1}{s_i^2}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii})^2 \\ & - \sum_{i=2}^{q} \sum_{\substack{\ell=2 \\ \ell \neq i}}^{q} \left(\frac{1}{s_i} - 1 \right) \left(\frac{1}{s_\ell} - 1 \right) \boldsymbol{\mathcal{G}}_{i\ell}\boldsymbol{\mathcal{G}}_{\ell i} \Bigg) + O(\Delta x^3). \end{aligned} \tag{21}$$

Once more, we take the limit for $s_1 \to 0$ in order to find the desired result on the remaining second-order terms in the development $\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})$

$$\begin{aligned} \lim_{s_1 \to 0} \mathrm{D}_{\boldsymbol{CC}} \left(\det(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{D}) = 2\Pi \frac{\Delta x^2}{\lambda^2} \Bigg(& \left(\sum_{\ell=2}^{q} \frac{1}{s_\ell} \right) \partial_{tt} + \left(\sum_{\ell=2}^{q} \frac{1}{s_\ell} \right) \boldsymbol{\mathcal{G}}_{11}\partial_t + \left(\sum_{i=2}^{q} \left(\frac{1}{s_i} - 1 \right) \boldsymbol{\mathcal{G}}_{ii} \right) \partial_t + \boldsymbol{\mathcal{G}}_{11} \left(\sum_{i=2}^{q} \left(\frac{1}{s_i} - 1 \right) \boldsymbol{\mathcal{G}}_{ii} \right) \\ & - \sum_{\ell=2}^{q} \left(\frac{1}{s_\ell} - 1 \right) \boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell 1} \Bigg) + O(\Delta x^3). \end{aligned} \tag{22}$$

Putting Equation (20) and Equation (22) together in Lemma 5.4, with expansion around $\boldsymbol{S}$, allows to write $\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})$ up to third order. This is

$$\begin{aligned} \lim_{s_1 \to 0} \det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) = \Pi \frac{\Delta x}{\lambda} \Bigg(& \partial_t + \boldsymbol{\mathcal{G}}_{11} + \frac{\Delta x}{\lambda} \Bigg(\left(\frac{1}{2} + \sum_{\ell=2}^{q} \frac{1}{s_\ell} \right) \partial_{tt} + \left(\sum_{\ell=2}^{q} \frac{1}{s_\ell} \right) \boldsymbol{\mathcal{G}}_{11}\partial_t + \left(\sum_{i=2}^{q} \left(\frac{1}{s_i} - 1 \right) \boldsymbol{\mathcal{G}}_{ii} \right) \partial_t \\ & - \frac{1}{2}\boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{11} - \sum_{\ell=2}^{q} \left(\frac{1}{s_\ell} - \frac{1}{2} \right) \boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell 1} + \boldsymbol{\mathcal{G}}_{11} \left(\sum_{i=2}^{q} \left(\frac{1}{s_i} - 1 \right) \boldsymbol{\mathcal{G}}_{ii} \right) \Bigg) \Bigg) + O(\Delta x^3). \end{aligned} \tag{23}$$

5.2 Adjugate

We now switch to the formal power series of the adjugate function of the inverse of the resolvent, in order to deal with the right hand side of the corresponding Finite Difference scheme given by Equation (9). Let us start by characterizing its derivatives.

Lemma 5.6 (Derivatives and expansion of the adjugate function)**.** *Let $\boldsymbol{C} \in GL_q(\mathfrak{R})$ and $\boldsymbol{D}, \boldsymbol{E} \in \mathscr{M}_q(\mathfrak{R})$, where $\mathfrak{R}$ is a commutative ring. Then the adjugate function*

$$\begin{aligned} adj \colon \mathscr{M}_q(\mathfrak{R}) &\to \mathscr{M}_q(\mathfrak{R}) \\ \boldsymbol{C} &\mapsto adj(\boldsymbol{C}), \end{aligned}$$

has the following derivatives.

$$D_{\boldsymbol{C}} \left(adj(\boldsymbol{C}) \right)(\boldsymbol{D}) = det(\boldsymbol{C}) \left(tr(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D} \right) \boldsymbol{C}^{-1}, \tag{24}$$

$$\begin{aligned} D_{\boldsymbol{CC}} \left(adj(\boldsymbol{C}) \right)(\boldsymbol{D})(\boldsymbol{E}) = det(\boldsymbol{C}) \Big(& \left(tr(\boldsymbol{C}^{-1}\boldsymbol{E})\, tr(\boldsymbol{C}^{-1}\boldsymbol{D}) - tr(\boldsymbol{C}^{-1}\boldsymbol{E}\boldsymbol{C}^{-1}\boldsymbol{D}) \right) \boldsymbol{C}^{-1} \\ & + \boldsymbol{C}^{-1} \left(\boldsymbol{E}\boldsymbol{C}^{-1}\boldsymbol{D} + \boldsymbol{D}\boldsymbol{C}^{-1}\boldsymbol{E} - tr(\boldsymbol{C}^{-1}\boldsymbol{E})\boldsymbol{D} - tr(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{E} \right) \boldsymbol{C}^{-1} \Big). \end{aligned} \tag{25}$$

Moreover, the second-order Taylor expansion of the adjugate function reads

$$adj(\boldsymbol{C}+\boldsymbol{D}) = adj(\boldsymbol{C}) + D_{\boldsymbol{C}}\left(adj(\boldsymbol{C})\right)(\boldsymbol{D}) + \frac{1}{2}D_{\boldsymbol{CC}}\left(adj(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{D}) + O(\|\boldsymbol{D}\|^3),$$

where the derivatives are given by Equation (24) *and Equation* (25)*.*

Proof. Since Equation (10) holds and $\boldsymbol{C}$ is invertible, we have that $\mathrm{adj}(\boldsymbol{C}) = \det(\boldsymbol{C})\boldsymbol{C}^{-1}$. Therefore

$$\begin{aligned}\mathrm{D}_{\boldsymbol{C}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D}) &= \mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\boldsymbol{C}^{-1}\right)(\boldsymbol{D}) = \mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\right)(\boldsymbol{D})\boldsymbol{C}^{-1} + \det(\boldsymbol{C})\mathrm{D}_{\boldsymbol{C}}\left(\boldsymbol{C}^{-1}\right)(\boldsymbol{D}),\\ &= \det(\boldsymbol{C})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{C}^{-1} - \det(\boldsymbol{C})\boldsymbol{C}^{-1}\boldsymbol{D}\boldsymbol{C}^{-1},\end{aligned}$$

where we have used the rule for the derivative of a product, the Jacobi formula Equation (17) and the identity $\mathrm{D}_{\boldsymbol{C}}\left(\boldsymbol{C}^{-1}\right)(\boldsymbol{D}) = -\boldsymbol{C}^{-1}\boldsymbol{D}\boldsymbol{C}^{-1}$. For the second derivative, we have

$$\begin{aligned}\mathrm{D}_{\boldsymbol{CC}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{E}) &:= \mathrm{D}_{\boldsymbol{C}}\left(\mathrm{D}_{\boldsymbol{C}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D})\right)(\boldsymbol{E}) = \mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{C}^{-1}\right)(\boldsymbol{E}),\\
&= \mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\right)(\boldsymbol{E})\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{C}^{-1} + \det(\boldsymbol{C})\mathrm{D}_{\boldsymbol{C}}\left(\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{C}^{-1}\right)(\boldsymbol{E}),\\
&= \det(\boldsymbol{C})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{E})\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{C}^{-1} + \det(\boldsymbol{C})\mathrm{D}_{\boldsymbol{C}}\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)(\boldsymbol{E})\boldsymbol{C}^{-1}\\
&+ \det(\boldsymbol{C})\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)\mathrm{D}_{\boldsymbol{C}}\left(\boldsymbol{C}^{-1}\right)(\boldsymbol{E}),\\
&= \det(\boldsymbol{C})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{E})\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{C}^{-1}\\
&+ \det(\boldsymbol{C})\left(\mathrm{tr}\left(\mathrm{D}_{\boldsymbol{C}}\left(\boldsymbol{C}^{-1}\right)(\boldsymbol{E})\boldsymbol{D}\right)\boldsymbol{I} - \mathrm{D}_{\boldsymbol{C}}\left(\boldsymbol{C}^{-1}\right)(\boldsymbol{E})\boldsymbol{D}\right)\boldsymbol{C}^{-1}\\
&- \det(\boldsymbol{C})\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{C}^{-1}\boldsymbol{E}\boldsymbol{C}^{-1},\\
&= \det(\boldsymbol{C})\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{E})\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{C}^{-1} - \det(\boldsymbol{C})\left(\mathrm{tr}\left(\boldsymbol{C}^{-1}\boldsymbol{E}\boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{E}\boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{C}^{-1}\\
&- \det(\boldsymbol{C})\left(\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{D}\right)\boldsymbol{C}^{-1}\boldsymbol{E}\boldsymbol{C}^{-1},\end{aligned}$$

where we have used the rule for the derivative of a product, the Jacobi formula Equation (17), the linearity of the derivative and the trace and the identity $\mathrm{D}_{\boldsymbol{C}}\left(\boldsymbol{C}^{-1}\right)(\boldsymbol{D}) = -\boldsymbol{C}^{-1}\boldsymbol{D}\boldsymbol{C}^{-1}$. Upon rearrangement, this yields the result. □

Remark 5.7. We observe that, looking at Equation (24) and Equation (25) compared to Equation (17) and Equation (18), we have that

$$\begin{aligned}\mathrm{D}_{\boldsymbol{C}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D}) &= \mathrm{D}_{\boldsymbol{C}}\left(\det(\boldsymbol{C})\right)(\boldsymbol{D})\boldsymbol{C}^{-1} - \det(\boldsymbol{C})\boldsymbol{C}^{-1}\boldsymbol{D}\boldsymbol{C}^{-1},\\
\mathrm{D}_{\boldsymbol{CC}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{E}) &= \mathrm{D}_{\boldsymbol{CC}}\left(\det(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{E})\boldsymbol{C}^{-1} + \det(\boldsymbol{C})\boldsymbol{C}^{-1}\left(\boldsymbol{E}\boldsymbol{C}^{-1}\boldsymbol{D} + \boldsymbol{D}\boldsymbol{C}^{-1}\boldsymbol{E} - \mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{E})\boldsymbol{D} - \mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{D})\boldsymbol{E}\right)\boldsymbol{C}^{-1}.\end{aligned}$$

This implies that we can reuse the computations we did for the determinant in the current treatment of the adjugate, as far as the first terms on the right hand sides are concerned. However, one must be careful that now they are multiplied by $\boldsymbol{C}^{-1}$.

