Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 18, No. 4/2026

TRANSIENT FLOW OF AN EXPONENTIALLY VISCOUS FLUID IN A RECTANGULAR DOMAIN*

Dumitru Vieru† Nehad Ali Shah‡

Dedicated to Professor Constantin Fetecau on the occasion of his 80th anniversary

DOI 10.56082/annalsarscimath.2026.4.245

Abstract

This paper investigates the motion of an advanced shear-thinning model for paint. The study examines the liquid's flow in a rectangular domain bounded by two parallel flat plates subjected to Robin type boundary conditions. The governing equations are integrated numerically for both pressure driven and pressure gradient free flows. Under certain conditions relating the material coefficients to the velocity gradient, exact analytical solutions are also derived. The impact of material properties on paint behavior was also assessed.

Keywords: paint, shear-thinning fluid, analytical solutions, finite sine-cosine Fourier transform.

MSC: 76Axx.

1 Introduction

Paints are non-Newtonian systems with complex rheological behavior consisting of a continuous liquid phase in which solid and liquid particles are

*Accepted for publication on September 22, 2026

†dumitru.vieru@academic.tuiasi.ro, Department of Theoretical Mechanics, *Gheorghe Asachi* Technical University of Iasi, Romania

‡nehadali199@yahoo.com, Department of Mechanical Engineering, Sejong University, Seoul 05006, Republic of Korea

dispersed. An elegant presentation of these materials is given by Tadros [16]. The rheological behavior of non-Newtonian fluids plays a critical role in numerous industrial and biological applications. Among these, fluids that exhibit a rapid increase in viscosity at low shear rates present unique modeling challenges. Traditional formulations, such as the classical Bingham plastic or Herschel Bulkley models, rely on a strict yield stress threshold. While these models are mathematically convenient, they introduce discontinuities that complicate numerical simulations. Furthermore, physical observations often reveal that many real world materials do not possess a true yield stress, but rather behave as highly viscous fluids under minimal deformation.

To address these limitations, recent research has focused on continuous viscous formulations that mimic yield behavior without mathematical discontinuities. Notably, Garimella et al. [10] proposed a novel two-parameter mathematical model specifically designed for these "apparently viscous" fluids. Under simple shear flow, this model captures a distinct physical response: as the shear rate approaches zero, the apparent viscosity tends to infinity, while the shear stress asymptotically approaches a finite value. The generalized viscosity function features a sufficiently steep gradient relative to the shear rate. This mathematical structure effectively replicates the flow characteristics of complex commercial and biological fluids, such as paints, and meat extracts.

Building upon their theoretical framework, Garimella et al. [10] investigated the steady flow of the apparently viscous fluids within a cylindrical pipe. Their study provided analytical and numerical insights into the velocity profiles, pressure drops, and shear stress distributions inherent to this fluid model. Furthermore, they validated their findings by comparing the results against established benchmarks solutions available in the specialized literature.

Almqvist et al. [1] derived a generalized mathematical model for pressure-driven flows of viscoplastic materials between parallel plates. The novelty of this approach lies in modeling fluid response using constitutive relations where the symmetric part of the velocity gradient is expressed directly as a function of the flow's additional stress. In related research, Anand and Gnaneshwarii [2] studied the generalized second Stokes problem for shear thinning paint fluids, providing numerical solutions for boundaries subjected to oscillatory velocity or shear stress.

The linearized stability of the flow of non-Newtonian fluids such as paints or food products was studied by Fusi and Rajagopal [9]. The flow occurs over a porous solid plate subjected to suction at the plate level. The authors showed that the suction has the effect of decreasing the thickness of

the boundary layer and stabilizing the flow. Also, the flows of fluids with shear thinning are more stable than the corresponding Navier-Stokes fluids. The transient rectilinear flow of a fluid whose constitutive equation mimics that of a viscoplastic material (boundary stress fluid) driven by a horizontal pressure gradient and the motion of the upper plate was studied by Fusi et al. [8] using the finite element method. They showed that the solution of the problem exists and is unique for a limited range of values of the material parameters.

Papaioannou et al. [14] examined the transient injection of a viscoplastic material, driven by a constant pressure drop, inside a pipe or between two parallel coaxial disks under creeping flow conditions by utilizing the papanastasiou constitutive equation. The problem solutions are determined using the mixed finite element/Galerkin formulation coupled with the implicit Euler method. Because the fluid is viscoplastic, the flow field changes qualitatively from that of a Newtonian fluid, primarily because the material detaches from the walls. To highlight the differences in particle motion between non-Newtonian and Newtonian fluids, Datt and Elfring [4] analyzed the dynamic changes in several problems involving rigid spherical particles in shear thinning fluids in the absence of inertia. The authors calculated the first correction for suspension viscosity-the Einstein viscosity-for a dilute suspension of spheres in a weak shear-thinning fluid.

Interesting results regarding the flow of fluids with variable properties, more precisely fluids with pressure-dependent viscosity, can be found in the articles [6, 7, 17]. The present paper formulates and evaluates a novel shear-thinning constitutive model for paint. We analyze the unsteady flow of the liquid constrained within the rectangular domain bounded by two flat parallel plates/ The governing equations are solved numerically subject to well-posed initial and boundary conditions. We investigate both cases: flow with and without an applied pressure gradient. Furthermore, under specific constraints on the material coefficients and the velocity gradient, exact analytical solutions are derived using the Laplace and finite sine cosine Fourier transforms.

To determine the range of validity for the proposed model, a comparative analysis between the analytical and numerical solutions was conducted. The numerical results were found to be in excellent agreement with the analytical predictions, serving as a robust validation of the obtained framework. Furthermore, the numerical solution provides a systematic method to identify the specific conditions and parametric thresholds under which certain approximations of the nonlinear governing equation can be justified. By applying these specific limits, the nonlinear system can be accurately

linearized and transformed into a partial differential equation enabling to analytical treatment.

2 Problem formulation

We consider the movement of paint (a shear thinning fluid) in the infinite domain delimited by two parallel plane plates

$$D = \{(X, Y, Z) \in R^3, -\infty < X, Z < \infty, 0 \leq Y \leq H\}$$

with respect to the Cartesian coordinate system OXYZ (the two plates that delimit the domain are located in the planes $Y = 0, Y = H$, respectively). Let $\vec{V} = (U, V, W)$ be the velocity field. The constitutive equation of the considered fluid is given by [2, 10]

$$\underset{\sim}{\overline{T}} = -\overline{p}\,\underset{\sim}{I} + 2\overline{\mu}\left(\left\|\underset{\sim}{\overline{D}}\right\|\right)\underset{\sim}{\overline{D}}, \tag{1}$$

where, $\underset{\sim}{\overline{T}}$ is the Chauchy stress tensor, $\overline{p}$ is the mechanical pressure, $\underset{\sim}{I}$ is the unit tensor, and $\overline{\mu}\left(\left\|\underset{\sim}{\overline{D}}\right\|\right)$ is the generalized fluid viscosity, which is a function of the Frobenius norm of the symmetric part of velocity gradient $\nabla\vec{V}$ defined as

$$\underset{\sim}{\overline{D}} = \frac{1}{2}\left(\nabla\vec{V} + \left(\nabla\vec{V}\right)^T\right), \qquad \left\|\underset{\sim}{\overline{D}}\right\| = \left(\operatorname{trace}\underset{\sim}{\overline{D}}^{\,2}\right)^{1/2}. \tag{2}$$

In this paper, we consider the generalized viscosity given by [2, 10]

$$\overline{\mu}\left(\left\|\underset{\sim}{\overline{D}}\right\|\right) = \frac{\overline{\alpha}_1}{\left\|\underset{\sim}{\overline{D}}\right\|}\left[1 - \exp\left(-\overline{\alpha}_2\left\|\underset{\sim}{\overline{D}}\right\|\right)\right], \tag{3}$$

where $\overline{\alpha}_1$ and $\overline{\alpha}_2$ are positive constants. Supposing the velocity field and pressure of the form

$$\vec{V} = U(Y, \bar{t})\,\vec{e}_z, \qquad \bar{p} = \bar{p}(X, Y, \bar{t}), \tag{4}$$

velocity gradient is given by [7]

$$\nabla\vec{V} = \begin{bmatrix} \dfrac{\partial U}{\partial X} & \dfrac{\partial U}{\partial Y} & \dfrac{\partial U}{\partial Z} \\ \dfrac{\partial V}{\partial X} & \dfrac{\partial V}{\partial Y} & \dfrac{\partial V}{\partial Z} \\ \dfrac{\partial W}{\partial X} & \dfrac{\partial W}{\partial Y} & \dfrac{\partial W}{\partial Z} \end{bmatrix} = \begin{bmatrix} 0 & \dfrac{\partial U}{\partial Y} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{5}$$

In addition, we assume that the density of the fluid is constant (ρ=constant). Taking into account these assumptions we obtain

$$\operatorname{div} \vec{V} = 0. \tag{6}$$

Therefore, the velocity field satisfies the mass balance equation (the fluid is incompressible), respectively

$$\underset{\sim}{D} = \frac{1}{2} \begin{bmatrix} 0 & \dfrac{\partial U}{\partial Y} & 0 \\ \dfrac{\partial U}{\partial Y} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{7}$$

The norm of the symmetric part of the velocity gradient is equal to

$$\left\| \underset{\sim}{\bar{\mathbf{D}}} \right\| = \frac{1}{\sqrt{2}} \left| \frac{\partial U}{\partial Y} \right| \tag{8}$$

and the Cauchy stress tensor of Eq. (1) is written as

$$\underset{\sim}{\overline{\mathbf{T}}} = \begin{bmatrix} -\overline{p} & \overline{\mu}\left(\left\| \underset{\sim}{\overline{\mathbf{D}}} \right\| \right) \dfrac{\partial U}{\partial Y} & 0 \\ \overline{\mu}\left(\left\| \underset{\sim}{\overline{\mathbf{D}}} \right\| \right) \dfrac{\partial U}{\partial Y} & -\overline{p} & 0 \\ 0 & 0 & -\overline{p} \end{bmatrix}. \tag{9}$$

The governing equations of the flow are [3]

$$\begin{aligned} \overline{\rho} \frac{\partial U}{\partial \bar{t}} &= -\frac{\partial \overline{p}}{\partial X} + \frac{\partial \underset{\sim}{\overline{\mathbf{T}}}_{YX}}{\partial Y}, \\ \frac{\partial \overline{p}}{\partial Y} &= 0, \\ \frac{\partial \overline{p}}{\partial Z} &= 0. \end{aligned} \tag{10}$$

Taking into account that the functions U and $\underset{\sim}{\overline{T}}$ are functions only of the variables Y and $\bar{t}$, from the equations (10) it follows that $\overline{p}$ is independent of Y and Z, respectively, $\dfrac{\partial \overline{p}}{\partial X} = \overline{c}(\bar{t})$, where $\overline{c}(\bar{t})$ is an arbitrary function. Using equations (3), (8) and (9), the equation governing the flow becomes

$$\begin{aligned}\bar{\rho}\frac{\partial U}{\partial \bar{t}} &= -\bar{c}(\bar{t}) + \frac{\partial}{\partial Y}\left\{\frac{\bar{\alpha}_1\sqrt{2}}{\left|\frac{\partial U}{\partial Y}\right|}\left[1-\exp\left(-\frac{\bar{\alpha}_2}{\sqrt{2}}\left|\frac{\partial U}{\partial Y}\right|\right)\right]\frac{\partial U}{\partial Y}\right\} \\ &= -\bar{c}(\bar{t}) + \frac{\partial}{\partial Y}\left\{\bar{\alpha}_1\sqrt{2}\left[1-\exp\left(-\frac{\bar{\alpha}_2}{\sqrt{2}}\left|\frac{\partial U}{\partial Y}\right|\right)\right]\operatorname{sign}\left(\frac{\partial U}{\partial Y}\right)\right\}.\end{aligned} \tag{11}$$

The function $U(Y,\bar{t})$ has to satisfy the initial and Robin type boundary conditions

$$U(Y,0) = 0, \qquad 0 \le Y \le H, \tag{12}$$

$$\begin{aligned}U(0,\bar{t}) - \bar{a}\left.\frac{\partial U(Y,\bar{t})}{\partial Y}\right|_{Y=0} &= U_0\bar{f}(\bar{t}), \qquad U_0 > 0,\ \bar{t} > 0, \\ U(H,\bar{t}) + \bar{b}\left.\frac{\partial U(Y,\bar{t})}{\partial Y}\right|_{Y=H} &= 0, \qquad \bar{t} > 0.\end{aligned} \tag{13}$$

By introducing the following dimensionless variables and functions

$$\begin{aligned}&u = \frac{U}{U_0}, \quad t = \frac{\bar{t}}{H/U_0}, \quad y = \frac{Y}{H}, \\ &\alpha_1 = \frac{\bar{\alpha}_1\sqrt{2}}{\bar{\rho}U_0^2}, \quad \alpha_2 = \frac{\bar{\alpha}_2 U_0}{H\sqrt{2}}, \quad P(t) = \frac{H}{\bar{\rho}U_0^2}\bar{c}\left(\frac{Ht}{U_0}\right),\end{aligned} \tag{14}$$

we obtain the dimensionless governing equation of the flow as

$$\frac{\partial u}{\partial t} = -P(t) + \alpha_1\frac{\partial}{\partial y}\left\{\left[1-\exp\left(-\alpha_2\left|\frac{\partial u}{\partial y}\right|\right)\right]\operatorname{sign}\left(\frac{\partial u}{\partial y}\right)\right\}. \tag{15}$$

The non-dimensional initial and boundary conditions are

$$u(y,0) = 0, \qquad 0 \le y \le 1. \tag{16}$$

$$\begin{aligned}u(0,t) - a\left.\frac{\partial u(y,t)}{\partial y}\right|_{y=0} &= f(t), \qquad t > 0, \\ u(1,t) + b\left.\frac{\partial u(y,t)}{\partial y}\right|_{y=1} &= 0, \qquad t > 0,\end{aligned} \tag{17}$$

where $a = \frac{\bar{a}}{H},\ b = \frac{\bar{b}}{H},\ f(t) = \bar{f}\left(\frac{Ht}{U_0}\right),\ f(0) = 0.$

3 Numerical solution

To obtain the numerical solution of the problem (15)-(17) we use the subroutine "*Pdesolve*" from the Mathcad 15 software package. This subroutine is suitable for the numerical integration of partial differential equations and is based on the method of lines [15]. In the method of lines for solving certain types of initial and boundary value problems, one of the variables, say y, is discretized, while the other variable t is left continuous. When the partial derivatives with respect to y are replaced by the appropriate finite difference approximations, the differential equation becomes a simultaneous algebraic system of ordinary differential equations in the variable t. An advantage of using the method of lines is that the partial differential equation is transformed into an ordinary differential equation, and therefore powerful numerical techniques for solving ordinary differential equations can be used, such as dynamically changing the step size and method to maintain stability and a desired accuracy of the integration over time. The subroutine $Pdesolve(u, y, yrange, t, trange, [ypts], [tpts]$ returns the function $u(y,t)$ which is the solution to the one dimensional nonlinear problem (15)-(17). The meaning of the parameters in the "$Pdesolve$" subroutine is as follows: "$yrange$" is a two element column vector containing the initial and final values of the spatial variable y. "$trange$" is a two element column vector containing the initial and final values of the temporal variable t, "$ypts$" (optional) is the integer number of spatial discretization points, and "$tpts$" (optional) is the integer number of temporal discretization points.

4 Analytical solution

In this section, the analytical solution of the problem (15)-(17) is determined under the following assumption:

$$\left|\alpha_2 \frac{\partial u(r,t)}{\partial r}\right| \ll 1. \tag{18}$$

In this case, by retaining the first two terms from the Taylor series of the exponential function and using the formula $\left|\frac{\partial u(r,t)}{\partial r}\right| = \frac{\partial u(r,t)}{\partial r} \operatorname{sign}\left(\frac{\partial u(r,t)}{\partial r}\right)$, equation given in (15) becomes

$$\frac{\partial u(y,t)}{\partial t} = -P(t) + \alpha_1\alpha_2 \frac{\partial^2 u(y,t)}{\partial y^2}. \tag{19}$$

The solution of equation (19), together with the initial and boundary conditions (16) and (17) will be determined using the Laplace and finite

sine-cosine Fourier transforms [5, 12, 13]. Applying the Laplace transform to equation (19) we obtain the ordinary differential equation

$$\alpha_1\alpha_2\frac{\partial^2\hat{u}(y,s)}{\partial y^2} = s\hat{u}(y,s) + \hat{P}(s), \tag{20}$$

where $\hat{u}(y,s)$ and $\hat{P}(s)$ are Laplace transforms of the functions $u(y,t)$ and $P(t)$, respectively. The finite sine-cosine Fourier transform of the function $\hat{u}(y,s)$ is defined as follows [3, 15]

$$\hat{u}_F(\zeta_k,s) = \int_0^1 \hat{u}(y,s)\left[\sin(\zeta_k y) + a\zeta_k\cos(\zeta_k y)\right]\,dy, \tag{21}$$

where ζ_k, $k = 1, 2, \ldots$, are the positive roots of the transcendental equation

$$\tan(\zeta) = \frac{(a+b)\zeta}{ab\zeta^2 - 1}. \tag{22}$$

Equation (22) is written in the equivalent form

$$\frac{\sin(\zeta_k) + a\zeta_k\cos(\zeta_k)}{a\zeta_k^2\sin(\zeta_k) - \zeta_k\cos(\zeta_k)} = b. \tag{23}$$

Using integration by parts, it is shown that the equality holds

$$\begin{aligned}
&\int_0^1 \frac{\partial^2\hat{u}(y,s)}{\partial y^2}\left[\sin(\zeta_k y) + a\zeta_k\cos(\zeta_k y)\right]dy \\
&= -\zeta_k^2\hat{u}_F(\zeta_k,s) + \zeta_k\left[\hat{u}(0,s) - a\left.\frac{\partial\hat{u}(y,s)}{\partial y}\right|_{y=0}\right] \\
&\quad + \zeta_k\cos(\zeta_k)\left(a\zeta_k\tan(\zeta_k) - 1\right) \\
&\qquad \times\left[\hat{u}(1,s) + \frac{\sin(\zeta_k) + a\zeta_k\cos(\zeta_k)}{a\zeta_k^2\sin(\zeta_k) - \zeta_k\cos(\zeta_k)}\left.\frac{\partial\hat{u}(y,s)}{\partial y}\right|_{y=1}\right].
\end{aligned} \tag{24}$$

Using (16), (17) and (23), equation (24) becomes

$$\int_0^1 \frac{\partial^2\hat{u}(y,s)}{\partial y^2}\left[\sin(\zeta_k y) + a\zeta_k\cos(\zeta_k y)\right]dy = -\zeta_k^2\hat{u}_F(\zeta_k,s) + \zeta_k\hat{f}(s). \tag{25}$$

Applying the finite sine-cosine Fourier transform to equation (20) and using the relationship (25) we obtain

$$\hat{u}_F(\zeta_k,s) = \frac{\alpha_1\alpha_2\zeta_k}{s + \alpha_1\alpha_2\zeta_k^2}\hat{f}(s) - \frac{a\zeta_k\sin(\zeta_k) + 1 - \cos(\zeta_k)}{\zeta_k\left(s + \alpha_1\alpha_2\zeta_k^2\right)}\hat{P}(s). \tag{26}$$

Let's consider the auxiliary function

$$\hat{h}(y,s)=\frac{1+b-y}{1+a+b}\,\hat{f}(s), \tag{27}$$

whose sine-cosine Fourier transform is

$$\hat{h}_F(\zeta_k,s)=H_k\hat{f}(s), \tag{28}$$

where

$$H_k=\frac{1}{(1+a+b)\zeta_k^2}\left[\left(ab\zeta_k^2-1\right)\sin(\zeta_k)-(a+b)\zeta_k\cos(\zeta_k)+(1+a+b)\zeta_k\right]. \tag{29}$$

The function (26) is written in equivalent form as

$$\begin{aligned}\hat{u}_F(\zeta_k,s)&=\hat{h}_F(\zeta_k,s)+\frac{\alpha_1\alpha_2\zeta_k}{s+\alpha_1\alpha_2\zeta_k^2}\hat{f}(s)\\&\quad-\frac{a\zeta_k\sin(\zeta_k)+1-\cos(\zeta_k)}{\zeta_k\left(s+\alpha_1\alpha_2\zeta_k^2\right)}\hat{P}(s)-\hat{h}_F(\zeta_k,s)\\&=\hat{h}_F(\zeta_k,s)+A_k\frac{1}{s+\alpha_1\alpha_2\zeta_k^2}\hat{f}(s)-B_k\frac{1}{s+\alpha_1\alpha_2\zeta_k^2}\hat{P}(s),\end{aligned} \tag{30}$$

where

$$\begin{aligned}A_k&=\alpha_1\alpha_2\zeta_k-H_k,\\B_k&=\frac{a\zeta_k\sin(\zeta_k)+1-\cos(\zeta_k)}{\zeta_k}.\end{aligned} \tag{31}$$

The inverse sine-cosine Fourier transform is given by [12]

$$\hat{u}(y,s)=\sum_{n=1}^{\infty}\frac{1}{I_{nn}}\hat{u}_F(\zeta_n,s)\left[\sin(\zeta_n y)+a\zeta_n\cos(\zeta_n y)\right], \tag{32}$$

where

$$I_{nn}=\frac{a^2\zeta_n^2-1}{4\zeta_n}\sin(2\zeta_n)+\frac{1}{2}\left[1+a+a^2\zeta_n^2-a\cos(2\zeta_n)\right]. \tag{33}$$

Using (27), (28), (30) and (32) and applying the inverse Laplace and inverse Fourier transforms, we obtain the following expression of the velocity field:

$$\begin{aligned}u(y,t)&=\frac{1+b-y}{1+a+b}f(t)+\sum_{n=1}^{\infty}\frac{A_n}{I_{nn}}\varphi_n(y)\int_0^t\exp\left[-\alpha_1\alpha_2\zeta_n^2(t-\tau)\right]f(\tau)\,d\tau\\&\quad-\sum_{n=1}^{\infty}\frac{B_n}{I_{nn}}\varphi_n(y)\int_0^t\exp\left[-\alpha_1\alpha_2\zeta_n^2(t-\tau)\right]P(\tau)\,d\tau,\end{aligned} \tag{34}$$

with

$$\varphi_n(y) = \sin(\zeta_n y) + a\zeta_n \cos(\zeta_n y). \tag{35}$$

With the help of simple calculations it is shown that the function given by (34) verifies the initial and boundary conditions (16) and (17).

To verify the working hypothesis (18), we determine the gradient of the function (34), namely:

$$\begin{aligned}\frac{\partial u(y,t)}{\partial y} = &-\frac{f(t)}{1+a+b} + \sum_{n=1}^{\infty} \frac{A_n\zeta_n}{I_{nn}} \left[\cos(\zeta_n y) - a\zeta_n \sin(\zeta_n y)\right] \\ &\times \int_0^t \exp\left[-\alpha_1\alpha_2\zeta_n^2(t-\tau)\right] f(\tau)\, d\tau \\ &- \sum_{n=1}^{\infty} \frac{B_n\zeta_n}{I_{nn}} \left[\cos(\zeta_n y) - a\zeta_n \sin(\zeta_n y)\right] \\ &\times \int_0^t \exp\left[-\alpha_1\alpha_2\zeta_n^2(t-\tau)\right] P(\tau)\, d\tau.\end{aligned} \tag{36}$$

5 Semi-analytical solution

In this section, we present a semi-analytic solution of the differential equation (19) that satisfies the initial and boundary conditions (16) and (17). This solution is determined using the Laplace transform and the general theory of linear differential equations with constant coefficients.

Applying the Laplace transform to equations (17) and (19), we obtain that the Laplace transform $\hat{u}(y,s)$ of the function $u(y,t)$ must satisfy equation (20) and the boundary conditions

$$\begin{aligned}\hat{u}(0,s) - a\left.\frac{\partial \hat{u}(y,s)}{\partial y}\right|_{y=0} &= \hat{f}(s), \\ \hat{u}(1,s) + b\left.\frac{\partial \hat{u}(y,s)}{\partial y}\right|_{y=1} &= 0.\end{aligned} \tag{37}$$

Equation (20) has the particular solution

$$\hat{u}_p(y,s) = -\frac{1}{s}\hat{P}(s) \tag{38}$$

and the general solution

$$\hat{u}(y,s) = C_1(s)\exp\left(cy\sqrt{s}\right) + C_2(s)\exp\left(-cy\sqrt{s}\right) - \frac{1}{s}\hat{P}(s), \tag{39}$$

where $c = \dfrac{1}{\sqrt{\alpha_1 \alpha_2}}$ and $C_1(s), C_2(s)$ will be determined using the boundary conditions. By imposing that function (39) satisfies conditions (37), the following values of the functions $C_1(s)$ and $C_2(s)$ are obtained:

$$C_1(s) = \frac{\Delta_1(s)}{\Delta(s)}, \qquad C_2(s) = \frac{\Delta_2(s)}{\Delta(s)} \tag{40}$$

with

$$\begin{aligned}
\Delta_1(s) &= (1 - bc\sqrt{s}) \left(\hat{f}(s) + \frac{\hat{P}(s)}{s} \right) \exp(-c\sqrt{s}) - (1 + ac\sqrt{s}) \frac{\hat{P}(s)}{s}, \\
\Delta_2(s) &= (1 - ac\sqrt{s}) \frac{\hat{P}(s)}{s} - (1 + bc\sqrt{s}) \left(\hat{f}(s) + \frac{\hat{P}(s)}{s} \right) \exp(c\sqrt{s}), \\
\Delta(s) &= (1 - ac\sqrt{s})(1 - bc\sqrt{s}) \exp(-c\sqrt{s}) - (1 + ac\sqrt{s})(1 + bc\sqrt{s}) \exp(c\sqrt{s}).
\end{aligned} \tag{41}$$

Even though the inverse Laplace transform of function (39) could be determined by complex analysis methods, the original form of the function is complicated and therefore difficult to use for calculating numerical values of the velocity field. For this reason, we will determine the real-domain values of function (39) using the Stehfest algorithm for numerical inversion of Laplace transforms [11]. Based on this algorithm, the inverse Laplace transform of the function $\hat{u}(y, s)$ is given by the relation

$$\begin{aligned}
u(y,t) \simeq \frac{\ln(2)}{t} \sum_{j=1}^{2p} (-1)^{j+p} \sum_{i=\left\lfloor \frac{j+1}{2} \right\rfloor}^{\min(j,p)} \\
\times \frac{i^p (2i)!}{(p-i)!\, i!\, (i-1)!\, (j-1)!\, (2i-j)!} \hat{u} \left(y,\, j \frac{\ln(2)}{t} \right),
\end{aligned} \tag{42}$$

where p is a positive integer and $[w]$ denotes the integer part of real number w.

At the end of this section we show, graphically, the equivalence of the analytical solution (34) with the semi-analytical solution (42). The ζ_k, $k = 1, 2, \ldots$, roots of equation (22), which are necessary for determining the numerical values of the function (34), are determined with the subroutine "root()" from the Mathcad 15 software package.

The curves in Figure 1 were plotted to compare the numerical values of functions (34) and (42). These curves were plotted for the values of the

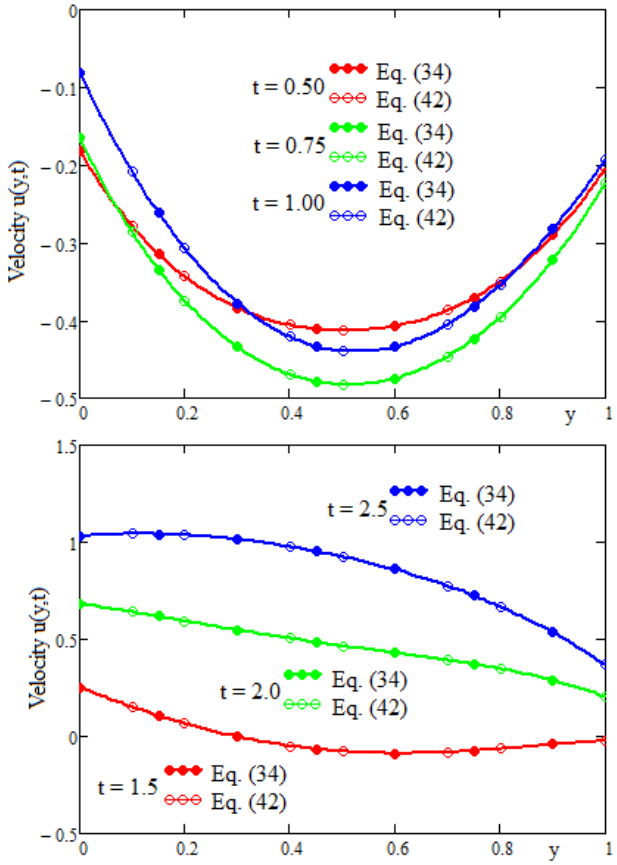


Figure 1: Velocity profiles corresponding to equations (34) and (42) for different values of the time t.

material parameters $\alpha_1 = 0.8$, $\alpha_2 = 0.2$, $a = 0.5$, $b = 0.2$, respectively, $f(t) = 1 - \exp(-t)$ for the boundary condition, and the pressure gradient

$$P(t) = \cos\left(\frac{\pi t}{2}\right).$$

It can be seen in Figure 1 that the curves corresponding to the expressions (34) and (42) of the fluid velocity overlap, so the two solutions are equivalent.

6 Time variation of fluid properties

Shear thinning of paints is a property in which the viscosity of a liquid decreases as it is mixed, brushed, or sprayed. At rest, paint has a high viscosity, so it behaves like a thick liquid. If a shear stress is applied to the paint, this force aligns the microscopic polymer chains in the paint, making the fluid thinner and easier to spread evenly.

Next, we present the evolution of the fluid velocity, velocity gradient and its viscosity over time. Numerical simulations were obtained for the boundary condition

$$f(t) = \frac{1 - \exp(-t)}{5},$$

for the pressure gradient

$$P(t) = \cos\left(\frac{\pi t}{2}\right),$$

in five positions of the spatial domain and for different values of the material coefficients α_1 and α_2.

In Figure 2, the curves representing the velocity profiles for three different values of the material parameter α_2 that essentially influences the fluid viscosity are plotted. It is observed that the fluid velocity changes significantly for small values of time t. With increasing time, the velocity tends towards a constant value specific to the considered spatial position.

Figure 3 shows the curves corresponding to the absolute values of the velocity gradient in the spatial positions $y \in \{0, 0.2, 0.5, 0.8, 1.0\}$ and for three different values of the material parameter α_2.

The velocity gradient has significant variations for small values of time t. After a certain value of time the values of the velocity gradient become constant for each spatial position considered.

Since the dimensionless viscosity of the fluid is given by the relation

$$\mu(y,t) = \alpha_1 \left[1 - \exp\left(-\alpha_2 \left|\frac{\partial u(y,t)}{\partial y}\right|\right)\right] \left|\frac{\partial u(y,t)}{\partial y}\right|^{-1}. \tag{43}$$

It results that the fluid viscosity values also vary significantly for small values of time and tend towards a constant value for large values of time t. This behavior of the fluid viscosity is highlighted by the curves shown in Figure 4.

It is important to note that the material parameter α_2 has a strong influence on the fluid behavior. It can be seen from Figure 2 that the

fluid velocity values decrease for increasing values of the material parameter α_2. This property is closely related to the evolution of the fluid viscosity. From relation (43), it follows that the fluid viscosity values increase for increasing values of the material parameter α_2, causing the fluid to thicken and therefore slow down its movement.

7 Steady-state solution

In this section, we show that if condition (18) is fulfilled and if the functions f(t) and p(t) have finite limits for t tending to infinity, then the fluid motion becomes stationary. To determine the stationary solution, we use the following property of the Laplace transform:

$$\lim_{s\to 0} s\hat{\psi}(s) = \lim_{t\to\infty} \psi(t), \tag{44}$$

where $\hat{\psi}(s)$ denotes the Laplace transform of the function $\psi(t)$. The Laplace transform of the transient velocity given by the equation (34) is written as

$$\hat{u}(y,s) = \frac{1+b-y}{1+a+b}\,\hat{f}(s) + \sum_{n=1}^{\infty} \frac{\varphi_n(y)}{I_{nn}\,(s+\alpha_1\alpha_2\xi_n^2)} \left[A_n\hat{f}(s) - B_n\hat{P}(s)\right]. \tag{45}$$

Assuming that $\lim\limits_{t\to\infty} f(t) = \ell_1 < \infty$, $\lim\limits_{t\to\infty} P(t) = \ell_2 < \infty$, and using (44) and (45), we obtain the following expression of the stationary solution:

$$u_{st}(y) = \frac{1+b-y}{1+a+b}\,\ell_1 + \sum_{n=1}^{\infty} \frac{\varphi_n(y)}{I_{nn}\alpha_1\alpha_2\xi_n^2} \left[A_n\ell_1 - B_n\ell_2\right]. \tag{46}$$

As illustrated in Figure 5, the profiles for the transient velocity curves (34) asymptotically approach the reference curve (46) with increasing values of time t. The numerical simulations utilized to generate these profiles were conducted using the material parameters $\alpha_1 = 0.8$, $\alpha_2 = 0.2$ in conjunction with the functions $f(t) = 2(1-\exp(-t))$ and $P(t) = \cos(\pi t/2)\exp(-t) - 1$.

8 Conclusions

This paper investigated flows of an apparently viscoelastic fluid modeled as paint in a rectangular domain bounded by two flat parallel plates with Robin-type conditions. The fluid is described by the constitutive relation proposed by Garimella et al. [10], wherein viscosity varies exponentially with

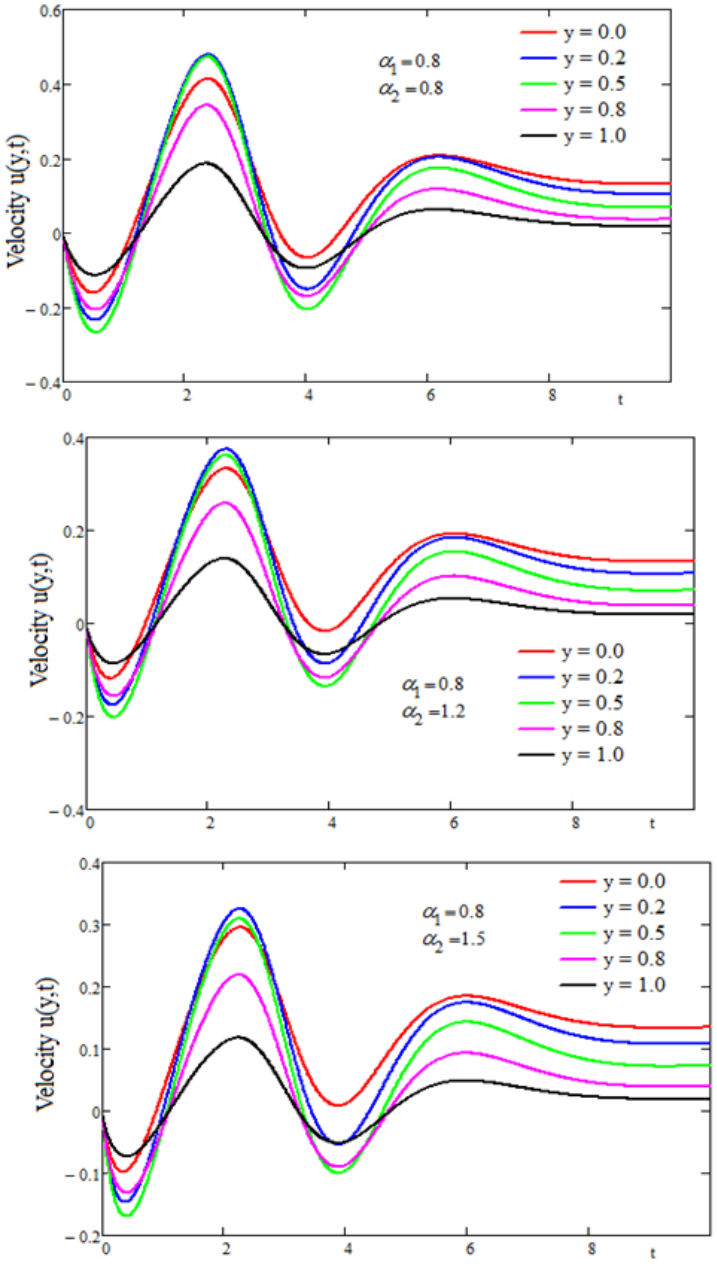


Figure 2: Time variation of the fluid velocity for different positions of spatial domain

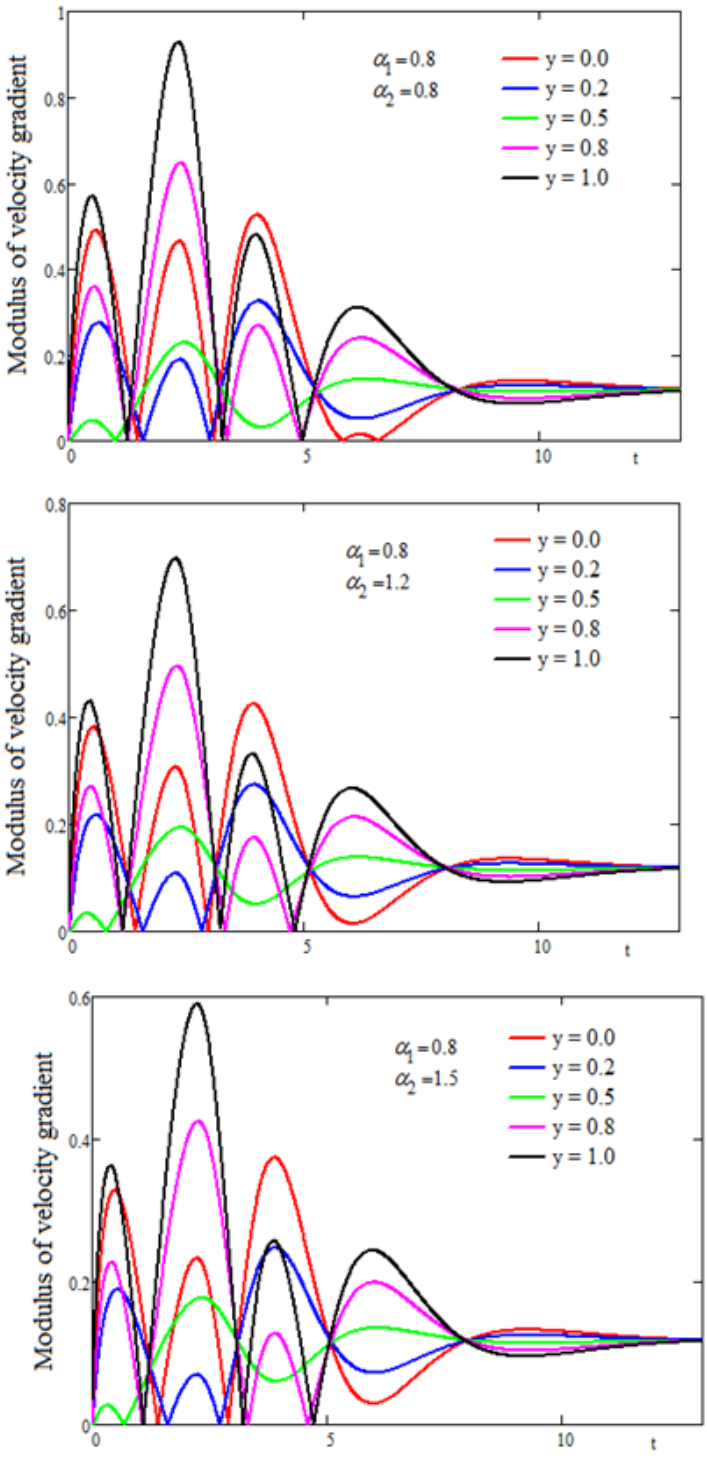


Figure 3: Time variation of absolute values of velocity gradient for different positions of spatial domain

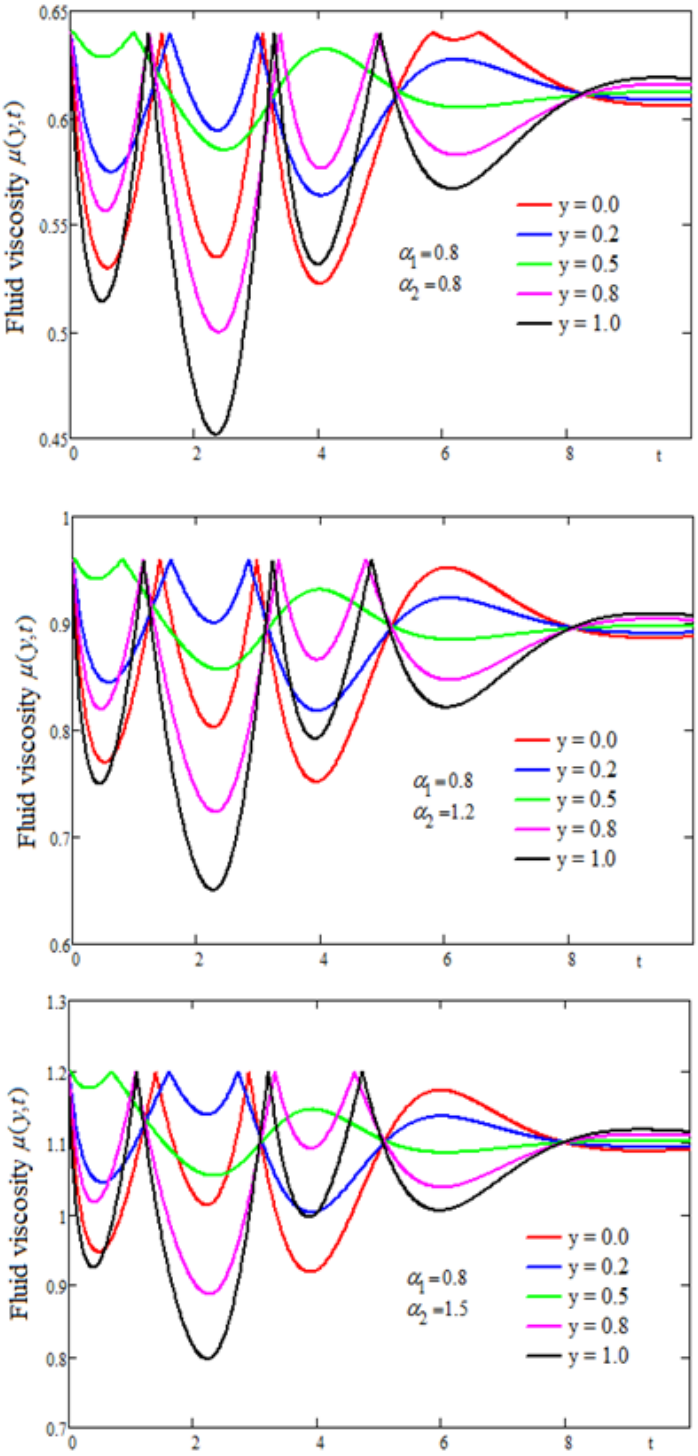


Figure 4: Time variation of the fluid viscosity for different positions of spatial domain

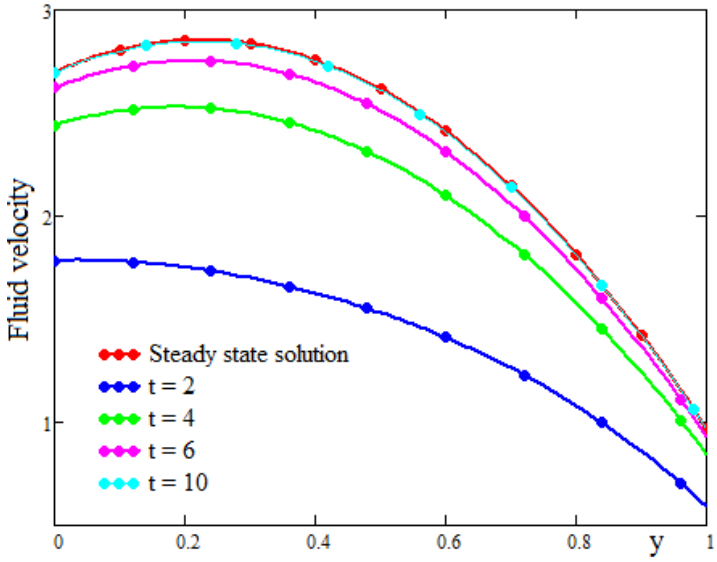


Figure 5: Convergence of the transient solution (34) to the steady-state solution (46).

the Frobenius norm of the symmetric part of the velocity gradient. A numerical solution to the governing nonlinear partial differential equation was obtained using the *Pdsolve* routine in Mathcad 15. Under specific conditions for the material coefficients and velocity gradient, the nonlinear equation simplifies to a linear form, allowing for exact analytical solutions via Laplace and finite sine-cosine Fourier transforms, as well as a semi-analytical solution. A comparative analysis between the numerical and analytical solutions validates the methodology. Investigations into the influence of material parameters show that for increasing values of the material parameter α_2, the fluid thickens and therefore its movement is slowed down.

References

[1] A. Almgvist, E. Burtseva, K. Rajagopal and P. Wall, Modeling pressure-driven flow between adjacent surfaces of viscoplastic and seemingly viscoplastic materials, *Appl. Math.* 71 (2026), 1-30.

[2] M. Anand and B. Gnaneshwari, A variant of Stokes' second problem for a new shear thinning model for paint, *Appl. Eng. Sci.* 25 (2026),

100288.

[3] R.B. Bird, R.C. Armstrong and O. Hassager, *Dynamics of Polymeric Liquids, Vol. 1: Fluid Mechanics*, John Wiley & Sons, New York, 1987.

[4] C. Datt and G. Elfring, Dynamics and rheology of particles in shear-thinning fluids, *J. Non-Newtonian Fluid Mech.* 262 (2018), 107-114.

[5] L. Debnath and D. Bhatta, *Integral Transforms and Their Applications*, Third Edition, CRC Press, Boca Raton, London, New York, 2015.

[6] C. Fetecau and H. Hanif, Long-time solutions of the modified MHD Stokes' problems for a class of Maxwell fluids with pressure-dependent viscosity, Applications, *Discrete Contin. Dyn. Syst. - Ser. S* 30 (2026), 161-176.

[7] C. Fetecau and C. Morosanu, General solutions for MHD motions of viscous fluids with viscosity linearly dependent on pressure in a planar channel, *IgMin Res.* 3 (2025), 104-112.

[8] L. Fusi, A. Giovinetto and K.R. Rajagopal, Unsteady one dimensional motions of a new class of seemingly viscoplastic materials, *Appl. Math. Comput.* 497 (2025), 129366.

[9] L. Fusi and K.R. Rajagopal, Flow past a porous plate of a new class of fluids with limiting stress: Analytical results and linear stability analysis, *Eur. J. Mech. B. Fluids* 112 (2025), 58-64.

[10] S.M. Garimella, M. Anand and K.R. Rajagopal, A new model to describe the response of a class of seemingly viscoplastic materials, *Appl. Math.* 67 (2022), 153-165.

[11] A. Kuznetsov and J. Miles, On the rate of convergence of the Gaver–Stehfest algorithm, *IMA J. Numer. Anal.* 42 (2022), 1645-1664.

[12] I.A. Mirza, M.S. Akram, N.A. Shah, W. Imtiaz and J.D. Chung, Analytical solutions to the advection-diffusion equation with Atangana–Baleanu time-fractional derivative and a concentrated loading, *Alexandria Eng. J.* 60 (2021), 1199-1208.

[13] I.A. Mirza, D. Vieru and N. Ahmed, Fractional advection-diffusion equation with memory and Robin-type boundary condition, *Math. Model. Nat. Phenom.* 14 (2019), 306.

[14] J. Papaioannou, G. Karapetsas, Y. Dimakopoulos and J. Tsamopoulos, Injection of a viscoplastic material inside a tube or between two parallel disks: Conditions for wall detachment of the advancing front, *J. Rheol.* 53 (2009), 1155.

[15] W.E. Schiesser, *Numerical Method of Lines Integration of Partial Differential Equations*, Academic Press, San Diego, CA, 1991.

