DIFFUSIVE LIMITS OF 2D WELL-BALANCED SCHEMES FOR KINETIC MODELS OF NEUTRON TRANSPORT

Gabriella Bretti, Laurent Gosse, Nicolas Vauchelet

Mathematical Modelling and Numerical Analysis Will be set by the publisher
Modélisation Mathématique et Analyse Numérique

DIFFUSIVE LIMITS OF 2D WELL-BALANCED SCHEMES FOR KINETIC MODELS OF NEUTRON TRANSPORT *

Gabriella Bretti[1], Laurent Gosse[2] and Nicolas Vauchelet[3]

Abstract. Two-dimensional dissipative and isotropic kinetic models, like the ones used in neutron transport theory, are considered. Especially, steady-states are expressed for constant opacity and damping, allowing to derive a scattering S-matrix and corresponding "truly 2D well-balanced" numerical schemes. A first scheme is obtained by directly implementing truncated Fourier-Bessel series, whereas another proceeds by applying an exponential modulation to a former, conservative, one. Consistency with the asymptotic damped parabolic approximation is checked for both algorithms. These findings are confirmed by means of practical benchmarks carried out on coarse Cartesian computational grids.

Résumé. On considere des modeles cinetiques bidimensionnels, dissipatifs et isotropes, comme ceux utilisés en théorie du transport des neutrons. En particulier, les états d'équilibre sont explicités pour des opacités et des amortissements constants, permettant de déduire une S-matrice de scattering et les schémas "2D well-balanced" correspondants. Un premier schéma est obtenu en implémentant directement des séries de Fourier-Bessel tronquées, alors qu'un second procede d'une modulation exponentielle d'un schéma précédemment publié dans le cas conservatif. La consistance avec l'approximation parabolique dissipative est vérifiée pour les deux algorithmes. Ces résultats sont confirmés grâce a des cas-tests pratiques conduits sur des grilles Cartésiennes grossieres.

1991 Mathematics Subject Classification. 65M06, 35K57, 82D75.

xx/xx/2020.

Introduction

Dissipative $(2+2)$-dimensional kinetic models

Consider a damped kinetic model, where $\sigma(\mathbf{x}) \geq 0$ is the opacity,

$$\partial_t f + \mathbf{v} \cdot \nabla f = \sigma(\mathbf{x}) \left(c(\mathbf{x}) \int_{\mathbb{S}^1} f(t, \mathbf{x}, \mathbf{v}') \frac{\mathrm{d}\mathbf{v}'}{2\pi} - f \right), \qquad 0 \leq c := 1 - \kappa \leq 1, \tag{1}$$

Keywords and phrases: Kinetic model of neutron transport, Two-dimensional well-balanced, Asymptotic-preserving scheme, Bessel functions, Laplace transforms, Pizzetti's formula.

* *Support of Italian Minister of Instruction, University and Research (MIUR) through PRIN Project "Innovative numerical methods for evolutionary partial differential equations and applications" #2017KKJP4X is acknowledged.*

[1] Istituto per le Applicazioni del Calcolo, via dei Taurini, 19, 00185 Rome, Italy

[2] Istituto per le Applicazioni del Calcolo, via dei Taurini, 19, 00185 Rome, Italy

[3] Univ. Sorbonne Paris Nord, Labo. Analyse, Géométrie et Applications, CNRS UMR 7539, F-93430, Villetaneuse, France

for $\mathbf{x} = (x, y) \in \mathbb{R}^2$ and $\mathbf{v} = (\cos\theta, \sin\theta)$. In the special case where $c(\mathbf{x}) \equiv 1$, the model studied in [24] is clearly recovered, hence our intention hereafter is to extend these findings toward the more general model (1). We shall be especially interested in its "four-stream approximation" [13, 25] with diagonal microscopic velocities,

$$\mathbf{v} \in \left\{ \frac{\pm 1}{\sqrt{2}}(1,1), \frac{\pm 1}{\sqrt{2}}(-1,1) \right\},$$

which yield the following system, (in standard notation)

$$\begin{aligned} \partial_t f^\pm \pm \frac{1}{\sqrt{2}}(\partial_x f^\pm + \partial_y f^\pm) &= \sigma(\mathbf{x})\big(c(\mathbf{x})\rho - f^\pm\big), \\ \partial_t g^\pm \pm \frac{1}{\sqrt{2}}(\partial_x g^\pm - \partial_y g^\pm) &= \sigma(\mathbf{x})\big(c(\mathbf{x})\rho - g^\pm\big), \\ \rho &:= f^+ + f^- + g^+ + g^-. \end{aligned} \tag{2}$$

This discrete kinetic model, within a convenient rescaling of variables, relaxes towards the damped heat equation,

$$\partial_t \rho + \sigma(\mathbf{x})\kappa(\mathbf{x})\rho = \operatorname{div}\left(\frac{\nabla\rho}{2\sigma(\mathbf{x})}\right), \qquad \text{or } = \frac{\Delta\rho}{2\sigma}, \text{ if } \sigma \text{ is a positive constant.} \tag{3}$$

Indeed, assuming

$$(t, \mathbf{x}) \to (\varepsilon^2 t, \varepsilon\mathbf{x}), \qquad \kappa \to \varepsilon^2\kappa,$$

and rotating the axes for simplicity, it comes

$$\varepsilon\partial_t f^\pm \pm \partial_x f^\pm = \frac{\sigma(\mathbf{x})}{\varepsilon}\Big((1 - \varepsilon^2\kappa(\mathbf{x}))\rho - f^\pm\Big), \qquad \varepsilon\partial_t g^\pm \pm \partial_y g^\pm = \frac{\sigma(\mathbf{x})}{\varepsilon}\Big((1 - \varepsilon^2\kappa(\mathbf{x}))\rho - g^\pm\Big).$$

Adding the four microscopic balance laws,

$$\partial_t \rho + \operatorname{div}\mathbf{J} + \kappa(\mathbf{x})\rho = 0, \qquad \mathbf{J} = \frac{1}{\varepsilon}\begin{pmatrix} f^+ - f^- \\ g^+ - g^- \end{pmatrix},$$

and subtracting the first (second) with the third (fourth) ones,

$$\varepsilon\partial_t(f^\pm - g^\pm) \pm (\partial_x f^\pm - \partial_y g^\pm) = -\frac{\sigma(\mathbf{x})}{\varepsilon}(f^\pm - g^\pm).$$

This last equation implies that, formally, $|f^\pm - g^\pm| = O(\varepsilon)$ and (3) holds: see [25, §5] for related rigorous results.

Plan of the paper

Following the roadmap proposed in [24], we intend to study numerical approximations of (2) endowed with both 2D well-balanced (WB) and asymptotic-preserving (AP) properties. To this end, two distinct numerical processes will be introduced: the first one, given in §2, strongly relies on the data of two-dimensional steady-states for (1). Such steady-states are derived in §1, following original ideas given in [5, 6]: in both papers, it is shown that solutions of stationary elliptic equations yield, thanks to Laplace transforms, microscopic steady-states of related kinetic models. A second numerical scheme is proposed in §3, being essentially an exponential modulation of the one given in [24] for the special case $\kappa \equiv 0$. Its WB and AP properties are studied, in the light of what was previously done in §2. A main difference between both schemes is that only the second one appears to be unconditionally positivity-preserving; numerical tests displayed in §4 reveal that it is slightly more diffusive, though, especially in kinetic regime with stiff parameters. Concluding remarks are given in §4.2.

1. Two-dimensional kinetic steady-states

1.1. **Conservative isotropic scattering and diffusion**

Consider first, for $\mathbf{x} = (x, y) \in \mathbb{R}^2$ and $\mathbf{v} = (\cos\theta, \sin\theta)$, the special case of (1) where $c(\mathbf{x}) \equiv 1$,

$$\partial_t f(t, \mathbf{x}, \mathbf{v}) + \mathbf{v} \cdot \nabla f = \sigma(\mathbf{x}) \left(\int_{\mathbb{S}^1} f(t, \mathbf{x}, \mathbf{v}') \frac{\mathrm{d}\mathbf{v}'}{2\pi} - f \right), \qquad \sigma(\mathbf{x}) \geq 0, \tag{4}$$

The heart of the matter in [5] (see also [12]) consists in showing that, by combining the method of characteristics

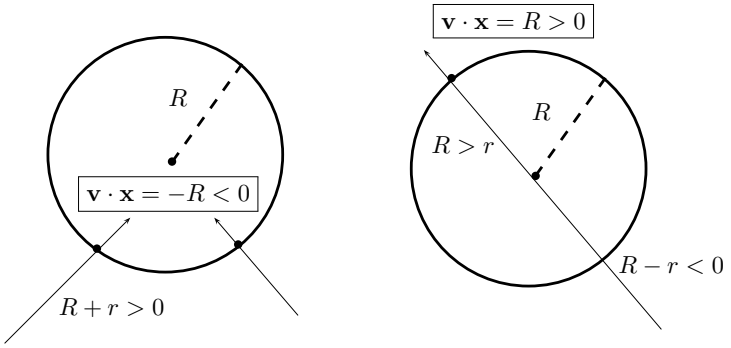

Figure 1. Incoming (left) and outgoing (right) state $f(\mathbf{x}, \mathbf{v})$

and the Laplace transform $\mathcal{L}[\cdot](p)$, stationary (microscopic) solutions of (4) can be retrieved from harmonic (i.e., stationary, macroscopic and smooth) functions $\rho(\mathbf{x})$:

$$f(\mathbf{x}, \mathbf{v}) = \int_0^\infty \exp(-r) \rho(\mathbf{x} - r\mathbf{v}) \, \mathrm{d}r = \mathcal{L}_r[\rho(\mathbf{x} - r\mathbf{v})](p = 1), \qquad \Delta\rho = 0. \tag{5}$$

Indeed, when $\sigma \equiv 1$, the method of characteristics yields, for any $\mathbf{v} \in \mathbb{S}^1$,

$$\frac{d}{ds}\Big(f(s, \mathbf{x} + s\mathbf{v}, \mathbf{v}) \exp(s) \Big) = \rho(s, \mathbf{x} + s\mathbf{v}) \exp(s), \quad \rho(s, \mathbf{x}) = \int_{\mathbb{S}^1} f(s, \mathbf{x}, \mathbf{v}') \frac{d\mathbf{v}'}{2\pi},$$

so that, by integrating on $s \in (0, t)$,

$$f(t, \mathbf{x}, \mathbf{v}) = f(0, \mathbf{x} - t\mathbf{v}, \mathbf{v}) \exp(-t) + \int_0^t \rho(\mathbf{x} - r\mathbf{v}) \exp(-r) \, \mathrm{d}r \to \mathcal{L}_r[\rho(\mathbf{x} - r\mathbf{v})](p = 1), \qquad t \to +\infty.$$

One way to justify that ρ is harmonic, goes as follows: by integrating in $\mathbf{v}$ and invoking Pizzetti's formula [14,26],

$$\begin{aligned}
\rho(\mathbf{x}) &= \int_{\mathbb{S}^1} f(\mathbf{x}, \mathbf{v}) \frac{\mathrm{d}\mathbf{v}}{2\pi} \\
&= \int_0^\infty \exp(-r) \left(\int_{\mathbb{S}^1} \rho(\mathbf{x} - r\mathbf{v}) \frac{\mathrm{d}\mathbf{v}}{2\pi} \right) \mathrm{d}r \\
&= \int_0^\infty \exp(-r) \left(\rho(\mathbf{x}) + \sum_{m \geq 1} \frac{\Delta^m \rho(\mathbf{x})}{\alpha_m} r^{2m} \right) \mathrm{d}r, \qquad 0 < \alpha_m, \text{ depending on dimension 2,} \\
&= \rho(\mathbf{x}) \int_0^\infty \exp(-r) \, \mathrm{d}r = \rho(\mathbf{x}) \Gamma(1),
\end{aligned}$$

and using that, for $\rho \in C^\infty(\mathbb{R}^2)$ harmonic, every term in the series vanishes. Accordingly, in order to build a "2D-WB-AP scheme" for (4), a 4×4 S-matrix is derived, which relates 4 "outgoing states" to 4 "incoming states" (which are available data): see Fig. 2. Such a S-matrix can be seen as the restriction to a finite set of velocities of a "continuous scattering operator" $\mathcal{S}$, defined on any circle of radius $R > 0$:

$$\boxed{\mathcal{S} : f\Big(R(\cos\theta, \sin\theta); -(\cos\theta, \sin\theta)\Big) \mapsto f\Big(R(\cos\theta, \sin\theta); (\cos\theta, \sin\theta)\Big).} \tag{6}$$

By definition, we call an "incoming" (respectively "outgoing") state, any state $f(\mathbf{x}, \mathbf{v})$ such that $\mathbf{x} \cdot \mathbf{v} = -R < 0$ (respectively $\mathbf{x} \cdot \mathbf{v} = R > 0$), see Fig. 1. By (5), they are,

$$f\Big(R(\cos\theta, \sin\theta); \mp(\cos\theta, \sin\theta)\Big) = \int_0^\infty \exp(-r)\rho\Big((R \pm r)(\cos\theta, \sin\theta)\Big)\, \mathrm{d}r, \qquad \theta \in (0, 2\pi).$$

Since any macroscopic steady-state $\rho(\mathbf{x})$ is harmonic, it reads in polar coordinates,

$$\begin{aligned} \rho(r, \theta) &= a_0 + \sum_{n \in \mathbb{N}_*} r^n \left(a_n \cos n\theta + b_n \sin n\theta\right) \\ &= a_0 + a_1\, x + b_1\, y + a_2(x^2 - y^2) + b_2\, xy + ..., \end{aligned}$$

and this determines the (first four) Fourier coefficients of the resulting f,

$$\begin{aligned} f\Big(R(\cos\theta, \sin\theta); \mp(\cos\theta, \sin\theta)\Big) &= \int_0^\infty \exp(-r)\rho\Big((R \pm r)(\cos\theta, \sin\theta)\Big)\, \mathrm{d}r. \\ &= \int_0^\infty \exp(-r)\Big[a_0 + (R \pm r)(a_1 \cos\theta + b_1 \sin\theta) + a_2(R \pm r)^2 \cos 2\theta\Big]\, \mathrm{d}r \\ &= \Gamma(1)\Big[a_0 + R(a_1 \cos\theta + b_1 \sin\theta) + R^2 a_2 \cos 2\theta] \\ &\qquad \pm \Gamma(2)\Big[(a_1 \cos\theta + b_1 \sin\theta) + 2Ra_2 \cos 2\theta\Big] + \Gamma(3) a_2 \cos 2\theta \\ &= a_0 + (R \pm 1)(a_1 \cos\theta + b_1 \sin\theta) + (R^2 \pm 2R + 2) a_2 \cos 2\theta, \end{aligned} \tag{7}$$

being $\Gamma(n+1) = n!$ the Gamma function. Accordingly, any (obviously periodic) "incoming state"

$$f(R\mathbf{v}, -\mathbf{v}) = f(R(\cos\theta, \sin\theta), -(\cos\theta, \sin\theta)), \qquad \theta \in (0, 2\pi),$$

rewrites as a Fourier series with (real) coefficients $(A_n, B_n)_n$, and, by identifying successive coefficients in (7),

$$A_0 = a_0,\ A_1 = (R+1)a_1,\ B_1 = (R+1)b_1,\ A_2 = (R^2 + 2R + 2)a_2,$$

which expression is very similar to the columns of matrix M, see [24, eqn (3.3)]. The first Fourier components of the corresponding "outgoing state" follow,

$$\begin{aligned} f(R\mathbf{v}, \mathbf{v}) &= f\Big(R(\cos\theta, \sin\theta); (\cos\theta, \sin\theta)\Big) \\ &= A_0 + \frac{R-1}{1+R}(A_1 \cos\theta + B_1 \sin\theta) + \frac{2 - 2R + R^2}{2 + 2R + R^2} A_2 \cos 2\theta + ... \end{aligned} \tag{8}$$

The scattering $\mathcal{S}$, non-local in physical space, becomes local[1] in Fourier space for isotropic collisions.

[1] The Fourier coordinates "diagonalize" the discrete scattering operator.

1.2. **Dissipative isotropic scattering with adsorption**

In the original model (1), the expression of the stationary regimes is obtained (by passing to the limit $t \to +\infty$ in the method of characteristics), see *e.g.*, [11, page 34] as a Laplace transform (again, letting $\sigma = 1$), see [6]:

$$f(\mathbf{x},\mathbf{v}) = c\int_0^\infty \exp(-r)\rho(\mathbf{x}-r\mathbf{v})\mathrm{d}r, \qquad \rho(\mathbf{x}) = \int_{\mathbb{S}^1} f(t,\mathbf{x},\mathbf{v})\frac{\mathrm{d}\mathbf{v}}{2\pi}.$$

Again, another integration in $\mathbf{v}$ produces a Fredholm integral equation on ρ,

$$\rho(\mathbf{x}) = c\int_0^\infty \exp(-r)\left(\int_{\mathbb{S}^1}\rho(\mathbf{x}-r\mathbf{v})\frac{\mathrm{d}\mathbf{v}}{2\pi}\right)\mathrm{d}r, \tag{9}$$

in which the circular integral on $\rho(\mathbf{x}-r\mathbf{v})$ will reduce to the pointwise value $\rho(\mathbf{x})$ if ρ solves a convenient elliptic differential operator [15, 29]. Yet, the diffusive approximation suggests the "modified Helmholtz problem",

$$-\Delta\rho + \lambda^2\rho = 0, \tag{10}$$

which solution satisfies (see *e.g.* [3, §2.4] for more details on Green and modified Bessel functions),

$$\rho(\mathbf{x}) = \frac{1}{\mathcal{I}_0(\lambda r)}\int_{\mathbb{S}^1}\rho(\mathbf{x}-r\mathbf{v})\frac{\mathrm{d}\mathbf{v}}{2\pi}, \qquad r>0. \tag{11}$$

By inserting (11) into (9), one gets that the constant λ (different of the 1d case, see [7]) must satisfy:

$$1 = c\int_0^\infty \exp(-r)\,\mathcal{I}_0(\lambda r)\mathrm{d}r = c\,\mathcal{L}_r[\mathcal{I}_0(\lambda r)](p=1) = \frac{c}{\sqrt{1-\lambda^2}}, \qquad \lambda^2 = 1-c^2 = \kappa(1+c).$$

Being a solution to (10), the macroscopic density ρ (expressed in polar coordinates and separating variables) rewrites as a Fourier-Bessel series in any disk of radius $R>0$,

$$\rho(r,\theta) = a_0\mathcal{I}_0(\lambda r) + \sum_{n\in\mathbb{N}_*}\mathcal{I}_n(\lambda r)\big(a_n\cos n\theta + b_n\sin n\theta\big), \qquad r\in(0,R),$$

so that the corresponding stationary microscopic density $f(\mathbf{x},\mathbf{v})$ follows by computing Laplace transforms of Bessel functions because trigonometric functions depend only on θ (and $\mathbf{v} = (\cos\theta,\sin\theta)\in\mathbb{S}^1$).

$$f(R\mathbf{v},\mp\mathbf{v}) = c\left(a_0X_0^\pm(\lambda R) + \sum_{n\in\mathbb{N}_*}X_n^\pm(\lambda R)\big(a_n\cos n\theta + b_n\sin n\theta\big)\right), \qquad X_n^\pm(\lambda R) = \mathcal{L}_r\big(\mathcal{I}_n(\lambda(R\pm r))\big)[1]. \tag{12}$$

To proceed, the "summation formula",

$$\forall (n,x,y)\in\mathbb{N}\times\mathbb{R}^2, \qquad \mathcal{I}_n(x+y) = \sum_{k\in\mathbb{Z}}\mathcal{I}_{n-k}(x)\,\mathcal{I}_k(y), \text{ with } \mathcal{I}_{-n}(x) = \mathcal{I}_n(x), \tag{13}$$

is recalled [2], along with usual Laplace transforms of Bessel functions:

$$\forall n\in\mathbb{N}, \qquad \mathcal{L}_r[\mathcal{I}_n(\lambda r)](p) = \frac{1}{\sqrt{p^2-\lambda^2}}\left(\frac{\lambda}{p+\sqrt{p^2-\lambda^2}}\right)^n.$$