If we had stopped the developments at first order, we could have used the first-order perturbation theory of the adjugate matrix as provided by Theorem 2.1 from [40]. However, to the best of our knowledge, no second-order perturbation theory for this matrix is available in the literature, thus we have been compelled to independently develop it using differential calculus. Lemma 5.6 is thus a generalization of the results from [40] and can therefore be used – beyond the application presented in this contribution – by researchers needing a second-order perturbation theory for the adjugate matrix.

Since we are ultimately interested, as one can notice from Equation (9), in multiplying the formal power series $\mathrm{adj}(\zeta\boldsymbol{I} - \boldsymbol{\mathcal{A}}) \in \mathscr{M}_q(\mathscr{S})$ by $\boldsymbol{\mathcal{B}} \in \mathscr{M}_q(\mathscr{S})$ in a Cauchy-like fashion (the standard product of formal power series) and select the first row, see Proposition 3.4 bis, we perform the computations only for the first row of $\mathrm{adj}(\zeta\boldsymbol{I} - \boldsymbol{\mathcal{A}})$.

- Using the definition of the adjugate matrix in combination with the Laplace formula or using the explicit formula for the adjugate of an upper triangular matrix, see [22], we have

$$\mathrm{adj}(\boldsymbol{S}) = \Pi\mathrm{diag}\left(1, \frac{s_1}{s_2}, \ldots, \frac{s_1}{s_q}\right), \qquad \text{thus} \qquad \lim_{s_1\to 0}\mathrm{adj}(\boldsymbol{S}) = \Pi\boldsymbol{e}_1 \otimes \boldsymbol{e}_1.$$

  Hence, contrarily to the determinant, the zero-order term in $\mathrm{adj}(\zeta\boldsymbol{I} - \boldsymbol{\mathcal{A}})$ is not zero for $s_1 = 0$ but is still a singular one-rank diagonal matrix.

- Let $\boldsymbol{C} = \boldsymbol{S} \in \mathrm{GL}_q(\mathbb{R}) \subset \mathrm{GL}_q(\mathscr{S})$ and $\boldsymbol{D} = \frac{\Delta x}{\lambda}(\partial_t \boldsymbol{I} + \boldsymbol{\mathcal{G}}(\boldsymbol{I}-\boldsymbol{S})) + \frac{\Delta x^2}{2\lambda^2}\left(\partial_{tt}\boldsymbol{I} - \boldsymbol{\mathcal{G}}^2(\boldsymbol{I}-\boldsymbol{S})\right) + O(\Delta x^3) \in \mathscr{M}_q(\mathscr{S})$ from Lemma 5.3. We utilize the previous computations from Equation (19) as suggested in Remark 5.7, plus Equation (24).

$$\begin{aligned}\mathrm{D}_{\boldsymbol{C}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D}) = s_1\Pi\Bigg(\frac{\Delta x}{\lambda}\Bigg(\frac{1}{s_1}\Big(\partial_t + \boldsymbol{\mathcal{G}}_{11}\Big) + \sum_{i=2}^{q}\frac{1}{s_i}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii})\Bigg) + \frac{\Delta x^2}{2\lambda^2}\Bigg(\frac{1}{s_1}\Big(\partial_{tt} - \boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{11} - \sum_{\ell=2}^{q}\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell 1}\Big)\\ + \sum_{i=2}^{q}\frac{1}{s_i}\Big(\partial_{tt} - (1-s_i)\sum_{\ell=1}^{q}\boldsymbol{\mathcal{G}}_{i\ell}\boldsymbol{\mathcal{G}}_{\ell i}\Big)\Bigg)\Bigg)\mathrm{diag}\Big(\frac{1}{s_1},\frac{1}{s_2},\ldots,\frac{1}{s_q}\Big)\\ - s_1\Pi\mathrm{diag}\Big(\frac{1}{s_1},\frac{1}{s_2},\ldots,\frac{1}{s_q}\Big)\boldsymbol{D}\mathrm{diag}\Big(\frac{1}{s_1},\frac{1}{s_2},\ldots,\frac{1}{s_q}\Big) + O(\Delta x^3).\end{aligned}$$

In this case, we do not even have to take the limit for $s_1 \to 0$, since all the terms in s_1 cancel. Therefore, for the very first component, we get

$$\begin{aligned}\left(\mathrm{D}_{\boldsymbol{C}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D})\right)_{11} = \Pi\frac{\Delta x}{\lambda}\Bigg(\Big(\sum_{\ell=2}^{q}\frac{1}{s_\ell}\Big)\partial_t + \sum_{i=2}^{q}\Big(\frac{1}{s_i}-1\Big)\boldsymbol{\mathcal{G}}_{ii} + \frac{\Delta x}{2\lambda}\Bigg(\Big(\sum_{\ell=2}^{q}\frac{1}{s_\ell}\Big)\partial_{tt} - \sum_{i=2}^{q}\Big(\frac{1}{s_i}-1\Big)\sum_{\ell=1}^{q}\boldsymbol{\mathcal{G}}_{i\ell}\boldsymbol{\mathcal{G}}_{\ell i}\Bigg)\Bigg)\\ + O(\Delta x^3).\end{aligned} \tag{26}$$

Now consider $j \in [2\,..\,q]$, then

$$\left(\mathrm{D}_{\boldsymbol{C}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D})\right)_{1j} = -\Pi\frac{\Delta x}{\lambda}\Big(\frac{1}{s_j}-1\Big)\Big(\boldsymbol{\mathcal{G}}_{1j} - \frac{\Delta x}{2\lambda}\Big(\boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{1j} + \sum_{\ell=2}^{q}\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell j}\Big)\Big) + O(\Delta x^3), \tag{27}$$

This gives all the first-order terms on the first row of $\mathrm{adj}(\zeta\boldsymbol{I} - \boldsymbol{\mathcal{A}})$ and part of the second-order terms.

- Let $\boldsymbol{C} = \boldsymbol{S} \in \mathrm{GL}_q(\mathbb{R}) \subset \mathrm{GL}_q(\mathscr{S})$ and $\boldsymbol{D} = \frac{\Delta x}{\lambda}(\partial_t \boldsymbol{I} + \boldsymbol{\mathcal{G}}(\boldsymbol{I}-\boldsymbol{S})) + O(\Delta x^2) \in \mathscr{M}_q(\mathscr{S})$ from Lemma 5.3. We reuse computations from Equation (21) as well as Equation (25).

$$\begin{aligned}\mathrm{D}_{\boldsymbol{CC}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{D}) = s_1\Pi\frac{\Delta x^2}{\lambda^2}\Bigg(\frac{2}{s_1}(\partial_t + \boldsymbol{\mathcal{G}}_{11})\sum_{i=2}^{q}\frac{1}{s_i}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii}) + \Big(\sum_{i=2}^{q}\frac{1}{s_i}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii})\Big)^2\\ - \frac{2}{s_1}\sum_{\ell=2}^{q}\Big(\frac{1}{s_\ell}-1\Big)\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell 1} - \sum_{i=2}^{q}\frac{1}{s_i^2}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii})^2\\ - \sum_{i=2}^{q}\sum_{\substack{\ell=2\\ \ell\neq i}}^{q}\Big(\frac{1}{s_i}-1\Big)\Big(\frac{1}{s_\ell}-1\Big)\boldsymbol{\mathcal{G}}_{i\ell}\boldsymbol{\mathcal{G}}_{\ell i}\Bigg)\mathrm{diag}\Big(\frac{1}{s_1},\frac{1}{s_2},\ldots,\frac{1}{s_q}\Big)\\ + 2s_1\Pi\mathrm{diag}\Big(\frac{1}{s_1},\frac{1}{s_2},\ldots,\frac{1}{s_q}\Big)\Big(\boldsymbol{D}\boldsymbol{S}^{-1}\boldsymbol{D} - \mathrm{tr}(\boldsymbol{S}^{-1}\boldsymbol{D})\boldsymbol{D}\Big)\mathrm{diag}\Big(\frac{1}{s_1},\frac{1}{s_2},\ldots,\frac{1}{s_q}\Big) + O(\Delta x^3).\end{aligned}$$

