UNIVERSIDAD DE LA FRONTERA

FACULTAD DE INGENIERÍA Y CIENCIAS

A NEW UNCONDITIONALLY POSITIVE AND CONSERVATIVE MODIFIED PATANKAR RUNGE KUTTA SCHEMES MPRK BASED ON OLIVER'S APPROACH

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor en Ciencias mención Matemática

AUTHOR: GALO JAVIER GONZÁLEZ HERNÁNDEZ

TEMUCO-CHILE

2021

Galo Javier González Hernández

Tutor: Dr. Andres Ávila Barrera

2000 Mathematics Subject Classification 65M12, 35K15, 35K65, 65M08, 76S05, 65M22

Keywords and phrases: Modified Patankar Runge Kutta methods, positivity, conservativity, second and third order, Oliver's approach.

INFORME DE APROBACIÓN

TESIS DE DOCTORADO
FACULTAD DE INGENIERIA Y CIENCIAS
UNIVERSIDAD DE LA FRONTERA

Se informa a la Vicerrectoría de Postgrado de la Universidad de La Frontera que la Tesis de Doctorado presentada por el candidato

GALO JAVIER GONZÁLEZ HERNÁNDEZ

ha sido aprobada por la Comisión de Evaluación de la tesis como requisito parcial para optar al grado de Doctor en Ciencias mención Matemáticas en el examen de Defensa de Tesis rendido el día xx/xx/2021

Nombre 1 (XXX) _____-

Tutor: Dr. Andres Ávila Barrera _____-

Nombre 2 (XXX) _____

Nombre 3 (XXX) _____-

Temuco, Chile, diciembre 2021

DEDICATORY

First to God.

To my princesses: Yara and Yerlis.

To my kings: Jeremy and Joel.

To my parents: Luisa and Vicente.

Galo Javier

ACKNOWLEDGEMENTS

I want to thank here starting for my thesis director, Professor Andrés Ávila, for his support both in the academic and personal sense, for his patience, for his way of working and above all for his advice, signs of support and words of encouragement, it is certainly very motivating.

Also, I want to thank Professor Meister and Dr. Kopecz from Kassel University for many fruitful discussions.

Finally, I want to thank to everyone the professors from Universidad de la Frontera for all their teaching.

To them: infinite thanks

ABSTRACT

The main objective of this thesis is to build MPRK–type schemes based on the Oliver's approach. Modified Patankar Runge Kutta (MPRK) schemes adapt explicit Runge Kutta schemes to ensure positivity and conservation of the solution of positive and conservative production–destruction systems irrespective of the time step size. These methods are highly stable and often outperform standard Runge–Kutta schemes.

Recently, Kopecz and Meister [23] give a general definition of MPRK schemes and based on the fundamental work of Burchard et al. [4] obtained the necessary and sufficient conditions for unconditionally positive and conservative first and second order scheme. Then, they also obtained conditions for third-order schemes and solved nonstiff and stiff systems of differential equations.

Inspired by the work of Kopecz and Meister, Huang et al. [18] modified the explicit Runge–Kutta scheme in the Shu and Osher form instead of the classical form and they developed another class of second and third order MPRK schemes, which have then been successfully applied to semi-discrete schemes arising from PDEs.

In this thesis, we extend MPRK methods, denoted MPRKO methods, using Oliver's [30] approach to improve the accuracy of these schemes in the field of nonautonomous systems. The approach does not require $\mathbf{Ae} = \mathbf{c}$ in the Butcher tableau $(\mathbf{A}, \mathbf{b}, \mathbf{c})$, where $\mathbf{e} = (1, \ldots, 1)^T$. Following the general analysis of MPRK schemes described by Kopecz and Meister, positivity and mass conservation fundamental properties are proven and even conditions concerning the Patankar weights are given to get second and third order accuracy of the MPRKO methods. Finally, we consider different linear models and a non-linear epidemiological SEIR problem as well as stiff Robertson test to confirm the theoretical results and to give reliable statements about the accuracy of the novel class of MPRKO methods.

RESUMEN

El principal objetivo de esta tesis es construir esquemas del tipo MPRK basados en el enfoque Oliver. Los esquemas Modified Patankar–Runge–Kutta (MPRK) adaptan los esquemas explícitos de Runge–Kutta para asegurar positividad y conservación de la solución de sistemas producción destrucción positivos y conservativos irrespectivamente del tamaño de paso de tiempo. Estos métodos son altamente estables y a menudo superan los esquemas estándar de Runge–Kutta.

Recientemente, Kopecz y Meister dan una definición general de los esquemas MPRK y basados en el trabajo fundamental de Burchard et al. obtuvieron las condiciones necesarias y suficientes para el esquema incondicionalmente positivo y conservativo de primer y segundo orden. También obtuvieron condiciones para esquemas de tercer orden y resuelven sistemas de ecuaciones diferenciales nonstiff y stiff.

Inspirado por el trabajo de Kopecz y Meister, Huang et al. modifican el esquema explícito de Runge–Kutta en la forma Shu y Osher en vez de la forma clásica y desarrollan otra clase de esquemas MPRK de segundo y tercer orden los cuales han sido satisfactoriamente aplicados a esquemas semi-discretos que surgen de ecuaciones diferenciales parciales.

En esta tesis, extendemos los métodos MPRK a los métodos denotados MPRKO utilizando el enfoque Oliver para mejorar la exactitud de estos esquemas en el campo de los sistemas no autónomos. El enfoque no requiere $\mathbf{Ae} = \mathbf{c}$ en la tabla de Butcher $(\mathbf{A}, \mathbf{b}, \mathbf{c})$, donde $\mathbf{e} = (1, \ldots, 1)^T$. Siguiendo el analisis general de los esquemas MPRK descritos por Kopecz y Meister, las propiedades fundamentales de positividad y conservación de masa son demostradas e incluso condiciones concernientes a los pesos Patankar son dadas para obtener segundo y tercer orden de exactitud de los métodos MPRKO. Finalmente, nosotros consideramos diferentes modelos lineales y un problema epidemiológico no lineal SEIR como también el test de Robertson altamente stiff para confirmar los resultados teóricos y dar pruebas confiables sobre la exactitud de la nueva clase de métodos MPRKO.

Content

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Scope of this thesis

Many science and engineering problems involve systems of ordinary differential equations, which are given in form of production–destruction systems

$$\frac{dy_i}{dt}(t) = \underbrace{\sum_{j=1}^{N} p_{ij}(t, \mathbf{y}(t))}_{=P_i(t,\mathbf{y}(t))} - \underbrace{\sum_{j=1}^{N} d_{ij}(t, \mathbf{y}(t))}_{=D_i(t,\mathbf{y}(t))}, \tag{1.1}$$

where $\mathbf{y}(t) = (y_1(t), \ldots, y_N(t))^T$ represents the solution vector, and the production terms p_{ij} and destruction terms d_{ij} are nonnegative for $t > 0$ and $\mathbf{y} > \mathbf{0}$, $i, j = 1, \ldots, N$. In the following, we refer to production–destruction systems as PDS. The term $p_{ij} \geq 0$ is the rate at which the jth component transforms into the ith component, while $d_{ij} \geq 0$ is the rate at which the ith component transforms into the jth component. Most of the PDS correspond to concentrations, which need the non negativity of the solution. We will use the following definitions.

Definition 1.1. *The PDS (1.1) is called* ***positive*** *if for $i = 1, \ldots, N$, positive initial values $y_i(0) > 0$ imply positive solutions $y_i(t) > 0$ for all times $t > 0$. The PDS (1.1) is called* ***conservative*** *if for all $i, j = 1, \ldots, N$ and $t \geq 0, \mathbf{y} > \mathbf{0}$, we get $p_{ij}(t, \mathbf{y}) = d_{ji}(t, \mathbf{y})$. In addition, the system is called* ***fully conservative*** *if we have $p_{ii}(t, \mathbf{y}) = d_{ii}(t, \mathbf{y}) = 0$ for all $\mathbf{y} > \mathbf{0}$ and $i = 1, \ldots, N$.*

If a PDS is positive and conservative, it is often essential to keep numerically properties, as negative approximations may lead to meaningless numerical solutions [3, 25, 32, 37] and disregarding conservation might cause accumulation of large errors over time [4, 36]. In these cases, numerical schemes are unconditionally positive and conservative in the following sense.

Definition 1.2. *Let* $\mathbf{y}^n > \mathbf{0}$ *denote an approximation of* $\mathbf{y}(t^n)$ *at time* t^n*, the scheme defined by*

$$\mathbf{y}^{n+1} = \mathbf{y}^n + \Delta t \Phi(\Delta t, t^n, \mathbf{y}^n, \mathbf{y}^{n+1})$$

is called ***unconditionally positive*** *if it guarantees* $\mathbf{y}^{n+1} > \mathbf{0}$ *for all* $\Delta t > 0$ *and* $\mathbf{y}^n > \mathbf{0}$*, and* ***unconditionally conservative*** *if*

$$\sum_{i=1}^{N} (y_i^{n+1} - y_i^n) = 0$$

for all $n \in \mathbb{N}$ *and* $\Delta t > 0$*.*

The standard time integration schemes to numerically solve are Runge–Kutta (RK), Rosenbrock, or multistep methods [7, 12, 13]. According to [12], an explicit s-stage RK method applied to (1.1) has the form

$$y_i^{(k)} = y_i^n + \Delta t \sum_{v=1}^{k-1} a_{kv} \left(P_i(t^n + c_v \Delta t, \mathbf{y}^{(v)}) - D_i(t^n + c_v \Delta t, \mathbf{y}^{(v)}) \right), \tag{1.2a}$$

$$y_i^{n+1} = y_i^n + \Delta t \sum_{k=1}^{s} b_k \left(P_i(t^n + c_k \Delta t, \mathbf{y}^{(k)}) - D_i(t^n + c_k \Delta t, \mathbf{y}^{(k)}) \right). \tag{1.2b}$$

Usually the parameters c_i satisfy the conditions

$$c_i = \sum_{j=1}^{i-1} a_{ij}, \quad i = 1, \ldots, s, \tag{1.3}$$

which particularly implies $c_1 = 0$. Oliver noticed in [30] that the conditions (1.3) are unnecessary for convergence and he constructed second and third order schemes with $c_1 \neq 0$. Recently, Oliver's work was continued in [38] by introducing a general formula for the order of such methods together with a 6-stage scheme of order 5.
For the general equation $y'(t) = f(t, y)$, at the first stage, $k_1 = f(t_0, y_0)$ is always used. Oliver [30], and recently Tsitouras [38], examined the case where $k_1 = f(t_0 + c_1 \Delta t, x_0)$ with $c_1 \neq 0$. As in [38], we refer to those methods of the form (1.2) that to do not necessarily satisfy conditions (1.3) as Runge–Kutta–Oliver (RKO) schemes and to those that do as RK methods. Clearly, RK schemes are contained in the larger class of RKO schemes.

Classical schemes usually generate conservative approximations, but cannot ensure unconditional positivity. Figure 1.1 show an example where the $RK22(\alpha)$ schemes produce negatives approximations with $\alpha = \frac{1}{2}$, when applied to solve the nonlinear test problem (5.6) and clearly showing the effect of the nonpositivity of this scheme.

In particular, it is shown in [2, 20] that no unconditionally positive RK or linear multistep method of order $p \geq 2$ exists. Hence, high order RK methods need to

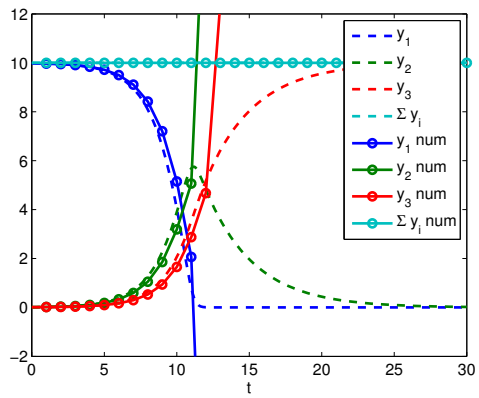


Figure 1.1: Negative solutions of numerical approximations of the nonlinear test problem (5.6) computed with RK($\frac{1}{2}$) scheme with $\Delta t = 1$.

Table 1.1: Barriers in the order [38]

No. of stages	1	2	3	4	5	6	7	8	9
Max. attained order for RK	1	2	3	4	4	5	6	6	7
Max. attained order for RKO	1	2	3	3	4	5	?	?	?

restrict the time step size to ensure positivity [1, 16, 17, 20]. If positivity of a forward Euler step can be guaranteed under some restriction on the time step size, then this is also true for high order strong stability preserving (SSP) RK schemes [10]. Among the Rosenbrock methods, there are some schemes which favor positivity [21, 33], but unconditional positivity cannot be guaranteed. Modified Patankar–Runge–Kutta (MPRK) schemes [4, 18, 19, 23, 24, 27, 28] adapt explicit Runge–Kutta methods such that they are positive irrespective of the time step size Δt, while maintaining their inherent property of being conservative. The key idea behind these methods is *Patankar's trick* [31], which is to multiply the destruction terms with weights making the scheme linearly implicit. As this procedure destroys conservation, the production terms must be weighted accordingly as well. MPRK schemes have been successfully employed for a large number of different applications [5, 6, 11, 14, 15, 22, 34, 40] and their success is particularly based on the fact that they are able to solve stiff PDS. All MPRK methods cited above can be used to integrate the Robertson test [13] with only a few steps.

In this tesis, we will extend the work of [23] to nonautonomous systems. Furthermore, we consider the more general class of RKO [30, 38] methods as underlying base scheme. Tsitouras [38,39] derived the arbitrary order conditions mentioning that there is no fourth order four stage scheme (see Table 1.1). Moreover, he constructed a 5th order method at the cost of six stages per step. This method outperforms other classical Runge–Kutta pairs with orders 5(4) when applied to problems with singu-

Table 1.2: Number of conditions for order p [38]

Order p	1	2	3	4	5	6	7	8	9
No. of conditions for RK	1	1	2	4	9	20	48	115	286
No. of conditions for RKO	1	2	5	13	37	108	332	1042	3360

larity at the beginning. In Table 1.1, we summarize the barriers with the maximum attained order with respect to the stage number. In Table 1.2, we show the order conditions [38]. For achieving a fourth order RKO we must satisfy 1 + 2 + 5 + 13 = 21 order conditions. For RK method of the same order, only 1 + 1 + 2 + 4 = 8 condition are required. This is a serious drawback for high order RKO.

1.2 Organization of this thesis

The thesis is organized as follows: the presentation of the work is described in this first introductory Chapter. Then, in Chapter 2, we introduce unconditionally positive and conservative modified Patankar Runge–Kutta (MPRK) schemes and we present necessary and sufficient conditions on the PWDs to get a second and third order accurate scheme. We present the one-parameter family of MPRK22(α) schemes, which are second order accurate two-stage MPRK schemes. Also, we present two third-order schemes, a two-parameter family of MPRK43I(α, β) and an one-parameter family of MPRK43I(γ) schemes.
Following the general analysis of MPRK schemes described in [24], positivity and mass conservation fundamental properties are proven and even conditions concerning the Patankar weights are given to get second order accuracy of the MPRKO methods and improve the accuracy of these schemes in the field of nonautonomous systems.

The results presented in Chapter 3 led to the following publication

- A. Ávila, G. González, S. Kopecz, and A. Meister, Extension of modified Patankar–Runge–Kutta schemes to nonautonomous production–destruction systems based on Oliver's approach, Jour. Compu. Appl. Math., 389(2020).

Then, we extend our previous work in [26] and develop a third-order unconditionally positivity preserving MPRKO method and the necessary and sufficient conditions for the method are presented in Chapter 4, which are presented in the following paper:

- A. Ávila, G. González, A Third-Order unconditionally positive and conservative modified Patankar Runge-Kutta schemes based on Oliver's approach, in the process of submit

Chapter 5 will show the numerical results of applying the novel MPRKO scheme to linear and nonlinear models as well as stiff Robertson test to confirm the theoretical results and to give reliable statements about the accuracy of the new class of

MPRKO methods. This chapter describes the nonautonomous and autonomous test problems used for numerical experiments, which include parameter studies to examine the dependence of the MPRKO22(α, β) schemes on α and β as well as numerical verification of the second order accuracy and also examine the MPRKO43(β) schemes to verification of the third order accuracy.

Finally in Chapter 6 we present conclusions, observations, and guidelines for future works.

Chapter 2

Modified Patankar Runge Kutta Schemes

Classical Runge–Kutta schemes usually generate conservative approximations, but cannot ensure unconditional positivity. Thus to obtain scientific reasonably results, a scheme must guarantee the positivity for all components and does not generate nor destruct matter.
The scheme (1.2) with two stages is denoted $RK(\alpha)$ and defined as

$$y_i^{(1)} = y_i^n, \tag{2.1a}$$

$$y_i^{(2)} = y_i^n + \alpha \Delta t \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{(1)}) - d_{ij}(\mathbf{y}^{(1)}) \right), \tag{2.1b}$$

$$\begin{aligned} y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[& \left((1 - \frac{1}{2\alpha}) p_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} p_{ij}(\mathbf{y}^{(2)}) \right) \\ & - \left((1 - \frac{1}{2\alpha}) d_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} d_{ij}(\mathbf{y}^{(2)}) \right) \Bigg] \end{aligned} \tag{2.1c}$$

for $i = 1, \ldots, N$.
To achieve a positive scheme, one can modify the original Runge Kutta scheme by the Patankar-trick [31]. This method consists of weighting the destruction term by the quotient of the corresponding constituent at the consecutive time steps. In particular, this procedure yields to the scheme

$$y_i^{(1)} = y_i^n, \tag{2.2a}$$

$$y_i^{(2)} = y_i^n + \alpha\Delta t \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{(1)}) - d_{ij}(\mathbf{y}^{(1)}) \frac{y_i^{(2)}}{y_i^{(n)}} \right), \tag{2.2b}$$

$$\begin{aligned} y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[& \left((1 - \frac{1}{2\alpha}) p_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} p_{ij}(\mathbf{y}^{(2)}) \right) \\ & - \left((1 - \frac{1}{2\alpha}) d_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} d_{ij}(\mathbf{y}^{(2)}) \right) \frac{y_i^{(n+1)}}{y_i^{(2)}} \Bigg] \end{aligned} \tag{2.2c}$$

for $i = 1, \ldots, N$.
In the following, we will refer to this family of schemes as PRK22(α) schemes. This scheme can be explicitly written as

$$y_i^{(1)} = y_i^n, \tag{2.3a}$$

$$y_i^{(2)} = \frac{y_i^n + \alpha\Delta t \sum_{j=1}^{N} p_{ij}(\mathbf{y}^{(1)})}{1 + \alpha\Delta t \sum_{j=1}^{N} d_{ij}(\mathbf{y}^{(1)}) \frac{1}{y_i^n}}, \tag{2.3b}$$

$$y_i^{(n+1)} = \frac{y_i^n + \Delta t \sum_{j=1}^{N} \left((1 - \frac{1}{2\alpha}) p_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} p_{ij}(\mathbf{y}^{(2)}) \right)}{1 + \Delta t \sum_{j=1}^{N} \left((1 - \frac{1}{2\alpha}) d_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} d_{ij}(\mathbf{y}^{(2)}) \right) \frac{1}{y_i^{(2)}}}, \tag{2.3c}$$

for $i = 1, \ldots, N$.
We immediately notice the positivity of the method if $\alpha \geq \frac{1}{2}$. However, we see from the numerical results depicted in Figure 2.1 that the scheme is not conservative when applied to solve the nonlinear test problem (5.6).
To overcome this serious disadvantage. Burchard et al. also weighted production terms and they built a scheme ensuring conservation and positivity. These schemes were named Modified Patankar Runge Kutta and will be described in the following section.

2.1 About modified Patankar–Runge–Kutta schemes

In [23] the following definition of MPRK schemes for autonomous PDS is given.

Definition 2.1. *A **modified Patankar–Runge–Kutta (MPRK)** scheme with s stages is defined as*

$$y_i^{(k)} = y_i^n + \Delta t \sum_{v=1}^{k-1} a_{kv} \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{(v)}) \frac{y_j^{(k)}}{\pi_j^{(k)}} - d_{ij}(\mathbf{y}^{(v)}) \frac{y_i^{(k)}}{\pi_i^{(k)}} \right), \tag{2.4a}$$

$$y_i^{n+1} = y_i^n + \Delta t \sum_{k=1}^{s} b_k \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{(k)}) \frac{y_j^{n+1}}{\sigma_j} - d_{ij}(\mathbf{y}^{(k)}) \frac{y_i^{n+1}}{\sigma_i} \right), \tag{2.4b}$$

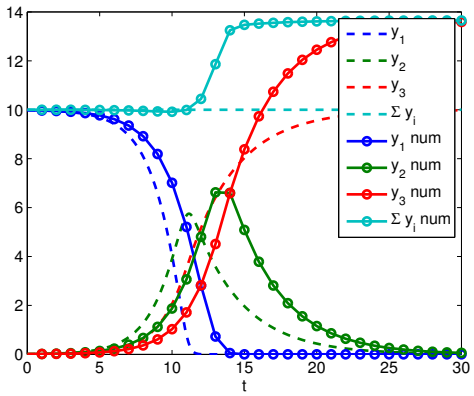


Figure 2.1: Negative solutions of numerical approximations of the nonlinear test problem (5.6) computed with PRK($\frac{1}{2}$) scheme.

for $i = 1, \ldots, N$ *with nonnegative coefficients* $a_{kv}, b_k \geq 0$ *for* $k = 1, \ldots, s$, $v = 1, \ldots, k-1$. *The denominators* σ_i *and* $\pi_i^{(k)}$ *are called* ***Patankar–weight denominators*** *(PWD) and must satisfy the conditions*

1. $\pi_i^{(k)}$ *and* σ_i *are unconditionally positive for* $k = 1, \ldots, s$ *and* $i = 1, \ldots, N$.
2. $\pi_i^{(k)}$ *is independent of* $y_i^{(k)}$ *and* σ_i *is independent of* y_i^{n+1} *for* $k = 1, \ldots, s$ *and* $i = 1, \ldots, N$.

To illustrate this definition we give the following remarks.

Remark 2.1.1. *The condition concerning the independence of PWDs enforces a linearly implicit scheme, in which the solution of* s *linear systems of size* $N \times N$ *is required in each time step.*

Remark 2.1.2. *The PWDs* σ_i *and* $\pi_i^{(k)}$ *are not constant during time integration. In all the schemes cited in the introduction. They depend on the previous stage values, i. e.*

$$\sigma_i = \sigma_i\big(y_i^{(1)}, \ldots, y_i^{(s)}\big), \quad \pi_i^{(k)} = \pi_i^{(k)}\big(y_i^{(1)}, \ldots, y_i^{(k-1)}\big).$$

Remark 2.1.3. *MPRK schemes as defined above require nonnegative parameters* a_{kv}, b_k, *but MPRK schemes with negative parameters can also be developed by treating production terms as destruction terms and vice versa as was done in [28].*

Remark 2.1.4. *The definition in [23] includes an additional parameter* $\delta \in \{0, 1\}$. *Definition 2.1 corresponds to the case* $\delta = 1$, *proven to be superior in numerical experiments.*

Kopecz and Meister [23] showed that s linear systems of size $N \times N$ need to be solved to obtain the stage values and the approximation at the next time level. Considering that $p_{ii} = d_{ii} = 0$ for $i = 1, ..., N$, the scheme (2.4) can be written in matrix-vector notation as

$$\mathbf{M}^{(k)}\mathbf{y}^{(k)} = \mathbf{y}^n, k = 1, ..., s, \tag{2.5}$$

$$\mathbf{M}\mathbf{y}^{n+1} = \mathbf{y}^n, \tag{2.6}$$

with $\mathbf{P}(\mathbf{y}^n) = (P_1(\mathbf{y}^n), ..., P_N(\mathbf{y}^n))^T$ and

$$m_{ii}^{(k)} = 1 + \Delta t \sum_{v=1}^{k-1} a_{kv} \sum_{j=1}^{N} d_{ij}(\mathbf{x}^{(v)}) \frac{1}{\pi_i^{(k)}} > 0, i = 1, ..., N, \tag{2.7a}$$

$$m_{ij}^{(k)} = -\Delta t \delta \sum_{v=1}^{k-1} a_{kv} p_{ij}(\mathbf{x}^{(v)}) \frac{1}{\pi_j^{(k)}} \leq 0, i, j = 1, ..., N, i \neq j, \tag{2.7b}$$

for $k = 1, ..., s$, and

$$m_{ii} = 1 + \Delta t \sum_{k=1}^{s} b_k \sum_{j=1}^{N} d_{ij}(\mathbf{x}^{(v)}) \frac{1}{\sigma_i} > 0, i = 1, ..., N, \tag{2.8a}$$

$$m_{ij} = -\Delta t \sum_{k=1}^{s} b_k p_{ij}(\mathbf{x}^{(k)}) \frac{1}{\sigma_j} \leq 0, i, j = 1, ..., N, i \neq j. \tag{2.8b}$$

The following two lemmas of [23] ensure that the MPRK scheme (2.4) are indeed unconditionally positive and conservative for the case $\delta = 1$.