As $p=1$, an interesting relation comes out,

$$\boxed{\mathcal{L}_r[\mathcal{I}_n(\lambda r)](p=1)=\frac{1}{\sqrt{1-\lambda^2}}\left(\frac{\lambda}{1+\sqrt{1-\lambda^2}}\right)^n=\frac{\lambda^n}{c(1+c)^n},}$$

so that the general term proceeds by the summation formula (13),

$$\begin{aligned}\forall n\in\mathbb{N},\qquad X_n^{\pm}(R)&=\int_0^\infty \exp(-r)\mathcal{I}_n\big(\lambda(R\pm r)\big)\mathrm{d}r=\mathcal{L}_r\big(\mathcal{I}_n(\lambda(R\pm r))\big)[1].\\
&=\sum_{k\in\mathbb{Z}}\left(\int_0^\infty \exp(-r)\mathcal{I}_{|n-k|}(\pm\lambda r)\mathrm{d}r\right)\mathcal{I}_k(\lambda R)\\
&=\sum_{k\in\mathbb{Z}}\mathcal{L}_r(\mathcal{I}_{|n-k|}(\pm\lambda r))[p=1]\,\mathcal{I}_k(\lambda R)\\
&=\mathcal{L}_r(\mathcal{I}_0(\pm\lambda r))[1]\sum_{k\in\mathbb{Z}}\left(\frac{\pm\lambda}{1+\sqrt{1-\lambda^2}}\right)^{|n-k|}\mathcal{I}_k(\lambda R)\\
&=\frac{1}{c}\sum_{k\in\mathbb{Z}}\left(\frac{\pm\lambda}{1+c}\right)^{|n-k|}\mathcal{I}_k(\lambda R)\\
&=\frac{1}{c}\left\{\left(\frac{\pm\lambda}{1+c}\right)^n\mathcal{I}_0(\lambda R)+\sum_{k\in\mathbb{N}^*}\left[\left(\frac{\pm\lambda}{1+c}\right)^{|n-k|}+\left(\frac{\pm\lambda}{1+c}\right)^{n+k}\right]\mathcal{I}_k(\lambda R)\right\}.\end{aligned}$$

For instance, since $\lambda^2=1-c^2=\kappa(1+c)$, we get the expression of the first terms in (12):

$$\begin{aligned}X_0^{\pm}(R)&=\frac{1}{c}\left(\mathcal{I}_0(\lambda R)\pm\frac{2\lambda}{1+c}\mathcal{I}_1(\lambda R)+\frac{2\lambda^2}{(1+c)^2}\mathcal{I}_2(\lambda R)+....\right),\\
X_1^{\pm}(R)&=\frac{1}{c}\left(\mathcal{I}_1(\lambda R)\left[1+\frac{\lambda^2}{(1+c)^2}\right]\pm\frac{\lambda\mathcal{I}_0(\lambda R)}{1+c}+....\right),\\
X_2^{\pm}(R)&=\frac{1}{c}\left(\mathcal{I}_2(\lambda R)\pm\frac{\lambda\mathcal{I}_1(\lambda R)}{(1+c)}\left[1+\frac{\lambda^2}{(1+c)^2}\right]+\frac{\lambda^2\mathcal{I}_0(\lambda R)}{(1+c)^2}+....\right).\end{aligned}\tag{14}$$

For $\kappa=1$, a well-known formula [2, (9.6.37-8)] implies that $X_0^{\pm}(R)$ has an exponential behavior,

$$\forall R\geq 0,\qquad \exp(\pm R)=\mathcal{I}_0(R)+2\sum_{k\in\mathbb{N}^*}(\pm 1)^k\,\mathcal{I}_k(R),$$

but in the general case $0<\kappa<1$, we simplify the damping factor $c>0$ and deduce from (14):

$$\begin{aligned}f\Big(R(\cos\theta,\sin\theta);\mp(\cos\theta,\sin\theta)\Big)&\simeq a_0\left[\mathcal{I}_0(\lambda R)\pm\frac{2\lambda}{1+c}\mathcal{I}_1(\lambda R)+\frac{2\kappa}{1+c}\mathcal{I}_2(\lambda R)\right]\\
&+\lambda(a_1\cos\theta+b_1\sin\theta)\left[\frac{\mathcal{I}_1(\lambda R)}{\lambda}(1+\frac{\kappa}{1+c})\pm\frac{\mathcal{I}_0(\lambda R)}{1+c}\right]\\
&+\lambda^2 a_2\cos 2\theta\left[\frac{\mathcal{I}_2(\lambda R)}{\lambda^2}\pm\frac{\mathcal{I}_1(\lambda R)}{\lambda(1+c)}(1+\frac{\kappa}{1+c})+\frac{\mathcal{I}_0(\lambda R)}{(1+c)^2}\right],\end{aligned}\tag{15}$$

which is the generalization of (7) when the dissipation parameter $\kappa\neq 0$.

Remark 1. *Both (7) and (15) express steady-states only in a special case where* $\mathbf{x} = R(\cos\theta, \sin\theta)$ *and* $\mathbf{v} = \pm(\cos\theta, \sin\theta)$ *are colinear. This is enough to build up the* S*-matrix (18) and to derive the corresponding 2D-WB scheme in the form (19). However, this is less general than what was found in [24, eq. (3.1)], where explicit steady-states were given for arbitrary arguments* $(\mathbf{x}, \mathbf{v}) \in \mathbb{R}^2 \times \mathbb{S}^1$.

To account for a (locally) constant opacity $\sigma > 0$, it suffices to rescale $R \to \sigma R$ in (15):

Lemma 1. *Let* $\bar{f}(\mathbf{x}, \mathbf{v})$ *be a steady-state of (1) with* $\sigma = 1$*, then* $\bar{f}(\sigma\mathbf{x}, \mathbf{v})$ *is a steady-state of (1) with any constant* $\sigma \in \mathbb{R}_*^+$

Proof. Let f_σ be a solution of $\mathbf{v}\cdot\nabla f_\sigma(\mathbf{x}, \mathbf{v}) = \sigma\, L[f_\sigma]$. Being the constant $\sigma > 0$, this yields that $\mathbf{v}\cdot f_\sigma(\mathbf{x}/\sigma, \mathbf{v}) = L[f_\sigma]$ and so, in particular, $f_\sigma(R/\sigma, \mathbf{v}) = \bar{f}(R, \mathbf{v})$ □

2. Two-dimensional well-balanced scheme

Hereafter, a uniform Cartesian grid is used, with $\Delta x = \Delta y$ and $\Delta t > 0$ a time-step. Standard notation is

$$\forall (n, i, j) \in \mathbb{N} \times \mathbb{Z}^2, \qquad f_{i,j}^{\pm,n} \simeq f^{\pm}(t^n = n\Delta t, i\Delta x, j\Delta y), \quad g_{i,j}^{\pm,n} \simeq g^{\pm}(t^n = n\Delta t, i\Delta x, j\Delta y).$$

2.1. Derivation of the dissipative S-matrix

Following the same procedure as [24, §3.1], a 2D S-matrix, $S_\kappa = \tilde{M}\, M^{-1}$, is deduced from the expression of kinetic (both incoming and outgoing) steady-states (15), where

$$M = \begin{pmatrix} X_0^+(R) & -X_1^+(R) & 0 & X_2^+(R) \\ X_0^+(R) & X_1^+(R) & 0 & X_2^+(R) \\ X_0^+(R) & 0 & -X_1^+(R) & -X_2^+(R) \\ X_0^+(R) & 0 & X_1^+(R) & -X_2^+(R) \end{pmatrix},$$

a 4×4 matrix which columns are orthogonal to each other, and

$$\tilde{M} = \begin{pmatrix} X_0^-(R) & X_1^-(R) & 0 & X_2^-(R) \\ X_0^-(R) & -X_1^-(R) & 0 & X_2^-(R) \\ X_0^-(R) & 0 & X_1^-(R) & -X_2^-(R) \\ X_0^-(R) & 0 & -X_1^-(R) & -X_2^-(R) \end{pmatrix}.$$

A Fourier series (15) of both incoming/outgoing states yields an analogue of (8),

$$\begin{aligned} &f\Big(R(\cos\theta, \sin\theta); (\cos\theta, \sin\theta)\Big) \\ &\quad = \frac{X_0^-(R)}{X_0^+(R)} A_0 + \frac{X_1^-(R)}{X_1^+(R)} (A_1 \cos\theta + B_1 \sin\theta) + \frac{X_2^-(R)}{X_2^+(R)} A_2 \cos 2\theta + ..., \end{aligned} \tag{16}$$

which gives back (8) when $\kappa \to 0$, so that $c \to 1$, because

$$\forall n \in \mathbb{N}, \qquad \mathcal{I}_n(\lambda R) \to \frac{1}{n!} \left(\frac{\lambda R}{2}\right)^n, \qquad \lambda R \to 0. \tag{17}$$

In order to express the S-matrix, the inverse of M is needed,

$$M^{-1} = \begin{pmatrix} \frac{1}{4X_0^+} & \frac{1}{4X_0^+} & \frac{1}{4X_0^+} & \frac{1}{4X_0^+} \\ \frac{-1}{2X_1^+} & \frac{1}{2X_1^+} & 0 & 0 \\ 0 & 0 & \frac{-1}{2X_1^+} & \frac{1}{2X_1^+} \\ \frac{1}{4X_2^+} & \frac{1}{4X_2^+} & \frac{-1}{4X_2^+} & \frac{-1}{4X_2^+} \end{pmatrix},$$

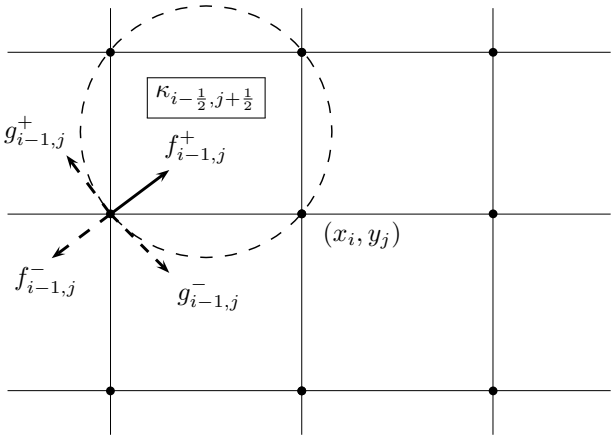

FIGURE 2. The S-matrix $(S_\kappa)_{i-\frac{1}{2},j+\frac{1}{2}}$ and an incoming state, $f^+_{i-1,j}$.

so that, by setting $Y_0 = X_0^-/X_0^+$, and so on. A discrete version of (6) is deduced,

$$S_\kappa = \tilde{M}\, M^{-1} = \frac{1}{4}\begin{pmatrix} Y_0 - 2Y_1 + Y_2 & Y_0 + 2Y_1 + Y_2 & Y_0 - Y_2 & Y_0 - Y_2 \\ Y_0 + 2Y_1 + Y_2 & Y_0 - 2Y_1 + Y_2 & Y_0 - Y_2 & Y_0 - Y_2 \\ Y_0 - Y_2 & Y_0 - Y_2 & Y_0 - 2Y_1 + Y_2 & Y_0 + 2Y_1 + Y_2 \\ Y_0 - Y_2 & Y_0 - Y_2 & Y_0 + 2Y_1 + Y_2 & Y_0 - 2Y_1 + Y_2 \end{pmatrix}, \tag{18}$$

(see Fig. 2) where each function $Y_i(R)$ rewrites as,

$$\begin{aligned}
Y_0(R) &= \frac{\mathcal{I}_0(\lambda R) - \frac{2\lambda}{1+c}\mathcal{I}_1(\lambda R) + \frac{2\kappa}{1+c}\mathcal{I}_2(\lambda R)}{\mathcal{I}_0(\lambda R) + \frac{2\lambda}{1+c}\mathcal{I}_1(\lambda R) + \frac{2\kappa}{1+c}\mathcal{I}_2(\lambda R)} \\
&= 1 - \frac{4\lambda\mathcal{I}_1(\lambda R)}{(1+c)\mathcal{I}_0(\lambda R) + 2\lambda\mathcal{I}_1(\lambda R) + 2\kappa\mathcal{I}_2(\lambda R)}, \\
Y_1(R) &= \frac{\frac{\mathcal{I}_1(\lambda R)}{\lambda}(1 + \frac{\kappa}{1+c}) - \frac{\mathcal{I}_0(\lambda R)}{1+c}}{\frac{\mathcal{I}_1(\lambda R)}{\lambda}(1 + \frac{\kappa}{1+c}) + \frac{\mathcal{I}_0(\lambda R)}{1+c}} \\
&= 1 - \frac{2\lambda\mathcal{I}_0(\lambda R)}{(1+c+\kappa)\mathcal{I}_1(\lambda R) + \lambda\mathcal{I}_0(\lambda R)}, \qquad \text{(notice that: } 1+c+\kappa = 2\text{)} \\
Y_2(R) &= \frac{\frac{\mathcal{I}_2(\lambda R)}{\lambda^2} - \frac{\mathcal{I}_1(\lambda R)}{\lambda(1+c)}(1 + \frac{\kappa}{1+c}) + \frac{\mathcal{I}_0(\lambda R)}{(1+c)^2}}{\frac{\mathcal{I}_2(\lambda R)}{\lambda^2} + \frac{\mathcal{I}_1(\lambda R)}{\lambda(1+c)}(1 + \frac{\kappa}{1+c}) + \frac{\mathcal{I}_0(\lambda R)}{(1+c)^2}} \\
&= 1 - \frac{4\lambda\mathcal{I}_1(\lambda R)}{(1+c)^2\mathcal{I}_2(\lambda R) + 2\lambda\mathcal{I}_1(\lambda R) + \lambda^2\mathcal{I}_0(\lambda R)},
\end{aligned}$$

2.2. **2D well-balanced scheme and consistency**

Accordingly, one gets the following time-marching scheme (similar to [24, eq. (3.5)]),

$$\begin{pmatrix} f^{+,n+1}_{i,j+1} \\ f^{-,n+1}_{i-1,j} \\ g^{+,n+1}_{i-1,j+1} \\ g^{-,n+1}_{i,j} \end{pmatrix} = \left(1-\frac{\Delta t}{2R}\right) \begin{pmatrix} f^{+,n}_{i,j+1} \\ f^{-,n}_{i-1,j} \\ g^{+,n}_{i-1,j+1} \\ g^{-,n}_{i,j} \end{pmatrix} + \frac{\Delta t}{2R} \left(S_\kappa\right)_{i-\frac12,j+\frac12} \begin{pmatrix} f^{+,n}_{i-1,j} \\ f^{-,n}_{i,j+1} \\ g^{+,n}_{i,j} \\ g^{-,n}_{i-1,j+1} \end{pmatrix}, \tag{19}$$

which naturally preserves any $2D$ stationary regime locally expressed as (15).

Theorem 1 (2D well-balanced)**.** *Let $\Delta t \le 2R$ and λR be small enough, then (19)–(18) is consistent with (2) and dissipates both L^1 and L^∞ norms. Moreover, it is "2D well-balanced" in the following sense: any steady-state locally expressed like (15) is invariant by (19)–(18).*

Proof. Under the CFL restriction, the marching scheme (19) is a convex combination, and preserves positivity as soon as all the entries of S_κ are nonnegative, hence λR small enough. The dissipation of L^1 and L^∞ norms is a consequence of the symmetry of S_κ where both lines and columns add to less than one. For consistency, when $\lambda R \ll 1$, by (17), it comes

$$\begin{aligned}
Y_0 &= \frac{X_0^-(R)}{X_0^+(R)} = \frac{\mathcal{I}_0(\lambda R) - \frac{2\lambda}{1+c}\mathcal{I}_1(\lambda R) + \frac{2\kappa}{1+c}\mathcal{I}_2(\lambda R)}{\mathcal{I}_0(\lambda R) + \frac{2\lambda}{1+c}\mathcal{I}_1(\lambda R) + \frac{2\kappa}{1+c}\mathcal{I}_2(\lambda R)} \\
&\simeq \frac{1-\kappa R}{1+\kappa R} \simeq 1 - 2\kappa R, \\
Y_1 &= \frac{X_1^-(R)}{X_1^+(R)} = \frac{\frac{\mathcal{I}_1(\lambda R)}{\lambda}(1+\frac{\kappa}{1+c}) - \frac{\mathcal{I}_0(\lambda R)}{1+c}}{\frac{\mathcal{I}_1(\lambda R)}{\lambda}(1+\frac{\kappa}{1+c}) + \frac{\mathcal{I}_0(\lambda R)}{1+c}} \simeq \frac{R-1}{R+1} \quad \text{(like conservative case } \kappa = 0\text{)}, \\
Y_2 &= \frac{X_2^-(R)}{X_2^+(R)} = \frac{\frac{\mathcal{I}_2(\lambda R)}{\lambda^2} - \frac{\mathcal{I}_1(\lambda R)}{\lambda(1+c)}(1+\frac{\kappa}{1+c}) + \frac{\mathcal{I}_0(\lambda R)}{(1+c)^2}}{\frac{\mathcal{I}_2(\lambda R)}{\lambda^2} + \frac{\mathcal{I}_1(\lambda R)}{\lambda(1+c)}(1+\frac{\kappa}{1+c}) + \frac{\mathcal{I}_0(\lambda R)}{(1+c)^2}} \\
&\simeq \frac{(1+c)^2R^2 - 8R + 8}{(1+c)^2R^2 + 8R + 8} = \frac{R^2 - 2R + 2}{R^2 + 2R + 2} + O(R^2) \quad \text{(like conservative case } \kappa = 0\text{)},
\end{aligned} \tag{20}$$

so that both Y_1, Y_2 behave like C, D in [24, eq. (3.4)], and

$$S_\kappa \simeq \mathrm{Id} - \frac{\kappa R}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} + \frac{R}{2(1+R)}\begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix},$$

which yields that (19) is consistent with (2) as $R \to 0$. To establish the "2D well-balanced" property, notice first that, in a given cell centered in $x_{i-\frac12}, y_{j+\frac12}$ a sufficient condition for a state to be invariant by (19) is,

$$\begin{pmatrix} f^{+,n}_{i,j+1} \\ f^{-,n}_{i-1,j} \\ g^{+,n}_{i-1,j+1} \\ g^{-,n}_{i,j} \end{pmatrix} = \left(S_\kappa\right)_{i-\frac12,j+\frac12} \begin{pmatrix} f^{+,n}_{i-1,j} \\ f^{-,n}_{i,j+1} \\ g^{+,n}_{i,j} \\ g^{-,n}_{i-1,j+1} \end{pmatrix},$$

for some "frozen values" $\bar\sigma = \sigma_{i-\frac12,j+\frac12}, \kappa_{i-\frac12,j+\frac12} \in (\mathbb{R}^+)^2$ in the prescribed computational cell. Assume now that there locally exists a truncated steady-state of the form (15),

$$\bar f(\bar\sigma R(\cos\theta,\sin\theta),\pm(\cos\theta,\sin\theta)), \qquad \theta\in(0,2\pi),$$

which samples for $\theta_k = \frac{\pi}{4} + \frac{k\,\pi}{2}$, $k\in\{0,1,2,3\}$ match all the 8 values located around the computational cell,

$$\begin{pmatrix} f^{\pm,n}_{i-1,j} \\ f^{\mp,n}_{i,j+1} \\ g^{\pm,n}_{i,j} \\ g^{\mp,n}_{i-1,j+1} \end{pmatrix} := \bar f\big(\bar\sigma R(\cos\theta_k,\sin\theta_k),\mp(\cos\theta_k,\sin\theta_k)\big),$$