Then we have, for the first matrix entry

$$\begin{aligned}\left(\mathrm{D}_{\boldsymbol{CC}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{D})\right)_{11} = \Pi\frac{\Delta x^2}{\lambda^2}\Bigg(\Big(\sum_{i=2}^{q}\frac{1}{s_i}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii})\Big)^2 - \sum_{i=2}^{q}\frac{1}{s_i^2}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii})^2\\ - \sum_{i=2}^{q}\sum_{\substack{\ell=2\\ \ell\neq i}}^{q}\Big(\frac{1}{s_i}-1\Big)\Big(\frac{1}{s_\ell}-1\Big)\boldsymbol{\mathcal{G}}_{i\ell}\boldsymbol{\mathcal{G}}_{\ell i}\Bigg) + O(\Delta x^3),\end{aligned} \tag{28}$$

independent from s_1. On the other hand, for $j \in [2\,..\,q]$

$$\begin{aligned}\left(\mathrm{D}_{\boldsymbol{CC}}\left(\mathrm{adj}(\boldsymbol{C})\right)(\boldsymbol{D})(\boldsymbol{D})\right)_{1j} = 2\Pi\frac{\Delta x^2}{\lambda^2}\Big(\frac{1}{s_j}-1\Big)\Bigg(\frac{1}{s_j}\boldsymbol{\mathcal{G}}_{1j}(\partial_t + (1-s_j)\boldsymbol{\mathcal{G}}_{jj}) + \sum_{\substack{\ell=2\\ \ell\neq j}}^{q}\Big(\frac{1}{s_\ell}-1\Big)\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell j}\\ - \boldsymbol{\mathcal{G}}_{1j}\sum_{i=2}^{q}\frac{1}{s_i}(\partial_t + (1-s_i)\boldsymbol{\mathcal{G}}_{ii})\Bigg) + O(\Delta x^3).\end{aligned} \tag{29}$$

Using Equation (26) and Equation (28), we have that the first entry on the first row of $\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})$ is

$$\lim_{s_1 \to 0} (\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}))_{11} = \Pi + \Pi \frac{\Delta x}{\lambda} \left(\left(\sum_{\ell=2}^{q} \frac{1}{s_\ell} \right) \partial_t + \sum_{i=2}^{q} \left(\frac{1}{s_i} - 1 \right) \boldsymbol{\mathcal{G}}_{ii} \right) + \Pi \frac{\Delta x^2}{2\lambda^2} \left(\left(\sum_{\ell=2}^{q} \frac{1}{s_\ell} \right) \partial_{tt} - \sum_{i=2}^{q} \left(\frac{1}{s_i} - 1 \right) \sum_{\ell=1}^{q} \boldsymbol{\mathcal{G}}_{i\ell} \boldsymbol{\mathcal{G}}_{\ell i} \right.$$
$$\left. + 2 \left(\sum_{i=2}^{q} \frac{1}{s_i} (\partial_t + (1 - s_i) \boldsymbol{\mathcal{G}}_{ii}) \right)^2 - 2 \sum_{i=2}^{q} \frac{1}{s_i^2} (\partial_t + (1 - s_i) \boldsymbol{\mathcal{G}}_{ii})^2 - 2 \sum_{i=2}^{q} \sum_{\substack{\ell=2 \\ \ell \neq i}}^{q} \left(\frac{1}{s_i} - 1 \right) \left(\frac{1}{s_\ell} - 1 \right) \boldsymbol{\mathcal{G}}_{i\ell} \boldsymbol{\mathcal{G}}_{\ell i} \right) + O(\Delta x^3). \tag{30}$$

On the other hand, using Equation (27) and Equation (29), for any $j \in [2 \,..\, q]$, we write

$$\lim_{s_1 \to 0} (\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}))_{1j} = -\Pi \frac{\Delta x}{\lambda} \left(\frac{1}{s_j} - 1 \right) \boldsymbol{\mathcal{G}}_{1j} + \Pi \frac{\Delta x^2}{2\lambda^2} \left(\frac{1}{s_j} - 1 \right) \left(\boldsymbol{\mathcal{G}}_{11} \boldsymbol{\mathcal{G}}_{1j} + \sum_{\ell=2}^{q} \boldsymbol{\mathcal{G}}_{1\ell} \boldsymbol{\mathcal{G}}_{\ell j} + \frac{2}{s_j} \boldsymbol{\mathcal{G}}_{1j} (\partial_t + (1 - s_j) \boldsymbol{\mathcal{G}}_{jj}) \right.$$
$$\left. + 2 \sum_{\substack{\ell=2 \\ \ell \neq j}}^{q} \left(\frac{1}{s_\ell} - 1 \right) \boldsymbol{\mathcal{G}}_{1\ell} \boldsymbol{\mathcal{G}}_{\ell j} - 2 \boldsymbol{\mathcal{G}}_{1j} \sum_{i=2}^{q} \frac{1}{s_i} (\partial_t + (1 - s_i) \boldsymbol{\mathcal{G}}_{ii}) \right) + O(\Delta x^3). \tag{31}$$

In general, we have written, for the first row, the leading terms in $\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})$. We shall take its product with $\boldsymbol{\mathcal{B}}$. Thus, one has

$$\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) \boldsymbol{\mathcal{B}} = \mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(0)} \boldsymbol{\mathcal{B}}^{(0)} + \Delta x \left(\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(0)} \boldsymbol{\mathcal{B}}^{(1)} + \mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(1)} \boldsymbol{\mathcal{B}}^{(0)} \right), \tag{32}$$
$$+ \Delta x^2 \left(\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(0)} \boldsymbol{\mathcal{B}}^{(2)} + \mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(1)} \boldsymbol{\mathcal{B}}^{(1)} + \mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(2)} \boldsymbol{\mathcal{B}}^{(0)} \right) + O(\Delta x^3),$$

generating products of terms in the fashion of the Cauchy product. This completes the preliminary results needed to prove Theorem 4.4 and Theorem 4.5.

5.3 Overall computation

We now put all the previous calculations together to prove Theorem 4.4 and Theorem 4.5. As previously pointed out, we can assume, without loss of generality, that $s_1 = 0$, passing to the limit. This allows to deal with simpler expressions with less terms.

5.3.1 First-order equations

Starting with Theorem 4.4, it is sufficient to truncate all the formal power series at $O(\Delta x^2)$. In particular, using the fact that the first column of $\boldsymbol{\mathcal{B}}$ is zero for $s_1 = 0$, we have that $\lim_{s_1 \to 0} (\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) \boldsymbol{\mathcal{B}})_{11} = 0$. Observe that if the relaxation parameter corresponding to the conserved moment were not equal to zero, we would have had $(\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) \boldsymbol{\mathcal{B}})_{11} = O(1)$. Still the matrix $\boldsymbol{S}$ would not have been singular thus we would have some non vanishing zero-order term in $\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})$ to compensate the one from the adjugate.

On the other hand, for any $j \in [2 \,..\, q]$, using Equation (31), Lemma 5.2 and Equation (32), entails

$$\lim_{s_1 \to 0} (\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) \boldsymbol{\mathcal{B}})_{1j} = \Delta x \left(\left(\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(0)} \boldsymbol{\mathcal{B}}^{(1)} \right)_{1j} + \left(\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(1)} \boldsymbol{\mathcal{B}}^{(0)} \right)_{1j} \right) + O(\Delta x^2),$$
$$= \Delta x \left(-\Pi \frac{1}{\lambda} s_j \boldsymbol{\mathcal{G}}_{1j} - \Pi \frac{1}{\lambda} s_j \left(\frac{1}{s_j} - 1 \right) \right) + O(\Delta x^2) = -\Pi \frac{\Delta x}{\lambda} \boldsymbol{\mathcal{G}}_{1j} + O(\Delta x^2).$$

On the other hand, Equation (23) directly yields

$$\lim_{s_1 \to 0} \det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) = \Pi \frac{\Delta x}{\lambda} (\partial_t + \boldsymbol{\mathcal{G}}_{11}) + O(\Delta x^2),$$

thus we obtain the modified equation (whatever the choice of $s_1 \in \mathbb{R}$)

$$\Pi \frac{\Delta x}{\lambda} \left(\partial_t m_1 + \boldsymbol{\mathcal{G}}_{11} m_1 + \sum_{j=2}^{q} \boldsymbol{\mathcal{G}}_{1j} m_j^{\mathrm{eq}} \right) = O(\Delta x^2),$$

which is the desired result from Theorem 4.4 for $N = 1$ upon dividing by the constant $\Pi = O(1)$. Observe that the term Π is never present in the computations by [15] because they are done on the original lattice Boltzmann scheme Equation (5) or Equation (8) which has only one time step. For instance, in [15], the multi-step

nature of the problem, generated by the non-conserved moments relaxing away from the equilibrium, is damped at the very beginning of the procedure by performing the Taylor expansions of the scheme on the non-conserved variables and then plugging them into the expansions for the conserved moments.