Lemma 2.1. *A MPRK scheme (2.4) applied to a conservative PDS is unconditionally conservative, that is* $\sum_{i=1}^{N}(y_i^{(k)} - y_i^n) = 0$ *for* $k = 1, \ldots, s$.

Lemma 2.2. *A MPRK scheme (2.4) is unconditionally positive. The same holds for all the stages of the scheme, that is, for all* $\Delta t > 0$ *and* $y^n > 0$, *we have* $y^{(k)} > 0$ *for* $k = 1, \ldots, s$.

2.2 Second order two-stage Modified Patankar Runge–Kutta schemes

In [23], the MPRK22(α) schemes were introduced. This one-parameter family of second order two-stage MPRK schemes is given by the Butcher tableau

$$\begin{array}{c|cc} 0 & & \\ \alpha & \alpha & \\ \hline & 1 - \frac{1}{2\alpha} & \frac{1}{2\alpha} \end{array}$$

where $\alpha \geq 1/2$ together with the PWDs

$$\pi_i = y_i^n, \quad \sigma_i = (y_i^n)^{1-1/\alpha}(y_i^{(2)})^{1/\alpha}, \qquad i = 1, \ldots, N.$$

Thereby, the choice $\alpha \geq \frac{1}{2}$ ensures that all parameters in the Butcher Tableau are nonnegative. So far, MPRK schemes have only been studied for autonomous PDS and defined as follows to the case $\delta = 1$:

$$y_i^{(1)} = y_i^n, \tag{2.9a}$$

$$y_i^{(2)} = y_i^n + \alpha \Delta t \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{(1)}) \frac{y_j^{(2)}}{y_j^{(1)}} - d_{ij}(\mathbf{y}^{(1)}) \frac{y_i^{(2)}}{y_i^{(1)}} \right), \tag{2.9b}$$

$$\begin{aligned} y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[& \left((1 - \frac{1}{2\alpha}) p_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} p_{ij}(\mathbf{y}^{(2)}) \right) \frac{y_j^{n+1}}{(y_j^{(2)})^{\frac{1}{\alpha}} (y_j^n)^{1-\frac{1}{\alpha}}} \\ & - \left((1 - \frac{1}{2\alpha}) d_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} d_{ij}(\mathbf{y}^{(2)}) \right) \frac{y_i^{n+1}}{(y_i^{(2)})^{\frac{1}{\alpha}} (y_i^n)^{1-\frac{1}{\alpha}}} \Bigg] \end{aligned} \tag{2.9c}$$

for $i = 1, \ldots, N$.
The MPRK22(α) scheme, with $\alpha \geq \frac{1}{2}$, have been successfully applied to solve physical, biogeochemical, and ecosystem models. They also proved the capability of the MPRK22 schemes to integrate stiff PDS like the Robertson problem. The MPRK22(1) scheme is equivalent to the original MPRK scheme introduced in [4].

The PWDs of the scheme 2.9 are not the only possible choices. In particular, we can use convex combinations of PWDs as

$$\pi_i = y_i^n, \quad \sigma_i = \omega y_i^n \left(\frac{y_i^{(2)}}{y_i^n} \right)^{s_1} + (1-\omega) y_i^n \left(\frac{y_i^{(2)}}{y_i^n} \right)^{s_2}, \qquad i = 1, \ldots, N.$$

with $0 \leq \omega \leq 1$, and $s_2 = \frac{\alpha\omega s_1 - 1}{\alpha(\omega - 1)}$.

.

2.3 Third-order four stage Modified Patankar Runge–Kutta schemes

In [24], a one-parameter family of third order MPRK schemes given by the Butcher tableau

$$\begin{array}{c|cccc} 0 & & & & \\ \frac{2}{3} & \frac{2}{3} & & & \\ \frac{2}{3} & \frac{2}{3} - \frac{1}{4\gamma} & \frac{1}{4\gamma} & & \\ \hline & \frac{1}{4} & \frac{3}{4} - \gamma & \gamma & \end{array}$$

with

$$\frac{3}{8} \leq \gamma \leq \frac{3}{4},$$

and a two-parameter family

$$\begin{array}{c|ccc}
0 & & & \\
\alpha & \alpha & & \\
\beta & \frac{3\alpha\beta(1-\alpha)-\beta^2}{\alpha(2-3\alpha)} & \frac{\beta(\beta-\alpha)}{\alpha(2-3\alpha)} & \\
\hline
 & 1+\frac{2-3(\alpha+\beta)}{6\alpha\beta} & \frac{3\beta-2}{6\alpha(\beta-\alpha)} & \frac{2-3\alpha}{6\alpha(\beta-\alpha)}
\end{array}$$

with

$$\left.\begin{array}{c} 2/3 \leq \beta \leq 3\alpha(1-\alpha) \\ 3\alpha(1-\alpha) \leq \beta \leq 2/3 \\ (3\alpha-2)/(6\alpha-3) \leq \beta \leq 2/3 \end{array}\right\} \quad for \quad \left\{\begin{array}{c} 2/3 \leq \alpha \leq \frac{2}{3}, \\ 2/3 \leq \alpha \leq \alpha_0, \\ \alpha > \alpha_0, \end{array}\right.$$

where $\alpha_0 = \frac{1}{6}(3+(3-2\sqrt{2})^{1/3}+(3+2\sqrt{2})^{1/3}) \approx 0.89255$, were introduced and called MPRK43I(γ), MPRK43(α, β) respectively.
They derived necessary and sufficient conditions for third order MPRK schemes and introduce the first family of such schemes defined as follows when $\delta = 1$:

$$y_i^{(1)} = y_i^n, \tag{2.10a}$$

$$y_i^{(2)} = y_i^n + a_{21}\Delta t \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{(1)}) \frac{y_j^{(2)}}{y_j^{(1)}} - d_{ij}(\mathbf{y}^{(1)}) \frac{y_i^{(2)}}{y_i^{(1)}} \right), \tag{2.10b}$$

$$\begin{aligned} y_i^{(3)} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[& \left(a_{31}p_{ij}(\mathbf{y}^{(1)}) + a_{32}p_{ij}(\mathbf{y}^{(2)})\right) \frac{y_j^{(3)}}{(y_j^{(2)})^{\frac{1}{p}}(y_j^n)^{1-\frac{1}{p}}} \\ & - \left(a_{31}d_{ij}(\mathbf{y}^{(1)}) + a_{32}d_{ij}(\mathbf{y}^{(2)})\right) \frac{y_i^{(3)}}{(y_i^{(2)})^{\frac{1}{p}}(y_i^n)^{1-\frac{1}{p}}} \Bigg] \end{aligned} \tag{2.10c}$$

$$\begin{aligned} \sigma_i = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[& \left(\beta_1 p_{ij}(\mathbf{y}^{(1)}) + \beta_2 p_{ij}(\mathbf{y}^{(2)})\right) \frac{\sigma_j}{(y_j^{(2)})^{\frac{1}{q}}(y_j^n)^{1-\frac{1}{q}}} \\ & - \left(\beta_1 d_{ij}(\mathbf{y}^{(1)}) + \beta_2 d_{ij}(\mathbf{y}^{(2)})\right) \frac{\sigma_i}{(y_i^{(2)})^{\frac{1}{q}}(y_i^n)^{1-\frac{1}{q}}} \Bigg] \end{aligned} \tag{2.10d}$$

$$\begin{aligned} y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[& \left(b_1 p_{ij}(\mathbf{y}^{(1)}) + b_2 p_{ij}(\mathbf{y}^{(2)}) + b_3 p_{ij}(\mathbf{y}^{(3)})\right) \frac{y_j^{n+1}}{\sigma_j} \\ & - \left(b_1 d_{ij}(\mathbf{y}^{(1)}) + b_2 d_{ij}(\mathbf{y}^{(2)}) + b_3 d_{ij}(\mathbf{y}^{(3)})\right) \frac{y_i^{n+1}}{\sigma_i} \Bigg] \end{aligned} \tag{2.10e}$$

with $p = 3a_{21}(a_{31}+a_{32})b_3$, $q = a_{21}$, $\beta_2 = \frac{1}{2a_{21}}$ and $\beta_1 = 1 - \beta_2$ for $i = 1, \ldots, N$.
The MPRK scheme (2.10) can be understood as a four-stage MPRK scheme with corresponding Butcher tableau

$$
\begin{array}{c|cccc}
0 & & & & \\
a_{21} & a_{21} & & & \\
a_{31}+a_{32} & a_{11} & a_{32} & & \\
1 & \beta_1 & \beta_2 & & \\
\hline
 & b_1 & b_2 & b_3 & 0
\end{array}
$$

The extra stage to compute the PWDs σ_i requires of the MPRK22(a_{21}) scheme, thus $a_{21} = \alpha \geq \frac{1}{2}$ for the two parameter family. In [27], the authors proved that it is impossible to construct third-order MPRK schemes with only three stages, when the usual practice, which takes products of powers of previous stage values as PWDs. Three specific MPRK43 schemes were used in [24] for the case $\delta = 1$. The MPRK43I($1, \frac{1}{2}$) scheme is based on the Butcher tableau

$$
\begin{array}{c|ccc}
0 & & & \\
1 & 1 & & \\
\frac{1}{2} & \frac{1}{4} & \frac{1}{4} & \\
\hline
 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{array}
$$

which is based on Heun's method. It is interesting that the above Butcher tableau belongs to the optimal third order strong stability preserving SSP(3,3) scheme introduced in [29]. The method MPRK43I($\frac{1}{2}, \frac{3}{4}$) utilizes the MPRK22($\frac{1}{2}$) scheme, which is adapted from the midpoint method and represented by the Butcher tableau

$$
\begin{array}{c|ccc}
0 & & & \\
\frac{1}{2} & \frac{1}{2} & & \\
\frac{3}{4} & 0 & \frac{3}{4} & \\
\hline
 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{array}
$$

The MPRK43II($\frac{2}{3}$) scheme is associated with the Butcher

$$
\begin{array}{c|ccc}
0 & & & \\
\frac{2}{3} & \frac{2}{3} & & \\
\frac{2}{3} & \frac{1}{6} & \frac{1}{2} & \\
\hline
 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2}
\end{array}
$$

and it employs Ralston's method MPRK22($\frac{2}{3}$) to calculate the PWDs.

2.4 Related work

In the literature there are other schemes of the MPRK type. We first mentioned Juntao Huang and Wang Shu [18], who constructed a family of modified Patankar Runge–Kutta methods using the RK schemes of the Shu–Osher form instead classic RK form.

2.4.1 Modified Patankar Runge–Kutta in the Shu–Osher form

This scheme is conservative and unconditionally positivity for production–destruction equations, and of second-order accuracy.
The scheme is defined as

$$y_i^{(1)} = y_i^n, \tag{2.11a}$$

$$y_i^{(2)} = \alpha_{10} y_i^{(1)} + \beta_{10} \Delta t \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{(1)}) \frac{y_j^{(2)}}{\pi_j} - d_{ij}(\mathbf{y}^{(1)}) \frac{y_i^{(2)}}{\pi_i} \right), \tag{2.11b}$$

$$\begin{aligned} y_i^{n+1} = \alpha_{20} y_i^{(1)} + \alpha_{21} y_i^{(2)} + \Delta t \sum_{j=1}^{N} \Bigg[& \left(\beta_{20} p_{ij}(\mathbf{y}^{(1)}) + \beta_{21} p_{ij}(\mathbf{y}^{(2)}) \right) \frac{y_j^{n+1}}{\sigma_j} \\ & - \left(\beta_{20} d_{ij}(\mathbf{y}^{(1)}) + \beta_{21} d_{ij}(\mathbf{y}^{(2)}) \right) \frac{y_i^{n+1}}{\sigma_i} \Bigg] \end{aligned} \tag{2.11c}$$

for $i = 1, \ldots, N$, with the coefficients satisfying the conditions

$$\alpha_{21} = \alpha, \quad \beta_{10} = \beta, \quad \alpha_{10} = 1, \quad \alpha_{20} = 1 - \alpha, \quad \alpha_{21} = \alpha, \quad \beta_{10} = \beta, \quad \beta_{20} = 1 - \frac{1}{2\beta}, \tag{2.12}$$

$$\beta_{21} = \frac{1}{2\beta}, \quad s = \frac{1 - \alpha\beta + \alpha\beta^2}{\beta(1 - \alpha\beta)}, \quad 0 \le \alpha \le 1, \quad \beta > 0, \quad \alpha\beta + \frac{1}{2\beta} \le 1 \tag{2.13}$$

and PWD

$$\pi_i = y_i^n, \quad \sigma i = \left(y_i^{(2)} \right)^s \left(y_i^n \right)^{1-s}.$$

This scheme was successfully applied to solve non-stiff and stiff problems of ODEs. In addition, the solver was extended to solve a class of semi-discrete schemes for PDEs.

Remark 2.4.1. *The scheme (2.11) is a generalization of the schemes in [4, 23]. Taking $\alpha = 0$ in (2.12,2.13), it reduces to the scheme in [23]. Furthermore, if we set $\beta = 1$, it reduces to the scheme in [4].*

Remark 2.4.2. *If $\alpha = \frac{1}{2}$ and $\beta = 1$, the coefficients of the optimal Strong Stability Preserving Runge Kutta method are recovered:*

$$\alpha_{10} = \beta_{10} = 1, \quad \alpha_{20} = \alpha_{21} = \beta_{21} = \frac{1}{2}, \quad \beta_{20} = 0 \tag{2.14}$$

and accordingly $s = 2$.

Then, they extended their previous work and developed a third-order unconditionally positivity preserving modified Patankar Runge–Kutta scheme. The necessary and sufficient conditions for the methods to be third-order accurate are proved. A variety of numerical examples were conducted to validate the performance of the time integration method. The third order scheme is defined by

$$y_i^{(1)} = y_i^n, \tag{2.15a}$$

$$y_i^{(2)} = \alpha_{10} y_i^{(1)} + \beta_{10} \Delta t \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{(1)}) \frac{y_j^{(2)}}{\pi_j} - d_{ij}(\mathbf{y}^{(1)}) \frac{y_i^{(2)}}{\pi_i} \right), \tag{2.15b}$$

$$y_i^{(3)} = \alpha_{20} y_i^{(1)} + \alpha_{21} y_i^{(2)} + \Delta t \sum_{j=1}^{N} \Bigg[\left(\beta_{20} p_{ij}(\mathbf{y}^{(1)}) + \beta_{21} p_{ij}(\mathbf{y}^{(2)}) \right) \frac{y_j^{n+1}}{\rho_j} - \left(\beta_{20} d_{ij}(\mathbf{y}^{(1)}) + \beta_{21} d_{ij}(\mathbf{y}^{(2)}) \right) \frac{y_i^{n+1}}{\rho_i} \Bigg] \tag{2.15c}$$

$$y_i^{n+1} = \alpha_{30} y_i^{(1)} + \alpha_{31} y_i^{(2)} + \alpha_{32} y_i^{(3)} + \Delta t \sum_{j=1}^{N} \Bigg[\left(\beta_{30} p_{ij}(\mathbf{y}^{(1)}) + \beta_{31} p_{ij}(\mathbf{y}^{(2)}) + \beta_{32} p_{ij}(\mathbf{y}^{(3)}) \right) \frac{y_j^{n+1}}{\sigma_j} - \left(\beta_{30} d_{ij}(\mathbf{y}^{(1)}) + \beta_{31} d_{ij}(\mathbf{y}^{(2)}) + \beta_{32} d_{ij}(\mathbf{y}^{(3)}) \right) \frac{y_i^{n+1}}{\sigma_i} \Bigg] \tag{2.15d}$$

for $i = 1, \ldots, N$, with the coefficients satisfying the conditions

$$\alpha_{10} = \beta_{10} = 1, \quad \alpha_{20} = \frac{3}{4}, \quad \alpha_{21} = \beta_{21} = \frac{1}{4}, \quad \beta_{20} = 0, \quad \alpha_{30} = \frac{1}{3}, \quad \alpha_{32} = \beta_{32} = \frac{2}{3},$$
$$\alpha_{31} = \beta_{30} = \beta_{31} = 0,$$

and PWD satisfying

$$\pi_i = y_i^n, \quad \frac{\sigma i}{y_i^n} = 4 \left(\frac{y_i^{(3)}}{y_i^n} - \frac{1}{2} \right)^2 \left(\frac{y_i^{(3)}}{y_i^n} - \frac{3}{2} \right)^2 + \frac{1}{2} \left(\frac{y_i^{(3)}}{y_i^n} \right)^4 + \frac{1}{4} \frac{c_i^{(2)}}{y_i^n} \frac{y_i^{(3)}}{\rho_i},$$

$$\rho_i = \frac{1}{2} y_i^n \left(\frac{y_i^{(2)}}{y_i^n} + \left(\frac{y_i^{(1)}}{y_i^n} \right)^2 \right), \text{or } y_i^n \left(\frac{1}{4} + \frac{3}{4} \left(\frac{y_i^{(2)}}{y_i^n} \right)^2 \right)$$

Another scheme of the MPRK type was introduced by Öffner et al. and named Modified Patankar Deferred Correction scheme. In the section next we will describe this scheme.

2.4.2 Modified Patankar Deferred Correction (MPDeC) scheme

The Deferred Correction (DeC) method is based on the Picard-Lindelöf theorem in the continuous setting. The DeC is an explicit, arbitrary high order method for ODEs. Further extensions of DeC can be found in the literature, including semi-implicit approaches. In [28], the authors proposed a method to solve PDS problems,

using the explicit Deferred Correction (DeC) process as a time integration method instead RK classic. Applying the modified Patankar approach to the DeC scheme results in provable conservative and positivity preserving methods. Furthermore, they demonstrated that these modified Patankar DeC schemes can be constructed up to arbitrarily high order. Finally, they validated their theoretical analysis through numerical simulations.
The MPDeC scheme is defined as

$$y_i^{m,(0)} = y_i(t^n), \tag{2.16a}$$

$$y_i^{m,(k)} = y_i^{(0)} + \sum_{r=0}^{M} \theta_r^m \Delta t \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{r,(k-1)}) \frac{y^{m,(k)}_{\gamma(j,i,\theta_r^m)}}{y^{m,(k-1)}_{\gamma(j,i,\theta_r^m)}} - d_{ij}(\mathbf{y}^{r,(k-1)}) \frac{y^{m,(k)}_{\gamma(i,j,\theta_r^m)}}{y^{m,(k-1)}_{\gamma(i,j,\theta_r^m)}} \right) \tag{2.16b}$$

for $k = 1, \dots, K$, $m = 1, \dots, M$, and $i = 1, \dots, N$, where $\gamma(a, b, \theta) = a$ if $\theta > 0$ and $\gamma(a, b, \theta) = b$ if $\theta < 0$.

Remark 2.4.3. *The modification of the scheme is done only through the coefficients* $\frac{y_j^{m,(k)}}{y_j^{m,(k-1)}}$ *on both the production and the destruction terms. These coefficients allow to choose each term* $\theta_r^m p_{i,j}$ *and* $\theta_r^m d_{i,j}$*, according to the sign of the* θ *coefficient. The index* γ *takes care of the sign of the destruction and production terms, when negative entries in the Butcher Tableau of the RK scheme appear interchanging the destruction terms with the production ones to guarantee the positivity preserving property.*

Öffner and Torlo proved that the proposed scheme is unconditionally conservative and positivity preserving.

Chapter 3

Second order MPRKO scheme

The goal of this chapter is the construction of unconditionally positive and conservative second order methods for the solution of positive and conservative nonautonomous PDS.

An explicit two-stage RKO scheme is given by the Butcher tableau

$$\begin{array}{c|cc} c_1 & & \\ c_2 & a_{21} & \\ \hline & b_1 & b_2 \end{array}$$

and it is second order accurate if and only if the three conditions

$$b_1 + b_2 = 1, \qquad b_1 c_1 + b_2 c_2 = \frac{1}{2}, \qquad b_2 a_{21} = \frac{1}{2}$$

are satisfied, see [30, 38]. Compared to RK schemes, there is one additional condition due to the possibility to choose $c_1 \neq 0$. All explicit two-stage second order RKO schemes can be parameterized by a family with two free parameters. To extend the notation given for MPRK(α), we will use the following notation: $\beta := c_1$ and $\alpha := a_{21}$. From [30], all explicit two-stage second-order Runge–Kutta–Oliver schemes can be represented by the Butcher tableau

$$\begin{array}{c|cc} \beta & & \\ \alpha - 2\alpha\beta + \beta & \alpha & \\ \hline & 1 - \frac{1}{2\alpha} & \frac{1}{2\alpha} \end{array} \tag{3.1}$$

with two parameters $\alpha \neq 0$ and $\beta \in \mathbb{R}$. For $\beta \neq 0$, the schemes are general RK methods with stages, which are not consistent. For $\beta = 0$, we obtain the usual explict two-stage second order RK schemes. Well known examples of explict two-stage second order RK schemes are Heun's method ($\alpha = 1, \beta = 0$), Ralston's method ($\alpha = 2/3, \beta = 0$), and the midpoint method ($\alpha = 1/2, \beta = 0$).

3.1 Second order MPRKO scheme

Introduced in [23]. The MPRK scheme (2.4) can be generalized to integrate nonautonomous systems in a natural way and to be as general as possible, we allow for the second order two-stage RKO schemes (3.1) as underlying base schemes. Altogether, we obtain the family of schemes in natural way.

$$y_i^{(1)} = y_i^n, \tag{3.2a}$$

$$y_i^{(2)} = y_i^n + \alpha\Delta t \sum_{j=1}^{N} \left(p_{ij}\big(t^n + \beta\Delta t, \mathbf{y}^{(1)}\big) \frac{y_j^{(2)}}{\pi_j} - d_{ij}\big(t^n + \beta\Delta t, \mathbf{y}^{(1)}\big) \frac{y_i^{(2)}}{\pi_i} \right), \tag{3.2b}$$

$$\begin{aligned} y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[& \left(\left(1 - \frac{1}{2\alpha}\right) p_{ij}\big(t^n + \beta\Delta t, \mathbf{y}^{(1)}\big) + \frac{1}{2\alpha} p_{ij}\big(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}\big) \right) \frac{y_j^{n+1}}{\sigma_j} \\ & - \left(\left(1 - \frac{1}{2\alpha}\right) d_{ij}\big(t^n + \beta\Delta t, \mathbf{y}^{(1)}\big) + \frac{1}{2\alpha} d_{ij}\big(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}\big) \right) \frac{y_i^{n+1}}{\sigma_i} \Bigg], \end{aligned} \tag{3.2c}$$

for $i = 1, \ldots, N$ with $\alpha \geq \frac{1}{2}$. We refer to this family of schemes as two-stage MPRKO schemes.