[16] T.F. Tadros, Rheology of paint formulations, *Industrial Applications II: Agrochemicals, Paints, Coatings and Food Systems*, Berlin, Boston: De Gruyter, 2018, pp. 259-288.

[17] D. Vieru, C. Fetecau and Z. Ismail, The effects of shear stress memory and variable viscosity on viscous fluids flowing between two horizontal parallel plates, *Mathematics* 13 (2025), 3043.

Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 18, No. 4/2026

WEAK DISCONTINUITY WAVES IN ISOTROPIC POROUS MEDIA FILLED BY A FLUID FLOW*

Alessio Famà† Liliana Restuccia‡

Dedicated to Professor Constantin Fetecau on the occasion of his 80th anniversary

DOI 10.56082/annalsarscimath.2026.4.265

Abstract

In this paper the propagation of weak discontinuity waves is investigated in a special isotropic porous medium filled by a fluid flow. To this aim we introduce a new variable related to the surface across which the first derivatives of the solutions of the governing equations undergo a jump. Following a Boillat's methodology for quasi-linear and hyperbolic systems of the first order, we have investigated, in an one-dimensional case, the propagation of one solution into an uniform unperturbed state in the first approximation, obtaining the Bernoulli's equation governing the evolution of the amplitude of the weak discontinuity wave, the wave front, the velocity along the characteristic rays, the approximated solution and the critical time. The results obtained can be applied in certain technical fields, such as biology, the medical sciences and geology.

Keywords: weak discontinuity waves, partial differential equations, approximate solutions, non-equilibrium thermodynamics, porous media filled by a fluid flow.

*Accepted for publication on October 02, 2026

†alessio.fama.04@gmail.com, University of Messina, Department of Mathematical and Computer Sciences, Physical Sciences and Earth Sciences, Viale F. Stagno d'Alcontres, Salita Sperone 31, 98166 Messina, Italy

‡lrestuccia@unime.it, University of Messina, Department of Mathematical and Computer Sciences, Physical Sciences and Earth Sciences, Viale F. Stagno d'Alcontres, Salita Sperone 31, 98166 Messina, Italy

MSC: 35A25, 74A15, 74F10, 80A17.

1 Introduction

The study of media with thin porous channels may have relevance in the description of phenomena accompanying fluid flows in porous structures and find applications in applied sciences, like biology, medical sciences, seismic engineering. Here, we use a thermodynamic theory (see [38], [39], [41] and [42]), formulated following the methods of the extended rational thermodynamics [7], [21], [22], [23], [30], [31], [33], [45], with internal variables (see [1], [23], [28], [29], [32] and also [40]).

More precisely, in [38] and [39] introducing as internal variable the structural permeability tensor à la Kubik [24], see also [38] and [41], describing the influence of thin porous channels on the other fields, occurring inside the considered media, a model was developed for porous media filled by a fluid flow. By the analysis of the entropy inequality, using Liu's theorem [27], the constitutive theory and the rate equations for the fluxes and the porosity field were constructed as objective functions, in a special case. Smith's theorem [43] was used. In [41] and [42] constitutive relations, rate equations and other results were derived for the same media in the anisotropic case. In [10] the isotropic case was treated. In [11] a thermodynamic theory for erosion and/or deposition in elastic porous media was developed by the authors.

In [9], [12], [13], in special isotropic cases the harmonic waves propagation of coupled porosity-fluid concentration flux fields, of coupled porosity-fluid concentration flux-temperature fields and of porosity-temperature fields were investigated, respectively. In this article utilizing a model derived in [9] for perfect isotropic porous media satured by fluid flow in a particular case, the propagation of weak discontinuity waves is investigated, confining us only to the porosity, fluid concentration fields and their fluxes.

Several physical problems are described by nonlinear hyperbolic PDEs. Following A. Jeffrey in [20], the solution hypersurfaces of these are referred to as waves, because they may be interpreted as representing propagating wavefronts. The solution ahead of the wavefront may called undisturbed, whilst the solution on the other side may be regarded as a propagating disturbance wave entering in the region of the undisturbed solution. Because the closed-form solutions of nonlinear PDEs are rare, solutions are investigated, called *nonlinear waves*, that can present various types of discontinuities. As some surface is crossed, the solution and/or some of its derivatives undergo a jump: in the first case the solution presents a *shock*,

or it is a *shock wave*, in the second case we are in presence of a *discontinuity waves* (jumps of the first order derivatives) [2–4, 8, 17–20, 25, 26]. Also we can have instead of the jump smooth solutions of the non linear hyperbolic PDEs, that present a steep variation in the normal direction to the associated wavefront, called *asymptotic waves* (see [5, 6, 14–16, 36]).

The paper is organized as follows. In section 2, we present the equations of the thermodynamical model formulated in [38], used to describe the behavior of porous media filled by a fluid flow in anisotropic and isotropic cases (see [39], [41] and [42]) and we consider a particular isotropic case, worked out in [9], where the model is confined only to porosity and the fluid concentration fields and their fluxes, and some interactions among some fields are neglected. In this paper some coefficients of the governing equations are assumed to be non-constant, depending on the fields involved, obtaining a non-linear PDEs system. In sections 3, 4 and 5, the methodology established by Boillat to determine the weak discontinuity waves for quasi-linear and hyperbolic systems of the first order is illustrated, introducing a new variable related to the surface across which the first derivatives of the solutions of the governing equations undergo a jump. In sections 6 and 7, in an one-dimensional case, we study the propagation of one solution into an uniform unperturbed state in the first approximation, obtaining the associated wave front, the velocity along the characteristic rays, the Bernoulli's equation governing the evolution of the associated weak discontinuity wave amplitude and the approximated solution. Also, we calculate the critical time, that represents the instant in which the solution ceases to exist or to be unique or, from the point of view of the non-linear waves, the weak discontinuity wave amplitude becomes unbounded, and this may correspond to the formation of a shock wave.

2 The model

In the following we use the standard Cartesian tensor notation in rectangular coordinate systems and we refer to a current configuration K_t at the time t. We take into consideration the theory for a porous medium filled by a fluid flow, formulated in [38], [39], [42], [41], in the framework of rational extended thermodynamics, where we have assumed that within the medium there are present the elastic field described by the symmetric stress tensor τ_{ij} and the small-strain tensor ε_{ij}; the thermal field described by the temperature T, its gradient $T_{,i}$ and the heat flux q_i; the field of the fluid concentration c, its gradient $c_{,i}$ and its flux j_i^c; the porosity field, whose

geometric description is given by a structural permeability tensor r_{ij} à la Kubik (see [24] and for instance [41]), its gradient $r_{ij,k}$ and its flux $\mathcal{V}_{ijk}$, since the thin porous channels sometimes can self propagate because of some surrounding favorable conditions. We have chosen the following thermodynamic state vector:

$$C = \{\varepsilon_{ij}, c, T, r_{ij}, j_i^c, q_i, c_{,i}, T_{,i}, r_{ij,k}, \mathcal{V}_{ijk}\}, \tag{1}$$

where $\varepsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$, with u_i the displacement field.

In the previous papers we have supposed that the elastic porous medium and the fluid, flowing inside it, constitute two-components mixture, ρ_1 being the fluid mass density and ρ_2 the density of the porous skeleton, with $\rho_1 \ll \rho_2$, so that we have $\rho = \rho_1 + \rho_2$, with ρ the density of the mixture as a whole, supposed constant.

We have considered two groups of laws. The first group is concerning the classical balance equations:

- *the continuity equation*

$$\rho\dot{c} + j^c_{i,i} = 0, \tag{2}$$

where the source term is neglected, a superimposed dot indicates the material derivative (i.e. $\frac{d}{dt} = \frac{\partial}{\partial t} + x_\gamma \frac{\partial}{\partial x_\gamma}$), a comma in lower indices indicates the spatial derivation and Einstein convention for repeated indices is used. In (2) the concentration of the fluid is defined by $c = \frac{\rho_1}{\rho}$ and its flux j_i^c is given by $j_i^c = \rho_1(v_{1i} - v_i)$, where v_{1i} is the fluid velocity and v_i is the barycentric velocity of the mixture, defined by $\rho v_i = \rho_1 v_{1i} + \rho_2 v_{2i}$, with v_{2i} the porous medium velocity.

- *the momentum balance*

$$\rho\dot{v}_i - \tau_{ji,j} - f_i = 0, \tag{3}$$

where f_i denotes a body force, disregarded in the following for the sake of simplicity;

- *the internal energy balance*

$$\rho\dot{e} - \tau_{ji}v_{i,j} + q_{i,i} - \rho h = 0, \tag{4}$$

where e is the internal energy density, h is the energy source density (neglected in the following) and $v_{i,j}$ is the gradient of the barycentric velocity.

The second group is concerning the rate equations for the porosity field r_{ij}, its flux V_{ijk}, the heat flux q_i and the fluid-concentration flux j_i^c, constructed in such a way that they obey the objectivity and frame-indifference

principles (see [34], [35], [44]). We have chosen for these evolution equations the following form

$$\overset{*}{r}_{ij} + \mathcal{V}_{ijk,k} - \mathcal{R}_{ij}(C) = 0, \tag{5}$$

$$\overset{*}{j}{}^c_i - J^c_i(C) = 0, \tag{6}$$

$$\overset{*}{q}_i - Q_i(C) = 0, \tag{7}$$

$$\overset{*}{\mathcal{V}}_{ijk} - V_{ijk}(C) = 0. \tag{8}$$

In equations (5)-(8) the quantities $\mathcal{R}_{ij}(C)$, $J^c_i(C)$, $Q_i(C)$ and $V_{ijk}(C)$ are the source-like terms which deal with the creation or annihilation of porous channels, the fluid concentration flux source, the heat source and the source term for the structural permeability field flux, respectively, and represent constitutive functions of the independent variables of the set (1), the superimposed asterisk indicates the Zaremba-Jaumann derivative (see [34], [44]), i.e.

$$\overset{*}{r}_{ij} = \dot{r}_{ij} - \Omega_{ik} r_{kj} - \Omega_{jk} r_{ik},$$

$$\overset{*}{j}{}^c_i = \dot{j}^c_i - \Omega_{ik} j^c_k, \quad \overset{*}{q}_i = \dot{q}_i - \Omega_{ik} q_k,$$

$$\overset{*}{\mathcal{V}}_{ijk} = \dot{\mathcal{V}}_{ijk} - \Omega_{il}\mathcal{V}_{ljk} - \Omega_{jl}\mathcal{V}_{ilk} - \Omega_{kl}\mathcal{V}_{ijl},$$

with $\Omega_{ij} = \frac{1}{2}(v_{i,j} - v_{j,i})$ the antisymmetric part of the velocity gradient and v_i the velocity field of the body. In [38], [39] the constitutive functions and the rate equations were obtained in special isotropic case to close the system of balance equations. The entropy inequality was analyzed by Liu's theorem (see [27]). In [42] (see also [41]) we have obtained the following rate equations in the anisotropic case for the fields r_{ij}, j^c_i, q_i and $\mathcal{V}_{ijk}$

$$\begin{aligned}\dot{r}_{ij} + \mathcal{V}_{ijk,k} = {} & \beta^1_{ijkl}\varepsilon_{kl} + \beta^2_{ijkl} r_{kl} + \beta^3_{ijk} j^c_k + \beta^4_{ijk} q_k + \beta^5_{ijklm}\mathcal{V}_{klm} + \beta^6_{ijk} c_{,k} \\ & + \beta^7_{ijk} T_{,k} + \beta^8_{ijklm} r_{kl,m},\end{aligned} \tag{9}$$

$$\tau^{j^c}_{ij}\dot{j}^c_j = -j^c_i + \xi^2_{ij} q_j + \xi^3_{ijkl}\mathcal{V}_{jkl} - \xi^4_{ij} c_{,j} + \xi^5_{ij} T_{,j} + \xi^6_{ijkl} r_{jk,l}, \tag{10}$$

$$\tau^q_{ij}\dot{q}_j = \chi^1_{ij} j^c_j - q_i + \chi^3_{ijkl}\mathcal{V}_{jkl} + \chi^4_{ij} c_{,j} - \chi^5_{ij} T_{,j} + \chi^6_{ijkl} r_{jk,l}, \tag{11}$$

$$\dot{\mathcal{V}}_{ijk} = \gamma^1_{ijkl} j^c_l + \gamma^2_{ijkl} q_l + \gamma^3_{ijklmn}\mathcal{V}_{lmn} + \gamma^4_{ijkl} c_{,l} + \gamma^5_{ijkl} T_{,l} + \gamma^6_{ijklmn} r_{lm,n}, \tag{12}$$

with $\tau^{j^c}_{ij}$ and τ^q_{ij} relaxation time tensors of the fields j^c_i and q_i, respectively. In equations (9)-(12) we have assumed the material derivative can be used instead of Zaremba-Jaumann derivative. Equations (9)-(12) describe disturbances, having finite velocity of propagation and own relaxation time to

reach the respective thermodynamic equilibrium values. Equations (9) and (12) describe the evolution of the porosity field, of its flux and in their right hand sides the source terms represent the contributions of the fields occurring inside these media. Equations (10) and (11) are the anisotropic Fick-Nonnenmacher equation for the fluid concentration flux and the anisotropic generalized Maxwell-Vernotte-Cattaneo equation for the heat flux, respectively.

2.1 Evolution equations for the concentration of the fluid and the porosity fields and their fluxes in a special case

In [9] we have treated the special case of perfect isotropic media, where their geometric, transport and thermal properties are invariant respect to all rotations and inversions of the frame axes (under orthogonal transformations). We have assumed the media into consideration are at rest and we have disregarded some cross effects among the fields confining us only to the porosity and the fluid concentration fields and their fluxes. In particular, taking into account equation (2), (10), (9) and (12), in [9] we have worked out the following system of equations

$$\rho \frac{\partial c}{\partial t} + j^c_{k,k} = 0, \tag{13}$$

$$\tau^{j^c} \frac{\partial j^c_i}{\partial t} = -j^c_i + \alpha_c r_{,i} - \rho D_c c_{,i}, \tag{14}$$

$$\frac{\partial r}{\partial t} + \mathcal{V}_{k,k} = 0, \tag{15}$$

$$\tau^{\nu} \frac{\partial \mathcal{V}_k}{\partial t} = -\mathcal{V}_k - D_\nu r_{,k} + \alpha_\nu c_{,k}, \tag{16}$$

with α_c and α_ν constant coupling coefficients, τ^{j^c} and τ^{ν} the constant relaxation times of the field j^c_i and $\mathcal{V}_{ijk}$, and D_c, D_ν the constant diffusion coefficients of the fluid concentration and porosity fields, respectively. Furthermore, in (13) the mass density was supposed constant and in equation (15) the number of the very thin porous channels was assumed locally conserved.

In this paper we suppose that in (14) and (16) the two coupling coefficients α_c and α_ν are functions of at least class C^1 of the type
$\alpha_c = \alpha_c(r)$ and $\alpha_\nu = \alpha_\nu(c)$, and we investigate the weak discontinuity waves of the PDEs system (13)-(16).

3 Weak discontinuity waves of the non-linear PDEs system into consideration

Following Boillat [2], let us observe that the system of equations (13)- (16) can be written in the following matrix form:

$$\boldsymbol{A}^{\alpha}(\boldsymbol{U})\boldsymbol{U}_{\alpha} = \boldsymbol{B}(\boldsymbol{U}) \qquad (\alpha = 0, 1, 2, 3), \tag{17}$$

where $x^0 = t$ indicates the time, x^1, x^2, x^3 are the space coordinates, $\boldsymbol{U}_{\alpha} = \dfrac{\partial \boldsymbol{U}}{\partial x^{\alpha}}$, $\boldsymbol{U}$ is the vector of the unknown functions (depending on x^{α})

$$\boldsymbol{U} = (c, j_1^c, j_2^c, j_3^c, r, \mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3)^T, \tag{18}$$

the vector $\boldsymbol{B}$ is defined by

$$\boldsymbol{B} = (0, -j_1^c, -j_2^c, -j_3^c, 0, -\mathcal{V}_1, -\mathcal{V}_2, -\mathcal{V}_3)^T, \tag{19}$$

and finally $\boldsymbol{A}^{\alpha}$ $(\alpha = 0, \ldots, 3)$ are the following 8×8 square matrices

$$\boldsymbol{A}^0 = \begin{pmatrix} \rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \tau^{j^c} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tau^{j^c} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tau^{j^c} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \tau^{\nu} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \tau^{\nu} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tau^{\nu} \end{pmatrix}, \tag{20}$$

$$\boldsymbol{A}^1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho D_c & 0 & 0 & 0 & -\alpha_c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -\alpha_{\nu} & 0 & 0 & 0 & D_{\nu} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{21}$$

$$\boldsymbol{A}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho D_c & 0 & 0 & 0 & -\alpha_c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\alpha_\nu & 0 & 0 & 0 & D_\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{22}$$

$$\boldsymbol{A}^3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho D_c & 0 & 0 & 0 & -\alpha_c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\alpha_\nu & 0 & 0 & 0 & D_\nu & 0 & 0 & 0 \end{pmatrix}. \tag{23}$$

In (18) and (19) the symbol $(...)^T$ means that $\boldsymbol{U}$ and $\boldsymbol{B}$ are column vectors.

Since $\boldsymbol{A}^\alpha = \boldsymbol{A}^\alpha(\boldsymbol{U})$, the PDEs system (17) is a quasi-linear system.

We suppose that the system admits a known solution in the *uniform unperturbed state* $\boldsymbol{U}^0$ that satisfy the following condition

$$\boldsymbol{A}^\alpha(\boldsymbol{U}^0)\boldsymbol{U}^0_\alpha = \boldsymbol{B}(\boldsymbol{U}^0) \qquad (\alpha = 0, 1, 2.3). \tag{24}$$

3.1 Introduction of a new variable

Let us assume (see [2], [20]) that the system (17) describes a perturbation propagating into an unperturbed state characterized by the vector $\boldsymbol{U}^0$ and the hypersurface $\mathcal{S}$

$$\varphi(t, x^1, x^2, x^3, x^4) = 0, \tag{25}$$

improperly called *wavefront* and supposed to be characterized by the fact that the first derivatives of the solution $\boldsymbol{U}$ of the system (17) undergo a jump crossing it. $\varphi(x^\alpha) = 0$ separates the region perturbed, $\varphi(x^\alpha) = 0^+$, by the unperturbed, $\varphi(x^\alpha) = 0^-$, and moves in the Euclidean space E^{3+1} (when the time flows). The solution $\boldsymbol{U}$ is supposed continuous.

We remind that the wavefront $\varphi(x^\alpha) = 0$ is still an unknown function. In order to determine it, we see that from (25) along $\mathcal{S}$ we have $\dfrac{d\varphi}{dt} = 0$, implying $\frac{\partial \varphi}{\partial t} + \boldsymbol{v} \cdot \operatorname{grad} \varphi = 0$, or, equivalently, $\frac{\partial \varphi / \partial t}{|\operatorname{grad} \varphi|} + \boldsymbol{v} \cdot \frac{\operatorname{grad} \varphi}{|\operatorname{grad} \varphi|} = 0$, with

$(\text{grad})_i = \dfrac{\partial}{\partial x^i}$ and, obviously, $\frac{\text{grad}\,\varphi}{|\text{grad}\,\varphi|} = \boldsymbol{n}$ the *normal unit vector* to the hypersurface $\mathcal{S}$, such that the previous equality reads

$$\frac{\partial\varphi/\partial t}{|\text{grad}\,\varphi|} + \boldsymbol{v}\cdot\boldsymbol{n} = 0. \tag{26}$$

Introducing the notations

$$\lambda = -\frac{\partial\varphi/\partial t}{|\text{grad}\,\varphi|}, \quad \Psi(\boldsymbol{U},\varphi_\alpha) = \varphi_t + \lambda(\boldsymbol{U},\boldsymbol{n})|\text{grad}\,\varphi|, \tag{27}$$

with

$$\varphi_\alpha = \frac{\partial\varphi}{\partial x^\alpha}, \quad \Lambda_i(\boldsymbol{U},\boldsymbol{n}) = \frac{\partial\Psi}{\partial\varphi_i},$$

we obtain the *velocity normal to the progressive wave*, λ, given by

$$\lambda(\boldsymbol{U},\boldsymbol{n}) = \boldsymbol{v}\cdot\boldsymbol{n}, \tag{28}$$

and the components of a speed, called *radial velocity*, $\boldsymbol{\Lambda}$, defined by

$$\Lambda_i(\boldsymbol{U},\boldsymbol{n}) = \frac{\partial\Psi}{\partial\varphi_i} = \lambda n_i + \frac{\partial\lambda}{\partial n_i} - \left(\boldsymbol{n}\cdot\frac{\partial\lambda}{\partial\boldsymbol{n}}\right)n_i \quad (i=1,2,3). \tag{29}$$

From equation (29) we have

$$\Lambda_i n_i = \lambda, \tag{30}$$

i.e. the velocity of propagation of the wavefront λ is the component of radial velocity $\boldsymbol{\Lambda}$ along the normal to the wavefront.

Introducing the function $\varphi = \varphi(x^\alpha)$ as new variable, continuous together with its first and second derivatives, the derivative with respect to x^α is given by

$$\boldsymbol{U}_\alpha = \boldsymbol{U}_\varphi\,\varphi_\alpha, \tag{31}$$

where $\boldsymbol{U}_\varphi = \dfrac{\partial\boldsymbol{U}}{\partial\varphi}$ and $\varphi_\alpha = \dfrac{\partial\varphi}{\partial x^\alpha}$.

Moreover, we introduce the *jump* symbol by $[\ \ ] = \lim_{\varphi\to0^+}(\ \ ) - \lim_{\varphi\to0^-}(\ \ )$, defined by the difference between the value of a quantity in the unperturbed state and the perturbed state calculated on the surface $\varphi(x^\alpha) = 0$. Thus, the jump $\boldsymbol{\Pi}$ of the normal derivative $\boldsymbol{U}_\varphi$ is given by

$$[\boldsymbol{U}] = \boldsymbol{0}, \qquad \boldsymbol{\Pi} = [\boldsymbol{U}_\varphi] = \lim_{\varphi\to0^+}(\boldsymbol{U}_\varphi) - \lim_{\varphi\to0^-}(\boldsymbol{U}_\varphi). \tag{32}$$

From equations (17), (24) and (31) we have the following relations

$$\boldsymbol{A}^\alpha(\boldsymbol{U})\boldsymbol{U}_\varphi\varphi_\alpha = \boldsymbol{B}(\boldsymbol{U}) \qquad \text{and} \qquad \boldsymbol{A}^\alpha(\boldsymbol{U}^0)\boldsymbol{U}^0_\varphi\varphi_\alpha = \boldsymbol{B}(\boldsymbol{U}^0). \tag{33}$$

Subtracting equation $(33)_2$ from equation $(33)_1$ and by computing on the surface $\varphi(x^\alpha) = 0$, where $\boldsymbol{U} = \boldsymbol{U}^0$, $\boldsymbol{A}^\alpha(\boldsymbol{U}^0) = \boldsymbol{A}^\alpha(\boldsymbol{U})$ and $\boldsymbol{B}(\boldsymbol{U}^0) = \boldsymbol{B}(\boldsymbol{U})$, we get

$$(\boldsymbol{A}^\alpha)_0\varphi_\alpha\,[\boldsymbol{U}_\varphi] = \boldsymbol{0}, \quad \text{i.e.} \quad (\boldsymbol{A}^\alpha)_0\varphi_\alpha\boldsymbol{\Pi} = \boldsymbol{0}, \tag{34}$$

where $(\boldsymbol{A}^\alpha)_0\varphi_\alpha$ represents a 8×8 matrix and equation $(34)_2$ is a homogeneous system in the 8 components of $\boldsymbol{\Pi}$.

Introducing the quantities $\boldsymbol{n}$ and λ, defined in equation $(27)_1$, the system $(34)_2$ takes the form

$$\left(\boldsymbol{A}^i n_i - \lambda \boldsymbol{A}^0\right)\boldsymbol{\Pi} = \boldsymbol{0}, \tag{35}$$

which has non-trivial solutions only if its determinant vanishes, i.e.

$$\det\left\|\boldsymbol{A}_n - \lambda \boldsymbol{A}^0\right\| = 0. \tag{36}$$

with $\boldsymbol{A}_n = \boldsymbol{A}^i n_i$.

Equation (35) shows that $\boldsymbol{\Pi}$ can be taken proportional to the right-eigenvector $\boldsymbol{d}$ of $\boldsymbol{A}_n$, corresponding to some eigenvalue λ. In the case where the eigenvalue λ has multiplicity 1, $\boldsymbol{\Pi}$ has the form [2]

$$\boldsymbol{\Pi} = \pi(x^i, t)\boldsymbol{d}. \tag{37}$$

Thus, in order to determine $\boldsymbol{\Pi}$ we have to determine the function $\pi = \pi(x^\alpha)$.

Since $\boldsymbol{A}^0$ is a non-singular matrix, the system (17) can be written in the form

$$\boldsymbol{U}_t + (\boldsymbol{A}^0)^{-1}\boldsymbol{A}^i\boldsymbol{U}_i = (\boldsymbol{A}^0)^{-1}\boldsymbol{B}(\boldsymbol{U}), \qquad (i = 1, 2, 3), \tag{38}$$

where

$$(\boldsymbol{A}^0)^{-1} = \begin{pmatrix} \frac{1}{\rho} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\tau^{j^c}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\tau^{j^c}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\tau^{j^c}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{\tau^\nu} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\tau^\nu} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\tau^\nu} \end{pmatrix} \tag{39}$$

is the inverse matrix of $\boldsymbol{A}^0$.

In the following we continue to call $\boldsymbol{A}^i$ the matrices $(\boldsymbol{A}^0)^{-1}\boldsymbol{A}^i$, and $\boldsymbol{B}$ the vector $(\boldsymbol{A}^0)^{-1}\boldsymbol{B}$, this allows the system to take the following form

$$\boldsymbol{U}_t + \boldsymbol{A}^i\boldsymbol{U}_i = \boldsymbol{B}(\boldsymbol{U}) \qquad (i = 1, 2, 3). \tag{40}$$

From equation (36) we have the following eigenvalues problem

$$\det\|\boldsymbol{A}_n - \lambda\boldsymbol{I}\| = 0. \tag{41}$$

4 Wave front and first approximation of $\boldsymbol{U}$

Let us consider the propagation in a uniform state $\boldsymbol{U}^0$. The quantity $(27)_2$ becomes zero on the wavefront $\varphi(x^\alpha) = 0$ having velocity $\lambda = \lambda^0$ (see (26)), i.e.

$$\Psi(\boldsymbol{U}^0, \varphi_\alpha) = \varphi_t + \lambda^0(\boldsymbol{U}^0, \boldsymbol{n}^0)|\mathrm{grad}\,\varphi| = \Psi^0 = 0. \tag{42}$$

To solve equation (42) the *characteristic rays*, called characteristic curves of the system (17), are introduced, defined by the following differential equations

$$\frac{dx^\alpha}{d\sigma} = \frac{\partial \Psi^0}{\partial \varphi_\alpha}, \quad \frac{d\varphi_\alpha}{d\sigma} = -\frac{\partial \Psi^0}{\partial x^\alpha}, \tag{43}$$

where σ is the time along the characteristic rays.

From equations (42) and $(43)_2$ we obtain $\dfrac{\partial \Psi^0}{\partial x^\alpha} = 0$, from which we obtain that φ_α are constant along the characteristic rays. By integration of equation $(43)_1$ one obtains

$$x^0 = t = \sigma, \tag{44}$$

$$x^i(\sigma) = x_0^i + \Lambda_i^0 \sigma, \quad (i = 1, 2, 3), \tag{45}$$

with

$$x_0^i = x^i|_{\sigma=0} \quad \text{and} \quad \Lambda_i^0 = \Lambda_i(\boldsymbol{U}^0, \boldsymbol{n}^0), \quad (i = 1, 2, 3). \tag{46}$$

If we denote by φ^0 the given initial surface, we have $\varphi|_{\sigma=0} = \varphi^0(x_0^i)$ and $\boldsymbol{n}^0$ represents the unit normal vector to the wavefront at the point x_0^i defined by

$$\boldsymbol{n}^0 = \left(\frac{\mathrm{grad}\,\varphi}{|\mathrm{grad}\,\varphi|}\right)_{\sigma=0} = \frac{\mathrm{grad}^0\,\varphi^0}{\left|\mathrm{grad}^0\,\varphi^0\right|}, \quad \text{with} \quad (\mathrm{grad}^0)_i \equiv \frac{\partial}{\partial x_0^i} \quad (i = 1, 2, 3). \tag{47}$$

Thus, $\boldsymbol{x} = \boldsymbol{x}|_{\sigma=0} + \boldsymbol{\Lambda}^0\sigma$ and since the Jacobian J of the transformation $\boldsymbol{x} \to \boldsymbol{x}|_{t=0}$ is non-vanishing, i.e.

$$J = \det\left|\frac{\partial \Lambda_k^0}{\partial x_0^i}\sigma + \delta_{ik}\right| \neq 0 \quad (i, k = 1, 2, 3), \tag{48}$$

it follows that in the first approximation we obtain

$$\varphi(\sigma, x^i) = \varphi^0(x^i - \Lambda_i^0\sigma). \tag{49}$$

Then, expanding the vector $\boldsymbol{U}$ up the first-order approximation in a neighborhood of the wavefront $\varphi(x^\alpha) = 0$ we have

$$\boldsymbol{U} = (\boldsymbol{U})_{\varphi=0^+} + \left(\frac{\partial \boldsymbol{U}}{\partial \varphi}\right)_{\varphi=0^+} \varphi + \mathcal{O}(\varphi^2), \tag{50}$$

$$\boldsymbol{U}^0 = \left(\boldsymbol{U}^0\right)_{\varphi=0^-} + \left(\frac{\partial \boldsymbol{U}^0}{\partial \varphi}\right)_{\varphi=0^-} \varphi + \mathcal{O}(\varphi^2). \tag{51}$$

Subtracting equation (51) from (50) we get

$$\boldsymbol{U} = \boldsymbol{U}^0 + \varphi \boldsymbol{\Pi} + \mathcal{O}(\varphi^2), \tag{52}$$

where $\mathcal{O}(\varphi^2)$ is the Landau's notation and represents infinitesimals of higher order respect to φ.

5 Bernoulli's equation for the amplitude of discontinuity π

In (52), following [2], the amplitude of discontinuity π satisfies *Bernoulli's equation* having the form

$$\left(\boldsymbol{l}^0 \cdot \boldsymbol{d}^0\right)\left[\frac{d\pi}{d\sigma} + \left(\nabla\Psi \cdot \boldsymbol{d}\right)^0 \pi^2 + \frac{d}{d\sigma}\ln\sqrt{J}\,\pi\right] = F^0\pi, \tag{53}$$

$$\left(\nabla\Psi \cdot \boldsymbol{d}\right)^0 = \left(|\mathrm{grad}\,\varphi|\nabla\lambda \cdot \boldsymbol{d}\right)^0, \qquad F^0 = \left(\nabla(\boldsymbol{l}\cdot\boldsymbol{B})\cdot\boldsymbol{d}\right)^0, \tag{54}$$

with $\nabla \equiv \left(\frac{\partial}{\partial c}, \frac{\partial}{\partial j_1^c}, \frac{\partial}{\partial j_2^c}, \frac{\partial}{\partial j_3^c}, \frac{\partial}{\partial r}, \frac{\partial}{\partial \mathcal{V}_1}, \frac{\partial}{\partial \mathcal{V}_2}, \frac{\partial}{\partial \mathcal{V}_3}\right)$ and $\boldsymbol{d}^0$ the right eigenvector corresponding to the eigenvalue λ^0. We remind that the quantity $\boldsymbol{l}^0 \cdot \boldsymbol{d}^0$ is always different from zero by virtue of the hyperbolicity of the system [8].

In [2] it was seen that is possible to solve the equation (53) with the position

$$\pi = \frac{h(\sigma)}{\sqrt{J}\,\Phi(\sigma)}, \quad \text{with} \quad h(\sigma) = h^0 \exp\left[-\int_0^\sigma \frac{F^0}{(\boldsymbol{l}\cdot\boldsymbol{d})_0} d\tau\right], \tag{55}$$

$$\Phi(\sigma) = 1 + \int_0^\sigma \frac{(\nabla\Psi\cdot\boldsymbol{d})_0}{\sqrt{J}}\, h d\tau$$

and $h^0 = h(0) = \pi(0)$.

From (55) follows that if there exists a time σ_c where $\Phi(\sigma_c) = 0$, then $\pi \to \infty$, and this may correspond to the formation of a shock wave [2].

6 One-dimensional case

In this section, we treat the one-dimensional case. Confining us to the porosity and the fluid concentration fields and their fluxes, assuming that the propagation of weak discontinuity waves is along the x axis, the involved quantities depend on x_1, denoted by x, and $x_2 = x_3 = 0$, the system (13)-(16) takes the following form:

$$\frac{\partial c}{\partial t} + \frac{1}{\rho}\frac{\partial j_1^c}{\partial x} = 0, \tag{56}$$

$$\frac{\partial j_1^c}{\partial t} - \frac{\alpha_c}{\tau^{j^c}}\frac{\partial r}{\partial x} + \frac{\rho D_c}{\tau^{j^c}}\frac{\partial c}{\partial x} = -\frac{j_1^c}{\tau^{j^c}}, \tag{57}$$

$$\frac{\partial j_2^c}{\partial t} = -\frac{j_2^c}{\tau^{j^c}}, \tag{58}$$

$$\frac{\partial j_3^c}{\partial t} = -\frac{j_3^c}{\tau^{j^c}}, \tag{59}$$

$$\frac{\partial r}{\partial t} + \frac{\partial \mathcal{V}_1}{\partial x} = 0, \tag{60}$$

$$\frac{\partial \mathcal{V}_1}{\partial t} + \frac{D_\nu}{\tau^\nu}\frac{\partial r}{\partial x} - \frac{\alpha_\nu}{\tau^\nu}\frac{\partial c}{\partial x} = -\frac{\mathcal{V}_1}{\tau^\nu}, \tag{61}$$

$$\frac{\partial \mathcal{V}_2}{\partial t} = -\frac{\mathcal{V}_2}{\tau^\nu}, \tag{62}$$

$$\frac{\partial \mathcal{V}_3}{\partial t} = -\frac{\mathcal{V}_3}{\tau^\nu}. \tag{63}$$

where we recall that we have assumed $\alpha_c = \alpha_c(r)$ and $\alpha_\nu = \alpha_\nu(c)$. From the system of equations (56)-(63) we have

$$j_2^c(x,t) = (j_2^c)^0(x)e^{-\frac{t}{\tau^{j^c}}}, \qquad j_3^c(x,t) = (j_3^c)^0(x)e^{-\frac{t}{\tau^{j^c}}}, \tag{64}$$

$$\mathcal{V}_2(x,t) = (\mathcal{V}_2)^0(x)e^{-\frac{t}{\tau^\nu}}, \qquad \mathcal{V}_3(x,t) = (\mathcal{V}_3)^0(x)e^{-\frac{t}{\tau^\nu}}. \tag{65}$$

Then, we consider the following reduced system, where we have indicated j_1^c and $\mathcal{V}_1$ by the symbols j^c and $\mathcal{V}$

$$\boldsymbol{U}_t + \boldsymbol{A}(\boldsymbol{U})\boldsymbol{U}_x = \boldsymbol{B}(\boldsymbol{U}), \tag{66}$$

where

$$\boldsymbol{U} = (c, j^c, r, \mathcal{V})^T, \tag{67}$$

$$\boldsymbol{B} = \left(0, -\frac{j^c}{\tau^{j^c}}, 0, -\frac{\mathcal{V}}{\tau^\nu}, \right)^T, \tag{68}$$

and

$$\boldsymbol{A} = \begin{pmatrix} 0 & \frac{1}{\rho} & 0 & 0 \\ \frac{\rho D_c}{\tau^{j^c}} & 0 & -\frac{\alpha_c}{\tau^{j^c}} & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{\alpha_\nu}{\tau^\nu} & 0 & \frac{D_\nu}{\tau^\nu} & 0 \end{pmatrix}. \tag{69}$$

In the considered one dimensional case, in equation (41), being $\boldsymbol{n} = (n_1, 0, 0) = (1, 0, 0)$, we have $\boldsymbol{A}_n(\boldsymbol{U}) = \boldsymbol{A}$.

6.1 Eigenvalues and eigenvectors of the matrix $\boldsymbol{A}$

The matrix $\boldsymbol{A}$ admits the following simple eigenvalues:

$$\lambda_1^{(\pm)} = \pm\sqrt{\frac{\rho D_\nu \tau^{j^c} + \rho D_c \tau^\nu - \mathcal{M}}{2\rho\tau^{j^c}\tau^\nu}}, \quad \lambda_2^{(\pm)} = \pm\sqrt{\frac{\rho D_\nu \tau^{j^c} + \rho D_c \tau^\nu + \mathcal{M}}{2\rho\tau^{j^c}\tau^\nu}}, \tag{70}$$

where

$$\mathcal{M} = \sqrt{\left(\rho D_\nu \tau^{j^c} - \rho D_c \tau^\nu\right)^2 + 4\rho\alpha_c\alpha_\nu\tau^{j^c}\tau^\nu}. \tag{71}$$

The radicand of the quantity $\mathcal{M}$ is always positive, so $\mathcal{M}$ is real, moreover the eigenvalues $\lambda_1^{(\pm)}$ are real when the condition $\rho D_\nu \tau^{j^c} + \rho D_c \tau^\nu - \mathcal{M} \geq 0$ is valid, i.e. for

$$\alpha_c\alpha_\nu \leq \rho D_c D_\nu. \tag{72}$$

Thus, we have the propagation of discontinuity waves, with the normal velocity $\lambda_1^{(\pm)}$ when the condition (72) is valid. The eigenvalues $\lambda_2^{(\pm)}$ are always real. We observe that $\mathcal{M} \neq 0$ (so that $\lambda_1^{(\pm)} \neq \lambda_2^{(\pm)}$) because the relation $\left(\rho D_\nu \tau^{j^c} - \rho D_c \tau^\nu\right)^2 \neq -4\rho\alpha_c\alpha_\nu\tau^{j^c}\tau^\nu$ is always verified, being the left-hand member positive whereas the right-hand member is negative (the two members are equal only if they are both null but the right-hand member is different from zero by hypothesis).

The left eigenvectors $\boldsymbol{l}_1^{(\pm)}$, $\boldsymbol{l}_2^{(\pm)}$, and the right eigenvectors $\boldsymbol{d}_1^{(\pm)}$, $\boldsymbol{d}_2^{(\pm)}$

corresponding to eigenvalues $\lambda_1^{(\pm)}$, $\lambda_2^{(\pm)}$, respectively, have the form

$$\boldsymbol{l}_1^{(\pm)} = \left(\frac{\lambda_1^{(\pm)}\mathcal{N}}{2\alpha_c\tau^\nu}, \frac{\mathcal{N}}{2\rho\alpha_c\tau^\nu}, \lambda_1^{(\pm)}, 1 \right), \quad \boldsymbol{l}_2^{(\pm)} = \left(\frac{\lambda_2^{(\pm)}\mathcal{T}}{2\alpha_c\tau^\nu}, \frac{\mathcal{T}}{2\rho\alpha_c\tau^\nu}, \lambda_2^{(\pm)}, 1 \right), \tag{73}$$

$$\boldsymbol{d}_1^{(\pm)} = \left(\frac{2\alpha_c(\tau^{j^c})^2\lambda_1^{(\pm)}}{\mathcal{I}}, -\frac{\alpha_c\tau^\nu\mathcal{F}}{\tau^{j^c}\mathcal{I}}, \frac{\lambda_1^{(\pm)}\tau^\nu\mathcal{T}}{\mathcal{I}}, 1 \right)^T, \tag{74}$$

$$\boldsymbol{d}_2^{(\pm)} = \left(\frac{2\alpha_c(\tau^{j^c})^2\lambda_2^{(\pm)}}{\mathcal{Z}}, -\frac{\alpha_c\tau^\nu\mathcal{G}}{\tau^{j^c}\mathcal{Z}}, \frac{\lambda_2^{(\pm)}\tau^\nu\mathcal{N}}{\mathcal{Z}}, 1 \right)^T, \tag{75}$$

with

$$\mathcal{N} = \rho D_\nu \tau^{j^c} - \rho D_c \tau^\nu + \mathcal{M}, \quad \mathcal{T} = \rho D_\nu \tau^{j^c} - \rho D_c \tau^\nu - \mathcal{M}, \tag{76}$$

$$\mathcal{F} = \rho D_\nu \tau^{j^c} + \rho D_c \tau^\nu - \mathcal{M}, \quad \mathcal{G} = \rho D_\nu \tau^{j^c} + \rho D_c \tau^\nu + \mathcal{M}, \tag{77}$$

$$\mathcal{I} = D_\nu \left(\rho D_c \tau^\nu + \mathcal{M}\right) - \rho D_\nu^2 \tau^{j^c} - 2\alpha_c\alpha_\nu\tau^\nu, \tag{78}$$

$$\mathcal{Z} = D_\nu \left(\rho D_c \tau^\nu - \mathcal{M}\right) - \rho D_\nu^2 \tau^{j^c} - 2\alpha_c\alpha_\nu\tau^\nu.$$

The quantities $\mathcal{I}$, $\mathcal{Z}$ are assumed different than zero and this leads to the condition $\alpha_c\alpha_\nu \neq \rho D_c D_\nu$ (according to result (72)).

The discontinuity waves which are propagating with the velocity given by $\lambda_1^{(\pm)}$ and $\lambda_2^{(\pm)}$ are not exceptional waves in the sense of Lax-Boillat [2] (i.e. they can degenerate into shock waves) in the case where

$$\nabla\lambda_1^{(\pm)} \cdot \boldsymbol{d}_1^{(\pm)} = \mp\frac{\tau^\nu}{2\mathcal{M}\mathcal{I}} \left[2\rho\alpha_c^2\tau^\nu \frac{\partial\alpha_\nu}{\partial c} + \alpha_\nu\mathcal{T}\frac{\partial\alpha_c}{\partial r} \right] \neq 0, \tag{79}$$

$$\nabla\lambda_2^{(\pm)} \cdot \boldsymbol{d}_2^{(\pm)} = \pm\frac{\tau^\nu}{2\mathcal{M}\mathcal{Z}} \left[2\rho\alpha_c^2\tau^\nu \frac{\partial\alpha_\nu}{\partial c} + \alpha_\nu\mathcal{N}\frac{\partial\alpha_c}{\partial r} \right] \neq 0, \tag{80}$$

being

$$\nabla\lambda_1^{(\pm)} = \left(\mp\frac{\alpha_c}{2\mathcal{M}\lambda_1^{(\pm)}}\frac{\partial\alpha_\nu}{\partial c}, 0, \mp\frac{\alpha_\nu}{2\mathcal{M}\lambda_1^{(\pm)}}\frac{\partial\alpha_c}{\partial r}, 0 \right), \tag{81}$$

$$\nabla\lambda_2^{(\pm)} = \left(\pm\frac{\alpha_c}{2\mathcal{M}\lambda_2^{(\pm)}}\frac{\partial\alpha_\nu}{\partial c}, 0, \pm\frac{\alpha_\nu}{2\mathcal{M}\lambda_2^{(\pm)}}\frac{\partial\alpha_c}{\partial r}, 0 \right). \tag{82}$$

The scalar product between eigenvectors left and right corresponding to the same eigenvalue satisfy the following relations

$$\boldsymbol{l}_1^{(\pm)} \cdot \boldsymbol{d}_1^{(\pm)} = 1 + \frac{\mathcal{F}(\mathcal{N}+2\mathcal{M})}{2\rho\tau^{j^c}\mathcal{I}}, \quad \boldsymbol{l}_2^{(\pm)} \cdot \boldsymbol{d}_2^{(\pm)} = 1 + \frac{\mathcal{G}(3\rho D_\nu\tau^{j^c} - 3\rho D_c\tau^\nu - \mathcal{M})}{2\rho\tau^{j^c}\mathcal{Z}}, \tag{83}$$

whose values are supposed different than zero and this assumption ensures hyperbolicity of the system (66) and the propagation of the weak discontinuity waves, see (53). In the following we take into consideration $\lambda = \lambda_2^{(+)}$, corresponding to a progressive fast wave traveling to the right. Analogous results are valid for the waves propagating with the other velocities.

7 Determination of the approximated solution of the PDEs system

In this section, we consider an uniform unperturbed state in which $\boldsymbol{U}^0$, solution of the system (66), has the form

$$\boldsymbol{U}^0 = (c^0, (j^c)^0, r, (\mathcal{V})^0). \tag{84}$$

In $\boldsymbol{U}^0$ we have

$$\left(\lambda_2^{(+)}\right)^0 = \sqrt{\frac{\rho D_\nu \tau^{j^c} + \rho D_c \tau^\nu + \mathcal{M}^0}{2\rho\tau^{j^c}\tau^\nu}}, \tag{85}$$

$$\left(\boldsymbol{l}_2^{(+)}\right)^0 = \left(\frac{\left(\lambda_2^{(+)}\right)^0 \mathcal{T}^0}{2\alpha_c^0 \tau^\nu}, \frac{\mathcal{T}^0}{2\rho\alpha_c^0\tau^\nu}, \left(\lambda_2^{(+)}\right)^0, 1 \right),$$

$$\left(\boldsymbol{d}_2^{(+)}\right)^0 = \left(\frac{2\alpha_c^0 (\tau^{j^c})^2 \left(\lambda_2^{(+)}\right)^0}{\mathcal{Z}^0}, -\frac{\alpha_c^0 \tau^\nu \mathcal{G}^0}{\tau^{j^c}\mathcal{Z}^0}, \frac{\left(\lambda_2^{(+)}\right)^0 \tau^\nu \mathcal{N}^0}{\mathcal{Z}^0}, 1 \right), \tag{86}$$

$$\left(\boldsymbol{l}_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 = 1 + \frac{\mathcal{G}^0(3\rho D_\nu \tau^{j^c} - 3\rho D_c \tau^\nu - \mathcal{M}^0)}{2\rho\tau^{j^c}\mathcal{Z}^0} \tag{87}$$

and

$$\left(\nabla\lambda_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 = \frac{\tau^\nu}{2\mathcal{M}^0\mathcal{Z}^0}\left[2\rho(\alpha_c^0)^2\tau^\nu\left(\frac{\partial\alpha_\nu}{\partial c}\right)^0 + \alpha_\nu^0\mathcal{N}^0\left(\frac{\partial\alpha_c}{\partial r}\right)^0\right], \tag{88}$$

where the symbols " ^{0} " indicates that the quantities are calculated in $\boldsymbol{U}^0$.

The radial velocity along the characteristic rays is

$$\boldsymbol{\Lambda}^0(\boldsymbol{U}^0, \boldsymbol{n}^0) = \left(\lambda_2^{(+)}\right)^0 \boldsymbol{n}^0 = \left(\left(\lambda_2^{(+)}\right)^0, 0, 0\right). \tag{89}$$

and the characteristic rays are

$$\frac{dt}{d\sigma} = 1, \qquad \frac{dx_i}{d\sigma} = \frac{\partial \Psi^0}{\partial \varphi_\alpha} = \left(\lambda_2^{(+)}\right)^0, \qquad \frac{d\varphi_\alpha}{d\sigma} = 0. \tag{90}$$

By integrating $(90)_1$ we have

$$x^0 = t = \sigma, \qquad x(\sigma) = (x)_0 + \left(\lambda_2^{(+)}\right)^0 \sigma, \tag{91}$$

and the wave front takes the form

$$\varphi(x, \sigma) = \varphi^0 \left(x - \left(\lambda_2^{(+)}\right)^0 \sigma \right). \tag{92}$$

We remark that in this particular case, since $\boldsymbol{n}^0$ is constant, Λ_i^0 do not depends on x_0^i, so from (48) we deduce $J = 1$.

The amplitude π satisfies the following equation (see equation (55) with $J = 1$)

$$\left(\boldsymbol{l}_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 \left[\frac{d\pi}{d\sigma} + \left|\varphi_x^0\right| \left(\nabla \lambda_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 \pi^2 \right] = F^0 \, \pi, \tag{93}$$

where $F^0 = \left[\nabla \left(\boldsymbol{l}_2^{(+)} \cdot \boldsymbol{B}\right) \cdot \boldsymbol{d}_2^{(+)}\right]^0$.