Before proceeding, let us utilize the previous equation to get rid of the time derivatives in the second order terms. This is in general not compulsory but constitutes the policy by [13, 15] and is common to all the approaches (Chapman-Enskog, equivalent equation, Maxwell iteration, *etc.*) in order to find the value of the diffusion coefficients from the second-order terms. Notice that in this case, where $N = 1$, $\gamma_{1,\cdot}$ is a scalar here denoted γ_1 for brevity.

$$\partial_t m_1 = -\boldsymbol{\mathcal{G}}_{11} m_1 - \sum_{j=2}^{q} \boldsymbol{\mathcal{G}}_{1j} m_j^{\mathrm{eq}} + O(\Delta x) = -\gamma_1 + O(\Delta x), \tag{33}$$

$$\partial_t \boldsymbol{m}^{\mathrm{eq}} = \frac{\mathrm{d}\boldsymbol{m}^{\mathrm{eq}}}{\mathrm{d}m_1}\partial_t m_1 = -\frac{\mathrm{d}\boldsymbol{m}^{\mathrm{eq}}}{\mathrm{d}m_1}\left(\boldsymbol{\mathcal{G}}_{11} m_1 + \sum_{j=2}^{q} \boldsymbol{\mathcal{G}}_{1j} m_j^{\mathrm{eq}}\right) + O(\Delta x) = -\frac{\mathrm{d}\boldsymbol{m}^{\mathrm{eq}}}{\mathrm{d}m_1}\gamma_1 + O(\Delta x), \tag{34}$$

$$\partial_{tt} m_1 = -\partial_t\left(\boldsymbol{\mathcal{G}}_{11} m_1 + \sum_{j=2}^{q} \boldsymbol{\mathcal{G}}_{1j} m_j^{\mathrm{eq}}\right) + O(\Delta x) = -\boldsymbol{\mathcal{G}}_{11}\partial_t m_1 - \sum_{j=2}^{q} \boldsymbol{\mathcal{G}}_{1j}\partial_t m_j^{\mathrm{eq}} + O(\Delta x), \tag{35}$$

$$= \boldsymbol{\mathcal{G}}_{11}\gamma_1 + \sum_{j=2}^{q} \boldsymbol{\mathcal{G}}_{1j}\frac{\mathrm{d}m_j^{\mathrm{eq}}}{\mathrm{d}m_1}\gamma_1 + O(\Delta x) = \boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{11} m_1 + \boldsymbol{\mathcal{G}}_{11}\sum_{j=2}^{q} \boldsymbol{\mathcal{G}}_{1j} m_j^{\mathrm{eq}} + \sum_{j=2}^{q} \boldsymbol{\mathcal{G}}_{1j}\frac{\mathrm{d}m_j^{\mathrm{eq}}}{\mathrm{d}m_1}\gamma_1 + O(\Delta x). \tag{36}$$

These equalities are formal and obtained by taking advantage either of the chain rule, since the moments at equilibrium are functions of the conserved moments, or of the re-injection of Equation (33) by assuming that the differentiation preserves the asymptotic relations from the symbol $O(\cdot)$. This way of proceeding is fully understood in the framework of Finite Difference schemes for PDEs, see [45, 1].

5.3.2 Second-order equations

We can now go to the proof of Theorem 4.5, which is more involved due to the presence of more terms to estimate. To make the link with the findings of [15], the increased complexity comes from the more intricate and entangled block structure of $\boldsymbol{\mathcal{G}}^2$. We have to treat the second-order term in Equation (32), made up of three products. For any $j \in [2\,..\,q]$ (once again, the first component vanishes for $s_1 = 0$)

- Using Lemma 5.2 and the zero-order expansion of the adjugate gives

$$\lim_{s_1\to 0}\left(\mathrm{adj}(\zeta\boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(0)}\boldsymbol{\mathcal{B}}^{(2)}\right)_{1j} = \Pi\boldsymbol{\mathcal{B}}_{1j}^{(2)} = \Pi\frac{1}{2\lambda^2}s_j\left(\boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{1j} + \sum_{\ell=2}^{q}\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell j}\right).$$

- Using Lemma 5.2 with Equation (30) and Equation (31)

$$\begin{aligned}\lim_{s_1\to 0}\left(\mathrm{adj}(\zeta\boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(1)}\boldsymbol{\mathcal{B}}^{(1)}\right)_{1j} &= -\frac{1}{\lambda}s_j\sum_{\ell=1}^{q}\left(\mathrm{adj}(\zeta\boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(1)}\right)_{1\ell}\boldsymbol{\mathcal{G}}_{\ell j},\\ &= -\Pi\frac{1}{\lambda^2}s_j\left(\boldsymbol{\mathcal{G}}_{1j}\left(\sum_{\ell=2}^{q}\frac{1}{s_\ell}\right)\partial_t + \boldsymbol{\mathcal{G}}_{1j}\sum_{\ell=2}^{q}\left(\frac{1}{s_\ell} - 1\right)\boldsymbol{\mathcal{G}}_{\ell\ell} - \sum_{\ell=2}^{q}\left(\frac{1}{s_\ell} - 1\right)\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell j}\right).\end{aligned}$$

- Using Lemma 5.2 and Equation (31)

$$\begin{aligned}\lim_{s_1\to 0}\left(\mathrm{adj}(\zeta\boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(2)}\boldsymbol{\mathcal{B}}^{(0)}\right)_{1j} &= s_j\left(\mathrm{adj}(\zeta\boldsymbol{I} - \boldsymbol{\mathcal{A}})^{(2)}\right)_{1j},\\ &= \Pi\frac{1}{\lambda^2}\left(1 - s_j\right)\left(\frac{1}{2}\boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{1j} + \sum_{\ell=2}^{q}\left(\frac{1}{s_\ell} - \frac{1}{2}\right)\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell j} - \left(\frac{1}{s_j} - 1\right)\boldsymbol{\mathcal{G}}_{1j}\boldsymbol{\mathcal{G}}_{jj}\right.\\ &\quad\left. + \frac{1}{s_j}\boldsymbol{\mathcal{G}}_{1j}(\partial_t + (1 - s_j)\boldsymbol{\mathcal{G}}_{jj}) - \boldsymbol{\mathcal{G}}_{1j}\sum_{i=2}^{q}\frac{1}{s_i}(\partial_t + (1 - s_i)\boldsymbol{\mathcal{G}}_{ii})\right).\end{aligned}$$

Summing these three contributions and after some straightforward but tedious computations, the second-order term in Equation (32) is given by

$$\lim_{s_1\to 0}\left((\mathrm{adj}(\zeta \boldsymbol{I}-\boldsymbol{\mathcal{A}})\boldsymbol{\mathcal{B}})^{(2)}\right)_{1j}=\Pi\frac{1}{\lambda^2}\left(\frac{1}{2}\boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{1j}+\sum_{\ell=2}^{q}\left(\frac{1}{s_\ell}-\frac{1}{2}\right)\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell j}-\boldsymbol{\mathcal{G}}_{1j}\left(1+\sum_{\substack{\ell=2\\ \ell\neq j}}^{q}\frac{1}{s_\ell}\right)\partial_t-\boldsymbol{\mathcal{G}}_{1j}\sum_{\ell=2}^{q}\left(\frac{1}{s_\ell}-1\right)\boldsymbol{\mathcal{G}}_{\ell\ell}\right).$$

Hence, using Equation (34) to get rid of the time derivative of the equilibria, we have

$$\begin{aligned}\lim_{s_1\to 0}\sum_{j=2}^{q}\left((\mathrm{adj}(\zeta \boldsymbol{I}-\boldsymbol{\mathcal{A}})\boldsymbol{\mathcal{B}})^{(2)}\right)_{1j}m_j^{\mathrm{eq}}=\Pi\frac{1}{\lambda^2}\sum_{j=2}^{q}\Bigg(&\frac{1}{2}\boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{1j}m_j^{\mathrm{eq}}+\sum_{\ell=2}^{q}\left(\frac{1}{s_\ell}-\frac{1}{2}\right)\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell j}m_j^{\mathrm{eq}}\\&+\boldsymbol{\mathcal{G}}_{1j}\left(1+\sum_{\substack{\ell=2\\ \ell\neq j}}^{q}\frac{1}{s_\ell}\right)\frac{\mathrm{d}m_j^{\mathrm{eq}}}{\mathrm{d}m_1}\gamma_1-\boldsymbol{\mathcal{G}}_{1j}\sum_{\ell=2}^{q}\left(\frac{1}{s_\ell}-1\right)\boldsymbol{\mathcal{G}}_{\ell\ell}m_j^{\mathrm{eq}}\Bigg)+O(\Delta x).\end{aligned}$$

