Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2018-0071

Vol. **26**, No. 1, pp. 1-34
July 2019

Capturing Near-Equilibrium Solutions: A Comparison between High-Order Discontinuous Galerkin Methods and Well-Balanced Schemes

Maria Han Veiga[1,2,*], David A. Velasco-Romero[1,3,4], Rémi Abgrall[2] and Romain Teyssier[1]

[1] *Institute of Computational Science, University of Zurich, Switzerland.*
[2] *Institute of Mathematics, University of Zurich, Switzerland.*
[3] *Universidad Autónoma del Estado de Morelos, Mexico.*
[4] *Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Mexico.*

Received 19 March 2018; Accepted (in revised version) 28 August 2018

Abstract. Equilibrium or stationary solutions usually proceed through the exact balance between hyperbolic transport terms and source terms. Such equilibrium solutions are affected by truncation errors that prevent any classical numerical scheme from capturing the evolution of small amplitude waves of physical significance. In order to overcome this problem, we compare two commonly adopted strategies: going to very high order and reduce drastically the truncation errors on the equilibrium solution, or design a specific scheme that preserves by construction the equilibrium exactly, the so-called well-balanced approach. We present a modern numerical implementation of these two strategies and compare them in details, using hydrostatic but also dynamical equilibrium solutions of several simple test cases. Finally, we apply our methodology to the simulation of a protoplanetary disc in centrifugal equilibrium around its star and model its interaction with an embedded planet, illustrating in a realistic application the strength of both methods.

AMS subject classifications: 65M60, 65Z05

Key words: Numerical methods, benchmark, well-balanced methods, discontinuous Galerkin methods.

1 Introduction

Hyperbolic balance laws are used to describe many dynamical problems in natural sciences. They are defined as a set of conservation laws with associated source terms, which

*Corresponding author. *Email addresses:* mariaioque.hanveiga@math.uzh.ch (M. Han Veiga), david.velasco@icf.unam.mx (D. Velasco), remi.abgrall@math.uzh.ch (R. Abgrall), romain.teyssier@uzh.ch (R. Teyssier)

model the production or destruction of the corresponding conserved quantity. Many physical systems of scientific interest can be described by a system of hyperbolic conservation laws with source terms, or in short, hyperbolic balance laws.

Hyperbolic balance laws are particularly challenging because they feature equilibrium solutions that result from the exact cancellation of the divergence of the flux and the source terms in these equations. Small truncation errors can perturb this equilibrium solution, leading to the production of spurious waves that can dominate over the real waves that control the physics of the problem at hand.

For example, for the case of the inviscid Euler equations with a gravity source term (also known as the Euler-Poisson system), hydrostatic steady states are important in, for example, hydraulics [3, 5, 6] and astrophysics [20, 21, 29]. The difficulty here is to capture properly sound waves, gravity waves or convective flows, whose amplitude can be comparable to the truncation errors of a second order method and a reasonable grid resolution.

General steady states with non constant velocity fields are also found to be important in planetary sciences, namely in the early stages of protoplanetary discs, where the source term models the gravity of a central star [28], and is balanced by the centrifugal and pressure forces. The challenge here is to be able to resolve the interaction of a small planet with the gaseous disc, leading to the formation of a small amplitude spiral wave that can be dominated by the truncation errors of the equilibrium solution. In this context, the classical approach is to use a cylindrical mesh, reducing drastically discretisation errors along circular orbits. It is however desirable to find a solution on a Cartesian mesh, as it allows to deal with more general cases which are not strictly axisymmetric.

In summary, solving for such flows which are close to equilibrium can be very challenging for a naive, low order numerical method on a mesh not necessarily adapted to the geometry of the equilibrium solution as the truncation error incurred while solving the steady state can be larger than the small amplitude waves of interest.

There are nowadays many practical numerical methods with very low truncation errors. A class of such methods are the so-called discontinuous Galerkin (DG) methods [1]. These methods, at least for smooth and regular problems, can be made as accurate as desired. This means that, at least in principle, the amplitude of the truncation errors can be reduced to an arbitrarily small value. This requires an appropriate way to implement the source terms in the DG formalism [7, 18]. This also requires the use of a high enough resolution mesh to capture the equilibrium solution, which translates into higher computational cost for higher order solutions.

There is another strategy that allows one to use a low-order method, while capturing almost exactly the equilibrium solution. This is called the well-balanced approach (introduced in detail [17]), which is concerned with numerical schemes that satisfy the discrete equivalent of an underlying steady state, effectively, taking into account the existence of a steady state (or near steady state) solution.

The natural question is thus whether exact well-balancedness is required in practice or if methods that solve the PDE (including the source term) need only to be very accu-

rate. This is the question we wish to explore in this paper on several examples of interest for natural sciences in general and astrophysics in particular.

Let d, $e\in\mathbb{N}$, Ω an open subset of $\mathbb{R}^e$ and $\mathbf{f}_j$ for $1\leq j\leq d$ be smooth functions from Ω into $\mathbb{R}^e$. A general e-size system of d-dimensional hyperbolic balance laws can be written in the following form:

$$\frac{\partial \mathbf{w}}{\partial t}+\sum_{j=1}^{d}\frac{\partial}{\partial x_j}\mathbf{f}_j(\mathbf{w})-\mathbf{s}(\mathbf{w},\mathbf{x})=0,\quad \mathbf{x}=(x_1,\cdots,x_d)\in\mathbb{R}^d,\quad t>0, \tag{1.1}$$

where the vector valued function $\mathbf{w}=(w_1,\cdots,w_e):\mathbb{R}^d\times[0,\infty)\to\Omega$ denotes the solution, the functions $\mathbf{f}_j=[f_{1j},\cdots,f_{ej}]^T$ are flux-functions and $\mathbf{s}(\mathbf{w},\mathbf{x})$ is the vector of source terms. We denote vectors in bold $\mathbf{v}$ and a component of the vector as v, where the index is omitted if not important.

In order to solve hyperbolic balance laws, one can use classical methods for hyperbolic conservation laws (i.e. when $\mathbf{s}(\mathbf{w},\mathbf{x})=\mathbf{0}$) in conjunction with an operator-split approach to add the source terms. However, problems can arise when one tries to model flows near equilibrium states, for which (1.1) admits a steady state solution such that:

$$\sum_{j=1}^{d}\frac{\partial}{\partial x_j}\mathbf{f}_j(\mathbf{w})-\mathbf{s}(\mathbf{w},\mathbf{x})=\mathbf{0}. \tag{1.2}$$

In this work, we are mainly interested in solving the Euler-Poisson system with an analytical gravitational potential Φ with moving steady states (where the velocity field $\mathbf{v}\not\equiv\mathbf{0}$). We restrict ourselves to one- or two-dimensional cases given by $e=3$ and $d=1$, or $e=4$ and $d=2$ respectively. We however present the main equations in the 2D case only, as given in (1.3)

$$\frac{\partial \mathbf{w}}{\partial t}+\sum_{j=1}^{2}\frac{\partial}{\partial x_j}\mathbf{f}_j(\mathbf{w})-s(\mathbf{w},\mathbf{x})=0, \tag{1.3}$$

where

$$\mathbf{w}=\begin{bmatrix}\rho\\ \rho v_x\\ \rho v_y\\ E\end{bmatrix},\quad \mathbf{f}_1(\mathbf{w})=\begin{bmatrix}\rho v_x\\ \rho v_x^2+p\\ \rho v_x v_y\\ v_x(E+p)\end{bmatrix},\quad \mathbf{f}_2(\mathbf{w})=\begin{bmatrix}\rho v_y\\ \rho v_x v_y\\ \rho v_y^2+p\\ v_y(E+p)\end{bmatrix},\quad \mathbf{s}(\mathbf{w})=\begin{bmatrix}0\\ -\rho\frac{\partial}{\partial x}\Phi\\ -\rho\frac{\partial}{\partial y}\Phi\\ -\rho\mathbf{v}\cdot\nabla\Phi\end{bmatrix}.$$

Here ρ is the mass density, $\mathbf{v}=(v_x,v_y)$ the velocity and E the total energy given by the sum of internal and kinetic energy.

$$E=\rho\epsilon+\frac{1}{2}\rho|\mathbf{v}|^2.$$

In addition, p denotes the pressure and we assume $p=p(\rho,e)$ is a known function. Furthermore, we assume an ideal gas, such that the system is closed with the equation of state:

$$p=\rho\epsilon(\gamma-1),$$

where γ denotes the adiabatic index. The source terms, shown on the right hand side of the momentum and energy equations, model the effect of the gravitational forces on the fluid, for a given potential Φ.

To correctly solve these equations numerically and capture small perturbations to the steady state, dedicated computational methods are required to solve the discrete version of the source-flux balance (1.2). For non well-balanced methods, there is no guarantee that the truncation errors induced by discretising the steady state solution are not greater than the small perturbations we want to describe.

The design of *well-balanced schemes* (i.e. schemes which satisfy exactly a discrete equivalent of the underlying steady state) has been an active field of research, first coined in [14]. There have been many attempts to deal with this aspect, in particular for the shallow water equations, where steady states can represent the lake at rest case (hydrostatic equilibrium) [4] or a running river (non-trivial velocity equilibrium state) [25].

For the Euler-Poisson system there have been several recent contributions. We do not intend to give an exhaustive account of all the work that has be done in this topic, but we refer to: [16] where the authors design a well-balanced first and second order accurate finite volume scheme for approximating the Euler equations with gravitation using a discretisation of the hydrostatic equilibrium for the pressure reconstruction, [18, 19] where a similar approach to treat hydrostatic equilibria, isothermal and polytropic equations of state achieves a high order well-balanced discontinuous Galerkin scheme, and [8] where a relaxation scheme is adopted.

In [28] a second order finite volume method dealing with non zero velocities is presented in the context of protoplanetary discs. Concerning more general classes of steady states, the survey [22] describes two classes of schemes, one based on high-order accurate, non-oscillatory finite difference operators which are well-balanced for a general class of equilibria, and another one based on well-balanced quadratures, showing the suitability of these methods on the Shallow Water equations, and the work in [23], describing a high order path-conservative scheme and well balanced reconstruction, however, the analysis for this work is restricted to 1-dimension quasi-linear hyperbolic systems.

On the other hand, due to the tractability of modern, very-high-order methods, one could ask whether these methods alone could be enough to solve the equilibrium solutions to a high enough accuracy.

In this paper, we make a comparative study between a new method which is truly well balanced and a popular class of very-high-order methods, namely the discontinuous Galerkin method. We would like to answer the following fundamental questions, considering both hydrostatic and moving equilibria solutions:

1. Are there cases where using a high order scheme is sufficient to capture solutions close to a steady state?
2. Under which circumstances is it necessary to use a well balanced method?
3. What is the cost associated to each approach and how does it balance with accuracy?

In particular, we compare a well-balanced, high-order discontinuous Galerkin method with a non well-balanced, high-order discontinuous Galerkin method under different steady state regimes, both hydrostatic and stationary (with a non-zero velocity).

The outline of this paper is as follows: a brief introduction on equilibrium solutions, as well as the description of the Runge-Kutta discontinuous Galerkin (RKDG) method is provided in Section 2. In Section 3, we describe our well balanced formulation of RKDG. In Section 4, a set of benchmark problems are defined, both in one and two dimensions, and quantitative results are presented, followed by our final discussion in Section 5.

2 Preliminaries

2.1 Steady state solutions

A solution $\mathbf{w}$ is said to be a steady state solution of (1.1) if it fulfils the following relation

$$\sum_{j=1}^{d} \frac{\partial}{\partial x_j} \mathbf{f}_j(\mathbf{w}) - \mathbf{s}(\mathbf{w},\mathbf{x}) = \mathbf{0}, \tag{2.1}$$

for $w:\mathbb{R}^d \times [0,\infty) \to \Omega$, $\mathbf{f}_j:\Omega \to \mathbb{R}^e$ and $s:\Omega \times \mathbb{R}^d \to \mathbb{R}^e$, $\Omega \subset \mathbb{R}^e$.

We call $\mathbf{w}$ a hydrostatic steady state of (1.3) if the pressure component fulfils the following relation

$$\nabla p = -\rho \nabla \Phi, \tag{2.2}$$

for $\rho, p:\mathbb{R}^d \times [0,\infty) \to \Omega$, and $\Phi:\Omega \times \mathbb{R}^d \to \mathbb{R}^e$, $\Omega \subset \mathbb{R}^e$ and a gravity potential $\Phi \in \mathcal{C}^1$.