By virtue of the hyperbolicity of the system, we have $\left(\boldsymbol{l}_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 \neq 0$.

Taking into account that

$$\boldsymbol{l}_2^{(+)} \cdot \boldsymbol{B} = -\frac{j^c \mathcal{T}}{2\rho \alpha_c \tau^{j^c} \tau^\nu} - \frac{\mathcal{V}}{\tau^\nu}, \tag{94}$$

(95)

$$\nabla \left(\boldsymbol{l}_2^{(+)} \cdot \boldsymbol{B}\right) = \Big(\frac{j^c}{\mathcal{M}} \frac{\partial \alpha_\nu}{\partial c}, \; -\frac{\mathcal{T}}{2\rho \alpha_c \tau^{j^c} \tau^\nu}, \\ \frac{j^c}{2\rho \alpha_c^2 \tau^{j^c} \tau^\nu} \left(\frac{2\rho \alpha_c \alpha_\nu \tau^{j^c} \tau^\nu}{\mathcal{M}} + \mathcal{T} \right) \frac{\partial \alpha_c}{\partial r}, \; -\frac{1}{\tau^\nu} \Big),$$

being $\nabla \equiv \left(\frac{\partial}{\partial c}, \frac{\partial}{\partial j^c}, \frac{\partial}{\partial r}, \frac{\partial}{\partial \mathcal{V}}\right)$, in $\boldsymbol{U}^0$ we obtain

$$\left[\nabla \left(\boldsymbol{l}_2^{(+)} \cdot \boldsymbol{B}\right)\right]^0 = \Big(\frac{(j^c)^0}{\mathcal{M}^0} \left(\frac{\partial \alpha_\nu}{\partial c} \right)^0, -\frac{\mathcal{T}^0}{2\rho \alpha_c^0 \tau^{j^c} \tau^\nu}, \\ \frac{(j^c)^0}{2\rho \left(\alpha_c^0\right)^2 \tau^{j^c} \tau^\nu} \left(\frac{2\rho \alpha_c^0 \alpha_\nu^0 \tau^{j^c} \tau^\nu}{\mathcal{M}^0} + \mathcal{T}^0 \right) \left(\frac{\partial \alpha_c}{\partial r} \right)^0, -\frac{1}{\tau^\nu} \Big) \tag{96}$$

and

$$F^0 = \frac{2\alpha_c^0(\tau^{j^c})^2(j^c)^0\left(\lambda_2^{(+)}\right)^0}{\mathcal{M}^0\mathcal{Z}^0}\left(\frac{\partial\alpha_\nu}{\partial c}\right)^0 + \frac{(j^c)^0\mathcal{N}^0\left(\lambda_2^{(+)}\right)^0}{2\rho\mathcal{Z}^0\left(\alpha_c^0\right)^2\tau^{j^c}}\left(\frac{2\rho\alpha_c^0\alpha_\nu^0\tau^{j^c}\tau^\nu}{\mathcal{M}^0}+\mathcal{T}^0\right)\left(\frac{\partial\alpha_c}{\partial r}\right)^0 + \frac{\mathcal{G}^0\mathcal{T}^0}{2\rho(\tau^{j^c})^2\mathcal{Z}^0}-\frac{1}{\tau^\nu}. \tag{97}$$

Thus, we derive $\pi = \dfrac{h(\sigma)}{\Phi(\sigma)}$ (see equation (55) with $J=1$), being

$$h(\sigma) = \pi^0\exp\left[-\frac{F^0}{\left(\boldsymbol{l}_2^{(+)}\cdot\boldsymbol{d}_2^{(+)}\right)^0}\sigma\right], \tag{98}$$

with $\pi^0 = \pi(0)$,

$$\Phi(\sigma) = 1-\frac{\pi^0}{H^0}\left|\varphi_x^0\right|\left(\nabla\lambda_2^{(+)}\cdot\boldsymbol{d}_2^{(+)}\right)^0(e^{-H^0\sigma}-1), \tag{99}$$

where $H^0 = \dfrac{F^0}{\left(\boldsymbol{l}_2^{(+)}\cdot\boldsymbol{d}_2^{(+)}\right)^0}$.

From equations (98) and (99) we obtain

$$\pi(\sigma) = \frac{\pi^0H^0e^{-H^0\sigma}}{H^0-\pi^0|\varphi_x^0|\left(\nabla\lambda_2^{(+)}\cdot\boldsymbol{r}_2^{(+)}\right)^0(e^{-H^0\sigma}-1)}. \tag{100}$$

We observe that the amplitude $\pi(\sigma)$ is a limited function when

$$\lim_{\sigma\to\infty}\pi(\sigma) = \begin{cases} -\frac{H^0}{|\varphi_x^0|\left(\nabla\lambda_2^{(+)}\cdot\boldsymbol{d}_2^{(+)}\right)^0}, & \text{if } H^0<0,\\ 0, & \text{if } H^0>0. \end{cases} \tag{101}$$

Finally, by virtue of relations (84), (87), (88), (93) and (100) we derive the explicit form of the first approximation of the solution $\boldsymbol{U}$ of the system (66)

$$(c, j^c, r, \mathcal{V})^T = (c^0, (j^c)^0, r, (\mathcal{V})^0)^T$$

$$+\Gamma^0 \left(\frac{2\alpha_c^0(\tau^n)^2 \left(\lambda_2^{(+)}\right)^0}{\mathcal{Z}^0}, -\frac{\alpha_c^0 \tau^\nu \mathcal{G}^0}{\tau^{j^c}\mathcal{Z}^0}, \frac{\left(\lambda_2^{(+)}\right)^0 \tau^\nu \mathcal{N}^0}{\mathcal{Z}^0}, 1 \right)^T, \tag{102}$$

with

$$\Gamma^0 = \frac{\varphi^0\left(x - \left(\lambda_2^{(+)}\right)^0 \sigma\right) \pi^0 H^0 e^{-H^0\sigma}}{H^0 - \pi^0|\varphi_x^0| \left(\nabla\lambda_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 (e^{-H^0\sigma} - 1)}. \tag{103}$$

Thus, from (102) and (103) we have

$$c(x,\sigma) = c^0 + \frac{2\varphi^0\left(x - \left(\lambda_2^{(+)}\right)^0 \sigma\right) \pi^0 H^0 \alpha_c^0 (\tau^n)^2 \left(\lambda_2^{(+)}\right)^0 e^{-H^0\sigma}}{\mathcal{Z}^0 \left[H^0 - \pi^0|\varphi_x^0| \left(\nabla\lambda_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 (e^{-H^0\sigma} - 1)\right]}, \tag{104}$$

$$j^c(x,\sigma) = (j^c)^0 - \frac{\varphi^0\left(x - \left(\lambda_2^{(+)}\right)^0 \sigma\right) \pi^0 H^0 \alpha_c^0 \tau^\nu \mathcal{G}^0 e^{-H^0\sigma}}{\tau^{j^c}\mathcal{Z}^0 \left[H^0 - \pi^0|\varphi_x^0| \left(\nabla\lambda_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 (e^{-H^0\sigma} - 1)\right]}, \tag{105}$$

$$r(x,\sigma) = r^0 + \frac{\left(\lambda_2^{(+)}\right)^0 \varphi^0\left(x - \left(\lambda_2^{(+)}\right)^0 \sigma\right) \pi^0 H^0 \tau^\nu \mathcal{N}^0 e^{-H^0\sigma}}{\mathcal{Z}^0 \left[H^0 - \pi^0|\varphi_x^0| \left(\nabla\lambda_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 (e^{-H^0\sigma} - 1)\right]}, \tag{106}$$

$$\mathcal{V}(x,\sigma) = (\mathcal{V})^0 + \frac{\varphi^0\left(x - \left(\lambda_2^{(+)}\right)^0 \sigma\right) \pi^0 H^0 e^{-H^0\sigma}}{H^0 - \pi^0|\varphi_x^0| \left(\nabla\lambda_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0 (e^{-H^0\sigma} - 1)}. \tag{107}$$

In the case where in (100) $\Phi(\sigma_c) = 0$, with $\Phi(\sigma_c)$ given by equation (99), there exists a critical time σ_c given by

$$\sigma_c = \frac{1}{H^0} \ln V^0, \tag{108}$$

with

$$V^0 = \left(\frac{\pi^0 \left|\varphi_x^0\right| \left(\nabla\lambda_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0}{H^0 + \pi^0|\varphi_x^0| \left(\nabla\lambda_2^{(+)} \cdot \boldsymbol{d}_2^{(+)}\right)^0} \right).$$

In this case the amplitude of the weak discontinuty wave becomes unbounded and this can correspond to the formation of a shock wave [2]. In equation (108) σ_c exists in the case where

$$\frac{1}{H^0}\ln V^0 > 0 \quad \text{and} \quad V^0 > 0. \tag{109}$$

8 Conclusions

In this paper, we have presented, in the frame of extended irreversible thermodynamics with internal variables, a model for perfect isotropic porous media satured by a fluid flow in a particular case, derived in a previous paper by the authors, taking into consideration only the porosity and the fluid concentration fields and their fluxes. But, here some coefficients of the governing equations are assumed to be non-constant, depending on the fields involved, obtaining a non-linear PDEs system. Since closed-form solutions of nonlinear PDEs are rare, we have investigated the propagation of weak discontinuity waves, as approximated solutions. To this aim we have introduced a new variable related to the surface across which the first derivatives of the solutions undergo a jump. Following a Boillat's methodology for quasi-linear and hyperbolic systems of the first order, we have investigated, in an one-dimensional case, the propagation of one solution into an uniform unperturbed state in the first approximation, obtaining the Bernoulli's equation governing the evolution of the amplitude of the weak discontinuity wave, the wave front, the velocity along the characteristic rays, the approximated solution and the critical time in which the amplitude of this wave becomes unbounded, and this may correspond to the formation of a shock wave. The worked out results can be applied in several technological sectors as medical sciences, biology and geology.

References

[1] A. Berezovski and P. Van, *Internal Variables in Thermoelasticity*, Berlin, Springer, 2017.

[2] G. Boillat, *La Propagation des Ondes*, Gauthier-Villars, Paris, 1965.

[3] G. Boillat, Ondes asymptotiques nonlinéaires, *Annali di Matematica Pura ed Applicata* 91(IV) (1976), 31-44.

[4] G. Boillat and T. Ruggeri, On the evolution law of weak discontinuities for a hyperbolic quasi linear systems, *Wave Motion* 1 (1979), 149-152.

[5] Y. Choquet-Bruhat, Ondes asymptotiques et approchées pour systèmes d'équations aux dérivées partielles nonlinéaires, *J. Math. Pures Appl.*, 1968.

[6] V. Ciancio and L. Restuccia, Asymptotic waves in anelastic media without memory (Maxwell media), *Phys. A* 131 (1985), 251-262.

[7] S.R. De Groot and P. Mazur, *Non-Equilibrium Thermodynamics*, Amsterdam, North-Holland Publishing Company, 1962.

[8] A. Donato and A.M. Greco, *Metodi Qualitativi per Onde Non Lineari*, Quaderni del C. N. R., Gruppo Nazionale di Fisica Matematica, 11th Scuola Estiva di Fisica Matematica, Ravello, September 8-20, 1986.

[9] A. Famà and L. Restuccia, Propagation of coupled porosity and fluid-concentration waves in isotropic porous media, *Electron. J. Differ. Equations* 73 (2020), 1-16.

[10] A. Famà and L. Restuccia, Non-equilibrium thermodynamics framework for fluid flow and porosity dynamics in porous isotropic media, *Ann. Acad. Rom. Sci. Ser. Math. Appl.* 12 (2020), 198-225.

[11] A. Famà, L. Restuccia and D. Jou, A simple model of porous media with elastic deformations and erosion or deposition, *Z. Angew. Math. Phys.* 71 (2020), 123-144.

[12] A. Famà and L. Restuccia, Coupled porosity-fluid concentration flux-temperature waves in isotropic porous media, *Electron. J. Differ. Equations* 88 (2021), 1-12.

[13] A. Famà and L. Restuccia, Propagation of coupled porosity-temperature waves in isotropic porous media, *Proc. Rom. Acad. Ser. A* 23 (2022), 9-16.

[14] A. Georgescu, *Asymptotic Treatment of Differential Equations*, Chapman and Hall, London, 1995.

[15] A. Georgescu and L. Restuccia, Asymptotic waves from the point of view of double-scale method, *Atti Accad. Peloritana dei Pericolanti di Messina* LXXXIV (2006), 1-9.

[16] A. Georgescu and L. Restuccia, An application of double-scale method to the study of non-linear dissipative waves in Jeffreys media, *Ann. Acad. Rom. Sci. Ser. Math. Appl.* 3 (2011), 1-20.

[17] A. Jeffrey, The development of jump discontinuities in nonlinear hyperbolic systems of equations in two independent variables, *Arch. Ration. Mech. Anal.* 14 (1963), 27-37.

[18] A. Jeffrey, The propagation of weak discontinuities in quasilinear symmetric hyperbolic systems, *Z. Angew. Math. Phys.* 14 (1963), 301-314.

[19] A. Jeffrey and T. Taniuti, *Nonlinear Wave Propagation*, Academic Press, New York, 1964.

[20] A. Jeffrey, *Quasilinear Hyperbolic Systems and Waves*, Pitman, London, 1976.

[21] D. Jou, J. Casas-Vázquez and G. Lebon, *Extended Irreversible Thermodynamics*, 4th ed., Berlin/Heidelberg, Germany, Springer, 2010.

[22] D. Jou and L. Restuccia, Mesoscopic transport equations and contemporary thermodynamics: an introduction, *Contemp. Phys.* 52 (2011), 465-474.

[23] G.A. Kluitenberg, *Plasticity and Non-Equilibrium Thermodynamics*, CISM Lecture Notes, Wien, NewYork, Springer-Verlag, 1984.

[24] J. Kubik, A macroscopic description of geometrical pore structure of porous solids, *Int. J. Eng. Sci.* 24 (1986), 971-980.

[25] P.D. Lax, *Contributions to the Theory of Partial Differential Equations*, Princeton University Press, Princeton, 1954.

[26] P.D. Lax, Nonlinear hyperbolic equations, *Commun. Pure Appl. Math.* 6 (1983), 231-258.

[27] I.S. Liu, The method of Lagrange multipliers for exploitation of the entropy principle, *Arch. Ration. Mech. Anal.* 46 (1972), 131-148.

[28] G.A. Maugin and W. Muschik, Thermodynamics with internal variables. Part I. General concepts, *J. Non-Equilib. Thermodyn.* 19 (1994), 217-249.

[29] G.A. Maugin and W. Muschik, Thermodynamics with internal variables. Part II. Applications, *J. Non-Equilib. Thermodyn.* 19 (1994), 250-289.

[30] I. Müller, *Thermodynamics*, Boston, MA, USA, Pitman Advanced Publishing Program, 1985.

[31] W. Muschik, *Aspects of Non-Equilibrium Thermodynamics*, Singapore, World Scientific, 1990.

[32] W. Muschik, Internal variables in non-equilibrium thermodynamics, *J. Non-Equilib. Thermodyn.* 15 (1990), 127-137.

[33] W. Muschik, C. Papenfuss and H. Ehrentraut, *Concepts of Continuum Thermodynamics*, Kielce, Poland, Kielce University of Technology, 1996.

[34] W. Muschik and L. Restuccia, Changing the observer and moving materials in continuum physics: objectivity and frame-indifference, *Technische Mechanik* 22 (2002), 152-160.

[35] W. Muschik and L. Restuccia, Systematic remarks on objectivity and frame-indifference, liquid crystal theory as an example, *Arch. Appl. Mech.* 78 (2008), 837-857.

[36] L. Restuccia and A. Georgescu, Determination of asymptotic waves in Maxwell media by double-scale method, *Technische Mechanik* 28 (2008), 140-151.

[37] L. Restuccia, Legge di evoluzione delle discontinuita‘ e tempo critico in un particolare modello di dielettrico non lineare, *Atti Accad. Peloritana dei Pericolanti* LVII (1979), 23-28.

[38] L. Restuccia, A thermodynamical model for fluid flow through porous solids, *Rend. Circ. Mat. Palermo* 77 (2005), 565-577.

[39] L. Restuccia, Thermomechanics of porous solids filled by fluid flow. In: *Applied and Industrial Mathematics in Italy*, Vol. 82, pp 485–495, Eds. De Bernardis, R. Spigler, and V. Valente, World Scientific, Singapore, 2010.

[40] L. Restuccia and D. Jou, Non-Local vectorial internal variables and generalized Guyer-Krumhansl evolution equations for the heat flux, *Entropy* 25 (2023), 1259-1280.

[41] L. Restuccia, L. Palese, M.T. Caccamo, A. Famà, Heat equation for porous nanostructures filled by a fluid flow, *Atti Accad. Peloritana dei Pericolanti* 97 (2019), 1-16.

[42] L. Restuccia, L. Palese, M.T. Caccamo, and A. Famà, A description of anisotropic porous nanocrystals filled by a fluid flow in the framework

of extended thermodynamics with internal variables, *Proc. Rom. Acad. Ser. A* 21 (2020), 123-130.

[43] G.F. Smith, On isotropic functions of symmetric tensors, skew-symmetric tensors and vectors, *Int. J. Eng. Sci.* 9 (1971), 899-916.

[44] C. Truesdell and R.A. Toupin, *The Classical Field Theories*, Springer Verlag, Berlin, 1960.

[45] J. Verhás, *Thermodynamics and Rheology*, Budapest, Akadémiai Kiadó and Kluwer Academic Publisher, Budapest, Hungary, 1997.

Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 18, No. 4/2026

AN EVOLUTION EQUATION WITH COMPETING p-LAPLACIAN OPERATORS*

Dumitru Motreanu†

Dedicated to Professor Constantin Fetecau on the occasion of his 80th anniversary

DOI 10.56082/annalsarscimath.2026.4.289

Abstract

The paper investigates an initial-boundary value problem involving the competing (p,q)-Laplacian $-\Delta_p + \Delta_q$. Under some conditions, which are different from the usual ellipticity and monotonicity conditions, the existence of a generalized solution is proven through a special Galerkin approach. An example illustrates the applicability of our result.

Keywords: initial-boundary value problem, time depending competing operators, p-Laplacian, generalized solution.
MSC: 35K92, 47J35, 35K20.

1 Introduction and main result

Our aim is to study the following initial-boundary value problem

$$\begin{cases} u_t - \Delta_p u + \Delta_q u = f(x,t,u) & \text{in } \Omega \times (0,T) \\ u = 0 & \text{on } \partial\Omega \times [0,T] \\ u(x,0) = u_0(x) & \text{on } \Omega, \end{cases} \tag{P}$$

*Accepted for publication on October 07, 2026
†motreanu@univ-perp.fr, Département de Mathématiques, Université de Perpignan Via Domitia, Perpignan, 66860, France

on a bounded domain Ω in $\mathbb{R}^N$ with Lipschitz boundary $\partial\Omega$ and with fixed real numbers $T > 0$, $p \geq 2$, $q \in (1, p)$, and a function $u_0 \in L^2(\Omega)$. The unknown in problem (P) is a function $u(x,t)$ that will be subsequently viewed as $u(t)(x)$, with $(x,t) \in \Omega \times (0,T)$. The notation u_t (or $u'(t)$) stands for the time derivative, i.e., the partial derivative of u with respect to the time variable t.

The main difficulty in dealing with problem (P) comes from the operator $-\Delta_p + \Delta_q$ that we call the competing (p,q)-Laplacian. This operator is not monotone and no ellipticity condition can be matched. We note that if $u \in W_0^{1,p}(\Omega)$ is a nonzero function and λ is a positive number, the expression

$$\langle -\Delta_p(\lambda u) + \Delta_q(\lambda u), \lambda u\rangle_{W_0^{1,p}(\Omega)} = \lambda^p \langle -\Delta_p u, u\rangle_{W_0^{1,p}(\Omega)} - \lambda^q \langle -\Delta_q u, u\rangle_{W_0^{1,q}(\Omega)}$$

is positive for λ sufficiently large and negative for λ small enough. Until now, problems involving competing operators like $-\Delta_p + \Delta_q$ have been examined only for stationary problems. Their study was initiated in [12] and after that developed in [6, 8, 11, 13, 15, 17, 20] (refer also to the recent book [10]).

Here, we consider for the first time evolution problems with competing operators. Due to the competing character, the monotonicity method of Browder and Minty as well as the theory of semigroups are not applicable (see, e.g., [9]).

Another major difficulty in treating problem (P) is the presence of the operator Δ_q which is not pseudomonotone. Indeed, if Δ_q were pseudomonotone on $W_0^{1,q}(\Omega)$, then for each sequence $\{u_n\}$ in $W_0^{1,q}(\Omega)$ with $u_n \rightharpoonup u$ in $W_0^{1,q}(\Omega)$ and

$$\limsup_{n\to\infty} \langle \Delta_q u_n, u_n - u\rangle_{W_0^{1,q}(\Omega)} \leq 0$$

it would hold $\lim_{n\to\infty} \langle \Delta_q u_n, u_n\rangle_{W_0^{1,q}(\Omega)} = \langle \Delta_q u, u\rangle_{W_0^{1,q}(\Omega)}$ or equivalently $u_n \to u$ in $W_0^{1,q}(\Omega)$ because the space $W_0^{1,q}(\Omega)$ is uniformly convex. Nevertheless, we always have

$$\limsup_{n\to\infty} \langle \Delta_q u_n, u_n - u\rangle_{W_0^{1,q}(\Omega)} = \limsup_{n\to\infty} \langle \Delta_q u_n - \Delta_q u, u_n - u\rangle_{W_0^{1,q}(\Omega)} \leq 0$$

due to the monotonicity of $-\Delta_q$. Therefore the pseudomonotonicity of Δ_q would lead to the contradiction that the weak convergence and strong convergence in $W_0^{1,q}(\Omega)$ are equivalent. We are thus prevented to use the surjectivity theorem for pseudomonotone operators relative to time-derivative (see [3, Theorem 2.152]) as well as the corresponding sub-supersolution approach (refer to [2, 4, 5, 14]).

The lower-order term in the equation of (P) is described by a Carathéodory function $f : \Omega \times (0,T) \times \mathbb{R} \to \mathbb{R}$ (i.e., $f(\cdot,\cdot,s)$ is measurable for all

$s \in \mathbb{R}$ and $f(x,t,\cdot)$ is continuous for a.e. $(x,t) \in \Omega \times (0,T)$). Due to the lack of needed convergence of the gradient ∇u of the solution (see [1]), we don't consider the gradient dependence in the lower-order term.

We require the following growth condition:

(H) There exist a function $\sigma \in L^{\frac{p}{p-1}}(\Omega \times (0,T))$ and a constant $a \in (0, \lambda_{1,p})$ such that

$$|f(x,t,s)| \leq \sigma(x,t) + a|s|^{p-1}$$

for a.e. $(x,t) \in \Omega \times (0,T)$ and all $s \in \mathbb{R}$,

where $\lambda_{1,p}$ denotes the first eigenvalue of $-\Delta_p$ on $W_0^{1,p}(\Omega)$, that is

$$\lambda_{1,p} = \inf_{u \in W_0^{1,p}(\Omega) \setminus \{0\}} \frac{\|\nabla u\|^p_{L^p(\Omega)}}{\|u\|^p_{L^p(\Omega)}}. \tag{1}$$

In order to overcome the lack of ellipticity and monotonicity in problem (P) we weaken the notion of solution as follows.

Definition 1. *By a generalized solution to problem (P) we mean a function $u \in L^p(0,T;W_0^{1,p}(\Omega))$ for which $u(0) = u_0$ such that there exists the generalized derivative $u' \in L^{\frac{p}{p-1}}(0,T;W^{-1,\frac{p}{p-1}}(\Omega))$ and there exists a sequence $\{u_n\}$ in $L^p(0,T;W_0^{1,p}(\Omega))$ with the properties:*

(i) *$u_n \rightharpoonup u$ in $L^p(0,T;W_0^{1,p}(\Omega))$ and $u_n(0) \rightharpoonup u_0$ and $u_n(T) \rightharpoonup u(T)$ in $L^2(\Omega)$;*

(ii) *there exists the generalized derivative $u_n' \in L^{\frac{p}{p-1}}(0,T;W^{-1,\frac{p}{p-1}}(\Omega))$ for every n and $u_n' \rightharpoonup u'$ in $L^{\frac{p}{p-1}}(0,T;W^{-1,\frac{p}{p-1}}(\Omega))$ as $n \to \infty$;*

(iii) *$-\Delta_p u_n + \Delta_q u_n \rightharpoonup f(\cdot,\cdot,u) - u'$ in $L^{\frac{p}{p-1}}(0,T;W^{-1,\frac{p}{p-1}}(\Omega))$;*

(iv)

$$\lim_{n\to\infty} \langle u_n' - \Delta_p u_n + \Delta_q u_n, u_n - u \rangle_{L^p(0,T;W_0^{1,p}(\Omega))} = 0.$$

We state our existence result.

Theorem 1. *Under hypothesis (H), problem (P) possesses a generalized solution in the sense of Definition 1.*

The proof of Theorem 1 proceeds along a special Galerkin-type approach without any use of monotonicity, ellipticity and coerciveness, relying on a

finite dimensional differential system associated with problem (P) that is resolved through the Carathéodory theorem for ordinary differential equations. In Section 5 we illustrate the applicability of Theorem 1 by an example.

The rest of the paper consists of the following material. Section 2 comprises preliminary results. Section 3 focuses on resolving the finite dimensional differential system associated with problem (P). Section 4 sets forth the analysis of the sequence required in Definition 1. Section 5 contains the proof of Theorem 1 as well as an example.

2 Preliminary results

For any normed space E we denote by $\|\cdot\|_E$ the norm of E, by E^* its dual and by $\langle\cdot,\cdot\rangle_X$ the duality pairing between E and E^*. The weak convergence is denoted by $\rightharpoonup$ and the convergence in norm by $\to$.

Throughout the paper, the space $W_0^{1,p}(\Omega)$ is endowed with the norm $\|u\|_{W_0^{1,p}} = (\int_\Omega |\nabla u|^p dx)^{\frac{1}{p}}$. The gradient of u with respect to $x \in \Omega$ is denoted ∇u, while $|\nabla u(x,t))|$ represents the Euclidean norm of the vector $\nabla u(x,t) \in \mathbb{R}^N$. The dual of $W_0^{1,p}(\Omega)$ is the space $W^{-1,\frac{p}{p-1}}(\Omega)$. Noting that $p \geq 2$, we have the evolution triple

$$W_0^{1,p}(\Omega) \subset L^2(\Omega) \subset W^{-1,\frac{p}{p-1}}(\Omega) \tag{2}$$

with each embedding dense and compact. Recall that $L^2(\Omega)$ is a Hilbert space with the scalar product $(u,v)_{L^2(\Omega)} = \int_\Omega u(x)v(x)dx$.

The space $X := L^p(0,T;W_0^{1,p}(\Omega))$ is equipped with the norm $\|u\|_X := \big(\int_0^T \|u(t)\|^p_{W_0^{1,p}} dt\big)^{\frac{1}{p}}$ making it a separable and uniformly convex Banach space. The dual of X is $X^* = L^{\frac{p}{p-1}}(0,T;W^{-1,\frac{p}{p-1}}(\Omega))$. There is a dense and continuous embedding

$$C([0,T],W_0^{1,p}(\Omega)) \subset X. \tag{3}$$

Moreover, the set P of polynomials $\sum_{i=0}^m t^i v_i$ with $v_i \in W_0^{1,p}(\Omega)$ and $m \geq 0$, is dense in X, that is,

$$\overline{P} = X. \tag{4}$$

In the statement of problem (P) there appears the time-dependent neg-

ative p-Laplacian $-\Delta_p : X \to X^*$ given by

$$\begin{aligned} \langle -\Delta_p u, v\rangle_X &= \int_0^T \langle -\Delta_p u(t), v(t)\rangle_{W_0^{1,p}(\Omega)} dt \\ &= \int_0^T \int_\Omega |\nabla u(x,t)|^{p-2}\nabla u(x,t)\nabla v(x,t) dxdt \end{aligned} \tag{5}$$

for all $u, v \in X$. The operator $-\Delta_p$ in (5) is the time realization on the space X of the ordinary negative p-Laplacian $-\Delta_p : W_0^{1,p}(\Omega) \to W^{-1,\frac{p}{p-1}}(\Omega)$. The fact that the operator in (5) is well defined is seen by

$$\int_0^T \| -\Delta_p u(t)\|^{\frac{p}{p-1}}_{W^{-1,\frac{p}{p-1}}(\Omega)} dt = \int_0^T \|u(t)\|^p_{W_0^{1,p}(\Omega)} dt \quad \text{for all } u \in X.$$

Likewise, the time-dependent negative q-Laplacian

$$-\Delta_q : L^q(0,T; W_0^{1,q}(\Omega)) \to L^{\frac{q}{q-1}}(0,T; W^{-1,\frac{q}{q-1}}(\Omega))$$

is defined by

$$\begin{aligned} \langle -\Delta_q u, v\rangle_{L^q(0,T;W_0^{1,q}(\Omega))} &= \int_0^T \langle -\Delta_q u(t), v(t)\rangle_{W_0^{1,q}(\Omega)} dt \\ &= \int_0^T \int_\Omega |\nabla u(x,t)|^{q-2}\nabla u(x,t)\nabla v(x,t) dxdt \end{aligned} \tag{6}$$

for all $u, v \in L^q(0,T; W_0^{1,p}(\Omega))$. Since $q < p$, Hölder's inequality produces a constant $C > 0$ independent of t such that

$$\left(\int_0^t \int_\Omega |\nabla u(x,s)|^q dxds\right)^{\frac{1}{q}} \le C \left(\int_0^t \int_\Omega |\nabla u(x,s)|^p dxds\right)^{\frac{1}{p}} \tag{7}$$

for all $u \in X$ and $t \in [0,T]$. Taking $t = T$, we obtain the continuous embedding $X \subset L^q(0,T; W_0^{1,q}(\Omega))$. Therefore the operator $-\Delta_p + \Delta_q$ is well defined on the space X, so the formulation of problem (P) is meaningful.

The following lemma provides important properties of the operator $-\Delta_p$ in (5). Similarly, we can proceed with the operator $-\Delta_q$ in (6).

Lemma 1. *The time-dependent negative p-Laplacian $-\Delta_p : X \to X^*$ is well defined, demicontinuous, strictly monotone, bounded (i.e., bounded on bounded sets), and satisfies*

$$\| -\Delta_p u\|_{X^*} = \|u\|_X^{p-1} \quad \textit{for all } u \in X. \tag{8}$$

Proof. From (5) and Hölder's inequality we infer that

$$\begin{aligned}|\langle -\Delta_p u, v\rangle_X| &\le \int_0^T \int_\Omega |\nabla u(x,t)|^{p-1} |\nabla v(x,t)| dxdt \\ &\le \left(\int_0^T \int_\Omega |\nabla u(x,t)|^p dxdt \right)^{\frac{p-1}{p}} \left(\int_0^T \int_\Omega |\nabla v(x,t)|^p dxdt \right)^{\frac{1}{p}} \\ &= \|u\|_X^{p-1} \|v\|_X\end{aligned}$$

for all $u, v \in X$. It turns out that the operator $-\Delta_p$ from X into X^* is well defined and bounded. Taking $v = u$ shows that (8) holds true.

Since $-\Delta_p$ in (5) is the time realization on the space X of the negative p-Laplacian $-\Delta_p : W_0^{1,p}(\Omega) \to W^{-1,p'}(\Omega)$, the operator $-\Delta_p : X \to X^*$ is demicontinuous (see [3, Theorem 2.153]). The strict monotonicity of $-\Delta_p$ in (5) is valid because the mapping $\xi \mapsto |\xi|^{p-2}\xi$ is strictly monotone on $\mathbb{R}^N$. The proof is complete. □

Recall that the Nemytskii operator N_f associated with the Carathéodory function $f : \Omega \times (0,T) \times \mathbb{R} \to \mathbb{R}$ is defined as $N_f(u) := f(\cdot, \cdot, u(\cdot, \cdot))$.

Lemma 2. *Assume condition (H) holds. Then the mapping $N_f : X \to X^*$ is well defined, bounded and continuous.*

Proof. The growth condition in (H) guarantees that

$$f(\cdot, \cdot, u(\cdot, \cdot)) \in L^{\frac{p}{p-1}}(\Omega \times (0,T)) \quad \text{for all } u \in X. \tag{9}$$

The continuity and density of the embedding $X \subset L^p(\Omega \times (0,T))$ implies that $L^{\frac{p}{p-1}}(\Omega \times (0,T)) \subset X^*$ continuously. According to (9), through Krasnoselskii's theorem (see [7]), the mapping N_f is bounded and continuous from $L^p(\Omega \times (0,T))$ to $L^{\frac{p}{p-1}}(\Omega \times (0,T))$. Then the restriction $N_f : X \to X^*$ is bounded and continuous, which completes the proof. □

Next we briefly discuss the generalized derivative u' associated with the evolution triple (2) for a function $u \in X$. More details can be found in [3, 16, 18, 19].

The function $u' \in X^*$ is called the generalized derivative of $u \in X$ if there holds

$$\int_0^T \varphi'(t) u(t) dt = -\int_0^T \varphi(t) u'(t) dt \quad \text{for all } \varphi \in C_c^\infty(0,T).$$

A necessary and sufficient condition is that

$$\int_0^T (u(t), v)_{L^2(\Omega)} \varphi'(t) dt = -\int_0^T \langle u'(t), v \rangle_{W_0^{1,p}(\Omega)} \varphi(t) dt \tag{10}$$

for all $\varphi \in C_c^\infty(0,T)$ and $v \in W_0^{1,p}(\Omega)$.

Definition 1 shows that each generalized solution to problem (P) belongs to

$$W := \{w \in X : w' \in X^*\}, \tag{11}$$

which is a Banach space endowed with the norm $\|w\|_W = \|w\|_X + \|w'\|_{X^*}$. There is the continuous embedding

$$W \subset C([0,T], L^2(\Omega)) \tag{12}$$

(see [18, Proposition 23.23]). Due to inclusion (12), the initial condition $u(0) = u_0 \in L^2(\Omega)$ in problem (P) makes sense.

Thanks to the Lions-Aubin theorem, we have the compact embedding

$$W \subset L^p(\Omega \times (0,T)) \tag{13}$$

(see [16, p. 106]).

Lemma 3. *Assume condition (H). Then the mapping $N_f : W \to X^*$ is completely continuous (i.e., $u_n \rightharpoonup u$ in W implies $N_f(u_n) \to N_f(u)$ in X^*).*

Proof. As noted in the proof of Lemma 2, the mapping N_f is bounded and continuous from $L^p(\Omega \times (0,T))$ to X^*. In view of the compact embedding (13), the restriction of N_f to W is completely continuous, so achieving the result. □

3 Finite dimensional approximation for problem (P)

Since the space $W_0^{1,p}(\Omega)$ is separable, there is a countable family $\{w_k\}_{k \geq 1}$ of linearly independent vectors in $W_0^{1,p}(\Omega)$ such that if E_n is the linear span of $\{w_1, \ldots, w_n\}$, we have

$$\overline{\bigcup_{n=1}^{\infty} E_n} = W_0^{1,p}(\Omega). \tag{14}$$

Hence E_n is a vector subspace of $W_0^{1,p}(\Omega)$ of finite dimension n with the basis $\{w_1, \ldots, w_n\}$ that we endow with the scalar product of $L^2(\Omega)$. Given the initial condition $u_0 \in L^2(\Omega)$, we fix a sequence $\{u_{0,n}\}_{n\geq 1}$ with $u_{0,n} \in E_n$ for each n satisfying $u_{0,n} \to u_0$ in $L^2(\Omega)$. This is possible according to (14) (note $p \geq 2$). We can uniquely express $u_{0,n} = \sum\limits_{1\leq k\leq n} c_{k,n} w_k$, with $c_{k,n} \in \mathbb{R}$, $1 \leq k \leq n$.

On the interval $[0, T]$ we consider the Cauchy problem

$$\begin{cases} \sum\limits_{k=1}^{n} \theta_k'(t)(w_k, w_i)_{L^2(\Omega)} \\ \qquad\qquad +\langle(-\Delta_p + \Delta_q - N_f)(\sum\limits_{k=1}^{n} \theta_k(t) w_k), w_i\rangle_{W_0^{1,p}(\Omega)} = 0, \\ \theta_i(0) = c_{i,n}, \quad 1 \leq i \leq n, \end{cases} \tag{S_P}$$

with the unknowns $(\theta_1(t), \ldots, \theta_n(t)) \in \mathbb{R}^n$.

Proposition 1. *Assume that condition (H) holds. Then the system of ordinary differential equations (S_P) admits an absolutely continuous solution $(\theta_1(t), \ldots, \theta_n(t))$ which is defined on the entire interval $[0, T]$.*

Proof. The Gram determinant $\det(w_k, w_i)_{L^2(\Omega)}$ is nonzero because the vectors $\{w_1, \ldots, w_n\}$ are linearly independent, so the matrix $((w_k, w_i)_{L^2})_{1\leq k,i\leq n}$ is invertible. Hypothesis (H) permits to apply Carathéodory's theorem (see [19, p. 1044]) to the ordinary differential system (S_P) noting that the term involving N_f can be locally bounded by an integrable function in t due to the imposed growth condition. Consequently, there exists an absolutely continuous solution $(\theta_1(t), \cdots, \theta_n(t))$ of system (S_P) on a maximal subinterval $[0, T_n)$ of $[0, T]$. We are going to show that $T_n = T$.

Set

$$u_n(t) = \sum_{k=1}^{n} \theta_k(t) w_k. \tag{15}$$

Then system (S_P) takes the form

$$\begin{cases} \sum\limits_{k=1}^{n} \theta_k'(t)(w_k, w_i)_{L^2(\Omega)} \\ \qquad\qquad +\langle(-\Delta_p + \Delta_q - N_f)(u_n(t)), w_i\rangle_{W_0^{1,p}(\Omega)} = 0 \\ \theta_i(0) = c_{i,n}, \quad 1 \leq i \leq n. \end{cases} \tag{16}$$

It follows that

$$(u_n'(t), u_n(t))_{L^2(\Omega)} + \langle(-\Delta_p + \Delta_q - N_f)(u_n(t)), u_n(t)\rangle_{W_0^{1,p}(\Omega)} = 0 \tag{17}$$

for almost all $t \in [0, T_n)$. Integration over the interval $[0, t]$ results in

$$\frac{1}{2}\|u_n(t)\|^2_{L^2(\Omega)} + \int_0^t \langle(-\Delta_p + \Delta_q - N_f)(u_n(s)), u_n(s)\rangle_{W_0^{1,p}(\Omega)} ds = \frac{1}{2}\|u_{0,n}\|^2_{L^2(\Omega)}$$

for all $t \in [0, T_n)$. By hypothesis (H) and Hölder's inequality, in conjunction with (1), it turns out

$$\begin{aligned}
& \frac{1}{2}\|u_n(t)\|^2_{L^2(\Omega)} + \int_0^t \|u_n(s)\|^p_{W_0^{1,p}(\Omega)} ds - \int_0^t \|u_n(s)\|^q_{W_0^{1,q}(\Omega)} ds \\
\leq\ & \frac{1}{2}\|u_{0,n}\|^2_{L^2(\Omega)} + \int_0^t \int_\Omega \sigma(x) u_n(x,s) dx ds \\
& + a \int_0^t \int_\Omega |u_n(x,s)|^p dx ds \\
\leq\ & \frac{1}{2}\|u_{0,n}\|^2_{L^2(\Omega)} + \|\sigma\|_{L^{\frac{p}{p-1}}(\Omega\times(0,T))} \lambda_{1,p}^{-\frac{1}{p}} \left(\int_0^t \|u_n(s)\|^p_{W_0^{1,p}(\Omega)} ds \right)^{\frac{1}{p}} \\
& + a\lambda_{1,p}^{-1} \int_0^t \|u_n(s)\|^p_{W_0^{1,p}(\Omega)} ds.
\end{aligned}$$

Taking into account (7), this amounts to saying that

$$\begin{aligned}
& \frac{1}{2}\|u_n(t)\|^2_{L^2(\Omega)} + \left(1 - a\lambda_{1,p}^{-1}\right) \int_0^t \|u_n(s)\|^p_{W_0^{1,p}(\Omega)} ds \\
\leq\ & \frac{1}{2}\|u_{0,n}\|^2_{L^2(\Omega)} + C_1 \left(\int_0^t \|u_n(s)\|^p_{W_0^{1,p}(\Omega)} ds \right)^{\frac{1}{p}} \\
& + C_2 \left(\int_0^t \|u_n(s)\|^p_{W_0^{1,p}(\Omega)} ds \right)^{\frac{q}{p}},
\end{aligned} \tag{18}$$

with positive constants C_1 and C_2 independent of t. Using that $1 - a\lambda_{1,p}^{-1} > 0$ and $p > q > 1$, it is seen from (18) that the set $\{u_n(t) : t \in [0, T_n)\}$ is uniformly bounded in $L^2(\Omega)$. Hence the solution $(\theta_1(t), \cdots, \theta_n(t))$ of system (16) is uniformly bounded on $[0, T_n)$, which renders $T_n = T$. This completes the proof. □

Corollary 1. *The function $u_n : [0, T] \to W_0^{1,p}(\Omega)$ in (15) fulfills $u_n \in W$, where W is the space described in (11).*

Proof. Proposition 1 combined with (3) entails that $u_n \in X$. From (15) and Proposition 1 we infer that the derivative $u_n' = \sum_{k=1}^{n} \theta_k' w_k$, defined almost everywhere on $[0, T]$, is an element of X^*. Altogether, we have $u_n \in W$. □

4 The sequence of approximate solutions

Now, we focus on the sequence $\{u_n\}_{n\geq 1} \subset W$ introduced in (15). The next proposition relies on key uniform estimates for the Galerkin approximations u_n and their time derivatives u_n'.

Proposition 2. *Assume that hypothesis (H) holds. Then there exists $u \in W$ such that*

(i) $u_n \rightharpoonup u$ *in* X;

(ii) $u_n' \rightharpoonup u'$ *in* X^*, *so* $u_n \rightharpoonup u$ *in* W *and in* $C([0,T], L^2(\Omega))$;

(iii) $-\Delta_p u_n + \Delta_q u_n \rightharpoonup N_f(u) - u'$ *in* X^*;

(iv) $u(0) = u_0$ *and* $u_n(0) \rightharpoonup u_0$ *and* $u_n(T) \rightharpoonup u(T)$ *in* $L^2(\Omega)$.

Proof. From (18) we infer for $t = T$ the estimate

$$\begin{aligned} & \frac{1}{2}\|u_n(T)\|^2_{L^2(\Omega)} + \left(1 - a\lambda_{1,p}^{-1}\right)\|u_n\|^p_X \\ \leq\ & \frac{1}{2}\|u_{0,n}\|^2_{L^2(\Omega)} + C_1\|u_n\|_X + C_2\|u_n\|^q_X, \end{aligned}$$

with positive constants C_1 and C_2. Since $1 - a\lambda_{1,p}^{-1} > 0$ and $p > q > 1$, it follows that the sequence $\{u_n\}_{n\geq 1}$ is bounded in X and the sequence $\{u_n(T)\}_{n\geq 1}$ is bounded in $L^2(\Omega)$. Thanks to the reflexivity of X and of $L^2(\Omega)$ we can pass to a subsequence for which there exist $u \in X$ and $h \in L^2(\Omega)$ such that

$$u_n \rightharpoonup u \text{ in } X \text{ and } u_n(T) \rightharpoonup h \text{ in } L^2(\Omega). \tag{19}$$

The first weak limit in (19) proves item (i).

By virtue of Lemmas 1 and 2, the sequences $\{-\Delta_p u_n\}_{n\geq 1}$, $\{-\Delta_q u_n\}_{n\geq 1}$ and $\{N_f(u_n)\}_{n\geq 1}$ are bounded in X^*. Due to the reflexivity of X^*, up to subsequences there hold

$$-\Delta_p u_n \rightharpoonup \xi,\ -\Delta_q u_n \rightharpoonup \eta,\ N_f(u_n) \rightharpoonup \zeta \quad \text{in } X^*, \tag{20}$$

with some $\xi \in X^*$, $\eta \in X^*$ and $\zeta \in X^*$.

From (16) we get for almost all $t \in |0,T]$ and all $\varphi \in C^1([0,T])$ and $v \in E_n$ that

$$(u_n'(t), \varphi(t)v)_{L^2(\Omega)} + \varphi(t)\langle(-\Delta_p + \Delta_q - N_f)(u_n(t)), v\rangle_{W_0^{1,p}(\Omega)} = 0. \tag{21}$$

Let $v \in \bigcup_{n=1}^{\infty} E_n$. Fix an m with $v \in E_m$ and thus (21) is valid for all $n \geq m$. Integration by parts over $|0,T]$ in (21) entails

$$-\int_0^T \varphi'(t)(u_n(t),v)_{L^2(\Omega)}dt + \varphi(T)(u_n(T),v)_{L^2(\Omega)} - \varphi(0)(u_{0,n},v)_{L^2(\Omega)} + \int_0^T \varphi(t)\langle(-\Delta_p + \Delta_q - N_f)(u_n(t)),v\rangle_{W_0^{1,p}(\Omega)}dt = 0, \quad \forall n \geq m. \tag{22}$$

Letting $n \to \infty$, by means of (19) and (20) we find

$$-\int_0^T \varphi'(t)(u(t),v)_{L^2(\Omega)}dt + \varphi(T)(h,v)_{L^2(\Omega)} - \varphi(0)(u_0,v)_{L^2(\Omega)} + \int_0^T \varphi(t)\langle\xi - \eta - \zeta, v\rangle_{W_0^{1,p}(\Omega)}dt = 0. \tag{23}$$

Thanks to the density in (14), we have that (23) holds for all $v \in W_0^{1,p}(\Omega)$. In particular, taking $\varphi \in C_c^{\infty}(0,T)$, we arrive at

$$\int_0^T \varphi'(t)(u(t),v)_{L^2(\Omega)}dt = -\int_0^T \varphi(t)\langle(-\xi + \eta + \zeta, v\rangle_{W_0^{1,p}(\Omega)}dt$$

for all $\varphi \in C_c^{\infty}(0,T)$ and $v \in W_0^{1,p}(\Omega)$. According to (10), there exists the generalized derivative $u' \in X^*$ of u with

$$u' = -\xi + \eta + \zeta. \tag{24}$$

From (11) we conclude that $u \in W$.

We pass now to uniformly estimate the time derivatives u_n'. Return to (22) that can be written in the form

$$\int_0^T (u_n'(t),v)_{L^2(\Omega)}\varphi(t)dt + \int_0^T \varphi(t)\langle(-\Delta_p + \Delta_q - N_f)(u_n(t)),v\rangle_{W_0^{1,p}(\Omega)}dt = 0$$

whenever $\varphi \in C^1([0,T])$, $v \in E_m$ and $n \geq m$. In the limit as $n \to \infty$ we obtain

$$\lim_{n\to\infty}\left[\int_0^T (u_n'(t),\varphi(t)v)_{L^2(\Omega)}dt + \int_0^T \langle(-\Delta_p + \Delta_q - N_f)(u_n(t)),\varphi(t)v\rangle_{W_0^{1,p}(\Omega)}dt\right] = 0$$

for all $\varphi \in C^1([0,T])$ and $v \in \bigcup_{n=1}^{\infty} E_n$. Using the density properties described in (4) and (14), we get

$$\lim_{n\to\infty} \left[\int_0^T (u_n'(t), w)_{L^2(\Omega)} dt + \int_0^T \langle (-\Delta_p + \Delta_q - N_f)(u_n(t)), w \rangle_{W_0^{1,p}(\Omega)} dt \right] = 0 \quad (25)$$

for all $w \in X$. On the basis of (20) and (24), this yields $u_n' \rightharpoonup u'$ in X^*. In view of (11) and assertion (i), we infer that $u_n \rightharpoonup u$ in W. The continuous embedding (12) implies $u_n \rightharpoonup u$ in $C([0,T], L^2(\Omega))$, thereby assertion (ii) follows.