In that case, by construction, the S-matrix (18) maps the 4 "incoming values" ($\mathbf{v}\cdot\mathbf{x}<0$) into 4 "outgoing values" ($\mathbf{v}\cdot\mathbf{x}>0$) which all correspond to $\bar f$, so that

$$(S_\kappa)_{i-\frac12,j+\frac12}\begin{pmatrix} f^{+,n}_{i-1,j} \\ f^{-,n}_{i,j+1} \\ g^{+,n}_{i,j} \\ g^{-,n}_{i-1,j+1} \end{pmatrix} = \bar f\big(\bar\sigma R(\cos\theta_k,\sin\theta_k),(\cos\theta_k,\sin\theta_k)\big) = \begin{pmatrix} f^{+,n}_{i,j+1} \\ f^{-,n}_{i-1,j} \\ g^{+,n}_{i-1,j+1} \\ g^{-,n}_{i,j} \end{pmatrix},$$

and this exactly means that the discrete "4 velocities" restriction of $\bar f$ given by (15) is left invariant by (19). □

Remark 2. *Oppositely, if the opacity is stiff, $\sigma\lambda R \gg 1$ (local hydrodynamic limit), the sign of the entries in the S-matrix isn't obvious. These expressions allow to get necessary conditions for positivity in (19): for instance, in order to get $Y_0 - Y_2 \geq 0$ in the limit of infinite stiffness $\sigma R\gg 1$ with $\kappa\simeq 1, c\simeq 0$, one should have*

$$\begin{aligned} 0 &\leq \frac{1}{(1+c)^2+(1-c)(1+c)} - \frac{1}{(1+c)+2(1-c)} \\ &\leq \frac{1}{2(1+c)} - \frac{1}{3-c} \\ &\leq \frac{1-c}{2} - \frac{1+c/3}{3} \quad\Leftrightarrow\quad c\geq\frac37 \text{ or } \kappa\leq\frac47. \end{aligned}$$

Corollary 1. *Under the same assumptions, if $\sigma(\mathbf{x}),\kappa(\mathbf{x})$ are constant in the whole computational domain, then any truncated steady-state of the form (15) restricted to four velocities, $\theta_k = \frac{\pi}{4}+\frac{k\,\pi}{2}$, $k\in\{0,1,2,3\}$ is preserved.*

Proof. Given $R>0$, if all parameters are constant, any steady-state $\bar f(\mathbf{x},\mathbf{v})$ can be split into "patches" of the form (15) inside the disks D_R of radius R and centered in $(x_{i-\frac12}, y_{j+\frac12})$, $i,j\in\mathbb{Z}^2$ (see Fig. 2) inside which the proof of Theorem 1 applies. □

2.3. **Consistency with damped diffusive limits in 2D**

In order to study its diffusive limit,

$$R\mapsto\frac{R}{\varepsilon}, \qquad \kappa\mapsto\varepsilon^2\kappa, \qquad \lambda^\varepsilon\frac{R}{\varepsilon} = R\sqrt{\kappa(2-\varepsilon^2\kappa)}\to\Delta x\sqrt{\kappa}, \tag{21}$$

we shall again decompose S^ε_κ in a form directly inspired by [23,24],

$$S^\varepsilon_\kappa = \begin{pmatrix} 0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix} + O(\lambda^\varepsilon\simeq\varepsilon\sqrt{2\kappa}), \qquad S_0 := \begin{pmatrix} 0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}.$$

Since $\lambda^\varepsilon = \sqrt{1-c^2} = \varepsilon\sqrt{\kappa(2-\varepsilon^2\kappa)}$, $1+c+\kappa = 2$, and $\lambda^\varepsilon R/\varepsilon \simeq \Delta x\sqrt{\kappa}$,

$$\begin{aligned}
Y_0 - Y_2 &= \frac{4\lambda^\varepsilon \mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon})}{(1+c)^2\mathcal{I}_2(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\lambda^\varepsilon\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + |\lambda^\varepsilon|^2\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})} \\
&\quad - \frac{4\lambda^\varepsilon \mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon})}{(1+c)\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\lambda^\varepsilon\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\varepsilon^2\kappa\mathcal{I}_2(\lambda^\varepsilon \frac{R}{\varepsilon})}, \\
Y_0 - 2Y_1 + Y_2 &= \frac{4\lambda^\varepsilon \mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})}{2\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + \lambda^\varepsilon\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})} - \frac{4\lambda^\varepsilon \mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon})}{(1+c)\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\lambda^\varepsilon\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\varepsilon^2\kappa\mathcal{I}_2(\lambda^\varepsilon \frac{R}{\varepsilon})} \\
&\quad - \frac{4\lambda^\varepsilon \mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon})}{(1+c)^2\mathcal{I}_2(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\lambda^\varepsilon\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + |\lambda^\varepsilon|^2\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})}, \\
Y_0 + 2Y_1 + Y_2 &= 4 - \frac{4\lambda^\varepsilon \mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})}{2\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + \lambda^\varepsilon\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})} - \frac{4\lambda^\varepsilon \mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon})}{(1+c)\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\lambda^\varepsilon\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\varepsilon^2\kappa\mathcal{I}_2(\lambda^\varepsilon \frac{R}{\varepsilon})} \\
&\quad - \frac{4\lambda^\varepsilon \mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon})}{(1+c)^2\mathcal{I}_2(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\lambda^\varepsilon\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + |\lambda^\varepsilon|^2\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})}.
\end{aligned}$$

By analogy with coefficients defined in [24, §4.1], and neglecting $2\varepsilon^2\kappa\mathcal{I}_2(...)$,

$$\alpha^\varepsilon = \frac{\sqrt{\kappa(2-\varepsilon^2\kappa)}\,\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})}{2\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + \lambda^\varepsilon\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})}, \tag{22}$$

$$\beta^\varepsilon = \frac{\sqrt{\kappa(2-\varepsilon^2\kappa)}\,\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon})}{(1+c)^2\mathcal{I}_2(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\lambda^\varepsilon\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon}) + |\lambda^\varepsilon|^2\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon})} \tag{23}$$

$$\gamma^\varepsilon = -\frac{\sqrt{\kappa(2-\varepsilon^2\kappa)}\,\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon})}{(1+c)\mathcal{I}_0(\lambda^\varepsilon \frac{R}{\varepsilon}) + 2\lambda^\varepsilon\mathcal{I}_1(\lambda^\varepsilon \frac{R}{\varepsilon})}, \tag{24}$$

and the decomposition of $S^\varepsilon_\alpha = S_0 + \varepsilon S^\varepsilon_1$ follows:

$$S^\varepsilon_\alpha = \begin{pmatrix} 0&1&0&0\\ 1&0&0&0\\ 0&0&0&1\\ 0&0&1&0 \end{pmatrix} + \varepsilon \left\{ \gamma^\varepsilon \begin{pmatrix} 1&1&1&1\\ 1&1&1&1\\ 1&1&1&1\\ 1&1&1&1 \end{pmatrix} + \begin{pmatrix} \alpha^\varepsilon-\beta^\varepsilon & -(\alpha^\varepsilon+\beta^\varepsilon) & \beta^\varepsilon & \beta^\varepsilon \\ -(\alpha^\varepsilon+\beta^\varepsilon) & \alpha^\varepsilon-\beta^\varepsilon & \beta^\varepsilon & \beta^\varepsilon \\ \beta^\varepsilon & \beta^\varepsilon & \alpha^\varepsilon-\beta^\varepsilon & -(\alpha^\varepsilon+\beta^\varepsilon) \\ \beta^\varepsilon & \beta^\varepsilon & -(\alpha^\varepsilon+\beta^\varepsilon) & \alpha^\varepsilon-\beta^\varepsilon \end{pmatrix} \right\}.$$

Accordingly, an IMEX scheme emerges from

$$\begin{pmatrix} f^{+,n+1}_{i,j+1} \\ f^{-,n+1}_{i-1,j} \\ g^{+,n+1}_{i-1,j+1} \\ g^{-,n+1}_{i,j} \end{pmatrix} + \frac{\Delta t}{2\varepsilon R}\left\{ \begin{pmatrix} f^{+,n+1}_{i,j+1} \\ f^{-,n+1}_{i-1,j} \\ g^{+,n+1}_{i-1,j+1} \\ g^{-,n+1}_{i,j} \end{pmatrix} - S_0 \begin{pmatrix} f^{+,n+1}_{i-1,j} \\ f^{-,n+1}_{i,j+1} \\ g^{+,n+1}_{i,j} \\ g^{-,n+1}_{i-1,j+1} \end{pmatrix} \right\} = \begin{pmatrix} f^{+,n}_{i,j+1} \\ f^{-,n}_{i-1,j} \\ g^{+,n}_{i-1,j+1} \\ g^{-,n}_{i,j} \end{pmatrix} + \frac{\Delta t}{2R} S^\varepsilon_1 \begin{pmatrix} f^{+,n}_{i-1,j} \\ f^{-,n}_{i,j+1} \\ g^{+,n}_{i,j} \\ g^{-,n}_{i-1,j+1} \end{pmatrix}. \tag{25}$$

According to (17), there are two different types of terms in that expression:

- the ones in $\mathcal{I}_n(\Delta x)/\mathcal{I}_{n+1}(\Delta x)$, behaving like $1/\Delta x$, render diffusion,

$$\alpha^\varepsilon \to \frac{\sqrt{2\kappa}\,\mathcal{I}_0(\sigma\Delta x\sqrt{\kappa})}{2\mathcal{I}_1(\sigma\Delta x\sqrt{\kappa})} \simeq \frac{\sqrt{2\kappa}}{\sigma\Delta x\sqrt{\kappa}}, \qquad \beta^\varepsilon \to \frac{\sqrt{2\kappa}\,\mathcal{I}_1(\sigma\Delta x\sqrt{\kappa})}{4\mathcal{I}_2(\sigma\Delta x\sqrt{\kappa})} \simeq \frac{\sqrt{2}}{\sigma\Delta x},$$

and if the dissipation isn't stiff, $\sigma\Delta x\sqrt{\kappa} \ll 1$, modified Bessel functions (17) ensure that

$$\begin{aligned}\alpha^\varepsilon - \beta^\varepsilon &\to \frac{\sqrt{2\kappa}}{4}\left(\frac{2\mathcal{I}_0(\sigma\Delta x\sqrt{\kappa})\,\mathcal{I}_2(\sigma\Delta x\sqrt{\kappa}) - \mathcal{I}_1(\sigma\Delta x\sqrt{\kappa})^2}{\mathcal{I}_1(\sigma\Delta x\sqrt{\kappa})\mathcal{I}_2(\sigma\Delta x\sqrt{\kappa})}\right) \qquad \varepsilon\to 0\\ &\simeq \frac{\sqrt{2}\,\Delta x\,\sigma\kappa}{12} = \frac{R\,\sigma\kappa}{6}, \quad \text{because} \quad \mathcal{I}_n(x) = \Big(\frac{x}{2}\Big)^n \sum_{k\in\mathbb{N}} \frac{(x/2)^{2k}}{k!\,(k+n)!}.\end{aligned}$$

- the only one in $\mathcal{I}_1(\Delta x)/\mathcal{I}_0(\Delta x)$, behaving like $O(\Delta x)$, renders dissipation,

$$\gamma^\varepsilon \to -\frac{\sqrt{2\kappa}\,\mathcal{I}_1(\sigma\Delta x\sqrt{\kappa})}{2\mathcal{I}_0(\sigma\kappa\Delta x\sqrt{2})} \simeq -\frac{\sigma\kappa\Delta x\sqrt{2}}{4} = -\frac{\sigma\kappa R}{2}, \qquad \sigma\Delta x\sqrt{\kappa}\ll 1.$$

After a shift of indexes, the IMEX scheme (25) rewrites,

$$\begin{pmatrix} 1+\frac{\Delta t}{2\varepsilon R} & -\frac{\Delta t}{2\varepsilon R} & 0 & 0\\ -\frac{\Delta t}{2\varepsilon R} & 1+\frac{\Delta t}{2\varepsilon R} & 0 & 0\\ 0 & 0 & 1+\frac{\Delta t}{2\varepsilon R} & -\frac{\Delta t}{2\varepsilon R}\\ 0 & 0 & -\frac{\Delta t}{2\varepsilon R} & 1+\frac{\Delta t}{2\varepsilon R}\end{pmatrix}\begin{pmatrix} f^{+,n+1}_{i,j}\\ f^{-,n+1}_{i,j}\\ g^{+,n+1}_{i,j}\\ g^{-,n+1}_{i,j}\end{pmatrix} = \tag{26}$$
$$\begin{pmatrix} f^{+,n}_{i,j}\\ f^{-,n}_{i,j}\\ g^{+,n}_{i,j}\\ g^{-,n}_{i,j}\end{pmatrix} + \frac{\Delta t}{2R}\left\{\begin{pmatrix} \gamma^\varepsilon_{i-\frac12,j-\frac12}(f^{+,n}_{i-1,j-1}+f^{-,n}_{i,j}+g^{+,n}_{i,j-1}+g^{-,n}_{i-1,j})\\ \gamma^\varepsilon_{i+\frac12,j+\frac12}(f^{+,n}_{i,j}+f^{-,n}_{i+1,j+1}+g^{+,n}_{i+1,j}+g^{-,n}_{i,j+1})\\ \gamma^\varepsilon_{i+\frac12,j-\frac12}(f^{+,n}_{i,j-1}+f^{-,n}_{i+1,j}+g^{+,n}_{i+1,j-1}+g^{-,n}_{i,j})\\ \gamma^\varepsilon_{i-\frac12,j+\frac12}(f^{+,n}_{i-1,j}+f^{-,n}_{i,j+1}+g^{+,n}_{i,j}+g^{-,n}_{i-1,j+1})\end{pmatrix} + \right.$$
$$\left.\begin{pmatrix} \alpha^\varepsilon_{i-\frac12,j-\frac12}(f^{+,n}_{i-1,j-1}-f^{-,n}_{i,j}) - \beta^\varepsilon_{i-\frac12,j-\frac12}(f^{+,n}_{i-1,j-1}+f^{-,n}_{i,j}-g^{+,n}_{i,j-1}-g^{-,n}_{i-1,j})\\ \alpha^\varepsilon_{i+\frac12,j+\frac12}(f^{-,n}_{i+1,j+1}-f^{+,n}_{i,j}) - \beta^\varepsilon_{i+\frac12,j+\frac12}(f^{+,n}_{i,j}+f^{-,n}_{i+1,j+1}-g^{+,n}_{i+1,j}-g^{-,n}_{i,j+1})\\ \alpha^\varepsilon_{i+\frac12,j-\frac12}(g^{+,n}_{i+1,j-1}-g^{-,n}_{i,j}) + \beta^\varepsilon_{i+\frac12,j-\frac12}(f^{+,n}_{i,j-1}+f^{-,n}_{i+1,j}-g^{+,n}_{i+1,j-1}-g^{-,n}_{i,j})\\ \alpha^\varepsilon_{i-\frac12,j+\frac12}(g^{-,n}_{i-1,j+1}-g^{+,n}_{i,j}) + \beta^\varepsilon_{i-\frac12,j+\frac12}(f^{+,n}_{i-1,j}+f^{-,n}_{i,j+1}-g^{+,n}_{i,j}-g^{-,n}_{i-1,j+1})\end{pmatrix}\right\}.$$

Denote $\mathfrak{f}^n_{i,j} = f^{+,n}_{i,j} + f^{-,n}_{i,j}$ and $\mathfrak{g}^n_{i,j} = g^{+,n}_{i,j} + g^{-,n}_{i,j}$, so $\rho^n_{i,j} = \mathfrak{f}^n_{i,j} + \mathfrak{g}^n_{i,j}$: adding equations in (26) gives,

$$\begin{aligned}\mathfrak{f}^{n+1}_{i,j} = \ \mathfrak{f}^n_{i,j} &+ \frac{\Delta t}{2R}\Big(\alpha^\varepsilon_{i-\frac12,j-\frac12}(f^{+,n}_{i-1,j-1}-f^{-,n}_{i,j}) + \alpha^\varepsilon_{i+\frac12,j+\frac12}(f^{-,n}_{i+1,j+1}-f^{+,n}_{i,j})\Big)\\ &- \frac{\Delta t}{2R}\Big(\beta^\varepsilon_{i-\frac12,j-\frac12}(f^{+,n}_{i-1,j-1}+f^{-,n}_{i,j}-g^{+,n}_{i,j-1}-g^{-,n}_{i-1,j})\\ &\qquad + \beta^\varepsilon_{i+\frac12,j+\frac12}(f^{+,n}_{i,j}+f^{-,n}_{i+1,j+1}-g^{+,n}_{i+1,j}-g^{-,n}_{i,j+1}) \qquad (27)\\ &\qquad\quad - \gamma^\varepsilon_{i-\frac12,j-\frac12}(f^{+,n}_{i-1,j-1}+f^{-,n}_{i,j}+g^{+,n}_{i,j-1}+g^{-,n}_{i-1,j})\\ &\qquad\qquad - \gamma^\varepsilon_{i+\frac12,j+\frac12}(f^{+,n}_{i,j}+f^{-,n}_{i+1,j+1}+g^{+,n}_{i+1,j}+g^{-,n}_{i,j+1})\Big),\end{aligned}$$

(28)

$$\begin{aligned}\mathfrak{g}_{i,j}^{n+1} = \ \mathfrak{g}_{i,j}^{n} &+ \frac{\Delta t}{2R}\Big(\alpha^{\varepsilon}_{i+\frac12,j-\frac12}(g^{+,n}_{i+1,j-1} - g^{-,n}_{i,j}) + \alpha^{\varepsilon}_{i-\frac12,j+\frac12}(g^{-,n}_{i-1,j+1} - g^{+,n}_{i,j})\Big)\\ &+ \frac{\Delta t}{2R}\Big(\beta^{\varepsilon}_{i+\frac12,j-\frac12}(f^{+,n}_{i,j-1} + f^{-,n}_{i+1,j} - g^{+,n}_{i+1,j-1} - g^{-,n}_{i,j})\\ &\qquad + \beta^{\varepsilon}_{i-\frac12,j+\frac12}(f^{+,n}_{i-1,j} + f^{-,n}_{i,j+1} - g^{+,n}_{i,j} - g^{-,n}_{i-1,j+1})\\ &\qquad\quad + \gamma^{\varepsilon}_{i+\frac12,j-\frac12}(f^{+,n}_{i,j-1} + f^{-,n}_{i+1,j} + g^{+,n}_{i+1,j-1} + g^{-,n}_{i,j})\\ &\qquad\qquad + \gamma^{\varepsilon}_{i-\frac12,j+\frac12}(f^{+,n}_{i-1,j} + f^{-,n}_{i,j+1} + g^{+,n}_{i,j} + g^{-,n}_{i-1,j+1})\Big).\end{aligned} \tag{29}$$

Corollary 2. *The IMEX scheme (25) is "2D well-balanced" in the sense that, if data are at steady-state,*

$$A^n_{i-\frac12,j+\frac12} = (S^{\varepsilon}_{\kappa})_{i-\frac12,j+\frac12}\, B^n_{i-\frac12,j+\frac12} \qquad \textit{where} \qquad A^n_{i-\frac12,j+\frac12} := \begin{pmatrix} f^{+,n}_{i,j+1}\\ f^{-,n}_{i-1,j}\\ g^{+,n}_{i-1,j+1}\\ g^{-,n}_{i,j}\end{pmatrix}, \quad B^n_{i-\frac12,j+\frac12} := \begin{pmatrix} f^{+,n}_{i-1,j}\\ f^{-,n}_{i,j+1}\\ g^{+,n}_{i,j}\\ g^{-,n}_{i-1,j+1}\end{pmatrix},$$

for all $i,j \in \mathbb{Z}^2$, then

$$\forall (i,j) \in \mathbb{Z}^2, \qquad f^{\pm,n+1}_{i,j} = f^{\pm,n}_{i,j}, \qquad g^{\pm,n+1}_{i,j} = g^{\pm,n}_{i,j}. \tag{30}$$

Proof. Denote $S := (S^{\varepsilon}_{\kappa})_{i-\frac12,j+\frac12}$ for simplicity; since $S = S_0 + \varepsilon S^{\varepsilon}_1$, (25) rewrites

$$A^{n+1}_{i-\frac12,j+\frac12} + \frac{\Delta t}{2\varepsilon R}\left(A^{n+1}_{i-\frac12,j+\frac12} - S_0 B^{n+1}_{i-\frac12,j+\frac12}\right) = A^n_{i-\frac12,j+\frac12} + \frac{\Delta t}{2\varepsilon R} \times \varepsilon S^{\varepsilon}_1\, B^n_{i-\frac12,j+\frac12}.$$