3.1.1 Matrix form of the MPRKO scheme

These schemes can be written in matrix-vector notation as

$$\mathbf{y}^{(1)} = \mathbf{y}^n, \tag{3.3a}$$

$$\mathbf{M}^{(2)}\mathbf{y}^{(2)} = \mathbf{y}^n, \tag{3.3b}$$

$$\mathbf{M}\mathbf{y}^{n+1} = \mathbf{y}^n, \tag{3.3c}$$

with matrix elements

$$m_{ii}^{(2)} = 1 + \alpha\Delta t \sum_{j=1}^{N} d_{ij}\big(t^n + \beta\Delta t, \mathbf{y}^{(1)}\big) \frac{1}{\pi_i} > 0, \tag{3.4a}$$

$$m_{ij}^{(2)} = -\alpha\Delta t p_{ij}\big(t^n + \beta\Delta t, \mathbf{y}^{(1)}\big) \frac{1}{\pi_j} \leq 0 \tag{3.4b}$$

and

$$m_{ii} = 1 + \Delta t \sum_{j=1}^{N} \left(\left(1 - \frac{1}{2\alpha}\right) d_{ij}\big(t^n + \beta\Delta t, \mathbf{y}^{(1)}\big) + \frac{1}{2\alpha} d_{ij}\big(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}\big) \right) \frac{1}{\sigma_i} > 0, \tag{3.5a}$$

$$m_{ij} = -\Delta t \left(\left(1 - \frac{1}{2\alpha}\right) p_{ij}\big(t^n + \beta\Delta t, \mathbf{y}^{(1)}\big) + \frac{\Delta t}{2\alpha} p_{ij}\big(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}\big) \right) \frac{1}{\sigma_j} \leq 0 \tag{3.5b}$$

for $i, j = 1, \ldots, N$, $i \neq j$. The next lemma states that the two-stage MPRKO schemes (3.2) are unconditionally positive and conservative as intended.

Lemma 3.1. *The two-stage MPRKO schemes (3.2) applied to a positive and conservative PDS are unconditionally positive and conservative numerical schemes.*

Proof. The proof follows the proofs of Lemmas 2.7 and 2.8 in [23] using the representation (3.3) with (3.4) and (3.5). □

Remark 3.1.1. *The matrices* $\mathbf{M}$ *and* $\mathbf{M}^{(2)}$ *from* (3.3b) *and* (3.3c) *are M-matrices. Therefore,* $\mathbf{M}$ *and* $\mathbf{M}^{(2)}$ *are non-singular.*

Next, the PWDs π_i and σ_i in (3.2) must guarantee second-order accuracy. We show that the conditions given in [23, Theorem 3.4] are also necessary and sufficient for the nonautonomous case.

3.2 Proof of second-order conditions

Theorem 3.2.1. *The two-stage MPRKO scheme (3.2) is of second order if and only if the conditions*

$$\pi_i = y_i^n + O(\Delta t), \tag{3.6a}$$

$$\sigma_i = y_i^n + \Delta t \left(P_i(t^n, \mathbf{y}^n) - D_i(t^n, \mathbf{y}^n)\right) + O(\Delta t^2), \tag{3.6b}$$

are satisfied for $i = 1, \ldots, N$.

Proof. To prove convergence, we study the local truncation errors and identify y_i^n and $y_i(t^n)$ for $i = 1, \ldots, N$ as usual. Furthermore, we assume $\mathbf{y}^n > \mathbf{0}$, since we are dealing with positive PDS. We use Landau symbol $O(\cdot)$ when $\Delta t \to 0$. To shorten the notation, we use ϕ^* as an abbreviation for $\phi(t^*, y(t^*))$ and omit the index $i = 1, \ldots, N$.
Before deriving the necessary and sufficient third order conditions, we consider a specific class of PDS,

$$\frac{dy_i}{dt}(t) = \widehat{P}_i(t, \mathbf{y}(t)) - \widehat{D}_i(t, \mathbf{y}(t)), \tag{3.7}$$

with initial values $y_i(0) = 1$, for $i = 1, ...N$, and

$$\widehat{p}_{ij}(t, \mathbf{y}) = \begin{cases} \mu y_I^k, & i = J,\ j = I, \\ 0, & \text{otherwise}, \end{cases} \qquad \widehat{d}_{ij}(t, \mathbf{y}) = \begin{cases} \mu y_I^k, & i = I,\ j = J, \\ 0, & \text{otherwise}, \end{cases} \tag{3.8}$$

where $I, J \in \{1, 2, ..., N\}$,$I \neq J$, $\mu > 0$ for all $t > 0$, $k \in \{1, 2\}$.
The PDS (3.7) is written as

$$\frac{dy_I}{dt} = -\mu y_I^k, \quad \frac{dy_J}{dt} = \mu y_I^k, \quad \frac{dy_i}{dt} = 0, \quad i \in \{1, 2, ..., N\}/\{I, J\}.$$

The exact solution for $k = 1$ is

$$y_I(t) = e^{-\mu t}, \quad y_J(t) = 2 - e^{-\mu t}, \quad y_i(t) = 1, \ i \in \{1, 2, ..., N\}/\{I, J\},$$

and for $k = 2$, the exact solution is

$$y_I(t) = \frac{1}{1+\mu t}, \quad y_J(t) = 2 - \frac{1}{1+\mu t}, \quad y_i(t) = 1, \ i \in \{1, 2, ..., N\}/\{I, J\}.$$

Thus, PDS (3.7) is positive and also fully conservative.

3.2.1 Necessary condition

Using (1.1), and denoting P_i^n and D_i^n by $P_i(t^n, \mathbf{y}^n)$ and $D_i(t^n, \mathbf{y}^n)$ respectively, the exact solution at time t^{n+1} can be written as

$$y_i(t^{n+1}) = y_i(t^n) + \Delta t(P_i^n - D_i^n) + \frac{1}{2}\Delta t^2\left(\frac{\partial(P_i^n - D_i^n)}{\partial t} + \frac{\partial(P_i^n - D_i^n)}{\partial \mathbf{y}}(P_i^n - D_i^n)\right) + O(\Delta t^3). \tag{3.9}$$

Since the solver (3.2) is second-order accurate, from (3.2c) and (3.9) we obtain

$$\begin{aligned}
&\Delta t \sum_{i=1}^{N} \Bigg[\left(\left(1 - \frac{1}{2\alpha}\right) p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + \frac{1}{2\alpha} p_{ij}(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}) \right) \frac{y_j^{n+1}}{\sigma_j} \\
&- \left(\left(1 - \frac{1}{2\alpha}\right) d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + \frac{1}{2\alpha} d_{ij}(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}) \right) \frac{y_i^{n+1}}{\sigma_i} \Bigg] \\
&- \Delta t\,(P_i^n - D_i^n) - \frac{\Delta t^2}{2}\left[\frac{\partial}{\partial \mathbf{y}}(P_i^n - D_i^n)(P^n - D^n) + \frac{\partial}{\partial t}(P_i^n - D_i^n) \right] = O(\Delta t^3).
\end{aligned} \tag{3.10}$$

Considering a Taylor expansion for $p_{ij}(t, y)$ and $d_{ij}(t, y)$ around point (t^n, y^n), then $p_{ij}(t^n + c_1\Delta t, y^n)$ and $d_{ij}(t^n + c_1\Delta t, y^n)$ can be written as

$$\begin{aligned}
p_{ij}(t^n + c_1\Delta t, \mathbf{y}^{(1)}) &= p_{ij}^n + \beta\Delta t\frac{\partial p_{ij}^n}{\partial t} + (\mathbf{y}^{(1)} - \mathbf{y}^n)\frac{\partial p_{ij}^n}{\partial \mathbf{y}} + O(\Delta t^2), \\
d_{ij}(t^n + c_1\Delta t, \mathbf{y}^{(1)}) &= d_{ij}^n + \beta\Delta t\frac{\partial d_{ij}^n}{\partial t} + (\mathbf{y}^{(1)} - \mathbf{y}^n)\frac{\partial d_{ij}^n}{\partial \mathbf{y}} + O(\Delta t^2).
\end{aligned} \tag{3.11}$$

Similarly,

$$\begin{aligned}
p_{ij}(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}) &= p_{ij}^n + (\alpha - 2\alpha\beta + \beta)\Delta t\frac{\partial p_{ij}^n}{\partial t} + (\mathbf{y}^{(2)} - \mathbf{y}^n)\frac{\partial p_{ij}^n}{\partial \mathbf{y}} + O(\Delta t^2), \\
d_{ij}(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}) &= d_{ij}^n + (\alpha - 2\alpha\beta + \beta)\Delta t\frac{\partial d_{ij}^n}{\partial t} + (\mathbf{y}^{(2)} - \mathbf{y}^n)\frac{\partial d_{ij}^n}{\partial \mathbf{y}} + O(\Delta t^2).
\end{aligned} \tag{3.12}$$

Replacing (3.2a), (3.11), and (3.12) in (3.10), we obtain

$$\Delta t\sum_{i=1}^{N}\left[\left(\left(1-\frac{1}{2\alpha}+\frac{1}{2\alpha}\right)p_{ij}^{n}+\left((1-\frac{1}{2\alpha})\beta+\frac{1}{2\alpha}(\alpha-2\alpha\beta+\beta)\right)\Delta t\frac{\partial p_{ij}^{n}}{\partial t}+\frac{1}{2\alpha}(\mathbf{y}^{(2)}-\mathbf{y}^{n})\frac{\partial p_{ij}^{n}}{\partial \mathbf{y}}\right)\frac{y_j^{n+1}}{\sigma_j}\right. \tag{3.13}$$

$$\left.-\left(\left(1-\frac{1}{2\alpha}+\frac{1}{2\alpha}\right)d_{ij}^{n}+\left((1-\frac{1}{2\alpha})\beta+\frac{1}{2\alpha}(\alpha-2\alpha\beta+\beta)\right)\Delta t\frac{\partial d_{ij}^{n}}{\partial t}+\frac{1}{2\alpha}(\mathbf{y}^{(2)}-\mathbf{y}^{n})\frac{\partial d_{ij}^{n}}{\partial \mathbf{y}}\right)\frac{y_i^{n+1}}{\sigma_i}\right]$$

$$-\Delta t\left(P_i^n-D_i^n\right)-\frac{\Delta t^2}{2}\left[\frac{\partial}{\partial \mathbf{y}}\left(P_i^n-D_i^n\right)\left(P^n-D^n\right)+\frac{\partial}{\partial t}\left(P_i^n-D_i^n\right)\right]=O(\Delta t^3). \tag{3.14}$$

We use PDS (3.7) from now on, with $I,J\in\{1,2,...,N\}$,$I\neq J$, $\mu>0$ for all $t>0$, $k\in\{1,2\}$. Equation (3.13) yields to

$$-\left(\widehat{D}_I^n+\frac{1}{2}\Delta t\frac{\partial\widehat{D}_I^n}{\partial t}+b_2(y_I^{(2)}-y_I^n)\frac{\partial\widehat{D}_I^n}{\partial y_I}\right)\frac{y_I^{n+1}}{\sigma_I}+\widehat{D}_I^n-\frac{\Delta t}{2}\left[\frac{\partial\widehat{D}_I^n}{\partial y_I}(\widehat{D}_I^n)-\frac{\partial\widehat{D}_I^n}{\partial t}\right]=O(\Delta t^2). \tag{3.15}$$

Owing to $\widehat{D}_I^n=\mu y_I^k$, we have $\frac{\partial\widehat{D}_I^n}{\partial y_I}=\mu k y_I^{k-1}$ and $\frac{\partial\widehat{D}_I^n}{\partial t}=0$, replacing into (3.15), it yields to

$$-\left(\mu y_I^k+\frac{1}{2}\Delta t\mu y_I^k+\frac{1}{2\alpha}(y_I^{(2)}-y_I^n)\mu k y_I^{k-1}\right)\frac{y_I^{n+1}}{\sigma_I}+\mu y_I^k-\frac{\Delta t}{2}\mu(t^n)k y_I^{k-1}\mu(t)y_I^k=O(\Delta t^2). \tag{3.16}$$

Owing to (3.2b), we have

$$y_I^{(2)}-y_I^n=\alpha\Delta t(-\widehat{D}_I^n)\frac{y_I^{(2)}}{\pi_I}=-\alpha\Delta t\mu y_I^k\frac{y_I^{(2)}}{\pi_I}. \tag{3.17}$$

Replacing (3.17) into (3.16),we get

$$-\left(\mu y_I^k+\frac{1}{2}\Delta t\mu y_I^k-\frac{1}{2}\Delta t\mu y_I^k\frac{y_I^{(2)}}{\pi_I}\mu k y_I^{k-1}\right)\frac{y_I^{n+1}}{\sigma_I}+\mu y_I^k-\frac{\Delta t}{2}\mu k y_I^{k-1}\mu y_I^k=O(\Delta t^2), \tag{3.18}$$

and subsequently since $y_I^k>0$, we simplify to

$$-\mu\left(\frac{y_I^{n+1}}{\sigma_I}-1\right)+\frac{1}{2}\Delta t\mu^2 k y_I^{k-1}\left(\frac{y_I^{n+1}}{\sigma_I}\frac{y_I^{(2)}}{\pi_I}-1\right)=O(\Delta t^2). \tag{3.19}$$

Since $\mu>0$, we get

$$-\left(\frac{y_I^{n+1}}{\sigma_I}-1\right)+\frac{1}{2}\Delta t\mu k y_I^{k-1}\left(\frac{y_I^{n+1}}{\sigma_I}\frac{y_I^{(2)}}{\pi_I}-1\right)=O(\Delta t^2), \tag{3.20}$$

with μ constant, we obtain the case of Theorem 3.4 of [23]. From Lemma 4 in [24], we conclude that

$$\begin{aligned} \frac{y_I^{n+1}}{\sigma_I} - 1 &= O(\Delta t^2), \\ \frac{y_I^{n+1}}{\sigma_I}\frac{y_I^{(2)}}{\pi_I} - 1 &= O(\Delta t), \end{aligned} \tag{3.21}$$

hold true. Using Theorem 3.1 in [23], we conclude

$$\begin{aligned} \pi_i &= y_i^n + O(\Delta t) \\ \sigma_i &= y_i^n + \Delta t(P_i^n - D_i^n) + O(\Delta t^2). \end{aligned} \tag{3.22}$$

Equation (3.22) shows that (3.6a) and (3.6b) are the necessary conditions.

3.2.2 Sufficient condition

Let $\mathbf{M}^{-1} = (\widetilde{m}_{ij})_{i,j=1,\dots,N}$ and $(\mathbf{M}^{(2)})^{-1} = (\widetilde{m}_{ij}^{(2)})_{i,j=1,\dots,N}$. Then, equations (3.3b) and (3.3c) can be written as

$$y_i^{(2)} = \sum_{j=1}^{N} \widetilde{m}_{ij}^{(2)} y_j^n, \qquad y_i^{n+1} = \sum_{j=1}^{N} \widetilde{m}_{ij} y_j^n.$$

Thus, Lemma 3.1 in [23] guarantees that $0 \le \widetilde{m}_{ij}, \widetilde{m}_{ij}^{(2)} \le 1$ for $i, j = 1, \dots, N$, and

$$\frac{y_i^{(2)}}{\pi_i} = \sum_{j=1}^{N} \widetilde{m}_{ij}^{(2)} \frac{y_j^n}{\pi_i} = O(1), \tag{3.23a}$$

$$\frac{y_i^{n+1}}{\sigma_i} = \sum_{j=1}^{N} \widetilde{m}_{ij} \frac{y_j^n}{\sigma_i} = O(1), \tag{3.23b}$$

since $y_i^n > 0$. Using (3.23a) in (3.2b), we obtain

$$y_i^{(2)} = y_i^n + \alpha\Delta t \sum_{j=1}^{N} \left(\left(p_{ij}^n + O(\Delta t)\right) \frac{y_j^{(2)}}{\pi_j} - (d_{ij}^n + O(\Delta t)) \frac{y_i^{(2)}}{\pi_i} \right) = y_i^n + O(\Delta t). \tag{3.24}$$

Inserting into (3.6a) and (3.2b), we get

$$\begin{aligned} y_i^{(2)} &= y_i^n + \alpha\Delta t \sum_{j=1}^{N} \left(\left(p_{ij}^n + O(\Delta t)\right)(1 + O(\Delta t)) - \left(d_{ij}^n + O(\Delta t)\right)(1 + O(\Delta t)) \right) \\ &= y_i^n + \alpha\Delta t(P_i^n - D_i^n) + O(\Delta t^2). \end{aligned} \tag{3.25}$$

Next, we focus on the expansion of the approximation step (3.2c). Using (3.24) in combination with $\mathbf{y}^{(1)} = \mathbf{y}^n$ and the fact that α, β are fixed numbers, we find

$$y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[\left(\left(1 - \frac{1}{2\alpha}\right)(p_{ij}^n + O(\Delta t)) + \frac{1}{2\alpha}\left(p_{ij}^n + O(\Delta t)\right) \right) \frac{y_j^{n+1}}{\sigma_j} \\ - \left(\left(1 - \frac{1}{2\alpha}\right)(d_{ij}^n + O(\Delta t)) + \frac{1}{2\alpha}(d_{ij}^n + O(\Delta t)) \right) \frac{y_i^{n+1}}{\sigma_i} \Bigg], \quad (3.26)$$

and together with (3.23b) we conclude

$$y_i^{n+1} = y_i^n + O(\Delta t).$$

Replacing in (3.26) while considering (3.6b), we simplify it to

$$\begin{aligned} y_i^{n+1} &= y_i^n + \Delta t \sum_{j=1}^{N} \left((p_{ij}^n + O(\Delta t))(1 + O(\Delta t)) - (d_{ij}^n + O(\Delta t))(1 + O(\Delta t)) \right) \\ &= y_i^n + \Delta t (P_i^n - D_i^n) + O(\Delta t^2) \end{aligned}$$

and hence

$$\frac{y_i^{n+1}}{\sigma_i} = 1 + O(\Delta t^2). \quad (3.27)$$

Due to (3.24), a higher order Taylor series expansion of p_{ij} and d_{ij} in (3.2c) yields

$$\begin{aligned} y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[& \left(1 - \frac{1}{2\alpha}\right) \left(p_{ij}^n + \beta \Delta t \frac{\partial p_{ij}^n}{\partial t} + O(\Delta t^2) \right) \frac{y_j^{n+1}}{\sigma_j} \\ & + \frac{1}{2\alpha} \left(p_{ij}^n + (\alpha - 2\alpha\beta + \beta)\Delta t \frac{\partial p_{ij}^n}{\partial t} + \frac{\partial p_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(2)} - \mathbf{y}^n) + O(\Delta t^2) \right) \frac{y_j^{n+1}}{\sigma_j} \\ & - \left(1 - \frac{1}{2\alpha}\right) \left(d_{ij}^n + \beta \Delta t \frac{\partial d_{ij}^n}{\partial t} + O(\Delta t^2) \right) \frac{y_i^{n+1}}{\sigma_i} \\ & - \frac{1}{2\alpha} \left(d_{ij}^n + (\alpha - 2\alpha\beta + \beta)\Delta t \frac{\partial d_{ij}^n}{\partial t} + \frac{\partial d_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(2)} - \mathbf{y}^n) + O(\Delta t^2) \right) \frac{y_i^{n+1}}{\sigma_i} \Bigg], \end{aligned}$$

which simplify to

$$y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[\left(p_{ij}^n + \frac{\Delta t}{2} \frac{\partial p_{ij}^n}{\partial t} + \frac{1}{2\alpha} \frac{\partial p_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(2)} - \mathbf{y}^n) + O(\Delta t^2) \right) \frac{y_j^{n+1}}{\sigma_j} \\ - \left(d_{ij}^n + \frac{\Delta t}{2} \frac{\partial d_{ij}^n}{\partial t} + \frac{1}{2\alpha} \frac{\partial d_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(2)} - \mathbf{y}^n) + O(\Delta t^2) \right) \frac{y_i^{n+1}}{\sigma_i} \Bigg].$$

Finally, inserting (3.25) and (3.27), we obtain

$$\begin{aligned}
y_i^{n+1} &= y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[\left(p_{ij}^n + \frac{\Delta t}{2}\frac{\partial p_{ij}^n}{\partial t} + \frac{\Delta t}{2}\frac{\partial p_{ij}^n}{\partial \mathbf{y}}(\mathbf{P}^n - \mathbf{D}^n) + O(\Delta t^2) \right)(1 + O(\Delta t^2)) \\
&\qquad - \left(d_{ij}^n + \frac{\Delta t}{2}\frac{\partial d_{ij}^n}{\partial t} + \frac{\Delta t}{2}\frac{\partial d_{ij}^n}{\partial \mathbf{y}}(\mathbf{P}^n - \mathbf{D}^n) + O(\Delta t^2) \right)(1 + O(\Delta t^2)) \Bigg] \\
&= y_i^n + \Delta t(P_i^n - D_i^n) + \frac{\Delta t^2}{2}\left(\frac{\partial (P_i^n - D_i^n)}{\partial t} + \frac{\partial (P_i^n - D_i^n)}{\partial \mathbf{y}}(\mathbf{P}^n - \mathbf{D}^n) \right) + O(\Delta t^3) \\
&= y_i(t^{n+1}) + O(\Delta t^3),
\end{aligned}$$

where $t^{n+1} = t^n + \Delta t$. Thus, the two-stage MPRKO scheme (3.2) is second order accurate. □

3.3 Computations the Patankar weight denominators

The PWDs π_i and σ_i must be specified for $i = 1, \ldots, N$. One choice that satisfies conditions (3.6) is given in the following theorem.