The fact that $u_n \rightharpoonup u$ in W allows us to apply Lemma 3 that provides the strong convergence $N_f(u_n) \to N_f(u)$ in X^*. Combining with (20) and (24) leads to item (iii) in the statement of proposition.

Insert (24) (with $\zeta = N_f(u)$) into (23), which gives

$$-\int_0^T \varphi'(t)(u(t), v)_{L^2(\Omega)} dt + \varphi(T)(h, v)_{L^2(\Omega)} - \varphi(0)(u_0, v)_{L^2(\Omega)} - \int_0^T \varphi(t) \langle u'(t)), v \rangle_{W_0^{1,p}(\Omega)} dt = 0$$

for all $\varphi \in C^1([0,T])$ and $v \in W_0^{1,p}(\Omega)$. The integration by parts provides

$$\varphi(T)(h, v)_{L^2(\Omega)} - \varphi(0)(u_0, v)_{L^2(\Omega)} = \varphi(T)(u(T), v)_{L^2(\Omega)} - \varphi(0)(u(0), v)_{L^2(\Omega)}.$$

Since $\varphi \in C^1([0,T])$ and $v \in W_0^{1,p}(\Omega)$ are arbitrary, and $W_0^{1,p}(\Omega)$ is dense in $L^2(\Omega)$, we deduce that $u(0) = u_0$ and $u(T) = h$. Hence item (iv) is proven. The proof is complete. □

Corollary 2. *Assume that condition (H) holds. Then the time derivative u' of the function $u \in W$ constructed in Proposition 2 can be estimated as follows*

$$\|u'\|_{X^*} \le \limsup_{n\to\infty} \| - \Delta_p u_n\|_{X^*} + \limsup_{n\to\infty} \| - \Delta_q u_n\|_{X^*} + \|N_f(u)\|_{X^*}.$$

Proof. From item (iii) in Proposition 2 and the weak lower semicontinuity of the norm we infer that

$$\begin{aligned} \|u'\|_{X^*} &\le \liminf_{n\to\infty} \| - \Delta_p u_n - \Delta_q u_n\|_{X^*} + \|N_f(u)\|_{X^*} \\ &\le \limsup_{n\to\infty} \| - \Delta_p u_n - \Delta_q u_n\|_{X^*} + \|N_f(u)\|_{X^*}. \end{aligned}$$

The stated conclusion holds true. □

5 The generalized solution

Proof of Theorem 1. Proposition 2 ensures the existence of an $u \in W$ for which assertions (i)-(iv) are satisfied with the sequence $\{u_n\}_{n\geq 1} \subset W$ constructed in Proposition 1. We check that the requirements in Definition 1 for a generalized solution to problem (P) are fulfilled. We know from Proposition 2 that $u \in L^p(0,T;W_0^{1,p}(\Omega))$ with $u' \in L^{\frac{p}{p-1}}(0,T;W^{-1,\frac{p}{p-1}}(\Omega))$. Item (iv) in Proposition 2 entails $u(0) = u_0$. Part (i) in Definition 1 is the direct consequence of items (i) and (iv) of Proposition 2. Parts (ii) and (iii) in Definition 1 are verified by the statements (ii) and (iii) of Proposition 2; respectively.

It remains to justify property (iv) in Definition 1. By (17) we have that

$$\langle u_n' - \Delta_p u_n + \Delta_q u_n - N_f(u_n), u_n\rangle_X = 0 \tag{26}$$

for all n. On the other hand, items (ii) and (iii) of Proposition 2 provide

$$\langle u_n' - \Delta_p u_n + \Delta_q u_n - N_f(u_n), u\rangle_X \to 0 \text{ as } n \to \infty. \tag{27}$$

Then (26) and (27) ensure that

$$\langle u_n' - \Delta_p u_n + \Delta_q u_n - N_f(u_n), u_n - u\rangle_X \to 0 \text{ as } n \to \infty.$$

Since $u_n \rightharpoonup u$ in W, Lemma 3 implies that

$$\langle N_f(u_n), u_n - u\rangle_X \to 0 \text{ as } n \to \infty.$$

Therefore property (iv) in Definition 1 holds true. The proof of Theorem 1 is complete.

Finally, we present an example illustrating the applicability of Theorem 1.

Example 1. Consider the initial-boundary value problem

$$\begin{cases} u_t - \Delta_p u + \Delta_q u = \rho(x,t) + \frac{|u|^{\alpha+1}}{1+|u|} & \text{in } \Omega \times (0,T) \\ u = 0 & \text{on } \partial\Omega \times (0,T) \\ u(x) = u_0(x) & \text{on } \partial\Omega \end{cases} \tag{28}$$

on a bounded domain Ω in $\mathbb{R}^N$ with Lipschitz boundary $\partial\Omega$ and for fixed real numbers $T > 0$, $p \geq 2$, $q \in (1,p)$, $\alpha \in [0, p-1)$, and given functions $\rho \in L^{\frac{p}{p-1}}(\Omega \times (0,T))$ and $u_0 \in L^2(\Omega)$.

Problem (28) matches the statement of (P) by setting

$$f(x,t,s) = \rho(x,t) + \frac{|s|^{\alpha+1}}{1+|s|}$$

whenever $(x,t,s) \in \Omega \times (0,T) \times \mathbb{R}$. Using the assumption $\alpha \in [0, p-1)$, we obtain through Young's inequality that for each $\varepsilon > 0$ there is a constant $c(\varepsilon) > 0$ such that

$$|f(x,t,s)| \le \rho(x,t) + |s|^{\alpha} \le |\rho(x,t)| + \varepsilon |s|^{p-1} + c(\varepsilon).$$

Choose $\varepsilon \in (0, \lambda_{1,p})$. It results that hypothesis (H) is verified with $\sigma(x,t) = |\rho(x,t)| + c(\varepsilon)$ and $a = \varepsilon$, which permits to apply Theorem 1 to problem (28).

References

[1] L. Boccardo and F. Murat, Almost everywhere convergence of the gradients of solutions to elliptic and parabolic equations, *Nonlinear Anal.* 19 (1992), 581-597.

[2] S. Carl and V.K. Le, On systems of parabolic variational inequalities with multivalued terms, *Monatsh. Math.* 194 (2021), 227-260.

[3] S. Carl, V.K. Le and D. Motreanu, *Nonsmooth Variational Problems and Their Inequalities. Comparison Principles and Applications*, Springer Monographs in Mathematics, Springer, New York, 2007.

[4] S. Carl and D. Motreanu, Extremal solutions for quasilinear parabolic systems in trapping regions, *Pure Appl. Funct. Anal.* 3 (2018), 57-74.

[5] S. Carl and P. Winkert, Parabolic double phase obstacle problems, *Nonlinear Anal. Real World Appl.* 80 (2024), 104169.

[6] J. Cen, S.A. Marano and S. Zeng, Differential inclusion systems with double phase competing operators, convection, and mixed boundary conditions, *Appl. Math. Lett.* 167 (2025), 109556.

[7] D.G. de Figueiredo, *Lectures on the Ekeland Variational Principle with Applications and Detours*, Springer-Verlag, Berlin, 1989.

[8] J. Diblík, M. Galewski, I. Kossowski and D. Motreanu, On competing (p,q)-Laplacian Dirichlet problem with unbounded weight, *Differ. Integr. Equations* 38 (2025), 23-42.

[9] L.C. Evans, *Partial Differential Equations*, Amer. Math. Soc., Providence, RI, 2010.

[10] M. Galewski and D. Motreanu, *Competing Operators and Their Applications to Boundary Value Problems*, Springer Briefs in Mathematics. Springer, Cham, 2026.

[11] L. Gambera, S.A. Marano and D. Motreanu, Quasilinear Dirichlet systems with competing operators and convection, *J. Math. Anal. Appl.* 530 (2024), 127718.

[12] Z. Liu, R. Livrea, D. Motreanu and S. Zeng, Variational differential inclusions without ellipticity condition, *Electron. J. Qual. Theory Differ. Equations* 43 (2020), 1-17.

[13] D. Motreanu, Quasilinear Dirichlet problems with competing operators and convection, *Open Math.* 18 (2020), 1510-1517.

[14] D. Motreanu and Z. Peng, Doubly coupled systems of parabolic hemivariational inequalities: existence and extremal solutions, *Nonlinear Anal.* 181 (2019), 101-118.

[15] A. Razani and G.M. Figueiredo, Degenerated and competing anisotropic (p, q)-Laplacians with weights, *Appl. Anal.* 102 (2023), 4471-4488.

[16] R.E. Showalter, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, Amer. Math. Soc., Providence, RI, 1997.

[17] F. Vetro and R. Efendiev, Systems of differential inclusions with competing operators and variable exponents, *Opuscula Math.* 45 (2025), 665-684.

[18] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/A – Linear Monotone Operators*, Springer, New York, 1990.

[19] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/B – Nonlinear Monotone Operators*, Springer-Verlag, New York, 1990.

[20] Z. Zhong, F. Vetro and B. Chen, A class of logarithmic double phase competing problems with convection term and mixed boundary conditions, *J. Fixed Point Theory Appl.* 27 (2025), 83.

Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 18, No. 4/2026

CONSISTENCY OF PORTFOLIO ALLOCATION STRATEGIES ACROSS INDEPENDENT ASSET UNIVERSES: EVIDENCE FROM DJIA BACKTESTING*

Nazish Shahid† Jiawei Guo‡ Haoyang Liu§

DOI 10.56082/annalsarscimath.2026.4.305

Abstract

To determine whether portfolio strategy outcomes are dataset-dependent or reflect broader trends, this study examines two distinct asset sets. We evaluate five allocation models—Equal Weighting, Maximum Sharpe Ratio, Inverse Variance, Markowitz Mean-Variance, and Robust Optimization—using two 15-stock groups derived from the Dow Jones Industrial Average. Strategy effectiveness is measured through a standardized backtest incorporating a 40-day warm-up window and regular rebalancing schedules. Evaluation criteria include historical returns, risk metrics, turnover frequency, drawdowns, and execution fees. Rather than relying on a single backtest, this work specifically tests whether strategic performance remains uniform across

*Accepted for publication on October 08, 2026

†nazishshahid@fccollege.edu.pk, Department of Mathematics, Forman Christian College (A Chartered University), Lahore, Pakistan; School of Statistics and Data Science, Capital University of Economics and Business, Beijing, China; nazishshahid@arizona.edu, Department of Mathematics, The University of Arizona, Tucson, United States

‡32023230041@cueb.edu.cn, School of Statistics and Data Science, Capital University of Economics and Business, Beijing, China

§32023230067@cueb.edu.cn, School of Statistics and Data Science, Capital University of Economics and Business, Beijing, China

different asset pools. The findings demonstrate broad structural consistency across both tested groups, though localized shifts occur in secondary strategy return rankings. Although the Equal Weighted model produces the top cumulative return, the Inverse Variance approach reliably delivers superior downside protection by minimizing volatility, curtailing drawdowns, and maintaining steady position weights. Conversely, optimization-driven techniques incur greater rebalancing activity and transaction expenses without yielding better out-of-sample performance. These results indicate that risk-focused, resilient allocation frameworks offer more dependable portfolio behavior than models dependent on historical return forecasts.

Keywords: portfolio allocation, backtesting, portfolio optimization, equal weighting, inverse variance strategy, maximum Sharpe ratio, robust optimization, risk management, diversification, Dow Jones Industrial Average (DJIA).

MSC: 91G10, 90C90, 62P05.

1 Introduction

Portfolio allocation is one of the most widely studied problems in finance because investment performance depends not only on the selection of assets but also on how capital is distributed among them. The foundations of modern portfolio theory were established by Markowitz [10], who formalized the trade-off between expected return and risk through mean-variance optimization. Subsequent developments by Sharpe [20,21] extended the analysis of risk-adjusted performance and provided a theoretical framework linking expected returns to systematic risk. These contributions continue to serve as the foundation for modern portfolio construction methods.

Over time, portfolio allocation research has expanded far beyond the original mean-variance framework. Important developments include improved covariance estimation techniques, robust optimization models, diversification-based allocation methods, and Bayesian approaches that incorporate investor views into the asset allocation process [1,5,7,8]. More recently, comprehensive reviews of portfolio optimization have documented the evolution of the field from classical financial models toward more sophisticated approaches that consider multiple objectives, uncertainty, and practical implementation constraints [16,17].

Despite their theoretical appeal, optimization-based portfolio models remain highly sensitive to estimation error. The quality of portfolio allocations depends heavily on expected return and covariance estimates, which

are inevitably subject to uncertainty. Michaud [14] argued that optimized portfolios often exhibit instability due to estimation error, while DeMiguel et al. [4] demonstrated that simple diversification rules can frequently outperform complex optimized portfolios in out-of-sample settings. These findings have raised important questions regarding whether the additional complexity of optimization models consistently translates into superior realized investment performance.

As a consequence, simple diversification methods have received increasing attention in both academic and practical portfolio management. Equal-weighted portfolios avoid estimation error entirely and provide broad diversification across securities. Previous research has shown that equal weighting often performs competitively relative to more sophisticated optimization techniques [4, 15]. Similarly, diversification and risk-allocation studies have shown that portfolio performance can often be improved by focusing on risk distribution rather than return forecasting [3, 9, 13]. These approaches have become increasingly attractive because of their transparency, simplicity, and implementation stability.

Among risk-based approaches, inverse variance and related risk-parity strategies allocate capital according to asset risk characteristics rather than expected returns. Such methods seek to reduce exposure to highly volatile assets while maintaining diversification benefits. Research on risk-based portfolio construction has shown that these approaches can improve downside-risk control and achieve attractive risk-adjusted performance, particularly when return forecasts are unreliable [3, 9]. In many practical settings, portfolios constructed using simple risk measures have exhibited performance that is comparable to, or better than, portfolios derived from more complex optimization procedures.

To address the limitations of classical optimization, robust portfolio optimization methods have been developed to explicitly account for uncertainty in model parameters. Rather than relying on a single point estimate of expected returns, robust approaches seek allocations that remain effective under a range of plausible market conditions [5]. Recent studies continue to demonstrate that robust optimization can improve allocation stability, mitigate estimation risk, and produce stronger out-of-sample performance. Examples include distributionally robust optimization frameworks, robust return prediction models, sparse robust portfolio selection approaches, and hybrid optimization techniques that combine robustness with decision-making algorithms [6, 23–25].

The rapid growth of computational finance has further expanded portfolio optimization research. Modern studies increasingly employ machine

learning, artificial intelligence, and reinforcement learning techniques to support dynamic asset allocation decisions [2, 22]. While these methods offer new opportunities for portfolio construction, recent literature reviews indicate that issues of robustness, diversification, transparency, and implementation costs remain central considerations in practical investment management [16, 17, 22]. Interestingly, many empirical studies continue to report that simple allocation strategies remain highly competitive despite advances in optimization technology [15].

Recent work has also highlighted the importance of understanding asset risk characteristics before portfolio construction. In [18], a volatility-based analysis of KSE-30 equities was conducted using a one-factor credit copula framework and portfolio optimization techniques. The study showed that highly volatile assets tended to be excluded from optimized portfolios, whereas assets with moderate volatility contributed more significantly to final allocations. In [19], portfolio optimization was investigated using the Black-Litterman framework for the Pakistan Stock Exchange KSE-30 index. By combining market equilibrium returns with analysts' views, the study demonstrated how forward-looking information can substantially alter portfolio composition and asset selection decisions.

While [18] and [19] examined optimized portfolio composition–demons trating that highly volatile assets are systematically excluded and showing how forward-looking methods can recalibrate diversified allocations—both studies relied exclusively on expected-return-based optimization. Consequently, they lacked a comprehensive performance evaluation and did not offer a direct comparison between expected-return optimization and explicit risk-based weighting schemes.

Although substantial research exists on portfolio optimization and portfolio allocation, the majority of comparative studies evaluate competing strategies within a single investment universe. Whether considering classical mean-variance optimization, equal weighting, risk-parity methods, robust optimization, or modern artificial intelligence approaches, conclusions are often drawn from a single set of assets and a single portfolio experiment. Consequently, it becomes difficult to determine whether observed performance differences reflect genuine characteristics of the allocation method or whether they are largely driven by the particular securities selected for analysis. A strategy that performs well in one portfolio universe may not necessarily exhibit similar performance when applied to a different set of assets.

The present study addresses this limitation by comparing five portfolio allocation strategies, namely Equal Weighted, Maximum Sharpe Ratio,

Inverse Variance, Markowitz Mean-Variance Optimization, and Robust Optimization, across two independent asset universes constructed from Dow Jones Industrial Average (DJIA) constituents. Using a unified backtesting framework, portfolio performance is evaluated through realized return, volatility, drawdown, turnover, and transaction-cost measures. Unlike many previous studies that rely on a single investment universe, the current research focuses on the consistency of strategy behaviour across two independent asset universes. The study therefore provides evidence on the robustness of strategy rankings and offers practical insight into which portfolio allocation methods provide the most reliable balance between return generation, risk control, diversification, and allocation stability. Throughout this paper, we use the term *consistency* to refer specifically to the preservation of the relative ranking of strategies on a given performance criterion—return, volatility, drawdown, or turnover—across the two independent asset universes, rather than to the numerical stability of the criterion values themselves.

2 Materials and Methods

2.1 Dataset and asset universes

The performance of investment strategies is evaluated and compared using historical data through backtesting. In this study, one year of asset price data from the Dow Jones Industrial Average (DJIA) market is considered. Specifically, we analyze the daily closing prices of the following DJIA constituent assets to examine the performance of different investment strategies:

Apple (AAPL), Amgen (AMGN), Amazon (AMZN), American Express (AXP), Boeing (BA), Caterpillar (CAT), Salesforce (CRM), Cisco (CSCO), Chevron (CVX), Walt Disney (DIS), Goldman Sachs (GS), Home Depot (HD), Honeywell International (HON), Johnson & Johnson (JNJ), International Business Machines (IBM), JPMorgan (JPM), Coca-Cola (KO), McDonald's (MCD), 3M (MMM), Merck (MRK), Microsoft (MSFT), NIKE (NKE), NVIDIA (NVDA), Procter & Gamble (PG), Travelers (TRV), UnitedHealth (UNH), Visa (V), Sherwin-Williams (SHW), Verizon (VZ), and Walmart (WMT).

The dataset comprises the DJIA index and 252 trading days of daily closing prices for these constituent assets.

The comparison of investment strategies is conducted separately on two subsets of assets: the first 15 and the last 15 asset price series. This split is based purely on ticker order among the 30 DJIA constituents and is not

stratified by sector or market capitalization; consequently, the two subsets may differ somewhat in sector composition, which can influence covariance structure and the behavior of variance-based strategies. The analysis based on the assets AAPL, AMGN, AMZN, AXP, BA, CAT, CRM, CSCO, CVX, DIS, GS, HD, HON, IBM, and JNJ is referred to as *Analysis A*. The analysis based on the assets JPM, KO, MCD, MMM, MRK, MSFT, NKE, NVDA, PG, SHW, TRV, UNH, V, VZ, and WMT is referred to as *Analysis B*.

2.2 Portfolio allocation strategies

In this study, five distinct investment strategies are examined. Each strategy determines portfolio weights according to a specific allocation rule, enabling a comparative analysis of their performance under a common backtesting framework. Within this framework, each strategy is evaluated at predetermined points in time, where it is allowed to update its portfolio composition using information derived from recent market data.

This study evaluates five portfolio allocation strategies that translate historical information on asset prices into portfolio weights under a unified backtesting framework. Let N denote the number of assets, and let $\omega = (\omega_1, \omega_2, \ldots, \omega_N)^\top$ represent the vector of portfolio weights, where ω_i denotes the fraction of total capital allocated to asset i. All strategies satisfy the full-investment constraint $\sum_{i=1}^{N} \omega_i = 1$ and impose non-negativity constraints to prohibit short selling. These full-investment and non-negativity (long-only) constraints are enforced uniformly across all five strategies considered in this study—including the Maximum Sharpe Ratio and Markowitz Mean-Variance strategies—even though these two formulations are sometimes implemented with short selling permitted in the broader literature; no short positions were allowed in any of the backtests reported here.

2.2.1 Equal-weighted strategy

The equal-weighted strategy allocates capital uniformly across all assets in the investment universe. Each asset receives the same weight, independent of its expected return or risk characteristics. Formally, the portfolio weights are defined as [11, 12]

$$\omega_i = \frac{1}{N}, \quad i = 1, \ldots, N. \tag{1}$$

This strategy serves as a simple benchmark, as it avoids estimation error and does not rely on model-based assumptions regarding asset behaviour.

2.2.2 Maximum Sharpe ratio strategy

The maximum Sharpe ratio strategy constructs a portfolio by explicitly optimizing risk-adjusted performance. The Sharpe ratio measures the expected excess portfolio return per unit of total risk, where risk is quantified by the standard deviation of portfolio returns. Let $\mathbf{r}$ denote the vector of expected asset returns and Q denote the covariance matrix of asset returns estimated from historical data. The portfolio weights are obtained by solving the optimization problem [12]

$$\boldsymbol{\omega}^{SR} = \arg\max_{\boldsymbol{\omega}} \frac{\mathbf{r}^\top \boldsymbol{\omega}}{\sqrt{\boldsymbol{\omega}^\top Q \boldsymbol{\omega}}} \tag{2}$$

subject to the constraints $\boldsymbol{\omega} \geq 0$ and $\sum_{i=1}^{N} \omega_i = 1$, and upper bounds on individual weights to prevent excessive concentration in a single asset.

This formulation selects the portfolio that lies on the efficient frontier with the highest slope when measured relative to the risk-free asset. In practice, the strategy emphasizes assets with high expected returns relative to their contribution to overall portfolio volatility. While theoretically appealing, the maximum Sharpe ratio solution is known to be sensitive to estimation error, especially in expected returns, which may lead to unstable or highly concentrated allocations when implemented using finite samples.

2.2.3 Inverse variance strategy

The inverse variance strategy determines asset weights using individual risk measures rather than expected returns. Let σ_{ii} denote the variance of asset i, corresponding to the diagonal elements of the covariance matrix Q. Portfolio weights are assigned inversely proportional to these variances and normalized to satisfy the budget constraint [12]

$$\omega_i = \frac{\sigma_{ii}^{-1}}{\sum_{j=1}^{N} \sigma_{jj}^{-1}}, \quad i = 1, \ldots, N. \tag{3}$$

By emphasizing relatively low-variance assets and ignoring cross-asset covariances, this approach produces portfolios that favour stability and are less sensitive to estimation noise in expected returns.

2.2.4 Markowitz mean–variance optimization strategy

The Markowitz mean–variance strategy formalizes portfolio allocation as a trade-off between expected return and risk. Instead of maximizing a ratio,

this approach maximizes a quadratic objective function in which expected portfolio return is reduced by a penalty proportional to portfolio variance. Using the same notation for expected returns and covariance, the optimization problem is given by [12]

$$\boldsymbol{\omega}^{MV} = \arg\max_{\boldsymbol{\omega}} \left(\mathbf{r}^\top \boldsymbol{\omega} - \lambda \, \boldsymbol{\omega}^\top Q \boldsymbol{\omega} \right), \tag{4}$$

subject to non-negativity and full-investment constraints.

The parameter $\lambda > 0$ represents the investor's degree of risk aversion. Larger values of λ place greater emphasis on variance reduction, leading to more conservative portfolios, while smaller values allow higher exposure to risky assets in pursuit of increased expected return. Unlike the maximum Sharpe ratio strategy, which produces a single tangency portfolio, the Markowitz framework generates a family of optimal portfolios along the efficient frontier, each corresponding to a different level of risk tolerance.

Although mean–variance optimization provides a foundational framework for portfolio construction, its effectiveness depends critically on the quality of the estimated input parameters. In particular, inaccuracies in expected return estimates can significantly affect the resulting optimal weights, motivating the consideration of alternative methods that incorporate robustness to estimation uncertainty.

2.2.5 Robust optimization with uncertain expected returns

The robust optimization strategy extends the mean–variance framework by explicitly accounting for uncertainty in expected returns. Rather than assuming that the expected return vector is known exactly, it is assumed to lie within a predefined uncertainty set centered around a nominal estimate $\mathbf{r}_0$. In this formulation, portfolio weights are selected to maximize performance under the least favourable realization of expected returns within the uncertainty set [12]

$$\max_{\boldsymbol{\omega}} \min_{\mathbf{r} \in \mathcal{S}(\mathbf{r}_0)} \left(\mathbf{r}^\top \boldsymbol{\omega} - \lambda \boldsymbol{\omega}^\top Q \boldsymbol{\omega} \right). \tag{5}$$

In this study, the uncertainty set $\mathcal{S}(\mathbf{r}_0)$ is specified as an ellipsoid characterized by an uncertainty-aversion parameter κ and a matrix capturing estimation error in expected returns [12]

$$\mathcal{S}(\mathbf{r}_0) = \left\{ \mathbf{r} : (\mathbf{r} - \mathbf{r}_0)^\top \Sigma_r^{-1} (\mathbf{r} - \mathbf{r}_0) \leq \kappa^2 \right\}. \tag{6}$$

This formulation leads to a robust counterpart of the Markowitz optimization problem, in which an additional penalty term is introduced to control sensitivity to return uncertainty. The resulting allocations aim to improve stability and performance robustness when return estimates are imprecise.

2.3 Analysis A

For this part of the analysis, the closing prices of AAPL, AMGN, AMZN, AXP, BA, CAT, CRM, CSCO, CVX, DIS, GS, HD, HON, IBM, and JNJ are used to compute the portfolio weights corresponding to different investment strategies.

2.3.1 Backtesting framework and strategy implementation

The portfolio strategies considered in this study are dynamic, meaning that portfolio weights are periodically revised as new market information becomes available. At each rebalancing date, the existing portfolio composition is compared with the allocation suggested by the strategy's decision rule, and the portfolio is adjusted accordingly. This process allows the strategies to respond to changing market conditions while maintaining practical investment constraints such as full capital allocation and the absence of short selling. By distinguishing changes caused by market movements from those resulting from deliberate reallocation decisions, the framework provides a realistic setting for evaluating portfolio performance over time.

Before the evaluation period begins, an initial calibration stage is required to place each strategy in a meaningful starting state. Without this step, portfolios would start fully in cash and remain uninvested until the first scheduled rebalancing date, which could distort early performance results. To address this issue, the first 40 trading days of historical price data are used as a warm-up period. During this stage, each strategy estimates the quantities needed for portfolio construction, including return and risk measures. The resulting allocations approximate the positions the strategies would have held had they already been operating before the start of the evaluation window.

To illustrate the effect of this initialization process and to highlight differences in initial capital allocation (Assets A) across strategies, Fig. 1 reports the portfolio weights obtained at the end of the 40-day warm-up period. Although all strategies use the same historical price data, they produce different starting allocations because of their distinct allocation rules. The equal-weighted strategy distributes capital evenly across all assets, whereas

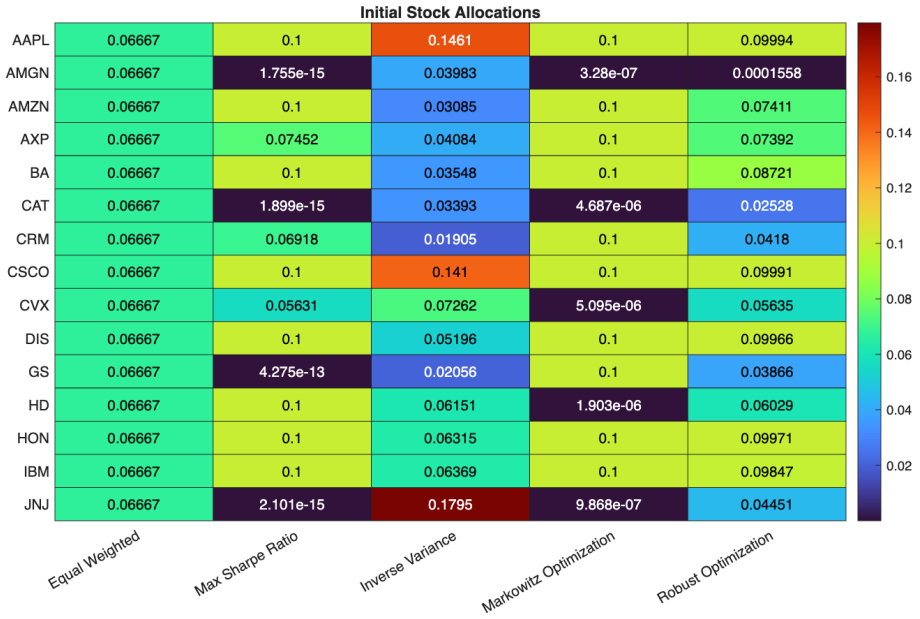


Figure 1: Initial portfolio weights (Assets A) after the 40-day warm-up period

the remaining strategies allocate capital according to their individual objectives. The inverse variance strategy places greater weight on assets with relatively lower estimated risk, such as JNJ, CSCO, and AAPL, while the optimization-based strategies produce more concentrated allocations based on their optimization criteria. These weights represent the initial portfolio positions used at the start of the backtest and should be viewed as calibration-based starting allocations rather than indicators of future performance.

After the warm-up period, the backtest is performed using the remaining data. Portfolio allocations are updated at discrete rebalancing dates rather than continuously, reflecting common portfolio management practice. In this study, portfolios are rebalanced every 21 trading days, approximately corresponding to a monthly adjustment cycle. Allocation decisions are based on a rolling look-back window of recent historical data, allowing the strategies to adapt to changing market conditions while avoiding reliance on outdated information.

All strategies are evaluated within a unified backtesting framework that applies the same market data, rebalancing schedule, and performance crite-

ria to every strategy. To ensure meaningful comparisons, the initial portfolio value and risk-free rate are specified at the framework level and remain identical across all simulations. The annualized risk-free rate is adjusted internally to match the daily frequency of the asset price series. By standardizing these settings, any differences in performance can be attributed to the allocation methodology rather than implementation choices.

The simulation begins only after the warm-up period has been completed so that every portfolio decision is based solely on information available at the time it is made. This prevents look-ahead bias and ensures that the reported results reflect realistic investment conditions. Throughout the evaluation period, the framework records portfolio values, asset holdings, realized returns, turnover, drawdowns, and trading costs. These outputs provide the basis for the subsequent comparison of portfolio performance across the competing investment strategies.

2.3.2 Evaluation of backtest performance outcomes (Assets A)

Following backtest execution, Table 1 outlines the empirical performance of each portfolio strategy across risk, return, and transaction metrics. Properly evaluating these outcomes requires accounting for the strictly out-of-sample nature of the decision process, as portfolio weights were computed using only historical data available at each rebalance date. For example, the Maximum Sharpe Ratio approach selects weights that maximize historical risk-adjusted returns over the preceding look-back window. Once fixed, these weights are evaluated against subsequent, unobserved price trajectories. Consequently, allocations optimized on historical windows can underperform during subsequent periods. This in-sample optimization drag is reflected in the lower realized returns and Sharpe ratios of the Maximum Sharpe Ratio strategy, reinforcing that optimized historical performance does not ensure superior out-of-sample results.

Table 1 compares the performance of the five portfolio strategies across return, risk, and trading characteristics. In terms of return, the Equal-Weighted strategy achieves the highest total (0.15462) and average return (0.00075), followed by Inverse Variance (total return 0.13567) and Robust Optimization (total return 0.12119). In contrast, the Maximum Sharpe Ratio (0.11600) and Markowitz Mean-Variance strategies (0.10239) yield comparatively lower realized performance. This demonstrates that strategies explicitly optimized on historical data do not necessarily deliver superior out-of-sample outcomes.

From a risk perspective, the Inverse Variance strategy exhibits the lowest

volatility (0.01162) and the smallest maximum drawdown (0.18125), proving its ability to stabilize fluctuations by allocating greater weight to lower-variance assets. Robust Optimization achieves a similarly low volatility (0.01194) and drawdown (0.19528). Conversely, the Markowitz strategy records the highest volatility (0.01439) and the largest maximum drawdown (0.23935), reflecting its sensitivity to parameter estimation error.

Trading activity and turnover further distinguish the performance across strategies. The Maximum Sharpe Ratio strategy records the highest average turnover (0.00776) and buy/sell costs (0.40795), driven by frequent position reallocations from unstable parameter estimates. The Markowitz strategy exhibits a similarly high average turnover (0.00726). In contrast, the Equal-Weighted (average turnover 0.00115) and Inverse Variance strategies (average turnover 0.00310) maintain substantially lower rebalancing activity, resulting in superior allocation stability over time.

In terms of portfolio structure, Equal Weighting guarantees uniform exposure across all constituent assets, whereas Inverse Variance achieves diversification through a risk-balanced allocation. Robust Optimization avoids extreme weight concentration while generating moderate returns. Overall, the Inverse Variance strategy provides the most balanced combination of risk control, reduced drawdown, and allocation stability, whereas the Equal-Weighted strategy stands out for its superior cumulative return and implementation simplicity.

Results	**Equal Weighted**	**Max Sharpe Ratio**	**Inverse Variance**	**Markowitz Opt.**	**Robust Opt.**
Total Return	0.15462	0.116	0.13567	0.10239	0.12119
Sharpe Ratio	0.057194	0.044587	0.053834	0.036221	0.04764
Volatility	0.012483	0.012399	0.011624	0.014396	0.011948
Average Turnover	0.0011468	0.0077634	0.0030982	0.0072563	0.0027509
Max Turnover	0.032336	0.30743	0.16096	0.30629	0.11499
Average Return	0.00075239	0.00059168	0.00066444	0.00056033	0.00060797
Max Drawdown	0.19851	0.2026	0.18125	0.23935	0.19528
Average Buy Cost	0.060207	0.40795	0.16084	0.36997	0.14165
Average Sell Cost	0.060207	0.40795	0.24126	0.36997	0.14165

Table 1: Strategy performance (Assets A)

2.3.3 Evolution of portfolio value (Assets A)

To illustrate the dynamic behavior of each allocation strategy over time, Fig. 2 displays the equity curves tracking portfolio value evolution across

the backtesting window. These curves capture the cumulative effect of periodic rebalancing and subsequent market movements, offering clear visual insight into wealth accumulation patterns, growth trajectories, and drawdowns. Specifically, the trajectories in Fig. 2 highlight that the Equal-Weighted portfolio achieves the steepest upward growth trajectory, while the Inverse Variance curve demonstrates the smoothest, most stable trajectory with minimal visible pullbacks. Conversely, the Markowitz and Maximum Sharpe Ratio curves exhibit noticeable structural dips and heightened fluctuations, visually confirming their higher drawdown vulnerability and lower out-of-sample stability compared to non-optimized approaches.

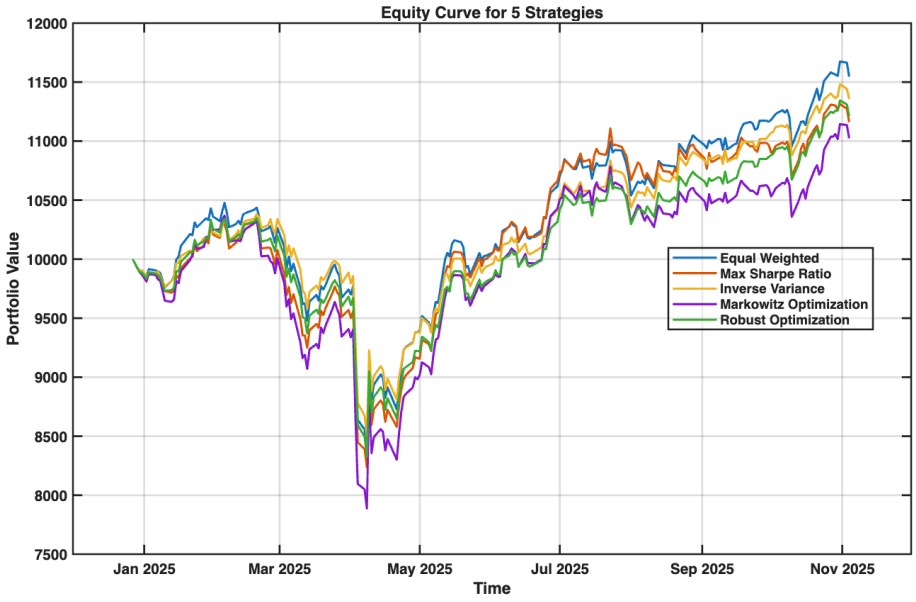


Figure 2: Evolution of portfolio value under competing investment strategies (Assets A)

The average turnover across strategies is shown in Fig. 3, providing insight into the frequency of portfolio rebalancing and associated trading activity. As illustrated, the Maximum Sharpe Ratio and Markowitz strategies exhibit the highest turnover levels, reflecting frequent weight reallocations driven by sensitivity to historical parameter estimates. Conversely, the Equal Weighted and Inverse Variance strategies maintain significantly lower turnover, demonstrating superior allocation stability over time. From a practical standpoint, lower rebalancing activity makes these passive and risk-balanced approaches more favorable because it minimizes execution drag and

preserves net returns. However, turnover alone is not the sole benchmark for evaluating portfolio performance; it must be assessed alongside net realized returns, maximum drawdowns, and volatility to capture the complete trade-off between implementation costs and risk-adjusted efficiency.

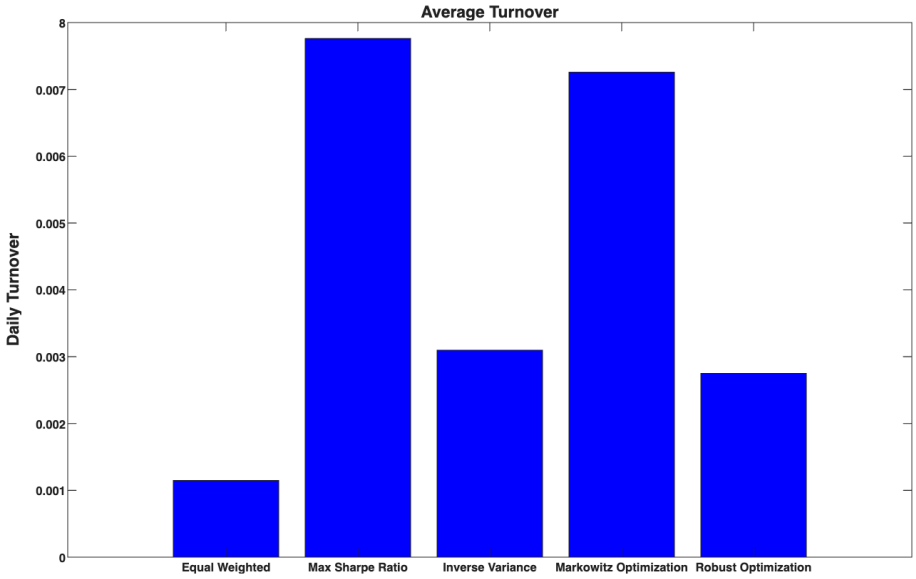


Figure 3: Average portfolio turnover across investment strategies (Assets A)

To examine how portfolio allocations evolve over time, the dynamic asset positions implied by each strategy are visualized using area charts of daily portfolio weights.

Fig. 4 through Fig. 8 present the evolution of asset-level portfolio weights over time for the five allocation strategies considered in Analysis A. Each figure is displayed as a stacked area plot, where individual components represent the proportion of total portfolio capital allocated to each asset at a given point in time, and the aggregate height reflects the full-investment constraint. The temporal variation in these stacked profiles captures both passive changes induced by asset price movements between rebalancing dates and active adjustments arising from periodic portfolio re-optimization. Across all figures, localized contractions in individual layers correspond to periods where the associated asset receives a reduced allocation, highlighting the dynamic redistribution of capital across the investment universe.

Equal-weighted strategy (A)—Asset allocation over time Fig. 4 presents the dynamic allocation path of the equal-weighted strategy, where asset weights evolve solely due to price fluctuations between rebalance dates. Capital remains evenly distributed across all constituent stocks, displaying smooth weight adjustments over time. This continuous stability accounts for the low trading turnover observed in Table 1 and confirms its role as a baseline benchmark for Analysis A.

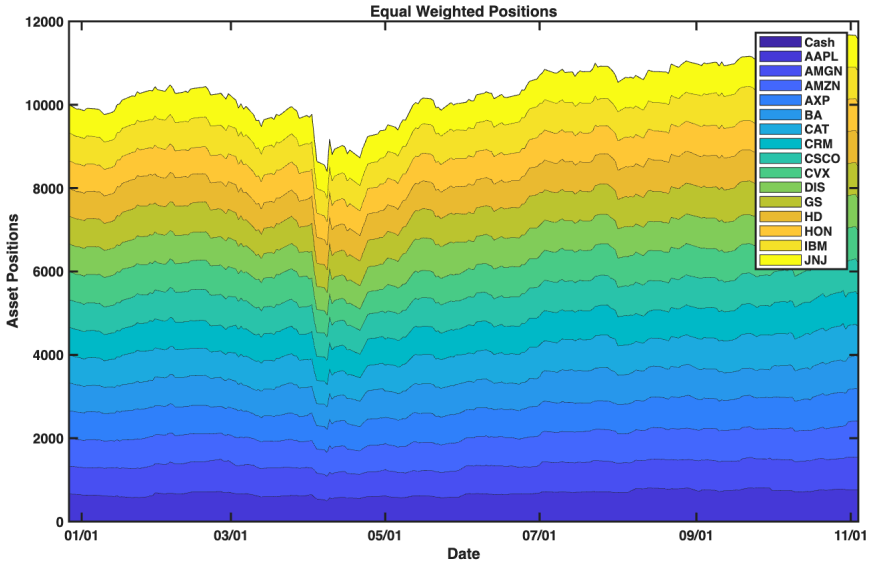


Figure 4: Asset-level portfolio positions over time (Equal-weighted strategy, A)

Maximum Sharpe ratio strategy (A)—Asset allocation over time Fig. 5 highlights the portfolio allocation dynamics under the maximum Sharpe ratio strategy, where weight adjustments are driven by periodic recalculations of risk-adjusted performance. The stacked area plot exhibits pronounced reallocation spikes, as capital is dynamically shifted toward top-performing assets while other position sizes contract or vanish completely. These notable weight swings align with the elevated turnover and higher trading costs reported in Table 1, reflecting the sensitivity of tangency optimization to input parameters.

Inverse variance strategy (A)—Asset allocation over time Fig. 6 illustrates the weight evolution for the inverse variance strategy, demonstrat-

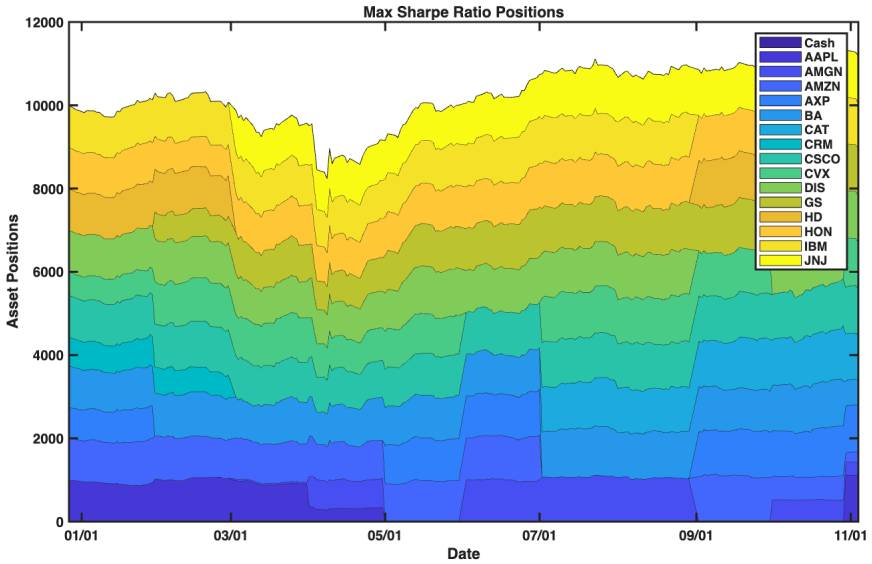


Figure 5: Asset-level portfolio positions over time (Maximum Sharpe ratio strategy, A)

ing the persistent allocation structure derived from risk-based weighting. The portfolio maintains stable position sizes in low-volatility constituents (such as JNJ, CSCO, and AAPL), with layer thickness adjusting gradually across rebalance points. This defensive asset distribution directly contributes to the lowest volatility and smallest maximum drawdown recorded in Table 1.

Markowitz optimization strategy (A)—Asset allocation over time Fig. 7 tracks the portfolio composition over time for the Markowitz mean–variance strategy, capturing the ongoing trade-off between return and risk. The allocation paths exhibit sharper position adjustments across assets as updated parameter estimates modify the optimal weighting decision. These variable allocation blocks help explain the strategy's increased turnover, higher trading expenditure, and larger drawdown in Table 1.

Robust optimization strategy (A)—Asset allocation over time. Fig. 8 depicts the allocation path generated by the robust optimization strategy, showing how asset weights adapt over time while incorporating parameter uncertainty. Compared to the standard Markowitz approach, the robust formulation leads to smoother and more balanced position tran-

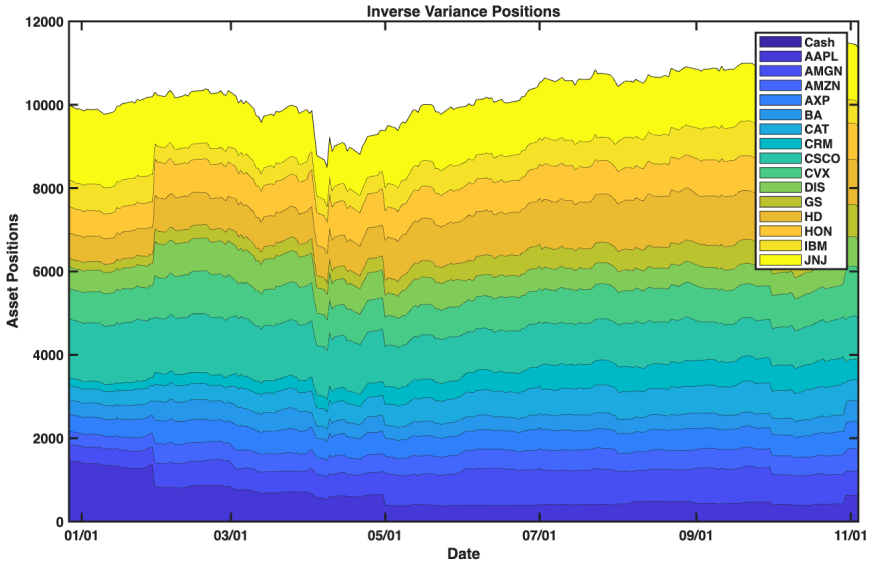


Figure 6: Asset-level portfolio positions over time (Inverse variance strategy, A)

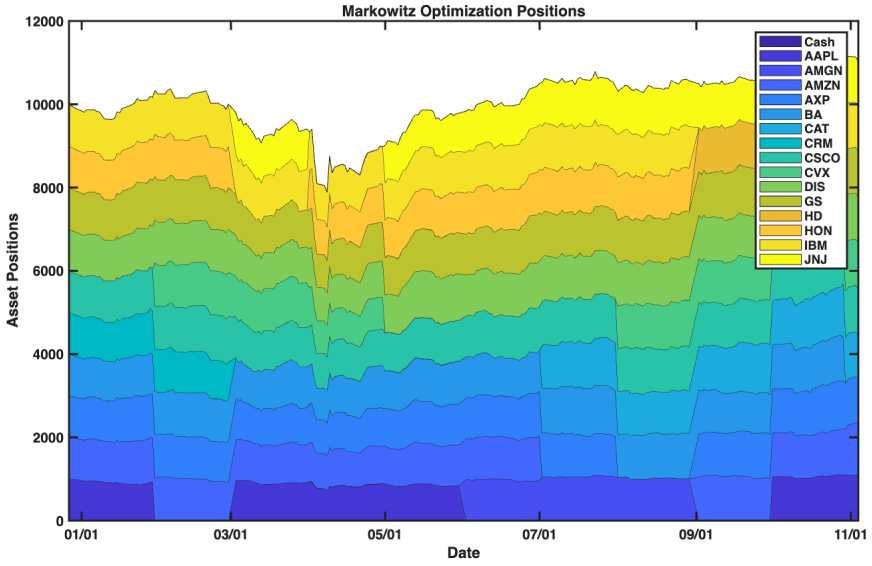


Figure 7: Asset-level portfolio positions over time (Markowitz optimization strategy, A)

sitions. By penalizing return estimation errors, it mitigates extreme capital shifts while preserving portfolio flexibility and maintaining controlled volatility (Table 1).