Notice that in this result, a reminder of order $O(\Delta x)$ appears. Once more, using Equation (33) and Equation (36) to eliminate the time derivatives in the second-order terms from Equation (23) gives

$$\begin{aligned}\lim_{s_1\to 0}(\det(\zeta \boldsymbol{I}-\boldsymbol{\mathcal{A}}))^{(2)}m_1=\Pi\frac{1}{\lambda^2}\Bigg(&\left(\frac{1}{2}+\sum_{\ell=2}^{q}\frac{1}{s_\ell}\right)\partial_{tt}m_1+\left(\sum_{\ell=2}^{q}\frac{1}{s_\ell}\right)\boldsymbol{\mathcal{G}}_{11}\partial_t m_1+\left(\sum_{i=2}^{q}\left(\frac{1}{s_i}-1\right)\boldsymbol{\mathcal{G}}_{ii}\right)\partial_t m_1\\&-\frac{1}{2}\boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{11}m_1-\sum_{\ell=2}^{q}\left(\frac{1}{s_\ell}-\frac{1}{2}\right)\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell 1}m_1+\boldsymbol{\mathcal{G}}_{11}\left(\sum_{i=2}^{q}\left(\frac{1}{s_i}-1\right)\boldsymbol{\mathcal{G}}_{ii}\right)m_1\Bigg),\\=\Pi\frac{1}{\lambda^2}\Bigg(&\left(\frac{1}{2}+\sum_{\ell=2}^{q}\frac{1}{s_\ell}\right)\left(\boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{11}+\boldsymbol{\mathcal{G}}_{11}\sum_{j=2}^{q}\boldsymbol{\mathcal{G}}_{1j}+\sum_{j=2}^{q}\boldsymbol{\mathcal{G}}_{1j}\frac{\mathrm{d}m_j^{\mathrm{eq}}}{\mathrm{d}m_1}\gamma_1\right)\\&-\left(\sum_{\ell=2}^{q}\frac{1}{s_\ell}\right)\boldsymbol{\mathcal{G}}_{11}\left(\boldsymbol{\mathcal{G}}_{11}m_1+\sum_{j=2}^{q}\boldsymbol{\mathcal{G}}_{1j}m_j^{\mathrm{eq}}\right)-\left(\sum_{i=2}^{q}\left(\frac{1}{s_i}-1\right)\boldsymbol{\mathcal{G}}_{ii}\right)\left(\boldsymbol{\mathcal{G}}_{11}m_1+\sum_{j=2}^{q}\boldsymbol{\mathcal{G}}_{1j}m_j^{\mathrm{eq}}\right)\\&-\frac{1}{2}\boldsymbol{\mathcal{G}}_{11}\boldsymbol{\mathcal{G}}_{11}m_1-\sum_{\ell=2}^{q}\left(\frac{1}{s_\ell}-\frac{1}{2}\right)\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell 1}m_1+\boldsymbol{\mathcal{G}}_{11}\left(\sum_{i=2}^{q}\left(\frac{1}{s_i}-1\right)\boldsymbol{\mathcal{G}}_{ii}\right)m_1\Bigg)+O(\Delta x).\end{aligned}$$

With this, after simplifications, we obtain the remaining term to master the second-order contributions in the expansion of the Finite Difference scheme Equation (9).

$$\begin{aligned}&(\det(\zeta \boldsymbol{I}-\boldsymbol{\mathcal{A}}))^{(2)}m_1-\sum_{j=2}^{q}\left((\mathrm{adj}(\zeta \boldsymbol{I}-\boldsymbol{\mathcal{A}})\boldsymbol{\mathcal{B}})^{(2)}\right)_{1j}m_j^{\mathrm{eq}}\\&=\Pi\frac{1}{\lambda^2}\left(-\sum_{j=2}^{q}\left(\frac{1}{s_j}-\frac{1}{2}\right)\boldsymbol{\mathcal{G}}_{1j}\boldsymbol{\mathcal{G}}_{j1}m_1-\sum_{j=2}^{q}\sum_{\ell=2}^{q}\left(\frac{1}{s_\ell}-\frac{1}{2}\right)\boldsymbol{\mathcal{G}}_{1\ell}\boldsymbol{\mathcal{G}}_{\ell j}m_j^{\mathrm{eq}}+\sum_{j=2}^{q}\left(\frac{1}{s_j}-\frac{1}{2}\right)\boldsymbol{\mathcal{G}}_{1j}\frac{\mathrm{d}m_j^{\mathrm{eq}}}{\mathrm{d}m_1}\gamma_1\right)+O(\Delta x).\end{aligned}$$

To wrap up, these computations yield, together with the ones from Section 5.3.1, the expected result for $N=1$, which reads

$$\frac{\Delta x}{\lambda}\Pi\left(\partial_t m_1+\gamma_1+\frac{\Delta x}{\lambda}\sum_{j=2}^{q}\left(\frac{1}{s_j}-\frac{1}{2}\right)\boldsymbol{\mathcal{G}}_{1j}\left(\frac{\mathrm{d}m_j}{\mathrm{d}m_1}\gamma_1-\boldsymbol{\mathcal{G}}_{j1}m_1-\sum_{\ell=2}^{q}\boldsymbol{\mathcal{G}}_{j\ell}m_\ell^{\mathrm{eq}}\right)\right)=O(\Delta x^3),$$

and thus proves Theorem 4.5.

6 Extension of the proofs to several conserved moments: key ideas

In this Section, we sketch the demonstration of Theorem 4.4 and Theorem 4.5 for any $N\geq 1$. For the sake of providing a quick and effective presentation of this matter, we limit ourselves to first-order in Δx. Select a conserved moment, which shall be indexed by $i\in[1\,..\,N]$.

Remark 6.1. The operation selecting rows and columns to yield $\mathbf{A}_i$ and $\mathbf{A}_i^\diamond$ from Proposition 3.5 bis does not change the orders of the expansions. This is, let $\mathbf{C} \in \mathscr{M}_q(\mathbb{R}[\mathsf{z}] \otimes_{\mathbb{R}} \mathsf{D})$ and $\mathcal{C} = \sum_{r=0}^{r=+\infty} \Delta x^r \mathcal{C}^{(r)} \in \mathscr{M}_q(\mathscr{S})$ such that $\mathbf{C} \asymp \mathcal{C}$ and $I \subset [1\,..\,q]$ a set of indices, then

$$\mathbf{C}_I \asymp \left(\sum_{r=0}^{+\infty} \Delta x^r \mathcal{C}^{(r)} \right)_I = \sum_{r=0}^{+\infty} \Delta x^r \left(\mathcal{C}^{(r)} \right)_I .$$

Thus we have the analogous of Lemma 5.3, where $\mathsf{z}\boldsymbol{I} - \mathbf{A}_i \asymp \zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i$, with

$$\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i = \sum_{r=0}^{+\infty} \frac{\Delta x^r}{\lambda^r r!} \left(\partial_t^r \boldsymbol{I} - (-1)^r \left(\boldsymbol{\mathcal{G}}^r (\boldsymbol{I} - \boldsymbol{S}) \right)_{\{i\} \cup [N+1\,..\,q]} \right).$$

The first two term in the expansion of the inverse of the resolvent are

$$(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)^{(0)} = \mathrm{diag}(1, \ldots, 1, s_i, 1, \ldots, 1, s_{N+1}, \ldots, s_q).$$

In the spirit of Remark 5.5 and the computations developed in Section 5, for the case $s_i = 0$, we introduce a regularization with $s_i \neq 0$ and then we pass to the limit. Moreover

$$(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)^{(1)} = \frac{1}{\lambda} \partial_t \boldsymbol{I} + \frac{1}{\lambda} \left(\begin{array}{ccc|c|ccc||c|c} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & (1-s_i)\boldsymbol{\mathcal{G}}_{ii} & 0 & \cdots & 0 & (1-s_{N+1})\boldsymbol{\mathcal{G}}_{i(N+1)} & \cdots & (1-s_q)\boldsymbol{\mathcal{G}}_{iq} \\ \hline 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline\hline 0 & \cdots & 0 & (1-s_i)\boldsymbol{\mathcal{G}}_{(N+1)i} & 0 & \cdots & 0 & (1-s_{N+1})\boldsymbol{\mathcal{G}}_{(N+1)(N+1)} & \cdots & (1-s_q)\boldsymbol{\mathcal{G}}_{(N+1)q} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & (1-s_i)\boldsymbol{\mathcal{G}}_{qi} & 0 & \cdots & 0 & (1-s_{N+1})\boldsymbol{\mathcal{G}}_{q(N+1)} & \cdots & (1-s_q)\boldsymbol{\mathcal{G}}_{qq} \end{array} \right).$$

We thus have

- As for the case $N = 1$ treated in detail, we have that $\lim_{s_i \to 0} \det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)^{(0)} = 0$. On the other hand, using the formula for the adjugate of an upper triangular matrix, see [22], we have $\lim_{s_i \to 0} \mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)^{(0)} = \Pi \boldsymbol{e}_i \otimes \boldsymbol{e}_i$, where in this Section $\Pi := \prod_{\ell=N+1}^{\ell=q} s_\ell$.