Being at steady-state,

$$A^n_{i-\frac12,j+\frac12} = S\, B^n_{i-\frac12,j+\frac12} = (S_0 + \varepsilon S^{\varepsilon}_1) B^n_{i-\frac12,j+\frac12}, \qquad S_0 = \begin{pmatrix} 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix},$$

so the IMEX scheme rewrites, for this particular type of data,

$$(A^{n+1}_{i-\frac12,j+\frac12} - A^n_{i-\frac12,j+\frac12}) + \frac{\Delta t}{2\varepsilon R}\left(A^{n+1}_{i-\frac12,j+\frac12} - S_0 B^{n+1}_{i-\frac12,j+\frac12}\right) = \frac{\Delta t}{2\varepsilon R}\left(A^n_{i-\frac12,j+\frac12} - S_0 B^n_{i-\frac12,j+\frac12}\right).$$

Yet, (30) is reached by shifting indexes like in (26), along with

- adding both first and last two equations, so as to get the preservation of averages,

$$\mathfrak{f}^{n+1}_{i,j} = \mathfrak{f}^n_{i,j}, \qquad \mathfrak{g}^{n+1}_{i,j} = \mathfrak{g}^n_{i,j}, \qquad \text{and} \qquad \rho^{n+1}_{i,j} = \rho^n_{i,j};$$

- subtracting both first and last two equations while defining macroscopic fluxes,

$$\mathbf{J}^n_{i,j} = \begin{pmatrix} f^+ - f^-\\ g^+ - g^-\end{pmatrix}^n_{i,j} \in \mathbb{R}^2, \qquad \left(1 + \frac{\Delta t}{\varepsilon R}\right)(\mathbf{J}^{n+1}_{i,j} - \mathbf{J}^n_{i,j}) = 0;$$

- and finally taking advantage of the elementary relation,

$$\begin{pmatrix} f^{\pm}\\ g^{\pm}\end{pmatrix} = \frac12\left(\begin{pmatrix} \mathfrak{f}\\ \mathfrak{g}\end{pmatrix} \pm \mathbf{J}\right).$$

□

Remark 3. *The assumptions of Corollary 2 are similar to the ones of Theorem 1. Indeed, to ensure that, in a given cell, $A^n_{i-\frac{1}{2},j+\frac{1}{2}} = S^\varepsilon_{i-\frac{1}{2},j+\frac{1}{2}}\, B^n_{i-\frac{1}{2},j+\frac{1}{2}}$ holds, it is sufficient that there exists a local steady-state of the form (15) for the rescaled equation*

$$\mathbf{v}\cdot\nabla_{\mathbf{x}} f = \frac{\sigma}{\varepsilon}\left((1-\varepsilon^2\kappa)\int_{\mathbb{S}^1} f(t,\mathbf{x},\mathbf{v}')\frac{\mathrm{d}\mathbf{v}'}{2\pi} - f\right), \qquad \mathbf{x}\in D_R(x_{i-\frac{1}{2}}, y_{j+\frac{1}{2}})$$

By sending formally $\varepsilon\to 0$, having divided by $4\lambda^\varepsilon$, coefficients split into

$$\begin{aligned}
\frac{Y_0 - Y_2}{4\lambda^\varepsilon} &= \frac{\mathcal{I}_1(\Delta x\sqrt{\kappa})}{4\mathcal{I}_2(\Delta x\sqrt{\kappa})} - \frac{\mathcal{I}_1(\Delta x\sqrt{\kappa})}{2\mathcal{I}_0(\Delta x\sqrt{\kappa})} + O(\varepsilon),\\
\frac{Y_0 - 2Y_1 + Y_2}{4\lambda^\varepsilon} &= \left(\frac{\mathcal{I}_0(\Delta x\sqrt{\kappa})}{2\mathcal{I}_1(\Delta x\sqrt{\kappa})} - \frac{\mathcal{I}_1(\Delta x\sqrt{\kappa})}{4\mathcal{I}_2(\Delta x\sqrt{\kappa})}\right) - \frac{\mathcal{I}_1(\Delta x\sqrt{\kappa})}{2\mathcal{I}_0(\Delta x\sqrt{\kappa})} + O(\varepsilon)\\
\frac{(Y_0 + 2Y_1 + Y_2) - 4}{4\lambda^\varepsilon} &= -\left(\frac{\mathcal{I}_0(\Delta x\sqrt{\kappa})}{2\mathcal{I}_1(\Delta x\sqrt{\kappa})} + \frac{\mathcal{I}_1(\Delta x\sqrt{\kappa})}{4\mathcal{I}_2(\Delta x\sqrt{\kappa})}\right) - \frac{\mathcal{I}_1(\Delta x\sqrt{\kappa})}{2\mathcal{I}_0(\Delta x\sqrt{\kappa})} + O(\varepsilon),
\end{aligned}$$

along with (26), which yields,

$$\begin{pmatrix} f^{+,n+1}_{i,j}\\ f^{-,n+1}_{i,j}\\ g^{+,n+1}_{i,j}\\ g^{-,n+1}_{i,j}\end{pmatrix} \in \mathrm{Span}\left\{\begin{pmatrix}1\\1\\0\\0\end{pmatrix},\ \begin{pmatrix}0\\0\\1\\1\end{pmatrix}\right\};$$

$$f^{+,n+1}_{i,j} = f^{-,n+1}_{i,j} = \frac{1}{2}\mathfrak{f}^{n+1}_{i,j}, \qquad g^{+,n+1}_{i,j} = g^{-,n+1}_{i,j} = \frac{1}{2}\mathfrak{g}^{n+1}_{i,j}. \tag{31}$$

Yet, invoking (31) in both (27) and (29) yields diffusive fluxes and dissipative terms,

$$\begin{aligned}
\mathfrak{f}^{n+1}_{i,j} = \mathfrak{f}^n_{i,j} + \frac{\Delta t}{4R}\Big(&-(\alpha_{i-\frac{1}{2},j-\frac{1}{2}} + \alpha_{i+\frac{1}{2},j+\frac{1}{2}} + \beta_{i-\frac{1}{2},j-\frac{1}{2}} + \beta_{i+\frac{1}{2},j+\frac{1}{2}})\mathfrak{f}^n_{i,j}\\
&+ (\alpha_{i-\frac{1}{2},j-\frac{1}{2}} - \beta_{i-\frac{1}{2},j-\frac{1}{2}})\mathfrak{f}^n_{i-1,j-1} + (\alpha_{i+\frac{1}{2},j+\frac{1}{2}} - \beta_{i+\frac{1}{2},j+\frac{1}{2}})\mathfrak{f}^n_{i+1,j+1}\\
&+ \beta_{i-\frac{1}{2},j-\frac{1}{2}}(\mathfrak{g}^n_{i,j-1} + \mathfrak{g}^n_{i-1,j}) + \beta_{i+\frac{1}{2},j+\frac{1}{2}}(\mathfrak{g}^n_{i+1,j} + \mathfrak{g}^n_{i,j+1})\Big)\\
+\frac{\Delta t}{4R}\Big(&\gamma_{i-\frac{1}{2},j-\frac{1}{2}}(\mathfrak{f}^n_{i-1,j-1} + \mathfrak{f}^n_{i,j} + \mathfrak{g}^n_{i,j-1} + \mathfrak{g}^n_{i-1,j})\\
&+ \gamma_{i+\frac{1}{2},j+\frac{1}{2}}(\mathfrak{f}^n_{i,j} + \mathfrak{f}^n_{i+1,j+1} + \mathfrak{g}^n_{i+1,j} + \mathfrak{g}^n_{i,j+1})\Big),\\
\mathfrak{g}^{n+1}_{i,j} = \mathfrak{g}^n_{i,j} + \frac{\Delta t}{4R}\Big(&-(\alpha_{i+\frac{1}{2},j-\frac{1}{2}} + \alpha_{i-\frac{1}{2},j+\frac{1}{2}} + \beta_{i+\frac{1}{2},j-\frac{1}{2}} + \beta_{i-\frac{1}{2},j+\frac{1}{2}})\mathfrak{g}^n_{i,j}\\
&+ (\alpha_{i+\frac{1}{2},j-\frac{1}{2}} - \beta_{i+\frac{1}{2},j-\frac{1}{2}})\mathfrak{g}^n_{i+1,j-1} + (\alpha_{i-\frac{1}{2},j+\frac{1}{2}} - \beta_{i-\frac{1}{2},j+\frac{1}{2}})\mathfrak{g}^n_{i-1,j+1}\\
&+ \beta_{i+\frac{1}{2},j-\frac{1}{2}}(\mathfrak{f}^n_{i,j-1} + \mathfrak{f}^n_{i+1,j}) + \beta_{i-\frac{1}{2},j+\frac{1}{2}}(\mathfrak{f}^n_{i-1,j} + \mathfrak{f}^n_{i,j+1})\Big)\\
+\frac{\Delta t}{4R}\Big(&\gamma_{i+\frac{1}{2},j-\frac{1}{2}}(\mathfrak{f}^n_{i,j-1} + \mathfrak{f}^n_{i+1,j} + \mathfrak{g}^n_{i+1,j-1} + \mathfrak{g}^n_{i,j})\\
&+ \gamma_{i-\frac{1}{2},j+\frac{1}{2}}(\mathfrak{f}^n_{i-1,j} + \mathfrak{f}^n_{i,j+1} + \mathfrak{g}^n_{i,j} + \mathfrak{g}^n_{i-1,j+1})\Big).
\end{aligned}$$

We choose to set up the horizontal/vertical discrete Laplace operator with $\beta^{\varepsilon}_{i\pm\frac{1}{2},j\pm\frac{1}{2}}$,

$$\begin{aligned}\mathfrak{f}^{n+1}_{i,j} = \mathfrak{f}^n_{i,j} + \frac{\Delta t}{4R}\Big((\alpha_{i-\frac12,j-\frac12}-\beta_{i-\frac12,j-\frac12})(\mathfrak{f}^n_{i-1,j-1}-\mathfrak{f}^n_{i,j})\\ +(\alpha_{i+\frac12,j+\frac12}-\beta_{i+\frac12,j+\frac12})(\mathfrak{f}^n_{i+1,j+1}-\mathfrak{f}^n_{i,j})\\ +\beta_{i-\frac12,j-\frac12}(\mathfrak{g}^n_{i,j-1}-2\mathfrak{f}^n_{i,j}+\mathfrak{g}^n_{i-1,j})\\ +\beta_{i+\frac12,j+\frac12}(\mathfrak{g}^n_{i+1,j}-2\mathfrak{f}^n_{i,j}+\mathfrak{g}^n_{i,j+1})\Big)\\ +\frac{\Delta t}{4R}\Big(\gamma_{i-\frac12,j-\frac12}(\mathfrak{f}^n_{i-1,j-1}+\mathfrak{f}^n_{i,j}+\mathfrak{g}^n_{i,j-1}+\mathfrak{g}^n_{i-1,j})\\ +\gamma_{i+\frac12,j+\frac12}(\mathfrak{f}^n_{i,j}+\mathfrak{f}^n_{i+1,j+1}+\mathfrak{g}^n_{i+1,j}+\mathfrak{g}^n_{i,j+1})\Big),\end{aligned}$$

$$\begin{aligned}\mathfrak{g}^{n+1}_{i,j} = \mathfrak{g}^n_{i,j} + \frac{\Delta t}{4R}\Big((\alpha_{i+\frac12,j-\frac12}-\beta_{i+\frac12,j-\frac12})(\mathfrak{g}^n_{i+1,j-1}-\mathfrak{g}^n_{i,j})\\ +(\alpha_{i-\frac12,j+\frac12}-\beta_{i-\frac12,j+\frac12})(\mathfrak{g}^n_{i-1,j+1}-\mathfrak{g}^n_{i,j})\\ +\beta_{i+\frac12,j-\frac12}(\mathfrak{f}^n_{i,j-1}-2\mathfrak{g}^n_{i,j}+\mathfrak{f}^n_{i+1,j})\\ +\beta_{i-\frac12,j+\frac12}(\mathfrak{f}^n_{i-1,j}-2\mathfrak{g}^n_{i,j}+\mathfrak{f}^n_{i,j+1})\Big)\\ +\frac{\Delta t}{4R}\Big(\gamma_{i+\frac12,j-\frac12}(\mathfrak{f}^n_{i,j-1}+\mathfrak{f}^n_{i+1,j}+\mathfrak{g}^n_{i+1,j-1}+\mathfrak{g}^n_{i,j})\\ +\gamma_{i-\frac12,j+\frac12}(\mathfrak{f}^n_{i-1,j}+\mathfrak{f}^n_{i,j+1}+\mathfrak{g}^n_{i,j}+\mathfrak{g}^n_{i-1,j+1})\Big),\end{aligned}$$

so that another "diagonal" Laplace operator, involving $\alpha^{\varepsilon}-\beta^{\varepsilon}=O(R)$ (for fine grid) appears. Clearly, both $\mathfrak{f}$ and $\mathfrak{g}$ solve similar damped diffusion equations, if both $\sigma(\mathbf{x})$ and $\kappa(\mathbf{x})$ are smooth. Accordingly, if initial data are "well-prepared", they are likely to stay so for later times because the Maxwellian gap $\mathfrak{f}^n_{i,j}-\mathfrak{g}^n_{i,j}$ satisfies a parabolic linear equation, too. As $\rho=\mathfrak{f}+\mathfrak{g}$, assuming $\mathfrak{f}_{i,j}=\mathfrak{g}_{i,j}$ for all $i,j\in\mathbb{Z}^2$ brings the limiting scheme,

$$\begin{aligned}\rho^{n+1}_{i,j} = \rho^n_{i,j} + \frac{\Delta t}{4R}\Big(&\frac{\beta_{i-\frac12,j-\frac12}+\beta_{i+\frac12,j-\frac12}}{2}(\rho^n_{i,j-1}-\rho^n_{i,j}) + \frac{\beta_{i-\frac12,j+\frac12}+\beta_{i-\frac12,j-\frac12}}{2}(\rho^n_{i-1,j}-\rho^n_{i,j})\\ &+\frac{\beta_{i+\frac12,j+\frac12}+\beta_{i-\frac12,j+\frac12}}{2}(\rho^n_{i,j+1}-\rho^n_{i,j}) + \frac{\beta_{i+\frac12,j-\frac12}+\beta_{i+\frac12,j+\frac12}}{2}(\rho^n_{i+1,j}-\rho^n_{i,j})\\ &+\frac{\alpha_{i-\frac12,j-\frac12}-\beta_{i-\frac12,j-\frac12}}{2}(\rho^n_{i-1,j-1}-\rho^n_{i,j}) + \frac{\alpha_{i+\frac12,j+\frac12}-\beta_{i+\frac12,j+\frac12}}{2}(\rho^n_{i+1,j+1}-\rho^n_{i,j})\\ &+\frac{\alpha_{i+\frac12,j-\frac12}-\beta_{i+\frac12,j-\frac12}}{2}(\rho^n_{i+1,j-1}-\rho^n_{i,j}) + \frac{\alpha_{i-\frac12,j+\frac12}-\beta_{i-\frac12,j+\frac12}}{2}(\rho^n_{i-1,j+1}-\rho^n_{i,j})\\ &+\frac{\gamma_{i-\frac12,j-\frac12}}{2}(\rho^n_{i-1,j-1}+\rho^n_{i-1,j}+\rho^n_{i,j-1}+\rho^n_{i,j}) + \frac{\gamma_{i-\frac12,j+\frac12}}{2}(\rho^n_{i-1,j}+\rho^n_{i-1,j+1}+\rho^n_{i,j}+\rho^n_{i,j+1})\\ &+\frac{\gamma_{i+\frac12,j-\frac12}}{2}(\rho^n_{i,j-1}+\rho^n_{i,j}+\rho^n_{i+1,j-1}+\rho^n_{i+1,j}) + \frac{\gamma_{i+\frac12,j+\frac12}}{2}(\rho^n_{i,j}+\rho^n_{i,j+1}+\rho^n_{i+1,j}+\rho^n_{i+1,j+1})\Big).\end{aligned} \tag{32}$$

Assuming that the parameters aren't stiff, meaning $\sigma\Delta x\sqrt{\kappa} \ll 1$,

$$\frac{\alpha^\varepsilon}{2\times 2R},\ \frac{\beta^\varepsilon}{2\times 2R} \simeq \frac{\sqrt{2}}{2\sigma\Delta x\times\Delta x\sqrt{2}} = \frac{1}{2\sigma\Delta x^2}, \qquad (2R=\Delta x\sqrt{2})$$
$$\alpha^\varepsilon-\beta^\varepsilon \simeq O(R), \qquad \frac{\gamma^\varepsilon}{4R}\simeq -\frac{\sigma\kappa}{8},$$

so the aforementioned asymptotic scheme is consistent, as $\Delta x \to 0$ with the diffusion approximation (3). Moreover, assuming that $\sigma(\mathbf{x}), \kappa(\mathbf{x})$ are positive constants, the scheme (32) rewrites as follows,

$$\begin{aligned}\rho_{i,j}^{n+1} =& \rho_{i,j}^n\left(1-\frac{\Delta t(\alpha^\varepsilon+\beta^\varepsilon-\gamma^\varepsilon)}{2R}\right)\\ &+\frac{\Delta t(\beta^\varepsilon+\gamma^\varepsilon)}{4R}\left(\rho_{i,j-1}^n+\rho_{i-1,j}^n+\rho_{i,j+1}^n+\rho_{i+1,j}^n\right)\\ &+\frac{\Delta t(\alpha^\varepsilon-\beta^\varepsilon+\gamma^\varepsilon)}{8R}\left(\rho_{i-1,j-1}^n+\rho_{i-1,j+1}^n+\rho_{i+1,j-1}^n+\rho_{i+1,j+1}^n\right),\end{aligned}$$

and under the fine-grid approximation (17), $\beta^\varepsilon+\gamma^\varepsilon \geq 0$. Consequently, an elementary manner to make it unconditionally positivity-preserving is to treat implicitly all the dissipative terms, the ones multiplied by γ^ε.

Remark 4. *Oppositely, in stiff regime, $\sigma\Delta x\sqrt{\kappa}\gg 1$, and Bessel functions balance each other:*

$$\alpha^\varepsilon \simeq \frac{\sqrt{\kappa}}{\sqrt{2}}, \qquad \beta^\varepsilon\simeq\frac{\sqrt{\kappa}}{2\sqrt{2}}, \qquad \gamma^\varepsilon\simeq-\frac{\sqrt{\kappa}}{\sqrt{2}},$$
$$\alpha^\varepsilon+\beta^\varepsilon\simeq\frac{3\sqrt{\kappa}}{2\sqrt{2}}\simeq 3\beta^\varepsilon, \qquad \alpha^\varepsilon-\beta^\varepsilon\simeq\frac{\sqrt{\kappa}}{2\sqrt{2}}\simeq\beta^\varepsilon, \qquad \gamma^\varepsilon\simeq-2\beta^\varepsilon.$$

The scheme (32) works on a 9-points "Moore stencil", hence doesn't match the "Steklov scheme", see [19, §4] and [4], which has only a 5-points stencil. Points on the diagonals $\rho^n_{i\pm1,j\pm1}$ are multiplied by $\beta^\varepsilon_{i\pm\frac12,j\pm\frac12}+\gamma^\varepsilon_{i\pm\frac12,j\pm\frac12}$.