Theorem 3.3.1. *The two-stage MPRKO scheme (3.2) is of second order if we choose*

$$\pi_i = y_i^n, \quad \sigma_i = (y_i^n)^{1-\frac{1}{\alpha}} \left(y_i^{(2)} \right)^{\frac{1}{\alpha}}. \tag{3.28}$$

Proof. The proof is similar to the one of Theorem 3.6 in [23] using $a_{21} = \alpha$. □

Theorem 3.3.1 introduces a new two-parameter family of second order two-stage MPRKO schemes given by

$$y_i^{(1)} = y_i^n, \tag{3.29a}$$

$$y_i^{(2)} = y_i^n + \alpha \Delta t \sum_{j=1}^{N} \left(p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) \frac{y_j^{(2)}}{y_j^{(1)}} - d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) \frac{y_i^{(2)}}{y_i^{(1)}} \right), \tag{3.29b}$$

$$\begin{aligned}
y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[&\left((1 - \frac{1}{2\alpha}) p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + \frac{1}{2\alpha} p_{ij}(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}) \right) \frac{y_j^{n+1}}{(y_j^{(2)})^{\frac{1}{\alpha}} (y_j^n)^{1-\frac{1}{\alpha}}} \\
&- \left((1 - \frac{1}{2\alpha}) d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + \frac{1}{2\alpha} d_{ij}(t^n + (\alpha - 2\alpha\beta + \beta)\Delta t, \mathbf{y}^{(2)}) \right) \frac{y_i^{n+1}}{(y_i^{(2)})^{\frac{1}{\alpha}} (y_i^n)^{1-\frac{1}{\alpha}}} \Bigg]
\end{aligned} \tag{3.29c}$$

for $i = 1, \ldots, N$ with $\alpha \geq \frac{1}{2}$. Although not necessary with respect to the method's order, we requiere

$$0 \leq \beta \leq 1, \quad 0 \leq \alpha - 2\alpha\beta + \beta = \alpha - (2\alpha - 1)\beta \leq 1$$

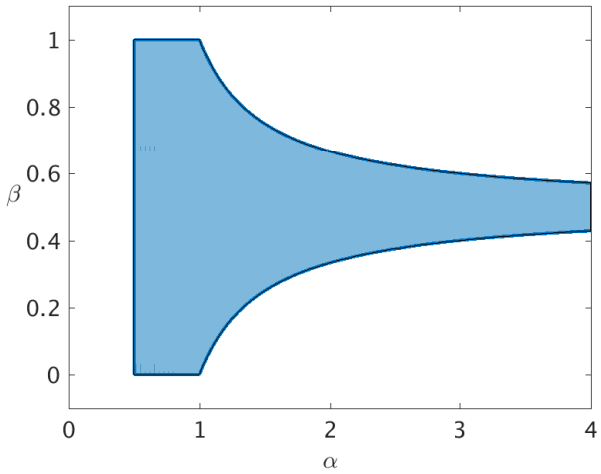


Figure 3.1: In blue, the feasible region $\mathcal{FR}$ containing all pairs (α, β) admissible for the MPRKO scheme (3.29).

to ensure that functions are only evaluated at times $t \in [t^n, t^n + \Delta t]$. It is particularly important for PDS in which the production and destruction terms P_i and D_i are only defined on a closed time interval $[0, T_{\max}]$ like problems (5.3) and (5.4) in Section 5.1. Straightforward calculations including the general requirement $\alpha \geq \frac{1}{2}$ show that the feasible region $\mathcal{FR}$ for admissible pairs (α, β) is given by

$$\mathcal{FR} = \left\{ (\alpha, \beta) \in \mathbb{R}^2 \,\middle|\, \frac{1}{2} \leq \alpha \leq 1,\, 0 \leq \beta \leq 1 \right\} \cup \left\{ (\alpha, \beta) \in \mathbb{R}^2 \,\middle|\, 1 \leq \alpha,\, \frac{\alpha - 1}{2\alpha - 1} \leq \beta \leq \frac{\alpha}{2\alpha - 1} \right\},$$

which is illustrated in Figure 3.1. We will refer to the family of schemes (3.29) with $(\alpha, \beta) \in \mathcal{FR}$ as MPRKO22(α, β) schemes.

We will base our methods using Oliver's Runge–Kutta method (RKO) [30] to construct the new class of MPRK scheme. These schemes must guarantee positivity and conservation irrespectively of the time step size.

Chapter 4

Third order MPRKO scheme

Third order modified Patankar–Runge–Kutta (MPRK) schemes were introduced in [24], which were developed to guarantee unconditional positivity and conservation, when integrating positive and conservative production-destruction systems. They introduced the first family of third-order MPRK schemes and its can be interpreted as four-stage methods named MPRK43 schemes. Recently, Kopecz and Meister proved that it is impossible to construct third-order MPRK schemes with only three stages in the usual way, which takes products of powers of previous stage values as Patankar-weight denominators [27].
We will base our methods using Oliver's Runge-Kutta method (RKO) [30] to construct the new class of MPRKO scheme. These schemes must guarantee positivity and conservation irrespective of the time step size. It is well known that an explicit three-stage RKO scheme can be represented by the following Butcher tableau

$$\begin{array}{c|ccc} c_1 \neq 0 & & & \\ c_2 & a_{21} & & \\ c_3 & a_{31} & a_{32} & \\ \hline & b_1 & b_2 & b_3 \end{array}$$

and explicit RKO method uses $f(t_0 + c_1\Delta t, x_0)$ with $c_1 \neq 0$ in the first stage contrary to the RK method that uses $f(t_0, x_0)$.
It is well known that for achieving a third order, the explicit RKO method must satisfy eigth conditions from [38].

$$b_1 + b_2 + b_3 = 1, \tag{4.1a}$$
$$b_2 a_{21} + b_3(a_{31} + a_{32}) = \frac{1}{2}, \tag{4.1b}$$
$$b_1 c_1 + b_2 c_2 + b_3 c_3 = \frac{1}{2}, \tag{4.1c}$$
$$b_1 c_1^2 + b_2 c_2^2 + b_3 c_3^2 = \frac{1}{3}, \tag{4.1d}$$
$$b_2 a_{21} c_2 + b_3(a_{31} + a_{32}) c_3 = \frac{1}{3}, \tag{4.1e}$$
$$b_2 a_{21}^2 + b_3(a_{31} + a_{32})^2 = \frac{1}{3}, \tag{4.1f}$$
$$b_2 c_1 a_{21} + b_3(c_1 a_{31} + c_2 a_{32}) = \frac{1}{6}, \tag{4.1g}$$
$$b_3 a_{21} a_{32} = \frac{1}{6}. \tag{4.1h}$$

The last condition implies $a_{21}, a_{32}, b_3 \neq 0$.

Lemma 4.1. *If*

$$c_j = \sum_{k=1}^{j-1} a_{jk}, \quad j = 2, 3, \tag{4.2}$$

then the equations (4.1) are equivalent to

$$b_1 + b_2 + b_3 = 1, \tag{4.3a}$$
$$b_2 c_2 + b_3 c_3 = \frac{1}{2}, \tag{4.3b}$$
$$b_1 c_1 + b_2 c_2 + b_3 c_3 = \frac{1}{2}, \tag{4.3c}$$
$$b_1 c_1^2 + b_2 c_2^2 + b_3 c_3^2 = \frac{1}{3}, \tag{4.3d}$$
$$b_2 c_2^2 + b_3 c_3^2 = \frac{1}{3}, \tag{4.3e}$$
$$b_3 a_{21} a_{32} = \frac{1}{6}, \tag{4.3f}$$
$$c_1(b_2 c_2 + b_3 a_{31}) = 0. \tag{4.3g}$$

Proof. Replacing a_{21} by c_2, $a_{31}+a_{32}$ by c_3 and using (4.1g) into (4.1f) we get (4.3). □

Remark 4.0.1. *The previous lemma includes the usual conditions of the third-order Runge-Kutta scheme [24] when $c_1 = 0$.*

Oliver [30] found the conditions of a family of third-order schemes with $c_1 \neq 0$ and are summarized in the following lemma

Lemma 4.2.

$$c_2 = a_{21}, \tag{4.4a}$$
$$c_3 = a_{31} + a_{32}, \tag{4.4b}$$
$$b_2 + b_3 = 1, \tag{4.4c}$$
$$b_2c_2 + b_3c_3 = \frac{1}{2}, \tag{4.4d}$$
$$b_2c_2^2 + b_3c_3^2 = \frac{1}{3}, \tag{4.4e}$$
$$b_3a_{21}a_{32} = \frac{1}{6}, \tag{4.4f}$$
$$c_1(b_2c_2 + b_3a_{31}) = 0. \tag{4.4g}$$

Proof. See [30]. □

To determine all third order Oliver-Runge-Kutta schemes, the following lemma will be helpful.

Lemma 4.3. *Given b_2 and b_3 non-negative parameters such that $b_2 + b_3 = 1$, the nonlinear system*

$$b_2x + b_3y = \frac{1}{2}, \tag{4.5a}$$
$$b_2x^2 + b_3y^2 = \frac{1}{3}, \tag{4.5b}$$

has two solutions

$$x_1 = \frac{1}{2} - \frac{1}{2}\frac{b_3}{b_2}\sqrt{\frac{b_2}{3b_3}}, \quad y_1 = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{b_2}{3b_3}}, \tag{4.6a}$$
$$x_2 = \frac{1}{2} + \frac{1}{2}\frac{b_3}{b_2}\sqrt{\frac{b_2}{3b_3}}, \quad y_2 = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{b_2}{3b_3}}, \tag{4.6b}$$

for $0 < b_2, b_3 < 1$.

Proof. (4.5a) and (4.5b) represents a line and an ellipse respectively in the x-y-plane. There are at most two intersections of the line and the ellipse, therefore, the system (4.5) has at most two solutions. Substituting (4.5a) into (4.5b) and solving for variable x we get

$$x_1 = \frac{b_2 - b_3\sqrt{\frac{b_2(4b_2+4b_3-3)}{3b_3}}}{2b_2(b_2+b_3)}, \quad x_2 = \frac{b_2 + b_3\sqrt{\frac{b_2(4b_2+4b_3-3)}{3b_3}}}{2b_2(b_2+b_3)}, \tag{4.7}$$

and using $b_2 + b_3 = 1$, it yields to

$$x_1 = \frac{1}{2} - \frac{1}{2}\frac{b_3}{b_2}\sqrt{\frac{b_2}{3b_3}}, \quad x_2 = \frac{1}{2} + \frac{1}{2}\frac{b_3}{b_2}\sqrt{\frac{b_2}{3b_3}}, \tag{4.8}$$

and substituting (4.8) in (4.5a) we get

$$y_1 = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{b_2}{3b_3}}, \quad y_2 = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{b_2}{3b_3}}. \tag{4.9}$$

If $b_2 = 0$ implies $b_3 = 1$ and (4.5) is inconsistent. Similarly for $b_3 = 0$. Thus, (4.8) and (4.9) are valid for all b_2, b_3 with $0 < b_2, b_3 < 1$ and (4.6) are the only solutions of (4.5). □

Figure 4.1 shows the three possible cases: $0 < b_2 < b_3 < 1$, $0 < b_2 = b_3 < 1$, and $0 < b_3 < b_2 < 1$.
All explicit three-stage third order Oliver-Runge-Kutta schemes are parameterized by family with at one-free parameter and it is summarized in the following lemma. We will use the following notation $\beta := c_1$.

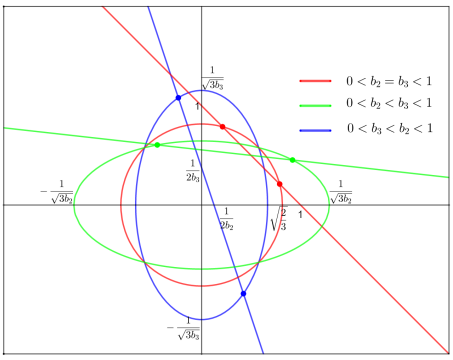


Figure 4.1: Possible intersections between the line (4.5a) and the ellipse (4.5b)

Lemma 4.4. *All explicit third-order Runge-Kutta-Oliver schemes can be represented by the following Butcher tableau:*

$$\begin{array}{c|ccc} \beta & & & \\ \frac{1}{3} & \frac{1}{3} & & \\ 1 & -1 & 2 & \\ \hline & 0 & \frac{3}{4} & \frac{1}{4} \end{array}$$

Proof. According to Lemma 3.1, the solutions of equations (4.4c), (4.4d), and (4.4e) are given by

$$c_{21} = \frac{1}{2} - \frac{1}{2}\frac{b_3}{b_2}\sqrt{\frac{b_2}{3b_3}}, \quad c_{31} = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{b_2}{3b_3}}, \tag{4.10a}$$

$$c_{22} = \frac{1}{2} + \frac{1}{2}\frac{b_3}{b_2}\sqrt{\frac{b_2}{3b_3}}, \quad c_{32} = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{b_2}{3b_3}}. \tag{4.10b}$$

As $c_1 \neq 0$, using (4.10a), (4.4f) and (4.4g) into (4.4b) we get

$$\frac{1}{2} + \frac{1}{2}\sqrt{\frac{b_2}{3b_3}} = -\frac{b_2}{2b_3} + \frac{1}{2}\sqrt{\frac{b_2}{3b_3}} + \left[6b_3\left(\frac{1}{2} - \frac{b_3}{2b_2}\sqrt{\frac{b_2}{3b_3}}\right)\right]^{-1} \tag{4.11}$$

and using $b_2 + b_3 = 1$, (4.11) yield to $\frac{b_2}{3b_3} = \sqrt{\frac{b_2}{3b_3}}$. Thus, $b_2 = \frac{3}{4}$, $b_3 = \frac{1}{4}$ implies $c_2 = a_{21} = \frac{1}{3}$, $c_3 = 1$, $a_{31} = -1$ and $a_{32} = 2$. On the other hand, using (4.10b), (4.4f) and (4.4g) into (4.4b) we get

$$\frac{1}{2} - \frac{1}{2}\sqrt{\frac{b_2}{3b_3}} = -\frac{b_2}{2b_3} - \frac{1}{2}\sqrt{\frac{b_2}{3b_3}} + \left[6b_3\left(\frac{1}{2} + \frac{b_3}{2b_2}\sqrt{\frac{b_2}{3b_3}}\right)\right]^{-1} \tag{4.12}$$

and using $b_2 + b_3 = 1$, (4.12) yield to $-\frac{b_2}{3b_3} = \sqrt{\frac{b_2}{3b_3}}$. Thus, (4.4) is inconsistent for (4.10b). Therefore, the only solution is represented by the Butcher tablue. □

4.1 Third order MPRKO scheme

Now, we will develop MPRKO schemes based on third-order three-stage explcit Runge-Kutta-Oliver schemes. The proposed MPRKO scheme with $\beta \neq 0$ is the following

$$y_i^{(1)} = y_i^n, \tag{4.13a}$$

$$y_i^{(2)} = y_i^n + \frac{\Delta t}{3}\sum_{j=1}^{N}\left(p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)})\frac{y_j^{(2)}}{\pi_j} - d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)})\frac{y_i^{(2)}}{\pi_i}\right), \tag{4.13b}$$

$$\begin{aligned} y_i^{(3)} = y_i^n + \Delta t\sum_{j=1}^{N}\Bigg[&\left(-p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + 2p_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)})\right)\frac{y_j^{(3)}}{\rho_j} \\ &\left(-d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + 2d_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)})\right)\frac{y_i^{(3)}}{\rho_i}\Bigg] \end{aligned} \tag{4.13c}$$

$$\begin{aligned} y_i^{n+1} = y_i^n + \Delta t\sum_{j=1}^{N}\Bigg[&\left(\frac{3}{4}p_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)}) + \frac{1}{4}p_{ij}(t^n + \Delta t, \mathbf{y}^{(3)})\right)\frac{y_j^{n+1}}{\sigma_j} \\ &- \left(\frac{3}{4}d_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)}) + \frac{1}{4}d_{ij}(t^n + \Delta t, \mathbf{y}^{(3)})\right)\frac{y_i^{n+1}}{\sigma_i}\Bigg] \end{aligned} \tag{4.13d}$$

for $i = 1, .., N$, where the Patankar weight denominators π_i, ρ_i and σ_i must be determined to ensure the third order of accuracy.

4.1.1 Matrix form of the scheme MPRKO

MPRKO equations (4.13) can be written in matrix-vector notation

$$\mathbf{y}^{(1)} = \mathbf{y}^n, \tag{4.14a}$$
$$\mathbf{M}^{(2)}\mathbf{y}^{(2)} = \mathbf{y}^n, \tag{4.14b}$$
$$\mathbf{M}^{(3)}\mathbf{y}^{(3)} = \mathbf{y}^n, \tag{4.14c}$$
$$\mathbf{M}\mathbf{y}^{n+1} = \mathbf{y}^n, \tag{4.14d}$$

with

$$m_{ii}^{(2)} = 1 + \frac{\Delta t}{3}\sum_{j=1}^{N} d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)})\frac{1}{\pi_i} > 0, i = 1, ..., N, \tag{4.15a}$$
$$m_{ij}^{(2)} = -\frac{\Delta t}{3} p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)})\frac{1}{\pi_j} \leq 0, i, j = 1, ..., N, i \neq j, \tag{4.15b}$$

$$m_{ii}^{(3)} = 1 + \Delta t\sum_{j=1}^{N}\left(-d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + 2d_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)})\right)\frac{1}{\rho_i}, i = 1, ..., N, \tag{4.16a}$$
$$m_{ij}^{(3)} = -\Delta t\left(-p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + 2p_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(1)})\right)\frac{1}{\rho_j}, i, j = 1, ..., N, i \neq j, \tag{4.16b}$$

and

$$m_{ii} = 1 + \Delta t\sum_{j=1}^{N}\left(\frac{3}{4}d_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)}) + \frac{1}{4}d_{ij}(t^n + \Delta t, \mathbf{y}^{(3)})\right)\frac{1}{\sigma_i} > 0, i = 1, ..., N, \tag{4.17a}$$
$$m_{ij} = -\Delta t\left(\frac{3}{4}p_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)}) + \frac{1}{4}p_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(3)})\right)\frac{1}{\sigma_j} \leq 0, i, j = 1, ..., N, i \neq j. \tag{4.17b}$$

Remark 4.1.1. *By Taylor expansion, the term* $-d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + 2d_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)}) = d_{ij}^{(n)} + ...$ *is positive for small time step* Δt*. Therefore,* $m_{ii}^{(3)} > 0$ *for* $i = 1, ..., N$*. Similarly, the term* $-p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + 2p_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)}) = p_{ij}^{(n)} + ...$ *is positive and* $m_{ij}^{(3)} < 0$ *for* $i, j = 1, ..., N$*.*

The next lemma states that MPRKO schemes applied to a conservative PDS is unconditionally positive and conservative.

Lemma 4.5. *A MPRKO scheme (4.13) applied to a conservative PDS is unconditionally conservative with* $\sum_{i=1}^{N}(y_i^{n+1}-y_i^n)=0$ *and also is unconditionally positive since for all* $\Delta t>0$ *and* $y_n>0$*, we get* $y_{n+1}>0$*.*

Proof. The proof can be done following Lemma 2.7 and 2.8 in [23]. □

Remark 4.1.2. *The transpose of the matrices* M*,* $M^{(2)}$ *and* $M^{(3)}$ *of the MPRKO scheme (4.13) are strictly diagonally dominant. Moreover, they are M-Matrix. Therefore,* M*,* $M^{(2)}$ *and* $M^{(3)}$ *are nonsingular. See Lemma 3.1 in [23] for more details.*

4.2 Proof of third-order conditions

The next theorem gives necessary and sufficient conditions for the PWDs of a third order three stage MPRKO scheme.

Theorem 4.2.1. *The MPRKO scheme (4.13) is of third-order, if and only if the conditions*

$$\pi_i = y_i^n + O(\Delta t),\ i=1,..,N, \tag{4.18a}$$

$$\rho_i = y_i^n + O(\Delta t),\ i=1,..,N, \tag{4.18b}$$

$$1 = \frac{1}{2}\frac{y_i^n+\frac{1}{3}\Delta t(P_i^n-D_i^n)}{\pi_i} + \frac{1}{2}\frac{y_i^n+\Delta t(P_i^n-D_i^n)}{\rho_i} + O(\Delta t^2),\ i=1,..,N, \tag{4.18c}$$

$$\sigma_i = y_i^n + \Delta t(P_i^n-D_i^n) + \frac{\Delta t^2}{2}\frac{\partial(P_i^n-D_i^n)}{\partial \mathbf{y}}(\mathbf{P}^n-\mathbf{D}^n) + O(\Delta t^3),\ i=1,..,N, \tag{4.18d}$$

are satisfied.

Proof. Denoting P_i^n and D_i^n by $P_i(t^n,\mathbf{y}^n)$ and $D_i(t^n,\mathbf{y}^n)$ respectively, the exact solution of (1.1) at time t^{n+1} can be written as

$$\begin{aligned} y_i(t^{n+1}) = y_i(t^n) + \Delta t(P_i^n-D_i^n) + \frac{\Delta t^2}{2}\frac{\partial(P_i^n-D_i^n)}{\partial \mathbf{y}}(\mathbf{P}^n-\mathbf{D}^n) + \frac{\Delta t^3}{6}(\mathbf{P}^n-\mathbf{D}^n)^T\mathbf{H}^n_{P_i^n-D_i^n}(\mathbf{P}^n-\mathbf{D}^n) \quad , \\ + \frac{\Delta t^3}{6}\sum_{i=1}^{N}\frac{\partial(P_i^n-D_i^n)}{\partial y_k}\frac{\partial(P_k^n-D_k^n)}{\partial \mathbf{y}}(\mathbf{P}^n-\mathbf{D}^n) + O(\Delta t^4) \end{aligned} \tag{4.19}$$

for $i=1,...,N$. $\mathbf{H}^n_{P_i^n-D_i^n}$ denotes the Hessian matrix of $\mathbf{P}-\mathbf{D}$ evaluated at y_n.