- Taking $\boldsymbol{C} = (\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)^{(0)} \in \mathrm{GL}_q(\mathbb{R}) \subset \mathrm{GL}_q(\mathscr{S})$ and $\boldsymbol{D} = \Delta x (\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)^{(1)} + O(\Delta x^2) \in \mathscr{M}_q(\mathscr{S})$ in the Jacobi formula Equation (17)

$$\begin{aligned} \lim_{s_i \to 0} \mathrm{D}_{\boldsymbol{C}} (\det(\boldsymbol{C}))(\boldsymbol{D}) &= \lim_{s_i \to 0} s_i \Pi \left((N-1)\partial_t + \frac{1}{s_i}(\partial_t + \boldsymbol{\mathcal{G}}_{ii}) + \sum_{\ell=N+1}^{q} \frac{1}{s_\ell}(\partial_t + (1-s_\ell)\boldsymbol{\mathcal{G}}_{\ell\ell}) \right) + O(\Delta x^2) \\ &= \Pi(\partial_t + \boldsymbol{\mathcal{G}}_{ii}) + O(\Delta x^2). \end{aligned}$$

To handle the term with the adjugate, observe that the first-order term is made up of the terms

$$(\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i)\boldsymbol{\mathcal{A}}_i^\diamond)^{(1)} = (\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i))^{(0)} (\boldsymbol{\mathcal{A}}_i^\diamond)^{(1)} + (\mathrm{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}_i))^{(1)} (\boldsymbol{\mathcal{A}}_i^\diamond)^{(0)}, \tag{37}$$

and in particular, we are interested in the i-th line of this matrix. Because of the fact that $(\boldsymbol{\mathcal{A}}_i^\diamond)^{(0)} = \mathrm{diag}(1 - s_1, \ldots, 1 - s_{i-1}, 0, 1 - s_{i+1}, \ldots, 1 - s_N, 0, \ldots, 0)$, the i-th line of the second term on the right hand side of Equa-

tion (37) is zero, thus we do not have do study it. For the remaining term, it can be easily seen that

$$(\boldsymbol{\mathcal{A}}_i^{\diamond})^{(1)} = -\frac{1}{\lambda}(\boldsymbol{I}-\boldsymbol{S}) \left(\begin{array}{ccc|c|ccc||ccc} \mathcal{G}_{11} & \cdots & \mathcal{G}_{1(i-1)} & \mathcal{G}_{1i} & \mathcal{G}_{1(i+1)} & \cdots & \mathcal{G}_{1N} & \mathcal{G}_{1(N+1)} & \cdots & \mathcal{G}_{1q} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{G}_{(i-1)1} & \cdots & \mathcal{G}_{(i-1)(i-1)} & \mathcal{G}_{(i-1)i} & \mathcal{G}_{(i-1)(i+1)} & \cdots & \mathcal{G}_{(i-1)N} & \mathcal{G}_{(i-1)(N+1)} & \cdots & \mathcal{G}_{(i-1)q} \\ \hline \mathcal{G}_{i1} & \cdots & \mathcal{G}_{i(i-1)} & 0 & \mathcal{G}_{i(i+1)} & \cdots & \mathcal{G}_{iN} & 0 & \cdots & 0 \\ \hline \mathcal{G}_{(i+1)1} & \cdots & \mathcal{G}_{(i+1)(i-1)} & \mathcal{G}_{(i+1)i} & \mathcal{G}_{(i+1)(i+1)} & \cdots & \mathcal{G}_{(i+1)N} & \mathcal{G}_{(i+1)(N+1)} & \cdots & \mathcal{G}_{(i+1)q} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{G}_{N1} & \cdots & \mathcal{G}_{N(i-1)} & \mathcal{G}_{Ni} & \mathcal{G}_{N(i+1)} & \cdots & \mathcal{G}_{NN} & \mathcal{G}_{N(N+1)} & \cdots & \mathcal{G}_{Nq} \\ \hline\hline \mathcal{G}_{(N+1)1} & \cdots & \mathcal{G}_{(N+1)(i-1)} & 0 & \mathcal{G}_{(N+1)(i+1)} & \cdots & \mathcal{G}_{(N+1)N} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{G}_{q1} & \cdots & \mathcal{G}_{q(i-1)} & 0 & \mathcal{G}_{q(i+1)} & \cdots & \mathcal{G}_{qN} & 0 & \cdots & 0 \end{array}\right),$$

thus we deduce that

$$\left((\mathrm{adj}(\zeta \boldsymbol{I}-\boldsymbol{\mathcal{A}}_i)\boldsymbol{\mathcal{A}}_i^{\diamond})^{(1)}\right)_{i,\cdot} = -\Pi\frac{1}{\lambda}\left((1-s_1)\mathcal{G}_{i1},\ldots,(1-s_{i-1})\mathcal{G}_{i(i-1)},0,(1-s_{i+1})\mathcal{G}_{i(i+1)},\ldots,(1-s_N)\mathcal{G}_{iN},0,\ldots,0\right).$$

Dealing with the zero and first order term in $\mathrm{adj}(\zeta \boldsymbol{I}-\boldsymbol{\mathcal{A}}_i)\boldsymbol{\mathcal{B}}$ works the same than $N=1$, thus we do not repeat it. Moreover, these terms allow for the compensation of the dependence on the choice of the relaxation parameter of the other conserved moments $s_1,\cdots,s_{i-1},s_{i+1},\ldots,s_N$ in the previous equation, as claimed in Section 3.4, thanks to Equation (1).

Putting all the previously discussed facts together into the truncated Equation (16) yields

$$\Pi\frac{\Delta x}{\lambda}\left(\partial_t m_i + \mathcal{G}_{ii}m_i + \sum_{\substack{j=1\\ j\neq i}}^{N}\mathcal{G}_{ij}m_j + \sum_{j=N+1}^{q}\mathcal{G}_{ij}m_j^{\mathrm{eq}}\right) = O(\Delta x^2),$$

which is the result from Theorem 4.4 for $N\geq 1$. Analogous reasonings yield Theorem 4.5.

7 Link with the existing approaches

To finish our contribution, we briefly sketch the links with previous works on the macroscopic equations like [46] and [13, 15]. A more complete study shall be the object of future investigations.

7.1 Equivalent equations

Our results Theorem 4.4 and Theorem 4.5 coincide with the analogous results in [15] up to second order. The substantial difference is that we apply the Taylor expansions to the solution of the corresponding Finite Difference scheme given either by Proposition 3.4 bis or Proposition 3.5 bis, where non-conserved moments have been removed. We therefore reasonably conjecture that the obtained macroscopic equations coincide at any order. The mathematical justification of this conjecture shall be the object of future investigations.

However, the quasi-equilibrium, which is extensively used in [15] can be somehow recovered in our previous discussion. Let $N=1$ to fix ideas. In the proof of Proposition 3.4 bis, nothing prevents us from selecting, instead of the first row, the $i\in[2\,..\,q]$ row, corresponding to a non-conserved moment. This is

$$\det(\mathrm{z}\boldsymbol{I}-\mathbf{A})\,m_i = \left(\mathrm{adj}(\mathrm{z}\boldsymbol{I}-\mathbf{A})\mathbf{B}\boldsymbol{m}^{\mathrm{eq}}\right)_i. \tag{38}$$

Let us stress that even if this could seem to be a viable Finite Difference scheme for the non-conserved variable m_i, it is not independent from the conserved moment m_1 the equilibria depend on and furthermore, this formulation

certainly depends on the choice of s_1, the relaxation parameter of the conserved moment. This is somehow unwanted since s_1 is *in fine* not present in the original lattice Boltzmann scheme. From the computations of Section 5, we see that

$$\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) = s_1 \Pi + O(\Delta x), \qquad \operatorname{adj}(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}) = \Pi \operatorname{diag}\left(1, \frac{s_1}{s_2}, \dots, \frac{s_1}{s_q}\right) + O(\Delta x), \qquad \boldsymbol{\mathcal{B}} = \boldsymbol{S} + O(\Delta x).$$

Using the asymptotic equivalents truncated at leading order in Equation (38) thus provides

$$s_1 \Pi m_i + O(\Delta x) = s_1 \Pi m_i^{\mathrm{eq}} + O(\Delta x), \qquad \text{hence also} \qquad m_i = m_i^{\mathrm{eq}} + O(\Delta x),$$

provided that $s_1 \neq 0$. This is the quasi-equilibrium of the non-conserved moments, which is re-injected in the lattice Boltzmann schemes to eliminate them in the procedure by [15]. However, in our framework, we are not really allowed to write Equation (38).

7.2 Maxwell iteration

In [46], the computations have been carried only for the $\mathrm{D_2Q_9}$ scheme by [31] with $N = 3$. In this part of our work, we are going to develop the computations until third-order for any lattice Boltzmann scheme under acoustic scaling, *i.e.* Assumptions 4.1. In this Section, it is crucial to assume that $\boldsymbol{S} \in \mathrm{GL}_q(\mathbb{R})$. Observe that this assumption ensures that $\det(\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}})$ is a unit (invertible) in the ring $\mathscr{S}$ or equivalently that $\zeta \boldsymbol{I} - \boldsymbol{\mathcal{A}}$ belongs to $\mathrm{GL}_q(\mathscr{S})$. The Maxwell iteration [46] at step $k \in \mathbb{N}$ reads, after simple computations

$$\boldsymbol{m}^{[k]} = \left(\sum_{r=0}^{k} (-\boldsymbol{S}^{-1}(\zeta \overline{\boldsymbol{\mathcal{T}}} - \boldsymbol{I}))^r \right) \boldsymbol{m}^{\mathrm{eq}}, \tag{39}$$

where the quasi-equilibrium is encoded in the choice $\boldsymbol{m}^{[0]} = \boldsymbol{m}^{\mathrm{eq}}$ and where we have taken, as for Equation (14)

$$\overline{\mathsf{T}} := \boldsymbol{M} \operatorname{diag}(\mathsf{t}_{-c_1}, \dots, \mathsf{t}_{-c_q}) \boldsymbol{M}^{-1} \asymp \boldsymbol{M} \left(\sum_{|\boldsymbol{\nu}| \geq 0} \frac{\Delta x^{|\boldsymbol{\nu}|}}{\boldsymbol{\nu}!} \operatorname{diag}\left(\boldsymbol{c}_{\boldsymbol{1}}^{\boldsymbol{\nu}}, \dots, \boldsymbol{c}_{\boldsymbol{q}}^{\boldsymbol{\nu}}\right) \partial^{\boldsymbol{\nu}} \right) \boldsymbol{M}^{-1} =: \overline{\boldsymbol{\mathcal{T}}} \in \mathscr{M}_q(\mathscr{S}).$$