Let the following be the standardised form of an time explicit numerical scheme for (1.3), where $H(\cdot)$ denotes the update function for each timestep n in a quantity indexed by k (e.g. cell average) and q, p denote the stencil size:

$$w_k^{n+1} = w_k^n + \frac{\Delta t}{\Delta x} H(w_{k-q}^n, \cdots, w_{k+p}^n). \tag{2.3}$$

The numerical scheme is exactly well-balanced if for a steady state solution w, the following holds

$$H(w_{k-q}^n, \cdots, w_{k+p}^n) = 0. \tag{2.4}$$

The scheme is said to be well-balanced with order N_p if, for a steady state solution w, the following holds

$$|H(w_{k-q}^n, \cdots, w_{k+p}^n)| = \mathcal{O}(\Delta x^{N_p+1}). \tag{2.5}$$

A formal definition of the above is given in [23].

Remark 2.1. In the astrophysics literature, in particular in the planet formation community [2], the following equilibrium relation between the centrifugal force associated to the cross-radial velocity v_θ, the gradients of the thermal pressure p and the gradient of the gravitational potential Φ is often referred to as a *dynamical equilibrium*:

$$\frac{v_\theta^2}{r} = \frac{1}{\rho}\nabla p + \nabla\Phi. \tag{2.6}$$

It is often the case that the pressure gradient ∇p is assumed to be small and thus can be neglected [2]. For the purpose of this work, we considered initial conditions which are strictly hydrostatic, and also more general *stationary solutions* or *steady states* of the Euler-Poisson equation.

2.2 Runge Kutta Discontinuous Galerkin (RKDG) method

Consider a regular domain $D \in \mathbb{R}$, approximated by K non-overlapping elements such that $\bigcup_{K\in D_h} K \approx D$. The 2-dimensional tessellation is given by the tensor product of the 1-dimensional discretizations, thus yielding square volumes (or cubic volumes in 3D). Let T_h denote the Cartesian tessellation of the domain D where our problem is defined.

We seek for the approximate solution $w_h(t)$ in the finite element space of discontinuous functions V_h:

$$V_h = \{v_h \in \mathcal{L}^\infty(D) : v_h|K \in V_h(K),\ \forall K \in T_h\}.$$

We take $V_h(K)$ to be the collection of polynomials of at most degree N_p.

Following the Runge Kutta discontinuous Galerkin (RKDG) method described in [9], we write the weak formulation for each component of (1.1) by multiplying the system by a smooth test function $v(x)$ and integrate over a control volume K:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_K w(\mathbf{x},t)v(\mathbf{x})\mathrm{d}\mathbf{x} + \sum_{e\in\partial K}\int_e f(w(\mathbf{x},t))\cdot n_{e,K}v(\mathbf{x})\mathrm{d}\Gamma - \int_K f(w(\mathbf{x},t))\cdot\nabla v(\mathbf{x})\mathrm{d}\mathbf{x} \\ = \int_K s(w(\mathbf{x},t))v(\mathbf{x})\mathrm{d}\mathbf{x}\end{aligned} \tag{2.7}$$

for any smooth $v(\mathbf{x})$. We denote the outward unit normal as $n_{e,K}$ and edge as e.

The following integrals are approximated with a suitable order numerical quadrature (where $\{\mathbf{x}_i,\omega_i\}_{i=0}^{M,L}$ denotes the set of quadrature points and weights):

$$\int_e f(w(\mathbf{x},t))\cdot n_{e,K}v(\mathbf{x})\mathrm{d}\Gamma \approx \sum_{i=0}^{L} f(w(\mathbf{x}_i,t))\cdot n_{e,K}v(\mathbf{x}_i)\omega_i|e|, \tag{2.8}$$

$$\int_K f(w(\mathbf{x},t))\cdot\nabla v(\mathbf{x})\mathrm{d}\mathbf{x}\approx\sum_{j=0}^{M} f(w(\mathbf{x}_j,t))\cdot\nabla v(\mathbf{x}_j)\omega_j|K|, \tag{2.9}$$

$$\int_K s(w(\mathbf{x},t))v(\mathbf{x})\mathrm{d}\mathbf{x}\approx\sum_{j=0}^{M} s(w(\mathbf{x}_j,t))v(\mathbf{x}_j)\omega_j|K|. \tag{2.10}$$

Note that there is an ambiguity in the definition of the flux $f(w(\mathbf{x},t))\cdot n_{e,K}$ since w can be multi-valued at the cell interface. In order to overcome this inconsistency, this term is replaced by a single-valued numerical flux $h_{e,K}(\mathbf{x},t)$ computed using a Riemann solver. The exact solution is replaced by the finite dimensional approximate solution $w_h=\sum_{i=0}\hat{w}_i(t)\psi_i(\mathbf{x})$, where $\hat{w}_i(t)$ is given by the L^2 inner product between $w(\mathbf{x},t)$ and ψ_i, a basis element of $V_h(K)$, and the test functions $v(\mathbf{x})$ are replaced by $v_h(\mathbf{x})\in V_h(K)$. This yields the following numerical scheme:

$$\begin{aligned}
&w_h(t=0)=P_{V_h}(w_0),\\
&\frac{\mathrm{d}}{\mathrm{d}t}\int_K w_h(\mathbf{x},t)v_h(x)\mathrm{d}\mathbf{x}=-\sum_{e\in\partial K}\sum_{i=1}^{L}h_{e,K}(\mathbf{x}_i,t)v_h(\mathbf{x}_i)\omega_i|e|+\sum_{j=1}^{M}f(w_h(\mathbf{x}_j,t))\cdot\nabla v_h(\mathbf{x}_j)\omega_j|K|\\
&\qquad\qquad+\sum_{j=1}^{M}S(w_h(\mathbf{x}_j,t))v_h(\mathbf{x}_j)\omega_j|K|,\quad\forall v_h(\mathbf{x})\in V(K),\ \forall K\in T_h,
\end{aligned}$$

where operator P_{V_h} denotes the L^2 projection of the initial data $w_0(x)$ into the space of finite elements V_h.

In addition, throughout this work, we make the following choices:

1. We denote by $\{\psi\}_{i=0}^{N_p}$ the Legendre basis vectors spanning $V_h(K)$, subject to the following normalisation:
$$\int_{-1}^{1}\psi_i(x)\psi_j(x)dx=\delta_{ij};$$
2. We take the numerical quadrature points and weights $\{x_i,\omega_i\}_{i=0}^{M}$ to be Gauss-Legendre quadrature points;
3. We use local Lax-Friedrichs flux as the numerical flux. We note that our analysis works for any consistent numerical flux function that is Lipschitz continuous in both arguments, non-decreasing in its first argument and non-increasing in its second argument.

2.2.1 Time discretisation

We use the TVD Runge-Kutta time discretization as in [13]. Let $\{t^n\}_{n=0}^{N}$ be a partition of $[0,T]$ and $\Delta t^n=t^{n+1}-t^n$, $n=0,\cdots,N-1$, then the time marching algorithm is given in Algorithm 1.

The parameters $a_{i,j}$, b_i and c_i can be found in Tables 2, see [13].

Data: $w_h^0 = P_{V_h}(w_0)$
Result: w_h^{n+1}
for *n = 0, ... N-1* **do**
  $w_h^{(0)} = w_h^n$
  for *i = 1, ... k+1* **do**
    $k_i = \mathcal{L}(t^n + c_i, w_h^{(0)} + h\sum_{j=1}^{i-1} a_{i,j}k_j)$
  end
  $w_h^{n+1} = w_h^{(0)} + h\sum_{i=1}^{s} b_i k_i$
end

Algorithm 1: TVD RK time marching algorithm.

Given that an explicit time integrator is used, the timestep Δt has to fulfill a Courant-Friedrich-Lewy (CFL) condition to achieve numerical stability. In this work, the timestep Δt^K at cell K is calculated as [26]. Furthermore, the introduction of a source term can introduce additional constraints on the timestep. As described in [31], the timestep for a solution approximation of degree at most N_p, we choose the minimum of the expression below:

$$\Delta t^K = \min\left(\frac{C}{2N_p+1}\left(\sum_{i=1}^{d}\frac{|v_i^K|+c_s^K}{\Delta x_i^K}\right)^{-1}, \frac{1}{\sqrt{2\gamma(\gamma-1)}}\frac{c_s^K}{|\nabla\Phi^K|}\right),$$

where $c_s = \sqrt{\gamma p/\rho}$ is the sound speed, v_i^K is the i^{th} component of the velocity average at cell K, Δx_i^K is the mesh-width in the i^{th} dimension and $|\nabla\Phi|^K$ the magnitude of the gradient of Φ at cell K. The constant C is chosen to be small, for example, **0.2**.

Table 1: Runge-Kutta Butcher tableaus for the TVDRK schemes.

0		
1/2	1/2	
	1/2	1/2

SSP(2,2)

0			
1	1		
3/4	1/4	1/4	
	1/6	1/6	1/3

SSP(3,3)

2.3 Discretisation error of DG for steady Euler system

In this short section we show that the traditional Runge Kutta discontinuous Galerkin method is inherently not well balanced and specify the source of approximation error for each conserved variable.

Table 2: Runge-Kutta Butcher tableaus for the TVDRK schemes.

0					
0.39175222700392	0.39175222700392				
0.58607968896779	0.21766909633821	0.36841059262959			
0.47454236302687	0.08269208670950	0.13995850206999	0.25189177424738		
0.93501063100924	0.06796628370320	0.11503469844438	0.20703489864929	0.54497475021237	
	0.14681187618661	0.24848290924556	0.10425883036650	0.27443890091960	0.22600748319395

SSP(4,5)

We consider the 1-dimensional Euler system (2.11) and the DG discretisation described above

$$\frac{\partial}{\partial t}\rho+\frac{\partial}{\partial x}\rho v=0, \tag{2.11a}$$

$$\frac{\partial}{\partial t}\rho v+\frac{\partial}{\partial x}(\rho v^2+p)=-\rho\frac{\partial}{\partial x}\Phi, \tag{2.11b}$$

$$\frac{\partial}{\partial t}E+\frac{\partial}{\partial x}(v(E+p))=-\rho v\frac{\partial}{\partial x}\Phi. \tag{2.11c}$$

Considering, for example, (2.11a) in some control volume K mapped to $[-1,1]$ interval, and modal coefficient i:

$$\partial_t\tilde{\rho}_i=\int_K \rho v\partial_x\psi_i \mathrm{d}x-\oint_{\partial K}\rho v\psi_i\cdot n\mathrm{d}\Gamma. \tag{2.12}$$

If the solution of the system is a steady state, then $\partial_t\tilde{\rho}_i=0 \quad \forall i$.

One can write the update H_ρ vector (as in (2.3)), where each component corresponds to i^{th} the modal update, with the associated test function ψ_i:

$$H_\rho^i=\sum_j \rho v\partial_x\psi_i(x_j)w_j-\hat{\rho v}\psi_i(1)+\hat{\rho v}\psi_i(-1).$$

To evaluate $\hat{\cdot}$ we need an approximate flux function to combine the left and right hand-side values of the flux. We consider the Lax-Friedrichs flux, for example, defined as:

$$\hat{f}(a,b)=\frac{f(a)+f(b)}{2}-\frac{\alpha(a-b)}{2},$$

where $\alpha=\max(v+c_s)$ and $c_s=\sqrt{\gamma\frac{p}{\rho}}$.

Assuming a simple steady state class of solutions, consider the non-moving equilibria, where $v\equiv 0$, (2.11a) has the following (component wise) update:

$$H_\rho^i=\frac{\alpha}{2}[[\rho(x_{k+1/2})]]\psi_i(1)-\frac{\alpha}{2}[[\rho(x_{k-1/2})]]\psi_i(-1),$$

where $[[f(x)]]=f(x^+)-f(x^-)$ denotes the jump in f between left and right states at x, and $\langle f(x)\rangle=\frac{f(x^+)+f(x^-)}{2}$ the average of f at x. This shows that the error comes from the jumps in the variable ρ at the interfaces of the control volume K.

Similarly, the update function $H_{\rho v}$ for (2.11b) is:

$$H^i_{\rho v} = -\langle p(x_{k+1/2})\rangle \psi_i(1) + \langle p(x_{k-1/2})\rangle \psi_i(-1) - \sum_j \rho \partial_x \Phi \psi_i(x_j) w_j + \sum_j p \partial_x \psi(x_j) w_j.$$

And for (2.11c):

$$H^i_E = \frac{\alpha}{2}[[E(x_{k+1/2})]]\psi_i(1) - \frac{\alpha}{2}[[E(x_{k-1/2})]]\psi_i(-1).$$

For the density and energy evolution, the error comes from the jump on the respective variable at the interfaces. For the momentum equation, the error will arise from the split treatment when discretising $\nabla \cdot f(w)$ and $s(w)$, which should exactly cancel out if w is a steady state solution.