4.2.1 Necessary Conditions

Since the solver (4.13) is third-order accurate, from (4.13d) and (4.19) we get

$$\Delta t\sum_{i=1}^{N}\left[\left(\frac{3}{4}p_{ij}(t^n+\frac{\Delta t}{3},\mathbf{y}^{(2)})+\frac{1}{4}p_{ij}(t^n+\Delta t,\mathbf{y}^{(3)})\right)\frac{y_j^{n+1}}{\sigma_j}-\left(\frac{3}{4}d_{ij}(t^n+\frac{\Delta t}{3},\mathbf{y}^{(2)})+\frac{1}{4}d_{ij}(t^n+\Delta t,\mathbf{y}^{(3)})\right)\frac{y_i^{n+1}}{\sigma_i}\right]$$
$$-\Delta t\,(P_i^n-D_i^n)-\frac{\Delta t^2}{2}\frac{\partial(P_i^n-D_i^n)}{\partial\mathbf{y}}(\mathbf{P}^n-\mathbf{D}^n)-\frac{\Delta t^3}{6}(\mathbf{P}^n-\mathbf{D}^n)^T\mathbf{H}^n_{P_i^n-D_i^n}(\mathbf{P}^n-\mathbf{D}^n)$$
$$-\frac{\Delta t^3}{6}\sum_{i=1}^{N}\frac{\partial(P_i^n-D_i^n)}{\partial y_k}\frac{\partial(P_k^n-D_k^n)}{\partial\mathbf{y}}(\mathbf{P}^n-\mathbf{D}^n)=O(\Delta t^4), \tag{4.20}$$

for $i=1,...,N$. Now, we apply the scheme to the PDS(3.7), with $I,J\in\{1,2,...,N\}$,$I\neq J$, $\mu>0$ for all $t>0$, $k\in\{1,2\}$. Recalling that $p_{Ij}(\mathbf{y})=0$ for any j (4.20), it yields

$$-\left(\frac{3}{4}D_I^{(2)}+\frac{1}{4}D_I^{(3)}\right)\frac{y_I^{n+1}}{\sigma_I}+D_I^n-\frac{\Delta t}{2}\frac{\partial(D_I^n)}{\partial y_I}(D_I^n)+\frac{\Delta t^2}{6}(D_I^n)^2\frac{\partial^2 D_I^n}{\partial y_I^2}+\frac{\Delta t^2}{6}D_I^n\left(\frac{\partial D_I^n}{\partial y_I}\right)^2=O(\Delta t^3). \tag{4.21}$$

Using Taylor expansion for $k=1,2$, the destruction terms D_I^k is written as

$$D_I^{(k)}=D_I^n+\frac{\partial D_I^n}{\partial y_I}(y_I^{(k)}-y_I^n)+\frac{1}{2}\frac{\partial^2 D_I^n}{\partial y_I^2}(y_I^{(k)}-y_I^n)^2, \tag{4.22}$$

where the higher two order terms vanish. Replacing (4.22) into (4.21), we get

$$D_I^n\left(1-\frac{y_I^{n+1}}{\sigma_I}\right)-\frac{\partial D_I^n}{\partial y_I}\left(\left(\frac{3}{4}(y_I^{(2)}-y_I^n)+\frac{1}{4}(y_I^{(3)}-y_I^n)\right)\frac{y_I^{n+1}}{\sigma_I}+\frac{\Delta t}{2}D_I^{(n)}\right)+\frac{\Delta t^2}{6}D_I^n\left(\frac{\partial D_I^n}{\partial y_I}\right)^2$$
$$-\frac{1}{2}\frac{\partial^2 D_I^n}{\partial y_I^2}\left(\left(\frac{3}{4}(y_I^{(2)}-y_I^n)^2+\frac{1}{4}(y_I^{(2)}-y_I^n)^2\right)\frac{y_I^{n+1}}{\sigma_I}-\frac{\Delta t^2}{3}(D_I^n)^2\right)=O(\Delta t^3). \tag{4.23}$$

By (4.13b), we have

$$y_I^{(2)}=y_I^n-\frac{1}{3}\Delta t D_I^n\frac{y_I^{(2)}}{\pi_I}. \tag{4.24}$$

and substituting (4.24), (4.22) into (3.2c) we get

$$y_I^{(3)}=y_I^n-\Delta t\left[D_I^n+2\frac{\partial D_I^n}{\partial y_I}\left(-\frac{1}{3}\Delta t D_I^n\frac{y_I^{(2)}}{\pi_I}\right)+\frac{\partial^2 D_I^n}{\partial y_I^2}\left(-\frac{1}{3}\Delta t D_I^n\frac{y_I^{(2)}}{\pi_I}\right)^2\right]\frac{y_I^{(3)}}{\rho_I}. \tag{4.25}$$

By $\widehat{D}_I^n=\mu y_I^k$, we have $\frac{\partial^2\widehat{D}_I^n}{\partial y_I^2}=\mu k(k-1)y_I^{k-2}$ for $k\in\{1,2\}$. First setting $k=1$, in this case $\frac{\partial^2\widehat{D}_I^n}{\partial y_I^2}=0$. Replacing (4.24), and (4.25) into (4.23), we have

$$D_I^n\left(1-\frac{y_I^{n+1}}{\sigma_I}\right)+\frac{\partial D_I^n}{\partial y_I}\Delta t D_I^n\left(\left(\frac{1}{4}\frac{y_I^{(2)}}{\pi_I}+\frac{1}{4}\frac{y_I^{(3)}}{\rho_I}\right)\frac{y_I^{n+1}}{\sigma_I}-\frac{1}{2}\right)+\frac{\Delta t^2}{6}D_I^n\left(\frac{\partial D_I^n}{\partial y_I}\right)^2\left(1-\frac{y_I^{(2)}}{\pi_I}\frac{y_I^{(3)}}{\rho_I}\frac{y_I^{n+1}}{\sigma_I}\right)=O(\Delta t^3). \tag{4.26}$$

Dividing by $D_I^n = \mu y_I > 0$, and using $\frac{\partial D_I^n}{\partial y_I} = \mu > 0$, we get

$$\left(1 - \frac{y_I^{n+1}}{\sigma_I}\right) + \Delta t \mu \left(\left(\frac{1}{4}\frac{y_I^{(2)}}{\pi_I} + \frac{1}{4}\frac{y_I^{(3)}}{\rho_I}\right)\frac{y_I^{n+1}}{\sigma_I} - \frac{1}{2}\right) + \frac{\Delta t^2}{6}\mu^2 \left(1 - \frac{y_I^{(2)}}{\pi_I}\frac{y_I^{(3)}}{\rho_I}\frac{y_I^{n+1}}{\sigma_I}\right) = O(\Delta t^3). \tag{4.27}$$

By Lemma 4 in [24], we have

$$\frac{y_I^{n+1}}{\sigma_I} - 1 = O(\Delta t^3), \tag{4.28a}$$

$$\left(\frac{1}{4}\frac{y_I^{(2)}}{\pi_I} + \frac{1}{4}\frac{y_I^{(3)}}{\rho_I}\right)\frac{y_I^{n+1}}{\sigma_I} - \frac{1}{2} = O(\Delta t^2), \tag{4.28b}$$

$$1 - \frac{y_I^{(2)}}{\pi_I}\frac{y_I^{(3)}}{\rho_I}\frac{y_I^{n+1}}{\sigma_I} = O(\Delta t). \tag{4.28c}$$

Thus, (4.28a) implies $\frac{y_I^{n+1}}{\sigma_I} \to 1$. Substituting this into (4.28b) and (4.28c), we get

$$\frac{1}{4}\frac{y_I^{(2)}}{\pi_I} + \frac{1}{4}\frac{y_I^{(3)}}{\rho_I} \to \frac{1}{2}, \tag{4.29a}$$

$$\frac{y_I^{(2)}}{\pi_I}\frac{y_I^{(3)}}{\rho_I} \to 1. \tag{4.29b}$$

It follows that

$$\frac{y_I^{(2)}}{\pi_I} \to 1, \tag{4.30a}$$

$$\frac{y_I^{(3)}}{\rho_I} \to 1. \tag{4.30b}$$

Next, we set $k = 2$, in this case $\frac{\partial^2 \widehat{D}_I^n}{\partial y_I^2} \neq 0$ and (4.25) is written as

$$y_I^{(3)} = y_I^n - \Delta t \left(D_I^n - \frac{2}{3}\Delta t \frac{\partial D_I^n}{\partial y_I} D_I^n \frac{y_I^{(2)}}{\pi_I} + \frac{1}{9}\Delta t^2 \frac{\partial^2 D_I^n}{\partial y_I^2}(D_I^n)^2 \left(\frac{y_I^{(2)}}{\pi_I}\right)^2\right)\frac{y_I^{(3)}}{\rho_I}. \tag{4.31}$$

Using (4.30a) into (4.31), it yields to

$$y_I^{(3)} = y_I^n - \Delta t D_I^n \frac{y_I^{(3)}}{\rho_I} + \frac{2}{3}\Delta t^2 \frac{\partial D_I^n}{\partial y_I} D_I^n \frac{y_I^{(2)}}{\pi_I}\frac{y_I^{(3)}}{\rho_I} + O(\Delta t^2). \tag{4.32}$$

Thus,

$$(y_I^{(3)} - y_I^n)^2 = \Delta t^2 (D_I^n)^2 \left(\frac{y_I^{(3)}}{\rho_I}\right)^2 + O(\Delta t^2). \tag{4.33}$$

Replacing (4.33) and (4.24) into (4.23) yields

$$D_I^n\left(1-\frac{y_I^{n+1}}{\sigma_I}\right)-\frac{\partial D_I^n}{\partial y_I}\left(\left(-\frac{1}{4}\Delta t D_I^n\left(\frac{y_I^{(2)}}{\pi_I}+\frac{y_I^{(3)}}{\rho_I}\right)+\frac{1}{6}\Delta t^2\frac{\partial D_I^n}{\partial y_I}D_I^n\frac{y_I^{(2)}}{\pi_I}\frac{y_I^{(3)}}{\rho_I}\right)\frac{y_I^{n+1}}{\sigma_I}+\frac{\Delta t}{2}D_I^n\right)$$
$$-\frac{1}{2}\frac{\partial^2 D_I^n}{\partial y_I^2}\left(\frac{1}{4}\Delta t^2(D_I^n)^2\left(\frac{1}{3}\left(\frac{y_I^{(2)}}{\pi_I}\right)^2+\left(\frac{y_I^{(3)}}{\rho_I}\right)^2\right)\frac{y_I^{n+1}}{\sigma_I}-\frac{\Delta t^2}{3}(D_I^n)^2\right)+\frac{\Delta t^2}{6}D_I^n\left(\frac{\partial D_I^n}{\partial y_I}\right)^2=O(\Delta t^3). \tag{4.34}$$

Furthermore,

$$D_I^n\left(1-\frac{y_I^{n+1}}{\sigma_I}\right)+\Delta t D_I^n\frac{\partial D_I^n}{\partial y_I}\left(\frac{1}{4}\left(\frac{y_I^{(2)}}{\pi_I}+\frac{y_I^{(3)}}{\rho_I}\right)\frac{y_I^{n+1}}{\sigma_I}-\frac{1}{2}\right)+\frac{\Delta t^2}{6}D_I^n\left(\frac{\partial D_I^n}{\partial y_I}\right)^2\left(1-\frac{y_I^{(2)}}{\pi_I}\frac{y_I^{(3)}}{\rho_I}\frac{y_I^{n+1}}{\sigma_I}\right)$$
$$-\frac{1}{2}\Delta t^2(D_I^n)^2\frac{\partial^2 D_I^n}{\partial y_I^2}\left(\left(\frac{1}{12}\left(\frac{y_I^{(2)}}{\pi_I}\right)^2+\frac{1}{4}\left(\frac{y_I^{(3)}}{\rho_I}\right)^2\right)\frac{y_I^{n+1}}{\sigma_I}-\frac{1}{3}\right)=O(\Delta t^3), \tag{4.35}$$

which implies

$$\frac{1}{12}\left(\frac{y_I^{(2)}}{\pi_I}\right)^2+\frac{1}{4}\left(\frac{y_I^{(3)}}{\rho_I}\right)^2=\frac{1}{3}+O(\Delta t), \tag{4.36}$$

by (4.30a) and (4.30b).
Substituting (4.30a) and (4.30b) into (4.24) and (4.25), we find

$$y_I^{(2)} = y_I^n+O(\Delta t), \tag{4.37a}$$
$$y_I^{(3)} = y_I^n+O(\Delta t), \tag{4.37b}$$

and thus

$$1+O(\Delta t)=\frac{y_I^{(2)}}{\pi_I}=\frac{y_I^n+O(\Delta t)}{\pi_I}=\frac{y_I^n}{\pi_I}+O(\Delta t), \tag{4.38a}$$
$$1+O(\Delta t)=\frac{y_I^{(3)}}{\rho_I}=\frac{y_I^n+O(\Delta t)}{\rho_I}=\frac{y_I^n}{\rho_I}+O(\Delta t), \tag{4.38b}$$

from which we conclude

$$\pi_I=y_I^n+O(\Delta t), \tag{4.39a}$$
$$\rho_I=y_I^n+O(\Delta t). \tag{4.39b}$$

As I was chosen arbitrary, we find that (4.18a) and (4.18b) are necessary conditions. Taking into account that (4.30a, 4.30b), from (4.13b, 4.13c) it yields to

$$y_i^{(2)} = y_i^n + \frac{\Delta t}{3}\sum_{j=1}^{N}\Big(p_{ij}(t^n+\beta\Delta t,\mathbf{y}^{(1)}) - d_{ij}(t^n+\beta\Delta t,\mathbf{y}^{(1)})\Big), \tag{4.40a}$$

$$y_i^{(3)} = y_i^n + \Delta t\sum_{j=1}^{N}\left[-p_{ij}(t^n+\beta\Delta t,\mathbf{y}^{(1)}) + 2p_{ij}(t^n+\frac{\Delta t}{3},\mathbf{y}^{(2)}) + d_{ij}(t^n+\beta\Delta t,\mathbf{y}^{(1)}) - 2d_{ij}(t^n+\frac{\Delta t}{3},\mathbf{y}^{(2)})\right], \tag{4.40b}$$

for $i = 1,..,N$. By Taylor expansions

$$y_i^{(2)} = y_i^n + \frac{\Delta t}{3}\left(P_i^n - D_i^n\right) + O(\Delta t^2), \tag{4.41a}$$

$$y_i^{(3)} = y_i^n + \Delta t\left(P_i^n - D_i^n\right) + O(\Delta t^2), \tag{4.41b}$$

for $i = 1,..,N$. Substituting (4.41a) and (4.41b) together with (4.28a) into (4.28b) we have

$$\frac{1}{2}\frac{y_I^n + \frac{\Delta t}{3}\left(P_I^n - D_I^n\right)}{\pi_I} + \frac{1}{2}\frac{y_I^n + \Delta t\left(P_I^n - D_I^n\right)}{\rho_I} = 1 + O(\Delta t^2).$$

Similarly to (4.38), we replace (4.28a) into (4.19) to get

$$\sigma_I = y_I^n + \Delta t(P_I^n - D_I^n) + \frac{\Delta t^2}{2}\frac{\partial(P_I^n - D_I^n)}{\partial\mathbf{y}}(\mathbf{P}^n - \mathbf{D}^n) + O(\Delta t^3), \quad i = 1,...,N. \tag{4.42}$$

Thus, letting I run from 1 to N, we also see that conditions (4.18c) and (4.18d) are necessary.

4.2.2 Sufficient Condition

Since matrices M, $M^{(2)}$ and $M^{(3)}$ of the MPRKO scheme (4.13) are nonsingular, and let $M^{-1} = (\widetilde{m}_{ij})$, $(M^{(k)})^{-1} = (\widetilde{m}_{ij}^{(k)})$, for $i,j = 1,...,N$, $k \in \{2,3\}$, (4.14) can be written as

$$y_i^{(2)} = \sum_{j=1}^{N}\widetilde{m}_{ij}^{(2)}y_j^n, \qquad y_i^{(3)} = \sum_{j=1}^{N}\widetilde{m}_{ij}^{(3)}y_j^n, \qquad y_i^{n+1} = \sum_{j=1}^{N}\widetilde{m}_{ij}y_j^n, \tag{4.43}$$

and the Lemma 3.1 in [23] ensures that $0 < \widetilde{m}_{ij}, \widetilde{m}_{ij}^{(2)}, \widetilde{m}_{ij}^{(3)} < 1$ for $i, j = 1, ..., N$. Then,

$$\frac{y_i^{(2)}}{\pi_i} = \sum_{j=1}^{N} \widetilde{m}_{ij}^{(2)} \frac{y_j^n}{\pi_i} = O(1), \tag{4.44a}$$

$$\frac{y_i^{(3)}}{\rho_i} = \sum_{j=1}^{N} \widetilde{m}_{ij}^{(3)} \frac{y_j^n}{\rho_i} = O(1), \tag{4.44b}$$

$$\frac{y_i^{n+1}}{\sigma_i} = \sum_{j=1}^{N} \widetilde{m}_{ij} \frac{y_j^n}{\sigma_i} = O(1), \tag{4.44c}$$

for $i = 1, ..., N$. Using (4.44a) in (4.13b), it yields to

$$y_i^{(2)} = y_i^n + \frac{1}{3}\Delta t \sum_{j=1}^{N} \left(p_{ij}(t^n + \beta\Delta t, \mathbf{y}^n) \frac{y_j^{(2)}}{\pi_j} - d_{ij}(t^n + \beta\Delta t, \mathbf{y}^n) \frac{y_i^{(2)}}{\pi_i} \right), \tag{4.45}$$

for $i = 1, ..., N$. By Taylor expansion, we have that

$$y_i^{(2)} = y_i^n + \frac{1}{3}\Delta t \sum_{j=1}^{N} \left(\left(p_{ij}^n + O(\Delta t)\right) \frac{y_j^{(2)}}{\pi_j} - \left(d_{ij}^n + O(\Delta t)\right) \frac{y_i^{(2)}}{\pi_i} \right) = y_i^n + O(\Delta t). \tag{4.46}$$

Replacing (4.46) and (4.18a) again in (4.13b), we get

$$\begin{aligned} y_i^{(2)} &= y_i^n + \frac{1}{3}\Delta t \sum_{j=1}^{N} \left(\left(p_{ij}^n + O(\Delta t)\right)(1 + O(\Delta t)) - \left(d_{ij}^n + O(\Delta t)\right)(1 + O(\Delta t)) \right), \\ &= y_i^n + \frac{1}{3}\Delta t (P_i^n - D_i^n) + O(\Delta t^2). \end{aligned} \tag{4.47}$$

Similarly, using (4.44b) and (4.18c) in (4.13c) it yields to

$$\begin{aligned} y_i^{(3)} &= y_i^n + \Delta t \sum_{i=1}^{N} \left(\left(-p_{ij}^n + 2p_{ij}^{(2)} + O(\Delta t)\right)(1 + O(\Delta t)) - \left(-d_{ij}^n + 2d_{ij}^{(2)} + O(\Delta t)\right)(1 + O(\Delta t)) \right) \\ &= y_i^n - \Delta t \left(P_i^n - D_i^n\right) + 2\Delta t \left(P_i^{(2)} - D_i^{(2)}\right) + O(\Delta t^2), \end{aligned} \tag{4.48}$$

for $i = 1, ..., N$.. According to (4.46), we have $\mathbf{y}^{(2)} - \mathbf{y}^{(n)} = O(\Delta t)$, and from (4.22) we see

$$P_i^{(2)} - D_i^{(2)} = P_i^n - D_i^n + O(\Delta t). \tag{4.49}$$

Hence

$$y_i^{(3)} = y_i^n + \Delta t \left(P_i^n - D_i^n\right) + O(\Delta t^2). \tag{4.50}$$

Now we compute an expansion of y^{n+1} using (4.13d) and (4.46)-(4.50) we get

$$y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \left[\left(\frac{3}{4} p_{ij}^n + \frac{3}{4} \frac{\partial p_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(2)} - \mathbf{y}^n) + O(\Delta t^2) + \frac{1}{4} p_{ij}^n + \frac{1}{4} \frac{\partial p_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(3)} - \mathbf{y}^n) + O(\Delta t^2) \right) \frac{y_j^{n+1}}{\sigma_j} \right.$$
$$\left. - \left(\frac{3}{4} d_{ij}^n + \frac{3}{4} \frac{\partial d_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(2)} - \mathbf{y}^n) + O(\Delta t^2) + \frac{1}{4} d_{ij}^n + \frac{1}{4} \frac{\partial d_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(3)} - \mathbf{y}^n) + O(\Delta t^2) \right) \frac{y_i^{n+1}}{\sigma_i} \right]. \tag{4.51}$$

By (4.44c), we can conclude

$$y_i^{n+1} = y_i^n + O(\Delta t), \tag{4.52}$$

for i =1, ..., N. From (4.18d) and (4.52) it follows that

$$y_i^{n+1} = \sigma_i + O(\Delta t). \tag{4.53}$$

Inserting into (4.51), we get

$$y_i^{n+1} = y_i^n + \Delta t (P_i^n - D_i^n) + O(\Delta t^2), \tag{4.54}$$

for i =1, ..., N. Now, we can conclude

$$y_i^{n+1} = \sigma_i + O(\Delta t^2) \tag{4.55}$$

by (4.18d). Introducing this relation into (4.51) yields

$$y_i^{n+1} = y_i^n + \Delta t (P_i^n - D_i^n) + \Delta t \frac{\partial \left(P_i^n - D_i^n \right)}{\partial \mathbf{y}} \left(\frac{3}{4} (\mathbf{y}^{(2)} - \mathbf{y}^n) + \frac{1}{4} (\mathbf{y}^{(3)} - \mathbf{y}^n) \right) + O(\Delta t^3) \tag{4.56}$$

for i =1, ..., N. Finally, replacing (4.47) and (4.50) into (4.56) results in

$$y_i^{n+1} = y_i^n + \Delta t (P_i^n - D_i^n) + \frac{\Delta t^2}{2} \frac{\partial \left(P_i^n - D_i^n \right)}{\partial \mathbf{y}} \left(\mathbf{P}^n - \mathbf{D}^n \right) + O(\Delta t^3) \tag{4.57}$$

and with (4.18d) we can conclude

$$y_i^{n+1} = \sigma_i + O(\Delta t^3). \tag{4.58}$$

By Taylor expansion, owing to (4.47) and (4.50) we have

$$p_{ij}(\mathbf{y}^k) = p_{ij}^n + \frac{\partial p_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(k)} - \mathbf{y}^n) + \frac{1}{2} (\mathbf{y}^{(k)} - \mathbf{y}^n)^T \mathbf{H}_{ij}^n (\mathbf{y}^{(k)} - \mathbf{y}^n) + O(\Delta t^3), \tag{4.59a}$$

$$d_{ij}(\mathbf{y}^k) = d_{ij}^n + \frac{\partial d_{ij}^n}{\partial \mathbf{y}} (\mathbf{y}^{(k)} - \mathbf{y}^n) + \frac{1}{2} (\mathbf{y}^{(k)} - \mathbf{y}^n)^T \mathbf{H}_{ij}^n (\mathbf{y}^{(k)} - \mathbf{y}^n) + O(\Delta t^3), \tag{4.59b}$$

for $i, j = 1, ..., N$ and $k = 2, 3$. Finally, substitution this and (4.58) into (4.13d) yields

$$y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[p_{ij}^n + \frac{\partial p_{ij}^n}{\partial \mathbf{y}} \left(\frac{3}{4}(\mathbf{y}^{(2)} - \mathbf{y}^n) + \frac{1}{4}(\mathbf{y}^{(3)} - \mathbf{y}^n) \right) + \frac{1}{2} \frac{\partial^2 p_{ij}^n}{\partial \mathbf{y^2}} \left(\frac{3}{4}(\mathbf{y}^{(2)} - \mathbf{y}^n)^2 + \frac{1}{4}(\mathbf{y}^{(3)} - \mathbf{y}^n)^2 \right)$$
$$- d_{ij}^n \frac{\partial p_{ij}^n}{\partial \mathbf{y}} \left(\frac{3}{4}(\mathbf{y}^{(2)} - \mathbf{y}^n) + \frac{1}{4}(\mathbf{y}^{(3)} - \mathbf{y}^n) \right) - \frac{1}{2} \frac{\partial^2 p_{ij}^n}{\partial \mathbf{y^2}} \left(\frac{3}{4}(\mathbf{y}^{(2)} - \mathbf{y}^n)^2 + \frac{1}{4}(\mathbf{y}^{(3)} - \mathbf{y}^n)^2 \right) \Bigg] + O(\Delta t^4), \tag{4.60}$$

Using (4.13b) with (4.44a), we get

$$y_i^{(2)} - y_i^n = \frac{1}{3}\Delta t \, (P_i^n - D_i^n) + \frac{1}{9}\Delta t^2 \frac{\partial (P_i^n - D_i^n)}{\partial \mathbf{y}} (\mathbf{P}^n - \mathbf{D}^n) + O(\Delta t^3). \tag{4.61}$$

Similarly with (4.13c) and (4.44b)

$$y_i^{(3)} - y_i^n = \Delta t \, (P_i^n - D_i^n) + \Delta t^2 \frac{\partial (P_i^n - D_i^n)}{\partial \mathbf{y}} (\mathbf{P}^n - \mathbf{D}^n) + O(\Delta t^3). \tag{4.62}$$

Inserting (4.61) and (4.62) into (4.60), we get

$$y_i^{n+1} = y_i(t^n) + \Delta t (P_i^n - D_i^n) + \frac{\Delta t^2}{2} \frac{\partial (P_i^n - D_i^n)}{\partial \mathbf{y}} (\mathbf{P}^n - \mathbf{D}^n) + \frac{\Delta t^3}{6} (\mathbf{P}^n - \mathbf{D}^n)^T \mathbf{H}_{P_i^n - D_i^n}^n (\mathbf{P}^n - \mathbf{D}^n)$$
$$+ \frac{\Delta t^3}{6} \sum_{i=1}^{N} \frac{\partial (P_i^n - D_i^n)}{\partial y_k} \frac{\partial (P_k^n - D_k^n)}{\partial \mathbf{y}} (\mathbf{P}^n - \mathbf{D}^n) + O(\Delta t^4), \tag{4.63}$$

for $i = 1, ..., N$. Thus, the MPRKO scheme (4.13) is third-order accurate. □

4.3 Computations the Patankar weight denominators

Based on the same idea used in [24] to compute the PDWs σ_i, we compute these weights using a second order MPRKO($\frac{1}{3}, \beta$) scheme, as condition (4.18d) of the Theorem 4.2.1 requires. The following theorem defines a family of third order MPRK schemes.