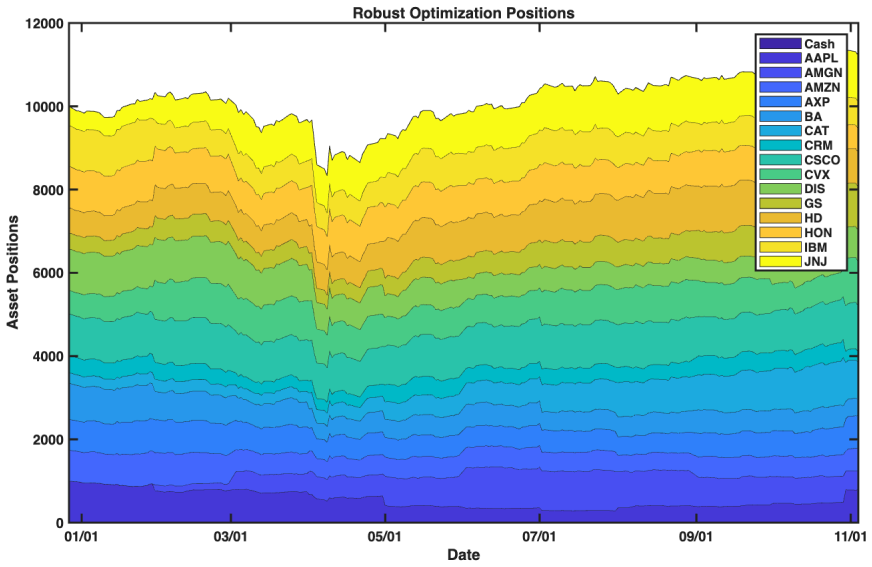


Figure 8: Asset-level portfolio positions over time (Robust optimization strategy, A)

2.4 Analysis B

For this part of the analysis, the closing prices of JPM, KO, MCD, MMM, MRK, MSFT, NKE, NVDA, PG, SHW, TRV, UNH, V, VZ, and WMT are used to compute the weights of the different investment strategies.

To highlight differences in initial capital allocation across strategies, Fig. 9 displays the portfolio weights computed at the end of the warm-up period, where each strategy translates the same historical price information into a starting allocation. The equal-weighted strategy gives uniform exposure to the full asset set and therefore serves as the broad diversification benchmark. Under the inverse variance strategy, assets such as KO, MCD, PG, V, VZ, and WMT receive comparatively larger weights, indicating that their estimated variances over the calibration window were relatively low compared to other assets. The maximum Sharpe ratio and Markowitz strategies are more selective, assigning substantial weight to a narrower set of assets while giving very small positions to others. The robust optimiza-

tion strategy keeps a wider spread of exposures than the more aggressive optimization rules, with visible emphasis on assets that combine moderate risk with stable historical performance. The strategies differ mainly in how they balance concentration, defensive exposure, and sensitivity to estimated returns. These initial weights should therefore be interpreted as calibration-period positions rather than evidence of long-term asset performance.

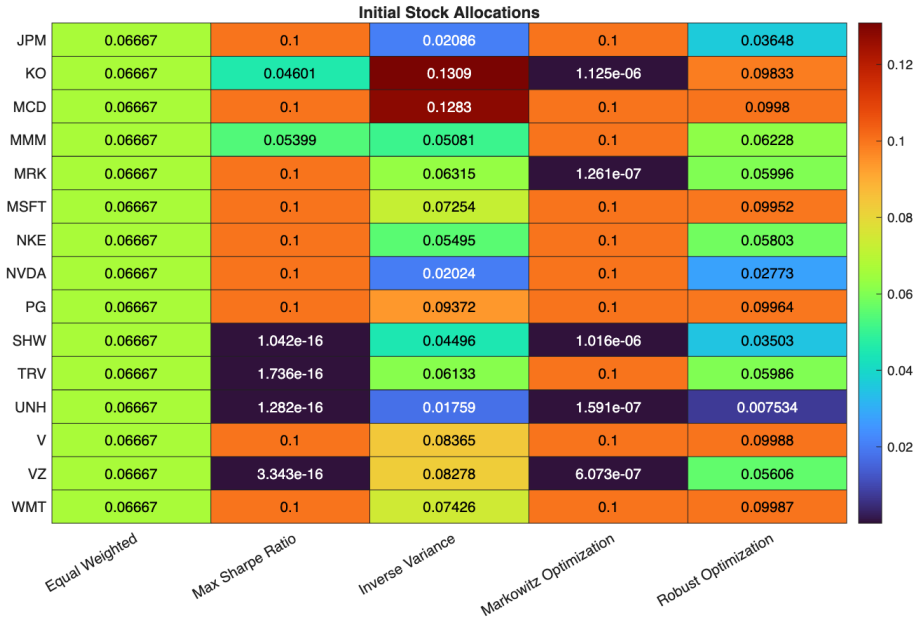


Figure 9: Initial portfolio weights (Assets B) after the 40-day warm-up period

2.4.1 Evaluation of backtest performance outcomes (Assets B)

After running the backtest, Table 2 summarizes the performance of each strategy for Assets B using standard evaluation metrics: realized return, volatility, drawdown, turnover, and trading cost. The market conditions and testing rules were kept the same for all methods.

The interpretation of Table 2 must account for the fact that portfolio decisions are made using only information available before each rebalance date. This point is especially relevant for Analysis B because the optimization-based strategies often form concentrated allocations from a limited estimation window. Their selected portfolios may appear efficient at the time of construction, yet their realized performance depends on price movements

Results	Equal Weighted	Max Sharpe Ratio	Inverse Variance	Markowitz Opt.	Robust Opt.
Total Return	0.07382	0.056951	0.046644	0.048307	0.052193
Sharpe Ratio	0.035303	0.028645	0.023888	0.023782	0.026636
Volatility	0.0096722	0.0091613	0.0089821	0.009571	0.0089964
Average Turnover	0.0011811	0.0056974	0.0027293	0.0066467	0.0022865
Max Turnover	0.032776	0.21149	0.10211	0.21104	0.090746
Average Return	0.00038078	0.00030194	0.00025419	0.00026722	0.0002792
Max Drawdown	0.11513	0.12427	0.10273	0.12852	0.10863
Average Buy Cost	0.061486	0.30048	0.14273	0.34909	0.1195
Average Sell Cost	0.061486	0.30048	0.21409	0.34909	0.1195

Table 2: Strategy performance (Assets B)

that occur after the weights are fixed. The empirical results show that greater optimization complexity can fail to translate into stronger realized outcomes in this asset universe.

In terms of return, the Equal-Weighted strategy achieves the highest total (0.07382) and average return (0.00038), demonstrating that simple, broad capital distribution effectively captures market recovery without relying on active selection or frequent reoptimization. The Maximum Sharpe Ratio (0.05695) and Robust Optimization strategies (0.05219) follow, whereas the Markowitz (0.04831) and Inverse Variance strategies (0.04664) yield weaker cumulative performance. Inverse Variance deliberately forfeits upside potential by concentrating on low-risk assets, while the Markowitz strategy is impacted by parameter instability and unfavorable reallocations.

From a risk perspective, the Inverse Variance strategy provides the strongest downside protection, recording the lowest volatility (0.00898) and the smallest maximum drawdown (0.10273). This confirms that the Assets B universe contains sufficient low-risk constituents for a variance-based rule to mitigate portfolio stress. Robust Optimization exhibits a similarly defensive profile with low volatility (0.00899) and a controlled maximum drawdown (0.10863). The Equal-Weighted strategy incurs slightly higher volatility (0.00967) but maintains a moderate drawdown (0.11513) due to its broad asset distribution, whereas the Markowitz strategy experiences the largest maximum drawdown (0.12852).

Trading activity and costs further separate the strategies. The Markowitz strategy records the highest average turnover (0.00665) and trading costs (0.34909), followed by the Maximum Sharpe Ratio strategy (average turnover 0.00570). These high turnover rates highlight the sensitivity of classical optimization to shifting historical estimates, where additional trading activity

fails to yield a performance advantage. Conversely, the Equal-Weighted (average turnover 0.00118) and Robust Optimization strategies (average turnover 0.00229) maintain high allocation stability and minimal transaction drag.

Overall, the results demonstrate a clear separation between broad capital distribution and risk-balanced diversification. The Equal-Weighted strategy offers the strongest total return and risk-adjusted efficiency through simple diversification, while the Inverse Variance strategy remains the most effective choice for capital preservation and downside control. Robust Optimization serves as a practical middle ground among optimized methods by maintaining allocation stability while capturing higher returns than purely defensive approaches.

2.4.2 Evolution of portfolio value (Assets B)

Fig. 10 traces the portfolio value trajectories for Assets B, illustrating how each strategy accumulates wealth over the backtest period. Consistent with the trends in Analysis A, the Equal-Weighted strategy maintains the highest equity curve throughout the horizon, confirming its superior cumulative return capability. The Inverse Variance curve displays the most consistent, low-volatility progression, maintaining a smooth upward slope even during market pullbacks. In contrast, the Markowitz and Maximum Sharpe Ratio curves lag behind, exhibiting sharp drawdowns and noticeable stagnation that visually illustrate their out-of-sample instability and sensitivity to estimation errors.

The average turnover across strategies is shown in Fig. 11, providing insight into the frequency of portfolio rebalancing and associated trading activity. Mirroring the behavior observed in Analysis A, the results for Assets B confirm that the Maximum Sharpe Ratio and Markowitz strategies generate the highest turnover, whereas the Equal Weighted and Inverse Variance models maintain the lowest rebalancing requirements. This consistency reinforces that parameter-dependent optimizations persistently incur higher trading frequency across different asset pools, while risk-weighted and static allocations offer greater operational stability. Consequently, while low turnover makes Equal Weighted and Inverse Variance allocations structurally more favorable for cost management, rebalancing frequency remains just one dimension of strategy selection and must be weighed against drawdown control and net realized returns.

To examine how portfolio allocations evolve over time, the dynamic asset positions implied by each strategy are visualized using area charts of daily

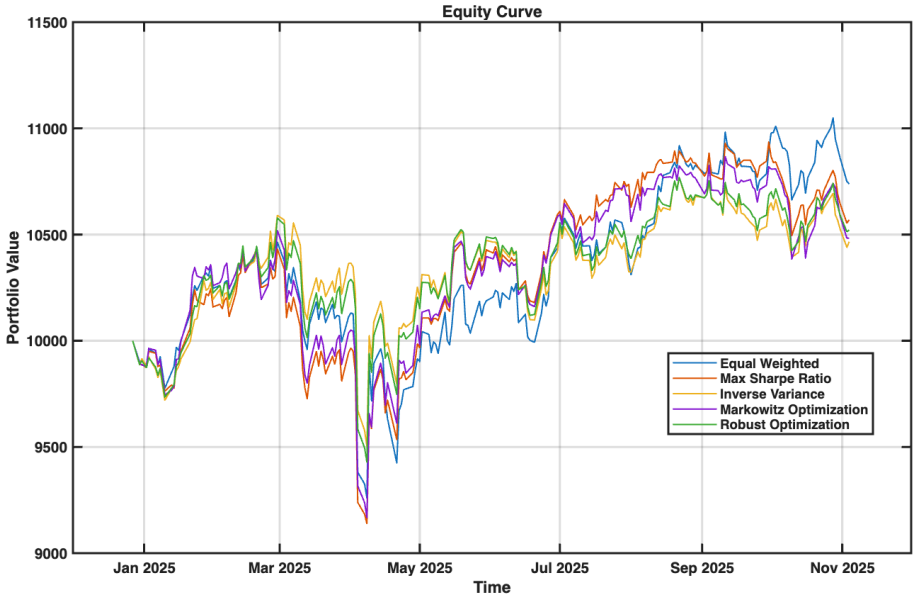


Figure 10: Evolution of portfolio value under competing investment strategies (Assets B)

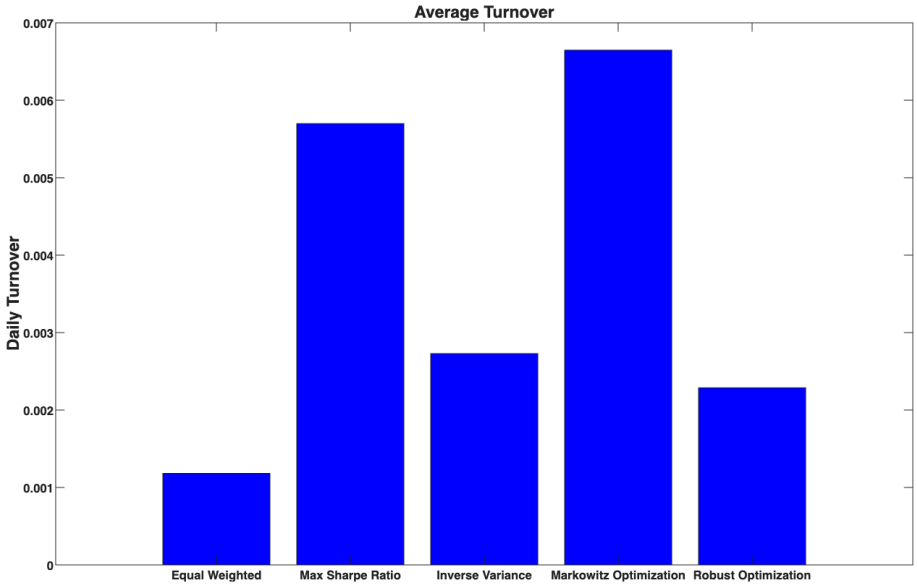


Figure 11: Average portfolio turnover across investment strategies (Assets B)

portfolio weights.

Fig. 12 through Fig. 16 present the evolution of asset-level portfolio weights over time for the allocation strategies considered in Analysis B. The stacked area plots show how each method distributes capital across the same asset universe and how those distributions change as prices move and rebalancing decisions occur.

Equal-weighted strategy (B)—Asset allocation over time Fig. 12 illustrates the time-varying portfolio composition of the equal-weighted strategy. The allocation remains broadly distributed, and changes in the stacked areas mainly reflect asset price movements between rebalancing dates. This pattern explains the strategy's low turnover and shows why it functions as a stable benchmark in Analysis B.

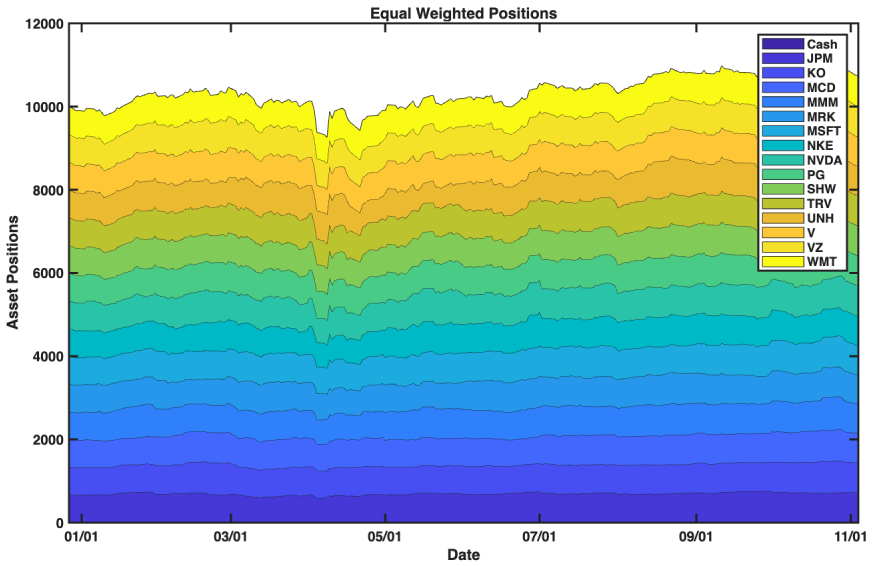


Figure 12: Asset-level portfolio positions over time (Equal-weighted strategy, B)

Maximum Sharpe ratio strategy (B)—Asset allocation over time Fig. 13 depicts the evolution of asset allocations under the maximum Sharpe ratio strategy. The area chart displays more abrupt changes than the equal-weighted and inverse variance strategies, with capital shifting toward selected assets as risk-adjusted return estimates are updated. Several layers

expand or contract sharply at rebalancing points, and some asset exposures become temporarily very small or nearly disappear from the stacked area plot. These visible reallocations are consistent with the higher turnover reported in Table 2 and with the sensitivity of Sharpe-ratio optimization to estimated returns and covariance inputs.

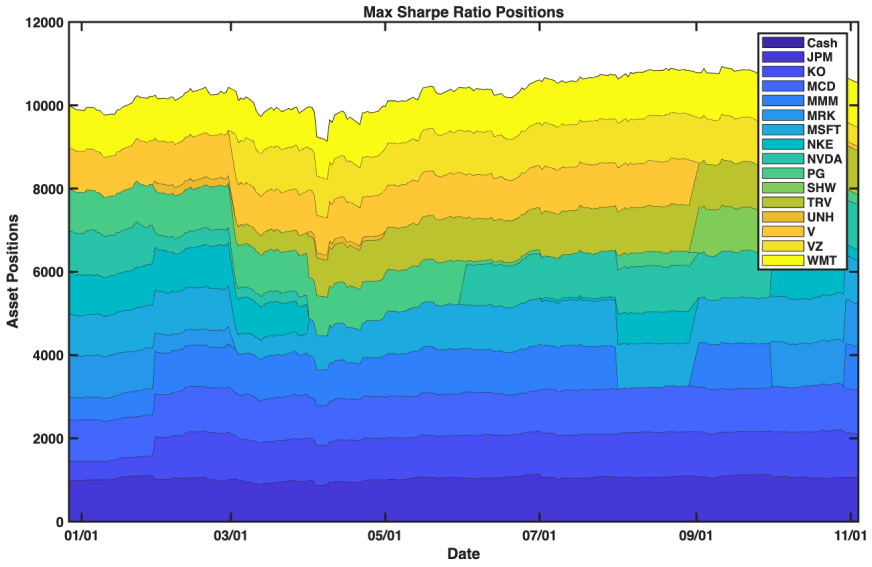


Figure 13: Asset-level portfolio positions over time (Maximum Sharpe ratio strategy, B)

Inverse variance strategy (B)—Asset allocation over time Fig. 14 presents the allocation dynamics of the inverse variance strategy. The portfolio keeps relatively persistent exposure to lower-volatility assets, and its weights adjust more gradually over time. This stable structure supports the low volatility and drawdown results observed in Table 2.

Markowitz optimization strategy (B)—Asset allocation over time Fig. 15 shows the evolution of portfolio weights for the Markowitz strategy. The allocation path contains sharper shifts across assets, reflecting the strategy's sensitivity to estimated expected returns and covariance relationships. This behaviour helps explain its higher turnover, larger trading cost, and weaker drawdown control.

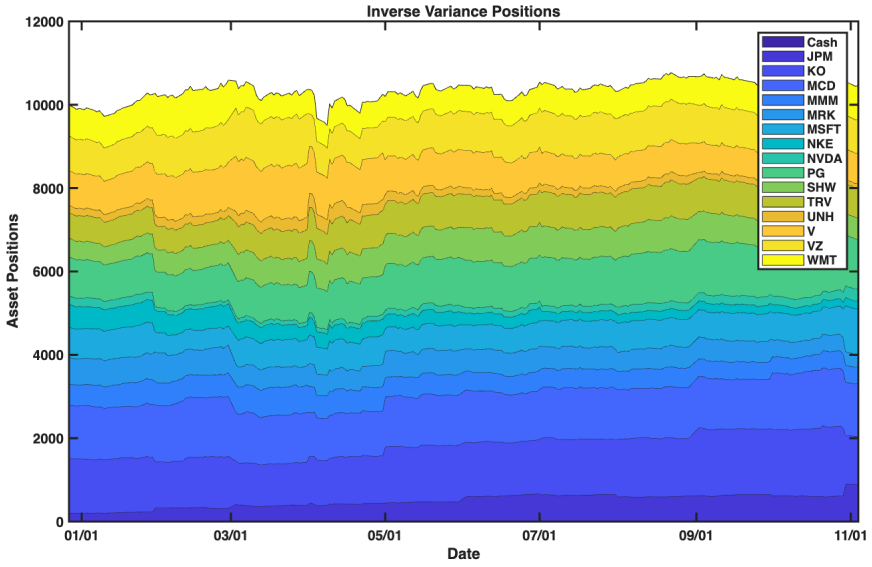


Figure 14: Asset-level portfolio positions over time (Inverse variance strategy, B)

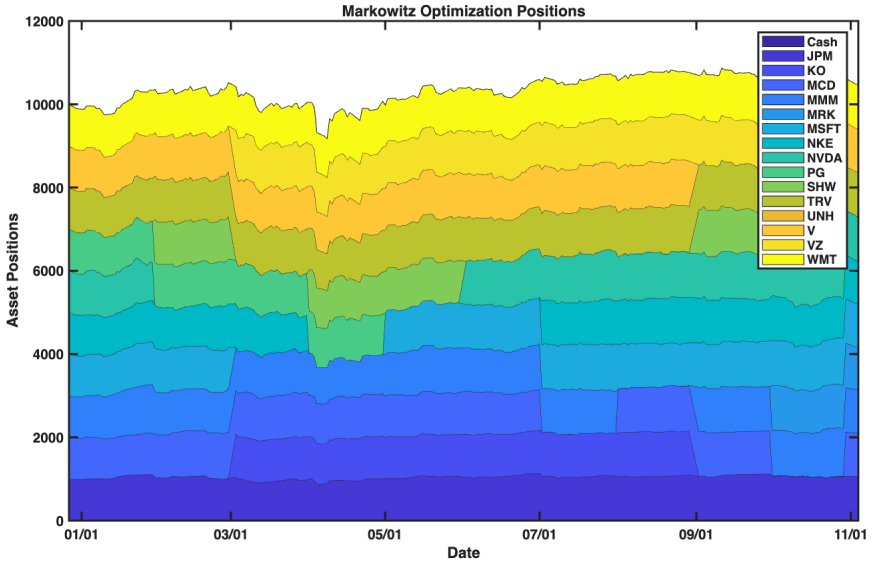


Figure 15: Asset-level portfolio positions over time (Markowitz optimization strategy, B)

Robust optimization strategy (B)—Asset allocation over time Fig. 16 illustrates the asset allocation path generated by the robust optimization strategy. Compared with the standard Markowitz allocation, the robust strategy produces smoother and more balanced position changes. Its allocation path reflects the effect of accounting for uncertainty in expected returns, which reduces extreme reallocations while preserving adaptability.

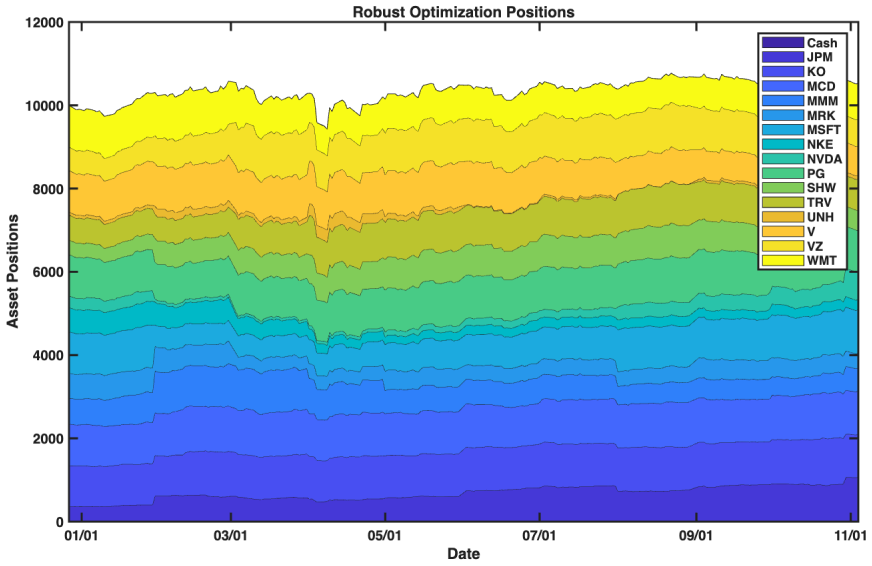


Figure 16: Asset-level portfolio positions over time (Robust optimization strategy, B)

Across these figures, the defining feature of Analysis B is the contrast between stable broad exposure and concentrated optimization. Equal weighting produces the smoothest benchmark allocation, inverse variance delivers a defensive risk-based profile, and robust optimization moderates the instability associated with standard mean–variance optimization. The maximum Sharpe ratio and Markowitz strategies show more visible reallocation blocks, which match their higher turnover and less stable realized performance. For Assets B, allocation stability and controlled risk exposure contribute more reliably to long-term performance than frequent re-optimization driven by unstable historical parameter estimates.

Criterion	**Equal Weighted (EW)**	**Inverse Variance (IV)**	**Robust Optimization (RO)**	**Max Sharpe Ratio (MSR)**	**Markowitz Mean-Variance (MV)**
Return performance	Highest in A and B	Second in A, lowest in B	Moderate	Moderate-low	Lowest in A, second lowest in B
Volatility control	Moderate	Lowest	Near-lowest	Moderate	Highest
Maximum drawdown	Moderate	Lowest	Second lowest	Higher	Highest
Turnover	Lowest	Low-moderate	Low	Highest	Highest
Transaction costs	Lowest	Low	Low	Highest	Highest
Allocation stability	Highest	High	High	Low	Low
Diversification	Maximum	Risk-based	Balanced	Concentrated	Concentrated
Sensitivity to estimation error	None	Low	Controlled	High	Very high
Practical portfolio construction	Excellent	Excellent	Good	Weak	Weak

Table 3: Comparative backtesting summary from Analysis A and Analysis B

3 Discussion

The results in Table 3 demonstrate broad structural consistency across both asset universes A and B. While the overall hierarchy between heuristic and optimization-driven models holds, specific return rankings vary slightly between the two pools (e.g., Inverse Variance yields the second-highest return in Analysis A but ranks lower in Analysis B, whereas Maximum Sharpe Ratio improves its return rank in Analysis B). Nevertheless, key performance characteristics—such as volatility control and turnover behavior—remain stable. While the Equal Weighting method yielded the top cumulative return in both groups, the Inverse Variance framework systematically achieved the lowest volatility, minimal maximum drawdowns, and modest turnover while preserving competitive return levels, making it the most balanced allocation overall. Conversely, despite their theoretical foundation, the Markowitz and Maximum Sharpe Ratio strategies exhibited elevated turnover rates, greater transaction drag, and lower realized returns. Robust Optimization enhanced weighting stability relative to classical mean–variance models, yet its performance remained less dependable than Inverse Variance.

While the Inverse Variance strategy appears to offer the most balanced overall performance, the results also demonstrate that portfolio allocation decisions depend strongly on the investor's objective. The Equal Weighted strategy consistently generated the highest realized returns in both analyses, making it attractive when return maximization is the primary goal. In contrast, the Inverse Variance strategy delivered the lowest volatility and smallest maximum drawdowns, indicating superior downside-risk control. It also maintained relatively low turnover and transaction costs compared with the Maximum Sharpe Ratio and Markowitz strategies, whose frequent reallocations increased implementation costs without producing superior realized outcomes. These findings suggest that investors seeking growth may prefer Equal Weighting, whereas investors emphasizing risk management, capital preservation, and allocation stability would likely favor the Inverse Variance approach. Overall, the latter provides the strongest combination of risk control, turnover management, and drawdown reduction, making it a compelling choice for practical portfolio allocation.

An additional implication of the comparative analysis concerns the identification of assets suitable for subsequent portfolio optimization. The volatility-based investigation [18] of KSE-30 equities showed that highly volatile stocks were generally excluded from optimized portfolios, whereas assets with lower or moderate sensitivity to market movements contributed more

meaningfully to portfolio construction. Similarly, the Black-Litterman analysis [19] retained only a limited set of assets with favorable return and risk characteristics after blending market equilibrium returns with analyst views. The appearance of assets such as HBL, LUCK, and POL—constituents of the Pakistan Stock Exchange KSE-30 index examined in [19]—within the final Black-Litterman allocation suggests that securities exhibiting both acceptable volatility characteristics and favorable expected-return profiles may represent promising candidates for future portfolio construction. Although a formal combined optimization framework would be required to confirm this observation, the current findings provide a preliminary indication that information from both analyses may assist in screening equities for subsequent optimal portfolio development and future Black-Litterman applications.

Furthermore, the comparison highlights an important limitation of historically popular optimization-based investment methods. In portfolio theory, the Maximum Sharpe Ratio strategy is frequently regarded as an attractive allocation rule because it seeks the highest risk-adjusted return on the efficient frontier. However, the backtesting results obtained in both analyses indicate that this theoretical advantage did not translate into consistently superior realized performance. The strategy exhibited substantially higher turnover and transaction costs than simpler allocation rules, while its realized returns and risk-adjusted outcomes remained comparatively modest. Similar behavior was observed for the Markowitz strategy. These findings emphasize that allocations derived from historical optimization may be highly sensitive to estimation error and changing market conditions. Consequently, relying exclusively on historically estimated inputs may not always be sufficient for index investment decisions, particularly when implementation costs and allocation stability are important considerations.

Ultimately, this analysis demonstrates that the Inverse Variance strategy provides the most balanced overall profile, while revealing broadly consistent structural performance patterns across two independent asset universes. Although asset-universe composition introduces slight variations in total return rankings for specific strategies, both evaluations yield identical overarching conclusions regarding return generation, risk control, turnover behavior, and allocation stability. Specifically, simpler diversification and risk-aware approaches consistently match or outperform complex optimization methods on realized portfolio outcomes. This indicates that the structural stability of the allocation mechanism is more vital to performance than aggressive optimization based on historical estimates. Because this pattern holds across multiple criteria and independent datasets, the observed behavior reflects a robust portfolio allocation principle rather than a dataset-specific artifact.

4 Conclusion

This study was conducted to examine whether different portfolio allocation strategies produce consistent investment outcomes when applied to different groups of equities within the same market environment. Using one year of daily closing price data from the Dow Jones Industrial Average (DJIA) constituents, two independent analyses were performed on separate sets of fifteen assets. Five portfolio allocation strategies, namely Equal Weighted, Maximum Sharpe Ratio, Inverse Variance, Markowitz Mean-Variance Optimization, and Robust Optimization, were evaluated under an identical backtesting framework with periodic rebalancing.

The main contribution of this work is the direct comparison of simple diversification approaches, risk-based allocation methods, and optimization-based portfolio strategies across two independent asset universes using the same evaluation methodology. This allows the robustness of strategy performance to be assessed rather than relying on results from a single asset set.

The evidence obtained from both analyses shows that strategy performance is highly consistent across the two asset universes. The Equal Weighted strategy produced the highest realized returns, while the Inverse Variance strategy consistently achieved the lowest volatility, the smallest drawdowns, and relatively low turnover. In contrast, the Maximum Sharpe Ratio and Markowitz strategies exhibited higher turnover, higher transaction costs, and weaker realized performance despite their theoretical optimization advantages. The Robust Optimization strategy provided a compromise between return generation and risk control but did not consistently outperform the Inverse Variance approach.

The principal conclusion supported by the results is that stable and risk-aware allocation methods can be more effective in practice than conventional strategies that depend heavily on estimates of expected returns. Across both analyses, simple diversification and variance-based risk control delivered more consistent portfolio outcomes than more complex optimization procedures. These findings suggest that portfolio allocation decisions may benefit more from robustness, stability, and risk management than from aggressive optimization of historical return estimates.

5 Limitations of the study

The analysis is based on a single year of historical data and two subsets of DJIA constituents. Consequently, the findings may be influenced by the specific market conditions observed during the study period. Although the consistency of results across both asset universes strengthens the conclusions, the performance of the strategies may differ under alternative market regimes or longer investment horizons. In addition, the two asset subsets were formed by ticker order rather than by a sector-stratified procedure; a sector-balanced split of the DJIA constituents would help confirm whether the observed consistency in strategy rankings is robust to differences in sector composition across the two universes.

6 Future research directions

Future work may extend the analysis to multiple years of data and additional equity markets to examine the robustness of the observed results across different market environments. The findings of this study may also be combined with volatility-based screening and Black-Litterman portfolio optimization frameworks to investigate whether assets consistently favored by risk-aware allocation strategies can improve portfolio construction and asset selection decisions.

References

[1] F. Black and R. Litterman, Global portfolio optimization, *Financ. Anal. J.* 48 (1992), 28-43.

[2] P. Chen and D. Luo, Dynamic optimization strategy of financial portfolios using deep reinforcement learning (DRL)-based neural networks, *Int. J. Decis. Support Syst. Technol.* 18 (2026), 1-20.

[3] R. Clarke, H. de Silva and S. Thorley, Risk parity, maximum diversification, and minimum variance: An analytic perspective, *J. Portf. Manag.* 39 (2013), 39-53.

[4] V. DeMiguel, L. Garlappi and R. Uppal, Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy?, *Rev. Financ. Stud.* 22 (2009), 1915-1953.

[5] F.J. Fabozzi, P.N. Kolm, D.A. Pachamanova and S.M. Focardi, Robust portfolio optimization, *J. Portf. Manag.* 33 (2007), 40-48.

[6] M. Kaucic, R. Pelessoni and F. Piccotto, An automated decision support system for portfolio allocation based on mutual information and financial criteria, *Entropy* 27 (2025), 480.

[7] P.N. Kolm, R. Tütüncü and F.J. Fabozzi, 60 years of portfolio optimization: Practical challenges and current trends, *Eur. J. Oper. Res.* 234 (2014), 356-371.

[8] O. Ledoit and M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, *J. Multivar. Anal.* 88 (2004), 365-411.

[9] S. Maillard, T. Roncalli and J. Teiletche, The properties of equally weighted risk contribution portfolios, *J. Portf. Manag.* 36 (2010), 60-70.

[10] H. Markowitz, Portfolio selection, *J. Finance* 7 (1952), 77-91.

[11] MathWorks, *Backtest Framework*, MATLAB R2026a documentation, The MathWorks, Inc., Natick, MA, 2026.

[12] MathWorks, *Backtest Investment Strategies*, MATLAB R2026a documentation, The MathWorks, Inc., Natick, MA, 2026.

[13] A. Meucci, Managing diversification, *Risk* 22 (2009), 74-79.

[14] R.O. Michaud, The Markowitz optimization enigma: Is 'optimized' optimal?, *Financ. Anal. J.* 45 (1989), 31-42.

[15] D. Palit and V.R. Prybutok, A comparative study of the equal-weight method and hierarchical risk parity in portfolio construction, *Finance Econ. Rev.* 6 (2024), 1-12.

[16] Ş. Rusu and M. Boloş, Bridging tradition and innovation: A literature review on portfolio optimization, *Ann. Univ. Oradea Econ. Sci.* 33 (2024), 337-349.

[17] A. Salo, M. Doumpos, J. Liesiö and C. Zopounidis, Fifty years of portfolio optimization, *Eur. J. Oper. Res.* 318 (2024), 1-18.

[18] N. Shahid, Algorithmically efficient identification of volatile KSE-30 equities and their role in optimized portfolio allocation, *Oper. Res. Forum* 6 (2025), 23.

[19] N. Shahid, Black-Litterman portfolio optimization for Pakistan Stock Exchange (PSX) KSE-30 index, *Discov. Anal.* 3 (2025), 4.

[20] W.F. Sharpe, Capital asset prices: A theory of market equilibrium under conditions of risk, *J. Finance* 19 (1964), 425-442.

[21] W.F. Sharpe, Mutual fund performance, *J. Bus.* 39 (1966), 119-138.

[22] K. Sutiene, P. Schwendner, C. Sipos, L. Lorenzo, M. Mirchev, P. Lameski, A. Kabasinskas, C. Tidjani, B. Ozturkkal and J. Cerneviciene, Enhancing portfolio management using artificial intelligence: Literature review, *Front. Artif. Intell.* 7 (2024), 1371502.

[23] X. Tu and B. Li, Robust portfolio selection with smart return prediction, *Econ. Model.* 135 (2024), 106719.

[24] R. Wu, Y. Yang, H.L. Shang and H. Zhu, Making distributionally robust portfolios feasible in high dimension, *J. Econometrics* 252 (2025), 106118.

[25] H. Zhao, Y. Jiang and Y. Yang, Robust and sparse portfolio: Optimization models and algorithms, *Mathematics* 11 (2023), 4925.

Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 18, No. 4/2026

THE MEAN-SQUARE STABILITY AND MEAN-SQUARE STABILIZABILITY OF THE DISCRETE-TIME MEAN-FIELD STOCHASTIC LINEAR SYSTEMS WITH PERIODIC COEFFICIENTS*

Vasile Drăgan†

DOI 10.56082/annalsarscimath.2026.4.339

Abstract

In this work, several aspects of the mean-square stability and of the mean-square stabilizability of a class of discrete-time mean-field stochastic linear systems (MFSLS) with periodic coefficients is addressed. To a discrete-time MFSLS one associates in a natural way a companion standard discrete-time stochastic linear system (DTSLS). Our aim is to compare the mean-field exponential stability in mean-square (MFESMS) property of the MFSLS under consideration with the exponential stability in mean-square (ESMS) property of its companion DTSLS. Even if in some already published works is asserted that the two concepts of exponential stability are equivalent, we show that this equivalence is not true. We also show that the MF mean-square stabilizability of a discrete-time controlled MFSLS is not equivalent to the mean-square stabilizability of its controlled companion DTSLS.

*Accepted for publication on October 19, 2026

†Vasile.Dragan@imar.ro, *Simion Stoilow* Institute of Mathematics of the Romanian Academy, P.O.Box 1-764, RO-014700, Bucharest, Romania; Academy of Romanian Scientists, Str. Ilfov, 3, Bucharest, Romania

Keywords: mean-field stochastic linear systems, mean-square stability, mean-square stabilizability, periodic coefficients.
MSC: 93E15, 93B05, 93B07, 93D15, 47B65.

1 Introduction

In recent years, mean-field control theory has received increasing interest, both from a theoretical perspective as well as from engineering and finance applications. The main distinctive feature of the mean-field setting when compared with classical stochastic optimal control is that the system dynamics involves state and control, as well as their expected values.
Since the pioneering works [10,12] there have been several important contributions related to this type of stochastic systems from the control perspective [1–5, 8, 11, 14, 16].

In the present work, we consider a class of discrete-time mean-field controlled stochastic linear systems with periodic coefficients. For this kind of stochastic systems we introduce the concept of mean-field exponential stability in mean-square (MF-ESMS) (Definition 1) and the concept of mean-field stabilizability in mean-square by linear state feedback (Definition 2).
The MF-ESMS property is characterized by exponential decay of the trajectories with random initial states of the mean-field stochastic system under consideration. We show that without loss of generality, we can characterize the MF-ESMS property only by exponential decay of the trajectories starting from deterministic initial states (Proposition 1). To a mean-field stochastic linear system one may associate, in a natural way, a companion discrete-time stochastic linear system with state multiplicative and /or control multiplicative independent random disturbances. In some already published papers ([13], [15]), is asserted that the MF-ESMS property of a mean-field stochastic linear system is equivalent to the standard exponential stability in mean-square (ESMS) of its discrete-time companion stochastic linear system. Through Example 1, we show that this equivalence is not always true. However, it remains true that the ESMS property of the companion discrete-time linear stochastic system implies the MF-ESMS property of the mean-field stochastic linear system under consideration. The exponential stability in mean-square of the mean-field stochastic linear system induced by the ESMS property of its companion discrete-time linear stochastic system is named, in this work, "mean-field strong exponential stability in mean-square (MF-SESMS)". The Lyapunov type criteria for exponential stability in mean-square of the companion discrete-time linear stochastic

system provide necessary and sufficient conditions for MF-SESMS property but only sufficient conditions for the MF-ESMS property of the mean-field stochastic linear system under consideration. By the Example 2, we show that the mean-field stabilizability in mean-square of a mean-field controlled stochastic linear system is not equivalent to the standard stochastic stabilizability in mean-square by a decentralized linear state feedback.
The mean-field stabilizability of the mean-field controlled stochastic linear system induced by the stochastic stabilizability by a decentralized linear state feedback of its companion controlled discrete-time stochastic linear system is named MF strong stabilizability in mean-square of the mean-field controlled stochastic linear system. This property may be used in future research, in order to establish the existence of the maximal solution and of the stabilizing solution of a Riccati type equation involved in the solution of some optimal control problems for mean-field controlled stochastic linear systems. The developments from sections 3 and 4, of the present work, are the discrete-time counterpart of the results from the recent paper [7].

2 The problem setting

2.1 Notation

In this work $(\Omega, \mathfrak{F}, \mathcal{P})$ is a probability space where a sequence of r-dimensional independent random vectors $\{w(t)\}_{t\in\mathbb{Z}_+}$, $w(t) = (w_1(t) \;\; w_2(t) \;\; \dots \;\; w_r(t))^\top$ is defined. Here and in the sequel $\mathbb{Z}_+ = \{0, 1, 2....\}$.

We assume that the following assumption holds:

H1). $\mathbb{E}[w(t)] = 0$ and $\mathbb{E}[w(t)w^\top(t)] = I_r$ for all $t \in \mathbb{Z}_+$, where I_r is the identity matrix of size $r \times r$.

Throughout this work $\mathbb{E}[\cdot]$ denotes the mathematical expectation of a random variable, while the superscript $\top$ stands for the transpose of a matrix or a vector. For each $t \in \mathbb{Z}_+$, $\mathcal{F}_t := \{\sigma(w(s)), 0 \le s \le t\} \subset \mathfrak{F}$ denotes the smallest σ-algebra with respect to which the random vectors $(w(s)$, $0 \le s \le t$, are measurable.

If $t_0 \in \mathbb{Z}_+$, $\ell^2_{\mathcal{F}}\{t_0, \infty; \mathbb{R}^d\}$ denotes the linear space of all sequences $\{v(t)\}_{t\ge t_0}$ of $\mathbb{R}^d$ valued random vectors with the properties:

a) for each $t \ge t_0$, $v(t)$ are $\mathcal{F}_t$-measurable;

b) $\sum\limits_{t=t_0}^{\infty} \mathbb{E}[|v(t)|^2] < \infty$.

Here and in the sequel $|\cdot|$ denotes the Euclidean norm of a vector i.e. $|v| = (v^\top v)^{1/2}$ for all $v \in \mathbb{R}^d$.
As usual $\mathbb{R}^{m\times n}$ denotes the linear space of $m \times n$ real matrices. For $p \ge 2$,

$\mathcal{S}_p \subset \mathbb{R}^{p\times p}$ stands for the space of symmetric matrices of size $p\times p$. $\mathcal{S}_p^+ \subset \mathcal{S}_p$ denotes the closed, solid, normal convex cone of the positive semidefinite matrices of size $p \times p$.
If $t_0 \in \mathbb{Z}_+$, we denote $\mathcal{X}_{t_0}$ the set of all n-dimensional random vectors x_0 which are independent of all random vectors $w(t)$, $t > t_0$ and are satisfying $\mathbb{E}[|x_0|^2] < \infty$.

Remark 1. *The set $\mathcal{X}_{t_0}$ contains the n-dimensional random vectors x_0 with finite second order moments which are either $\mathcal{F}_{t_0}$-measurable, or, are independent of all random vectors $w(t)$, $t \in \mathbb{Z}_+$.*

2.2 Model description

We consider the discrete-time controlled stochastic linear system having the state space representation of the form:

$$x(t+1) = A_0(t)x(t) + \overline{A}_0(t)\mathbb{E}[x(t)] + B_0(t)u(t) + \overline{B}_0(t)\mathbb{E}[u(t)] + \quad (1)$$
$$\sum_{k=1}^{r}(A_k(t)x(t) + \overline{A}_k(t)\mathbb{E}[x(t)] + B_k(t)u(t) + \overline{B}_k(t)\mathbb{E}[u(t)])w_k(t), t \geq t_0,$$

$x(t_0) = x_0 \in \mathcal{X}_{t_0}$, where $x(t) \in \mathbb{R}^n$ are the vectors of the state parameters and $u(t) \in \mathbb{R}^m$ are the vectors of the control parameters.
In (1), $\{A_k(t)\}_{t\in\mathbb{Z}}$, $\{\overline{A}_k(t)\})_{t\in\mathbb{Z}}$ are matrix valued sequences from $\mathbb{R}^{n\times n}$ and $\{B_k(t)\}_{t\in\mathbb{Z}}$, $\{\overline{B}_k(t)\}_{t\in\mathbb{Z}}$, $0 \leq k \leq r$ are matrix valued sequences in $\mathbb{R}^{n\times m}$ which are periodic of period $\theta \geq 1$.
The set $\mathcal{U}_{ad}(t_0)$ of the admissible controls consists of all sequences $\{u(t)\}_{t\geq t_0}$ of $\mathbb{R}^m$ valued random vectors with the properties:

a) for each $t \geq t_0$, $u(t)$ is $\mathcal{F}_t$-measurable;

b) $\mathbb{E}[|u(t)|^2] < \infty$.

The controlled stochastic linear system (1) is the discrete-time controlled linear system of McKean-Vlasov type which is also known as a discrete-time mean-field controlled stochastic linear system (MFCSLS).

The study of the stochastic differential equations of McKean-Vlasov type was initiated in [10, 12]. Among the most recent papers which studied the control problems involving discrete-time stochastic linear equations of McKean-Vlasov type, are [13, 15].

For each $0 \leq t_0 < t \in \mathbb{Z}_+$, $\mathcal{G}_{t_0,t} \subset \mathfrak{F}$ stands for the smallest σ-algebra with respect to which the random vectors $x_0 \in \mathcal{X}_{t_0}$ and $w(s)$, $t_0 \leq s \leq t-1$ are measurable.

The next lemma collects several properties of the solutions of (1). Their proofs may be done recursively with respect to $t \geq t_0$.

Lemma 1. *For each $t_0 \in \mathbb{Z}_+$, $x_0 \in \mathcal{X}_{t_0}$ and $\mathbf{u} = \{u(t)\}_{t\geq t_0} \in \mathcal{U}_{ad}(t_0)$, the problem with given initial values (IVP) (1) has a unique solution $x(t; t_0, x_0, \mathbf{u})$, with the properties:*

a) $x(t; t_0, x_0, \mathbf{u})$ is $\mathcal{G}_{t_0,t}$-measurable;

b) $\mathbb{E}[|x(t; t_0, x_0, \mathbf{u})|^2] < \infty$;

c) $x(t_0; t_0, x_0, \mathbf{u}) = x_0$.