Now we consider a general class of steady state solutions, where $v \not\equiv 0$. From (2.11a), follows that $\rho v = \text{const}$.

Then, for (2.11a) one can write the following update function:

$$H^i_\rho = \sum_j \rho v \partial_x \psi_i(x_j) w_j - \frac{\alpha}{2}[[\rho(x_{k+1/2})]]\psi_i(1) + \frac{\alpha}{2}[[\rho(x_{k-1/2})]]\psi_i(-1) - \langle \rho v\rangle \psi_i(1) + \langle \rho v\rangle \psi_i(-1).$$

Rewriting H_ρ, using the fact that $\rho v = \text{const}$ and that Legendre polynomials have the property $\psi_n(-x) = (-1)^n \psi_n(x)$, one can arrive at:

$$H^i_\rho = \frac{\alpha}{2}[[\rho(x_{k+1/2})]]\psi_i(1) - \frac{\alpha}{2}[[\rho(x_{k-1/2})]]\psi_i(-1).$$

Independently of the order of the polynomial ψ_n, the volume integral part cancels out either due to the numerical flux contribution or due to the fact ρv is constant.

For (2.11b):

$$\begin{aligned} H^i_{\rho v} = &-\langle \rho v^2 + p(x_{k+1/2})\rangle \psi_i(1) + \langle \rho v^2 + p(x_{k-1/2})\rangle \psi_i(-1) \\ &- \sum_j \rho \partial_x \phi \psi_i(x_j) w_j + \sum_j (\rho v^2 + p) \partial_x \psi(x_j) w_j. \end{aligned}$$

And for (2.11c):

$$\begin{aligned} H^i_E = &\frac{\alpha}{2}[[E(x_{k+1/2})]]\psi_i(1) - \frac{\alpha}{2}[[E(x_{k-1/2})]]\psi_i(-1) \\ &+ \langle v(E+p)(x_{k+1/2})\rangle \psi_i(1) - \langle v(E+p)(x_{k-1/2})\rangle \psi_i(-1) \\ &- \rho v \sum_j \partial_x \phi \psi_i(x_j) w_j + \sum_j (v(E+p)) \partial_x \psi(x_j) w_j. \end{aligned}$$

While update H_ρ remains unchanged, $H_{\rho v}$ and H_E have additional terms from the velocity contribution, and thus one can observe that the error arises from splitting the flux term in surface and volume terms, and the separated treatment of the source term and $\nabla \cdot f(w)$.

3 Well-balanced RKDG method

We now present our implementation of a well balanced method for RKDG. Using the formulation presented in (2.2), we follow an approach similar to [11], where we represent the solution of (1.1) as a sum of a steady state (or equilibrium) solution $w_{eq}(\mathbf{x})$ and a perturbation $\delta w(\mathbf{x},t)$:

$$w(\mathbf{x},t) = w_{eq}(\mathbf{x}) + \delta w(\mathbf{x},t) \quad a.e.$$

We note that if (1.1) admits a steady state solution w_{eq}, the flux-source balance relation holds:

$$\nabla \cdot f(w_{eq}(\mathbf{x})) = s(w_{eq}(\mathbf{x})). \tag{3.1}$$

And weakly, for a suitable test function $v(\mathbf{x})$:

$$\int \nabla \cdot f(w_{eq}(\mathbf{x}))v(\mathbf{x})d\mathbf{x} = \int s(w_{eq}(\mathbf{x}))v(\mathbf{x})d\mathbf{x}. \tag{3.2}$$

Subtracting (3.2) from (2.7), and noting that a state state solution satisfies $\frac{\partial}{\partial t} w_{eq} = 0$, we can write:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_K (\delta w(\mathbf{x},t))v(\mathbf{x})\mathrm{d}\mathbf{x} = &-\sum_{e\in\partial K}\int_e \delta f(w(\mathbf{x},t))\cdot n_{e,K}v(\mathbf{x})\mathrm{d}\Gamma \\ &+\int_K \delta f(w(\mathbf{x},t))\cdot\nabla v(\mathbf{x})\mathrm{d}\mathbf{x} \\ &+\int_K \delta s(w(\mathbf{x},t))v(\mathbf{x})\mathrm{d}\mathbf{x},\end{aligned}$$

where we use the following notation:

1. $\int_e \delta f(w(\mathbf{x},t))\cdot n_{e,K}v(\mathbf{x})\mathrm{d}\Gamma = \int_e \big(f(w(\mathbf{x},t)) - f(w_{eq}(\mathbf{x}))\big)\cdot n_{e,K}v(\mathbf{x})\mathrm{d}\Gamma$;
2. $\int_K \delta f(w(\mathbf{x},t))\cdot\nabla v(\mathbf{x})\mathrm{d}\mathbf{x} = \int_K (f(w(\mathbf{x},t)) - f(w_{eq}(\mathbf{x})))\cdot\nabla v(\mathbf{x})\mathrm{d}\mathbf{x}$;
3. $\int_K \delta S(w(\mathbf{x},t))v(\mathbf{x})\mathrm{d}\mathbf{x} = \int_K (s(w(\mathbf{x},t)) - s(w_{eq}(\mathbf{x})))v(\mathbf{x})\mathrm{d}\mathbf{x}$.

Note again that there is an ambiguity in the definition of the flux $f(w(\mathbf{x},t))\cdot n_{e,K}$ since w can be multi-valued at the cell interface. To overcome this inconsistency, the ambiguous term is here again replaced by a single-valued numerical flux $h_{e,K}(\mathbf{x},t)$ computed using the Lax Friedrich Riemann solver.

Let our numerical solution be represented as:

$$w_{num}(\mathbf{x},t) = w_{eq}(\mathbf{x}) + \delta w_h(\mathbf{x},t),$$

where $\delta w_h \in V_h(K)$. Furthermore, we approximate the integrals with a quadrature, which yields the following well balanced DG numerical scheme:

$$\begin{aligned}
&\delta w_h(t=0) = P_{V_h}(\delta w_0),\\
&\frac{\mathrm{d}}{\mathrm{d}t}\int_K \delta w_h(\mathbf{x},t)v_h(\mathbf{x})\mathrm{d}\mathbf{x} = -\sum_{e\in\partial K}\sum_{i=0}^{L}\delta f_{e,K}(w_{num}(\mathbf{x}_i,t))v_h(\mathbf{x}_i)\omega_i|e|\\
&\qquad\qquad +\sum_{j=0}^{M}\delta f(w_{num}(\mathbf{x}_j,t))\cdot\nabla v_h(\mathbf{x}_j)\omega_j|K|\\
&\qquad\qquad +\sum_{j=0}^{M}\delta s(w_{num}(\mathbf{x}_j,t))v_h(\mathbf{x}_j)\omega_j|K|, \quad \forall v_h(\mathbf{x})\in V(K),\ \forall K\in T_h.
\end{aligned}$$

Note that this reformulation is only suitable for problems where the solution w is close *enough* to the prescribed steady state solution w_{eq}. In fact, if the initial condition is exactly equal to the steady state solution ($w_0 = w_{eq}$), the scheme will capture the equilibrium solution exactly. If the initial condition is close to the steady state solution, this scheme is able to evolve the perturbation without being dominated by the truncation error on the steady state solution. However, if the initial condition is very far from the adopted steady state, this scheme might not be suitable and the traditional RKDG scheme will be more robust. If this is the case, setting the steady state w_{eq} to 0 and the perturbation δw to the full solution, one simply recovers the traditional RKDG scheme [9].

4 Numerical experiments

In this section, several benchmark problems will be introduced. These will be the basis of our discussion in Section 5. An introduction and description of code used to perform the numerical experiments can be found in [30]. Additional results can be found in Appendix C.

4.1 Error estimate

The empirical error estimates are calculated using the $\mathcal{L}_1$-error norm:

$$||w_h(\mathbf{x}) - w(\mathbf{x})||_1 = \int_D |w_h(\mathbf{x}) - w(\mathbf{x})| d\mathbf{x}.$$

It is shown in [27] that a convergence rate of N_p+1 for a N_p degree polynomial approximation of the solution in $\mathcal{L}_1$-error norm is expected for smooth enough functions. This quantity is computed with a numerical quadrature and computed the following manner:

$$||w_h(\mathbf{x}) - w(\mathbf{x})||_1 \approx \sum_{K\in D}\sum_{i=0}^{M}\sum_{j=0}^{M} |\, w_h(v^K(x_i,y_j)) - w(v^K(x_i,y_j))\,|\,\omega_i\omega_j\frac{\Delta x\Delta y}{4}, \tag{4.1}$$

where $\{x_i,y_j\}_{i,j=0}^M$ are Gauss Legendre quadrature points, $\{\omega_i,\omega_j\}_{i,j=0}^M$ the corresponding weights and $\nu^K(\chi,\upsilon)$ a linear transformation mapping element K to the canonical element $[-1,1]\times[-1,1]$,

$$\nu^K(\chi,\upsilon)=\Big(x_l-\chi\frac{\Delta x}{2},\, y_l-\upsilon\frac{\Delta y}{2}\Big),$$

and (x_l,y_l) the center of element K.

4.2 Well-balanced property

In this section, the well-balanced property of the schemes is evaluated. To this end, we first evolve a hydrostatic equilibrium solution. What should be observed, in this case, is that the solution does not change for any time $T>0$. However, due to the failure of perfectly balancing the discrete version of $\nabla\cdot f(w)$ and $s(w)$, the state at some time T might deviate from the initial condition. We then solve for the propagation of perturbations of the equilibrium solution, that we call here *waves*, using various amplitudes, and measure whether the schemes can capture these perturbations without being affected by the truncation errors of the equilibrium solution. In the last set of test cases, we evaluate the quality of our schemes using a dynamical equilibrium state, meaning that the velocity $\mathbf{v}=(v_x,v_y)$ is non-zero for the steady state solution.

4.2.1 Hydrostatic equilibrium

1-dimensional case. Considering an ideal gas γ=1.4 and a linear gravitational potential $\Phi_x=gx$, we are interested in preserving the following isothermal equilibrium state:

$$\begin{aligned}
&\rho_{eq}(x)=\rho_0\exp\Big(-\frac{\rho_0 g}{p_0}x\Big),\\
&u_{eq}(x)=0,\\
&p_{eq}(x)=p_0\exp\Big(-\frac{\rho_0 g}{p_0}x\Big),
\end{aligned}\tag{4.2}$$

with $\rho_0=1.0$, $p_0=1.0$ and $g=1.0$.

Because we are interested in preserving the equilibrium state, we impose the boundary condition as the extension of the domain in ∂D, as follows:

$$\rho(\mathbf{x})|_{\mathbf{x}\in\partial D}=\rho_{eq}(\mathbf{x}),\quad v_x(\mathbf{x})|_{\mathbf{x}\in\partial D}=v_{x,eq}(\mathbf{x}),\quad v_y(\mathbf{x})|_{\mathbf{x}\in\partial D}=v_{x,eq}(\mathbf{x}),\quad p(\mathbf{x})|_{\mathbf{x}\in\partial D}=p_{eq}(\mathbf{x}),\tag{4.3}$$

where $\mathbf{x}=(x)$ in 1-dimension and $\mathbf{x}=(x,y)$ in 2-dimensions.

The numerical errors for the density are shown in Fig. 1 for the following resolutions $N=8$, 16, 32, 64 at time $T=10.0$ for the second and third order well-balanced scheme (WBDG2 and WBDG3, respectively) and the traditional discontinuous Galerkin method with orders 2, 3 and 4 (DG2, DG3, DG4 respectively). One can observe that by increasing the resolution or the order, the truncation error can be reduced, even for long time

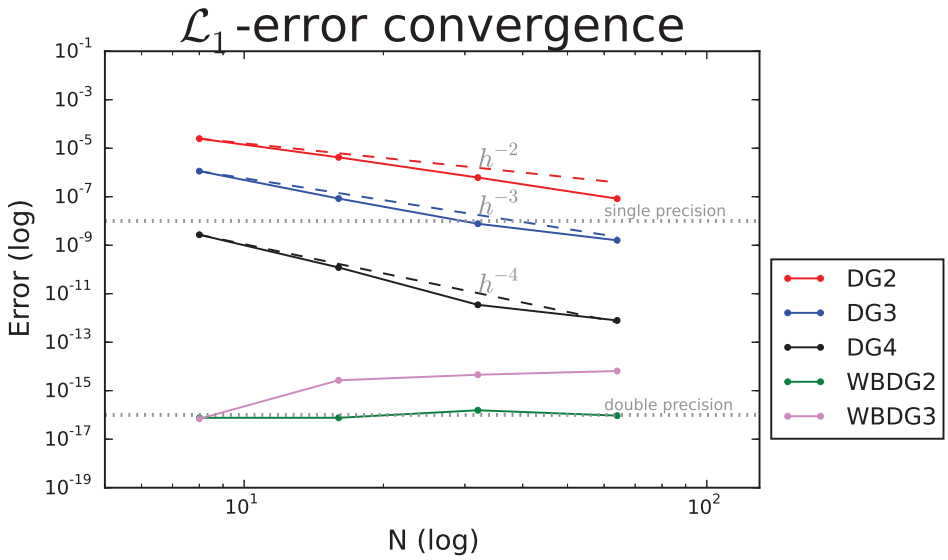

Figure 1: $\mathcal{L}_1$ error convergence for the 1-dimensional hydrostatic test case (4.2).

evolution. Furthermore, at $N=64$ at 4^{th} order, we reach a similar absolute error as in our well-balanced methods.