3. Positive, composite two-dimensional scheme

Consider again the following kinetic system with adsorption (1),

$$\partial_t f(t,\mathbf{x},\mathbf{v})+\mathbf{v}\cdot\nabla_{\mathbf{x}}f=\bar\sigma\left((1-\bar\kappa)\int_{\mathbb{S}^1}f(t,\mathbf{x},\mathbf{v}')\,\frac{d\mathbf{v}'}{2\pi}-f\right), \qquad (\mathbf{x},\mathbf{v})\in\mathbb{R}^2\times\mathbb{S}^1,$$

in the particular case where $\bar\sigma,\bar\kappa\in(0,1)$ are positive constants.Thus, the change of variable $u=fe^{\bar\sigma\bar\kappa t}$ allows to convert (1) into the conservative equation (formerly studied in [24]),

$$\partial_t u(t,\mathbf{x},\mathbf{v})+\mathbf{v}\cdot\nabla_{\mathbf{x}}u=\bar\sigma(1-\bar\kappa)\left(\int_{\mathbb{S}^1}u(t,\mathbf{x},\mathbf{v}')\,\frac{d\mathbf{v}'}{2\pi}-u\right).$$

In the general case, $\sigma(\mathbf{x})$ and $\kappa(\mathbf{x})$ vary, so this computation cannot be made on a global scale. However, by assuming that σ and κ are frozen in each disc of radius $R=\Delta x/\sqrt{2}$ and centered in $(x_{i-\frac12},y_{j+\frac12})$, see Fig. 2, this trick can still be applied locally and this is enough for our purposes. Accordingly, by means of an exponential

modulation of the scheme given in [24, eq. (3.5)],

$$\begin{pmatrix} f^{+,n+1}_{i,j+1} \\ f^{-,n+1}_{i-1,j} \\ g^{+,n+1}_{i-1,j+1} \\ g^{-,n+1}_{i,j} \end{pmatrix} e^{\Delta t(\sigma\kappa)_{i-\frac{1}{2},j+\frac{1}{2}}} = (1-\frac{\Delta t}{2R}) \begin{pmatrix} f^{+,n}_{i,j+1} \\ f^{-,n}_{i-1,j} \\ g^{+,n}_{i-1,j+1} \\ g^{-,n}_{i,j} \end{pmatrix} + \frac{\Delta t}{2R} \tilde{S}(\sigma_{i-\frac{1}{2},j+\frac{1}{2}}(1-\kappa_{i-\frac{1}{2},j+\frac{1}{2}})) \begin{pmatrix} f^{+,n}_{i-1,j} \\ f^{-,n}_{i,j+1} \\ g^{+,n}_{i,j} \\ g^{-,n}_{i-1,j+1} \end{pmatrix}, \qquad (33)$$

a time-marching process for (2) can be deduced, where

$$\sigma_{i-\frac{1}{2},j+\frac{1}{2}} = \sigma(x_{i-\frac{1}{2}}, y_{j+\frac{1}{2}}), \qquad \kappa_{i-\frac{1}{2},j+\frac{1}{2}} = \kappa(x_{i-\frac{1}{2}}, y_{j+\frac{1}{2}}),$$

and the S-matrix $\tilde{S}(\alpha)$, $\alpha \geq 0$ corresponds to the limit $\kappa \to 0$ in (18) and reads (see [24, eq. (3.4)]):

$$\tilde{S}(\alpha) = \mathrm{Id} + \alpha R \left\{ \frac{1}{1+\alpha R} \begin{pmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} + \frac{1}{1+(1+\alpha R)^2} \begin{pmatrix} -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix} . \right\},$$

3.1. **Consistency and positivity**

Lemma 2. *Under the CFL $\Delta t \leq 2R$, the scheme (33) is positivity-preserving and consistent with (2) as $R \to 0$.*

Proof. Under the CFL restriction, the scheme (33) realizes a nonnegative combination; moreover, all the entries in the symmetric matrix $\tilde{S}(\sigma(1-\kappa))$ are nonnegative and positivity is preserved. Yet, assume R is small enough so that the CFL restriction gives a time-step small enough to linearize the exponential modulation,

$$e^{-\sigma_{i-\frac{1}{2},j+\frac{1}{2}}\kappa_{i-\frac{1}{2},j+\frac{1}{2}}\Delta t} \simeq 1 - \sigma_{i-\frac{1}{2},j+\frac{1}{2}}\kappa_{i-\frac{1}{2},j+\frac{1}{2}}\Delta t$$

and

$$\tilde{S}(\sigma_{i-\frac{1}{2},j+\frac{1}{2}}(1-\kappa_{i-\frac{1}{2},j+\frac{1}{2}})) \simeq \mathrm{Id} + \frac{R}{2}\sigma_{i-\frac{1}{2},j+\frac{1}{2}}(1-\kappa_{i-\frac{1}{2},j+\frac{1}{2}}) \begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix}.$$

Hence, since $R \ll 1$ and $\Delta t \ll 1$, the time-marching scheme (33) is approximated by

$$\begin{aligned} \begin{pmatrix} f^{+,n+1}_{i,j+1} \\ f^{-,n+1}_{i-1,j} \\ g^{+,n+1}_{i-1,j+1} \\ g^{-,n+1}_{i,j} \end{pmatrix} \simeq \Big(1 - \frac{\Delta t}{2R} - \sigma_{i-\frac{1}{2},j+\frac{1}{2}}\kappa_{i-\frac{1}{2},j+\frac{1}{2}}\Delta t\Big) \begin{pmatrix} f^{+,n}_{i,j+1} \\ f^{-,n}_{i-1,j} \\ g^{+,n}_{i-1,j+1} \\ g^{-,n}_{i,j} \end{pmatrix} + \frac{\Delta t}{2R} \begin{pmatrix} f^{+,n}_{i-1,j} \\ f^{-,n}_{i,j+1} \\ g^{+,n}_{i,j} \\ g^{-,n}_{i-1,j+1} \end{pmatrix} \\ + \frac{\Delta t}{4}\sigma_{i-\frac{1}{2},j+\frac{1}{2}}(1-\kappa_{i-\frac{1}{2},j+\frac{1}{2}}) \begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix} \begin{pmatrix} f^{+,n}_{i-1,j} \\ f^{-,n}_{i,j+1} \\ g^{+,n}_{i,j} \\ g^{-,n}_{i-1,j+1} \end{pmatrix}, \end{aligned}$$

and this expression is consistent with system (2),

$$\begin{aligned} \partial_t f^{\pm} \pm \frac{1}{\sqrt{2}}(\partial_x f^{\pm} + \partial_y f^{\pm}) &= \sigma(1-\kappa)(\frac{\rho}{4} - f^{\pm}) - \sigma\kappa f^{\pm} = \sigma((1-\kappa)\frac{\rho}{4} - f^{\pm}) \\ \partial_t g^{\pm} \mp \frac{1}{\sqrt{2}}(\partial_x g^{\pm} - \partial_y g^{\pm}) &= \sigma((1-\kappa)\frac{\rho}{4} - g^{\pm}) \end{aligned}$$

□

3.2. **Numerical steady-states**

Since it involves an exponential modulation in the time variable, the ability of (33) to preserve continuous steady-states is not obvious. Accordingly, denote outgoing and incoming states, respectively, as follows:

$$\forall (i,j,n) \in \mathbb{Z}^2 \times \mathbb{N}, \qquad A^n_{i-\frac{1}{2},j+\frac{1}{2}} := \begin{pmatrix} f^{+,n}_{i,j+1} \\ f^{-,n}_{i-1,j} \\ g^{+,n}_{i-1,j+1} \\ g^{-,n}_{i,j} \end{pmatrix}, \qquad B^n_{i-\frac{1}{2},j+\frac{1}{2}} := \begin{pmatrix} f^{+,n}_{i-1,j} \\ f^{-,n}_{i,j+1} \\ g^{+,n}_{i,j} \\ g^{-,n}_{i-1,j+1} \end{pmatrix}.$$

Accordingly, the scheme (33) rewrites simply

$$A^{n+1}_{i-\frac{1}{2},j+\frac{1}{2}} \exp(\sigma\kappa\Delta t) = (1 - \frac{\Delta t}{2R}) A^n_{i-\frac{1}{2},j+\frac{1}{2}} + \frac{\Delta t}{2R} \tilde{S}(\sigma(1-\kappa)) B^n_{i-\frac{1}{2},j+\frac{1}{2}}.$$

At numerical steady-state, $A^{n+1}_{i-\frac{1}{2},j+\frac{1}{2}} = A^n_{i-\frac{1}{2},j+\frac{1}{2}}$, so that

$$A^n_{i-\frac{1}{2},j+\frac{1}{2}} \left(\frac{\exp(\sigma\kappa\Delta t) - 1}{\Delta t} + \frac{1}{2R} \right) = \frac{1}{2R} \tilde{S}(\sigma(1-\kappa)) B^n_{i-\frac{1}{2},j+\frac{1}{2}}. \tag{34}$$

Yet, assume $\sigma\kappa\Delta t$ is small enough to linearize,

$$(1 + 2R\sigma\kappa) A^n_{i-\frac{1}{2},j+\frac{1}{2}} = \tilde{S}(\sigma(1-\kappa)) B^n_{i-\frac{1}{2},j+\frac{1}{2}},$$

thus another scattering matrix appears,

$$\mathbf{S}_\kappa := (1 - 2R\sigma\kappa) \tilde{S}(\sigma(1-\kappa)).$$

An interesting question is to relate it with the "exact" S-matrix in (19), at least for small values of R. Notice that at this stage, we can assume that $\sigma = 1$ without loss of generality thanks to a simple rescaling of R. First, the entries of the S-matrix in [24] are the limits when $\kappa \to 0$ of (19), recalled in (8):

$$Y_0 \to 1, \qquad Y_1 \to \frac{R-1}{R+1}, \qquad Y_2 \to \frac{R^2 - 2R + 2}{R^2 + 2R + 2}.$$

Yet, in $\mathbf{S}_\kappa$, the R's must be multiplied by $1 - \kappa$, so its entries agree with (20):

$$\begin{aligned} (1 - 2\kappa R) \times 1 &\simeq \frac{1 - \kappa R}{1 + \kappa R} \times 1 \simeq Y_0 \qquad (\text{when } R \ll 1) \\ \frac{1 - \kappa R}{1 + \kappa R} \times \frac{(1-\kappa)R - 1}{(1-\kappa)R + 1} &\simeq \frac{(1-\kappa)R - 1 + \kappa R}{(1-\kappa)R + 1 + \kappa R} = \frac{R-1}{R+1} \simeq Y_1 \\ \frac{1 - \kappa R}{1 + \kappa R} \times \frac{(1-\kappa)^2 R^2 - 2(1-\kappa)R + 2}{(1-\kappa)^2 R^2 + 2(1-\kappa)R + 2} &\simeq \frac{R^2 - 2R + 2 + 2\kappa R - 2\kappa R}{R^2 + 2R + 2 - 2\kappa R + 2\kappa R} = \frac{R^2 - 2R + 2}{R^2 + 2R + 2} \simeq Y_2. \end{aligned}$$

This computation shows that the steady-states of (33), at least when $\sigma R \ll 1$ (so exponential and Bessel functions behave like polynomials), are consistent with the "exact ones" preserved by (19) as long as necessary linearizations are licit.

Lemma 3. *Both schemes (19) and (33) preserve the same steady-states in the fine-grid regime, $R \ll 1$.*

3.3. **Consistency with the damped diffusion limit**

The diffusive scaling of (2) corresponds to (21) at the discrete level; thus, applying the exponential modulation to the IMEX scheme given in [24, eq. (4.2)] with shorthand notation $\gamma_{i-\frac{1}{2},j+\frac{1}{2}} := (\sigma\kappa)_{i-\frac{1}{2},j+\frac{1}{2}}$ brings,

$$\begin{aligned}\begin{pmatrix} f^{+,n+1}_{i,j+1} \\ f^{-,n+1}_{i-1,j} \\ g^{+,n+1}_{i-1,j+1} \\ g^{-,n+1}_{i,j} \end{pmatrix} + \frac{\Delta t}{2\varepsilon R}\left\{ \begin{pmatrix} f^{+,n+1}_{i,j+1} \\ f^{-,n+1}_{i-1,j} \\ g^{+,n+1}_{i-1,j+1} \\ g^{-,n+1}_{i,j} \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} f^{+,n+1}_{i-1,j} \\ f^{-,n+1}_{i,j+1} \\ g^{+,n+1}_{i,j} \\ g^{-,n+1}_{i-1,j+1} \end{pmatrix} \right\} \\ = \begin{pmatrix} f^{+,n}_{i,j+1} \\ f^{-,n}_{i-1,j} \\ g^{+,n}_{i-1,j+1} \\ g^{-,n}_{i,j} \end{pmatrix} e^{-\gamma_{i-\frac{1}{2},j+\frac{1}{2}}\Delta t} + \frac{\Delta t}{2R}\tilde{S}^{1,\varepsilon}_{i-\frac{1}{2},j+\frac{1}{2}} \begin{pmatrix} f^{+,n}_{i-1,j} \\ f^{-,n}_{i,j+1} \\ g^{+,n}_{i,j} \\ g^{-,n}_{i-1,j+1} \end{pmatrix} e^{-\gamma_{i-\frac{1}{2},j+\frac{1}{2}}\Delta t}\end{aligned} \tag{35}$$

where the scattering S-matrix in (33) splits into Maxwellian and diffusive parts,

$$\tilde{S}_{i-\frac{1}{2},j+\frac{1}{2}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} + \varepsilon\tilde{S}^{1,\varepsilon}_{i-\frac{1}{2},j+\frac{1}{2}}.$$

The matrix $\tilde{S}^{1,\varepsilon}_{i-\frac{1}{2},j+\frac{1}{2}}$ contains diffusive incremental coefficients and reads

$$\tilde{S}^{1,\varepsilon}_{i-\frac{1}{2},j+\frac{1}{2}} = \begin{pmatrix} \alpha^\varepsilon - \beta^\varepsilon & -(\alpha^\varepsilon + \beta^\varepsilon) & \beta^\varepsilon & \beta^\varepsilon \\ -(\alpha^\varepsilon + \beta^\varepsilon) & \alpha^\varepsilon - \beta^\varepsilon & \beta^\varepsilon & \beta^\varepsilon \\ \beta^\varepsilon & \beta^\varepsilon & \alpha^\varepsilon - \beta^\varepsilon & -(\alpha^\varepsilon + \beta^\varepsilon) \\ \beta^\varepsilon & \beta^\varepsilon & -(\alpha^\varepsilon + \beta^\varepsilon)s & \alpha^\varepsilon - \beta^\varepsilon \end{pmatrix}_{i-\frac{1}{2},j+\frac{1}{2}},$$

with its entries given by

$$\begin{aligned}\alpha^\varepsilon_{i-\frac{1}{2},j+\frac{1}{2}} &= \frac{1}{\varepsilon + \sigma_{i-\frac{1}{2},j+\frac{1}{2}}(1-\varepsilon^2\kappa_{i-\frac{1}{2},j+\frac{1}{2}})R} \xrightarrow[\varepsilon\to 0]{} \frac{1}{\sigma_{i-\frac{1}{2},j+\frac{1}{2}}R}, \\ \beta^\varepsilon_{i-\frac{1}{2},j+\frac{1}{2}} &= \frac{\sigma_{i-\frac{1}{2},j+\frac{1}{2}}(1-\varepsilon^2\kappa_{i-\frac{1}{2},j+\frac{1}{2}})R}{\varepsilon^2 + (\varepsilon + \sigma_{i-\frac{1}{2},j+\frac{1}{2}}(1-\varepsilon^2\kappa_{i-\frac{1}{2},j+\frac{1}{2}})R)^2} \xrightarrow[\varepsilon\to 0]{} \frac{1}{\sigma_{i-\frac{1}{2},j+\frac{1}{2}}R}.\end{aligned}$$

Oppositely to (22) and (23), here, both coefficients α^ε and β^ε have the same limit, so that the asymptotic scheme will have a 5-points stencil, like the one of [24]. An index-shift in (35) yields:

$$\begin{pmatrix} 1+\frac{\Delta t}{2\varepsilon R} & -\frac{\Delta t}{2\varepsilon R} & 0 & 0 \\ -\frac{\Delta t}{2\varepsilon R} & 1+\frac{\Delta t}{2\varepsilon R} & 0 & 0 \\ 0 & 0 & 1+\frac{\Delta t}{2\varepsilon R} & -\frac{\Delta t}{2\varepsilon R} \\ 0 & 0 & -\frac{\Delta t}{2\varepsilon R} & 1+\frac{\Delta t}{2\varepsilon R} \end{pmatrix} \begin{pmatrix} f_{i,j}^{+,n+1} \\ f_{i,j}^{-,n+1} \\ g_{i,j}^{+,n+1} \\ g_{i,j}^{-,n+1} \end{pmatrix} =$$
$$\begin{pmatrix} e^{-\gamma_{i-\frac12,j-\frac12}\Delta t} & 0 & 0 & 0 \\ 0 & e^{-\gamma_{i+\frac12,j+\frac12}\Delta t} & 0 & 0 \\ 0 & 0 & e^{-\gamma_{i+\frac12,j-\frac12}\Delta t} & 0 \\ 0 & 0 & 0 & e^{-\gamma_{i-\frac12,j+\frac12}\Delta t} \end{pmatrix} \left\{ \begin{pmatrix} f_{i,j}^{+,n} \\ f_{i,j}^{-,n} \\ g_{i,j}^{+,n} \\ g_{i,j}^{-,n} \end{pmatrix} \right. \tag{36}$$
$$\left. + \frac{\Delta t}{2R} \begin{pmatrix} \alpha^\varepsilon_{i-\frac12,j-\frac12}(f^{+,n}_{i-1,j-1} - f^{-,n}_{i,j}) - \beta^\varepsilon_{i-\frac12,j-\frac12}(f^{+,n}_{i-1,j-1} + f^{-,n}_{i,j} - g^{+,n}_{i,j-1} - g^{-,n}_{i-1,j}) \\ \alpha^\varepsilon_{i+\frac12,j+\frac12}(f^{-,n}_{i+1,j+1} - f^{+,n}_{i,j}) - \beta^\varepsilon_{i+\frac12,j+\frac12}(f^{+,n}_{i,j} + f^{-,n}_{i+1,j+1} - g^{+,n}_{i+1,j} - g^{-,n}_{i,j+1}) \\ \alpha^\varepsilon_{i+\frac12,j-\frac12}(g^{+,n}_{i+1,j-1} - g^{-,n}_{i,j}) + \beta^\varepsilon_{i+\frac12,j-\frac12}(f^{+,n}_{i,j-1} + f^{-,n}_{i+1,j} - g^{+,n}_{i+1,j-1} - g^{-,n}_{i,j}) \\ \alpha^\varepsilon_{i-\frac12,j+\frac12}(g^{-,n}_{i-1,j+1} - g^{+,n}_{i,j}) + \beta^\varepsilon_{i-\frac12,j+\frac12}(f^{+,n}_{i-1,j} + f^{-,n}_{i,j+1} - g^{+,n}_{i,j} - g^{-,n}_{i-1,j+1}) \end{pmatrix} \right\}.$$

Letting again

$$\mathfrak{f}^n_{i,j} = f^{+,n}_{i,j} + f^{-,n}_{i,j}, \qquad \mathfrak{g}^n_{i,j} = g^{+,n}_{i,j} + g^{-,n}_{i,j},$$

when $\varepsilon \to 0$, the same situation as (31) is recovered because the left-hand side is identical. Moreover, $\mathfrak{f}, \mathfrak{g}$ again solve identical damped diffusion equations, so the Maxwellian gap $|\mathfrak{f}^n_{i,j} - \mathfrak{g}^n_{i,j}|$ corresponding to well-prepared initial data will remain small. Inserting $f^\pm_{i,j} = \mathfrak{f}_{i,j}/2$ and $g^\pm_{i,j} = \mathfrak{g}_{i,j}/2$ in (36) and adding leads to,

$$\begin{aligned} \mathfrak{f}^{n+1}_{i,j} =& \mathfrak{f}^n_{i,j} \frac{e^{-\gamma_{i-\frac12,j-\frac12}\Delta t} + e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}}{2} \\ &+ \frac{\Delta t}{4R^2} \left(\frac{e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}}{\sigma_{i-\frac12,j-\frac12}} \Big((\mathfrak{g}^n_{i-1,j} - \mathfrak{f}^n_{i,j}) + (\mathfrak{g}^n_{i,j-1} - \mathfrak{f}^n_{i,j}) \Big) \right. \\ &\left. + \frac{e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}}{\sigma_{i+\frac12,j+\frac12}} \Big((\mathfrak{g}^n_{i+1,j} - \mathfrak{f}^n_{i,j}) + (\mathfrak{g}^n_{i,j+1} - \mathfrak{f}^n_{i,j}) \Big) \right), \end{aligned}$$

$$\begin{aligned} \mathfrak{g}^{n+1}_{i,j} =& \mathfrak{g}^n_{i,j} \frac{e^{-\gamma_{i+\frac12,j-\frac12}\Delta t} + e^{-\gamma_{i-\frac12,j+\frac12}\Delta t}}{2} \\ &+ \frac{\Delta t}{4R^2} \left(\frac{e^{-\gamma_{i+\frac12,j-\frac12}\Delta t}}{\sigma_{i-\frac12,j+\frac12}} \Big((\mathfrak{f}^n_{i,j+1} - \mathfrak{g}^n_{i,j}) + (\mathfrak{f}^n_{i-1,j} - \mathfrak{g}^n_{i,j}) \Big) \right. \\ &\left. + \frac{e^{-\gamma_{i-\frac12,j+\frac12}\Delta t}}{\sigma_{i+\frac12,j-\frac12}} \Big((\mathfrak{f}^n_{i+1,j} - \mathfrak{g}^n_{i,j}) + (\mathfrak{f}^n_{i,j-1} - \mathfrak{g}^n_{i,j}) \Big) \right). \end{aligned}$$

Summing former equations and passing to the limit in the coefficients $\alpha^\varepsilon, \beta^\varepsilon$ gives the asymptotic scheme,

$$\begin{aligned}\rho_{i,j}^{n+1} = \rho_{i,j}^n \frac{e^{-\gamma_{i-\frac12,j+\frac12}\Delta t} + e^{-\gamma_{i+\frac12,j+\frac12}\Delta t} + e^{-\gamma_{i+\frac12,j-\frac12}\Delta t} + e^{-\gamma_{i-\frac12,j+\frac12}\Delta t}}{4} \\ + \frac{\Delta t}{8R^2}\left\{ \frac{e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}}{\sigma_{i-\frac12,j-\frac12}}(\rho_{i,j-1}^n + \rho_{i-1,j}^n - 2\rho_{i,j}^n) + \frac{e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}}{\sigma_{i+\frac12,j+\frac12}}(\rho_{i+1,j}^n + \rho_{i,j+1}^n - 2\rho_{i,j}^n) \right. \\ \left. + \frac{e^{-\gamma_{i+\frac12,j-\frac12}\Delta t}}{\sigma_{i+\frac12,j-\frac12}}(\rho_{i,j-1}^n + \rho_{i+1,j}^n - 2\rho_{i,j}^n) + \frac{e^{-\gamma_{i-\frac12,j+\frac12}\Delta t}}{\sigma_{i-\frac12,j+\frac12}}(\rho_{i-1,j}^n + \rho_{i,j+1}^n - 2\rho_{i,j}^n) \right\}. \end{aligned} \tag{37}$$

which is clearly consistent with (3).