Theorem 4.3.1. *Given an explicit three-stage third order Runge-Kutta-Oliver scheme with nonnegative PWDs, the MPRKO scheme (4.13) is of third-order, if we choose*

$$\pi_i = y_i^n, \ i = 1, \ldots, N, \tag{4.64a}$$

$$\rho_i = y_i^n \left(\frac{y_i^{(2)}}{y_i^n} \right)^4, \ i = 1, \ldots, N, \tag{4.64b}$$

$$\mu_i = y_i^n \left(\frac{y_i^{(2)}}{y_i^n} \right)^3, \ i = 1, \ldots, N, \tag{4.64c}$$

$$\begin{aligned} \sigma_i = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg(\Bigg(& -\frac{1}{2} p_{ij} \left(t^n + \beta \Delta t, \mathbf{y}^n\right) + \frac{3}{2} p_{ij}(t^n + \frac{\beta+1}{3} \Delta t, \mathbf{y}^{(2)}) \Bigg) \frac{\sigma_j}{\mu_j} \\ & - \Bigg(-\frac{1}{2} d_{ij}(t^n + \beta \Delta t, \mathbf{y}^n) + \frac{3}{2} d_{ij}(t^n + \frac{\beta+1}{3} \Delta t, \mathbf{y}^{(2)}) \Bigg) \Bigg) \frac{\sigma_i}{\mu_i}, \ i = 1, \ldots, N. \end{aligned} \tag{4.64d}$$

Proof. Note that (4.18a) is obvious with condition (4.64a), which implies

$$y_i^{(2)} = y_i^n + \frac{1}{3} \Delta t (P_i^n - D_i^n) + O(\Delta t^2), \tag{4.65}$$

for $i = 1, ..., N$. Introducing this into (4.64b), we show

$$\rho_i = y_i^n \left(\frac{y_i^n + \frac{1}{3} \Delta t (P_i^n - D_i^n) + O(\Delta t^2)}{y_i^n} \right)^4. \tag{4.66}$$

By Newton binomial theorem, we have

$$\rho_i = y_i^n \frac{(y_i^n)^4 + \frac{4}{3} (y_i^n)^3 \Delta t (P_i^n - D_i^n) + O(\Delta t^2)}{(y_i^n)^4}. \tag{4.67}$$

Furthermore,

$$\rho_i = y_i^n + \frac{4}{3} \Delta t (P_i^n - D_i^n) + O(\Delta t^2), \tag{4.68}$$

for $i = 1, ..., N$. Thus, condition (4.18b) holds true as well. Now, we verify that (4.64b) and (4.64c) satisfy the condition (4.18c). Defining $f(\Delta t) = 1/(a + b\Delta t)$ for some constants a and b, we have that $f(\Delta t) = f(0) + f(0)\Delta t + O(\Delta t^2) = \frac{1}{a} - \frac{b}{a^2} + O(\Delta t^2)$. Thus (4.68) implies

$$\frac{1}{\rho_i} = \frac{1}{y_i^n} - \frac{4}{3} \frac{\Delta t (P_i^n - D_i^n)}{(y_i^n)^2} + O(\Delta t^2), \tag{4.69}$$

for $i = 1, ..., N$. Substituting this and (4.64a) into condition (4.18c), we have

$$\begin{aligned}
&\frac{1}{2}\frac{y_i^n + \frac{1}{3}\Delta t(P_i^n - D_i^n)}{\pi_i} + \frac{1}{2}\frac{y_i^n + \Delta t(P_i^n - D_i^n)}{\rho_i} \\
&= \frac{1}{2}\frac{y_i^n + \frac{1}{3}\Delta t(P_i^n - D_i^n)}{y_i^n} + \frac{1}{2}\left(y_i^n + \Delta t(P_i^n - D_i^n)\right)\left(\frac{1}{y_i^n} - \frac{4}{3}\frac{\Delta t(P_i^n - D_i^n)}{(y_i^n)^2} + O(\Delta t^2)\right), \\
&= \frac{1}{2} + \frac{1}{6}\frac{\Delta t(P_i^n - D_i^n)}{y_i^n} + \frac{1}{2} - \frac{2}{3}\frac{\Delta t(P_i^n - D_i^n)}{y_i^n} + \frac{1}{2}\frac{\Delta t(P_i^n - D_i^n)}{y_i^n} + O(\Delta t^2), \\
&= 1 + O(\Delta t^2),
\end{aligned}$$

for $i = 1, ..., N$. And (4.18b) is satisfied. Finally (4.64c) and (4.64d) satisfy the conditions of a second order MPRKO($\frac{1}{3}, \beta$) scheme of [26] and thus the condition (4.18d) is satisfied as well. □

Assuming β is not restricted, Theorem 3.3.1 introduces a new one-parameter family of third-order four-stage MPRKO schemes given by

$$y_i^{(1)} = y_i^n, \tag{4.70a}$$

$$y_i^{(2)} = y_i^n + \frac{1}{3}\Delta t \sum_{j=1}^{N}\left(p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)})\frac{y_j^{(2)}}{y_j^{(1)}} - d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)})\frac{y_i^{(2)}}{y_i^{(1)}}\right), \tag{4.70b}$$

$$\begin{aligned}
y_i^{(3)} = y_i^n + \Delta t \sum_{j=1}^{N}\Bigg[&\left(-p_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + 2p_{ij}(t^n + \frac{1}{3}\Delta t, \mathbf{y}^{(2)})\right)\frac{y_j^{(3)}}{(y_j^{(2)})^4(y_j^n)^{-3}} \\
&- \left(-d_{ij}(t^n + \beta\Delta t, \mathbf{y}^{(1)}) + 2d_{ij}(t^n + \frac{1}{3}\Delta t, \mathbf{y}^{(2)})\right)\frac{y_i^{(3)}}{(y_i^{(2)})^4(y_i^n)^{-3}}\Bigg]
\end{aligned} \tag{4.70c}$$

$$\begin{aligned}
\sigma_i = y_i^n + \Delta t \sum_{j=1}^{N}\Bigg[&\left(-\frac{1}{2}p_{ij}\left(t^n + \beta\Delta t, \mathbf{y}^n\right) + \frac{3}{2}p_{ij}(t^n + \frac{\beta+1}{3}\Delta t, \mathbf{y}^{(2)})\right)\frac{\sigma_j}{(y_j^{(2)})^3(y_j^n)^{-2}} \\
&\left(-\frac{1}{2}d_{ij}(t^n + \beta\Delta t, \mathbf{y}^n) + \frac{3}{2}d_{ij}(t^n + \frac{\beta+1}{3}\Delta t, \mathbf{y}^{(2)})\right)\frac{\sigma_i}{(y_i^{(2)})^3(y_i^n)^{-2}}\Bigg]
\end{aligned} \tag{4.70d}$$

$$\begin{aligned}
y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N}\Bigg[&\left(\frac{3}{4}p_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)}) + \frac{1}{4}p_{ij}(t^n + \Delta t, \mathbf{y}^{(3)})\right)\frac{y_j^{n+1}}{\sigma_j} \\
&- \left(\frac{3}{4}d_{ij}(t^n + \frac{\Delta t}{3}, \mathbf{y}^{(2)}) + \frac{1}{4}d_{ij}(t^n + \Delta t, \mathbf{y}^{(3)})\right)\frac{y_i^{n+1}}{\sigma_i}\Bigg]
\end{aligned} \tag{4.70e}$$

for $i = 1, .., N$.

Although not necessary with respect to the method's order, we require $0 \leq \beta \leq 1$ to

ensure that functions are only evaluated at times $t \in [t^n, t^n + \Delta t]$. This is particularly important for PDS in which the production and destruction terms P_i and D_i are only defined on a closed time interval $[0, T_{max}]$ like problems (5.2) and (5.4) in Chapter 5.1. In the following, we will refer to this family of schemes as MPRKO43(β) schemes. Numerical experiments confirm the theoretical convergence order of the MPRKO43 scheme are presented in Chapter 5. Additionally, numerical solutions of the Robertson problems will show that these schemes have the ability to integrate stiff PDS.

Chapter 5

Numerical results

In this chapter, we apply the second-order MPRKO22(α, β) schemes (3.2) and the third-order MPRKO43(β) schemes (4.13) to six numerical tests given in the literature to compare numerical methods. To obtain reference solutions, we use the Matlab function **ode15s**, see [35], with tolerances **AbsTol** $= 10^{-12}$ and **RelTol** $= 10^{-10}$ in all computations of this chapter. To compare different MPRKO22(α, β) schemes from the chapter 3 and MPRKO43(β) schemes from the chapter 4, we introduce a relative error E taken over all state variables and all time steps, which is defined as

$$E = \frac{1}{N}\sum_{i=1}^{N} E_i, \quad E_i = \left(\frac{1}{M}\sum_{m=1}^{M}\left(y_{\mathrm{ref},i}^{m} - y_i^m\right)^2\right)^{\frac{1}{2}}\left(\frac{1}{M}\sum_{m=1}^{M}\left(y_{\mathrm{ref},i}^{m}\right)^2\right)^{-\frac{1}{2}}, \tag{5.1}$$

where M denotes the number of executed time steps and $\mathbf{y}_{\mathrm{ref}}^m$ is the value of a reference solution at time t^m.
In the next section, we introduce the test problems used to assess the accuracy and performance of our proposed schemes of second (3.29) and third (4.70) order.

5.1 Nonautonomous test

We consider a simple linear nonautonomous test modeling, salt exchange between two very large tanks partly filled with brine, a two-compartment model describing the interaction between the cytoplasm and the nucleus, and a Susceptible-Exposed-Infectious-Recovered (SEIR) epidemiological model.

5.1.1 Simple linear test

A simple linear nonautonomous test from [41, Sec. 3.3, Ex. 7] is given by

$$y_1'(t) = \frac{ay_2(t)}{100 + (b-a)t} - \frac{by_1(t)}{100 + (a-b)t}, \tag{5.2a}$$

$$y_2'(t) = \frac{by_1(t)}{100 + (a-b)t} - \frac{ay_2(t)}{100 + (b-a)t}, \tag{5.2b}$$

with constant parameters $a > b > 0$ and initial conditions $y_i(0) = y_i^0 > 0$ for $i = 1, 2$ on the time interval $[0, T]$ with $T < 100/(a-b)$. The system (5.2) can be written as a fully conservative production-destruction system (1.1) with

$$p_{12}(t, \mathbf{y}(t)) = d_{21}(t, \mathbf{y}(t)) = \frac{ay_2(t)}{100 + (b-a)t}, \quad p_{21}(t, \mathbf{y}(t)) = d_{12}(t, \mathbf{y}(t)) = \frac{by_1(t)}{100 + (a-b)t}$$

and $p_{ii}(t, \mathbf{y}(t)) = d_{ii}(t, \mathbf{y}(t)) = 0$ for $i = 1, 2$.

Equations (5.2) describe a closed system exchanging brine between two 100 gallons tanks A and B. Concentration $y_i(t)$ corresponds to the salt amount in tank i for $i = 1, 2$ at any time. Initially, y_1^0 pounds of salt are dissolved in tank A and y_2^0 pounds of salt are dissolved in tank B. The well-stirred liquid is pumped between tanks with exchange rates a [gal/min] and b [gal/min], respectively.

In our numerical experiments we use $a = 3$ and $b = 2$ on the time interval $[0, 90]$ with $\Delta t = 10$ and initial values $y_1(0) = 0.01$ and $y_2(0) = 99.99$ [26].

5.1.2 Linear JAK2/STAT5 model

This positive and conservative two-compartment model from [9] describes interactions between cytoplasm and nucleus inside a cell. In this model y_1 and y_2 denote the unphosphorylated and phosphorylated STAT5 concentrations in the cytoplasm, while y_3 and y_4 denote the unphosphorylated and phosphorylated STAT5 concentrations in the nucleus. The variables $y_5, \ldots, y_8$ describe processes in the nucleus.

The model is given by

$$y_1'(t) = -\frac{r_a}{v_c}\mathrm{pJAK}(t)y_1(t) - \frac{r_i}{v_c}y_1(t) + \frac{r_e}{v_n}y_3(t), \tag{5.3a}$$

$$y_2'(t) = -\frac{r_{i2}}{v_c}y_2(t) + \frac{r_a}{v_c}\mathrm{pJAK}(t)y_1(t), \tag{5.3b}$$

$$y_3'(t) = -\frac{r_e}{v_n}y_3(t) + \frac{r_i}{v_c}y_1(t) + \frac{r_d}{v_n}y_8(t), \tag{5.3c}$$

$$y_4'(t) = -\frac{r_d}{v_n}y_4(t) + \frac{r_{i2}}{v_c}y_2(t), \tag{5.3d}$$

$$y_k'(t) = -\frac{r_d}{v_n}y_k(t) + \frac{r_d}{v_n}y_{k-1}(t), \quad k = 5, \ldots, 8, \tag{5.3e}$$

Table 5.1: Parameter values used in model (5.3).

Parameter	Unit	Value
r_a	min^{-1}	11
r_i	min^{-1}	39
r_{i2}	min^{-1}	58
r_e	min^{-1}	265
r_d	min^{-1}	225
v_c	μm^3	429
v_n	μm^3	268

Table 5.2: Interpolation points (x_k, y_k), $k = 1, \ldots, 10$ used to obtain the function pJAK in (5.3).

k	1	2	3	4	5	6	7	8	9	10
x_k	0	20	40	60	80	100	120	140	160	180
y_k	0.25	1.90	1.50	1.10	0.85	0.68	0.58	0.50	0.45	0.44

on the time interval $[0, 180]$ with initial values

$$y_1(0) = 50v_c, \quad y_2(0) = 0, \quad y_3(0) = 18v_n, \quad y_k(0) = 0,\ k = 4, \ldots, 8.$$

The parameters $r_a, v_c, r_i, r_e, r_{i2}, v_n$ and r_d are constants with values given in Table 5.1. The function pJAK is given as the cubic spline interpolant, which interpolates the points given in Table 5.2 and is shown in Figure 5.1. The interpolation points were taken from measurements provided by E. Friedmann through personal communication.

Summing up all the equations (5.3), we see

$$\sum_{k=1}^{8} y_k'(t) = 0 \quad \text{such that} \quad \sum_{k=1}^{8} y_k(t) = \sum_{k=1}^{8} y_k(0),$$

which proves that the system (5.3) is conservative. Moreover, it can be written in the form (1.1) with

$$\begin{aligned}
&p_{13}(t, \mathbf{y}(t)) = d_{31}(t, \mathbf{y}(t)) = \frac{r_e}{v_n} y_3(t), && p_{21}(t, \mathbf{y}(t)) = d_{12}(t, \mathbf{y}(t)) = \frac{r_a}{v_c} \text{pJAK}(t) y_1(t), \\
&p_{31}(t, \mathbf{y}(t)) = d_{13}(t, \mathbf{y}(t)) = \frac{r_i}{v_c} y_1(t), && p_{38}(t, \mathbf{y}(t)) = d_{83}(t, \mathbf{y}(t)) = \frac{r_d}{v_n} y_8(t), \\
&p_{42}(t, \mathbf{y}(t)) = d_{24}(t, \mathbf{y}(t)) = \frac{r_{i2}}{v_c} y_2(t), && p_{k,k-1}(t, \mathbf{y}(t)) = d_{k-1,k}(t, \mathbf{y}(t)) = \frac{r_d}{v_n} y_{k-1}(t), \quad k = 5, \ldots, 8
\end{aligned}$$

and $p_{ij}(t, \mathbf{y}(t)) = d_{ji}(t, \mathbf{y}(t)) = 0$ in other cases.

In the numerical experiments, later presented in Section 5, we use $\Delta t = 10$.

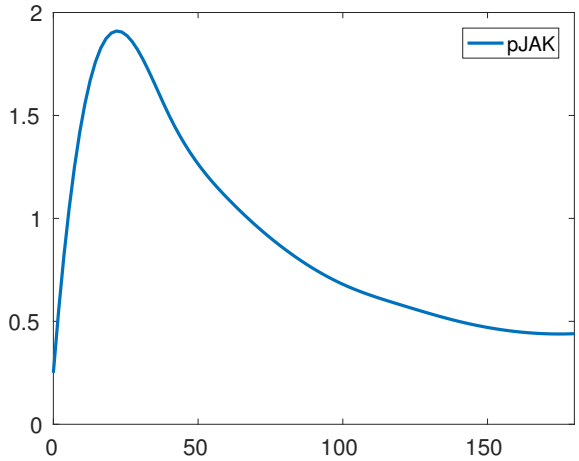


Figure 5.1: Graph of the function pJAK used in model [9].

5.1.3 Epidemiological SEIR model

In [8] vaccination strategies for a Susceptible-Exposed-Infectious-Recovered (SEIR) model are presented. The model is given by

$$y_1'(t) = -\frac{\beta y_1(t) y_3(t)}{N} + (\mu+\omega) y_4(t) + \mu y_2(t) + \mu y_3(t) - \mu N V(t), \tag{5.4a}$$

$$y_2'(t) = \frac{\beta y_1(t) y_3(t)}{N} - (\mu+\sigma) y_2(t), \tag{5.4b}$$

$$y_3'(t) = -(\mu+\gamma) y_3(t) + \sigma y_2(t), \tag{5.4c}$$

$$y_4'(t) = -(\mu+\omega) y_4(t) + \gamma y_3(t) + \mu N V(t), \tag{5.4d}$$

with constant parameters N, μ, ω, β, γ and σ given in Table 5.3. The variables y_1, y_2, y_3 and y_4 denote the susceptible (S), exposed (E), infectious (I) and recovered (R) compartments, respectively. By $V(t)$ we denote the vaccination strategy which is used to decrease the S, E and I populations and to increase the R population. Different from [8] we consider $V(t)$ as a given function. In particular, we use the vaccination strategy

$$V(t) = \frac{22500}{\mu N} \times \exp\Big(-\frac{1}{4}t\Big),$$

which models the progression depicted in [8, Fig. 5]. The initial conditions are

$$y_1(0) = 9.8 \times 10^5, \quad y_2(0) = 1.5 \times 10^4, \quad y_3(0) = 5 \times 10^3, \quad y_4(0) = 0,$$

and the time interval of interest is $[0, 60]$.

Table 5.3: SEIR model parameters (5.4).

Parameter	Unit	Value
N	—	10^6
μ	day^{-1}	5.48×10^{-5}
ω	day^{-1}	$1/7$
β	day^{-1}	3.288
γ	day^{-1}	0.274
σ	day^{-1}	9.82×10^{-2}

The model can be written as a PDS with

$$p_{12}(t,\mathbf{y}(t)) = d_{21}(t,\mathbf{y}(t)) = \mu y_2(t), \qquad p_{13}(t,\mathbf{y}(t)) = d_{31}(t,\mathbf{y}(t)) = \mu y_3(t),$$
$$p_{14}(t,\mathbf{y}(t)) = d_{41}(t,\mathbf{y}(t)) = (\mu+\omega) y_4(t), \qquad p_{21}(t,\mathbf{y}(t)) = d_{12}(t,\mathbf{y}(t)) = \frac{\beta y_1(t) y_3(t)}{N},$$
$$p_{32}(t,\mathbf{y}(t)) = d_{23}(t,\mathbf{y}(t)) = \sigma y_2(t), \qquad p_{41}(t,\mathbf{y}(t)) = d_{14}(t,\mathbf{y}(t)) = \mu N V(t),$$
$$p_{43}(t,\mathbf{y}(t)) = d_{34}(t,\mathbf{y}(t)) = \gamma y_3(t)$$

and $p_{ij}(t,\mathbf{y}(t)) = d_{ji}(t,\mathbf{y}(t)) = 0$ in other cases.

Furthermore, we use $\Delta t = 2$ within the numerical computations presented in the following section.

5.2 Autonomous test

We consired a linear model of mass exchange between two constituents, a non-linear model for a phytoplankton bloom, the Brusselator problem for multimolecular reactions as well as Stiff Robertson test, one of the most prominent examples of stiff ODEs.

5.2.1 Simple linear test

A linear model of mass exchange between two constituents from [4] is given by

$$y_1'(t) = y_2(t) - a y_1(t), \tag{5.5a}$$
$$y_2'(t) = a y_1(t) - y_2(t), \tag{5.5b}$$

with the non-dimensional constant $a > 0$ initial conditions $y_i(0) = y_i^0 > 0$ for $i = 1, 2$. The system (5.5) can be written as a fully conservative production-destruction system (1.1) with

$$p_{12}(t,\mathbf{y}(t)) = d_{21}(t,\mathbf{y}(t)) = y_2(t), \quad p_{21}(t,\mathbf{y}(t)) = d_{12}(t,\mathbf{y}(t)) = a y_1(t),$$

and $p_{ii}(t,\mathbf{y}(t)) = d_{ii}(t,\mathbf{y}(t)) = 0$ for $i = 1, 2$.

In our numerical experiments we use $a = 5$ and $\Delta t = 0.25$ and initial values $y_1(0) = 0.9$ and $y_2(0) = 0.1$. The solution is approximated on the time interval $[0, 1.75]$.

5.2.2 Nonlinear test

A simple non-linear geobiochemical non-dimensional model from [4] is presented. It consists of the three constituents y_1, y_2 y y_3, which might be interpreted as nutrients, phytoplankton and detritus, respectively. The model is given by

$$y_1'(t) = -\frac{y_1(t)y_2(t)}{y_1(t)+1}, \tag{5.6a}$$

$$y_2'(t) = \frac{y_1(t)y_2(t)}{y_1(t)+1} - ay_2(t), \tag{5.6b}$$

$$y_3'(t) = ay_2(t). \tag{5.6c}$$

Here the production and destruction terms are defined as follows:

$$p_{21}(t, \mathbf{y}(t)) = d_{12}(t, \mathbf{y}(t)) = \frac{y_1(t)y_2(t)}{y_1(t)+1}, \quad p_{32}(t, \mathbf{y}(t)) = d_{23}(t, \mathbf{y}(t)) = ay_2(t),$$

The system represents a biogeochemical model for the description of an algal bloom, that transforms nutrients (y_1) via phytoplankton (y_2) into detritus (y_3). Phytoplankton (y_2) is then lost by mortality and zooplankton grazing at a fixed non-dimensional rate a, which is collected in the detritus pool. This simple system of equations might be interpreted as a geobiochemical model for the upper oceanic layer in spring, when nutrient rich surface water is captured in the euphotic zone.
In our numerical simulations we use $a = 0.3$ and the initial conditions $y_1(0) = 9.98$, $y_2(0) = 0.01$ and $y_3(0) = 0.01$. The solution is approximated on the time interval $[0, 30]$.

5.2.3 Robertson test problem

One of the most prominent examples of stiff ODEs is the Robertson test problem [13], which describes the chemical reactions

$$y_1'(t) = 10^4 y_2(t)y_3(t) - 4 \times 10^{-2} y_1(t), \tag{5.7a}$$

$$y_2'(t) = 4 \times 10^{-2} y_1(t) - 10^4 y_2(t)y_3(t) - 3 \times 10^7 y_2^2(t) \tag{5.7b}$$

$$y_3'(t) = 3 \times 10^7 y_2^2(t) \tag{5.7c}$$

with initial values $y_i(0) = y_i^0 > 0$ for $i = 1, 2, 3$. In this problem, the production and destruction terms are

$$p_{12}(t, \mathbf{y}(t)) = d_{21}(t, \mathbf{y}(t)) = 10^4 y_2(t)y_3(t), \quad p_{21}(t, \mathbf{y}(t)) = d_{12}(t, \mathbf{y}(t)) = 4 \times 10^{-2} y_1(t),$$
$$p_{32}(t, \mathbf{y}(t)) = d_{23}(t, \mathbf{y}(t)) = 3 \times 10^7 y_2^2(t),$$

and $p_{ij}(t, \mathbf{y}(t)) = d_{ji}(t, \mathbf{y}(t)) = 0$ for other sets of i, j
In the numerical simulations, we take $y_1(0) = 1$ and $y_2(0) = y_3(0) = 0$. Following [23], the time step size in the kth time step is chosen as $\Delta t_k = 2^{k-1}\Delta t_0$ with $\Delta t_0 = 10^{-6}$. The time interval of interest is $[10^{-6}, 10^{10}]$ and to visualize the evolution of y_2, it was multiplied by 10^4.