It is important to stress that the well-balanced scheme requires either the storage of additional arrays or requires to perform many additional computation every time step. Indeed, at each Runge-Kutta timestep, we can either recompute face nodal values or we can store the equilibrium solution once and for all, requiring $\mathcal{O}((4+m)N_xN_ym)$ of additional memory, where N_x and N_y denotes the number of cells in $x-$,$y-$ direction, respectively, and m the order of the method[†]. Shown in Fig. 2, we show the total time it takes to run the 1-dimensional hydrostatic equilibrium test case (4.2) when performing the well-balanced reconstruction (denoted as WBDG2(Rec)) versus precomputing and storing the equilibrium variables (denoted as WBDG2(Mem)), compared to the traditional discontinuous Galerkin methods with order 2, 3 and 4 (denoted as DG2, DG3, DG4 respectively).

A perturbation is now added to the pressure state of the equilibrium solution described in (4.2), as shown below:

$$p(x,t=0)=p_{eq}(x)+\eta\exp\left(-\frac{\rho_0 g}{p_0}\frac{(x-0.5)^2}{0.01}\right). \tag{4.4}$$

The initial condition (4.4) is run until $T=0.25$ with different pulse amplitudes: $\eta=1\times10^{-2}$, 1×10^{-4}, 1×10^{-6} and 1×10^{-8}. In Fig. 3 we show the pointwise L_1 error of between the solution attained with different orders of the non-well balanced discontinuous Galerkin scheme and a high resolution solution which captures the pulse. We note

[†]Further optimization is possible, by storing the resulting volume/surface integral for each cell, further reducing the necessary storage to $\mathcal{O}(N_xN_y)$.

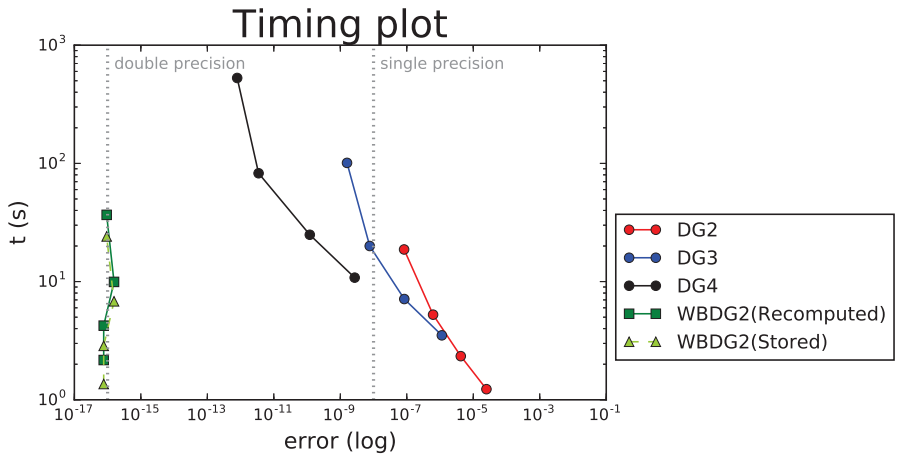

Figure 2: Total time to achieve a particular accuracy for the 1-dimensional hydrostatic test case (4.2).

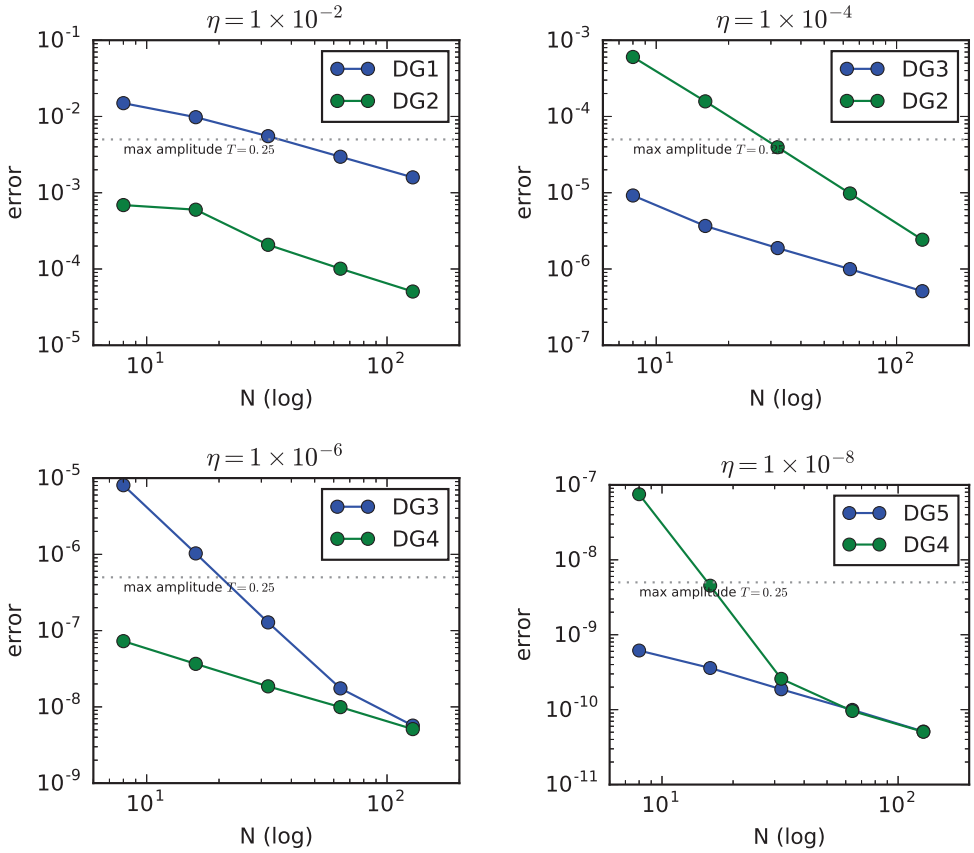

Figure 3: Convergence of the non well-balanced methods for initial conditions (4.4) with perturbation sizes of $\eta = 1 \times 10^{-2}$, 1×10^{-4}, 1×10^{-6} and 1×10^{-8} respectively.

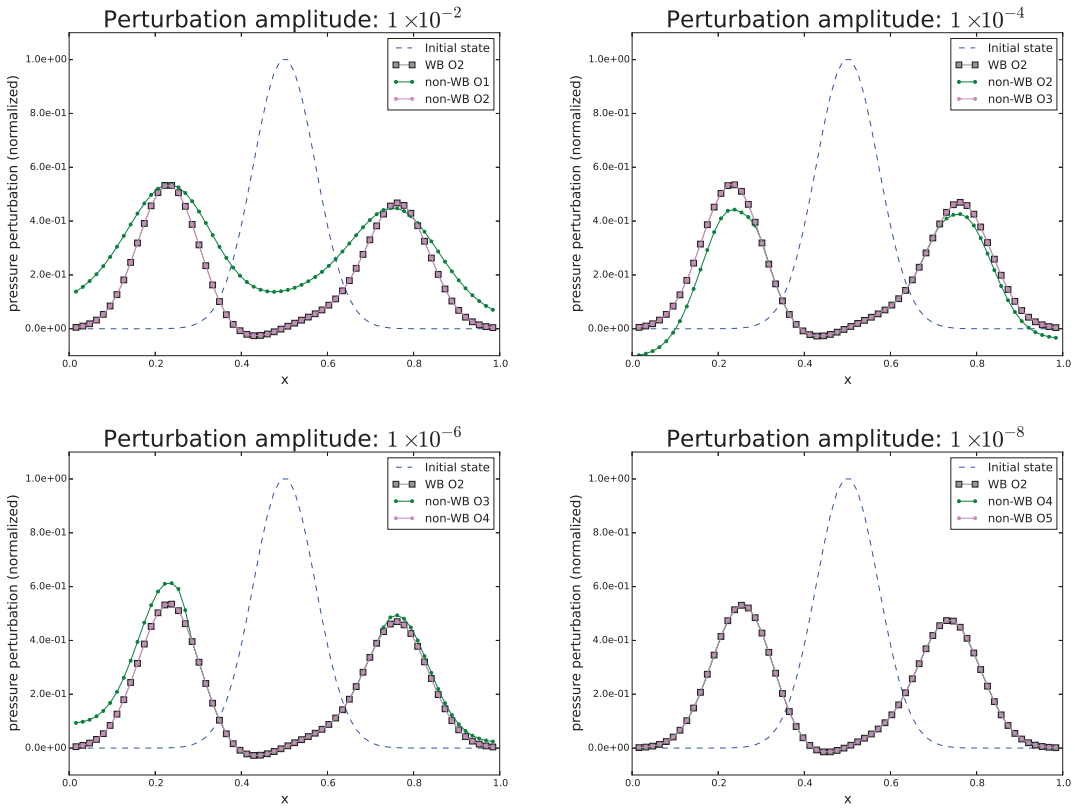

Figure 4: Non well-balanced method versus well-balanced method for hydrostatic equilibrium with varying amplitude perturbation on the pressure field for initial conditions (4.4).

that if the error is larger than the perturbation size, then it is clear that a particular combination of order and resolution is not enough to capture the perturbation. A qualitative depiction of this is shown in Fig. 4, for a fixed grid-size of $N=64$ and different orders. Furthermore, we note that for $\eta=1\times10^{-2}$, the difference between a second-order well balanced and a second-order non-well-balanced scheme is impossible to see. However, when the perturbation's amplitude η decreases below the truncation error of the scheme, the wave is no longer well captured. As shown in Fig. 3, the error for a non-well-balanced scheme can be reduced by increasing the order of the scheme or the resolution of the grid, effectively reducing the approximation error. Note that for the well-balanced method, we always capture the correct wave solution. The time to solution for the experiment with perturbation size 1×10^{-8} is shown in Table 3. Additional times to solution can be found in Appendix C.

2-dimensional case. We consider an ideal gas $\gamma=1.4$ and a linear gravitational potential $\Phi=g(x+y)$. We are interested in preserving the following isothermal equilibrium state

Table 3: Time to solution for initial conditions (4.4) for $\eta = 1\times10^{-8}$ in seconds (s).

N	DG4	DG5	WBDG2	WBDG3
8	0.37	0.56	0.05	0.14
16	0.62	1.28	0.10	0.29
32	2.05	4.79	0.20	0.69
64	12.8	32.4	0.62	3.25
128	103	270	3.84	24.1

on a unit square domain $\mathbf{x}\in[0,1]\times[0,1]$:

$$\begin{aligned}
&\rho_{eq}(x,y) = \rho_0 \exp\left(-\frac{\rho_0 g}{p_0}(x+y)\right),\\
&u_{eq}(x,y) = 0,\\
&v_{eq}(x,y) = 0,\\
&p_{eq}(x,y) = p_0 \exp\left(-\frac{\rho_0 g}{p_0}(x+y)\right),
\end{aligned} \tag{4.5}$$

with $\rho_0 = 1$, $p_0 = 1$ and $g = 1$.

The numerical errors for the pressure are reported in Fig. 5 for the following resolutions $N = 8$, 16, 32, 64, evaluated at final time $T = 10.0$. Similarly to the 1-dimensional case, one can observe that the truncation error can be reduced again by increasing the number of cells or the order, as expected.