Lemma 4. *Under the parabolic CFL restriction $2\Delta t \le \min(\sigma)R^2$, the scheme (37) preserves positivity and is consistent with (3) when $R \to 0$.*

Proof. Since $\sigma \ge 0$, the only negative incremental coefficients in (37) are the ones acting on $\rho_{i,j}^n$, hence a sufficient condition for positivity preservation reads

$$\frac14 - \frac{\Delta t}{8\min(\sigma)R^2} \ge 0.$$

Consistency with (3) proceeds again by assuming $\Delta t = O(R^2) \ll 1$ and linearizing exponentials, (37) rewrites

$$\begin{aligned}\rho_{i,j}^{n+1} = \rho_{i,j}^n \left(1 - \Delta t \frac{\gamma_{i+\frac12,j+\frac12} + \gamma_{i-\frac12,j-\frac12} + \gamma_{i+\frac12,j-\frac12} + \gamma_{i-\frac12,j+\frac12}}{4}\right) \\ + \frac{\Delta t}{4R^2}\left(\Big(\frac{1}{2\sigma_{i+\frac12,j+\frac12}} + \frac{1}{2\sigma_{i+\frac12,j-\frac12}}\Big)(\rho_{i+1,j}^n - \rho_{i,j}^n) \right. \\ + \Big(\frac{1}{2\sigma_{i+\frac12,j+\frac12}} + \frac{1}{2\sigma_{i-\frac12,j+\frac12}}\Big)(\rho_{i,j+1}^n - \rho_{i,j}^n) \\ - \Big(\frac{1}{2\sigma_{i-\frac12,j+\frac12}} + \frac{1}{2\sigma_{i-\frac12,j-\frac12}}\Big)(\rho_{i,j}^n - \rho_{i-1,j}^n) \\ \left. - \Big(\frac{1}{2\sigma_{i+\frac12,j-\frac12}} + \frac{1}{2\sigma_{i-\frac12,j-\frac12}}\Big)(\rho_{i,j}^n - \rho_{i,j-1}^n)\right) + O(R^2),\end{aligned}$$

after neglecting $O(\Delta t^2/R^2 \simeq R^2)$ terms. The scheme given in [24, eq. (4.8)] is recovered when $\kappa \to 0$. □

3.4. **Rigorous diffusive limit for constant parameters**

For $(u_{i,j})$ a real sequence, we shall use the notations

$$\delta u_{i+\frac12,j} = u_{i+1,j} - u_{i,j}, \quad \delta u_{i,j+\frac12} = u_{i,j+1} - u_{i,j}, \quad \|u\|_1 = \sum_{i,j} \Delta x^2 |u_{i,j}|,$$

$$TV(u) = \sum_{i,j} \Delta x(|\delta u_{i+\frac12,j}| + |\delta u_{i,j+\frac12}|), \quad \|\Delta u\|_1 = \sum_{i,j} |u_{i+1,j} + u_{i,j+1} + u_{i-1,j} + u_{i,j-1} - 4u_{i,j}|.$$

In order to make the above formal computation rigorous, we first derive estimates for the IMEX scheme (36).

Lemma 5. *Let us assume that the CFL parabolic condition*

$$\Delta t \le \frac23 R^2 \sigma_{min}(1 - \varepsilon^2 \kappa_{max}) \tag{38}$$

holds. Then, the scheme is nonnegative. Moreover, if the initial data are bounded in $L^1\cap L^\infty$, then the sequences $f^{\varepsilon\pm,n}_{i,j}$ and $g^{\varepsilon\pm,n}_{i,j}$ are bounded in L^∞, and we have the dissipation inequality

$$\|f^{\varepsilon+,n+1}\|_1+\|f^{\varepsilon-,n+1}\|_1+\|g^{\varepsilon+,n+1}\|_1+\|g^{\varepsilon-,n+1}\|_1\le\sum_{i,j}(f^{\varepsilon+,n}_{i,j}+f^{\varepsilon-,n}_{i,j}+g^{\varepsilon+,n}_{i,j}+g^{\varepsilon-,n}_{i,j})e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}.$$

Proof. Inverting the block-diagonal matrix in the left hand side in (36), we get

$$\begin{aligned}
f^{\varepsilon+,n+1}_{i,j} &= \frac{1}{2\varepsilon R+2\Delta t}\Big((2\varepsilon R+\Delta t-\frac{\Delta t^2}{2R}(\alpha^\varepsilon_{i+\frac12,j+\frac12}+\beta^\varepsilon_{i+\frac12,j+\frac12}))f^{\varepsilon+,n}_{i,j}e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}\\
&+(\Delta t-\frac{\Delta t}{2R}(2\varepsilon R+\Delta t)(\alpha^\varepsilon_{i-\frac12,j-\frac12}+\beta^\varepsilon_{i-\frac12,j-\frac12}))f^{\varepsilon-,n}_{i,j}e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}\\
&+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}(\alpha^\varepsilon_{i-\frac12,j-\frac12}-\beta^\varepsilon_{i-\frac12,j-\frac12})f^{\varepsilon+,n}_{i-1,j-1}e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}\\
&+\frac{\Delta t^2}{2R}(\alpha^\varepsilon_{i+\frac12,j+\frac12}-\beta^\varepsilon_{i+\frac12,j+\frac12})e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}f^{\varepsilon-,n}_{i+1,j+1}\\
&+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}\beta^\varepsilon_{i-\frac12,j-\frac12}e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}(g^{\varepsilon+,n}_{i,j-1}+g^{\varepsilon-,n}_{i-1,j})\\
&+\frac{\Delta t^2}{2R}\beta^\varepsilon_{i+\frac12,j+\frac12}e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}(g^{\varepsilon+,n}_{i+1,j}+g^{\varepsilon-,n}_{i,j+1})\Big),
\end{aligned}$$

$$\begin{aligned}
f^{\varepsilon-,n+1}_{i,j} &= \frac{1}{2\varepsilon R+2\Delta t}\Big((\Delta t-(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}(\alpha^\varepsilon_{i+\frac12,j+\frac12}+\beta^\varepsilon_{i+\frac12,j+\frac12}))e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}f^{\varepsilon+,n}_{i,j}\\
&+(2\varepsilon R+\Delta t-\frac{\Delta t^2}{2R}(\alpha^\varepsilon_{i-\frac12,j-\frac12}+\beta^\varepsilon_{i-\frac12,j-\frac12}))e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}f^{\varepsilon-,n}_{i,j}\\
&+\frac{\Delta t^2}{2R}(\alpha^\varepsilon_{i-\frac12,j-\frac12}-\beta^\varepsilon_{i-\frac12,j-\frac12})e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}f^{\varepsilon+,n}_{i-1,j-1}\\
&+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}(\alpha^\varepsilon_{i+\frac12,j+\frac12}-\beta^\varepsilon_{i+\frac12,j+\frac12})e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}f^{\varepsilon-,n}_{i+1,j+1}\\
&+\frac{\Delta t^2}{2R}\beta^\varepsilon_{i-\frac12,j-\frac12}e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}(g^{\varepsilon+,n}_{i,j-1}+g^{\varepsilon-,n}_{i-1,j})\\
&+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}\beta^\varepsilon_{i+\frac12,j+\frac12}e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}(g^{\varepsilon+,n}_{i+1,j}+g^{\varepsilon-,n}_{i,j+1})\Big),
\end{aligned}$$

and similar expressions for $g^{\varepsilon+,n+1}_{i,j}$ and $g^{\varepsilon-,n+1}_{i,j}$. We notice that

$$\alpha^\varepsilon_{i+\frac12,j+\frac12}-\beta^\varepsilon_{i+\frac12,j+\frac12}=\frac{2\varepsilon^2+\sigma_{i+\frac12,j+\frac12}\varepsilon(1-\varepsilon^2\kappa_{i+\frac12,j+\frac12})R}{(\varepsilon+\sigma_{i+\frac12,j+\frac12}(1-\varepsilon^2\kappa_{i+\frac12,j+\frac12})R)(\varepsilon^2+(\varepsilon+\sigma_{i+\frac12,j+\frac12}(1-\varepsilon^2\kappa_{i+\frac12,j+\frac12})R)^2)}\ge0.$$

Hence, $f^{\varepsilon\pm,n+1}_{i,j}$ and $g^{\varepsilon\pm,n+1}_{i,j}$ are a positive combination of $f^{\varepsilon\pm,n}_{i,j}$ and $g^{\varepsilon\pm,n}_{i,j}$ iff we have

$$\frac{\Delta t^2}{2R}(\alpha^\varepsilon_{i+\frac12,j+\frac12}+\beta^\varepsilon_{i+\frac12,j+\frac12})\le2\varepsilon R+\Delta t,\quad\text{ and }\quad\frac{\Delta t}{2R}(2\varepsilon R+\Delta t)(\alpha^\varepsilon_{i+\frac12,j+\frac12}+\beta^\varepsilon_{i+\frac12,j+\frac12})\le\Delta t.$$

The second condition being more restrictive than the first one, we only have to verify this latter condition, i.e.

$$(\varepsilon+\frac{\Delta t}{2R})(\alpha^\varepsilon_{i+\frac12,j+\frac12}+\beta^\varepsilon_{i+\frac12,j+\frac12})\le1.$$

We have

$$\alpha^{\varepsilon}_{i+\frac12,j+\frac12}+\beta^{\varepsilon}_{i+\frac12,j+\frac12}=\psi\big(\sigma_{i+\frac12,j+\frac12}(1-\varepsilon^2\kappa_{i+\frac12,j+\frac12})R\big),$$

where $\psi(x)=\frac{1}{\varepsilon+x}+\frac{x}{\varepsilon^2+(\varepsilon+x)^2}$ is a nonincreasing function on $(0,+\infty)$. The above inequality is satisfied provided

$$(\varepsilon+\frac{\Delta t}{2R})\psi(\sigma_{\min}(1-\varepsilon^2\kappa_{\max})R)\le 1.$$

Since $\frac{1}{\psi(x)}-\varepsilon\ge\frac{x}{3}$ for $x>0$, the latter inequality is satisfied provided condition (38) holds. Adding each line of (36) and summing over i and j, we deduce the dissipation of the L^1-norm as stated in the Lemma. Moreover, we verify easily from above expressions that if $0\le f^{\varepsilon\pm,n}\le M$ and $0\le g^{\varepsilon\pm,n}\le M$, then for all i,j,

$$\begin{aligned}
f^{\varepsilon+,n+1}_{i,j}&\le\frac{M}{2\varepsilon R+2\Delta t}\Big((2\varepsilon R+\Delta t)e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}+\Delta t\,e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}\Big)\le M,\\
f^{\varepsilon-,n+1}_{i,j}&\le\frac{M}{2\varepsilon R+2\Delta t}\Big(\Delta t\,e^{-\gamma_{i+\frac12,j+\frac12}\Delta t}+(2\varepsilon R+\Delta t)e^{-\gamma_{i-\frac12,j-\frac12}\Delta t}\Big)\le M,
\end{aligned}$$

and similarily $0\le g^{pm,n+1}_{i,j}\le M$. It provides the L^∞ bound. □

Constant coefficients. We assume now that $\sigma(\mathbf{x})$ and $\kappa(\mathbf{x})$ are constant and we study rigorously the diffusive limit when $\varepsilon\to 0$. We recall the definitions

$$\alpha^{\varepsilon}=\frac{1}{\varepsilon+\sigma(1-\varepsilon^2\kappa)R},\qquad \beta^{\varepsilon}=\frac{\sigma(1-\varepsilon^2\kappa)R}{\varepsilon^2+(\varepsilon+\sigma(1-\varepsilon^2\kappa)R)^2},\qquad \gamma=\sigma\kappa. \tag{39}$$

Then the scheme (36) rewrites

$$\begin{aligned}
&\begin{pmatrix}1+\frac{\Delta t}{2\varepsilon R}&-\frac{\Delta t}{2\varepsilon R}&0&0\\-\frac{\Delta t}{2\varepsilon R}&1+\frac{\Delta t}{2\varepsilon R}&0&0\\0&0&1+\frac{\Delta t}{2\varepsilon R}&-\frac{\Delta t}{2\varepsilon R}\\0&0&-\frac{\Delta t}{2\varepsilon R}&1+\frac{\Delta t}{2\varepsilon R}\end{pmatrix}\begin{pmatrix}f^{\varepsilon+,n+1}_{i,j}\\f^{\varepsilon-,n+1}_{i,j}\\g^{\varepsilon+,n+1}_{i,j}\\g^{\varepsilon-,n+1}_{i,j}\end{pmatrix}=\begin{pmatrix}f^{\varepsilon+,n}_{i,j}\\f^{\varepsilon-,n}_{i,j}\\g^{\varepsilon+,n}_{i,j}\\g^{\varepsilon-,n}_{i,j}\end{pmatrix}e^{-\sigma\kappa\Delta t}\\
&+\frac{\Delta t}{2R}\begin{pmatrix}\alpha^{\varepsilon}(f^{\varepsilon+,n}_{i-1,j-1}-f^{\varepsilon-,n}_{i,j})-\beta^{\varepsilon}(f^{\varepsilon+,n}_{i-1,j-1}+f^{\varepsilon-,n}_{i,j}-g^{\varepsilon+,n}_{i,j-1}-g^{\varepsilon-,n}_{i-1,j})\\
\alpha^{\varepsilon}(f^{\varepsilon-,n}_{i+1,j+1}-f^{\varepsilon+,n}_{i,j})-\beta^{\varepsilon}(f^{\varepsilon+,n}_{i,j}+f^{\varepsilon-,n}_{i+1,j+1}-g^{\varepsilon+,n}_{i+1,j}-g^{\varepsilon-,n}_{i,j+1})\\
\alpha^{\varepsilon}(g^{\varepsilon+,n}_{i+1,j-1}-g^{\varepsilon-,n}_{i,j})+\beta^{\varepsilon}(f^{\varepsilon+,n}_{i,j-1}+f^{\varepsilon-,n}_{i+1,j}-g^{\varepsilon+,n}_{i+1,j-1}-g^{\varepsilon-,n}_{i,j})\\
\alpha^{\varepsilon}(g^{\varepsilon-,n}_{i-1,j+1}-g^{\varepsilon+,n}_{i,j})+\beta^{\varepsilon}(f^{\varepsilon+,n}_{i-1,j}+f^{\varepsilon-,n}_{i,j+1}-g^{\varepsilon+,n}_{i,j}-g^{\varepsilon-,n}_{i-1,j+1})\end{pmatrix}e^{-\sigma\kappa\Delta t}.
\end{aligned} \tag{40}$$

Lemma 6. *Let $\sigma,\kappa\in\mathbb{R}^+$ be constants; under the parabolic CFL condition* (38)*, the scheme* (40) *is TVD.*

$$\begin{aligned}
&TV(f^{\varepsilon+,n+1})+TV(f^{\varepsilon-,n+1})+TV(g^{\varepsilon+,n+1})+TV(g^{\varepsilon-,n+1})\\
&\qquad\le e^{-\sigma\kappa\Delta t}\left(TV(f^{\varepsilon+,n})+TV(f^{\varepsilon-,n})+TV(g^{\varepsilon+,n})+TV(g^{\varepsilon-,n})\right).
\end{aligned}$$