5.3 Numerical results for second-order scheme

We apply the MPRKO22(α, β) (3.29) schemes to the three test problems introduced in section 5.1. For each problem a parameter study is provided to find a close approximation to the optimal (α, β) pair. In all three cases the optimal pair contains a nonzero parameter β, which proves that MPRKO schemes can be more accurate than MPRK schemes. In addition, we confirm numerically that MPRKO22(α, β) schemes are second order accurate.

To avoid division by zero errors whenever zero initial values are given, we replace the PWDs (3.28) by

$$\pi_i = y_i^n + C_0\Delta t, \quad \sigma_i = (y_i^n)^{1-\frac{1}{\alpha}} \left(y_i^{(2)}\right)^{\frac{1}{\alpha}} + C_0\Delta t^2,$$

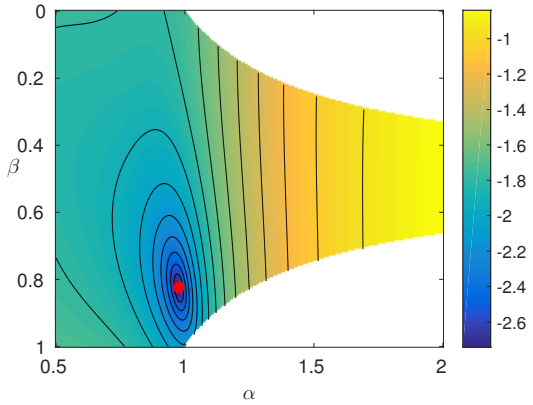
with some constant C_0. Thereby, multiplying by Δt and Δt^2 is necessary to keep the second-order accuracy of the scheme as shown from Theorem 1. In all numerical computations of this section we use $C_0 = \texttt{eps} \approx 2.2204 \times 10^{-16}$.

5.3.1 Parameter study

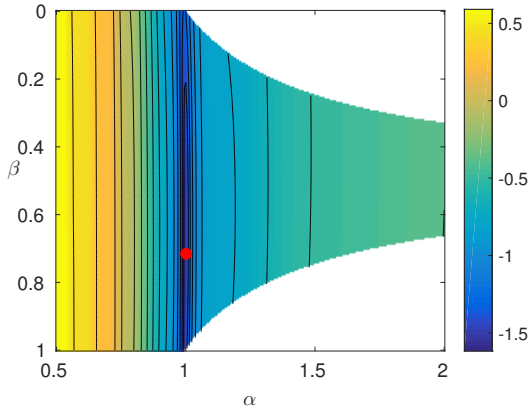
Table 5.4: Relative errors (E) and numerical convergence order (p) of optimal MPRKO schemes applied to the simple linear test problem (5.2).

Δt	MPRKO22$(0.975, 0.825)$		MPRKO22$(1, 0.715)$		MPRKO22$(0.69, 0.5)$	
	E	p	E	p	E	p
7.0312e-01	2.9742e-05	–	6.3958e-05	–	1.1706e-04	–
3.5156e-01	7.9360e-06	1.9060	1.6821e-05	1.9268	3.2521e-05	1.8478
1.7578e-01	2.0525e-06	1.9510	4.3231e-06	1.9602	8.8918e-06	1.8708
8.7891e-02	5.2143e-07	1.9769	1.0967e-06	1.9789	2.3952e-06	1.8923
4.3945e-02	1.3094e-07	1.9936	2.7612e-07	1.9899	6.3632e-07	1.9123
2.1973e-02	3.2487e-08	2.0109	6.9070e-08	1.9991	1.6690e-07	1.9308
1.0986e-02	7.8337e-09	2.0521	1.7057e-08	2.0177	4.3285e-08	1.9470
5.4932e-03	1.6863e-09	2.2159	4.0221e-09	2.0843	1.1151e-08	1.9568

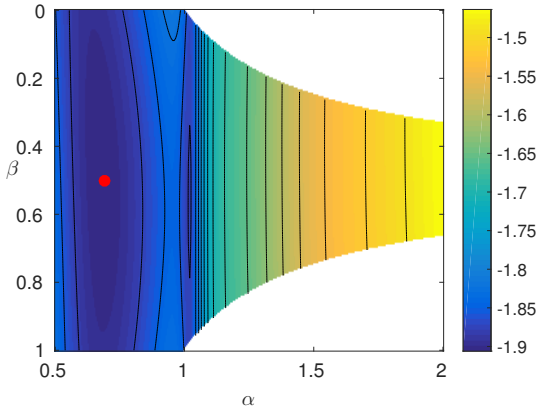
In order to see the dependence of the error E of an MPRKO22(α, β) scheme (5.1) on the parameters α and β, we perform parameter studies for the three test cases



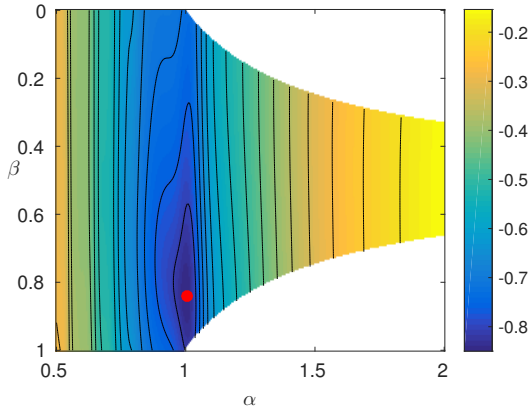
(a) Simple linear test (5.2): $(\alpha_{\text{opt}}, \beta_{\text{opt}}) = (0.975, 0.825)$.



(b) JAK2/STAT5 model (5.3): $(\alpha_{\text{opt}}, \beta_{\text{opt}}) = (1, 0.715)$



(c) SEIR model (5.4): $(\alpha_{\text{opt}}, \beta_{\text{opt}}) = (0.69, 0.5)$



(d) Relative mean error: $(\alpha_{\text{opt}}, \beta_{\text{opt}}) = (1.005, 0.84)$

Figure 5.2: $\log_{10} E$ for the three test cases (5.2), (5.3) and (5.4) with $(\alpha, \beta) \in \left([\frac{1}{2}, 2] \times [0, 1]\right) \cap \mathcal{FR}$. The red dots indicate the pairs (α, β) with the lowest relative errors.

(5.2), (5.3) and (5.4). For this purpose, we discretize the rectangle $[\frac{1}{2}, 2] \times [0, 1]$ with 301 equidistant grid points in direction of α and 201 equidistant grid points in direction of β, resulting in a mesh with 60501 grid points. Out of these, 42213 grid points are contained in the feasible region $\mathcal{FR}$. For each of these grid points (α, β) we compute the relative error of MPRKO22(α, β) applied to a specific test cases. In the following, we refer to the pair (α, β) with the lowest relative error as the optimal parameter pair $(\alpha_{\text{opt}}, \beta_{\text{opt}})$. Indeed, this optimality is based on the selected grid.

Figures 5.2 (a)–(c) show the relative errors for the three test cases (5.2), (5.3) and (5.4). The red dots indicate the optimal pairs $(\alpha_{\text{opt}}, \beta_{\text{opt}})$. We clearly see a dependence on the parameter β, which shows that MPRKO schemes are more accurate than the MPRK schemes, which are restricted to $\beta = 0$. For the sim-

ple linear test (5.2), we have $(\alpha_{\text{opt}}, \beta_{\text{opt}}) = (0.975, 0.825)$ with relative error $E = 0.0018$. The MPRK22 scheme with the lowest relative error is MPRK22(0.855) which generates a relative error $E = 0.01580$, which is about 8.7 times larger than the one of MPRKO22(0.975, 0.825). In case of the JAK2/STAT5 model (5.3) we find $(\alpha_{\text{opt}}, \beta_{\text{opt}}) = (1, 0.715)$. The relative error of MPRKO22(1, 0.715) is $E = 0.0241$. The MPRK22 scheme with the smallest relative error is MPRK22(0.995) for which we have $E = 0.0438$. Thus, the optimal relative error of the MPRKO schemes is about half the size of the smallest error with respect to MPRK22. The optimal parameter pair for the SEIR model (5.4) is given by $(\alpha_{\text{opt}}, \beta_{\text{opt}}) = (0.69, 0.5)$ for which we obtain a relative error $E = 0.0124$, for MPRK22(0.65) the relative error is $E = 0.0127$. In this test case there is no essential benefit by using MPRKO instead of MPRK schemes.

From 5.2 (a)–(c), we observe that the optimal pair (α, β) depends on the test case under consideration. To find an appropriate pair (α, β) for all test cases, we also consider a relative mean error defined as

$$E_{\text{mean}}(\alpha, \beta) = \frac{1}{3}\left(\frac{\widehat{E}_1(\alpha, \beta)}{\max_{(\alpha,\beta)} \widehat{E}_1(\alpha, \beta)} + \frac{\widehat{E}_2(\alpha, \beta)}{\max_{(\alpha,\beta)} \widehat{E}_2(\alpha, \beta)} + \frac{\widehat{E}_3(\alpha, \beta)}{\max_{(\alpha,\beta)} \widehat{E}_3(\alpha, \beta)}\right)$$

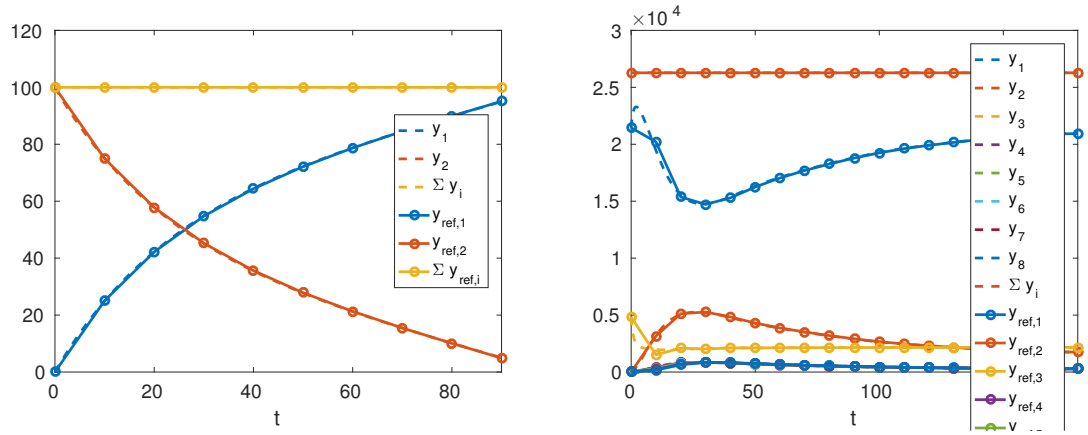
where $\widehat{E}_1, \widehat{E}_2, \widehat{E}_3$ denote the relative errors of the three test cases. This mean error is depicted in Figure 5.2 (d) and minimized for $(\alpha_{\text{opt}}, \beta_{\text{opt}}) = (1.005, 0.84)$. Hence, $\beta \neq 0$ is a proper choice for all test cases.

5.3.2 Convergence order

Next, we confirm the theoretical convergence order numerically. In order to do so, we compute approximations of the simple linear test (5.2) with the three MPRKO schemes identified in the parameter study of Section 5.3.1 and varying time step sizes Δt. Let $\widetilde{E}_i$ represent the local error corresponding to the step size Δt_i, then we compute the numerical convergence order p_i as

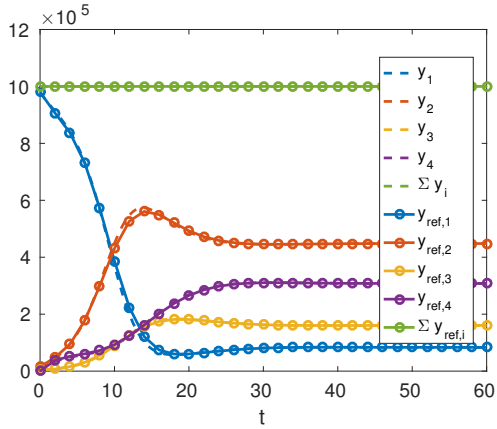
$$p_i = \log\left(\frac{\widetilde{E}_{i+1}}{\widetilde{E}_i}\right) \Big/ \log\left(\frac{\Delta t_{i+1}}{\Delta t_i}\right).$$

Table 5.4 lists the relative errors and numerical convergence orders obtained with the specified time step sizes. We see that all three MPRKO schemes are second order accurate as expected. Moreover, among these three test problems, MPRKO22(0.975, 0.825) for the linear test generates the least relative error for all time step sizes.
We apply the optimal MPRKO schemes to each of the three test cases (5.2), (5.3) and (5.4) , to show that the numerical solutions are reasonable, positive and conservative approximations. The results are depicted in Figure 5.3. As expected, the numerical solutions are positive and conservative. Moreover, we see that in each case the numerical solutions provide good approximations of the reference solution.



(a) Simple linear test (5.2) integrated with MPRKO22$(0.975, 0.825)$.

(b) JAK2/STAT5 model (5.3) integrated with MPRKO22$(1, 0.715)$.



(c) SEIR model (5.4) integrated with MPRKO22$(0.69, 0.5)$.

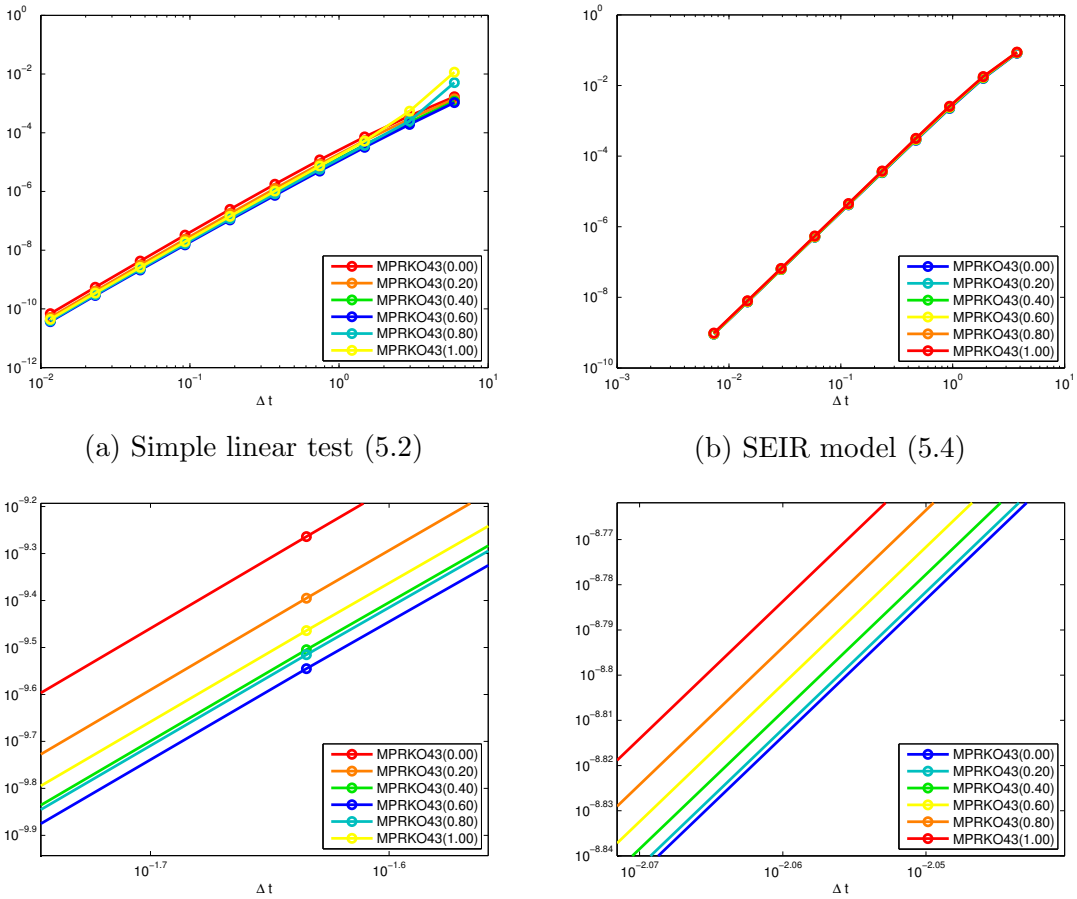
Figure 5.3: Numerical solutions for the three test cases (5.2), (5.3) and (5.4) computed with the corresponding optimal MPRKO22(α, β) schemes.

5.4 Numerical results for third-order scheme

In this section, we confirm the theoretical convergence order of the MPRKO43 schemes introduced in the preceding sections. To assess the order of the MPRKO43 schemes, we use a relative error E defined in section 5.1

5.4.1 Convergence order

Figure 5.4 shows error plots of six MPRKO43 schemes applied to the simple linear test problem (5.2) Fig. 5.4a and SEIR model (5.4) Fig. 5.4b. The parameter β takes the values $\beta_k = 0.2k$, with $k = 0, 1, \ldots, 5$, in two cases, In all cases the third



(a) Simple linear test (5.2)

(b) SEIR model (5.4)

(c) Simple linear test (5.2) zoom in

(d) SEIR model (5.4) zoom in

Figure 5.4: $\log_{10} E$ error plots of MPRKO43 schemes for various values of β for the two test cases (5.2) and (5.4).

order accuracy is confirmed. Moreover, Fig. 5.4d shows that MPRKO43(1) is less accurate than all MPRKO43 schemes in the case of the SEIR model problem 5.4 and Fig. 5.4c shows that MPRKO43(0.6) is more accurate when applied to the simple linear test problem. Fig. 5.4d shows error plots of all MPRKO43 schemes applied to the SEIR model problem (5.4),we see that the error seems to increase monotonically with the value of β. This property is not shared when applied to the simple linear test problem. We additionally show numerical solutions of the six MPRK43 schemes applied to the SEIR model problem (5.4) in Figure 5.5.

5.4.2 Stiff problems

In the case of a highly stiff problem, we obtained excellent accuracy results of our scheme for all β_k values. For saving space, we present only the numerical results of

Table 5.5: Error table for simple linear test (5.2). MPRKO43(β_a) scheme versus the schemes MPRK43I(1, $\frac{1}{2}$), MPRK43I($\frac{1}{2}$, $\frac{2}{3}$) and MPRK43II($\frac{1}{2}$) in [24]

Δt	MPRKO43(β_a)		MPRK43I(1, $\frac{1}{2}$)		MPRK43I($\frac{1}{2}$, $\frac{2}{3}$)		MPRK43II($\frac{1}{2}$)	
	Error	p	Error	p	Error	p	Error	p
5,63	**9,73e-04**	–	1,79e-03	–	1,30e-03	–	1,41e-03	–
2,81	**1,75e-04**	2,48	4,09e-04	2,13	2,65e-04	2,30	3,00e-04	2,24
1,41	**2,87e-05**	2,60	7,59e-05	2,43	4,58e-05	2,53	5,32e-05	2,49
0,70	**4,46e-06**	2,69	1,20e-05	2,66	7,02e-06	2,71	8,21e-06	2,70
0,35	**6,64e-07**	2,75	1,70e-06	2,81	9,97e-07	2,82	1,16e-06	2,83
0,18	**9,63e-08**	2,79	2,28e-07	2,90	1,36e-07	2,88	1,55e-07	2,90
0,09	**1,37e-08**	2,82	2,95e-08	2,95	1,81e-08	2,91	2,03e-08	2,94
0,04	**1,90e-09**	2,85	3,75e-09	2,97	2,38e-09	2,93	2,61e-09	2,96
0,02	**2,56e-10**	2,89	4,73e-10	2,99	3,09e-10	2,95	3,32e-10	2,97
0,01	**3,25e-11**	2,97	5,88e-11	3,01	3,86e-11	3,00	4,13e-11	3,01

an MPRKO43(β_a) scheme with β_a equal to the average of the β_k for $k = 0, 1, \ldots, 5$. Figure 5.6 shows numerical approximations of MPRKO43(β_a) scheme applied to the stiff Robertson problem (5.7). We chosen, the time step size in the kth time step as $\Delta t_k = 2^{k-1}\Delta t_0$ with $\Delta t_0 = 10^{-6}$. Hence, only 55 time steps are necessary to cover the time interval $[10^{-6}, 10^{10}]$. The small initial time step was chosen to obtain an adequate resolution of the component y_2 in the starting phase. To visualize the evolution of y_2, it is multiplied by 10^4.

5.4.3 Comparison to MPRK43 schemes

In this section we compares our scheme with the three MPRK43 schemes introduced at the section 2 from [24]. We used the simple linear test (5.5) as a test problem and the Table 5.5 shows that all four schemes generate adequate solutions and the third-order accuracy is clearly obtained. The MPRKO43(β_a) scheme generates the most accurate approximations with the lowest relative error, but we cannot affirm that our scheme is better than their schemes, because there may be parameters in the feasible region of section 2 of [[24], see Figure 2] that could generate better numerical results than our MPRKO43(β_a) scheme. Therefore, such a search for optimal parameters is of great interest and will be a major research topic in the future.

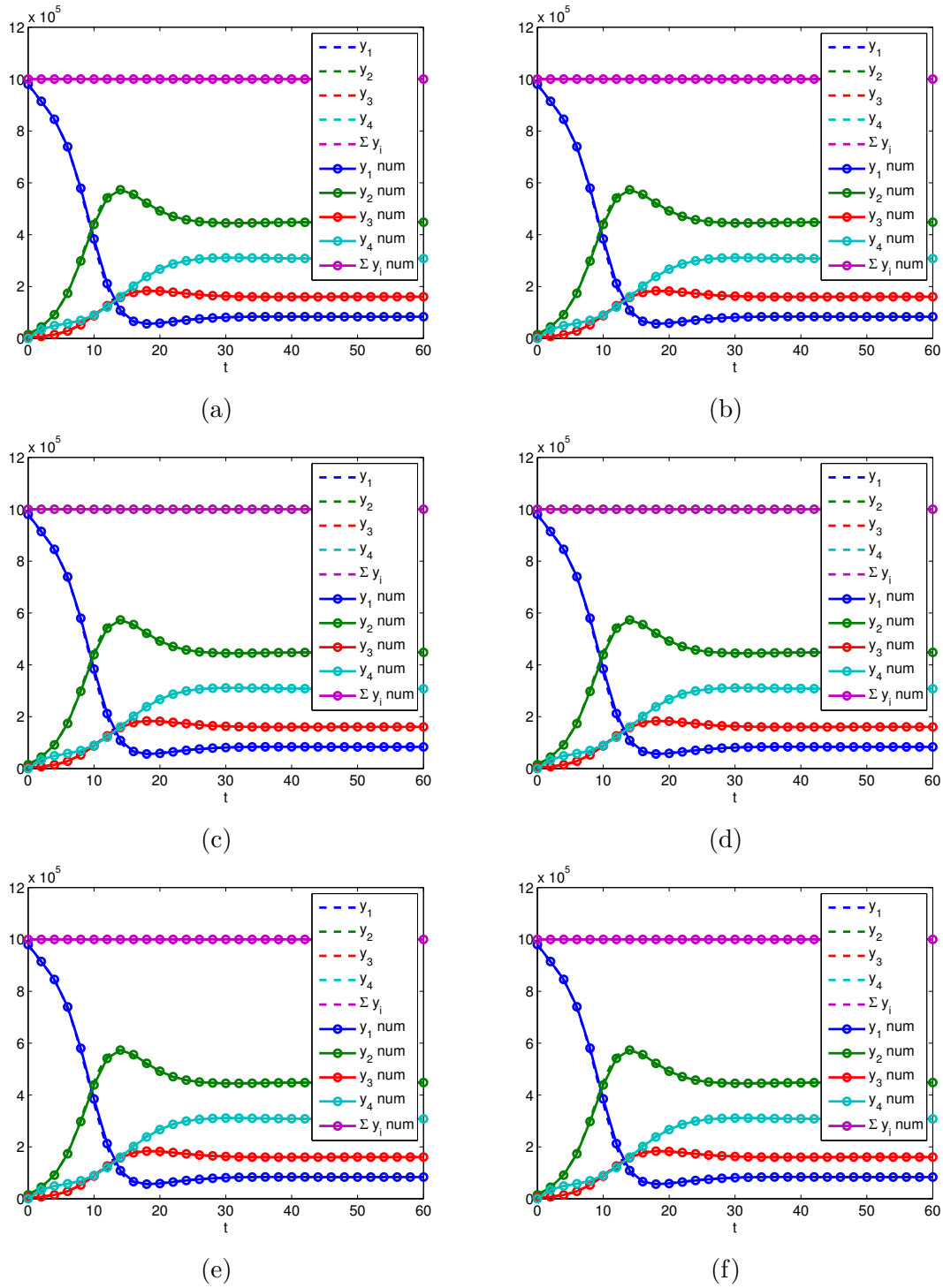


Figure 5.5: Numerical solutions of the SEIR model problem (5.4) for different MPRKO43 schemes with time step size $\Delta t = 0.2$. **(a)** MPRKO43(0). **(b)** MPRKO43(0.2). **(c)** MPRKO43(0.4). **(d)** MPRKO43(0.6). **(e)** MPRKO43(0.8). **(f)** MPRKO43(1).