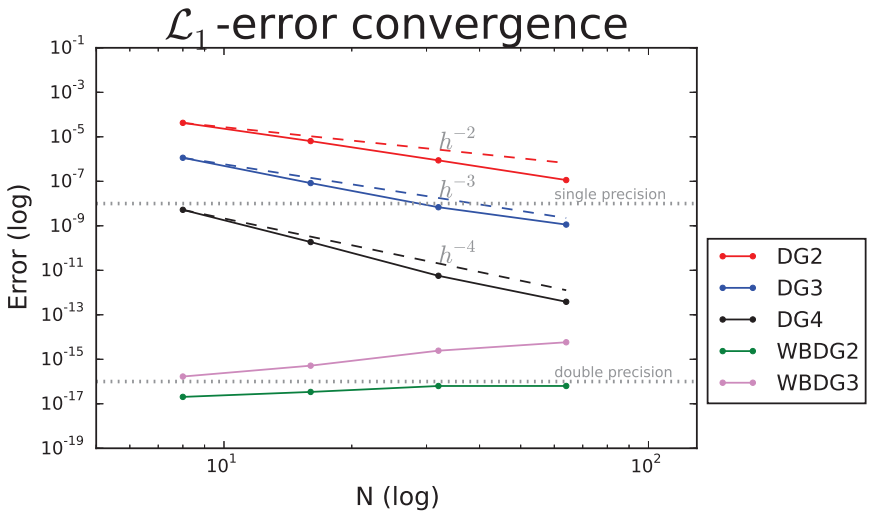

Figure 5: $\mathcal{L}_1$ error convergence for the 2-dimensional hydrostatic test case (4.5).

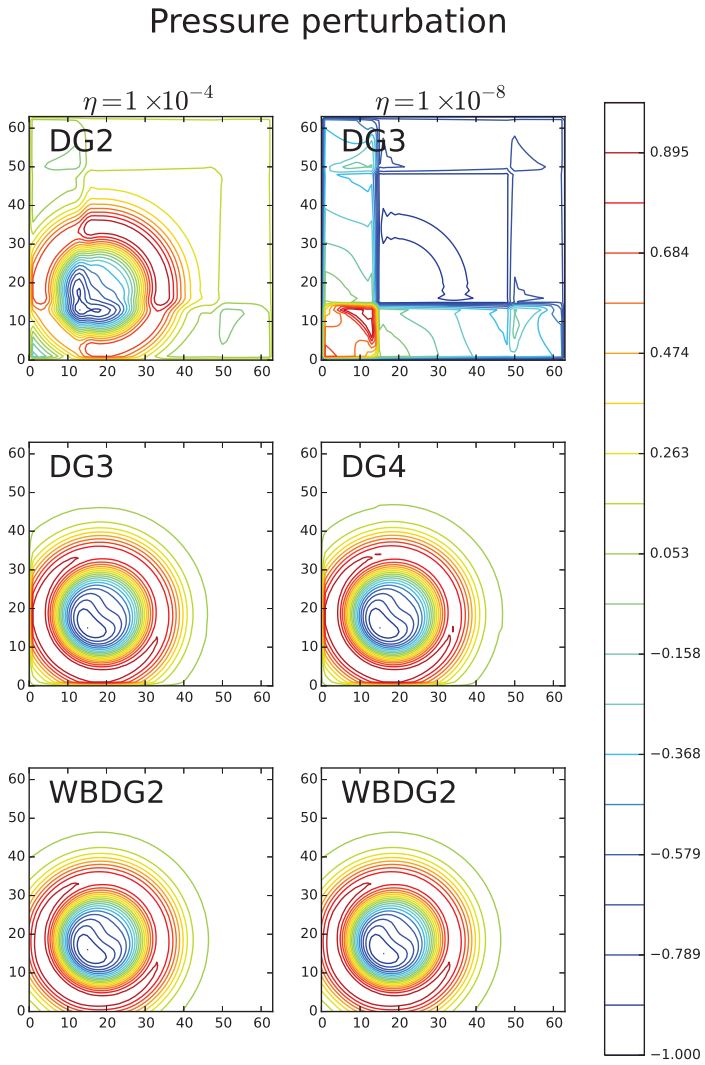

Figure 6: Non well-balanced method vs well-balanced method for hydrostatic equilibrium with varying amplitude perturbation on the pressure field for initial conditions (4.6).

Again, as in 1-dimension, a perturbation is added to the pressure state of the isothermal equilibrium solution:

$$p(x,y,t=0) = p_{eq}(x,y) + \eta \exp\left(-\frac{\rho_0 g}{p_0}\left(\frac{(x-0.3)^2+(y-0.3)^2}{0.01}\right)\right). \tag{4.6}$$

The initial condition (4.6) is run with different pulse amplitudes: $\eta = 1\times10^{-4}$ and 1×10^{-8}. The results are shown in Fig. 6. Again, we observe that by increasing the order, we can resolve for small perturbations, but we have to choose the resolution carefully

to guarantee that the pulse is captured accurately. As before, the well-balanced methods capture the wave solution correctly, even with a second-order scheme. Further analysis, such as the pointwise L_1 error of between the solution attained with different orders of the non-well balanced discontinuous Galerkin scheme and a high resolution solution and simulation time to solution can be found in Appendix C.

4.2.2 Non-hydrostatic steady state

1-dimensional case. We consider the manufactured example[‡] of an ideal steady gas $\gamma = 1.4$ with a nonzero velocity field and a gravitational field which balances the flux term exactly. We are interested in preserving the following *moving* equilibrium state:

$$\begin{aligned} \rho_{eq}(x) &= \rho_0 \exp\left(-\frac{\rho_0 g}{p_0} x\right), \\ u_{eq}(x) &= \exp(x), \\ p_{eq}(x) &= \exp\left(-\frac{\rho_0 g}{p_0} x\right)^{\gamma}, \end{aligned} \tag{4.7}$$

with $\rho_0 = 1$, $p_0 = 1$ and a non linear potential $\phi = \exp(x)(-\exp(x) + \gamma \exp(-\gamma x))$. The boundary values are imposed as in (4.3). The results are shown in Fig. 7 for $T = 10.0$.

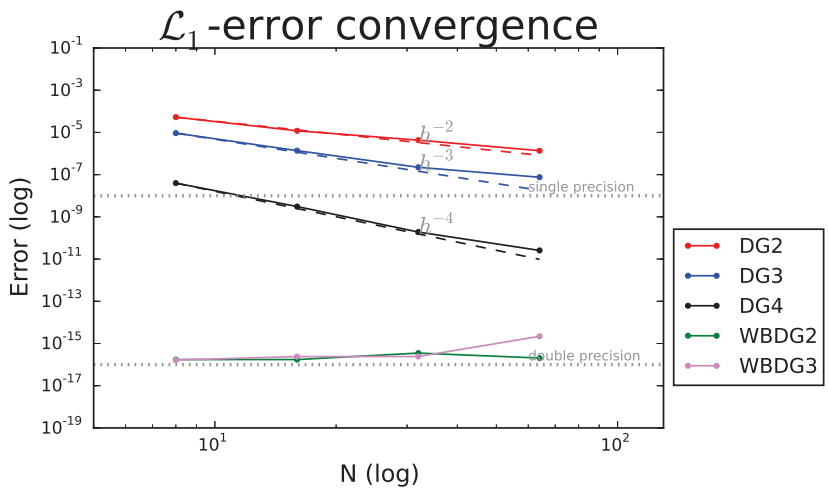

Figure 7: $\mathcal{L}_1$-error convergence for the 1-dimensional dynamic test case (4.7).

Now, just as in the hydrostatic equilibrium case (4.4), a perturbation is added to the pressure field:

$$p(x, t=0) = p_{eq}(x) + \eta \exp\left(-\frac{\rho_0 g}{p_0} \frac{(x-0.3)^2}{0.01}\right). \tag{4.8}$$

[‡]For details of this initial condition, refer to Appendix A.

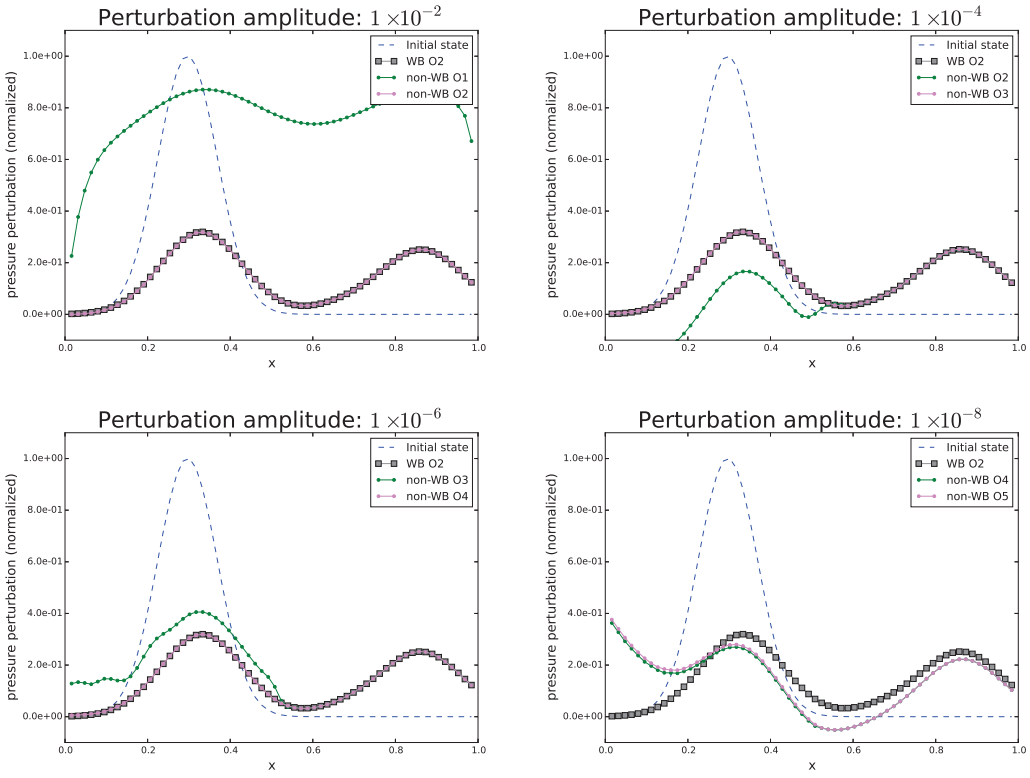

Figure 8: Non well-balanced method vs well-balanced method for dynamic equilibrium with varying amplitude perturbation on the pressure field as described in (4.8).

We run the numerical experiment with different pulse amplitudes: $\eta = 1\times10^{-2}$, 1×10^{-4}, 1×10^{-6} and 1×10^{-8}. The results are shown in Fig. 8. Our conclusions remain the same as for the hydrostatic case: for non-well-balanced methods, only a very high order scheme can capture the low amplitude wave correctly. It appears from Fig. 8 that the largest truncation error arises from the left boundary and propagates in the direction of the flow. On the contrary, our second-order, well balance method can deal with vanishingly small amplitude waves. Further analysis, such as the pointwise L_1 error of between the solution attained with different orders of the non-well balanced discontinuous Galerkin scheme and a high resolution solution and simulation time to solution can be found in Appendix C.

2-dimensional case. Modified steady vortex. We consider a modified gresho vortex, where the pressure is modified to balance exactly a gravity source term. The initial conditions for the primitive variables are:

$$\rho = 1.0, \quad v_x = -v_\theta \frac{(y-y_c)}{r}, \quad v_y = v_\theta \frac{(x-x_c)}{r}, \quad p = p(r), \tag{4.9}$$

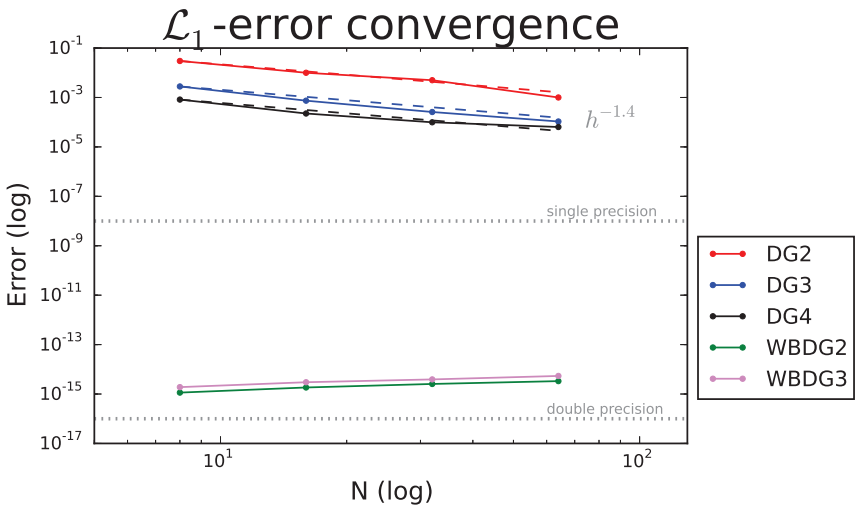

Figure 9: $\mathcal{L}_1$-error convergence for the 2-dimensional modified gresho vortex (see Eq. (4.9)).

with the cross-radial velocity v_θ and pressure p:

$$v_\theta(r) = \begin{cases} 5r, & r < 0.2, \\ 2-5r, & 0.2 \le r < 0.4, \\ 0, & r \ge 0.4, \end{cases}$$

$$p(r) = \begin{cases} 5+\frac{25}{2}r^2 - \alpha\Phi, & r < 0.2, \\ 9-4\log(0.2)+\frac{25}{2}r^2 - 20r + 4\log(r) - \alpha\Phi, & 0.2 \le r < 0.4, \\ 3+4\log(2) - \alpha\Phi, & r \ge 0.4, \end{cases}$$

where $\alpha = 0.01$ and $\Phi = \frac{1}{r}$. One can easily verify that adding a gravity source term with the potential $\alpha\Phi$ recovers the gresho vortex analytically and that this is a steady state solution of the Euler system. We run this initial condition until $T = 1.0$. In Fig. 9 we can see the measured empirical $\mathcal{L}_1$-norm between the well balanced discretisation and the traditional discontinuous Galerkin method. We recover the expected convergence rate is of $\mathcal{O}(1.4)$ for the traditional Gresho vortex test case. Including the gravity source term does not alter the convergence properties of the scheme. We also see that the well-balance scheme maintain the equilibrium state to machine precision accuracy.