Bearing in mind the assumption **H1**) and Lemma 1 a), we obtain that the random vectors $x(t; t_0, x_0, \mathbf{u})$ and $w(t)$ are mutually independent for any $t \geq t_0$. This allows us to obtain from (1) that if $x(t) \triangleq x(t; t_0, x_0, \mathbf{u})$, $t \geq t_0$, $\{\mathbb{E}[x(t)]\}_{t\geq t_0}$ is satisfying the following IVP:

$$\mathbb{E}[x(t+1)] = (A_0(t) + \overline{A}_0(t))\mathbb{E}[x(t)] + (B_0(t) + \overline{B}_0(t))\mathbb{E}[u(t)] \tag{2a}$$

$$\mathbb{E}[x(t_0)] = \mathbb{E}[x_0]. \tag{2b}$$

Subtracting (2) from (1), we obtain:

$$\begin{aligned} &x(t+1) - \mathbb{E}[x(t+1)] = A_0(t)(x(t) - \mathbb{E}[x(t)]) + B_0(t)(u(t) - \mathbb{E}[u(t)]) \\ &+ \sum_{k=1}^{r}[A_k(t)(x(t) - \mathbb{E}[x(t)]) + (A_k(t) + \overline{A}_k(t))\mathbb{E}[x(t)] \\ &+ B_k(t)(u(t) - \mathbb{E}[u(t)]) + (B_k(t) + \overline{B}_k(t))\mathbb{E}[u(t)]]w_k(t), t \geq t_0, \end{aligned} \tag{3a}$$

$$x(t_0) - \mathbb{E}[x(t_0)] = x_0 - \mathbb{E}[x_0]. \tag{3b}$$

We set

$$x^1(t) \triangleq x(t) - \mathbb{E}[x(t)] \tag{4a}$$

$$x^2(t) \triangleq \mathbb{E}[x(t)] \tag{4b}$$

$$u^1(t) \triangleq u(t) - \mathbb{E}[u(t)] \tag{5a}$$

$$u^2(t) \triangleq \mathbb{E}[u(t)] \tag{5b}$$

$t \geq t_0$. With these notations, we obtain from (2) and (3) the following IVP:

$$\begin{aligned} &\begin{pmatrix} x^1(t+1) \\ x^2(t+1) \end{pmatrix} = \mathbb{A}_0(t)\begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} + \mathbb{B}_0(t)\begin{pmatrix} u^1(t) \\ u^2(t) \end{pmatrix} \\ &+ \sum_{k=1}^{r}(\mathbb{A}_k(t)\begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} + \mathbb{B}_k(t)\begin{pmatrix} u^1(t) \\ u^2(t) \end{pmatrix})w_k(t), t \geq t_0 \end{aligned} \tag{6a}$$

$$\begin{pmatrix} x^1(t_0) \\ x^2(t_0) \end{pmatrix} = \begin{pmatrix} x_0 - \mathbb{E}[x_0] \\ \mathbb{E}[x_0] \end{pmatrix} \tag{6b}$$

with:

$$\mathbb{A}_0(t) = \begin{pmatrix} A_0(t) & 0 \\ 0 & A_0(t) + \overline{A}_0(t) \end{pmatrix} \tag{7a}$$

$$\mathbb{A}_k(t) = \begin{pmatrix} A_k(t) & A_k(t) + \overline{A}_k(t) \\ 0 & 0 \end{pmatrix}, 1 \leq k \leq r, \tag{7b}$$

$$\mathbb{B}_0(t) = \begin{pmatrix} B_0(t) & 0 \\ 0 & B_0(t) + \overline{B}_0(t) \end{pmatrix} \tag{8a}$$

$$\mathbb{B}_k(t) = \begin{pmatrix} B_k(t) & B_k(t) + \overline{B}_k(t) \\ 0 & 0 \end{pmatrix}, 1 \leq k \leq r. \tag{8b}$$

Along with the IVP (6), we introduce the following IVP on the space $\mathbb{R}^n \times \mathbb{R}^n$:

$$\mathbf{x}(t+1) = \mathbb{A}_0(t)\mathbf{x}(t) + \mathbb{B}_0(t)\mathbf{u}(t) + \sum_{k=1}^{r} (\mathbb{A}_k(t)\mathbf{x}(t) + \mathbb{B}_k(t)\mathbf{u}(t))w_k(t), \tag{9a}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0 = \begin{pmatrix} x_0^1 \\ x_0^2 \end{pmatrix} \tag{9b}$$

where $\mathbf{x}(t) = \begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^n$ and $\mathbf{u}(t) = \begin{pmatrix} u^1(t) \\ u^2(t) \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^m$, $x_0^k \in \mathcal{X}_{t_0}$, $k = 1, 2$. The IVP (6) is the particular case of the IVP (9), when $x_0^1 = x_0 - \mathbb{E}[x_0]$ and $x_0^2 = \mathbb{E}[x_0]$. In the sequel, the discrete-time controlled linear stochastic system, having the state space representation described by (6a) (or (9a), respectively) will be named the companion controlled difference stochastic linear system (CDSLS), associated to the MFCSLS (1).

When $B_k(t) = \overline{B}_k(t) = 0$, $t \in \mathbb{Z}$, $0 \leq k \leq r$, the MFCSLS (1) reduces to the uncontrolled mean-field stochastic linear system (MFSLS):

$$x(t+1) = A_0(t)x(t) + \overline{A}_0(t)\mathbb{E}[x(t)] + \sum_{k=1}^{r} (A_k(t)x(t) + \overline{A}_k(t)\mathbb{E}[x(t)])w_k(t), \tag{10a}$$

$$x(t_0) = x_0 \in \mathcal{X}_{t_0}. \tag{10b}$$

The IVPs (6) and (9) take the form:

$$\begin{pmatrix} x^1(t+1) \\ x^2(t+1) \end{pmatrix} = \mathbb{A}_0(t) \begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} + \sum_{k=1}^{r} \mathbb{A}_k(t) \begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} w_k(t), \tag{11a}$$

$$\begin{pmatrix} x^1(t_0) \\ x^2(t_0) \end{pmatrix} = \begin{pmatrix} x_0 - \mathbb{E}[x_0] \\ \mathbb{E}[x_0] \end{pmatrix} \tag{11b}$$

$$\mathbf{x}(t+1) = \mathbb{A}_0(t)\mathbf{x}(t) + \sum_{k=1}^{r} \mathbb{A}_k(t)\mathbf{x}(t)w_k(t), t \geq t_0 \tag{12a}$$

$$\mathbf{x_0}(t_0) = \mathbf{x}_0 = \begin{pmatrix} x_0^1 \\ x_0^2 \end{pmatrix} \tag{12b}$$

where $x_0^k \in \mathcal{X}_{t_0}$, $k = 1, 2$. The discrete-time stochastic linear system (11a) ((12a), respectively), will be named the companion discrete-time stochastic linear system (DTSLS) associated to the MFSLS (10a).

2.3 Basic definitions and the problem setting

Definition 1. *We say that the zero solution of the MFSLS (10a) or equivalently the MFSLS (10) is mean-field exponentially stable in mean square (MF-ESMS) if for all initial time $t_0 \in \mathbb{Z}_+$ and any initial state $x_0 \in \mathcal{X}_{t_0}$, the corresponding solution $x(\cdot; t_0, x_0)$ of (10a) has a decay of the form*

$$\mathbb{E}[|x(t; t_0, x_0)|^2] \leq \beta\delta^{t-t_0}\mathbb{E}[|x_0|^2] \tag{13}$$

for all $t \geq t_0$, where $\beta \geq 1$, $\delta \in (0, 1)$ are constants not depending upon t, t_0 and x_0.

Remark 2. *Unlike the nonlinear case, where the stability property refers to a given solution, in the case of a linear system, the stability property of a solution is equivalent to the stability of all solutions of that system. That is why, we may say that the MFSLS (10a) is MF-ESMS if its solutions are satisfying (13).*

Definition 2. *We say that the MFCSLS (1) is mean-field stabilizable in mean-square, or equivalently mean-field stabilizable by linear state feedback if there exist θ-periodic $\mathbb{R}^{m\times n}$ valued matrix sequences $\{F(t)\}_{t\in\mathbb{Z}}$, $\{\overline{F}(t)\}_{t\in\mathbb{Z}}$ with the property that the corresponding closed-loop MFSLS*

$$\begin{aligned} x(t+1) &= (A_0(t) + B_0(t)F(t))x(t) + [\overline{A}_0(t) + (B_0(t) + \overline{B}_0(t))\overline{F}(t) \\ &- B_0(t)F(t)]\mathbb{E}[x(t)] + \sum_{k=1}^{r}\{(A_k(t) + B_k(t)F(t))x(t) + [\overline{A}_k(t) + (B_k(t) \\ &+ \overline{B}_k(t))\overline{F}(t) - B_k(t)F(t)]\mathbb{E}[x(t)]\}w_k(t), t \geq t_0 \end{aligned} \tag{14a}$$

$$x(t_0) = x_0 \in \mathcal{X}_{t_0} \tag{14b}$$

is MF-ESMS in the sense of Definition 1.

The previous definition is a time-varying version of the Definition 3.2 from [15].

Our aim is to highlight several useful properties of the MF-ESMS of a MFSLS (10) and of the mean-field stabilizability in mean-square of a MFCSLS of the form 1). We shall emphasize the relationship between the MF-ESMS property of the MFSLS (10) and ESMS property of the corresponding companion DTSLS (12a).

We shall see that the ESMS property of the companion DTSLS (12) implies the MF-ESMS of the MFSLS (10) but the converse implication is not always true. Also, we shall show that the stochastic stabilizability in mean-square of the corresponding companion discrete-time controlled stochastic linear system implies the mean-field stabilizability in mean-square of the MFCSLS (1), but the converse implication is not always true.

Based on the particular form of the matrix coefficients of the discrete-time stochastic linear system (12a) given in (7)-(8), we obtain Lyapunov type criteria for the ESMS property of the companion DTSLS (12) of lower dimension than those obtained applying directly the already existing results from the literature.

3 The mean-field exponential stability in mean-square

3.1 An equivalent characterization of the MF-ESMS property

First we show that without loss of generality, we can characterize the MF-ESMS property of the MFSLS (10a) based only on exponential decay of its solutions starting from deterministic initial states.

Proposition 1. *Under the assumption* **H1**) *the following are equivalent:*
(i) the MFSLS (10a) is MF-ESMS in the sense of Definition 1;
(ii)

$$\mathbb{E}[|x(t;t_0,x_0)|^2] \leq \beta\delta^{t-t_0}|x_0|^2 \tag{15}$$

for all $t \geq t_0 \geq 0$, $x_0 \in \mathbb{R}^n$, *where* $\beta \geq 1$, $\delta \in (0,1)$ *are constants not depending upon* t, t_0, x_0.

Proof. The implication $(i) \Rightarrow (ii)$ is obvious because $\mathbb{R}^n \subset \mathcal{X}_{t_0}$ for any $t_0 \in \mathbb{Z}_+$. Further, we proof $(ii) \Rightarrow (i)$. Let $t_0 \in \mathbb{Z}_+$ and $\tilde{x}_0 \in \mathcal{X}_{t_0}$ be arbitrary but fixed. We have to show that the solution $x(\cdot;t_0,\tilde{x}_0)$ of the MFSLS (10a)

corresponding to this initial pair has an exponential decay as that from (13). For each $1 \leq j \leq n$, we consider $\xi_j(\cdot, t_0)$ the solution of the following IVP:

$$\xi_j(t+1) = A_0(t)\xi_j(t) + \overline{A}_0(t)\mathbb{E}[\xi_j(t)] + \sum_{k=1}^{r}(A_k(t)\xi_j(t) + \overline{A}_k(t)\mathbb{E}[\xi_j(t)])w_k(t), t \geq t_0 \tag{16a}$$

$$\xi_j(t_0) = e_j \tag{16b}$$

where $e_j \in \mathbb{R}^n$ is the j-th column of the identity matrix I_n. Comparing the IVP (16) to the IVP (10), we obtain based on the unicity of the solution of an IVP, that

$$\xi_j(t, t_0) = x(t; t_0, e_j)$$

for all $t \geq t_0$. If (ii) holds, we obtain from (15) applied in the case of the previous equality that

$$\mathbb{E}[|\xi_j(t; t_0)|^2] \leq \beta\delta^{t-t_0} \tag{17}$$

for all $t \geq t_0$.

We define the matrix $\Xi(t; t_0) = \begin{pmatrix} \xi_1(t, t_0) & \xi_2(t, t_0) & \dots & \xi_n(t, t_0) \end{pmatrix}$. Bearing in mind (16) we may infer that $\{\Xi(t; t_0)\}_{t \geq t_0}$ is the matrix valued solution of the following IVP:

$$\Xi(t+1) = A_0(t)\Xi(t) + \overline{A}_0\mathbb{E}[\Xi(t)] + \sum_{k=1}^{r}(A_k(t)\Xi(t) + \overline{A}_k(t)\mathbb{E}[\Xi(t)])w_k(t), \tag{18a}$$

$$\Xi(t_0; t_0) = I_n. \tag{18b}$$

Starting from the definition of the euclidian norm of a symmetric positive semidefinite matrix, we may write successively:

$$\|\mathbb{E}[\Xi^T(t, t_0)\Xi(t, t_0)]\|_2 = \lambda_{max}(\mathbb{E}[\Xi^T(t, t_0)\Xi(t, t_0)]) \leq Tr[\mathbb{E}[\Xi^T(t, t_0)\Xi(t, t_0)]] = \sum_{j=1}^{n}\mathbb{E}[|\xi_j(t, t_0)|^2] \tag{19}$$

where $\lambda_{max}(\cdot)$ denotes the maximal eigenvalue of a positive semidefinite symmetric matrix and $Tr[\cdot]$ stands for the trace of a matrix.

Further, (17) and (19) yield

$$\|\mathbb{E}[\Xi^T(t, t_0)\Xi(t, t_0)]\|_2 \leq n\beta\delta^{t-t_0} \tag{20}$$

for all $t \geq t_0 \geq 0$.

Let $\tilde{x}(t) := \Xi(t;t_0)\tilde{x}_0$. Employing (18), we deduce that $\tilde{x}(t)$ is the solution of the IVP (10) corresponding to the initial pair $(t_0, \tilde{x}_0)$. Hence,

$$x(t;t_0,\tilde{x}_0) = \Xi(t,t_0)\tilde{x}_0 \tag{21}$$

for all $t \geq t_0 \geq 0$, $\tilde{x}_0 \in \mathcal{X}_{t_0}$. Let $\sigma[\tilde{x}_0] \subset \mathfrak{F}$ be the smallest σ-algebra with respect to which the random vector $\tilde{x}_0$ is measurable.

Based on the properties of the conditional expectation, we obtain from (21) that

$$\begin{aligned}\mathbb{E}[|x(t;t_0,\tilde{x}_0)|^2|\sigma[\tilde{x}_0]] = \mathbb{E}[\tilde{x}_0^T\Xi^T(t,t_0)\Xi(t,t_0)\tilde{x}_0|\sigma[\tilde{x}_0]] = \\ \tilde{x}_0^T\mathbb{E}[\Xi^T(t,t_0)\Xi(t,t_0)|\sigma[\tilde{x}_0]]\tilde{x}_0.\end{aligned} \tag{22}$$

Bearing in mind the properties of the random vectors from $\mathcal{X}_{t_0}$, we may infer via (18) that the components of the matrix $\Xi^T(t,t_0)\Xi(t,t_0)$ are random vectors independent with respect to the σ-algebra $\sigma[\tilde{x}_0]$. Hence

$$\mathbb{E}[\Xi^T(t,t_0)\Xi(t,t_0)|\sigma[\tilde{x}_0]] = \mathbb{E}[\Xi^T(t,t_0)\Xi(t,t_0)] \ \ a.s. \tag{23}$$

Since $0 \leq \mathbb{E}[\Xi^T(t,t_0)\Xi(t,t_0)] \leq \|E[\Xi^T(t,t_0)\Xi(t,t_0)]\|_2 I_n$ we obtain via (20), (22), (23) that

$$\mathbb{E}[|x(t;t_0,\tilde{x}_0)|^2|\sigma[\tilde{x}_0]] \leq n\beta\delta^{t-t_0}|\tilde{x}_0|^2 \ \ a.s.$$

for all $t \geq t_0 \geq 0$, $\tilde{x}_0 \in \mathcal{X}_t$.

Taking the expectation in the previous inequality, we may conclude that $\mathbb{E}[|x(t;t_0,\tilde{x}_0)|^2] \leq n\beta\delta^{t-t_0}\mathbb{E}[|\tilde{x}_0|^2]$ for all $t \geq t_0 \geq 0$, $\tilde{x}_0 \in \mathcal{X}_{t_0}$. Thus the proof ends.

3.2 The connection between the MF-ESMS property of the MFSLS and the ESMS property of its companion DTSLS

Let $x(\cdot;t_0,x_0)$ be the solution of the MFSLS (10a) corresponding to an arbitrary initial pair (t_0,x_0), where $t_0 \in \mathbb{Z}_+$ and $x_0 \in \mathcal{X}_{t_0}$.

We set

$$x^1(t;t_0,x_0) \triangleq x(t;t_0,x_0) - \mathbb{E}[x(t;t_0,x_0)] \tag{24a}$$

$$x^2(t;t_0,x_0) \triangleq \mathbb{E}[x(t;t_0,x_0)] \tag{24b}$$

$t \geq t_0$ and obtain the solution of the IVP (11) generated by the solution $x(\cdot; t_0, x_0)$ of the IVP (10). Conversely, if $\begin{pmatrix} x^1(\cdot; t_0, x_0) \\ x^2(\cdot; t_0, x_0) \end{pmatrix}$ is a solution of the IVP (11), then $x(\cdot, t_0, x_0)$ defined by

$$x(t; t_0, x_0) \triangleq x^1(t; t_0, x_0) + x^2(t; t_0, x_0), t \geq t_0 \tag{25}$$

is the solution of the IVP (10).

Employing the equalities (24) and (25) we stabilize an equivalence between the MF-ESMS property of the MFSLS (10a) and the exponential decay of the solutions of the IVP (11) for an arbitrary $x_0 \in \mathcal{X}_{t_0}$.

Proposition 2. *Under the assumption* **H1**), *the following are equivalent:*
(i) the MFSLS (10a) is MF-ESMS in the sense of the Definition 1;
(ii) for each initial time $t_0 \in \mathbb{Z}_+$ and any initial state $x_0 \in \mathcal{X}_{t_0}$, the solution $\begin{pmatrix} x^1(\cdot; t_0, x_0) \\ x^2(\cdot; t_0, x_0) \end{pmatrix}$ of the corresponding IVP (11) has a decay of the form:

$$\mathbb{E}[|x^1(t; t_0, x_0)|^2] + \mathbb{E}[|x^2(t; t_0, x_0)|^2] \leq \tilde{\beta}\tilde{\delta}^{t-t_0}(\mathbb{E}[|x_0 - \mathbb{E}[x_0]|^2] + |\mathbb{E}[x_0]^2), \tag{26}$$

where $\tilde{\beta} \geq 1$, $\tilde{\delta} \in (0, 1)$ are constants not depending upon t, t_0, x_0.

Proof: Assuming that (i) holds, we obtain from (13), (24b), that:

$$\mathbb{E}[|x^2(t; t_0, x_0)|^2] \leq \beta\delta^{t-t_0}\mathbb{E}[|x_0|^2] \tag{27}$$

for all $t \geq t_0 \in \mathbb{Z}_+$ and $x_0 \in \mathcal{X}_{t_0}$, with $\beta \geq 1$, $\delta \in (0, 1)$ not depending upon t, t_0, x_0. Further, (25) gives

$$\mathbb{E}[|x(t; t_0, x_0)|^2] = \mathbb{E}[|x^1(t; t_0, x_0)|^2] + \mathbb{E}[|x^2(t; t_0, x_0)|^2] \tag{28}$$

for all $t \geq t_0 \in \mathbb{Z}_+$ because $\mathbb{E}[x^1(t; t_0, x_0)(x^2(t; t_0, x_0))^T] = 0$ for all $t \geq t_0$. Employing (13), (24a), (27), (28) we deduce that:

$$\mathbb{E}[|x^1(t; t_0, x_0)|^2] \leq 2\beta\delta^{t-t_0}\mathbb{E}[x_0|^2].$$

Combining this inequality to (27), we obtain:

$$\mathbb{E}[|x^1(t; t_0, x_0)|^2] + \mathbb{E}[|x^2(t; t_0, x_0)|^2] \leq 3\beta\delta^{t-t_0}\mathbb{E}[|x_0|^2] \tag{29}$$

for all $t \geq t_0 \in \mathbb{Z}_+$, $x_0 \in \mathcal{X}_{t_0}$. Bearing in mind that

$$\mathbb{E}[|x_0|^2] = \mathbb{E}[|x_0 - \mathbb{E}[x_0]|^2] + |\mathbb{E}[x_0]|^2$$

we may conclude that (29) coincides to (26), when $\tilde{\beta} = 3\beta$, $\tilde{\delta} = \delta$. Thus the validity of the implication $(i) \Rightarrow (ii)$ is confirmed. In order to check the validity of the implication $(ii) \Rightarrow (i)$ we remark that (28) allows us to deduce that (13) holds if (26) is true. Thus the proof ends.

Remark 3. *a) Even if $x^2(t;t_0,x_0)$ are deterministic vectors, we preferred to use the notation $\mathbb{E}[|x^2(t;t_0,x_0)|^2]$ in (26) in order to preserve the unit of the presentation.*

b) From (24b) we obtain, via (2) written for $B_0(t)=\overline{B}_0(t)=0$ for all $t\in\mathbb{Z}$, that $\{x^2(t;t_0,x_0)\}_{t\geq t_0}$ solves the deterministic discrete-time linear equation (DTLE) on the euclidian space $\mathbb{R}^n$:

$$x^2(t+1)=(A_0(t)+\overline{A}_0(t))x^2(t). \tag{30}$$

Invoking the implication $(i)\Rightarrow(ii)$ from Proposition 2 we may infer that the exponential stability of the zero solution of the deterministic DTLE (30) is a necessary condition for the MF-ESMS property of the MFSLS (10a).

For the reader convenience, we recall here the definition of the ESMS property of the DTSLS (12a).

Definition 3. *We say that the zero solution of the DTSLS (12a) or equivalently, the system (12a) is exponentially stable in mean-square (ESMS) if its solutions have a decay of the form*

$$\mathbb{E}[|x^1(t;t_0,\mathbf{x}_0)|^2]+\mathbb{E}[|x^2(t;t_0,\mathbf{x}_0)|^2]\leq\beta_1\delta_1^{t-t_0}(\mathbb{E}[|x_0^1|^2]+\mathbb{E}[|x_0^2|^2]) \tag{31}$$

for all $t\geq t_0\in\mathbb{Z}_+$ and any initial states $\mathbf{x}_0=\begin{pmatrix}x_0^1\\x_0^2\end{pmatrix}\in\mathcal{X}_{t_0}\times\mathcal{X}_{t_0}$. In (31), $\beta_1\geq 1$, $\delta_1\in(0,1)$ are constants not depending upon t,t_0,x_0^k, $k=1,2$.

Remark 4. *Bearing in mind that the initial states (11) are particular case of (12b) we may conclude via the implication $(ii)\Rightarrow(i)$ from Proposition 2 that the MFSLS (10a) is MF-ESMS if its companion DTSLS (12a) is ESMS. The next example shows that the converse implication is not always true.*

Example 1. *We consider the 1-dimensional discrete-time, time-invariant MFSLS:*

$$x(t+1)=a_0x(t)+\overline{a}_0\mathbb{E}[x(t)]+\sum_{k=1}^{r}(a_kx(t)+\overline{a}_k\mathbb{E}[x(t)])w_k(t),t\geq t_0 \tag{32a}$$

$$x(t_0)=x_0\in\mathbb{R} \tag{32b}$$

where $a_k,\overline{a_k}$, $0\leq k\leq r$ are real numbers subject to the following constraints:

c_1) $a_k+\overline{a}_k=0$, $1\leq k\leq r$;

c_2) $|a_0+\overline{a}_0|<1$;

$c_3)$ $\sum\limits_{k=0}^{r} a_k^2 \geq 1$,

$\{w_k(t)\}_{t\geq 0}$, $1 \leq k \leq r$, *are sequences of random variables which are satisfying the conditions from the assumption* $\mathbf{H}1$). *According to Proposition 1, for the study of the MF-ESMS property of the MFSLS (32) we may assume without loss of generality that the initial states* x_0 *are deterministic. Under the constraint* c_1), *the corresponding IVP (11) becomes:*

$$\begin{aligned} x^1(t+1) &= (a_0 + \sum_{k=1}^{r} a_k w_k)x^1(t), t \geq t_0 \\ x^1(t_0) &= 0 \end{aligned} \tag{33}$$

$$\begin{aligned} x^2(t+1) &= (a_0 + \overline{a}_0)x^2(t), t \geq t_0 \\ x^2(t_0) &= x_0. \end{aligned} \tag{34}$$

Based on the unicity of the solution of an IVP we obtain that the unique solution of the IVP (33) is $x^1(t) = 0$ *a.s.* $t \geq t_0$ *and the unique solution of (34) is* $x^2(t) = (a_0 + \overline{a}_0)^{t-t_0}x_0$, $t \geq t_0$. *Invoking* c_2) *we may infer that* $|x^2(t)|^2 \leq \delta^{t-t_0}|x_0|^2$ *for all* $t \geq t_0 \in \mathbb{Z}_+$, *where* $\delta \in (0,1)$. *In this case, the inequality (26) becomes:*

$$\mathbb{E}[|x^1(t)|^2] + \mathbb{E}[|x^2(t)|^2] \leq \delta^{t-t_0}|x_0|^2, \ \ \forall t \geq t_0.$$

The implication $(ii) \Rightarrow (i)$ *from Proposition 2 applied in this particular case, allows us to conclude that the MFSLS (32) is MF-ESMS. The companion discrete-time stochastic linear system of the form (12), associated to (32) is:*

$$\begin{aligned} x^1(t+1) &= a_0 x^1(t) + \sum_{k=1}^{r} a_k x^1(t) w_k(t), t \geq t_0 \\ x^1(t_0) &= x_0^1 \end{aligned} \tag{35}$$

$$\begin{aligned} x^2(t+1) &= (a_0 + \overline{a}_0)x^2(t), t \geq t_0 \\ x^2(t_0) &= x_0^2 \end{aligned} \tag{36}$$

where $x_0^k \in \mathcal{X}_{t_0}$, $k = 1, 2$. *Hence, the companion DTSLS associated to the MFSLS (32) is ESMS if and only if the discrete-time stochastic linear equations (35) and (36) are simultaneously ESMS.*

Employing a Lyapunov-type criterion for the characterization of the ESMS property of a discrete-time stochastic linear equation (see section 3.5.2 from

[6]) we obtain that the DTSLS (35) is ESMS if and only if its coefficients are satisfying the condition $\sum_{k=0}^{r} a_k^2 < 1$. *This is in contradiction to* c_3). *Thus, we may conclude that the MFSLS (32) is MF-ESMS but its companion discrete-time stochastic linear system is not ESMS. This contradicts Theorem 2.1 from [13] where is asserted that the MF-ESMS property of a MFSLS is equivalent to the ESMS property of its companion discrete-time standard stochastic linear system.*

Let us consider the analogous for the MFSLS (32) of the operator $\mathcal{L}$ *defined in (24) from [15]. This is* $\mathcal{L}\mathbf{X} = \begin{pmatrix} (a_0 + \overline{a}_0)^2 \overline{x} & 0 \\ 0 & \sum_{k=0}^{r} a_k^2 x \end{pmatrix}$ *for all* $\mathbf{X} = \begin{pmatrix} \overline{x} & 0 \\ 0 & x \end{pmatrix}$. *One sees that* $\lambda = \sum_{k=0}^{r} a_k^2$ *is the eigenvalue of the operator* $\mathcal{L}$ *corresponding to the eigenvector* $\mathbf{X} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

Bearing in mind the constraint c_3) *we deduce that* $\lambda \geq 1$. *Hence, the spectral radius of the operator* $\mathcal{L}$ *satisfies* $\rho(\mathcal{L}) \geq 1$. *So, the MFSLS (32) is MF-ESMS, but the spectral radius of the Lyapunov-type operator defined by its companion standard discrete-time stochastic linear system is not strictly less than 1. This shows that the equivalence stated in Theorem 3.1 from [15] is not always true.*

In the next definition, we introduce a new kind of mean-field exponential stability in mean-square.

Definition 4. *We say that the MFSLS (10) is mean-field strong exponentially stable in mean-square (MF-SESMS) if its companion discrete-time SLS (12) is ESMS in the sense of Definition 3.*

Remark 5. *a) From Remark 4 and Definition 4, we may infer that the MF-SESMS property of the MFSLS (10) implies its MF-ESMS property. The Example 1 and Definition 4 shows that the converse implication is not always true.*

b) Bearing in mind that for the characterization of the ESMS property of a discrete-time stochastic linear system there exist a lot of spectral criteria and Lyapunov-type criteria, we may infer via a) that these criteria provide a set of useful sufficient conditions for characterization of the MF-ESMS property of a MFSLS.

Employing the special structure of the matrix coefficients of the DTSLS described in (7), we shall derive in the next section, a set of necessary and

sufficient conditions for the MF-SESMS property of the MFSLS (10) of lower dimension than those obtained applying directly the already existing results from the literature (see [6], [17]).

4 Criteria for MF-SESMS

The main result of this section is in the next theorem.

Theorem 1. *Under the assumption* **H1**) *the following are equivalent:*

(i) the MFSLS (10) is MF-SESMS;

(ii) the discrete-time stochastic linear equation (DTSLE):

$$x(t+1) = (A_0(t) + \sum_{k=1}^{r} A_k(t)w_k(t))x(t) \tag{37}$$

(obtained from (10a) taking $\overline{A}_k(t) = 0$, $t \in \mathbb{Z}$, $0 \leq k \leq r$*) is ESMS and the discrete-time deterministic linear equation (DTDLE):*

$$x(t+1) = (A_0(t) + \overline{A}_0(t))x(t), \quad t \in \mathbb{Z} \tag{38}$$

(obtained from (10) taking $A_k(t) = \overline{A}_k(t) = 0$, $t \in \mathbb{Z}$, $1 \leq k \leq r$*) is exponentially stable;*

(iii) there exist θ*-periodic matrix valued sequences* $\{x(t)\}_{t\in\mathbb{Z}} \subset \mathcal{S}_n^+$, $\{\overline{X}(t)\}_{t\in\mathbb{Z}} \subset \mathcal{S}_n^+$ *which are satisfying the backward time-varying linear matrix inequalities (LMIs):*

$$\begin{gathered} X(t) - \sum_{k=0}^{r} A_k^\top(t)X(t+1)A_k(t) > 0 \\ X(t) > 0, \end{gathered} \tag{39}$$

$t \in \mathbb{Z}$,

$$\begin{gathered} \overline{X}(t) - (A_0(t) + \overline{A}_0(t))^\top \overline{X}(t+1)(A_0(t) + \overline{A}_0(t)) > 0 \\ \overline{X}(t) > 0, \end{gathered} \tag{40}$$

$t \in \mathbb{Z}$;

(iv) there exist θ*-periodic matrix valued sequences* $\{Y(t)\}_{t\in\mathbb{Z}} \subset \mathcal{S}_n^+$, $\{\overline{Y}(t)\}_{t\in\mathbb{Z}} \subset \mathcal{S}_n^+$ *solving the following forward time-varying LMIs:*

$$\begin{gathered} Y(t+1) - \sum_{k=0}^{r} A_k(t)Y(t)A_k^\top(t) > 0 \\ Y(t) > 0, \end{gathered} \tag{41}$$

$t \in \mathbb{Z}$

$$\begin{gathered} \overline{Y}(t+1) - (A_0(t) + \overline{A}_0^{(}t))\overline{Y}(t)(A_0(t) + \overline{A}_0(t))^\top > 0 \\ \overline{Y}(t) > 0, \end{gathered} \tag{42}$$

$t \in \mathbb{Z}$.

Proof. Let us assume that the MFSLS (10) is MF-SESMS. This means that the solutions $\begin{pmatrix} x^1(\cdot; t_0, \mathbf{x}_0) \\ x^2(\cdot; t_o, \mathbf{x}_0) \end{pmatrix}$ of the IVP (12) have a decay rate of the form (31) for all initial time $t_0 \in \mathbb{Z}_+$ and initial state $\mathbf{x}_0 = \begin{pmatrix} x_0^1 \\ x_0^2 \end{pmatrix}$ where $x_0^k \in \mathcal{X}_{t_0}$, $k = 1, 2$.

In the particular case $\mathbf{x}_0 = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$ with arbitrary $x_0 \in \mathbb{R}^n$ we have $\begin{pmatrix} x^1(\cdot; t_0, \mathbf{x}_0) \\ x^2(\cdot; t_0, \mathbf{x}_0) \end{pmatrix} = \begin{pmatrix} x^1(\cdot; t_0, x_0) \\ 0 \end{pmatrix}$ where $x^1(\cdot; t_0, x_0)$ is the solution of the DTSLE (37) corresponding to the initial pair $(t_0, x_0) \in \mathbb{Z}_+ \times \mathbb{R}^n$.
In this case, (31) gives:

$$\mathbb{E}[|x^1(t; t_0, x_0)|^2] \leq \beta_1 \delta_1^{t-t_0} |x_0|^2 \tag{43}$$

$\forall t \geq t_0 \in \mathbb{Z}_+$, $x_0 \in \mathbb{R}^n$, $\beta_1 \geq 1$, $\delta_1 \in (0, 1)$ are not depending upon t, t_0, x_0.
On the other hand, in the particular case when $\mathbf{x}_0 = \begin{pmatrix} 0 \\ x_0 \end{pmatrix}$ with arbitrary $x_0 \in \mathbb{R}^n$, we have:
$\begin{pmatrix} x^1(\cdot; t_0, \mathbf{x}_0) \\ x^2(\cdot; t_0, \mathbf{x}_0) \end{pmatrix} = \begin{pmatrix} x^1(\cdot; t_0, \mathbf{x}_0) \\ x^2(\cdot; t_0, x_0) \end{pmatrix}$ where $x^2(\cdot; t_0, x_0)$ is the solution of the DTDLE (38) corresponding to the initial pair $(t_0, x_0) \in \mathbb{Z}_+ \times \mathbb{R}^n$.
In this case, (31) becomes:

$$\mathbb{E}[|x^1(t; t_0, \mathbf{x}_0)|^2] + \mathbb{E}[|x^2(t; t_0, x_0)|^2] \leq \beta_1 \delta_1^{t-t_0} |x_0|^2$$

for all $t \geq t_0 \in \mathbb{Z}_+$, $x_0 \in \mathbb{R}^n$. Hence,

$$\mathbb{E}[|x^2(t; t_0, x_0)|^2] \leq \beta_1 \delta_1^{t-t_0} |x_0|^2 \tag{44}$$

for all $t \geq t_0 \in \mathbb{Z}_+$, $x_0 \in \mathbb{R}^n$. Bearing in mind (43) and (44), we may infer that (ii) holds if (i) is true. Further, we show that the implication $(ii) \Rightarrow (i)$

holds. Based on (7) we rewrite (12a) as:

$$x^1(t+1) = A_0(t)x^1(t) + \sum_{k=1}^{r}(A_k(t)x^1(t) + f_k(t))w_k(t) \tag{45a}$$

$$x^2(t+1) = (A_0(t) + \overline{A}_0(t))x^2(t) \tag{45b}$$

$$x^1(t_0) = x_0^1, \quad x^2(t_0) = x_0^2, \tag{45c}$$

where $x_0^k \in \mathcal{X}_{t_0}$, $k = 1, 2$, and

$$f_k(t) \triangleq (A_k(t) + \overline{A}_k(t))x^2(t) \tag{46}$$

$t \in \mathbb{Z}_+$, $1 \leq k \leq r$. If (ii) holds, then, (43) and (44) hold too.
In this case, the solutions of (45b) satisfy

$$\mathbb{E}[|x^2(t; t_0, x_0^2)|^2] \leq \tilde{\beta}\tilde{\delta}^{t-t_0}\mathbb{E}[|x_0^2|^2] \tag{47}$$

for all $t \geq t_0 \in \mathbb{Z}_+$, $x_0^2 \in \mathcal{X}_{t_0}$.

Combining (46) and (47) we deduce that the forcing terms from (45a) are satisfying

$$\mathbb{E}[|f_k(t)|^2] \leq \gamma_k\tilde{\delta}^{t-t_0}\mathbb{E}[|x_0^2|^2] \tag{48}$$

for all $t \geq t_0 \in \mathbb{Z}_+$, where $\gamma_k = \tilde{\beta}\sup\limits_{t\in\mathbb{Z}}\|A_k(t) + \overline{A}_k(t)\|$ and $\tilde{\delta} \in (0, 1)$.
Bearing in mind (43), we may infer via Corollary 3.9 (iii) from [6] applied in the case of non-homogeneous DTSLE (45a) that its solutions are satisfying:

$$\lim_{t\to\infty}\mathbb{E}[|x^1(t; t_0, x_0^1)|^2] = 0 \tag{49}$$

for any initial time $t_0 \in \mathbb{Z}_+$ and the initial state $x_0^1 \in \mathcal{X}_{t_0}$. Employing (48) and (49) we may conclude that the solutions of the IVP (12) are satisfying

$$\lim_{t\to\infty}(\mathbb{E}[|x^1(t; t_0, \mathbf{x}_0)|^2] + \mathbb{E}[|x^2(t; t_0, \mathbf{x}_0)|^2] = 0$$

for all $t_0 \in \mathbb{Z}_+$ and any $\mathbf{x}_0 = \begin{pmatrix} x_0^1 \\ x_0^2 \end{pmatrix}$ with $x_0^k \in \mathcal{X}_{t_0}$, $k = 1, 2$.

Since (12a) is a discrete-time linear equation with periodic coefficients, we may deduce via the equivalence $(i) \Leftrightarrow (v)$ from Theorem 3.10 in [6], applied in the case of the equation (12a), that its solutions are satisfying (31). According to Definition 4 we may conclude that the MFSLS (10) is MF-SESMS. Thus the validity of the implication $(ii) \Leftrightarrow (i)$ is confirmed. The equivalences $(ii) \Leftrightarrow (iii)$ and $(ii) \Leftrightarrow (iv)$ follow applying Theorem 3.11 from [6] in the case of the DTSLE (37) and DTDLE (38), respectively. Thus the proof is complete.

Remark 6. *The Lyapunov-type criteria (39)-(40) and (41)-(42), respectively, are based on the solutions of some time-varying LMIs with periodic coefficients with lower dimension than those obtained applying directly, for example, Theorem 3.11 from [6] in the case of the DTSLE (12a). Based on Remark 5 we may infer that (39)-(40) and (41)-(42) provide sets of sufficient conditions to test if a MFSLS is MF-ESMS.*

The next proposition shows that the conclusion of the Example 1 is still valid also in the time-varying context.

Proposition 3. *We assume: a) the assumption* **H1**) *holds;*

b) $A_k(t) + \overline{A}_k(t) = 0$, $t \in \mathbb{Z}$, $1 \leq k \leq r$;

c) the DTDLS (38) is exponentially stable, but the DTSLDE (37) is not ESMS.

Under these conditions, the MFSLS (10) is MF-ESMS but it is not MF-SESMS.

Proof. According to Proposition 1, without loss of generality, we may assume that $x_0 \in \mathbb{R}^n$ because $\mathbb{R}^n \subset \mathcal{X}_{t_0}$ for any $t_0 \in \mathbb{Z}_+$. Under the assumption *b*), the IVP (12) becomes:

$$\begin{gathered} x^1(t+1) = A_0(t)x^1(t) + \sum_{k=1}^{r} A_k(t)x^1(t)w_k(t), \quad t \geq t_0, \\ x^1(t_0) = 0, \end{gathered} \tag{50}$$

$$\begin{gathered} x^2(t+1) = (A_0(t) + \overline{A}_0(t))x^2(t), \quad t \geq t_0 \\ x^2(t_0) = x_0. \end{gathered} \tag{51}$$

Based on the unicity of the solution of an IVP, we may conclude that the unique solution of the IVP (50) is $x^1(t; t_0, x_0) = 0$ a.s. $t \geq t_0$ and the unique solution of the IVP (51) satisfies

$$\mathbb{E}[|x^2(t; t_0, x_0)|^2] \leq \beta\delta^{t-t_0}|x_0|^2. \tag{52}$$

In this case, (52) gives

$$\mathbb{E}[|x^1(t; t_0, x_0)|^2] + \mathbb{E}[|x^2(t; t_0, x_0)|^2] \leq \beta\delta^{t-t_0}|x_0|^2 \tag{53}$$

for all $t \geq t_0 \geq 0$, $x_0 \in \mathbb{R}^n$.

Since, $\mathbb{E}[|x(t; t_0, x_0)|^2] = \mathbb{E}[|x^1(t; t_0, x_0)|^2] + \mathbb{E}[|x^2(t; t_0, x_0)|^2]$ we obtain from (53) that

$$\mathbb{E}[|x(t; t_0, x_0)|^2] \leq \beta\delta^{t-t_0}|x_0|^2$$

for all $t \geq t_0 \in \mathbb{Z}_+$, $x_0 \in \mathbb{R}^n$, where $\beta \geq 1$, $\delta \in (0,1)$ are not depending upon t, t_0, x_0. Hence, the implication $(ii) \Rightarrow (i)$ from Proposition 2 allows us to conclude that under the considered assumptions the MFSLS (10) is MF-ESMS. It can not be MF-SESMS because the implication $(i) \Rightarrow (ii)$ from Theorem 1 would imply that the DTSLDE (37) would be ESMS, which contradicts the assumption c). Thus the proof ends.

5 The mean-fields stabilizability in mean-square

In this section we shall use the results derived in the previous sections to characterize the MF stabilizability in mean-square of the MFCSLS of the form (1).

Remark 7. *Based on Proposition 1 applied in the case of the closed-loop MFSLS (14), we may infer that, without loss of generality, in order to define the MF stabilizability in mean-square property of the MFCSLS (1) we may assume that its initial states x_0 are deterministic vectors.*

Taking the expectation in (14) we obtain:

$$\mathbb{E}[x(t+1)] = [A_0(t) + \overline{A}_0(t) + (B_0(t) + \overline{B}_0(t))\overline{F}(t)]\mathbb{E}[x(t)] \tag{54a}$$

$$\mathbb{E}[x(t_0)] = \mathbb{E}[x_0] \tag{54b}$$

for any initial time $t_0 \in \mathbb{R}^n$ and any initial state $x_0 \in \mathcal{X}_{t_0}$. To this end, we took into account that based of Lemma 1 applied in the case of the MFSLS (14), we may deduce that the random vectors $x(t)$ and $w(t)$ are mutually independent. Then, subtracting (54) from (14) we obtain:

$$\begin{aligned} &x(t+1) - \mathbb{E}[x(t+1)] = (A_0(t) + B_0(t)F(t))(x(t) - \mathbb{E}[x(t)]) + \\ &\sum_{k=1}^{r}[(A_k(t) + B_k(t)F(t))(x(t) - \mathbb{E}[x(t)]) + (A_k(t) + \overline{A}_k(t) + \\ &(B_k(t) + \overline{B}_k(t))\overline{F}(t))\mathbb{E}[x(t)]]w_k(t), \quad t \geq t_0 \end{aligned} \tag{55a}$$

$$x(t_0) - \mathbb{E}[x(t_0)] = x_0 - \mathbb{E}[x_0]. \tag{55b}$$

Employing the notation (4), we obtain from (54) and (55), the following

IVP:

$$\begin{pmatrix} x^1(t+1) \\ x^2(t+1) \end{pmatrix} = (\mathbb{A}_0(t) + \mathbb{B}_0(t)\mathbb{F}(t)) \begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} + \sum_{k=1}^{r} (\mathbb{A}_k(t) + \mathbb{B}_k(t)\mathbb{F}(t)) \begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} w_k(t), \quad t \geq t_0, \tag{56a}$$

$$\begin{pmatrix} x^1(t_0) \\ x^2(t_0) \end{pmatrix} = \begin{pmatrix} x_0 - \mathbb{E}[x_0] \\ \mathbb{E}[x_0] \end{pmatrix} \tag{56b}$$

for all $t_0 \in \mathbb{Z}_+$ and $x_0 \in \mathcal{X}_{t_0}$, where $\mathbb{A}_k(t), \mathbb{B}_k(t)$, $0 \leq k \leq r$ were defined in (7), (8) and

$$\mathbb{F}(t) = \begin{pmatrix} F(t) & 0 \\ 0 & \overline{\mathbb{F}}(t) \end{pmatrix} \tag{57}$$

$t \in \mathbb{Z}$. Applying Proposition 2 in the case of the pair formed by the IVP (14) and the IVP (56), we obtain via Definition 2, the following equivalence whose proof is omitted.

Corollary 1. *Under the assumption* **H1**) *the following are equivalent:*

(i) the MFCSLS (1) is MF stabilizable in mean-square;

(ii) there exist $\mathbb{R}^{m\times n}$ *valued* θ*-periodic matrix sequences* $\{F(t)\}_{t\in\mathbb{Z}}$, $\{\tilde{F}(t)\}_{t\in\mathbb{Z}}$ *with the properties that for each initial time* $t_0 \in \mathbb{Z}_+$ *and any initial state* $x_0 \in \mathcal{X}_{t_0}$, *the solutions* $\begin{pmatrix} x^1(\cdot;t_0,x_0) \\ x^2(\cdot;t_0,x_0) \end{pmatrix}$ *of the corresponding IVP (56) have a decay of the form:*

$$\mathbb{E}[|x^1(t;t_0,x_0)|^2] + \mathbb{E}[|x^2(t;t_0,x_0)|^2] \leq \beta^1 \hat{\delta}^{t-t_0} (\mathbb{E}[|x_0 - \mathbb{E}[x_0]|^2] + |\mathbb{E}[x_0]|^2) \tag{58}$$

for all $t \geq t_0$, *where* $\hat{\beta} \geq 1$, $\hat{\delta} \in (0,1)$ *are constants which are not depending upon* t, t_0, x_0.

In order to compare the mean-field stabilizability in mean-square property with a kind of stochastic stabilizability property of the DTCSLS (9) we introduce the following definition:

Definition 5. *We say that the DTCSLS (9) is* **stochastic stabilizable in mean-square by a decentralized linear state feedback** *if there exists a control of the form*

$$\mathbf{u}(t) = \mathbb{F}(t)\mathbf{x}(t) = \begin{pmatrix} F(t)x^1(t) & 0 \\ 0 & \overline{F}(t)x^2(t) \end{pmatrix} \tag{59}$$

with the properties: a) $\{F(t)\}_{t\in\mathbb{Z}}$, $\{\overline{F}(t)\}_{t\in\mathbb{Z}}$ *are* θ*-periodic* $\mathbb{R}^{m\times n}$ *matrix valued sequences; b) the closed-loop system*

$$\mathbf{x}(t+1)=(\mathbb{A}_0(t)+\mathbb{B}_0(t)\mathbb{F}(t))\mathbf{x}(t)+\sum_{k=1}^{r}(\mathbb{A}_k(t)+\mathbb{B}_k(t)\mathbb{F}(t))\mathbf{x}(t)w_k(t), \quad (60a)$$

$$\mathbf{x}(t_0)=\begin{pmatrix} x_0^1 \\ x_0^2 \end{pmatrix}\in\mathcal{X}_{t_0}\times\mathcal{X}_{t_0} \quad (60b)$$

(obtained plugging the control (59) to the DTCSLS (9)) is ESMS.

Remark 8. *For a given* $\mathbb{F}(t)$ *as in (57), the IVP (56) is the special case of the IVP (60) for* $x_0^1\leftarrow x_0-\mathbb{E}[x_0]$ *and* $x_0^2\leftarrow\mathbb{E}[x_0]$*. Hence, if the DTCSLS (9) is stochastic stabilizable in mean-square by a decentralized linear state feedback of the form (59), then the solutions of the IVP (56) obtained when in (6) is used the control (59), have a decay of the form (58).*

Corollary 2. *Under the assumption* **H1**) *the MFCSLS (1) is MF stabilizable in mean-square if its companion DTCSLS (9) is stochastic stabilizable in mean-square by a decentralized linear state feedback.*

Proof follows immediately combining Remark 8 with the implication $(ii)\Rightarrow(i)$ from Corollary 1.

The next example shows that the converse implication of the one from Corollary 2 is not always true.

Example 2. *We consider the 1-dimensional discrete-time time-invariant MFCSLS:*

$$\begin{aligned} x(t+1)&=a_0x(t)+\overline{a}_0\mathbb{E}[x(t)]+b_0u(t)+\overline{b}_0\mathbb{E}[u(t)]+\\ &\quad+(a_1x(t)+\overline{a}_1\mathbb{E}[x(t)]+b_1u(t)+\overline{b}_1\mathbb{E}[u(t)])w_1(t), t\geq t_0 \quad (61a)\\ x(t_0)&=x_0\in\mathbb{R} \quad (61b)\end{aligned}$$

where $a_k,\overline{a}_k,b_k,\overline{b}_k$ *are real numbers subject to the constraints:*

c_1) $a_1+\overline{a}_1=b_1+\overline{b}_1=0$, $b_0+\overline{b}_0\neq 0$;

c_2) $a_0b_0+a_1b_1=0$;

c_3) $a_0^2+a_1^2>1$.