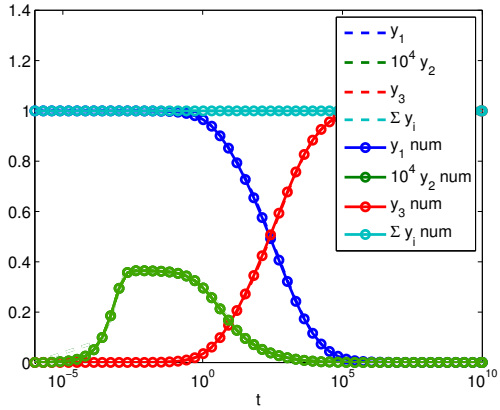


Figure 5.6: Numerical solutions of the Robertson problem (5.7) for MPRKO43(β_a) with $\beta_a = 0.5$.

Chapter 6

Concluding remarks and future work

In this thesis we have studied Modified Patankar Runge Kutta schemes that are numerical methods for the solution of positive and conservative production–destruction systems. The first schemes were introduced in [4] adapting explicit Runge–Kutta schemes to ensure positivity and conservation irrespective of the time step size. Next in [18], instead of using the Runge Kutta schemes in the classical form, they apply the Runge Kutta schemes of the Shu–Osher form and develop another class of MPRK scheme. Another scheme of the MPRK type uses the Deferred Correction method. They proved that the obtained mPDeC schemes are positive preserving, conservative and arbitrary high-order accurate. In contrast, we use approach Oliver to improve the accuracy of these schemes in the field of nonautonomous systems.

6.1 Concluding remarks

We introduced novel two-stage MPRKO schemes, which generalize the MPRK schemes of [23] to integrate nonautonomous PDS and proved their unconditionally positivity and conservation property as well as necessary and sufficient conditions with respect to the PWDs to obtain second and third order accuracy. Additionally, we followed Oliver [30, 38] and allowed Butcher tableaus $(\mathbf{A}, \mathbf{b}, \mathbf{c})$, which do not satisfy $\mathbf{c} = \mathbf{A}\mathbf{e}$ with $\mathbf{e} = (1, \ldots, 1)^T$. Altogether, this led to the introduction of the two-parameter family of MPRKO22(α, β).
The numerical experiments confirm the second order accuracy of MPRKO22(α, β) schemes. Indeed, the set of MPRK22(α) schemes is contained within the set of MPRKO22(α, β) by choosing $\beta = 0$. Hence, we have shown the benefit of the additional β parameter since in all considered test cases the optimal pair (α, β) has a nonzero β component. Furthermore, numerical experiments with autonomous problem did not show improvements respect to MPRK(α).
Since the order conditions of the RKO scheme increase very fast respect to the order,

building schemes of the MPRKO type would not be encouraging for schemes of order higher than four, for example, we must solve 166 equations for the six-satges schemes. We still do not known if there are sixth-order or higher RKO schemes .
We have extended the work of [26] to third-order by deriving necessary and sufficient conditions for three-stage third order schemes. Moreover, this led to the introduction of the one-parameter family of MPRKO43(β) scheme. Similarly to [24], these schemes can be regarded as four stage third order MPRK schemes.
This ODE solver is successfully applied to integrate nonautonomous and autonomous PDS and proved their unconditionally positivity and conservation property. The numerical experiments have shown that the MPRKO43 schemes are capable of integrating stiff ODEs, such as the Robertson problem. The numerical experiments confirm the third-order accuracy of MPRKO43(β) schemes.

6.2 Future Work

In general, we are interested in the stability of MPRK schemes. Since the usual approach by means of Dahlquist's equation is not feasible, an analytic investigation of the stability is still missing. One probable scenario that are of interest in future research in the context of the analytic investigation of the stability problems is to study the Lyapunov stability of the MPRK scheme.
Another interesting research topic is to use a convex combination of the PWDs to find other second order MPRK schemes. Following [23], we take

$$\pi_i = y_i^n, \quad \sigma_i = \omega y_i^n \left(\frac{y_i^{(2)}}{y_i^n}\right)^{\beta_1} + (1-\omega) y_i^n \left(\frac{y_i^{(2)}}{y_i^n}\right)^{\beta_2}, \qquad i = 1, \ldots, N,$$

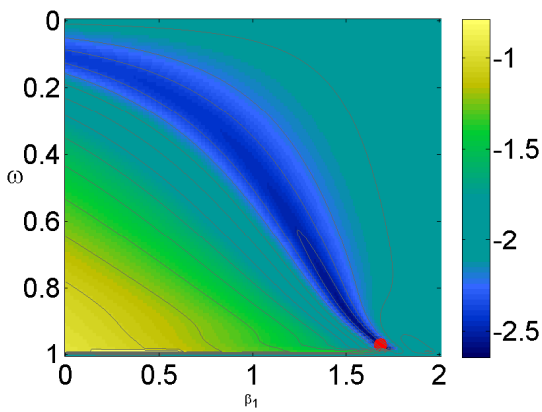
with $0 \le \omega \le 1$, and $\beta_2 = \frac{\alpha\omega\beta_1 - 1}{\alpha(\omega-1)}$.
Combined with the MPRK(α) scheme, we get a two-parameter family of second-order schemes, denoted by WMPRK22(β_1, ω) and defined by

$$y_i^{(1)} = y_i^n, \tag{6.1a}$$

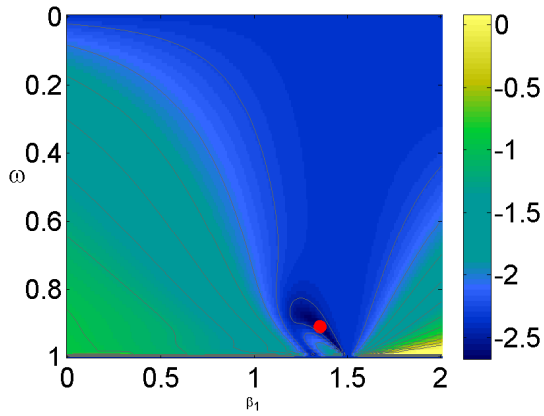
$$y_i^{(2)} = y_i^n + \alpha\Delta t \sum_{j=1}^{N} \left(p_{ij}(\mathbf{y}^{(1)}) \frac{y_j^{(2)}}{y_j^{(1)}} - d_{ij}(\mathbf{y}^{(1)}) \frac{y_i^{(2)}}{y_i^{(1)}} \right), \tag{6.1b}$$

$$\begin{aligned} y_i^{n+1} = y_i^n + \Delta t \sum_{j=1}^{N} \Bigg[& \left((1 - \frac{1}{2\alpha}) p_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} p_{ij}(\mathbf{y}^{(2)}) \right) \frac{y_j^{n+1}}{\omega (y_j^{(2)})^{\beta_1} (y_j^n)^{1-\beta_1} + (1-\omega)(y_j^{(2)})^{\beta_2} (y_j^n)^{1-\beta_2}} \\ & - \left((1 - \frac{1}{2\alpha}) d_{ij}(\mathbf{y}^{(1)}) + \frac{1}{2\alpha} d_{ij}(\mathbf{y}^{(2)}) \right) \frac{y_i^{n+1}}{\omega (y_j^{(2)})^{\beta_1} (y_j^n)^{1-\beta_1} + (1-\omega)(y_j^{(2)})^{\beta_2} (y_j^n)^{1-\beta_2}} \Bigg] \end{aligned} \tag{6.1c}$$

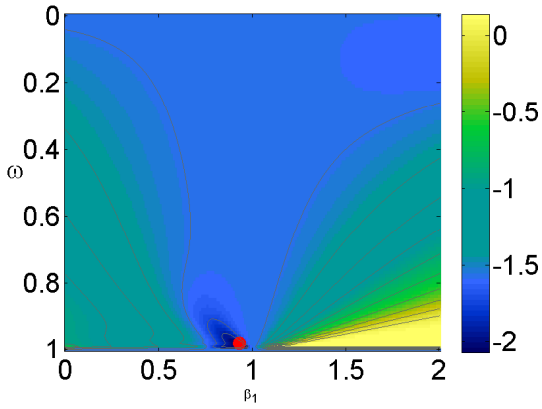
for $i = 1, \ldots, N$.
To understand the dependence of the error E (5.1) of an WMPRK22(β_1, ω) scheme



(a) $(\beta_1\text{opt}, \omega\text{opt}) = (1.68, 0.97)$ for $\alpha = \frac{1}{2}$



(b) $(\beta_1\text{opt}, \omega\text{opt}) = (1.35, 0.91)$ for $\alpha = \frac{2}{3}$



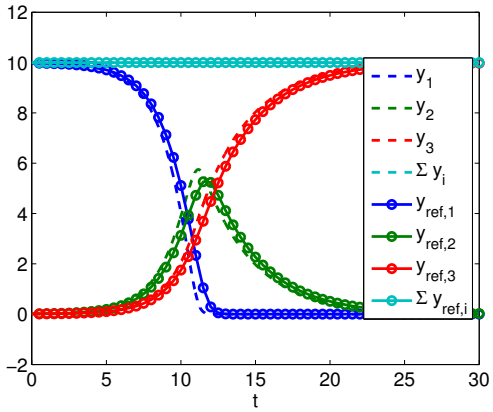
(c) $(\beta_1\text{opt}, \omega\text{opt}) = (0.93, 0.98)$ for $\alpha = 1$

Figure 6.1: $\log_{10} E$ error for the test case (5.5) for $\alpha \in \{\frac{1}{2}, \frac{2}{3}, 1\}$ with $(\beta_1, \omega) \in \big([0,2] \times [0,1]\big)$. The red dots indicate the pairs (β_1, ω) with the lowest relative errors.

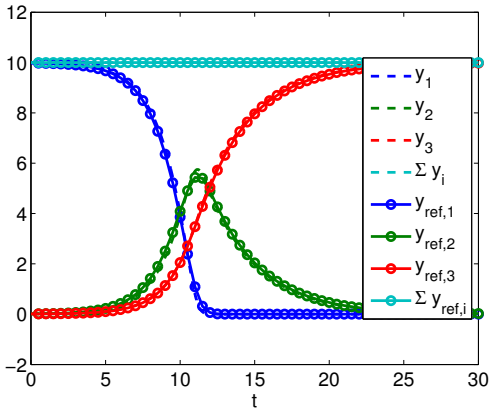
on the parameters β_1 and ω, we perform parameter studies for the test case (5.5) for $\alpha \in \{\frac{1}{2}, \frac{2}{3}, 1\}$. For this purpose, we discretize the rectangle $[0,2] \times [0,1]$ with 201 equidistant grid points in direction of β_1 and 101 equidistant grid points in direction of ω resulting in a mesh with 20301 grid points. We compute the relative error of WMPRK22(β_1, ω) for each of these grid points (β_1, ω) applied to a specific test case. In the following, we refer to the pair (β_1, ω) with the lowest relative error as the optimal parameter pair $(\beta_{1\text{opt}}, \omega_{\text{opt}})$. Of course, this optimality is based on the grid used. Figures 6.1 (a)–(c) show the relative errors for the test case (5.5), for $\alpha \in \{\frac{1}{2}, \frac{2}{3}, 1\}$. The red dots indicate the optimal pairs $(\beta_{1\text{opt}}, \omega_{\text{opt}})$. We clearly see a dependence on the parameters (β_1, ω), which shows that WMPRK22 schemes can be more accurate than the MPRK22 schemes.

We apply the optimal MPRKW22 schemes to nolinear test case (5.6) for $\alpha \in \{\frac{1}{2}, \frac{2}{3}, 1\}$, to show that the numerical solutions are positive and conservative approximations.

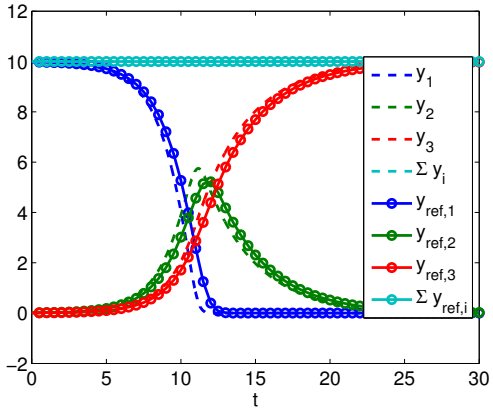
The results are shown in Figures 6.2, we see that in each case the numerical solutions of WMPRK22 (right column) is more accurate than numerical solutions of MPRK22 (left column) for $\alpha \in \{\frac{1}{2}, \frac{2}{3}, 1\}$. Therefore, we propose to study: The effect of the convex combination of PWDs in the MPRKO22(α, β) scheme.



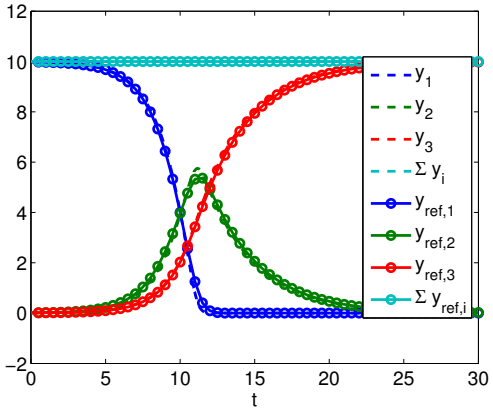
(a) $MPRK22(\frac{1}{2})$



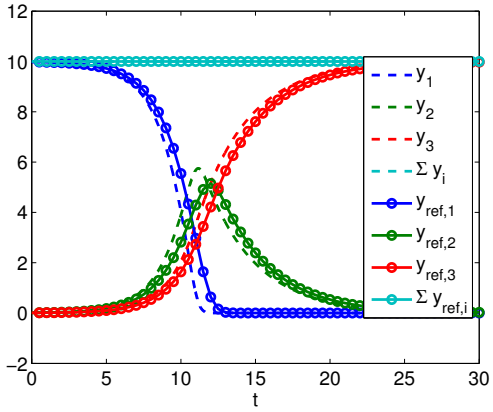
(b) $WMPRK22(1.68, 0.97)$ for $\alpha = \frac{1}{2}$



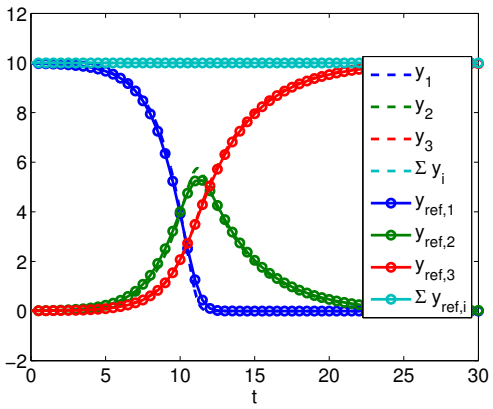
(c) $MPRK22(\frac{2}{3})$



(d) $WMPRK22(1.35, 0.91)$ for $\alpha = \frac{2}{3}$



(e) $MPRK22(1)$



(f) $WMPRK22(0.93, 0.98)$ for $\alpha = 1$

Figure 6.2: Numerical solutions of the nonlinear test problem (5.6) for different MPRK22 and WMPRK22 schemes

References

[1] E. Bertolazzi, Positive and Conservative Schemes for Mass Action Kinetics, Computers Math. Applic. 32 (1996) 29–43.

[2] C. Bolley, M. Crouzeix, Conservation de la positivité lors de la discrétisation des problèmes d'évolution paraboliques, ESAIM Math. Model. Numer. Anal. 12 (1978) 237–245.

[3] J. Bruggeman, H. Burchard, B. W. Kool, B. Sommelier, A second-order, unconditionally positive, mass conserving integration scheme for biochemical systems, Appl. Numer. Math. 57 (2007) 36–58.

[4] H. Burchard, E. Deleersnijder, A. Meister, A high-order conservative Patankar-type discretisation for stiff systems of production–destruction equations, Appl. Numer. Math. 47 (2003) 1–30.

[5] H. Burchard, E. Deleersnijder, A. Meister, Application of modified Patankar schemes to stiff biogeochemical models for the water column, Ocean Dyn. 55 (2005) 326–337.

[6] H. Burchard, K. Bolding, W. Kuhn, A. Meister, T. Neumann, L. Umlauf, Description of a flexible and extendable physical–biogeochemical model system for the water column, J. Marine Systems 61 (2006) 180–211.

[7] J. C. Butcher, Numerical methods for ordinary differential equations, Third Edition, Wiley, (2016).

[8] M. De la Sen, S. Alonso-Quesada, Vaccination strategies based on feedback control techniques for a general SEIR-epidemic model, App. Math. and Comp. 208 (2011) 3888–3904.

[9] E. Friedmann, F. Neumann, R. Rannacher, Well-Posedness of a Linear Spatio-Temporal Model of the JAK2/STAT5 Signaling Pathway, Comm. in Math. Anal. 15 (2013) 76–102.

[10] S. Gottlieb, D. I. Ketcheson, C.-W. Shu, Strong stability preserving Runge–Kutta and multistep time discretizations, World Scientific (2011).

[11] O. Gressel, Toward realistic simulations of magneto-thermal winds from weakly-ionized protoplanetary disks, J. Phys. Conf. Ser. 837 (2017) 012008.

[12] E. Hairer, S. P. Norsett, G. Wanner, Solving ordinary differential equations I. Nonstiff problems, 8th Edition, vol. 8 of Springer Series in Comp. Math. Springer-Verlag, Berlin, (2000).

[13] E. Hairer, G. Wanner, Solving ordinary differential equations II. Stiff problems, 8th Edition, volume 14 of Springer Series in Comp. Math. Springer-Verlag, Berlin, (2004).

[14] I. Hense, A. Beckmann, The representation of cyanobacteria life cycle processes in aquatic ecosystem models, Ecol. Model. 221 (2010) 2330–2338.

[15] I. Hense, H. Burchard, Modelling cyanobacteria in shallow coastal seas, Ecol. Model. 221 (2010) 238–244.

[16] Z. Hórvath, Positivity of Runge-Kutta and diagonally split Runge-Kutta methods, Appl. Numer. Math. 28 (1998) 309–326.

[17] Z. Hórvath, On the positivity step size threshold of Runge-Kutta methods, Appl. Numer. Math. 53 (2005) 341–356.

[18] J. Huang, CW. Shu, Positivity-preserving time discretizations for production–destruction equations with applications to non-equilibrium flows, J. of Sci. Comp. 78 (2018) 1811–1839.

[19] J. Huang, W. Zhao, CW. Shu, A Third-Order Unconditionally Positivity-Preserving Scheme for production–destruction Equations with Applications to Non-equilibrium Flows, J. of Sci. Comp. 79 (2019) 1015–1056.

[20] W. Hundsdorfer, J. Verwer, Numerical solution of time-dependent advection-diffusion-reaction equations, Springer (2003).

[21] W. Hundsdorfer, J. Verwer, J. Blom, Numerical time integration for air pollution models, Surv. Math. Ind. 10 (2002) 107–174.

[22] J. S. Klar, J. P. Mücket, A detailed view of filaments and sheets in the warm-hot intergalactic medium, Astronom. Astrophys. 522 (2010) A114.

[23] S. Kopecz, A. Meister, On order conditions for modified Patankar Runge–Kutta schemes. App. Num. Math. 123 (2018) 159–179.

[24] S. Kopecz, A. Meister, Unconditional positive and conservative third order modified Patankar Runge Kutta discretizations of production destruction systems, BIT Num. Math. 58 (2018) 694–728.

[25] S. Kopecz, A. Meister, A comparison of numerical methods for conservative and positive advection-diffusion-production–destruction systems, Proc. in App. Math. and Mech. 19 (2019) 02–09.

[26] A. Ávila, G. González, S. Kopecz, A. Meister, Extension of modified Patankar–Runge–Kutta schemes to nonautonomous production–destruction systems based on Oliver's approach, Jour. Compu. Appl. Math. 389 (2020) 20–33.

[27] S. Kopecz, A. Meister, On the existence of three-stage third-order modified Patankar–Runge–Kutta schemes, Numer. Algorithms 81 (2019) 1473–1484.

[28] P. Öffner, D. Torlo,Arbitrary high-order, conservative and positivity preserving Patankar-type deferred correction schemes, Appl. Numer. Math. 153 (2020) 15–34.

[29] C. Shu, S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, J. Comput. Phys. 77 (1988) 439–471.

[30] J. Oliver, A curiosity of low-order explicit Runge–Kutta methods, Math. Comput. 29 (1975) 1032–1036.

[31] S.V. Patankar, Numerical Heat Transfer and Fluid Flow, McGraw-Hill, New York, (1980).

[32] A. Sandu, Positive Numerical Integration Methods for Chemical Kinetic Systems, J. Comput. Phys., 170 (2001) 589-602.

[33] A. Sandu, Time-stepping methods that favor positivity for atmospheric chemistry modeling, Atmospheric Modeling, IMA Vol. Math. Appl. 130, Springer (2002) 21–37.

[34] K. Semeniuk, A. Dastoor, Development of a global ocean mercury model with a methylation cycle: outstanding issues, Glob. Biogeochem. Cycles 31 (2017) 400–433.

[35] L.F. Shampine, M.W. Reichelt, The MATLAB ODE suite. SIAM J. Sci. Comput. 18 (1997) 1–22.

[36] L.F. Shampine, Conservation laws and the numerical solution of ODEs, Comp. & Maths. with Appls. 12 (1986) 1287 – 1296

[37] L.F. Shampine, S. Thompson, J. A. Kierzenka, G.D. Byrne, Non-negative solutions of ODEs, Appl. Math. Comput. 170 (2005) 556–569.

[38] Ch. Tsitouras, Explicit Runge–Kutta methods for starting integration of Lane–Emden problem, Appl. Math. and Comp. 354 (2019) 353–364.

[39] I. Famelis, S. Papakostas, Ch. Tsitouras, Symbolic derivation of Runge–Kutta order conditions, Journal of Symbolic Computation. Elsevier, 37 (2004) 311–327.

[40] A. Warns, I. Hense, A. Kremp, Modelling the life cycle of dinoflagellates: a case study with biecheleria baltica, J. Plankton Res. 35 (2013) 379–392.

[41] D. Zill, M. Cullen, Differential Equations with Boundary-Value Problems, 4th Edition, Brooks Cole Publishing Company. An International Thomson Publishing Company, (2012).