It is easy to see that $\boldsymbol{\mathcal{T}} \overline{\boldsymbol{\mathcal{T}}} = \overline{\boldsymbol{\mathcal{T}}} \boldsymbol{\mathcal{T}} = \boldsymbol{I}$ and moreover, in analogy with Lemma 5.3

$$\zeta \overline{\boldsymbol{\mathcal{T}}} - \boldsymbol{I} = \frac{\Delta x}{\lambda} (\partial_t \boldsymbol{I} + \boldsymbol{\mathcal{G}}) + \frac{\Delta x^2}{2\lambda^2} (\partial_{tt} \boldsymbol{I} + 2\boldsymbol{\mathcal{G}} \partial_t + \boldsymbol{\mathcal{G}}^2) + O(\Delta x^3). \tag{40}$$

The Maxwell iteration works by assuming that $\boldsymbol{m} = \boldsymbol{m}^{[k]} + O(\Delta x^{k+1})$. Taking $k = 1$ in Equation (39) and using Equation (40), we have

$$\boldsymbol{m} = \boldsymbol{m}^{\mathrm{eq}} - \boldsymbol{S}^{-1} \frac{\Delta x}{\lambda} (\partial_t \boldsymbol{I} + \boldsymbol{\mathcal{G}}) \boldsymbol{m}^{\mathrm{eq}} + O(\Delta x^2).$$

Let $i \in [\![1 .. N]\!]$, then taking advantage of Equation (1)

$$m_i = m_i - \frac{\Delta x}{\lambda s_i} \left(\partial_t m_i + \sum_{j=1}^{N} \boldsymbol{\mathcal{G}}_{ij} m_j + \sum_{j=N+1}^{q} \boldsymbol{\mathcal{G}}_{1j} m_j^{\mathrm{eq}} \right) + O(\Delta x^2),$$

which upon division, is the same result than Theorem 4.4. Going up to order two considering $k = 2$, we have

$$\begin{aligned} \boldsymbol{m} = \boldsymbol{m}^{\mathrm{eq}} &- \boldsymbol{S}^{-1} \frac{\Delta x}{\lambda} (\partial_t \boldsymbol{I} + \boldsymbol{\mathcal{G}}) \boldsymbol{m}^{\mathrm{eq}} \\ &+ \frac{\Delta x^2}{2\lambda^2} \boldsymbol{S}^{-1} \left((2\boldsymbol{S}^{-1} - \boldsymbol{I}) \partial_{tt} + 2(\boldsymbol{S}^{-1} \boldsymbol{\mathcal{G}} + \boldsymbol{\mathcal{G}} \boldsymbol{S}^{-1} - \boldsymbol{\mathcal{G}}) \partial_t + \boldsymbol{\mathcal{G}} (2\boldsymbol{S}^{-1} - \boldsymbol{I}) \boldsymbol{\mathcal{G}} \right) \boldsymbol{m}^{\mathrm{eq}} + O(\Delta x^3). \end{aligned}$$

Once more, selecting the i-th row provides

$$\begin{aligned} m_i = m_i - \frac{\Delta x}{\lambda s_i} \Bigg(\partial_t m_i + \gamma_{1,i} + \frac{\Delta x}{\lambda} \Bigg(\left(\frac{1}{s_i} - \frac{1}{2} \right) \partial_{tt} m_i + \sum_{j=1}^{q} \left(\frac{1}{s_i} + \frac{1}{s_j} - 1 \right) \boldsymbol{\mathcal{G}}_{ij} \partial_t m_j^{\mathrm{eq}} + \sum_{j=1}^{q} \sum_{\ell=1}^{q} \left(\frac{1}{s_\ell} - \frac{1}{2} \right) \boldsymbol{\mathcal{G}}_{i\ell} \boldsymbol{\mathcal{G}}_{\ell j} m_j^{\mathrm{eq}} \Bigg) \Bigg) \\ + O(\Delta x^3). \end{aligned}$$

Using relations analogous to Equation (33), Equation (34) and Equation (36) for $N \geq 1$, formally obtained by differentiating the result from Theorem 4.4, we finally obtain, after tedious but elementary computations

$$m_i = m_i - \frac{\Delta x}{\lambda s_i}\left(\partial_t m_i + \gamma_{1,i} + \frac{\Delta x}{\lambda}\sum_{j=N+1}^{q}\left(\frac{1}{s_j}-\frac{1}{2}\right)\mathcal{G}_{ij}\left(\sum_{\ell=1}^{N}\frac{\mathrm{d}m_j^{\mathrm{eq}}}{\mathrm{d}m_\ell}\gamma_{1,\ell} - \sum_{\ell=1}^{N}\mathcal{G}_{j\ell}m_\ell - \sum_{\ell=N+1}^{q}\mathcal{G}_{j\ell}m_\ell^{\mathrm{eq}}\right)\right) + O(\Delta x^3),$$

which coincides with the result from Theorem 4.5. Therefore, up to order two, our approach yields results consistent with those from [46].

To intuitively illustrate why it is reasonable to believe that we recover the same result at any order, let us assume $N = 1$. Then we have, using that $\boldsymbol{S} \in \mathrm{GL}_q(\mathbb{R})$, $\boldsymbol{\mathcal{T}}\overline{\boldsymbol{\mathcal{T}}} = \overline{\boldsymbol{\mathcal{T}}}\boldsymbol{\mathcal{T}} = \boldsymbol{I}$, the rule for the inverse of a product of matrix and the identity relative to geometric series in the context of formal power series, that

$$\begin{aligned}
\mathbf{0} &= \det(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}})\boldsymbol{m} - \mathrm{adj}(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}})\boldsymbol{\mathcal{B}}\boldsymbol{m}^{\mathrm{eq}} = \det(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}})\left(\boldsymbol{m} - (\zeta\boldsymbol{I}-\boldsymbol{\mathcal{T}}(\boldsymbol{I}-\boldsymbol{S}))^{-1}\boldsymbol{\mathcal{T}}\boldsymbol{S}\boldsymbol{m}^{\mathrm{eq}}\right),\\
&= \det(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}})\left(\boldsymbol{m} - (\boldsymbol{S}^{-1}\overline{\boldsymbol{\mathcal{T}}}(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{T}}(\boldsymbol{I}-\boldsymbol{S})))^{-1}\boldsymbol{m}^{\mathrm{eq}}\right) = \det(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}})\left(\boldsymbol{m} - (\boldsymbol{I}+\boldsymbol{S}^{-1}(\zeta\overline{\boldsymbol{\mathcal{T}}}-\boldsymbol{I}))^{-1}\boldsymbol{m}^{\mathrm{eq}}\right),\\
&= \det(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}})\left(\boldsymbol{m} - \left(\sum_{r=0}^{+\infty}(-\boldsymbol{S}^{-1}(\zeta\overline{\boldsymbol{\mathcal{T}}}-\boldsymbol{I}))^r\right)\boldsymbol{m}^{\mathrm{eq}}\right) = \det(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}})\left(\boldsymbol{m} - \lim_{k\to+\infty}\boldsymbol{m}^{[k]}\right).
\end{aligned}$$

Therefore the expansion of the Finite Difference scheme from Proposition 3.4 bis and the non-truncated Maxwell iteration method on the lattice Boltzmann scheme coincide up to a multiplication by a formal power series of time-space differential operators, *i.e.* $\det(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}}) \in \mathscr{S}$. *A priori*, the resulting macroscopic equations are not the same, but since $\det(\zeta\boldsymbol{I}-\boldsymbol{\mathcal{A}}) = \det(\boldsymbol{S}) + O(\Delta x) = s_1\Pi + O(\Delta x)$, thus we "pay" only a constant factor we can divide by at dominant order, the macroscopic equations at leading order are the same. Then, at each order, the result must be the same because we re-inject, in a recursive fashion, the solution truncated at the previous order to eliminate the higher-order time derivatives, see for instance Equation (34) and Equation (36).

8 Conclusions and perspectives

In this original paper, we have rigorously derived the macroscopic equations up to order two for any lattice Boltzmann scheme under acoustic scaling by restating it as a multi-step macroscopic Finite Difference scheme on the conserved moments [3]. Since the passage from the kinetic to the macroscopic standpoint is fully discrete, our analysis can handle any type of time-space scaling and be pushed forward to reach higher orders of accuracy in the discretization parameters. Contrarily to the existing techniques, the quasi-equilibrium of the non-conserved moments in the limit of small discretization parameters is not the key to eliminate the non-conserved variables from the macroscopic equations. The obtained results confirm, going beyond empirical evidence, that the formal Taylor expansion by [13, 15] and the Maxwell iteration by [46] are well-grounded from the perspective of numerical analysts and traditional numerical methods for PDEs, such as Finite Difference.