In (61), $\{w_1(t)\}_{t\geq 0}$ *is a sequence of independent random variables with zero mean and variance 1. We show that under these conditions, the MFCSLS (61) is MF stabilizable in mean-square but its companion DTCSLS if type (9) is not stochastic stabilizable in mean-square by a decentralized linear*

state feedback. Based on Remark 7, we may infer that, without loss of generality, we can assume that the initial state x_0 is deterministic. Let $\delta \in (0,1)$ fixed and $\overline{f}$ defined from:

$$a_0 + \overline{a}_0 + (b_0 + \overline{b}_0)\overline{f} = \delta. \tag{62}$$

Let $f \in \mathbb{R}$ be arbitrary but fixed. We show that the closed-loop system obtained by coupling the control

$$u(t) = fx(t) + (\overline{f} - f)\mathbb{E}[x(t)] \tag{63}$$

$t \geq t_0$, to the MFCSLS (61) is MF-ESMS, which means that the MFCSLS (61) is MF stabilizable in mean-square.
The IVP of the form (56) obtained by coupling the control (63) to the MFCSLS (61) subject to the constraints c_1) is

$$\begin{aligned} x^1(t+1) &= (a_0 + b_0 f)x^1(t) + (a_1 + b_1 f)x^1(t)w_1(t), \quad t \geq t_0 \\ x^1(t_0) &= 0, \end{aligned} \tag{64}$$

$$\begin{aligned} x^2(t+1) &= (a_0 + \overline{a}_0 + (b_0 + \overline{b}_0)\overline{f})x^2(t), \quad t \geq t_0 \\ x^2(t_0) &= x_0. \end{aligned} \tag{65}$$

From the unicity of the solution of an IVP, we have that the unique solution of (64) is $x^1(t) = 0$ a.s. $t \geq t_0 \geq 0$. From (62) and (65) we obtain that

$$x^2(t) = \delta^{t-t_0}x_0, \quad t \geq t_0 \geq 0. \tag{66}$$

Hence, (66) gives

$$\mathbb{E}[|x^1(t;t_0,x_0)|^2] + \mathbb{E}[|x^2(t;t_0,x_0)|^2] = \delta^{t-t_0}|x_0|^2,$$

$t \geq t_0 \geq 0$, $x_0 \in \mathbb{R}$. Bearing in mind that $\delta \in (0,1)$ we may conclude via $(ii) \Rightarrow (i)$ from Corollary 1 applied in the case of the MFCSLS (61) that this system is MF stabilizable in mean-square. Further, we show that the corresponding companion of the form (9) associated to (61) is not stochastic stabilizable in mean-square by a decentralized linear state feedback, in the sense of the Definition 5.

Let

$$\begin{pmatrix} u^1(t) \\ u^2(t) \end{pmatrix} = \begin{pmatrix} f & 0 \\ 0 & \overline{f} \end{pmatrix} \begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} \tag{67}$$

be an arbitrary control of the form (59) where $f, \overline{f}$ are unknown real numbers which should be determined such that the corresponding closed-loop system to be ESMS.

The IVP of the form (60) associated to the MFCSLS (61) and to the control (67) is:

$$\begin{aligned} x^1(t+1) &= (a_0 + b_0 f)x^1(t) + (a_1 + b_1 f)x^1(t)w_1(t), \quad t \geq t_0 \\ x^1(t_0) &= x_0^1, \end{aligned} \tag{68}$$

$$\begin{aligned} x^2(t+1) &= (a_0 + \overline{a}_0 + (b_0 + \overline{b}_0)\overline{f})x^2(t), \quad t \geq t_0 \\ x^2(t_0) &= x_0^2. \end{aligned} \tag{69}$$

According to Definition 5, the DTCSLS of type (2.9) associated to the MFCSLS (61) is stochastic stabilizable in mean-square by a decentralized linear state feedback if the closed-loop DTCSLSs (68), (69) are simultaneously ESMS. Employing a Lyapunov type criteria we may infer that the DTCSLS (68) is ESMS if and only if

$$(a_0 + b_0 f)^2 + (a_1 + b_1 f)^2 - 1 < 0. \tag{70}$$

Bearing in mind the constraints c_2) we deduce that (70) is equivalent to

$$a_0^2 + a_1^2 - 1 + (b_0^2 + b_1^2)f^2 < 0. \tag{71}$$

Employing c_3) we may conclude that does not exist $f \in \mathbb{R}$ for which (71) to be satisfied. Thus, we have shown that the companion DTCSLS of the form (9) associated to the MFCSLS (61) is not stochastic stabilizable in mean-square by a decentralized linear state feedback.

Motivated by the implication stated in Corollary 2, we introduce a new kind of stabilizability in mean-square of the MFCSLS (1).

Definition 6. *We say that the MFCSLS (1) is MF strong stabilizable in mean-square if its companion DTCSLS (9) is stochastic stabilizable in mean-square by a decentralized linear state feedback.*

Remark 9. *From Corollary 2 and Definition 6, we may infer that the MF strong stabilizability in mean-square of the MFCSLS (1) implies its MF stabilizability in mean-square. From Example 2 and Definition 6 it follows that the converse implication is not always true. Hence, the necessary and sufficient conditions for MF strong stabilizability in mean-square of the MFCSLS (54) serves as sufficient conditions for the MF stabilizability in mean-square of this system.*

Theorem 2. *Under the assumption* **H1**), *the following are equivalent: (i) the MFCSLS (1) is MF strong stabilizable in mean-square;*

(ii) the discrete-time controlled stochastic linear system (DTCSLS):

$$x(t+1) = A_0(t)x(t) + B_0(t)u(t) + \sum_{k=1}^{r}(A_k(t)x(t) + B_k(t)u(t))w_k(t), \quad t \geq 0 \tag{72}$$

(obtained from (1) ignoring the mean-field part) is stochastic stabilizable in mean-square and the discrete-time controlled deterministic linear system (DTCDLS):

$$x(t+1) = (A_0(t) + \overline{A}_0(t))x(t) + (B_0(t) + \overline{B}_0(t))u(t) \tag{73}$$

(obtained from (1) taking $A_k(t) = \overline{A}_k(t) = 0$, $B_k(t) = \overline{B}_k(t) = 0$, $1 \leq k \leq r$, $t \in \mathbb{Z}$*) is stabilizable by linear state feedback.*

Proof. To prove the implication $(i) \Rightarrow (ii)$, let us remark that if (i) holds, there exist a control law of the form (59) with the property that the solutions $\begin{pmatrix} x^1(\cdot;t_0,\mathbf{x}_0) \\ x^2(\cdot;t_0,\mathbf{x}_0) \end{pmatrix}$ of the corresponding closed-loop system (60) have a decay of the form:

$$\mathbb{E}[|x^1(t;t_0,\mathbf{x}_0)|^2] + \mathbb{E}[|x^2(t;t_0,\mathbf{x}_0)|^2] \leq \beta\delta^{t-t_0}(\mathbb{E}[|x_0^1|^2] + \mathbb{E}[|x_0^2|^2]) \tag{74}$$

for all $t \geq t_0 \geq 0$, $\mathbf{x}_0 = \begin{pmatrix} (x_0^1)^\top & (x_0^2)^\top \end{pmatrix}^\top \in \mathcal{X}_{t_0} \times \mathcal{X}_{t_0}$, where $\beta \geq 1$, $\delta \in (0,1)$ are constants not depending upon t, t_0, x_0^k. Based on the unicity of the solution of an IVP we may infer that if $\mathbf{x}_0 \leftarrow \hat{\mathbf{x}}_0 = \begin{pmatrix} x_0^\top & 0^\top \end{pmatrix}^\top$ where $x_0 \in \mathbb{R}^n$, we have $\begin{pmatrix} x^1(t;t_0,\mathbf{x}_0) \\ x^2(t;t_0,\mathbf{x}_0 \end{pmatrix} = \begin{pmatrix} x^1(t;t_0,x_0) \\ 0 \end{pmatrix}$ where $t \to x^1(t;t_0,x_0)$ is the solution of the closed-loop DTSLS:

$$x^1(t+1) = (A_0(t)+B_0(t)F(t))x^1(t)+\sum_{k=1}^{r}(A_k(t)+B_k(t)F(t))x^1(t)w_k(t) \tag{75}$$

corresponding to the initial pair $(t_0, x_0) \in \mathbb{Z}_+ \times \mathbb{R}^n$. In this case (74) becomes:

$$\mathbb{E}[|x^1(t;t_0,x_0)|^2] \leq \beta\delta^{t-t_0}|x_0|^2.$$

This means that the DTSLS (75) is ESMS, or equivalently, the DTCSLS (72) is stochastic stabilizable in the sense of Definition **??** from [6]. On the other hand, if $\mathbf{x}_0 \leftarrow \check{\mathbf{x}}_0 = \begin{pmatrix} 0^\top & x_0^\top \end{pmatrix}^\top$ where $x_0 \in \mathbb{R}^n$, we have that

$\begin{pmatrix} x^1(t;t_0,\mathbf{x}_0) \\ x^2(t;t_0,\mathbf{x}_0) \end{pmatrix} = \begin{pmatrix} x^1(t;t_0,\check{\mathbf{x}}_0) \\ x^2(t;t_0,x_0) \end{pmatrix}$ where $t \to x^2(t;t_0,x_0)$ is the solution of the closed-loop deterministic linear equation:

$$x^2(t+1) = [A_0(t) + \overline{A}_0(t) + (B_0(t) + \overline{B}_0(t))\overline{F}(t)]x^2(t) \tag{76}$$

$t \geq t_0$ corresponding to the initial pair $(t_0, x_0) \in \mathbb{Z}_+ \times \mathbb{R}^n$. Employing (74) we obtain that in this case we have:

$$|x(t;t_0,x_0)|^2 \leq \beta\delta^{t-t_0}|x_0|^2$$

for all $t \geq t_0 \geq 0$, $x_0 \in \mathbb{R}^n$. This means that the discrete-time linear equation (76) is exponentially stable, or equivalent, the DTCDLE (73) is stabilizable. Thus we have shown that (ii) holds if (i) is true. The implication $(ii) \Rightarrow (i)$ may be proved in a similar way to the proof of the implication $(ii) \Rightarrow (i)$ from Theorem 1 replacing the DTSLS (37) by the DTSLS (75) and the DTDLE (38) by the closed-loop discrete-time equation (76). The details are omitted. Thus the proof ends.

References

[1] S. Aberkane and V. Drăgan, An addendum to the problem of zero-sum LQ stochastic mean-field dynamic games, *Automatica* 153 (2023), 111007.

[2] N.U. Ahmed and X. Ding. Controlled McKean–Vlasov equations, *Commun. Appl. Anal.* 5 (2001), 183-206.

[3] N.U. Ahmed, Nonlinear diffusion governed by McKean–Vlasov equation on Hilbert space and optimal control, *SIAM J. Control Optim.* 46 (2007), 356-378.

[4] D. Andersson and B. Djehiche, A maximum principle for SDEs of mean-field type, *Appl. Math. Optim.* 63 (2011), 341-356.

[5] R. Buckdahn, B. Djehiche and J. Li, A general maximum principle for SDEs of mean-field type, *Appl. Math. Optim.* 64 (2011), 197-216.

[6] V. Drăgan, T. Morozan and A.M. Stoica, *Mathematical Methods in Robust Control of Discrete-Time Linear Stochastic Systems*, Springer, 2010.

[7] V. Drăgan, S. Aberkane, On the exponential stability of mean-field stochastic linear differential equations, *Automatica*, accepted for publication October 2026.

[8] M. Huang, R.P. Malhame and P.E. Caines, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, *Commun. Inf. Syst.* 6 (2006), 221-251.

[9] J. Huang, X. Li and J. Yong, A linear-quadratic optimal control problem for mean-field stochastic differential equations in infinite-horizon, *Math. Control Relat. Fields* 5 (2015), 97-139.

[10] M. Kac. Foundations of kinetic theory, In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, III: (1956) 171-197.

[11] J.M. Lasry and P.L. Lions, Mean field games, *Jpn. J. Math.* 2 (2007), 229-260.

[12] H.P. McKean. Propagation of chaos for a class of non-linear parabolic equations, *Lect. Ser. Differ. Equations* 7 (1967), 41-57.

[13] H. Ma, Y. Cui and Z. Wang, Spectral criteria to stability and observability of mean-field stochastic periodic systems, *Automatica* 142 (2022), 110354.

[14] T. Meyer-Brandis, B. Oksendal and X.Y. Zhou, A mean-field stochastic maximum principle via Malliavin calculus, *Stochastics* 84 (2012), 643-666.

[15] Y.H. Ni, R. Elliott and X. Li, Discrete-time mean-field Stochastic linear–quadratic optimal control problems, II: Infinite horizon case, *Automatica* 57 (2015), 65-77.

[16] J.Y. Park, P. Balasubramaniam and Y.H. Kang, Controllability of McKean–Vlasov stochastic integrodifferential evolution equation in Hilbert spaces, *Numer. Funct. Anal. Optim.* 29 (2008), 1328-1346.

[17] W. Zhang, H. Zhang and B. Chen, Detectability, stabilizability, Lyapunov type theorems of linear discrete time-varying stochastic systems, with multiplicative noise, *Int. J. Control* 90 (2017), 2490-2507.

Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 18, No. 4/2026

ON EKELAND VARIATIONAL PRINCIPLE FOR PREORDERED UNIFORM SPACE*

Lucas Fresse† Viorica V. Motreanu‡

DOI 10.56082/annalsarscimath.2026.4.365

Abstract

We introduce a notion of preordered uniform space which extends the classical notion of uniform space and uniformizable topological space. We establish two versions of Ekeland variational principle for the present setting. The first version is for a real-valued submonotone map in a general preordered uniform space while the second version is for a vector-valued map but within an at most countable uniform structure.

Keywords: uniform space, preordered space, submonotone map, variational principle.

MSC: 58E30, 06A75, 54E15, 49J52.

1 Introduction

Uniform topological spaces appear in general topology in relation with Tychonoff's separation axiom ($T_{3\frac{1}{2}}$). Specifically, a topological space X is said to satisfy ($T_{3\frac{1}{2}}$) if for every $x \in X$ and every closed subset $C \subset X$ with $x \notin C$, there is a continuous function $f : X \to [0, 1]$ such that $f(x) = 0$ and

*Accepted for publication on October 18, 2026

†lucas.fresse@univ-lorraine.fr, Université de Lorraine, Institut Élie Cartan, 54506 Vandoeuvre-lès-Nancy, France

‡vmotreanu@gmail.com, Lycée Varoquaux, 10 rue Jean Moulin, 54510 Tomblaine, France

$f|_C \equiv 1$. It is shown in [1] that this axiom is equivalent to the property of X to be uniformizable, i.e., the topology of X can be induced by a so-called *uniform structure*. A uniform structure consists of a family of écarts, also called pseudometrics. Endowed with a uniform structure, X is called a *uniform space*. We refer to [1] for the precise definition of uniform structure and uniform space, or to Section 3.2 below.

The setting of the present paper is in fact more general: we consider a set P endowed with a *preorder* $\leq$ (a binary relation which is reflexive and transitive). We first extend the notion of uniform structure and uniform space to the case of a preordered set $(P, \leq)$, by defining a notion of *preordered uniform space* $(P, \leq, \mathcal{D})$ parametrized by a family $\mathcal{D}$ of $\leq$*-pseudometrics* (see Section 3.2). When the preorder $\leq$ is trivial (i.e., such that $x \leq y$ for all $x, y \in P$), we retrieve the classical notion of uniform space.

Instead of focussing on purely topological properties, our aim in this paper is to do variational analysis on preordered uniform spaces. Our main results are two versions of the classical Ekeland variational principle (which is usually formulated for metric spaces) valid for a preordered uniform space $(P, \leq, \mathcal{D})$:

- The first result (Theorem 1) is a version of Ekeland variational principle for an arbitrary complete preordered uniform space $(P, \leq, \mathcal{D})$ and concerns a map $f : P \to \mathbb{R} \cup \{+\infty\}$ with a submonotonicity property (it is not assumed continuous and not even lower semicontinuous).

- The second result (Theorem 2) is a version of Ekeland's principle for a family of maps $\mathcal{F} = \{f_d\}_{d \in \mathcal{D}}$ with $f_d : P \to \mathbb{R} \cup \{+\infty\}$, instead of just a map f, but with the assumption that the family $\mathcal{D}$ of $\leq$-pseudometrics is (at most) countable.

For showing these results, the main ingredient is a version of Ekeland variational principle over preordered monoids; see [3]. For the first theorem (Theorem 1), we apply this result to the monoid $\mathbb{K} = \mathbb{R} \cup \{+\infty\}$; for the second theorem (Theorem 2), we use the monoid $\mathbb{K} = (\mathbb{R} \cup \{+\infty\})^{\mathbb{N}}$. In order to be able to apply the general result in [3], we rely on a thorough analysis of preordered uniform spaces and their properties: key lemmas in this respect are Lemmas 1 (b), 4, and 5.

The paper is organized as follows. In Section 2, we introduce a general notion of preordered topological space. In Section 3, we first introduce the notions of $\leq$-pseudometric and preordered pseudometric space (Section 3.1); then we introduce the more general notion of preordered uniform space

(Section 3.2). Our main results are stated in Section 4 and the proofs are given in Section 5.

2 Preordered topological space

Throughout this paper, we fix a nonempty set P endowed with a preorder $\leq$, i.e., a binary relation on P which is reflexive and transitive (but not necessarily antisymmetric).

A sequence $(x_n) \subset P$ is said to be $\leq$-*ascending* if it satisfies $x_n \leq x_{n+1}$ for all n. Moreover, we say that x is an *upper bound* of a sequence $(x_n) \subset P$ if $x_n \leq x$ for all n.

Assume that the set P is equipped with a topology τ. We associate to this topology a second topology $\tilde{\tau}_{\leq}$ which takes the preorder $\leq$ into account and is defined in the following way. Hereafter, the notation $x_n \xrightarrow{\tau} x$ means that x is a limit of (x_n) with respect to τ (the limit is a priori not unique).

- A subset $A \subset P$ is said to be $(\leq, \tau)$-*closed* if

$$L_{(\leq,\tau)}(A) := \{x \in P : \exists (x_n) \subset A \leq\text{-ascending},\ x_n \xrightarrow{\tau} x\} = A$$

  (note that the inclusion $A \subset L_{(\leq,\tau)}(A)$ is always valid).

- We say that $U \subset P$ is $(\leq, \tau)$-*open* if $P \setminus U$ is $(\leq, \tau)$-closed.

Then it is straightforward to check that

$$\tilde{\tau}_{\leq} := \{U \subset P : U \text{ is } (\leq, \tau)\text{-open}\}$$

is indeed a topology on P. We call $\tilde{\tau}_{\leq}$ the $\leq$-*ascending extension* of τ.

Lemma 1. (a) $\tau \subset \tilde{\tau}_{\leq}$.

(b) *Let $(x_n) \subset P$ be a $\leq$-ascending sequence and let $x \in P$. Then*

$$x_n \xrightarrow{\tilde{\tau}_{\leq}} x \quad \Leftrightarrow \quad x_n \xrightarrow{\tau} x.$$

Proof. (a) Let $A \subset P$ be a closed subset for the topology τ and let us show that A is also closed for the topology $\tilde{\tau}_{\leq}$, i.e., that A is $(\leq, \tau)$-closed. So let $(x_n) \subset A$ be a $\leq$-ascending sequence and let $x \in P$ be such that $x_n \xrightarrow{\tau} x$. The latter fact, combined with the assumption that A is closed for the topology τ, already ensures that $x \in A$, and the claim is established.

(b) The implication $\Rightarrow$ easily follows from part (a). Let us prove the reversed implication $\Leftarrow$. Arguing by contradiction, assume that there exists $U \in \tilde{\tau}_{\leq}$ such that $x \in U$ and

$$\forall n_0 \in \mathbb{N},\ \exists n \geq n_0,\ x_n \notin U.$$

This yields a subsequence (x_{n_k}) of (x_n) such that $x_{n_k} \notin U$ for all k. Note that (x_{n_k}) is a $\leq$-ascending sequence (since (x_n) is $\leq$-ascending) and $x_{n_k} \xrightarrow{\tau} x$ (since $x_n \xrightarrow{\tau} x$). The subset $A := P \setminus U$ is $(\leq, \tau)$-closed (since $U \in \tilde{\tau}_{\leq}$) and $x_{n_k} \in A$ for all k. Whence $x \in A = P \setminus U$, a contradiction. □

Remark 1. (a) In Lemma 1 (b), the assumption that (x_n) is $\leq$-ascending is not needed for the implication $\Rightarrow$, but is required for the reversed implication $\Leftarrow$, as shown by the following example. Let $P = \mathbb{R}$ be endowed with the standard topology τ and standard order $\leq$. Let $(x_n)_{n\geq 1}$ be the sequence given by $x_n = \frac{1}{n}$ and let $x = 0$. Thus $x_n \xrightarrow{\tau} x$. However, (x_n) does not converge to x with respect to $\tilde{\tau}_{\leq}$. Indeed, it is easy to see that $(0, +\infty)$ is $(\leq, \tau)$-closed, hence $(-\infty, 0]$ is an open neighborhood of $x = 0$ in $\tilde{\tau}_{\leq}$ which contains no term of the sequence (x_n).

(b) If we define $\sigma = \tilde{\tau}_{\leq}$, then Lemma 1 implies that $\tilde{\sigma}_{\leq} = \sigma\,(= \tilde{\tau}_{\leq})$. Indeed, Lemma 1 (a) already yields the inclusion $\sigma \subset \tilde{\sigma}_{\leq}$. For the reversed inclusion, let $A \subset P$ be $(\leq, \sigma)$-closed and let us show that A is $(\leq, \tau)$-closed. So let $(x_n) \subset A$ be $\leq$-ascending and $x \in P$ such that $x_n \xrightarrow{\tau} x$. By Lemma 1 (b) we get $x_n \xrightarrow{\sigma} x$, whence $x \in A$ (since A is $(\leq, \sigma)$-closed), and the argument is complete.

(c) A subset $A \subset P$ is *sequentially closed* with respect to τ if for every sequence $(x_n) \subset A$ and $x \in P$ such that $x_n \xrightarrow{\tau} x$, we have $x \in A$. We say that $U \subset P$ is *sequentially open* with respect to τ if $P \setminus U$ is sequentially closed. The collection $\tilde{\tau}$ of all sequentially open subsets is a topology on P which contains τ and differs from τ in general. The equality $\tilde{\tau} = \tau$ holds if every $x \in P$ has a countable local basis in the topology τ, i.e., the topological space (P, τ) satisfies the axiom of countability (C1).

The inclusion of Lemma 1 (a) can be refined as the chain of inclusions $\tau \subset \tilde{\tau} \subset \tilde{\tau}_{\leq}$. If $\leq$ is the *trivial preorder* on P, i.e., such that $x \leq y$ for all $x, y \in P$, then $\tilde{\tau}_{\leq} = \tilde{\tau}$.

Remark 2. It can be noted that the $\leq$-ascending extension preserves inclusions between topologies, in the following sense. If τ and τ' are two topologies on P such that $\tau \subset \tau'$, then we obtain that $\tilde{\tau}_{\leq} \subset \tilde{\tau}'_{\leq}$. Indeed, if $A \subset P$ is $(\leq, \tau)$-closed, then whenever $(x_n) \subset A$ is a $\leq$-ascending sequence such that $x_n \xrightarrow{\tau'} x \in P$, then the inclusion $\tau \subset \tau'$ yields $x_n \xrightarrow{\tau} x$, whence

$x \in A$ (using that A is $(\leq, \tau)$-closed), and finally we have shown that A is $(\leq, \tau')$-closed.

We say that the preorder $\leq$ is *self-closed* for the topology τ if the set

$$P(\geq x) := \{y \in P : \ x \leq y\}$$

is closed (in τ) for all $x \in P$.

Lemma 2. (a) *If $\leq$ is self-closed for the topology τ, then it is self-closed for the topology $\tilde{\tau}_{\leq}$.*

(b) *The following conditions are equivalent:*

(i) *$\leq$ is self-closed for the topology $\tilde{\tau}_{\leq}$;*

(ii) *for every $\leq$-ascending sequence $(z_n) \subset P$ and every $z \in P$ such that $z_n \overset{\tau}{\to} z$ we have $z_n \leq z$ for all n (i.e., any limit of a $\leq$-ascending sequence is an upper bound of this sequence).*

Proof. (a) This follows from Lemma 1 (a).

(b) First we show the implication (ii) $\Rightarrow$ (i). Let $x \in P$ and let us show that $P(\geq x)$ is closed for the topology $\tilde{\tau}_{\leq}$, i.e., is $(\leq, \tau)$-closed. Let $(z_n) \subset P(\geq x)$ be a $\leq$-ascending sequence such that $z_n \overset{\tau}{\to} z \in P$. For any fixed n, condition (ii) yields $z_n \leq z$, while we know that $x \leq z_n$ (since $z_n \in P(\geq x)$), whence $z \in P(\geq x)$, which establishes (i).

Next let us show the implication (i) $\Rightarrow$ (ii). Let $(z_n) \subset P$ be a $\leq$-ascending sequence such that $z_n \overset{\tau}{\to} z$. For every n, the sequence $(z_m)_{m \geq n}$ is $\leq$-ascending, contained in $P(\geq z_n)$, and has limit z. Since $P(\geq z_n)$ is $(\leq, \tau)$-closed (by virtue of (i)), we obtain that $z \in P(\geq z_n)$; whence $z_n \leq z$ for all n, and the lemma is established. □

Remark 3. (a) Clearly, $\leq$ is self-closed for $\tilde{\tau}_{\leq}$ whenever the subset $P(\geq x)$ is sequentially closed with respect to the topology τ for all $x \in P$, i.e., $\leq$ is self-closed for the topology $\tilde{\tau}$ (see Remark 1 (c)). The latter condition is a priori weaker than requiring that $\leq$ is self-closed with respect to τ, though both notions coincide if for instance (P, τ) satisfies the axiom of countability (C1).

(b) Another sufficient condition for having that $\leq$ is self-closed with respect to $\tilde{\tau}_{\leq}$ is that $P(\leq x) := \{y \in P : \ y \leq x\} \in \tau$ for all $x \in P$. Indeed, if $(x_n) \subset P$ is a $\leq$-ascending sequence such that $x_n \overset{\tau}{\to} x$, then we obtain that $x_n \in P(\leq x)$ for all large enough n, and in fact $x_n \leq x$ for all n, since (x_n) is $\leq$-ascending. The desired conclusion now comes from Lemma 2 (b).

(c) The implication in Lemma 2 (a) is not an equivalence even if (P,τ) satisfies axiom (C1) (see part (a)). Take for instance $P=[0,1]\cup\{2\}$ endowed with the standard metric topology τ and the total order $\leq$ whose restriction to $(0,1]$ is the standard total order and such that $0<2<x$ for all $x\in(0,1]$. Then $\leq$ is not self-closed with respect to τ since $P(\geq 2)=(0,1]\cup\{2\}$ is not closed for the metric topology. However, $P(\geq 2)$ is $(\leq,\tau)$-closed, as well as $P(\geq x)$ for all $x\in[0,1]$, hence $\leq$ is self-closed with respect to $\tilde{\tau}_{\leq}$.

3 Preordered uniform space

In this section, we still consider a nonempty set P endowed with a preorder $\leq$. First, we apply the definitions and results of Section 2 to the case where the topology of P is induced by a pseudometric – in fact, a $\leq$-pseudometric. Then we replace the single $\leq$-pseudometric by a family of $\leq$-pseudometrics.

3.1 Preordered pseudometric space

We call *$\leq$-pseudometric on P* a map

$$d : P^2_{\leq} := \{(x,y)\in P\times P : x\leq y\} \to [0,+\infty]$$

such that

- $d(x,x)=0$ for all $x\in P$,
- $d(x,z)\leq d(x,y)+d(y,z)$ for all $x,y,z\in P$ such that $x\leq y\leq z$.

Then, we say that $(P,\leq,d)$ is a *preordered pseudometric space*.

For every $\eta\in\mathbb{R}^*_+$ and $x\in P$, we set

$$P^d_\eta(\leq x)=\{y\in P : y\leq x,\ d(y,x)<\eta\}.$$

The subsets $P^d_\eta(\leq x)$, for all η and x, form the basis of a topology τ_d on P.

Given a sequence $(x_n)\subset P$, we write $x_n \overset{d}{\to} x$ if x is a limit (possibly not unique) of (x_n) with respect to the topology τ_d. Note that this implies that $x_n\in P^d_{\eta_0}(\leq x)$ for all n large enough, for any fixed $\eta_0\in\mathbb{R}^*_+$, whence in particular $x_n\leq x$ for all n large. If (x_n) is $\leq$-ascending, then it follows that $x_n\leq x$ for all n, i.e., any limit of (x_n) in the topology τ_d is also an upper bound of this sequence.

The $\leq$-pseudometric d allows us to define two notions of distance from a point $x \in P$ to a subset $A \subset P$, namely

$$\begin{aligned} &\mathrm{dist}_d(\leq x, A) := \inf\{d(y,x) : y \in A,\ y \leq x\} \\ \text{and}\quad &\mathrm{dist}_d(\geq x, A) := \inf\{d(x,y) : y \in A,\ y \geq x\}. \end{aligned}$$

The former one is more closely related to the topology τ_d, as pointed out by the following lemma.

Lemma 3. *The closure of a subset $A \subset P$ with respect to τ_d is*

$$S_{(\leq,d)}(A) := \{x \in P : \mathrm{dist}_d(\leq x, A) = 0\}.$$

Proof. Let $\overline{A}^d$ denote the closure of A with respect to τ_d and first let us show that $\overline{A}^d \subset S_{(\leq,d)}(A)$. Let $x \in \overline{A}^d$. For every $\eta \in \mathbb{R}_+^*$, the set $P_\eta^d(\leq x)$ is an open neighborhood of x with respect to τ_d. Hence there exists $y \in A \cap P_\eta^d(\leq x)$, so $y \leq x$ and $d(y,x) < \eta$. We deduce that $\mathrm{dist}_d(\leq x, A) < \eta$ for all $\eta \in \mathbb{R}_+^*$, whence $\mathrm{dist}_d(\leq x, A) = 0$ and so $x \in S_{(\leq,d)}(A)$.

Conversely, let $x \in S_{(\leq,d)}(A)$. For showing that $x \in \overline{A}^d$, it is sufficient to show that any neighborhood V of x with respect to τ_d intersects A. By definition of τ_d, there is $\eta \in \mathbb{R}_+^*$ such that $P_\eta^d(\leq x) \subset V$. On the other hand, since $\mathrm{dist}_d(\leq x, A) = 0$, we can find $y \in A$ such that $y \leq x$ and $d(y,x) < \eta$. Thus $y \in A \cap P_\eta^d(\leq x) \subset A \cap V$, so that $A \cap V \neq \emptyset$, and the proof of the lemma is complete. □

Remark 4. (a) The topology τ_d satisfies the axiom (C1) (see Remark 1 (c)).

(b) In view of Lemma 3, the topology τ_d actually coincides with the $S_\leq$-topology considered in [2, §2] (with the slight difference that in [2], d is assumed to be defined on $P \times P$ instead of $P_\leq^2$, with values in $[0, +\infty)$ instead of $[0, +\infty]$).

Now let us denote by $\tilde{\tau}_{(\leq,d)}$ the $\leq$-ascending extension of τ_d, in the sense of Section 2. The closed subsets for the topology $\tilde{\tau}_{(\leq,d)}$ are the so-called $(\leq, \tau_d)$-closed subsets of Section 2; here, for simplicity, we will call them *$(\leq, d)$-closed subsets*. Also recall that $\tau_d \subset \tilde{\tau}_{(\leq,d)}$, hence $P_\eta^d(\leq x) \in \tilde{\tau}_{(\leq,d)}$ for all $\eta \in \mathbb{R}_+^*$, $x \in P$.

By $\overline{A}^{(\leq,d)}$ we denote the closure of A with respect to the topology $\tilde{\tau}_{(\leq,d)}$, i.e., $\overline{A}^{(\leq,d)}$ is the smallest $(\leq, d)$-closed subset in P which contains A, and A is itself $(\leq, d)$-closed if and only if $\overline{A}^{(\leq,d)} = A$.

In the next lemma, we write

$$L_{(\leq,d)}(A) := L_{(\leq,\tau_d)}(A) = \{x \in P : \exists (x_n) \subset A \leq\text{-ascending with } x_n \overset{d}{\to} x\}$$

(see, again, Section 2) while $S_{(\leq,d)}(A)$ comes from Lemma 3 above. Recall that $P(\leq x) = \{y \in P : y \leq x\}$ and $P(\geq x) = \{y \in P : x \leq y\}$.

Lemma 4. (a) *For every $x \in P$, the subset $P(\leq x)$ is open for the topology τ_d, thus also for $\tilde{\tau}_{(\leq,d)}$; the subset $P(\geq x)$ is closed for the topology τ_d, thus also for $\tilde{\tau}_{(\leq,d)}$. In particular, $\leq$ is self-closed with respect to τ_d and to $\tilde{\tau}_{(\leq,d)}$.*

(b) *For every $A \subset P$, we have the inclusions*

$$A \subset L_{(\leq,d)}(A) \subset \overline{A}^{(\leq,d)} \subset S_{(\leq,d)}(A).$$

(c) *The subset A is closed with respect to $\tilde{\tau}_{(\leq,d)}$ if and only if the first two inclusions of (b) are equalities. The subset A is closed with respect to τ_d if and only if all inclusions of (b) are equalities.*

Proof. (a) Let $x \in P$. Let $y \in P(\leq x)$ and choose any $\eta_0 \in \mathbb{R}_+^*$. Since $\leq$ is transitive, we easily have the inclusion $P^d_{\eta_0}(\leq y) \subset P(\leq x)$, which shows that y is an interior point of $P(\leq x)$, by the definition of the topology τ_d; hence $P(\leq x) \in \tau_d \subset \tilde{\tau}_{(\leq,d)}$.

Next let $y \in P \setminus P(\geq x)$ and, again, we choose any $\eta_0 \in \mathbb{R}_+^*$. For any $z \in P^d_{\eta_0}(\leq y)$, since $x \not\leq y$ and $z \leq y$, we must have $x \not\leq z$, hence $z \in P \setminus P(\geq x)$. This shows that $P(\geq x)$ is closed with respect to τ_d as well as with respect to $\tilde{\tau}_{(\leq,d)}$.

(b) The inclusion $A \subset L_{(\leq,d)}(A)$ is immediate. Since $A \subset \overline{A}^{(\leq,d)}$, we easily have $L_{(\leq,d)}(A) \subset L_{(\leq,d)}(\overline{A}^{(\leq,d)}) = \overline{A}^{(\leq,d)}$, where the last equality comes from the fact that $\overline{A}^{(\leq,d)}$ is $(\leq,d)$-closed. Finally, note that $S_{(\leq,d)}(A)$ contains A and is closed with respect to τ_d (due to Lemma 3) thus with respect to $\tilde{\tau}_{(\leq,d)}$ (since $\tau_d \subset \tilde{\tau}_{(\leq,d)}$); this yields $\overline{A}^{(\leq,d)} \subset S_{(\leq,d)}(A)$ as asserted.

(c) The first assertion comes from (b) and the fact that A is closed with respect to $\tilde{\tau}_{(\leq,d)}$ if and only if $A = \overline{A}^{(\leq,d)}$. The second assertion follows from (b) and the fact that A is closed with respect to τ_d if and only if $A = S_{(\leq,d)}(A)$ (due to Lemma 3). □

Remark 5. (a) All the inclusions in Lemma 4 (b) may be strict in general. Take for instance $P = [0,1] \times [0,1]$ equipped with the preorder $\leq$ given by

$$(x,y) \leq (x',y') \quad \text{if} \quad \begin{cases} x = x' < 1 \ \text{ and } \ y \leq y' < 1, & \text{or} \\ x \leq x' < 1 \ \text{ and } \ y' = 1, & \text{or} \\ x' = 1, \end{cases}$$

and the map $d : P_{\leq}^2 \to [0, +\infty)$ defined as the restriction of the standard metric on P. If $A = [0,1) \times [0,1)$, then it is easy to see that

$$A \subsetneq L_{(\leq,d)}(A) = [0,1) \times [0,1] \subsetneq \overline{A}^{(\leq,d)} = ([0,1) \times [0,1]) \cup \{(1,1)\} \subsetneq S_{(\leq,d)}(A) = P = [0,1] \times [0,1].$$

(b) For $x \in P$, note that

$$L_{(\leq,d)}(\{x\}) = \overline{\{x\}}^{(\leq,d)} = S_{(\leq,d)}(\{x\}) = \{y \in P : x \leq y,\ d(x,y) = 0\}.$$

In particular, a singleton $\{x\}$ is $(\leq, d)$-closed if and only if every $y \in P(\geq x) \setminus \{x\}$ is such that $d(x,y) \neq 0$. Thus the topology $\tilde{\tau}_{(\leq,d)}$ satisfies the separation axiom (T_1) (that is, singletons are $(\leq, d)$-closed in P) if and only if d is *nondegenerate*, that is, $(x,y) \in P_{\leq}^2$ and $x \neq y$ imply $d(x,y) \neq 0$.

Note also that if x is minimal with respect to the preorder $\leq$, which equivalently means that $P(\leq x) = \{x\}$, then Lemma 4 (a) implies that the singleton $\{x\}$ is $(\leq, d)$-open whatever d we consider.

3.2 Preordered uniform space

We call $\leq$-*uniform structure* on $(P, \leq)$ a (nonempty) family $\mathcal{D}$ of $\leq$-pseudometrics on P, and then we call the triple $(P, \leq, \mathcal{D})$ a *preordered uniform space*.

This gives rise to topologies on P:

- Let $\tau_{\mathcal{D}}$ be the coarsest topology on P which contains τ_d for all $d \in \mathcal{D}$. In other words, a basis for this topology is formed by finite intersections $U_1 \cap \ldots \cap U_k$ with $U_1, \ldots, U_k \in \bigcup_{d \in \mathcal{D}} \tau_d$. We will use the notation $x_n \xrightarrow{\mathcal{D}} x$ if $(x_n) \subset P$ converges to x with respect to the topology $\tau_{\mathcal{D}}$.
- Let $\tilde{\tau}_{(\leq,\mathcal{D})}$ denote the $\leq$-ascending extension of $\tau_{\mathcal{D}}$ in the sense of Section 2. In other words, the closed subsets for the topology $\tilde{\tau}_{(\leq,\mathcal{D})}$ are the $(\leq, \tau_{\mathcal{D}})$-closed subsets of P – which we will rather call $(\leq, \mathcal{D})$-closed for simplicity. Recall that $A \subset P$ is $(\leq, \mathcal{D})$-closed if, for every $\leq$-ascending sequence $(x_n) \subset A$ such that $x_n \xrightarrow{\mathcal{D}} x$, we get $x \in A$.

Remark 6. (a) We have the inclusion $\tau_{\mathcal{D}} \subset \tilde{\tau}_{(\leq,\mathcal{D})}$ (see Lemma 1 (a)). Moreover, since $\bigcup_{d \in \mathcal{D}} \tau_d \subset \tau_{\mathcal{D}}$, we have $\bigcup_{d \in \mathcal{D}} \tilde{\tau}_{(\leq,d)} \subset \tilde{\tau}_{(\leq,\mathcal{D})}$ (due to Remark 2).

(b) The preorder $\leq$ is self-closed with respect to the topologies $\tau_{\mathcal{D}}$ and $\tilde{\tau}_{(\leq,\mathcal{D})}$. Indeed, for every $x \in P$, taking any $d \in \mathcal{D}$, we know that $P(\geq x)$ is closed for the topologies τ_d and $\tilde{\tau}_{(\leq,d)}$ (see Lemma 4 (a)). In view of the

inclusions pointed out in part (a) above, we conclude that $P(\geq x)$ is also closed for $\tau_{\mathcal{D}}$ and for $\tilde{\tau}_{(\leq,\mathcal{D})}$.

(c) Part (b) of the present remark tells us that if $(x_n) \subset P$ is $\leq$-ascending and $x_n \overset{\mathcal{D}}{\to} x$, then $x_n \leq x$ for all n (we apply Lemma 2 (b) with $\tau = \tau_{\mathcal{D}}$).

Remark 7. If $\leq$ is the trivial preorder on P (i.e., $x \leq y$ for all $x, y \in P$), then a $\leq$-pseudometric $d : P^2_{\leq} = P \times P \to [0, +\infty]$ is an *écart* in the sense of [1, Chapitre IX]. For a family $\mathcal{D}$ of écarts on P, the pair $(P, \mathcal{D})$ is called a uniform space, and $\tau_{\mathcal{D}}$ is the topology on P considered in [1, Chapitre IX]. Thus we retrieve the classical notion of uniform space when we restrict to the case of the trivial preorder on P.

Lemma 5. *Let $(x_n) \subset P$ be a $\leq$-ascending sequence and let $x \in P$. The following conditions are equivalent:*

(i) $x_n \overset{\mathcal{D}}{\to} x$;

(ii) $x_n \to x$ *with respect to the topology* $\tilde{\tau}_{(\leq,\mathcal{D})}$;

(iii) $x_n \overset{d}{\to} x$ *for all* $d \in \mathcal{D}$.

Proof. The equivalence between (i) and (ii) comes from Lemma 1 (b). The implication (i)⇒(iii) is easily implied by the fact that $\tau_d \subset \tau_{\mathcal{D}}$ for all $d \in \mathcal{D}$. It remains to show (iii)⇒(i). So assume that $x_n \to x$ with respect to τ_d for all $d \in \mathcal{D}$. Let $V \subset P$ be a neighborhood of x for the topology $\tau_{\mathcal{D}}$. By definition of that topology, V contains a subset of the form

$$U_1 \cap \ldots \cap U_k$$

where, for every $i \in \{1, \ldots, k\}$, there is $d_i \in \mathcal{D}$ such that U_i is an open neighborhood of x for the topology τ_{d_i}. For every i, since $x_n \overset{d_i}{\to} x$, there is a rank $n_i \in \mathbb{N}$ such that $x_n \in U_i$ whenever $n \geq n_i$. Setting $n_0 = \max\{n_1, \ldots, n_k\}$, we obtain that $x_n \in U_1 \cap \ldots \cap U_k \subset V$ for all $n \geq n_0$. This shows that $x_n \to x$ with respect to $\tau_{\mathcal{D}}$, and (i) is established. □

We need a notion of completeness for the preordered pseudometric spaces $(P, \leq, d)$ and the preordered uniform space $(P, \leq, \mathcal{D})$. Let us present it now.

- A $\leq$-ascending sequence (x_n) is said to be *d-Cauchy* if, for every $\eta \in \mathbb{R}^*_+$, there is a rank n_0 such that, whenever $n_0 \leq m \leq n$, we have $d(x_m, x_n) < \eta$.

- We say that P is $(\leq, d)$-*complete* if every d-Cauchy $\leq$-ascending sequence $(x_n) \subset P$ is convergent with respect to the topology τ_d (or, equivalently, with respect to $\tilde{\tau}_{(\leq,d)}$; see Lemma 1 (b)).
- A $\leq$-ascending sequence (x_n) is said to be $\mathcal{D}$-*Cauchy* if it is d-Cauchy for all $d \in \mathcal{D}$.
- We say that P is $(\leq, \mathcal{D})$-*complete* if every $\mathcal{D}$-Cauchy $\leq$-ascending sequence is convergent with respect to $\tau_{\mathcal{D}}$ (or equivalently to $\tilde{\tau}_{(\leq,\mathcal{D})}$; see Lemma 5).

Remark 8. (a) The fact that P is $(\leq, \mathcal{D})$-complete does not imply that P is $(\leq, d)$-complete for every or even for some $d \in \mathcal{D}$, as shown by the following example. Let $P = (0, +\infty)$ be endowed with the trivial preorder $\leq$ and let $\mathcal{D} = \{d_1, d_2\}$ with

$$d_1(x, y) = |x - y| \quad \text{and} \quad d_2(x, y) = \Big|\frac{1}{x} - \frac{1}{y}\Big| \quad \text{for all } x, y \in P.$$

Thus P is neither $(\leq, d_1)$-complete nor $(\leq, d_2)$-complete. However, we claim that P is $(\leq, \mathcal{D})$-complete. Indeed, if $(x_n) \subset P$ is $\mathcal{D}$-Cauchy, then there must be $n_0 \in \mathbb{N}$ such that $|x_n - x_{n_0}| < 1$ and $|\frac{1}{x_n} - \frac{1}{x_{n_0}}| < 1$ for all $n \geq n_0$, which easily implies that there must be $a, b \in (0, +\infty)$ such that $a \leq x_n \leq b$ for all n. The metrics d_1, d_2 are equivalent on the interval $[a, b]$, and we obtain that there is $x \in [a, b]$ with $x_n \to x$ with respect to both d_1 and d_2, thus also $x_n \overset{\mathcal{D}}{\to} x$ by Lemma 5.

(b) The fact that P is $(\leq, d)$-complete for all $d \in \mathcal{D}$ is also not sufficient for ensuring that P is $(\leq, \mathcal{D})$-complete. Take for instance $P = [0, 1] \cup \{1'\}$ endowed with the partial order $\leq$ whose restriction to $[0, 1]$ is the standard total order and such that $x \leq 1'$ for all $x \in [0, 1)$, and $1, 1'$ are not comparable. Let $d, d' : P^2_{\leq} \to [0, +\infty)$ be the $\leq$-pseudometrics whose restrictions to $[0, 1)$ are given by the standard metric and such that

$$d(x, 1) = d'(x, 1') = 1 - x, \quad d(x, 1') = d'(x, 1) = 2 - x \quad \text{for all } x \in [0, 1).$$

It is easy to see that every $\leq$-ascending sequence $(x_n) \subset P$ converges with respect to both d and d', and this guarantees that P is both $(\leq, d)$-complete and $(\leq, d')$-complete. However the sequence (x_n) given by $x_n = 1 - \frac{1}{n}$ is $\mathcal{D}$-Cauchy (since it is both d-Cauchy and d'-Cauchy) but not convergent with respect to $\tau_{\mathcal{D}}$: this follows from Lemma 5 combined with the observation that 1 (resp., $1'$) is the only limit of (x_n) with respect to d (resp., d') and $1 \neq 1'$.

(c) Here we assume that $\mathcal{D}$ is directed in the sense that, for all $d_1, d_2 \in \mathcal{D}$ there is $d_3 \in \mathcal{D}$ with $d_1 \leq d_3$ and $d_2 \leq d_3$. We assume also that, for all $d \in \mathcal{D}$, every non-stationary $\leq$-ascending sequence (x_n) has at most one limit with respect to d. Then, in this case, the implication

$$(P \text{ is } (\leq, d)\text{-complete for all } d \in \mathcal{D}) \quad \Rightarrow \quad P \text{ is } (\leq, \mathcal{D})\text{-complete} \tag{1}$$

holds. For showing this, let $(x_n) \subset P$ be a $\mathcal{D}$-Cauchy $\leq$-ascending sequence. If (x_n) is stationary then it is immediately convergent with respect to $\mathcal{D}$, so we can assume that (x_n) is not stationary. We have that (x_n) is d-Cauchy for every $d \in \mathcal{D}$. Since P is $(\leq, d)$-complete, there is $x^d \in P$ with $x_n \overset{d}{\to} x^d$: this limit a priori depends on d, but we claim that $x^{d_1} = x^{d_2}$ for all $d_1, d_2 \in \mathcal{D}$. The assumption that $\mathcal{D}$ is directed yields $d_3 \in \mathcal{D}$ such that $d_1 \leq d_3$ and $d_2 \leq d_3$. We have $x_n \overset{d_3}{\to} x^{d_3}$ thus also $x_n \overset{d_1}{\to} x^{d_3}$ and $x_n \overset{d_2}{\to} x^{d_3}$, due to the previous inequalities. Since (x_n) is not stationary, it has at most one limit with respect to d_1, whence $x^{d_1} = x^{d_3}$. Similarly, $x^{d_2} = x^{d_3}$, and finally $x^{d_1} = x^{d_2}$, our claim is established. That claim yields $x \in P$ such that $x_n \overset{d}{\to} x$ for all $d \in \mathcal{D}$ and, therefore, $x_n \overset{\mathcal{D}}{\to} x$ by virtue of Lemma 5. This proves that P is $(\leq, \mathcal{D})$-complete.