Simplified protoplanetary disc. In the context of planet formation, it is customary to consider a stationary disc rotating around a single star, which is a steady state solution of the Euler-Poisson equations.

In this paper, we consider a constant density disc defined in a $[-6,6]\times[-6,6]$ box,

with the following initial conditions:

$$\rho_{eq}=1.0,\quad u_{eq}=-\frac{v_\theta}{r}y,\quad v_{eq}=\frac{v_\theta}{r}x,\quad p_{eq}=c_s^2\rho_{eq}, \tag{4.10}$$

where $v_\theta=\sqrt{\frac{1}{r}(1-\alpha^2)}$ is the orbital velocity (slightly sub-Keplerian), $c_s=\alpha v_K$ the speed of sound, given by the product of the Keplerian velocity v_K and the disk aspect ratio $\alpha=0.03$, and the gravity potential of a unit point mass given by $\Phi=-\frac{1}{r}$.

We now describe in details how we set up our boundary conditions, for which great care is required in order to preserve the correct geometry of the problem and to stabilise the solution:

- For the domain boundary conditions (on the box $[-6,6]\times[-6,6]$), the steady state solution is just imposed in ghost elements, as shown in (4.3).

- To minimise spurious effects due to the rotation of the disk near the end of the box domain, the constant density field ρ_{eq} is multiplied with a tampering function $d(r)$. The following tampering function is taken:

$$d(r)=\frac{1}{1+\left(\frac{r}{r_0}\right)^q},$$

  setting $q=20$, $r_0=4.2$. This function was adopted after several other functions have been tried. Note that for the stability of the RKDG method it is important to consider functions which have well behaved derivatives at all orders. Another good candidate we have tried is the sigmoid function (not shown here).

- The disc is an isolated system with no mass inflow and a (tampered) sharp edge. We need to introduce a buffer region near the disc edge where propagating waves are damped to reduce wave reflection. We use a methodology similar to [10], which smoothly relaxes the numerical solution to the equilibrium solution at the edge of the buffer zone using a function $R(r)$ so that

$$\tilde{H}(u)=H(u)R(r).$$

  Note that this function must leave the solution unaltered outside of the buffer region. In our experiments we set $R(r)=\frac{1}{1+\exp(r^2-15.0)}$ where the parameter 15.0 was chosen to set the size of the buffer region. In [10], the authors used a parabolic function $R(r)$ instead.

- Similarly, at the centre of the disk, around $r=0.0$, we use an inner buffer region where the numerical solution is set to the steady state solution. An inner radius of $r<0.75$ is considered for the size of the inner buffer region.

A perturbation is then added to the gravity field of the star. Physically, this perturbation can be interpreted as a planet. As such, the magnitude of the gravitational force exerted by the planet is very small in comparison to the gravitational force exerted by the star. We introduce this perturbation in the second term of Eq. (4.11)

$$\nabla\Phi(\mathbf{x}) = \frac{\mathbf{x}}{(r^2+\epsilon^2)^{\frac{3}{2}}} + \eta\frac{\mathbf{x}-\mathbf{x}_p}{(r_p^2+\epsilon^2)^{\frac{3}{2}}}, \tag{4.11}$$

where $r_p = \sqrt{||\mathbf{x}-\mathbf{x_p}||}$, $\mathbf{x}_p$ denotes the position of the perturbation

$$\mathbf{x}_p = \begin{pmatrix} x_p \\ y_p \end{pmatrix} = \begin{pmatrix} r_c\cos\left(\frac{v_K}{r_c}t\right) \\ r_c\sin\left(\frac{v_K}{r_c}t\right) \end{pmatrix},$$

fixed to be a circular orbit at $r=2.2$ with Keplerian velocity v_K and $\epsilon=0.01$ is the softening length for the planet. By varying η, we control the size of the perturbation. We test different sizes of η to denote different sized planets, namely, $\eta = 3.1\times10^{-6}$, 9×10^{-5} and 9.5×10^{-4} which correspond to Earth, Neptune and Jupiter sized planets, respectively.

The system is evolved until 20 rotations are performed at $r = 2.2$, corresponding to approximately $T = 410$ in our normalised units.

The results after the planet has performed only one rotation can be seen in Fig. 10 and after 10 rotations in Fig. 11. For the smallest perturbation ($\eta = 3.1\times10^{-6}$), we note that already after one full rotation, the solution of the DG2 method has interacted with the waves generated by the mismatch between the inner boundary condition and the evolved solution, and this effect disappears when increasing the method to 3^{rd} order or when using the WBDG2 scheme. After 10 rotations it's clear that the perturbation has been lost in the numerical errors when using DG2, whereas for DG3 the solution remains very clean, both in the perturbation and the steady state background solution. Similarly, when using WBDG2, we observe a very clean perturbation on top of the unperturbed steady state background, although the resolution on the perturbation is lower than in the DG3 case.

A similar behaviour is observed for the medium amplitude perturbation ($\eta=9\times10^{-5}$), after one rotation. After 10 rotations, although the spiral density wave can be seen in all methods, both when using DG2 and DG3, artefacts are observed in the gap opened by the planet, whereas when using WBDG2 the gap remains cleaner. For the larger perturbation ($\eta=9.5\times10^{-4}$), even though the effect from the boundary is still present, we see virtually no difference between DG2 and WBDG2. For such large sized planets, it is expected for a gap to be carved in the disc, and the regime of the study is very different. Indeed, after 10 rotations the disk is visibly unstable, and the solution has deviated enough from the steady state background solution that there's virtually no difference between DG2 and WBDG2. Indeed, for the simulation to reach 10 rotations, we had to stabilise all methods by using a positivity preserving limiter. Note that the large amplitude case is particularly

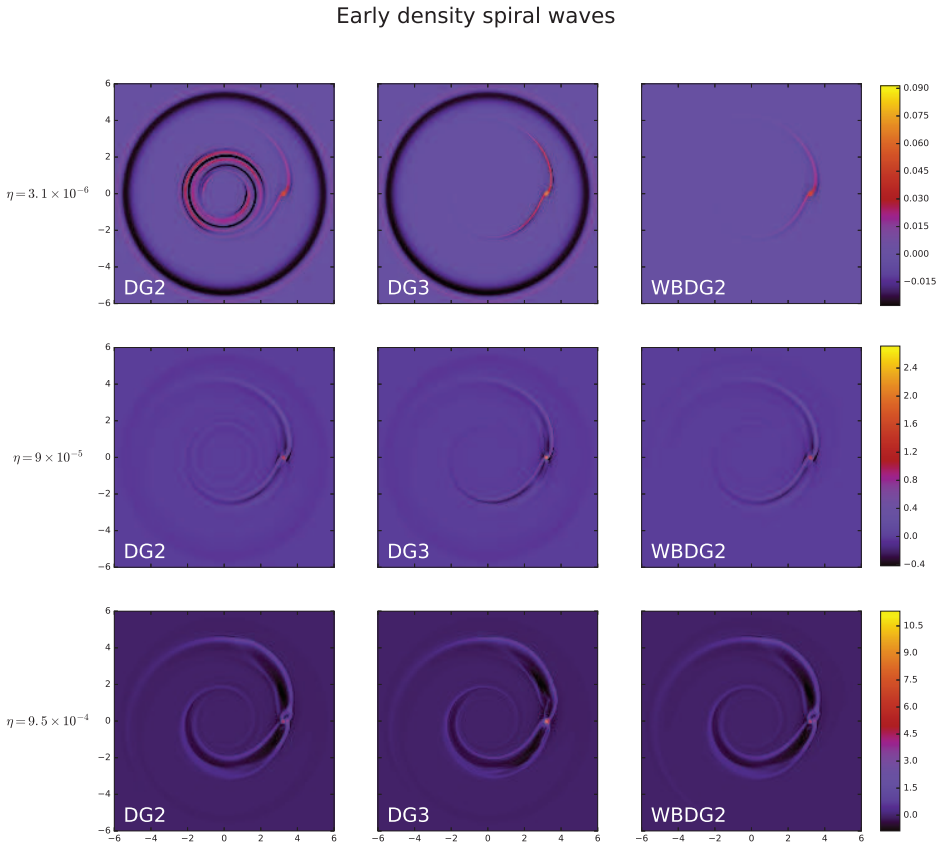

Figure 10: Density perturbations for non well-balanced method versus well-balanced method from dynamic equilibrium for varying perturbation of sizes on the gravity field, after 1 rotation, at approximately $T = 21$.

interesting, because it demonstrates that our well-balanced scheme is robust enough to sustain large deviations from the adopted equilibrium state, recovering the properties of the corresponding non-well-balanced scheme.

Lastly, as denoted in Table 4, we show the time to solution required for different non well-balanced and well-balanced methods. In this example, it becomes clear the advantage of using a well-balanced scheme for long term evolution of small perturbations, as

Table 4: Time to solution for the protoplanetary disc case after 10 rotations, for varying planet sizes.

η	DG2	DG3	WBDG2
3.1×10^{-6}	1h42m	16h30m6	3h18m
9×10^{-5}	1h42m	16h10m	3h21m
9.5×10^{-4}	1h30m	15h30m	3h30m

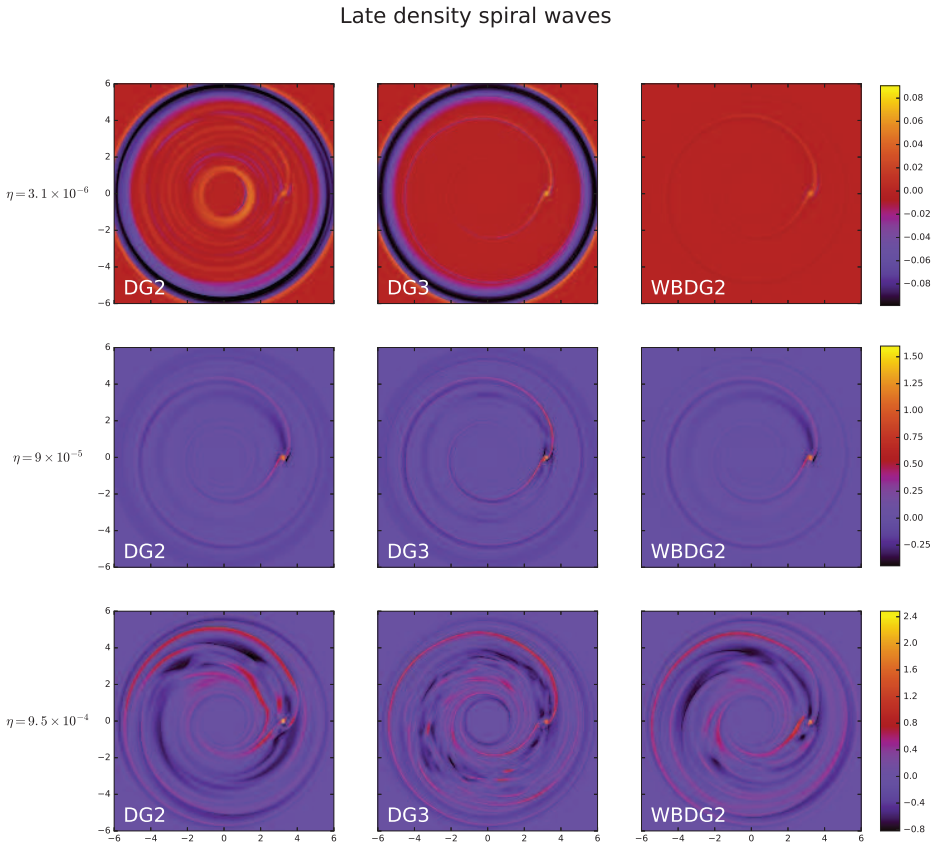

Figure 11: Density perturbations for non well-balanced method versus well-balanced method from dynamic equilibrium for varying perturbation of sizes on the gravity field, after 10 rotations, at approximately $T = 210$.

the necessary increase in order and resolution in the non well-balanced case can be translated into much longer simulation times.