Proof. As in the proof of [24, Lemma 5.4], inverting the block-diagonal matrix in the left hand side in (40) gives

$$\begin{aligned}
f^{\varepsilon+,n+1}_{i,j}&=\frac{e^{-\sigma\kappa\Delta t}}{2\varepsilon R+2\Delta t}\Big((2\varepsilon R+\Delta t-\frac{\Delta t^2}{2R}(\alpha^{\varepsilon}+\beta^{\varepsilon}))f^{\varepsilon+,n}_{i,j}\\
&+(\Delta t-\frac{\Delta t}{2R}(2\varepsilon R+\Delta t)(\alpha^{\varepsilon}+\beta^{\varepsilon}))f^{\varepsilon-,n}_{i,j}+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}(\alpha^{\varepsilon}-\beta^{\varepsilon})f^{\varepsilon+,n}_{i-1,j-1}\\
&+\frac{\Delta t^2}{2R}(\alpha^{\varepsilon}-\beta^{\varepsilon})f^{\varepsilon-,n}_{i+1,j+1}+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}\beta^{\varepsilon}(g^{\varepsilon+,n}_{i,j-1}+g^{\varepsilon-,n}_{i-1,j})+\frac{\Delta t^2}{2R}\beta^{\varepsilon}(g^{\varepsilon+,n}_{i+1,j}+g^{\varepsilon-,n}_{i,j+1})\Big),
\end{aligned}$$

$$
\begin{aligned}
f^{\varepsilon\,-,n+1}_{i,j} \;&= \frac{e^{-\sigma\kappa\Delta t}}{2\varepsilon R+2\Delta t}\Big((\Delta t-(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}(\alpha^\varepsilon+\beta^\varepsilon))f^{\varepsilon\,+,n}_{i,j}\\
&\quad+(2\varepsilon R+\Delta t-\frac{\Delta t^2}{2R}(\alpha^\varepsilon+\beta^\varepsilon))f^{\varepsilon\,-,n}_{i,j}+\frac{\Delta t^2}{2R}(\alpha^\varepsilon-\beta^\varepsilon)f^{\varepsilon\,+,n}_{i-1,j-1}\\
&\quad+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}(\alpha^\varepsilon-\beta^\varepsilon)f^{\varepsilon\,-,n}_{i+1,j+1}+\frac{\Delta t^2}{2R}\beta^\varepsilon(g^{\varepsilon\,+,n}_{i,j-1}+g^{\varepsilon\,-,n}_{i-1,j})\\
&\quad+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}\beta^\varepsilon(g^{\varepsilon\,+,n}_{i+1,j}+g^{\varepsilon\,-,n}_{i,j+1})\Big),
\end{aligned}
$$

and similar expressions for $g^{\varepsilon\,+,n+1}_{i,j}$ and $g^{\varepsilon\,-,n+1}_{i,j}$. Under the condition (38), each coefficient in the left hand side are nonnegative. Then by linearity, and after using a triangle inequality, we obtain

$$
\begin{aligned}
|\delta f^{+,n+1}_{i+\frac12,j}| \;&\le \frac{e^{-\sigma\kappa\Delta t}}{2\varepsilon R+2\Delta t}\Big((2\varepsilon R+\Delta t-\frac{\Delta t^2}{2R}(\alpha^\varepsilon+\beta^\varepsilon))|\delta f^{+,n}_{i+\frac12,j}|\\
&\quad+(\Delta t-\frac{\Delta t}{2R}(2\varepsilon R+\Delta t)(\alpha^\varepsilon+\beta^\varepsilon))|\delta f^{-,n}_{i+\frac12,j}|+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}(\alpha^\varepsilon-\beta^\varepsilon)|\delta f^{+,n}_{i-\frac12,j-1}|\\
&\quad+\frac{\Delta t^2}{2R}(\alpha^\varepsilon-\beta^\varepsilon)|\delta f^{-,n}_{i+\frac32,j+1}|+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}\beta^\varepsilon(|\delta g^{+,n}_{i+\frac12,j-1}|+|\delta g^{-,n}_{i-\frac12,j}|)\\
&\quad+\frac{\Delta t^2}{2R}\beta^\varepsilon(|\delta g^{+,n}_{i+\frac32,j}|+|\delta g^{-,n}_{i+\frac12,j+1}|)\Big),
\end{aligned}
$$

$$
\begin{aligned}
|\delta f^{-,n+1}_{i+\frac12,j}| \;&\le \frac{e^{-\sigma\kappa\Delta t}}{2\varepsilon R+2\Delta t}\Big((\Delta t-(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}(\alpha^\varepsilon+\beta^\varepsilon))|\delta f^{+,n}_{i+\frac12,j}|\\
&\quad+(2\varepsilon R+\Delta t-\frac{\Delta t^2}{2R}(\alpha^\varepsilon+\beta^\varepsilon))|\delta f^{-,n}_{i+\frac12,j}|+\frac{\Delta t^2}{2R}(\alpha^\varepsilon-\beta^\varepsilon)|\delta f^{+,n}_{i-\frac12,j-1}|\\
&\quad+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}(\alpha^\varepsilon-\beta^\varepsilon)|\delta f^{-,n}_{i+\frac32,j+1}|+\frac{\Delta t^2}{2R}\beta^\varepsilon(|\delta g^{+,n}_{i+\frac12,j-1}|+|\delta g^{-,n}_{i-\frac12,j}|)\\
&\quad+(2\varepsilon R+\Delta t)\frac{\Delta t}{2R}\beta^\varepsilon(|\delta g^{+,n}_{i+\frac32,j}|+|\delta g^{-,n}_{i+\frac12,j+1}|)\Big),
\end{aligned}
$$

with similar expressions for $|\delta g^{+,n}_{i+\frac12,j}|$ and $|\delta g^{-,n}_{i+\frac12,j}|$. Adding all these expressions, we obtain

$$
\begin{aligned}
&|\delta f^{+,n+1}_{i+\frac12,j}|+|\delta f^{-,n+1}_{i+\frac12,j}|+|\delta g^{+,n+1}_{i+\frac12,j}|+|\delta g^{-,n+1}_{i+\frac12,j}|\le e^{-\sigma\kappa\Delta t}\times\\
&\quad\Big[(1-\frac{\Delta t}{2R}(\alpha^\varepsilon+\beta^\varepsilon))(|\delta f^{+,n}_{i+\frac12,j}|+|\delta f^{-,n}_{i+\frac12,j}|+|\delta g^{+,n}_{i+\frac12,j}|+|\delta g^{-,n}_{i+\frac12,j}|)\\
&\qquad+\frac{\Delta t}{2R}(\alpha^\varepsilon-\beta^\varepsilon)(|\delta f^{+,n}_{i-\frac12,j}|+|\delta f^{-,n}_{i+\frac32,j}|+|\delta g^{+,n}_{i+\frac12,j}|+|\delta g^{-,n}_{i+\frac12,j}|)\\
&\qquad+\frac{\Delta t}{2R}\beta^\varepsilon(|\delta f^{+,n}_{i+\frac12,j-1}|+|\delta f^{-,n}_{i+\frac32,j}|+|\delta f^{+,n}_{i-\frac12,j}|+|\delta f^{-,n}_{i+\frac12,j+1}|)\\
&\qquad+\frac{\Delta t}{2R}\beta^\varepsilon(|\delta g^{+,n}_{i+\frac12,j-1}|+|\delta g^{+,n}_{i+\frac32,j}|+|\delta g^{-,n}_{i-\frac12,j}|+|\delta g^{-,n}_{i+\frac12,j+1}|)\Big].
\end{aligned}
$$

Summing over i and j, we get after shifting the indexes,

$$
\begin{aligned}
&\sum_{i,j}\Big(|\delta f^{+,n+1}_{i+\frac12,j}|+|\delta f^{-,n+1}_{i+\frac12,j}|+|\delta g^{+,n+1}_{i+\frac12,j}|+|\delta g^{-,n+1}_{i+\frac12,j}|\Big)\\
&\qquad\le e^{-\sigma\kappa\Delta t}\sum_{i,j}\Big(|\delta f^{+,n}_{i+\frac12,j}|+|\delta f^{-,n}_{i+\frac12,j}|+|\delta g^{+,n}_{i+\frac12,j}|+|\delta g^{-,n}_{i+\frac12,j}|\Big).
\end{aligned}
$$

□

Theorem 2 (Asymptotic preserving property)**.** *Let us assume that the parabolic stability condition*

$$\Delta t \leq \frac{1}{2}\sigma R^2, \tag{41}$$

holds and that initial data, independent of ε, are smooth enough so that,

$$\|\Delta f^{+,0}\|_1 + \|\Delta f^{-,0}\|_1 + \|\Delta g^{+,0}\|_1 + \|\Delta g^{-,0}\|_1 \leq C.$$

for some constant $C \geq 0$. Then, the sequences $(f^{\varepsilon\pm,n}_{i,j})$ and $(g^{\varepsilon\pm,n}_{i,j})$ converge, as $\varepsilon \to 0$, towards limits, denoted respectively $(f^{\pm,n}_{i,j})$ and $(g^{\pm,n}_{i,j})$ which satisfy:

$$f^{+,n}_{i,j} = f^{-,n}_{i,j} = \frac{1}{2}\mathfrak{f}^n_{i,j}, \qquad g^{+,n}_{i,j} = g^{-,n}_{i,j} = \frac{1}{2}\mathfrak{g}^n_{i,j},$$

where

$$\mathfrak{f}^{n+1}_{i,j} = \mathfrak{f}^n_{i,j}e^{-\sigma\kappa\Delta t} + \frac{\Delta t}{4\sigma R^2}(\mathfrak{g}^n_{i,j-1} + \mathfrak{g}^n_{i-1,j} + \mathfrak{g}^n_{i+1,j} + \mathfrak{g}^n_{i,j+1} - 4\mathfrak{f}^n_{i,j})e^{-\sigma\kappa\Delta t} \tag{42}$$

$$\mathfrak{g}^{n+1}_{i,j} = \mathfrak{g}^n_{i,j}e^{-\sigma\kappa\Delta t} + \frac{\Delta t}{4\sigma R^2}(\mathfrak{f}^n_{i,j-1} + \mathfrak{f}^n_{i+1,j} + \mathfrak{f}^n_{i-1,j} + \mathfrak{f}^n_{i,j+1} - 4\mathfrak{g}^n_{i,j})e^{-\sigma\kappa\Delta t}. \tag{43}$$

Moreover, we have, for all $n \in \mathbb{N}^*$,

$$\boxed{\|\mathfrak{f}^n - \mathfrak{g}^n\|_1 \leq e^{-(2/(\sigma R^2)+\sigma\kappa)n\Delta t}\,\|\mathfrak{f}^0 - \mathfrak{g}^0\|_1 + CR^2.}$$

By summing the equations (42)-(43), we deduce the following result:

Corollary 3. *Let us denote $\rho^n = \frac{1}{4}(\mathfrak{f}^n + \mathfrak{g}^n)$. Under the same assumptions as in Theorem 2, we have*

$$\boxed{\rho^{n+1}_{i,j} = \rho^n_{i,j}e^{-\sigma\kappa\Delta t} + \frac{\Delta t}{4\sigma R^2}\Big(\rho^n_{i,j-1} + \rho^n_{i,j+1} + \rho^n_{i-1,j} + \rho^n_{i+1,j} - 4\rho^n_{i,j}\Big)e^{-\sigma\kappa\Delta t},}$$

and $f^{+,n} = \rho^n + O(R^2)$, $f^{-,n} = \rho^n + O(R^2)$, $g^{+,n} = \rho^n + O(R^2)$, $g^{-,n} = \rho^n + O(R^2)$.

Proof. By the same token as in the proof of Lemma 6, we deduce that the sequences $(f^{\varepsilon\pm,n}_{i,j})$, and $(g^{\varepsilon\pm,n}_{i,j})$ are Cauchy sequences with respect to ε in ℓ^1. Thus, we deduce that, when $\varepsilon \to 0$, they converge to some limit sequences denoted respectively $(f^{\pm,n}_{i,j})$, and $(g^{\pm,n}_{i,j})$. We may pass into the limit $\varepsilon \to 0$ in (40), we deduce

$$\begin{pmatrix} f^{+,n+1}_{i,j} \\ f^{-,n+1}_{i,j} \\ g^{+,n+1}_{i,j} \\ g^{-,n+1}_{i,j} \end{pmatrix} \in \mathrm{Ker}(H_0) = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\},$$

where we recall the definition of the matrix

$$H_0 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

Thus, the limit verifies $f^{+,n+1}_{i,j} = f^{-,n+1}_{i,j}$ and $g^{+,n+1}_{i,j} = g^{-,n+1}_{i,j}$ for all $i \in \mathbb{Z}$, $j \in \mathbb{Z}$ and $n \in \mathbb{N}$. Denoting $\mathfrak{f}^{\varepsilon\,n}_{i,j} = f^{\varepsilon+,n}_{i,j} + f^{\varepsilon-,n}_{i,j}$ and $\mathfrak{g}^{\varepsilon\,n}_{i,j} = g^{\varepsilon+,n}_{i,j} + g^{\varepsilon-,n}_{i,j}$, we obtain an equation for $\mathfrak{f}^{\varepsilon\,n}_{i,j}$ and $\mathfrak{g}^{\varepsilon\,n}_{i,j}$ by adding the first two and the last two equations in (40) :

$$\begin{aligned}
\mathfrak{f}^{\varepsilon\,n+1}_{i,j} &= \mathfrak{f}^{\varepsilon\,n}_{i,j} e^{-\sigma\kappa\Delta t} + \frac{\Delta t}{2R}(\alpha^\varepsilon(f^{\varepsilon+,n}_{i-1,j-1} - f^{\varepsilon-,n}_{i,j}) + \alpha^\varepsilon(f^{\varepsilon-,n}_{i+1,j+1} - f^{\varepsilon+,n}_{i,j}))e^{-\sigma\kappa\Delta t} \\
&\quad - \frac{\Delta t}{2R}(\beta^\varepsilon(f^{\varepsilon+,n}_{i-1,j-1} + f^{\varepsilon\,-,n}_{i,j} - g^{\varepsilon+,n}_{i,j-1} - g^{\varepsilon-,n}_{i-1,j}) \\
&\qquad + \beta^\varepsilon(f^{\varepsilon+,n}_{i,j} + f^{\varepsilon-,n}_{i+1,j+1} - g^{\varepsilon+,n}_{i+1,j} - g^{\varepsilon-,n}_{i,j+1}))e^{-\sigma\kappa\Delta t}; \qquad (44) \\
\mathfrak{g}^{\varepsilon\,n+1}_{i,j} &= \mathfrak{g}^{\varepsilon\,n}_{i,j} e^{-\sigma\kappa\Delta t} + \frac{\Delta t}{2R}(\alpha^\varepsilon(g^{\varepsilon+,n}_{i+1,j-1} - g^{\varepsilon-,n}_{i,j}) + \alpha^\varepsilon(g^{\varepsilon-,n}_{i-1,j+1} - g^{\varepsilon+,n}_{i,j}))e^{-\sigma\kappa\Delta t} \\
&\quad + \frac{\Delta t}{2R}(\beta^\varepsilon(f^{\varepsilon+,n}_{i,j-1} + f^{\varepsilon-,n}_{i+1,j} - g^{\varepsilon+,n}_{i+1,j-1} - g^{\varepsilon-,n}_{i,j}) \\
&\qquad + \beta^\varepsilon(f^{\varepsilon+,n}_{i-1,j} + f^{\varepsilon-,n}_{i,j+1} - g^{\varepsilon+,n}_{i,j} - g^{\varepsilon-,n}_{i-1,j+1}))e^{-\sigma\kappa\Delta t}. \qquad (45)
\end{aligned}$$

From the expressions of α^ε and β^ε in (39), we deduce that $\alpha^\varepsilon \to \frac{1}{\sigma R}$ and $\beta^\varepsilon \to \frac{1}{\sigma R}$ when $\varepsilon \to 0$. Then passing into the limit in (44)–(45), recalling that $f^{\pm,n}_{i,j} = \frac{1}{2}\mathfrak{f}^n_{i,j}$ and $g^{\pm,n}_{i,j} = \frac{1}{2}\mathfrak{g}^n_{i,j}$, we obtain (42)-(43). Then, denoting $D^n_{i,j} = \mathfrak{f}^n_{i,j} - \mathfrak{g}^n_{i,j}$, we get straightforwardly from (42)-(43)

$$D^{n+1}_{i,j} = D^n_{i,j}(1 - \frac{2\Delta t}{\sigma R^2})e^{-\sigma\kappa\Delta t} + \frac{\Delta t}{4\sigma R^2}(4D^n_{i,j} - D^n_{i,j-1} - D^n_{i-1,j} - D^n_{i+1,j} - D^n_{i,j+1})e^{-\sigma\kappa\Delta t}, \qquad (46)$$

and

$$\begin{aligned}
\|\Delta\mathfrak{f}^{n+1}\|_1 &\le \|\Delta\mathfrak{f}^n\|_1\big(1 - \frac{\Delta t}{\sigma R^2}\big)e^{-\sigma\kappa\Delta t} + \frac{\Delta t}{\sigma R^2}\|\Delta\mathfrak{g}^n\|_1\, e^{-\sigma\kappa\Delta t} \\
\|\Delta\mathfrak{g}^{n+1}\|_1 &\le \|\Delta\mathfrak{g}^n\|_1\big(1 - \frac{\Delta t}{\sigma R^2}\big)e^{-\sigma\kappa\Delta t} + \frac{\Delta t}{\sigma R^2}\|\Delta\mathfrak{f}^n\|_1\, e^{-\sigma\kappa\Delta t}.
\end{aligned}$$

From the assumptions on the initial data in Theorem 2, we have $\|\Delta\mathfrak{f}^0\|_1 + \|\Delta\mathfrak{g}^0\|_1 \le C$. Then, for all $n \in \mathbb{N}$, we have $\|\Delta\mathfrak{f}^n\|_1 + \|\Delta\mathfrak{g}^n\|_1 \le Ce^{-\sigma n\kappa\Delta t}$, we deduce

$$\sum_{i,j} |4D^n_{i,j} - D^n_{i,j-1} - D^n_{i-1,j} - D^n_{i+1,j} - D^n_{i,j+1}| \le Ce^{-\sigma n\kappa\Delta t}.$$

We may inject this latter inequality into (46). Taking the absolute value, under the condition (41), and summing over i and j, we deduce

$$\|D^{n+1}\|_1 = \sum_{i,j} \Delta x^2 |D^{n+1}_{i,j}| \le \|D^n\|_1(1 - \frac{2\Delta t}{\sigma R^2})e^{-\sigma\kappa\Delta t} + C\frac{\Delta t}{\sigma},$$

for some nonnegative constant C. Applying a discrete Gronwall inequality, we get

$$\|D^n\|_1 \ \le \|D^0\|_1 e^{-(2/(\sigma R^2)+\sigma\kappa)n\Delta t} + C\frac{\Delta t}{\sigma}\sum_{k=0}^{n-1}\Big(1 - \frac{2\Delta t}{\sigma R^2}\Big)^k \le \|D^0\|_1 e^{-(2/(\sigma R^2)+\sigma\kappa)n\Delta t} + \frac{C}{2}R^2.$$

□

4. Numerical results

All the practical tests displayed hereafter are conducted on the unit square coarsely gridded with 32×32 points in order to demonstrate both the robustness and accuracy of the time-marching schemes.

4.1. **Kinetic scaling**

Both schemes (19) and (33) are iterated up to $T = 7$, with $\Delta t = 0.975\Delta x$ and

$$\sigma(\mathbf{x}) = 149.5 \max(|x - 0.5|, |y - 0.5|) < 0.25) + 0.5,$$
$$\kappa(\mathbf{x}) = 0.9 \max(|x - 0.5|, |y - 0.5|) < 0.25) + 0.05.$$

Initial data are null but an inflow boundary condition is specified on the left side,

$$f^+(x = 0, \cdot) = g^-(x = 0, \cdot) = 1.$$

along with specular reflection on all the other sides of the computational domain. The velocity field $\vec{V}$ reads

$$\vec{V}(t, \mathbf{x}) := \begin{pmatrix} \frac{f^+(t,\mathbf{x}) - f^-(t,\mathbf{x})}{\rho(t,\mathbf{x})} \\ \frac{g^+(t,\mathbf{x}) - g^-(t,\mathbf{x})}{\rho(t,\mathbf{x})} \end{pmatrix}.$$

Results are displayed on Fig. 3: despite the stiffness of the benchmark, results are quite similar, except for (33) showing slightly more numerical viscosity (compare the macroscopic densities on the first row of Fig. 3). Both time-marching algorithms succeed in stabilizing correctly, being a consequence of 2D well-balanced properties.

4.2. **Diffusive scaling**

Hereafter, both IMEX schemes (26) and (35) were set up on the same computational grid with $\varepsilon = 10^{-5}$,

$$\sigma(\mathbf{x}) = 35 \exp(-25((x - 0.5)^2 + (y - 0.5)^2)) + 15,$$
$$\kappa(\mathbf{x}) = 0.9 \exp(-25((x - 0.5)^2 + (y - 0.5)^2)) + 0.05,$$

and Maxwellian initial data,

$$f^{\pm}(\mathbf{x}) = g^{\pm}(\mathbf{x}) = \frac{1}{4}\Big(\exp(-250(x - 0.375)^2) + \exp(-250(x - 0.625)^2) \Big) \otimes \Big(\exp(-250(y - 0.375)^2) + \exp(-250(y - 0.635)^2 \Big).$$

Results at time $T = 0.15$ are given in Fig.4: both schemes appear to be quite similar. Especially, the Maxwellian gap $\mathfrak{f} - \mathfrak{g}$ is practically identical for both algorithms, the one produced by (26) being slightly bigger. Macroscopic densities obtained from (26) are 5% higher compared to (35), which might confirm that the exponential modulation is endowed with a slightly higher numerical dissipation.