An improvement of the present work could be the establishment of the equivalence between different consistency analyses for higher orders and ideally for any order. Even if more involved from the standpoint of computations, the extension can be easily done by considering derivatives of higher order for the determinant and adjugate functions, in the spirit of Lemma 5.4 and Lemma 5.6. In this work, all the computations have been done by hand but one could envision to seek some help from symbolic computations. This is a current path of investigation which final aim is to provide the computation – inside the package `pylbm`[5] – of the equivalent equations of any lattice Boltzmann scheme either by the corresponding Finite Difference scheme or using the Maxwell iteration. Another interesting track for future researches could be the re-state of our results using a parabolic scaling between time and space.

Acknowledgments

The author deeply thanks his PhD advisors, M. Massot and B. Graille, for the fruitful discussions and advice on the subject, as well as his brother P. Bellotti for the useful tips to improve the style of manuscript. The author is supported by a PhD funding (year 2019) from the Ecole polytechnique.

[5] `https://pylbm.readthedocs.io`

References

[1] ALLAIRE, G. *Numerical analysis and optimization: an introduction to mathematical modelling and numerical simulation*. Oxford University Press, 2007.

[2] BELLOTTI, T., GOUARIN, L., GRAILLE, B., AND MASSOT, M. High accuracy analysis of adaptive multiresolution-based lattice Boltzmann schemes via the equivalent equations. *arXiv preprint arXiv:2105.13816, Submitted to SMAI Journal of Computational Mathematics* (2021).

[3] BELLOTTI, T., GRAILLE, B., AND MASSOT, M. Finite difference formulation of any lattice Boltzmann scheme. *arXiv preprint arXiv:2201.05354* (2021).

[4] BOUCHUT, F., GUARGUAGLINI, F. R., AND NATALINI, R. Diffusive BGK approximations for nonlinear multidimensional parabolic equations. *Indiana University Mathematics Journal* (2000), 723–749.

[5] BREWER, J. W., BUNCE, J. W., AND VAN VLECK, F. S. *Linear systems over commutative rings*. CRC Press, 1986.

[6] CAIAZZO, A., JUNK, M., AND RHEINLÄNDER, M. Comparison of analysis techniques for the lattice Boltzmann method. *Computers & Mathematics with Applications 58*, 5 (2009), 883–897.

[7] CHAPMAN, S., AND COWLING, T. G. *The mathematical theory of non-uniform gases: an account of the kinetic theory of viscosity, thermal conduction and diffusion in gases*. Cambridge university press, 1990.

[8] CHEN, S., AND DOOLEN, G. D. Lattice Boltzmann method for fluid flows. *Annual Review of Fluid Mechanics 30*, 1 (1998), 329–364.

[9] CHENG, S. S. *Partial difference equations*, vol. 3. CRC Press, 2003.

[10] DELLACHERIE, S. Construction and analysis of lattice Boltzmann methods applied to a 1D convection-diffusion equation. *Acta Applicandae Mathematicae 131*, 1 (2014), 69–140.

[11] D'HUMIÈRES, D. *Generalized Lattice-Boltzmann Equations*. American Institute of Aeronautics and Astronautics, Inc., 1992, pp. 450–458.

[12] DING, J., AND ZHOU, A. Eigenvalues of rank-one updated matrices with some applications. *Applied Mathematics Letters 20*, 12 (2007), 1223–1226.

[13] DUBOIS, F. Equivalent partial differential equations of a lattice Boltzmann scheme. *Computers & Mathematics with Applications 55*, 7 (2008), 1441–1449.

[14] DUBOIS, F. General third order Chapman-Enskog expansion of lattice Boltzmann schemes. In *16th International Conference for Mesoscopic Methods in Engineering and Science, Edinburgh, 22–26 July 2019*. (Edimburgh, United Kingdom, July 2019).

[15] DUBOIS, F. Nonlinear fourth order Taylor expansion of lattice Boltzmann schemes. *Asymptotic Analysis*, 1 Jan. 2021 (2021), 1–41.

[16] DUBOIS, F., AND LALLEMAND, P. Towards higher order lattice Boltzmann schemes. *Journal of Statistical Mechanics: Theory and Experiment 2009*, 06 (2009), P06006.

[17] DUBOIS, F., AND LALLEMAND, P. Quartic parameters for acoustic applications of lattice Boltzmann scheme. *Computers & Mathematics with Applications 61*, 12 (2011), 3404–3416.

[18] FUČÍK, R., AND STRAKA, R. Equivalent finite difference and partial differential equations for the lattice Boltzmann method. *Computers & Mathematics with Applications 90* (2021), 96–103.

[19] GUO, Z., AND SHU, C. *Lattice Boltzmann method and its application in engineering*, vol. 3. World Scientific, 2013.

[20] HÉNON, M. Viscosity of a lattice gas. *Lattice Gas Methods for Partial Differential Equations* (1987), 179–207.

[21] HIGUERA, F. J., AND JIMÉNEZ, J. Boltzmann approach to lattice gas simulations. *EPL (Europhysics Letters) 9*, 7 (1989), 663.

[22] HORN, R. A., AND JOHNSON, C. R. *Matrix analysis*. Cambridge university press, 2012.

[23] HUANG, K. *Statistical Mechanics*, 2 ed. John Wiley & Sons, 1987.

[24] JOHNSON, W. P. The curious history of Faà di Bruno's formula. *The American Mathematical Monthly 109*, 3 (2002), 217–234.

[25] JUNK, M., KLAR, A., AND LUO, L.-S. Asymptotic analysis of the lattice Boltzmann equation. *Journal of Computational Physics 210*, 2 (2005), 676–704.

[26] JUNK, M., AND YANG, Z. Convergence of lattice Boltzmann methods for Navier–Stokes flows in periodic and bounded domains. *Numerische Mathematik 112*, 1 (2009), 65–87.

[27] JUNK, M., AND YONG, W.-A. Rigorous Navier–Stokes limit of the lattice Boltzmann equation. *Asymptotic Analysis 35*, 2 (2003), 165–185.

[28] JURY, E. I. *Theory and Application of the z-Transform Method*. Krieger Publishing Co., 1964.

[29] KASSEL, C. *Quantum Groups*, 1 ed. Graduate Texts in Mathematics. Springer-Verlag New York, 1995.

[30] KRÜGER, T., KUSUMAATMAJA, H., KUZMIN, A., SHARDT, O., SILVA, G., AND VIGGEN, E. M. The lattice Boltzmann method. *Springer International Publishing 10*, 978-3 (2017).

[31] LALLEMAND, P., AND LUO, L.-S. Theory of the lattice Boltzmann method: Dispersion, dissipation, isotropy, Galilean invariance, and stability. *Physical Review E 61*, 6 (2000), 6546.

[32] LANG, S. *Algebra*, 3 ed. Graduate Texts in Mathematics. Springer-Verlag New York, 2002.

[33] MCNAMARA, G. R., AND ZANETTI, G. Use of the Boltzmann equation to simulate lattice-gas automata. *Physical Review Letters 61*, 20 (1988), 2332.

[34] MILLER, K. S. *An Introduction to the Calculus of Finite Differences and Difference Equations*. Dover Publications, 1960.

[35] MONFORTE, A. A., AND KAUERS, M. Formal Laurent series in several variables. *Expositiones Mathematicae 31*, 4 (2013), 350–367.

[36] NIVEN, I. Formal power series. *The American Mathematical Monthly 76*, 8 (1969), 871–889.

[37] QIAN, Y.-H., AND ZHOU, Y. Higher-order dynamics in lattice-based models using the Chapman-Enskog method. *Physical Review E 61*, 2 (2000), 2103.

[38] RHEINLÄNDER, M. K. *Analysis of lattice-Boltzmann methods: asymptotic and numeric investigation of a singularly perturbed system*. PhD thesis, 2007.

[39] ROMAN, S. *The Umbral Calculus*. Dover Publications, 2005.

[40] STEWART, G. On the adjugate matrix. *Linear Algebra and its Applications 283*, 1-3 (1998), 151–164.

[41] STRIKWERDA, J. C. *Finite difference schemes and partial differential equations*. SIAM, 2004.

[42] SUCCI, S. *The lattice Boltzmann equation: for fluid dynamics and beyond*. Oxford University Press, 2001.

[43] SUGA, S. An accurate multi-level finite difference scheme for 1D diffusion equations derived from the lattice Boltzmann method. *Journal of Statistical Physics 140*, 3 (2010), 494–503.

[44] VAN LEEMPUT, P., RHEINLÄNDER, M., AND JUNK, M. Smooth initialization of lattice Boltzmann schemes. *Computers & Mathematics with Applications 58*, 5 (2009), 867–882.

[45] WARMING, R. F., AND HYETT, B. The modified equation approach to the stability and accuracy analysis of finite-difference methods. *Journal of Computational Physics 14*, 2 (1974), 159–179.

[46] YONG, W.-A., ZHAO, W., LUO, L.-S., ET AL. Theory of the lattice Boltzmann method: Derivation of macroscopic equations via the Maxwell iteration. *Physical Review E 93*, 3 (2016), 033310.

[47] ZHAO, W., AND YONG, W.-A. Maxwell iteration for the lattice Boltzmann method with diffusive scaling. *Physical Review E 95*, 3 (2017), 033311.

[48] ZWILLINGER, D. *CRC standard mathematical tables and formulas*. Chapman and Hall - CRC, 2018.