5 Conclusion

The motivation of this paper was to address the following three research questions:

- **RQ 1:** Are there cases where using a high order scheme is sufficient to capture solutions close to a steady state?
- **RQ 2:** Under which circumstances is it necessary to use a well balanced method?
- **RQ 3:** What is the cost associated to each approach and how does it balance with accuracy?

To address these questions, we compared a classical RKDG scheme with a novel well-balanced RKDG scheme using different numerical examples. We study the performance of these two numerical schemes under the regime of hydrostatic equilibrium and *dynamic* equilibrium. This last type of equilibrium is of interest for many studies and simulations of proto-planetary systems. A summary of our results from the numerical experiments shown in Section 4 follows:

- When considering **hydrostatic equilibrium in one space dimension**, the non-well balanced high order method behaves very well. In particular, for waves with amplitude larger than the scheme's truncation error, the method was able to resolve the perturbation accurately as expected. For example, when using the 3^{rd} or 4^{th} order method, we were able to reduce the truncation error down to $\mathcal{O}(10^{-8})$ and $\mathcal{O}(10^{-12})$ for a resolution of $N=64$, respectively.

- Studying the **hydrostatic equilibrium case in a 2-dimensional setting**, we were able to reduce the error to $\mathcal{O}(10^{-8})$ and $\mathcal{O}(10^{-12})$ only for a resolution of $N_x = 64$, $N_y = 64$ when using a 3^{rd} or 4^{th} order method, respectively. However, it was observed that when using a well-balanced scheme, the required resolution (either in space or in polynomial degree) could be lowered without affecting the ability of the scheme to capture the waves.

- Considering a **steady state with a non-trivial velocity in a 1-dimensional setting** (*dynamic* equilibrium), we are able to reduce the error to $\mathcal{O}(10^{-6})$ and $\mathcal{O}(10^{-10})$ for a resolution of $N=64$, for a 3^{rd} or 4^{th} order, respectively. When considering the same initial condition with a smaller perturbation in the pressure field, high order methods appear to fail in capturing properly the wave (in particular for perturbation $\eta = 1\times 10^{-8}$). Only our well-balanced scheme is robust enough to capture the wave dynamics properly.

- For the **modified Gresho vortex, a 2-dimensional steady state solution with a non-trivial velocity**, we observe that the non well-balanced method converges with the expected numerical accuracy, even in the presence of the source terms. The well-balanced scheme is preserving the stationary solution down to machine precision accuracy.

- For the **idealised disc-planet interaction case**, a 2-dimensional steady state solution with a non-trivial velocity, we observe that for small and medium amplitude perturbations, the well-balanced scheme is clearly superior, being able to keep the background solution for longer than the non well-balanced methods, although the increase in the approximation order improves the solution significantly. In particular, for the medium amplitude perturbation we observe the opening of a shallow gap. In both DG2 and DG3 this gap is not very clean. In the context of planet migration, the gap opening is important because it decreases the angular momentum exchange between the planet and disc, which in turn is translated into a decrease

of the migration rate of the planet [24]. For both the small and medium amplitude perturbations, we are able to observe the expected density spiral arms without the strong numerical artefacts that we see when using classical high order schemes. Indeed, in this case the density wave seems to be corrupted mostly from the inner and outer boundaries. In order to stabilise the solution when using DG2 and DG3, in addition to sophisticated boundary conditions, we have to use a positivity preserving limiter [15]. For the large amplitude perturbation, we see no difference between DG2 and WBDG2. Our hypothesis is that the solution can't be represented as a simple superposition of a steady state solution and a time dependent small perturbation, however, we note that the WBDG2 scheme behaves as DG2 (also requiring the positivity preserving limiter), which points towards our well-balanced method being robust for large perturbations.

The use of a well-balanced scheme for the rotating disk case seems a good compromise (**RQ2**), while for simpler 1-dimensional hydrostatic equilibrium problems, one can beat down the truncation error fairly easily by raising either the order or the resolution of the scheme, without using the well balance correction (**RQ1**).

We note that the well-balanced correction does not come without a cost (**RQ3**). As shown in Fig. 2, the well-balanced correction can slow down the code significantly. This can be alleviated by pre-computing and storing all the variables from the steady state solution. However, this means that the memory requirements for this algorithm almost double (in comparison to the classical RKDG scheme). Due to the compute intensity of DG methods, GPUs are usually the appropriate hardware to run these methods [12], which are often limited in memory so this is something worth considering when choosing the appropriate implementation.

Note than one very restrictive condition for our well-balanced scheme is to know the exact form of the equilibrium solution everywhere, either in analytical or tabulated form. Moreover, a high-order scheme, if it is robust enough to capture the wave dynamics, will always deliver higher accuracy than a well-balanced, low order scheme. In [30], we show that the DG method can be a competitive method in planet-disc interaction studies, even without the well balanced correction, if one uses enough grid points, a conservative slope limiter and a careful boundary condition strategy.

In conclusion, we have shown that the well balanced property in numerical schemes is important for two reasons: 1- we are able to capture low amplitude waves propagating in non-trivial equilibrium states without resorting to complex boundary conditions strategies and 2- we are able to solve for very small perturbations using lower order methods, which requires significantly less computations per time step, in particular when considering multi-dimensional problems. Both these points are relevant for setups like the one discussed here, namely the long term evolution of the sightly perturbed multi-dimensional equilibrium disc solution. As further steps, we hope to test and potentially study an extension of this scheme to capture equilibrium solution with discontinuities and to consider arbitrary general equilibrium states along the lines of [25].

Acknowledgments

MHV is supported by the UZH Candoc Scholarship. The computing resources were provided by the S3IT cluster at University of Zurich. The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

Appendices

A One dimensional moving steady state solution

In this section we are interested in the construction of a simple 1-dimensional test case which is a steady state solution to the 1-dimensional Euler equations with a non trivial velocity field. The objective is to find the quartet of functions (ρ,v,p,ϕ) such that they fulfil the following:

$$\frac{\partial}{\partial x}(\rho v)=0, \tag{A.1a}$$

$$\frac{\partial}{\partial x}(\rho v^2+p)=-\rho\frac{\partial}{\partial x}\Phi, \tag{A.1b}$$

$$\frac{\partial}{\partial x}\big((E+p)v\big)=-\rho v\frac{\partial}{\partial x}\Phi. \tag{A.1c}$$

From (A.1a), we have $\rho v=$const, whereas for (A.1b) and (A.1c):

$$\rho v\frac{\partial}{\partial x}v+\frac{\partial}{\partial x}p=-\rho\frac{\partial}{\partial x}\Phi,$$
$$\frac{\partial}{\partial x}\big((E+p)v\big)=-\rho v\frac{\partial}{\partial x}\Phi.$$

Noting that $E=\frac{p}{\gamma-1}+\frac{1}{2}\rho v^2$, (A.1c) yields:

$$\rho v^2\frac{\partial}{\partial x}v+\frac{\partial}{\partial x}\left(pv\frac{\gamma}{\gamma-1}\right)=-\rho v\frac{\partial}{\partial x}\Phi.$$

Substituting (A.1b) into (A.1c), one can solve find p if we assume some form for ρ (and consequently for v).

Setting $\rho=\exp(-x)$, thus $v=\exp(x)$ and $p=exp(-\gamma x)$. An expression for Φ can be written by solving the differential equation in (A.1b), yielding: $\frac{\partial}{\partial x}\Phi=\exp(x)(-\exp(x)+\gamma\exp(-\gamma x))$.

B A simple equilibrium solution for proto-planetary discs

The orbital speed for a gas can be calculated from the Euler-Poisson equations:

$$\frac{\partial \mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}=-\frac{1}{\rho}\nabla p-\nabla\Phi,$$

where p denotes the pressure, ρ the density and Φ the gravitational potential. We can rewrite the second term as[§]:

$$(\mathbf{v}\cdot\nabla)\mathbf{v}=\frac{1}{2}\nabla\mathbf{v}^T\mathbf{v}-\mathbf{v}\times(\nabla\times\mathbf{v}).$$

Assuming a steady state ($\frac{\partial}{\partial t}=0$) axisymmetric solution, we derive the orbital velocity:

$$\frac{v_\phi^2}{r}=\frac{1}{\rho}\frac{\partial}{\partial r}p+\frac{\partial}{\partial r}\Phi.$$

Furthermore, defining the Keplerian velocity: $v_{\rm K}(r)=\sqrt{r\frac{\partial}{\partial r}\Phi}$ and the constant disk aspect ratio [2] with

$$\alpha=\frac{\sqrt{\frac{p(r)}{\rho(r)}}}{v_{\rm K}(r)},$$

we can deduce the relation for the pressure to be $p(r)=\alpha^2\rho(r)v_{\rm K}^2$. Finally, we obtain the equilibrium tangential velocity v_ϕ knowing the constant α and the profile $\rho(r)$.

C Supplementary results

In this section we provide the time to solution for the numerical experiments performed in Section 4 and convergence plots associated to the perturbation tests.

C.1 1-dimensional hydrostatic

For convenience, we restate the initial conditions: an ideal gas $\gamma=1.4$ in isothermal equilibrium state and a linear gravitational potential $\Phi_x=gx$ is considered:

$$\begin{aligned}
&\rho_{eq}(x)=\rho_0\exp\left(-\frac{\rho_0 g}{p_0}x\right),\\
&u_{eq}(x)=0,\\
&p(x,t=0)=p_{eq}(x)+\eta\exp\left(-\frac{\rho_0 g}{p_0}\frac{(x-0.5)^2}{0.01}\right),
\end{aligned}\tag{C.1}$$

with $\rho_0=1.0$, $p_0=1.0$ and $g=1.0$.

[§]Using the following identity $\nabla(\mathbf{A}\cdot\mathbf{B})=\mathbf{A}\times(\nabla\times\mathbf{B})-(\nabla\times\mathbf{A})\times\mathbf{B}+(\mathbf{A}\cdot\nabla)\mathbf{B}+(\mathbf{B}\cdot\nabla)\mathbf{A}$.

Table 5: Time to solution for the 1-dimensional hydrostatic equilibrium (C.1) (s) for perturbation size $\eta = 1\times10^{-2}$, 1×10^{-4} and 1×10^{-6}, respectively.

N	DG1	DG2	WBDG2	WBDG3
8	0.37	0.56	0.05	0.14
16	0.62	1.28	0.10	0.29
32	2.05	4.79	0.20	0.69
64	12.8	32.4	0.62	3.25
128	103	270	3.84	24.1

N	DG2	DG3	WBDG2	WBDG3
8	0.37	0.56	0.05	0.14
16	0.62	1.28	0.10	0.29
32	0.18	4.79	0.20	0.69
64	12.8	32.4	0.62	3.25
128	103	270	3.84	24.1

N	DG2	DG3	WBDG2	WBDG3
8	0.37	0.56	0.05	0.14
16	0.62	1.28	0.10	0.29
32	2.05	4.79	0.20	0.69
64	12.8	32.4	0.62	3.25
128	103	270	3.84	24.1

C.2 2-dimensional hydrostatic

Ideal gas $\gamma=1.4$, in isothermal equilibrium and a linear gravitational potential $\Phi=g(x+y)$. Unit square domain $\mathbf{x}\in[0,1]\times[0,1]$:

$$\begin{aligned}
&\rho_{eq}(x,y)=\rho_0\exp\left(-\frac{\rho_0 g}{p_0}(x+y)\right),\\
&u_{eq}(x,y)=0,\\
&v_{eq}(x,y)=0,\\
&p_{eq}(x,y)=p_0\exp\left(-\frac{\rho_0 g}{p_0}(x+y)\right),
\end{aligned} \tag{C.2}$$

with $\rho_0=1$, $p_0=1$ and $g=1$. The time to solution is shown on Tables 6, 7 and error convergence plots in Fig. 12.