Conclusion and outlook

Two numerical schemes endowed with both 2D well-balanced (WB) and asymptotic-preserving (AP) properties, extending the one previously in [24], were studied in this paper. The first one involves a S-matrix directly built from the expression of exact steady-states for (1), namely truncated Fourier-Bessel series (15). The resulting scheme was proved to be 2D-WB; a novelty is that, in Corollary 2, it is also proved that its IMEX reformulation (26) is endowed with similar properties, too. A drawback of (19) is that it isn't unconditionally positivity-preserving. Accordingly, a simpler scheme (33), relying on an exponential modulation of the one

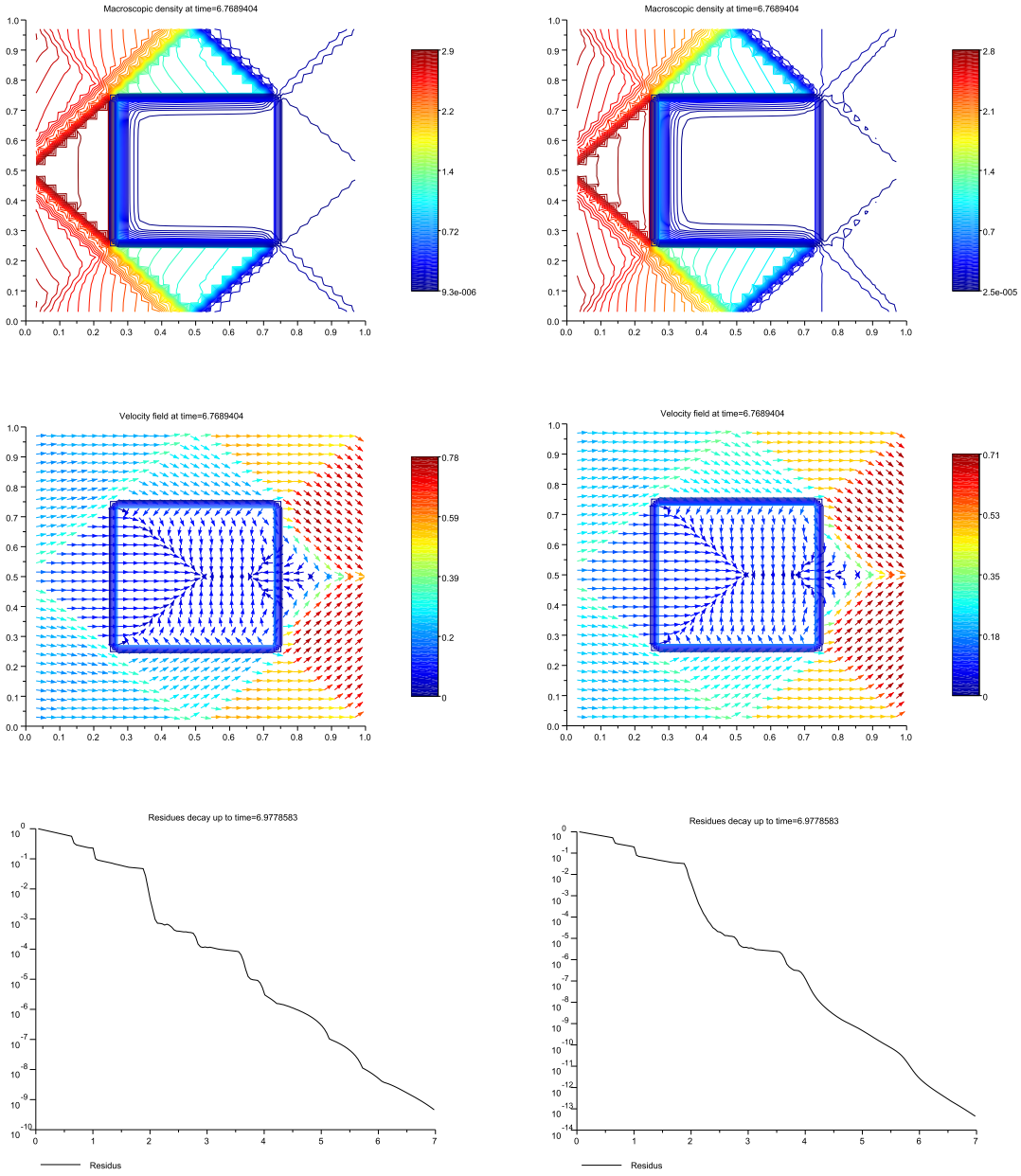

FIGURE 3. Bessel (left) and Composite (right) schemes: ρ, $\vec{V}$, and residues (top to bottom).

given in [24], was proved to be positivity-preserving, but endowed with weaker well-balancing features. Its IMEX reformulation (35) is again shown to be consistent with asymptotic damped diffusion behavior, see (37). Such diffusive limit is even rigorously established by means of convenient estimates.

The next stepping stone consists in applying the same program to two-dimensional kinetic models rendering biased velocity redistribution, like in semiconductor and chemotaxis dynamics modeling. Particular equations

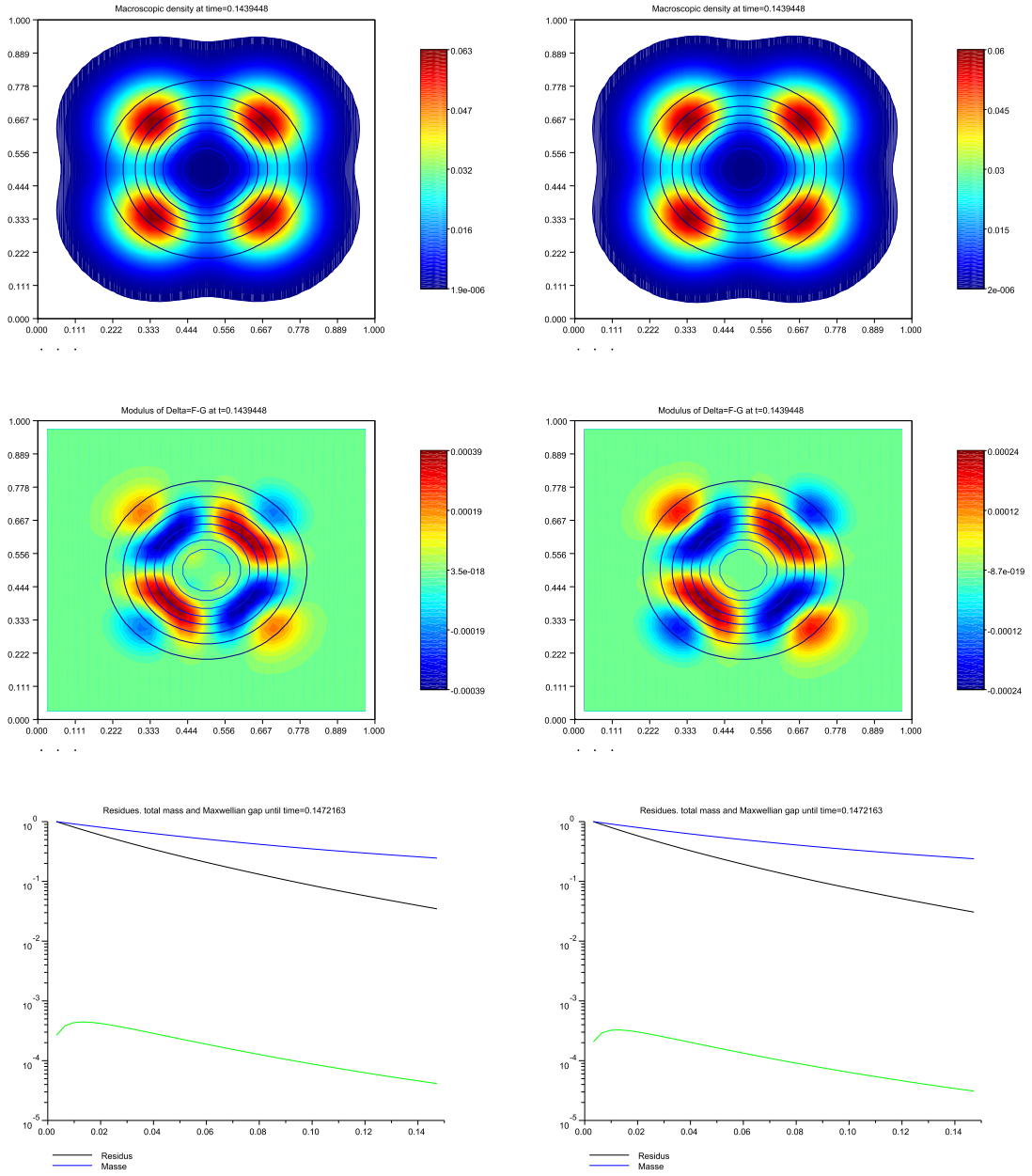

FIGURE 4. Bessel (left) and Composite (right) schemes: ρ, $\mathfrak{f} - \mathfrak{g}$, and residues (top to bottom).

belonging to this class of models are still simple enough to admit steady-states which can be again expressed as Fourier-Bessel series, hence permitting to derive 2D-WB numerical schemes. This would somehow extend the one-dimensional ideas of [22, 23] toward problems in higher dimensions.

Appendix A. Second order IMEX "Midpoint rule" scheme

Following [8,27], other IMEX time-integrators, possibly high-order can be substituted to (25). A well-known example is the second order in time "midpoint IMEX rule", which in our notation, rewrites as follows:

$$A^{n+\frac12}_{i-\frac12,j+\frac12} + \frac{\Delta t}{4\varepsilon R}\left(A^{n+\frac12}_{i-\frac12,j+\frac12} - S_0 B^{n+\frac12}_{i-\frac12,j+\frac12}\right) = A^{n}_{i-\frac12,j+\frac12} + \frac{\Delta t}{4R}\, S_1^\varepsilon\, B^{n}_{i-\frac12,j+\frac12}$$
$$A^{n+1}_{i-\frac12,j+\frac12} = A^{n}_{i-\frac12,j+\frac12} - \frac{\Delta t}{2\varepsilon R}\left(A^{n+\frac12}_{i-\frac12,j+\frac12} - (S_\kappa^\varepsilon)_{i-\frac12,j+\frac12}\, B^{n+\frac12}_{i-\frac12,j+\frac12}\right).$$

Corollary 4. *Under the assumptions of Corollary 2, and for Δt small enough to belong to the stability region of the "midpoint IMEX rule", any steady-state initial data is kept invariant,*

$$\forall (i,j)\in\mathbb{Z}^2,\qquad A^n_{i-\frac12,j+\frac12} = (S_\kappa^\varepsilon)_{i-\frac12,j+\frac12}\, B^n_{i-\frac12,j+\frac12} \quad\Longrightarrow\quad f^{\pm,n+1}_{i,j} = f^{\pm,n}_{i,j},\quad g^{\pm,n+1}_{i,j} = g^{\pm,n}_{i,j}.$$

Proof. Recall $(S_\kappa^\varepsilon)_{i-\frac12,j+\frac12} = S_0 + \varepsilon\, S_1^\varepsilon$: the "midpoint IMEX scheme" splits into two distinct steps:

- the first matches the IMEX scheme (25) where Δt is changed into $\Delta t/2$, hence if

$$A^n_{i-\frac12,j+\frac12} = (S_\kappa^\varepsilon)_{i-\frac12,j+\frac12}\, B^n_{i-\frac12,j+\frac12},$$

  then, by Corollary 2,

$$\forall (i,j)\in\mathbb{Z}^2,\qquad f^{\pm,n+\frac12}_{i,j} = f^{\pm,n}_{i,j},\qquad g^{\pm,n+\frac12}_{i,j} = g^{\pm,n}_{i,j}.$$

- the second is an explicit discretization, so that it corresponds to Theorem 1 because:

$$A^{n+\frac12}_{i-\frac12,j+\frac12} - (S_\kappa^\varepsilon)_{i-\frac12,j+\frac12}\, B^{n+\frac12}_{i-\frac12,j+\frac12} = A^{n}_{i-\frac12,j+\frac12} - (S_\kappa^\varepsilon)_{i-\frac12,j+\frac12}\, B^{n}_{i-\frac12,j+\frac12}.$$

Accordingly, the second order midpoint IMEX scheme is 2D well-balanced inside its domain of stability. □

In order to illustrate the behavior of this second-order IMEX integrator, the same example as §4.2 was set up, with $\varepsilon = 10^{-2}$ though. Results are displayed on Fig. 5: the mass is slightly bigger for the midpoint rule IMEX scheme, but the main differences show up in the Maxwellian gaps $\mathfrak{f}-\mathfrak{g}$ which have a quite different shape, despite both have roughly the same amplitude. The second-order IMEX scheme isn't stable when $\varepsilon \ll \Delta x$.

References

[1] R.E. Aamodt, K.M. Case, *Useful identities for half-space problems in linear transport theory*, Ann. Physics **21** (1963) 284–301.

[2] M. Abramovitz,. I. Stegun, **Handbook of Mathematical Functions**, National Bureau of Standards, Washington, D.C., 1972

[3] R. Bianchini, L. Gosse, *A truly two-dimensional discretization of drift-diffusion equations on Cartesian grids*, SIAM J. Numer. Anal. **56** (2018) 2845–2870.

[4] R. Bianchini, L. Gosse, E. Zuazua, *A two-dimensional "flea on the elephant" phenomenon and its numerical visualization*, SIAM Mult. Model. Simul. **17** (2019) 137–166.

[5] G. Birkhoff, I. Abu-Shumays, *Harmonic solutions of transport equations*, J. Math. Anal. Applic. **28** (1969) 211–221.

[6] G. Birkhoff, I. Abu-Shumays, *Exact analytic solutions of transport equations*, J. Math. Anal. Applic. **32** (1970) 468–481.

[7] G. Bretti, L. Gosse, N. Vauchelet, *$\mathcal{L}$-splines as diffusive limits of dissipative kinetic models*, submitted (2020).

[8] M. Briani, R. Natalini, G. Russo, *Implicit–explicit numerical schemes for jump–diffusion processes*, Calcolo **44** (2007) 33–57.

[9] Ch. Buet, B. Despres, G. Morel, *Trefftz Discontinuous Galerkin basis functions for a class of Friedrichs systems coming from linear transport*, 2018. ¡hal-01964528¿.

[10] Kenneth M. Case, *Elementary solutions of the transport equation and their applications*, Ann. Physics **9** (1960) 1–23.

[11] K.M. Case, P.F. Zweifel, **Linear transport theory**, Addison-Wesley series in nuclear engineering (1967).

[12] J.G. Conlon, *Fundamental solutions for the anisotropic neutron transport equation*, Proc. Royal Society of Edinburgh **82A** (1978) 325–350.

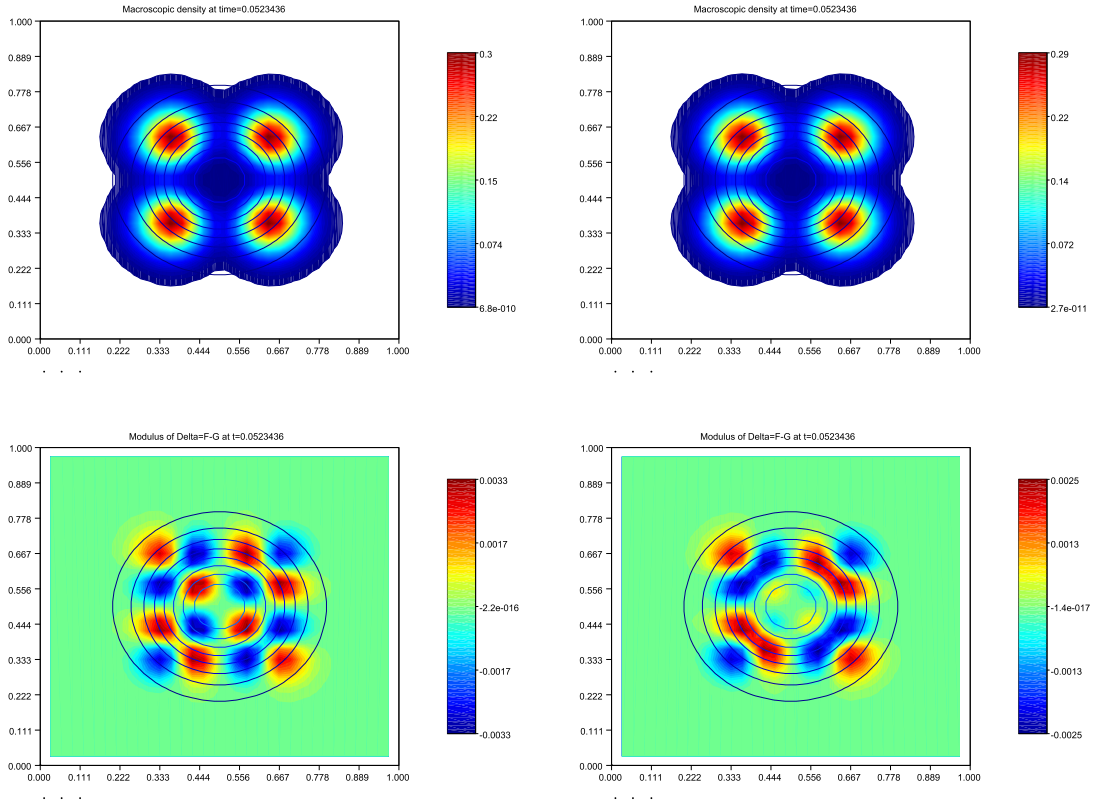

FIGURE 5. Second- (left) and first-order (right) IMEX schemes: ρ, $\mathfrak{f} - \mathfrak{g}$ (top to bottom).

[13] B. DESPRES, CH. BUET, *The structure of well-balanced schemes for Friedrichs systems with linear relaxation*, Applied Math. Comput. **272** (2016) 440–459.

[14] R. Estrada, *On Pizzetti's formula*, Asymptotic Analysis, **111** (2019) 1–14

[15] L. Flatto, *Functions with a mean value property*, J. Math. Mech. **10** (1961), 11–18.

[16] L. GOSSE, **Computing qualitatively correct approximations of balance laws**: Exponential-fit, well-balanced and asymptotic-preserving, SIMAI Springer Series **2** (2013).

[17] L. Gosse, *A well-balanced and asymptotic-preserving scheme for the one-dimensional linear Dirac equation*, BIT Numer Math. **55** (2015) 433–458.

[18] L. Gosse, *A well-balanced scheme able to cope with hydrodynamic limits for linear kinetic models*, Applied Math. Letters **42** (2015) 15–21.

[19] L. Gosse, *Viscous equations treated with $\mathcal{L}$-splines and Steklov-Poincaré operator in two dimensions*, in **Innovative Algorithms & Analysis**, DOI: 10.1007/978-3-319-49262-9_6.

[20] L. Gosse, *$\mathcal{L}$-splines and viscosity limits for well-balanced schemes acting on linear parabolic equations*, Acta Applic. Math. **153** (2018) 101–124.

[21] L. Gosse, G. Toscani, *An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations*, C.R. Math. Acad. Sci. Paris **334** (2002) 337–342.

[22] L. Gosse, N. Vauchelet, *Numerical high-field limits in two-stream kinetic models and 1D aggregation equations*, SIAM J. Sci. Comput. **38** (2016) A412–A434.

[23] L. Gosse, N. Vauchelet, *Some examples of kinetic schemes whose diffusion limit is Il'in's exponential-fitting*, Numer. Math. **141** (2019) 627–680

[24] L. Gosse, N. Vauchelet, *A truly two-dimensional, asymptotic-preserving scheme for a discrete model of radiative transfer*, SIAM J. Numer. Anal. (2020), DOI: 10.1137/19M1239829

[25] P.-L. Lions, G. Toscani, *Diffusive limit for finite velocity Boltzmann kinetic models*, Riv. Math. Iberoamericana **13** (1997) 473–513.
[26] Sawomir Michalik, *Summable solutions of some partial differential equations and generalised integral means*, J. Math.Anal.Appl. **444** (2016) 1242–1259
[27] L. Pareschi., G. Russo, *Implicit–explicit Runge-Kutta schemes and applications to hyperbolic system with relaxation*, J. Sci. Comput. **25** (2005) 129155.
[28] Xu Yang, F. Golse, Z. Huang, Shi Jin, *Numerical Study of a Domain Decomposition Method for a Two-Scale Linear Transport Equation*, Networks Heter. Media **1** (2006) 143–166.
[29] L. Zalcman, *Mean values and differential equations*, Israel J. Math. **14** (1973) 339–352.