Table 6: Time to solution for hydrostatic equilibrium (C.2) (s) at $T=10.0$.

N_x	DG2	DG3	DG4	WBDG2	WBDG3
8	1.53	3.36	10.8	1.85	4.63
16	2.92	7.32	24.6	3.74	9.52
32	6.88	19.9	82.1	8.74	24.7
64	19.3	101	525	2.46	130
128	119	777	4310	154	981

Table 7: Time to solution for hydrostatic equilibrium (C.2) (s) for perturbation sizes $\eta=1\times10^{-4}$, 1×10^{-8} at $T=0.25$.

N_x	DG2	DG3	WBDG2	WBDG3
8	0.04	0.08	0.05	0.10
16	0.08	0.17	0.09	0.22
32	0.17	0.49	0.22	0.64
64	0.50	2.51	0.69	3.26
128	2.98	19.5	3.84	24.5

N_x	DG3	DG4	WBDG2	WBDG3
8	0.08	0.27	0.03	0.10
16	0.17	0.62	0.07	0.22
32	0.50	2.06	0.17	0.62
64	2.50	13.2	0.61	3.26
128	19.5	108	3.86	24.5

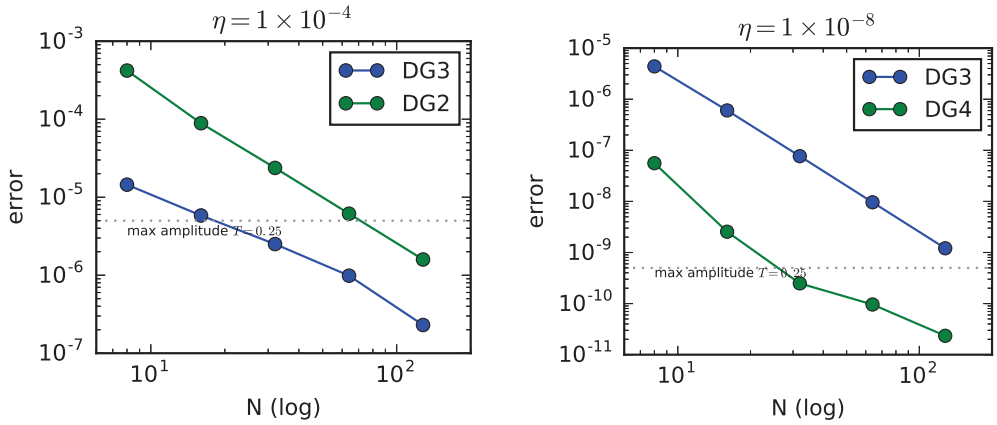

Figure 12: Non well-balanced method vs well-balanced method for hydrostatic equilibrium with varying amplitude perturbation on the pressure field for initial conditions (C.2).

C.3 1-dimensional dynamic

Ideal steady gas $\gamma=1.4$ with a nonzero velocity field and a non linear gravitational field

$$\begin{aligned} \rho_{eq}(x) &= \rho_0 \exp\left(-\frac{\rho_0 g}{p_0}x\right), \\ u_{eq}(x) &= \exp(x), \\ p_{eq}(x) &= \exp\left(-\frac{\rho_0 g}{p_0}x\right)^{\gamma}, \end{aligned} \tag{C.3}$$

with $\rho_0=1$, $p_0=1$ and a non linear potential $\phi=\exp(x)(-\exp(x)+\gamma\exp(-\gamma x))$. The time to solution for this test case is very similar to the 1-dimensional hydrostatic equilibrium case, and is thus omitted. The error convergence plots are shown in Fig. 13.

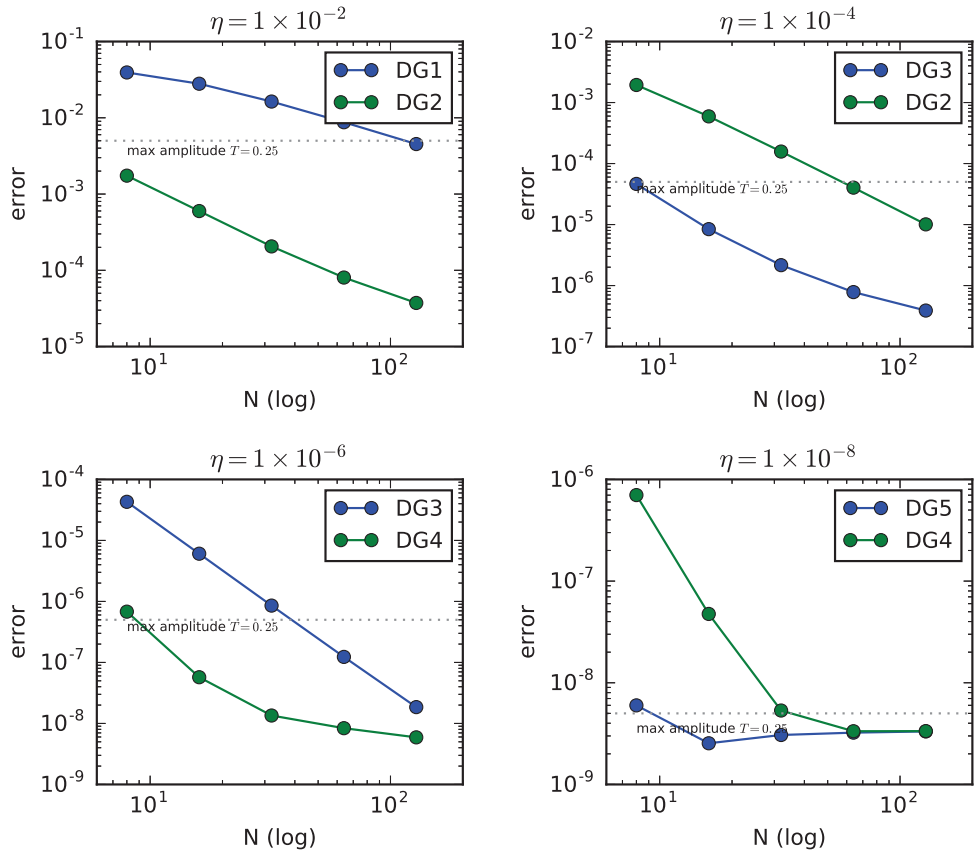

Figure 13: Non well-balanced method versus well-balanced method for hydrostatic equilibrium with varying amplitude perturbation on the pressure field for initial conditions (C.3).

References

[1] D. Antonio Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*, volume 69. 01 2012.

[2] P. J. Armitage. Dynamics of protoplanetary disks. *Annual Review of Astronomy and Astrophysics*, 49(1):195–236, 2011.

[3] A. Bermúdez, X. López, and M. E. Vázquez-Cendón. Numerical solution of non-isothermal non-adiabatic flow of real gases in pipelines. *J. Comput. Phys.*, 323(C):126–148, October 2016.

[4] A. Bermúdez and M. E. Vazquez. Upwind methods for hyperbolic conservation laws with source terms. *Computers & Fluids*, 23(8):1049–1071, 1994.

[5] C. Berthon and F. Foucher. Efficient well-balanced hydrostatic upwind schemes for shallow-water equations. 231:49935015, 06 2012.

[6] M. J. Castro, A. Pardo Milanés, and C. Parés. Well-balanced numerical schemes based on a generalized hydrostatic reconstruction technique. *Mathematical Models and Methods in Applied Sciences*, 17(12):2055–2113, 2007.

[7] P. Chandrashekar and M. Zenk. Well-balanced nodal discontinuous Galerkin method for

Euler equations with gravity. *ArXiv e-prints*, November 2015.

[8] P. Chandrashekar and C. Klingenberg. A second order well-balanced finite volume scheme for Euler equations with gravity. *SIAM Journal on Scientific Computing*, 37(3):B382–B402, 2015.

[9] B. Cockburn and C.-W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws v. *J. Comput. Phys.*, 141(2):199–224, April 1998.

[10] M. De Val-Borro, R. G. Edgar, P. Artymowicz, P. Ciecielag, P. Cresswell, G. D'Angelo, E. J. Delgado-Donate, G. Dirksen, S. Fromang, A. Gawryszczak, H. Klahr, W. Kley, W. Lyra, F. Masset, G. Mellema, R. P. Nelson, S.-J. Paardekooper, A. Peplinski, A. Pierens, T. Plewa, K. Rice, C. Schäfer, and R. Speith. A comparative study of discplanet interaction. *Monthly Notices of the Royal Astronomical Society*, 370(2):529–558, 2006.

[11] A. Dedner, I.L. Sofronov, and M. Wesenberg. Transparent boundary conditions for mhd simulations in stratified atmospheres. *Journal of Computational Physics*, 171(2):448–478, 2001.

[12] M. Fuhry, A. Giuliani, and L. Krivodonova. Discontinuous Galerkin methods on graphics processing units for nonlinear hyperbolic conservation laws. *CoRR*, abs/1601.07944, 2016.

[13] S. Gottlieb and C.-W. Shu. Total variation diminishing Runge-Kutta schemes. *Math. Comput.*, 67(221):73–85, January 1998.

[14] J. M. Greenberg and A. Y. Leroux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM Journal on Numerical Analysis*, 33(1):1–16, 1996.

[15] X. Y. Hu, N. A. Adams, and C.-W. Shu. Positivity-preserving method for high-order conservative schemes solving compressible Euler equations. *Journal of Computational Physics*, 242(Supplement C):169–180, 2013.

[16] R, Käppeli and S. Mishra. A well-balanced finite volume scheme for the Euler equations with gravitation – the exact preservation of hydrostatic equilibrium with arbitrary entropy stratification. *A&A*, 587:A94, 2016.

[17] R.J. LeVeque, O. Steiner, A. Gautschy, D. Mihalas, E.A. Dorfi, and E. Müller. *Computational Methods for Astrophysical Fluid Flow: Saas-Fee Advanced Course 27. Lecture Notes 1997 Swiss Society for Astrophysics and Astronomy*. Saas-Fee Advanced Course. Springer Berlin Heidelberg, 2006.

[18] G. Li and Y. Xing. Well-balanced discontinuous Galerkin methods for the Euler equations under gravitational fields. *Journal of Scientific Computing*, 67(2):493–513, May 2016.

[19] G. Li and Y. Xing. Well-balanced discontinuous Galerkin methods with hydrostatic reconstruction for the Euler equations with gravitation. *Journal of Computational Physics*, 352:445–462, 2018.

[20] L.A. McFadden, T. Johnson, and P. Weissman. *Encyclopedia of the Solar System*. Encyclopedia of the Solar System Series. Elsevier Science, 2006.

[21] H. Mo, F. van den Bosch, and S. White. *Galaxy Formation and Evolution*. Galaxy Formation and Evolution. Cambridge University Press, 2010.

[22] S. Noelle, Y. Xing, and C.-W. Shu. High-order well-balanced schemes. In *Numerical methods for balance laws*, pages 1–66. Caserta: Dipartimento di Matematica, Seconda Università di Napoli, 2009.

[23] C. Parés. Numerical methods for nonconservative hyperbolic systems: a theoretical framework. *SIAM Journal on Numerical Analysis*, 44(1):300–321, 2006.

[24] R. R. Rafikov. Planet migration and gap formation by tidally induced shocks. *The Astrophysical Journal*, 572(1):566, 2002.

[25] M. Ricchiuto. An explicit residual based approach for shallow water flows. *J. Comput. Phys.*, 280(C):306–344, January 2015.

[26] C.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-

capturing schemes, II. *J. Comput. Phys.*, 83(1):32–78, July 1989.

[27] T. Sun and D. Rumsey. Numerical smoothness and error analysis for RKDG on the scalar nonlinear conservation laws. *Journal of Computational and Applied Mathematics*, 241:68–83, 2013.

[28] C. Surville, L. Mayer, and D. N. C. Lin. Dust capture and long-lived density enhancements triggered by vortices in 2D protoplanetary disks. *The Astrophysical Journal*, 831(1):82, 2016.

[29] F.-K. Thielemann, K. Nomoto, and M.-A. Hashimoto. Core-Collapse Supernovae and Their Ejecta. *Applied Physics Journal*, 460:408, March 1996.

[30] D. A. Velasco, M. Han Veiga, F. Masset, and R. Teyssier. Planet-disc interactions with discontinuous Galerkin methods using GPUs. *MNRAS (submitted)*, 42(2):641–666, 2004.

[31] X. Zhang and C.-W. Shu. Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms. *Journal of Computational Physics*, 230(4):1238–1248, 2011.