Even under the additional assumptions considered in the present part (c), the converse of (1) is not true. Take for instance $(P, \leq)$, d_1, d_2 as in part (a) above, and set in addition $d_3(x, y) = \max\{d_1(x, y), d_2(x, y)\}$ for all $x, y \in P$, and $\mathcal{D} = \{d_1, d_2, d_3\}$. Thus $\mathcal{D}$ is directed and every sequence $(x_n) \subset P$ has at most one limit with respect to d_1, d_2, d_3 (since these are metrics). With the same argument as in part (a), P is $(\leq, \mathcal{D})$-complete. However, as shown in (a), it is neither $(\leq, d_1)$-complete nor $(\leq, d_2)$-complete.

4 Statement of main results

Our first main result concerns a map $f : P \to \mathbb{R} \cup \{+\infty\}$. Let $\mathrm{dom}(f) = f^{-1}(\mathbb{R})$ denote the *domain* of f.

- We say that f is $(\leq, \mathcal{D})$-*submonotone* if, for every $\leq$-ascending sequence $(x_n) \subset P$ and $x \in P$ such that $x_n \overset{\mathcal{D}}{\to} x$ and $(f(x_n))$ is nonincreasing, we have $f(x) \leq f(x_n)$ for all n.

- We say that f is $\geq$-*locally bounded from below* if, for every $x \in \mathrm{dom}(f)$, there is $\mu \in \mathbb{R}$ such that

$$\forall y \in P, \quad x \leq y \quad \Rightarrow \quad f(y) \geq \mu.$$

Theorem 1. *Let $(P, \leq, \mathcal{D})$ be a preordered uniform space which is $(\leq, \mathcal{D})$-complete. Let $f : P \to \mathbb{R} \cup \{+\infty\}$ be $(\leq, \mathcal{D})$-submonotone and $\geq$-locally bounded from below. Let $x \in \mathrm{dom}(f)$. Then there is $y \in \mathrm{dom}(f)$ such that*

(a) *$x \leq y$;*

(b) *$d(x, y) \leq f(x) - f(y)$ for all $d \in \mathcal{D}$;*

(c) *whenever $z \in \mathrm{dom}(f)$ and $z' \in P$ are such that $y \leq z \leq z'$, we have the implication*

$$\exists d \in \mathcal{D},\ d(z, z') > 0$$
$$\Rightarrow \quad \exists d \in \mathcal{D},\ (d(y, z) > f(y) - f(z) \quad or \quad d(z, z') > f(z) - f(z')).$$

For our second main result, instead of a single map f, we consider a family $\mathcal{F} = \{f_d\}_{d \in \mathcal{D}}$ of maps $f_d : P \to \mathbb{R} \cup \{+\infty\}$.

- We say that $\mathcal{F}$ is $(\leq, \mathcal{D})$-*submonotone* if, for every $\leq$-ascending sequence $(x_n) \subset P$ and $x \in P$ such that $x_n \overset{\mathcal{D}}{\to} x$ and $(f_d(x_n))$ is non-increasing for all $d \in \mathcal{D}$, we have $f_d(x) \leq f_d(x_n)$ for all n and all $d \in \mathcal{D}$.

- We say that $\mathcal{F}$ is $\geq$-*locally bounded from below* if, for every $x \in \mathrm{dom}(\mathcal{F}) := \bigcap_{d \in \mathcal{D}} \mathrm{dom}(f_d)$, there is a family $\{\mu_d\}_{d \in \mathcal{D}} \subset \mathbb{R}$ such that

$$\forall y \in P, \quad x \leq y \quad \Rightarrow \quad (\forall d \in \mathcal{D},\ f_d(y) \geq \mu_d).$$

Theorem 2. *Let $(P, \leq, \mathcal{D})$ be a preordered uniform space which is $(\leq, \mathcal{D})$-complete. Assume that $\mathcal{D}$ is (at most) countable. Let $\mathcal{F} = \{f_d\}_{d \in \mathcal{D}}$ be a family of maps $f_d : P \to \mathbb{R} \cup \{+\infty\}$ which is $(\leq, \mathcal{D})$-submonotone and $\geq$-locally bounded from below. Let $x \in \mathrm{dom}(\mathcal{F})$. Then there is $y \in \mathrm{dom}(\mathcal{F})$ such that*

(a) *$x \leq y$;*

(b) *$d(x, y) \leq f_d(x) - f_d(y)$ for all $d \in \mathcal{D}$;*

(c) *whenever $z \in \mathrm{dom}(\mathcal{F})$ and $z' \in P$ are such that $y \leq z \leq z'$, we have the implication*

$$\exists d \in \mathcal{D},\ d(z, z') > 0$$
$$\Rightarrow \quad \exists d \in \mathcal{D},\ (d(y, z) > f_d(y) - f_d(z) \quad or \quad d(z, z') > f_d(z) - f_d(z')).$$

The proofs of Theorems 1 and 2 are given in Section 5.

5 Proofs

5.1 Key tool: Ekeland variational principle over monoids

The key tool for proving our main results is a version of the Ekeland variational principle over monoids, established in [3]. We give the full statement here for making this paper as self-contained as possible (also, the result presented below is slightly rephrased).

- Let $(\mathbb{K}, +)$ be a commutative monoid, with neutral element 0 and we denote by $\mathbb{K}'$ its subset of invertible elements. The set $\mathbb{K}$ is also equipped with a preorder $\leq$ and a countable collection of subsets $\mathcal{V}$ which is stable under finite intersections and such that

$$\bigcap_{V \in \mathcal{V}} V = \{0\}.$$

  We will write

$$\mathbb{K}_+ := \{\alpha \in \mathbb{K} : 0 \leq \alpha\}, \quad \mathbb{K}'_+ := \mathbb{K}' \cap \mathbb{K}_+.$$

  For $(\alpha_n)_{n\in\mathbb{N}} \subset \mathbb{K}$, we write $\alpha_n \to 0$ if every $V \in \mathcal{V}$ contains α_n for all but finitely many n. We assume the following:

  (A) $\forall \alpha \in \mathbb{K}', \quad \{\beta \in \mathbb{K} : \beta \leq \alpha\} \subset \mathbb{K}'$.

  (B) $\forall \alpha, \beta, \gamma \in \mathbb{K}, \quad \alpha \leq \beta \quad \Rightarrow \quad \alpha + \gamma \leq \beta + \gamma$.

  (C) $\forall V \in \mathcal{V}, \quad \forall \alpha \in \mathbb{K}' \cap V, \quad \{\beta \in \mathbb{K} : 0 \leq \beta \leq \alpha\} \subset V$.

  (D) $\forall V \in \mathcal{V}, \ \forall \alpha \in \mathbb{K}', \ \exists N \geq 1, \ \forall \beta_1, \ldots, \beta_N \in \mathbb{K}'_+ \setminus V, \ \beta_1 + \ldots + \beta_N \not\leq \alpha$.

  (E) $\forall \alpha \in \mathbb{K}, \ \forall (\beta_n)_{n\in\mathbb{N}} \subset \mathbb{K}_+, \ (\alpha \leq \beta_n \ \forall n \text{ and } \beta_n \to 0) \ \Rightarrow \ \alpha \leq 0$.

- Let P be a nonempty set endowed with a preorder $\leq$ and a map

$$\delta : P^2_{\leq} = \{(x, y) \in P \times P : x \leq y\} \to \mathbb{K}_+$$

  such that $\delta(x, x) = 0$ for all x and $\delta(x, z) \leq \delta(x, y) + \delta(y, z)$ whenever $x \leq y \leq z$.

- In this context, a $\leq$-ascending sequence $(x_n) \subset P$ is said to be

  - *δ-convergent to x* if $x_n \leq x$ for all n and $\delta(x_n, x) \to 0$;
  - *δ-Cauchy* if for every $V \in \mathcal{V}$ there is n_0 such that $\delta(x_m, x_n) \in V$ whenever $n_0 \leq m \leq n$.

Finally P is called $(\leq, \delta)$-*complete* if every δ-Cauchy $\leq$-ascending sequence $(x_n) \subset P$ is δ-convergent to some x.

- Given a map $f : P \to \mathbb{K}$,
  - we write $\mathrm{dom}(f) = f^{-1}(\mathbb{K}')$;
  - we say that f is $(\leq, \delta)$-*submonotone* if, for every $\leq$-ascending sequence $(x_n) \subset \mathrm{dom}(f)$ which is δ-convergent to x,

$$(f(x_n) \leq f(x_m)\ \forall m, n \ \text{ with } \ m \leq n) \quad \Rightarrow \quad f(x) \leq f(x_m)\ \forall m;$$

  - we say that f is *locally bounded from below* if, for every $x \in \mathrm{dom}(f)$, there is $\mu \in \mathbb{K}'$ such that

$$\forall y \in P, \quad (x \leq y \quad \text{and} \quad f(y) \leq f(x)) \quad \Rightarrow \quad f(y) \geq \mu.$$

  (The latter notion is a special case of the notion of "locally finitely bounded from below" map considered in [3].)

In this setting, we have (see [3, Theorem 1]):

Theorem 3. *Let $\mathbb{K}$, P be as above, assume that P is $(\leq, \delta)$-complete and let $f : P \to \mathbb{K}$ be $(\leq, \delta)$-submonotone and locally bounded from below. Then, for every $x \in \mathrm{dom}(f)$, there is $y \in \mathrm{dom}(f)$ such that*

(a) $x \leq y$;

(b) $\delta(x, y) \leq f(x) - f(y)$;

(c) *whenever $z, z' \in \mathrm{dom}(f)$ satisfy $y \leq z \leq z'$ and $\delta(z, z') \neq 0$, we have*

$$\delta(y, z) \not\leq f(y) - f(z) \quad \text{or} \quad \delta(z, z') \not\leq f(z) - f(z').$$

Example 1. (a) The typical example of monoid to which the result applies is $\mathbb{K} = \mathbb{R} \cup \{+\infty\}$ endowed with the standard addition $+$ and total order $\leq$, and with the countable collection of subsets $\mathcal{V} := \{(-r, r)\}_{r \in \mathbb{Q}_+^*}$. Then conditions (A)–(E) are easily fulfilled. Note also that $\mathbb{K}' = \mathbb{R}$, $\mathbb{K}_+ = [0, +\infty]$, and $\mathbb{K}'_+ = [0, +\infty)$.

(b) As another example, consider the set of sequences

$$\mathbb{K} = (\mathbb{R} \cup \{+\infty\})^{\mathbb{N}} = \{(\alpha_n)_{n \in \mathbb{N}} : \alpha_n \in \mathbb{R} \cup \{+\infty\}\}$$

which is a commutative monoid for the standard componentwise addition, the neutral element is the zero constant sequence. We endow $\mathbb{K}$ with the partial order given by

$$(\alpha_n) \leq (\beta_n) \quad \text{if} \quad \alpha_n \leq \beta_n \ \forall n \in \mathbb{N}.$$

We consider the countable collection of subsets $\mathcal{V} = \{V_{I,r}\}$ parametrized by pairs (I, r) formed by a nonempty finite subset $I \subset \mathbb{N}$ and a map $r : I \to \mathbb{Q}_+^*$, $i \mapsto r_i$, and with

$$V_{I,r} := \{(\alpha_n) \in \mathbb{K} : \alpha_i \in (-r_i, r_i) \ \forall i \in I\}.$$

Here we have $\mathbb{K}' = \{(\alpha_n) : \alpha_n < +\infty \ \forall n\} = \mathbb{R}^{\mathbb{N}}$, $\mathbb{K}_+ = [0, +\infty]^{\mathbb{N}}$, and $\mathbb{K}'_+ = [0, +\infty)^{\mathbb{N}}$. Again it is straightforward to check that (A)–(C) hold true. For showing (D), let $V = V_{I,r} \in \mathcal{V}$, with $I \subset \mathbb{N}$ nonempty finite and $r : I \to \mathbb{Q}_+^*$. Let $\alpha = (\alpha_n) \in \mathbb{K}' = \mathbb{R}^{\mathbb{N}}$. For every $i \in I$, since $r_i \in \mathbb{Q}_+^*$ and $\alpha_i \in \mathbb{R}$, we can find $N_i \in \mathbb{N}^*$ such that $N_i > \frac{\alpha_i}{r_i}$. We set $N_0 = \max\{N_i : i \in I\}$ and $N = N_0 \operatorname{card}(I)$. Now let $\beta_1, \ldots, \beta_N \in \mathbb{K}'_+ \setminus V$. For $m \in \{1, \ldots, N\}$, we write $\beta_m = (\beta_{m,n})_{n\in\mathbb{N}}$. For every $m \in \{1, \ldots, N\}$, the assumption that $\beta_m \in \mathbb{K}'_+ \setminus V$ yields

$$I_m := \{i \in I : \beta_{m,i} \geq r_i\} \neq \emptyset.$$

Fix any $i_m \in I_m$, so that we obtain a map

$$F : \{1, \ldots, N\} \to I, \ m \mapsto i_m.$$

Since $N = N_0 \operatorname{card}(I)$, there is $i \in I$ with $\operatorname{card}(F^{-1}(i)) \geq N_0$. Since $\beta_{m,i} \geq 0$ for all $m \in \{1, \ldots, N\}$ (because $\beta_m \in \mathbb{K}'_+$), we obtain

$$\sum_{m=1}^{N} \beta_{m,i} \geq \sum_{m\in F^{-1}(i)} \beta_{m,i} \geq N_0 r_i > \alpha_i.$$

Whence $\beta_1 + \ldots + \beta_N \not\leq \alpha$. This shows (D).

For showing (E), let $\alpha \in \mathbb{K}$ and $(\beta_n)_{n\in\mathbb{N}} \subset \mathbb{K}_+$ be such that $\alpha \leq \beta_n$ for all n and $\beta_n \to 0$. We write $\beta_n = (\beta_{n,i})_{i\in\mathbb{N}}$. For every $i \in \mathbb{N}$, we aim to show that $\alpha_i \leq 0$. To this end, let $\varepsilon \in \mathbb{Q}_+^*$, set $I = \{i\}$ and $r : I \to \mathbb{Q}_+^*$, $i \mapsto r_i = \varepsilon$. Since $V := V_{I,r} \in \mathcal{V}$ and $\beta_n \to 0$, there is $n_0 \in \mathbb{N}$ such that $\beta_n \in V$ for all $n \geq n_0$. Using in addition that $\beta_n \in \mathbb{K}_+$, this yields $0 \leq \beta_{n,i} < \varepsilon$ for any $n \geq n_0$ and, therefore, $\alpha_i \leq \beta_{n,i} < \varepsilon$. Since this holds for any $\varepsilon \in \mathbb{Q}_+^*$, we conclude that $\alpha_i \leq 0$ and the argument is complete.

5.2 Proof of Theorem 1

We aim to apply Theorem 3 to the monoid $\mathbb{K} = \mathbb{R} \cup \{+\infty\}$ equipped with the standard addition $+$ and total order $\leq$, as in Example 1 (a) and with $\mathcal{V} = \{(-\varepsilon, \varepsilon)\}_{\varepsilon \in \mathbb{Q}_+^*}$. As explained in Example 1 (a), conditions (A)–(E) are fulfilled. For applying Theorem 3, we also consider our preordered set $(P, \leq)$ endowed with the map

$$\delta : P_{\leq}^2 \to [0, +\infty], \quad (x, y) \mapsto \sup_{d \in \mathcal{D}} d(x, y).$$

We clearly have $\delta(x, x) = 0$ for all $x \in P$, moreover it is straightforward to check that $\delta(x, z) \leq \delta(x, y) + \delta(y, z)$ whenever $x \leq y \leq z$.

For checking the applicability of Theorem 3 to the considered setting, we need to show that

- P is $(\leq, \delta)$-complete;
- f is $(\leq, \delta)$-submonotone.

For checking the first assertion, let $(x_n) \subset P$ be a $\leq$-ascending sequence which is δ-Cauchy. For every $d \in \mathcal{D}$, the fact that $d(x, y) \leq \delta(x, y)$ for all $(x, y) \in P_{\leq}^2$ easily implies that (x_n) is d-Cauchy. Hence (x_n) is $\mathcal{D}$-Cauchy. Knowing that P is $(\leq, \mathcal{D})$-complete, we find $x \in P$ such that $x_n \overset{\mathcal{D}}{\to} x$, in particular $x_n \leq x$ for all n (see Remark 6 (c)). We claim that (x_n) is δ-convergent to x. To see this, fix $V \in \mathcal{V}$, $V = (-\varepsilon, \varepsilon)$ with $\varepsilon \in \mathbb{Q}_+^*$. Using that (x_n) is δ-Cauchy, there is $n_0 \in \mathbb{N}$ such that $\delta(x_m, x_n) < \frac{\varepsilon}{2}$ whenever $n_0 \leq m \leq n$. We claim that

$$\delta(x_m, x) < \varepsilon \quad \text{whenever } m \geq n_0. \tag{2}$$

Indeed, let $m \geq n_0$ and let $d \in \mathcal{D}$. Since $x_n \overset{\mathcal{D}}{\to} x$, by virtue of Lemma 5 we have also $x_n \overset{d}{\to} x$, so there is $n_0' \in \mathbb{N}$ such that $d(x_n, x) < \frac{\varepsilon}{2}$ whenever $n \geq n_0'$. Now we set $n = \max\{m, n_0'\}$ and we obtain

$$d(x_m, x) \leq d(x_m, x_n) + d(x_n, x) \leq \delta(x_m, x_n) + d(x_n, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since this holds for all $d \in \mathcal{D}$, we get (2). The fact that (x_n) is δ-convergent to x is now justified, and we have established that P is $(\leq, \delta)$-complete.

For showing the second assertion, let $(x_n) \subset \mathrm{dom}(f)$ be a $\leq$-ascending sequence, which is δ-convergent to $x \in P$ (in particular, $x_n \leq x$ for all n) and such that $f(x_n) \leq f(x_m)$ for all m, n with $m \leq n$. For every $d \in \mathcal{D}$,

since $d(x,y) \leq \delta(x,y)$ for all $(x,y) \in P^2_{\leq}$, we obtain that $x_n \overset{d}{\to} x$. Whence $x_n \overset{\mathcal{D}}{\to} x$ (see Lemma 5). Using that f is $(\leq, \mathcal{D})$-submonotone, we conclude that $f(x) \leq f(x_m)$ for all m, and the fact that f is $(\leq, \delta)$-submonotone is established.

We are in position to apply Theorem 3, which, given $x \in \mathrm{dom}(f)$, provides us with an element $y \in \mathrm{dom}(f)$ satisfying conditions (a)–(c) therein. Let us deduce that the same element y satisfies conditions (a)–(c) of Theorem 1.

Condition (a) of Theorem 1 coincides with condition (a) of Theorem 3.

Condition (b) of Theorem 3 tells us that $\delta(x,y) \leq f(x) - f(y)$. By definition of δ, this means that $d(x,y) \leq f(x) - f(y)$ for all $d \in \mathcal{D}$, which is exactly condition (b) of Theorem 1.

Finally, let $z \in \mathrm{dom}(f)$ and $z' \in P$ be such that $y \leq z \leq z'$ and assume that $d(z,z') > 0$ for some $d \in \mathcal{D}$. By definition of δ, this yields $\delta(z,z') \neq 0$. If $f(z') = +\infty$ then the inequality $d'(z,z') > f(z) - f(z')$ is immediate for all $d' \in \mathcal{D}$, so we can assume that $z' \in \mathrm{dom}(f)$. Condition (c) of Theorem 3 now implies that $\delta(y,z) > f(y) - f(z)$ or $\delta(z,z') > f(z) - f(z')$ (recall that in our situation the monoid $\mathbb{K} = \mathbb{R} \cup \{+\infty\}$ is endowed with its *total* order $\leq$). Invoking again the definition of δ, these formulas mean that there is $d \in \mathcal{D}$ such that $d(y,z) > f(y) - f(z)$ or there is $d \in \mathcal{D}$ such that $d(z,z') > f(z) - f(z')$, which matches with the conclusion of part (c) of Theorem 1. The proof of Theorem 1 is complete.

5.3 Proof of Theorem 2

We write $\mathcal{D} = \{d_n\}_{n\in\mathbb{N}}$ (that is, we fix a surjective map $\mathbb{N} \to \mathcal{D}$, which is possible since $\mathcal{D}$ is at most countable). Also we write for simplicity $f_n = f_{d_n}$ for all $n \in \mathbb{N}$, so that $\mathcal{F} = \{f_n\}_{n\in\mathbb{N}}$. We consider the monoid $\mathbb{K} = (\mathbb{R} \cup \{+\infty\})^{\mathbb{N}}$ equipped with the addition $+$, the partial order $\leq$, and the set $\mathcal{V}$ as in Example 1 (b). Furthermore, let $\delta : P^2_{\leq} \to \mathbb{K}_+ = [0,+\infty]^{\mathbb{N}}$, $(x,y) \mapsto (d_n(x,y))_{n\in\mathbb{N}}$. Clearly, $\delta(x,x) = 0$ for all $x \in P$, where 0 denotes the neutral element in $\mathbb{K}$ (i.e., the zero constant sequence). Moreover, whenever $x \leq y \leq z$, we have $d_n(x,z) \leq d_n(x,y) + d_n(y,z)$ for all $n \in \mathbb{N}$, whence $\delta(x,z) \leq \delta(x,y) + \delta(y,z)$.

We aim to apply Theorem 3 with $\mathbb{K}$ and δ as above and to the map

$$f : P \to \mathbb{K},\ x \mapsto (f_n(x))_{n\in\mathbb{N}}.$$

We need to check that

- P is $(\leq, \delta)$-complete;

- f is $(\leq, \delta)$-submonotone;
- f is locally bounded from below.

Let us prove the first assertion. Let $(x_n) \subset P$ be a $\leq$-ascending sequence and assume that (x_n) is δ-Cauchy. Let us show that (x_n) is $\mathcal{D}$-Cauchy. To do this, for every $d = d_i \in \mathcal{D}$ with $i \in \mathbb{N}$, we have to show that (x_n) is d-Cauchy. So let $\varepsilon \in \mathbb{Q}_+^*$. Set $I = \{i\}$ and $r : I \to \mathbb{Q}_+^*$, $i \mapsto r_i = \varepsilon$. Considering $V = V_{I,r}$ (see Example 1 (b)), we find $n_0 \in \mathbb{N}$ such that whenever $n_0 \leq m \leq n$, we have $\delta(x_m, x_n) \in V$. The latter inclusion means that $d(x_m, x_n) < \varepsilon$. We obtain that (x_n) is d-Cauchy for all $d \in \mathcal{D}$, and so $\mathcal{D}$-Cauchy as claimed. Knowing that P is $(\leq, \mathcal{D})$-complete, it follows that there is $x \in P$ such that $x_n \overset{\mathcal{D}}{\to} x$, in particular $x_n \leq x$ for all n (see Remark 6 (c)).

We verify that (x_n) is δ-convergent to x and this will conclude the argument for showing that P is $(\leq, \delta)$-complete. Let $V = V_{I,r} \in \mathcal{V}$ with $I \subset \mathbb{N}$ finite and $r : I \to \mathbb{Q}_+^*$. For every $i \in I$, we know that $x_n \overset{d_i}{\to} x$ (see Lemma 5). Hence there is $n_i \in \mathbb{N}$ such that $d_i(x_n, x) < r_i$ whenever $n \geq n_i$. Set $n_0 = \max\{n_i : i \in I\}$. For every $n \geq n_0$, we then obtain that $d_i(x_n, x) < r_i$ for all $i \in I$, whence $\delta(x_n, x) \in V$. Therefore, (x_n) is δ-convergent to x.

Next we show that f is $(\leq, \delta)$-submonotone. To do this, let $(x_n) \subset \mathrm{dom}(f)$ be a $\leq$-ascending sequence which is δ-convergent to $x \in P$ (thus $x_n \leq x$ for all n) and such that $f(x_n) \leq f(x_m)$ for all $m, n \in \mathbb{N}$ with $m \leq n$. By definition of f, the latter inequality tells us that $f_i(x_n) \leq f_i(x_m)$ for all $m, n \in \mathbb{N}$ with $m \leq n$ and all $i \in \mathbb{N}$, i.e., $(f_i(x_n))_{n \in \mathbb{N}}$ is nonincreasing for all $i \in \mathbb{N}$.

Moreover, note that $x_n \overset{\mathcal{D}}{\to} x$. Indeed, by Lemma 5, it suffices to see that $x_n \overset{d_i}{\to} x$ for every $i \in \mathbb{N}$. For seeing this, let $\varepsilon \in \mathbb{Q}_+^*$. Set $I = \{i\}$ and $r : I \to \mathbb{Q}_+^*$, $i \mapsto \varepsilon$. Since (x_n) is δ-convergent to x, there is $n_0 \in \mathbb{N}$ such that $\delta(x_n, x) \in V_{I,r}$ whenever $n \geq n_0$; whence $d_i(x_n, x) < \varepsilon$ whenever $n \geq n_0$. Therefore, $x_n \overset{\mathcal{D}}{\to} x$.

Now we can invoke the assumption that $\mathcal{F}$ is $(\leq, \mathcal{D})$-submonotone, from which we obtain that $f_i(x) \leq f_i(x_m)$ for all $i \in \mathbb{N}$ and all $m \in \mathbb{N}$, hence $f(x) \leq f(x_m)$ for all $m \in \mathbb{N}$. The argument for showing that f is $(\leq, \delta)$-submonotone is complete.

Let us now establish the third assertion saying that f is locally bounded from below. Let $x \in \mathrm{dom}(f)$. This means that $f_i(x) < +\infty$ for all $i \in \mathbb{N}$, i.e., $x \in \bigcap_{i \in \mathbb{N}} \mathrm{dom}(f_i) = \mathrm{dom}(\mathcal{F})$. Using that $\mathcal{F}$ is $\geq$-locally bounded from below, there is a family $\{\mu_i\}_{i \in \mathbb{N}} \subset \mathbb{R}$ such that $\mu_i \leq f_i(y)$ whenever $y \in P$ satisfies $x \leq y$. Set $\mu = (\mu_i)_{i \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}} = \mathbb{K}'$. We obtain that for every $y \in P$ such that $x \leq y$, we have $\mu \leq f(y)$. The argument is complete.

The above assertions allow us to apply Theorem 3. Let $x \in \mathrm{dom}(\mathcal{F}) = \mathrm{dom}(f)$. There exists $y \in \mathrm{dom}(f) = \mathrm{dom}(\mathcal{F})$ fulfilling (a)–(c) of Theorem 3. We claim that the same element y also fulfils (a)–(c) of Theorem 2. Condition (a) is exactly the same for both theorems so we only have to consider conditions (b) and (c).

Theorem 3 (b) tells us that the relation $\delta(x,y) \leq f(x) - f(y)$ holds in the monoid $\mathbb{K}$. This means that $d_i(x,y) \leq f_i(x) - f_i(y)$ for all $i \in \mathbb{N}$, in other words $d(x,y) \leq f_d(x) - f_d(y)$ for all $d \in \mathcal{D}$. Condition (b) of Theorem 2 is established.

Finally, let $z \in \mathrm{dom}(\mathcal{F}) = \mathrm{dom}(f)$ and $z' \in P$ be such that $y \leq z \leq z'$ and $d(z,z') > 0$ for some $d = d_i \in \mathcal{D}$. Since $\delta(z,z') = (d_j(z,z'))_{j\in\mathbb{N}}$, we obtain that $\delta(z,z')$ is different from the zero constant sequence (the neutral element of $\mathbb{K}$). If $z' \notin \mathrm{dom}(f)$, then there is $d' = d_j \in \mathcal{D}$ such that $f_{d'}(z') = f_j(z') = +\infty$, thus $d'(z,z') > f_{d'}(z) - f_{d'}(z')$. Now assume that $z' \in \mathrm{dom}(f)$. By Theorem 3 (c), we then have $\delta(y,z) \not\leq f(y) - f(z)$ or $\delta(z,z') \not\leq f(z) - f(z')$. Due to the definition of the partial order $\leq$ of $\mathbb{K}$, we conclude that there is $d' = d_j \in \mathcal{D}$ such that $d'(y,z) > f_{d'}(y) - f_{d'}(z)$ or there is $d' \in \mathcal{D}$ such that $d'(z,z') > f_{d'}(z) - f_{d'}(z')$. Thus condition (c) of Theorem 2 holds, and the proof of the theorem is complete.

References

[1] N. Bourbaki, *Éléments de Mathématiques. Topologie Générale. Chapitres 5 à 10*. Hermann, Paris, 1974.

[2] L. Fresse and V.V. Motreanu, Nonlinear error bounds for maps on preordered pseudometric spaces, *Ann. Acad. Rom. Sci., Math. Appl.* 17 (2025), 169-191.

[3] L. Fresse and V.V. Motreanu, Ekeland variational principle over preordered monoids, *Ann. Acad. Rom. Sci., Math. Appl.* 18 (2026), 133-151.

Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 18, No. 4/2026

STRUCTURAL STABILITY OF MICROPOLAR FLUID FLOW IN A CURVED PIPE WITH DARCY–LAPWOOD–FORCHHEIMER–BRINKMAN MODEL*

A. Prabal† M. Devakar‡

DOI 10.56082/annalsarscimath.2026.4.385

Abstract

In this article, we address the structural stability of the Darcy–Lapwood–Forchheimer–Brinkman model for a micropolar fluid flowing through a curved pipe. A rigorous analysis is carried out to prove the continuous dependence of solutions of the model on several physical and geometric parameters such as the Reynolds number, curvature ratio, Forchheimer coefficient, and micropolarity parameter.

Keywords: Darcy–Lapwood–Forchheimer–Brinkman model, micropolar fluid, curved pipe, structural stability, continuous dependence.
MSC: 35Q35, 76A05, 76S05, 35B30.

1 Introduction and mathematical model

A central question in mathematical modelling is whether solutions depend continuously on variations in coefficients, boundary conditions, or the gov-

*Accepted for publication on October 20, 2026

†ds20mth001@students.vnit.ac.in, Department of Mathematics, Visvesvaraya National Institute of Technology Nagpur, 440010 India

‡m_devakar@mth.vnit.ac.in, Department of Mathematics, Visvesvaraya National Institute of Technology Nagpur, 440010 India

erning partial differential equations themselves. This continuous dependence of the solutions with respect to the model's structural components is known as structural stability. For real life problems, variations in coefficients of the governing partial differential equations correspond to changes in the constitutive parameters, therefore, establishing the continuous dependence is crucial for validating the mathematical model's real-world applicability. A comprehensive treatment of structural stability and related continuous dependence results can be found in the monographs by Ames and Straughan [1] and Straughan [19]. For Newtonian fluids, structural stability of the flows through porous media with the Darcy and Brinkman model has been extensively investigated by Morro and Straughan [13], Ames and Payne [2], and Payne and Straughan [14–16]. For the Brinkman–Forchheimer model, related stability results of Newtonian fluid flows were established by Payne and Straughan [17], Çelebi et al. [6,7], and Liu [11]. Łukaszewicz [12] established continuous dependence of two-dimensional micropolar fluid flows on the microrotation viscosity coefficient. To the best knowledge of the authors, structural stability of the flows in curved pipe is not yet explored. Therefore, in this paper, we investigate the structural stability of micropolar fluid flow in a curved pipe with Darcy-Lapwood-Forchheimer-Brinkman model on the pertinent physical and geometric parameters. The mathematical model of the flow problem under consideration is presented in Subsection 1.1. *A priori* estimates that are required for establishing structural stability are presented in Section 2. Main theorems establishing the structural stability of the flow model are presented, along with their proofs, in Section 3.

1.1 Mathematical model

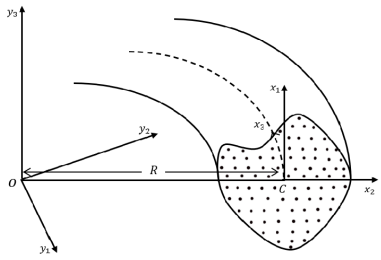


Figure 1: Flow Geometry.

We consider the flow of a micropolar fluid within a stationary curved pipe of arbitrary cross-section of diameter $2a$, which is characterized by a

constant centerline radius $R > a$. A typical section of the curved pipe is shown in Figure 1. We assume that the pipe's cross-section, $\Omega \subset \mathbb{R}^2$, is a bounded domain with Lipschitz boundary $\partial\Omega$. The flow is driven by a constant pressure gradient along the axis of the pipe and it is assumed to be axisymmetric. Under these assumptions, using toroidal coordinate system [3], the partial differential equations governing the flow of micropolar fluid through porous medium with Darcy-Lapwood-Forchheimer-Brinkman model are given by [8–10, 18]

$$\nabla \cdot (H\mathbf{u}) = 0, \tag{1}$$

$$\begin{aligned} &ReH(\mathbf{u}\cdot\nabla)\mathbf{u} - Re\delta u_3^2\hat{a}_2 + Re\delta u_2u_3\hat{a}_3 \\ &= \mathbf{f} - ReH\nabla p + ReG\hat{a}_3 + mH\nabla\times\boldsymbol{\nu} + m\delta\nu_3\hat{a}_1 + (1+m)\nabla\cdot(H\nabla\mathbf{u}) \\ &\quad -(1+m)\frac{\delta^2}{H}u_2\hat{a}_2 - (1+m)\frac{\delta^2}{H}u_3\hat{a}_3 - \sigma^2H\mathbf{u} - Re\sigma C_FH|\mathbf{u}|\mathbf{u}, \end{aligned} \tag{2}$$

$$\begin{aligned} &ReHJ(\mathbf{u}\cdot\nabla)\boldsymbol{\nu} - ReJ\delta u_3\nu_3\hat{a}_2 + ReJ\delta u_3\nu_2\hat{a}_3 \\ &= \mathbf{l} - 2mH\boldsymbol{\nu} + mH\nabla\times\mathbf{u} + m\delta u_3\hat{a}_1 + n_1\nabla\cdot(H\nabla\boldsymbol{\nu}) - n_1\frac{\delta^2}{H}(\nu_2\hat{a}_2+\nu_3\hat{a}_3) \\ &\quad +(n_2+n_3)H\nabla(\nabla\cdot\boldsymbol{\nu}) + (n_2+n_3)\delta\nabla\nu_2 - (n_2+n_3)\frac{\delta^2}{H}\nu_2\hat{a}_2 \end{aligned} \tag{3}$$

in Ω,

where, the velocity vector $\mathbf{u} = (u_1, u_2, u_3)$ and the microrotation vector $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3)$ are independent of x_3, $Re = \rho Ua/\mu$ is the Reynolds number, $\delta = a/R \in [0, 1)$ is the curvature ratio, $\sigma = a/\sqrt{K}$ is the porosity parameter, $H = 1 + \delta x_2$ is the scale factor, G is the constant axial pressure gradient so that the total pressure is $p(x_1, x_2) - G\,x_3$, $m = \frac{k}{\mu}$ is the micropolarity parameter, J is the nondimensional gyration parameter, and n_1, n_2 and n_3 are nondimensional gyro-viscosity parameters. We assume that $\min\{n_1, n_2+n_3\} > 0$. The constants μ and k represents the usual Newtonian viscosity and the microrotation viscosity. ρ is the density of the fluid, K is the permeability of the porous medium and C_F is Forchheimer coefficient. The functions $\mathbf{f}$ and $\mathbf{l}$ are the prescribed dimensionless body force and dimensionless bode couple, respectively, U is the characteristic velocity and $\{\hat{a}_1, \hat{a}_2, \hat{a}_3\}$ is the orthonormal basis of the toroidal coordinate system [3,18]. The differential operators $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, 0\right)$ and $\nabla\cdot = \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}$.

Since the curved pipe is stationary, the no-slip and the hyperstick boundary conditions to be satisfied are, respectively,

$$u_i = 0 \quad \text{and} \quad \nu_i = 0 \quad \text{on } \partial\Omega \quad \text{for } i = 1, 2, 3. \tag{4}$$

2 *A priori* estimates

The authors have established the existence and uniqueness of a solutions to the boundary value problem (BVP) (1)-(4) in [18]. More precisely, it was proved that if $\mathbf{f}, \mathbf{l} \in (L^2(\Omega))^3$ then for sufficiently small values of δ, there exists $(\mathbf{u}, p, \boldsymbol{\nu}) \in \mathcal{X} \times L_0^2(\Omega) \times \left(H_0^1(\Omega)\right)^3$ such that the triplet $(\mathbf{u}, p, \boldsymbol{\nu})$ satisfy (1)-(4) in the sense of distributions, where

$$\mathcal{X} = \left\{ \mathbf{u} = (u_1, u_2, u_3) \in \left(H_0^1(\Omega)\right)^3 : \frac{\partial (H u_1)}{\partial x_1} + \frac{\partial (H u_2)}{\partial x_2} = 0 \right\}.$$

It was also proved that the solution $(\mathbf{u}, p, \boldsymbol{\nu})$ is unique for sufficiently small Re and m. *A priori* estimates for $\mathbf{u}$ and $\boldsymbol{\nu}$, which were presented in [18], can be simplified as

$$\|\mathbf{u}\|_{1,2,\Omega} = \|\nabla \mathbf{u}\|_2 \leq \frac{\mathfrak{C}_1}{(1+m)H_0}\Big(m(1+\delta)\|\boldsymbol{\nu}\|_{1,2,\Omega} + \|\mathbf{F}\|_2 + \|\mathbf{f}\|_2\Big) \equiv E_{\mathbf{u}} \quad (5)$$

and

$$\|\boldsymbol{\nu}\|_{1,2,\Omega} = \|\nabla \boldsymbol{\nu}\|_2 \leq \frac{\mathfrak{C}_2 \left(\frac{m(1+\delta)}{(1+m)H_0}\Big(\|\mathbf{F}\|_2 + \|\mathbf{f}\|_2\Big) + \|\mathbf{l}\|_2\right)}{\Big(n_1 H_0 - \mathfrak{C}_3 \delta (n_2 + n_3)\Big)} \equiv E_{\boldsymbol{\nu}}, \quad (6)$$

where $\mathfrak{C}_1$, $\mathfrak{C}_2$ and $\mathfrak{C}_3$ are positive constants depending only on Ω, $\mathbf{F} = (0, 0, Re\, G)$ and $H_0 = \inf_{\Omega} H$.

To derive an estimate for p, we express equation (2) as

$$Re \nabla p = (1+m)\nabla^2 \mathbf{u} + \mathfrak{F}, \quad (7)$$

in which

$$\mathfrak{F} = \frac{\mathbf{f}}{H} + \frac{\mathbf{F}}{H} - Re(\mathbf{u} \cdot \nabla)\mathbf{u} + \frac{Re\delta}{H}(u_3^2 \hat{a}_2 - u_2 u_3 \hat{a}_3) + \frac{\delta}{H}\frac{\partial \mathbf{u}}{\partial x_2}$$
$$-(1+m)\frac{\delta^2}{H^2}(u_2 \hat{a}_2 + u_3 \hat{a}_3) - \sigma^2 \mathbf{u} - Re\sigma C_F |\mathbf{u}|\mathbf{u} + m\nabla \times \boldsymbol{\nu} + \frac{m\delta}{H}\nu_3 \hat{a}_1. \quad (8)$$

From the expression of $\mathfrak{F}$, it is clear that $\mathfrak{F} \in \left(H^{-1}(\Omega)\right)^3$; therefore, using Lemma IV.1.9 of [4], we get

$$Re\|p\|_2 \leq \mathfrak{C}_4\Big((1+m)\|\mathbf{u}\|_{1,2,\Omega} + \|\mathfrak{F}\|^*\Big), \quad (9)$$

where $\mathfrak{C}_4$ is a positive constant depending only on Ω and $\|\cdot\|^*$ denotes the operator norm in $\left(H^{-1}(\Omega)\right)^3$.
Now, we find $\|\mathfrak{F}\|^*$ to get appropriate estimate for p. For $\boldsymbol{v} \in \left(H_0^1(\Omega)\right)^3$,

$$\langle \mathfrak{F}, \boldsymbol{v} \rangle_{H^{-1},H_0^1} = \int_\Omega \left(\frac{\mathbf{f}}{H} + \frac{\mathbf{F}}{H} - Re(\mathbf{u}\cdot\nabla)\mathbf{u} + \frac{Re\delta}{H}(u_3^2\hat{a}_2 - u_2u_3\hat{a}_3) + \frac{\delta}{H}\frac{\partial \mathbf{u}}{\partial x_2} \right.$$
$$\left. -(1+m)\frac{\delta^2}{H^2}(u_2\hat{a}_2 + u_3\hat{a}_3) - \sigma^2\mathbf{u} - Re\sigma C_F|\mathbf{u}|\mathbf{u} + m\nabla\times\boldsymbol{\nu} + \frac{m\delta}{H}\nu_3\hat{a}_1 \right) \cdot \boldsymbol{v}\, dx \tag{10}$$

Using Sobolev embedding [5] and Hölder's inequality, we obtain

$$\left|\langle \mathfrak{F}, \boldsymbol{v} \rangle_{H^{-1},H_0^1}\right| \le \mathfrak{C}_5 \left[\frac{1}{H_0}\left(\|\mathbf{f}\|_2 + \|\mathbf{F}\|_2\right) + Re\left(1 + \sigma C_F + \frac{\delta}{H_0}\right)\|\nabla\mathbf{u}\|_2^2 \right.$$
$$\left. + \left(\sigma^2 + \frac{\delta}{H_0} + (1+m)\frac{\delta^2}{H_0^2}\right)\|\nabla\mathbf{u}\|_2 + m\left(1 + \frac{\delta}{H_0}\right)\|\nabla\boldsymbol{\nu}\|_2 \right] \|\nabla\boldsymbol{v}\|_2,$$

and hence

$$\|\mathfrak{F}\|^* \le \mathfrak{C}_5 \left[\frac{1}{H_0}\left(\|\mathbf{f}\|_2 + \|\mathbf{F}\|_2\right) + Re\left(1 + \sigma C_F + \frac{\delta}{H_0}\right) E_{\mathbf{u}}^2 \right.$$
$$\left. + \left(\sigma^2 + \frac{\delta}{H_0} + (1+m)\frac{\delta^2}{H_0^2}\right) E_{\mathbf{u}} + m\left(1 + \frac{\delta}{H_0}\right) E_{\boldsymbol{\nu}} \right], \tag{11}$$

where $\mathfrak{C}_5$ is a positive constant depending only on Ω.

Using (11) in (9), we get an estimate for p as

$$\|p\|_2 \le E_p, \tag{12}$$

$$E_p = \mathfrak{C}_6 \left[\frac{1}{H_0 Re}\left(\|\mathbf{f}\|_2 + \|\mathbf{F}\|_2\right) + \left(1 + \sigma C_F + \frac{\delta}{H_0}\right) E_{\mathbf{u}}^2 \right.$$
$$\left. + \frac{1}{Re}\left(\sigma^2 + \frac{\delta}{H_0} + (1+m)\left(1 + \frac{\delta^2}{H_0^2}\right)\right) E_{\mathbf{u}} + \frac{m}{Re}\left(1 + \frac{\delta}{H_0}\right) E_{\boldsymbol{\nu}} \right],$$

where $\mathfrak{C}_6$ is a positive constant depending only on Ω.

Using *a priori* estimates (5), (6) and (12), in the next section, we investigate the structural stability of the flow problem (1)-(4); particularly, we establish the continuous dependence of the solution on the key physical and geometric parameters, namely, the Reynolds number (Re), the curvature ratio (δ), the Forchheimer coefficient (C_F), and the micropolarity parameter (m).

3 Structural stability

We assume that $\mathbf{f}, \mathbf{l} \in \left(L^2(\Omega)\right)^3$ and the values of δ, Re and m are sufficiently small to ensure the existence and uniqueness of the weak solutions of the problem (1)-(4). Throughout this section, we denote $\mathfrak{C}$ as a generic positive constant depending only on Ω.

The main results of the paper are presented in the following four subsections; indeed, Theorems 1, 2, 3 and 4 present the continuous dependence of the solution on parameters Re, δ, C_F and m, respectively.

3.1 Continuous dependence on the Reynolds number (Re)

Let $(\mathbf{u}, p, \boldsymbol{\nu})$ and $(\mathbf{v}, q, \boldsymbol{\eta})$ be the solutions of the BVP (1)-(4) corresponding to the Reynolds numbers Re_1 and Re_2, respectively. Then the difference $(\mathbf{w}, \pi, \boldsymbol{\phi}) := (\mathbf{u} - \mathbf{v}, p - q, \boldsymbol{\nu} - \boldsymbol{\eta})$ satisfies the BVP

$$\nabla \cdot (H\mathbf{w}) = 0, \tag{13}$$

$$\begin{aligned}
&Re_1 H\Big((\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{v}\Big) + \widetilde{Re}H(\mathbf{v}\cdot\nabla)\mathbf{v} - Re_1\delta(u_3+v_3)w_3\hat{a}_2 \\
&-\widetilde{Re}\delta v_3^2\hat{a}_2 + Re_1\delta(w_2u_3 + v_2w_3)\hat{a}_3 + \widetilde{Re}\delta v_2v_3\hat{a}_3 \\
&= \widetilde{Re}\, G\hat{a}_3 - Re_1 H\nabla\pi - \widetilde{Re}H\nabla q + mH\nabla\times\boldsymbol{\phi} + m\delta\phi_3\hat{a}_1 \\
&+(1+m)\nabla\cdot(H\nabla\mathbf{w}) - (1+m)\frac{\delta^2}{H}(w_2\hat{a}_2 + w_3\hat{a}_3) - \sigma^2 H\mathbf{w} \\
&-Re_1\sigma C_F H(|\mathbf{u}|\mathbf{u} - |\mathbf{v}|\mathbf{v}) - \widetilde{Re}\sigma C_F H|\mathbf{v}|\mathbf{v},
\end{aligned} \tag{14}$$

$$\begin{aligned}
&Re_1 HJ\Big((\mathbf{u}\cdot\nabla)\boldsymbol{\phi} + (\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\Big) + \widetilde{Re}HJ(\mathbf{v}\cdot\nabla)\boldsymbol{\eta} - Re_1 J\delta u_3\phi_3\hat{a}_2 \\
&-Re_1 J\delta w_3\eta_3\hat{a}_2 - \widetilde{Re}J\delta v_3\eta_3\hat{a}_2 + Re_1 J\delta u_3\phi_2\hat{a}_3 + Re_1 J\delta w_3\eta_2\hat{a}_3 \\
&+\widetilde{Re}J\delta v_3\eta_2\hat{a}_3 \\
&= -2mH\boldsymbol{\phi} + mH\nabla\times\mathbf{w} + m\delta w_3\hat{a}_1 + n_1\nabla\cdot(H\nabla\boldsymbol{\phi}) - \frac{n_1\delta^2}{H}(\phi_2\hat{a}_2 + \phi_3\hat{a}_3) \\
&+(n_2+n_3)H\nabla(\nabla\cdot\boldsymbol{\phi}) + (n_2+n_3)\delta\nabla\boldsymbol{\phi} - (n_2+n_3)\frac{\delta^2}{H}\phi_2\hat{a}_2
\end{aligned} \tag{15}$$

in Ω,

and

$$\mathbf{w} = \mathbf{0} = \boldsymbol{\phi} \text{ on } \partial\Omega, \tag{16}$$

where $\mathbf{w} = (w_1, w_2, w_3)$, $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)$, $w_i = u_i - v_i$, $\phi_i = \nu_i - \eta_i$ $(i = 1, 2, 3)$ and $\widetilde{Re} = Re_1 - Re_2$.

The main result regarding the continuous dependence on the Reynolds number (Re) is stated in the following theorem.

Theorem 1. *If $(\mathbf{w}, \pi, \boldsymbol{\phi})$ is the solution of (13)-(16), then, for sufficiently small values of Re_1, m and δ, $\mathbf{w} \to \mathbf{0}$ in $\mathcal{X}$ and $\boldsymbol{\phi} \to \mathbf{0}$ in $(H_0^1(\Omega))^3$ as $Re_2 \to Re_1$.*

Proof. Taking the $L^2(\Omega)$-inner product of (14) with $\mathbf{w}$, we get

$$\begin{aligned}
&Re_1 \int_\Omega H\big((\mathbf{u}\cdot\nabla)\mathbf{w}\big)\cdot\mathbf{w}\,dx + Re_1 \int_\Omega H\big((\mathbf{w}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx \\
&+\widetilde{Re}\int_\Omega H\big((\mathbf{v}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx - Re_1\delta\int_\Omega v_3 w_3 w_2\,dx - \widetilde{Re}\delta\int_\Omega v_3^2 w_2\,dx \\
&+Re_1\delta\int_\Omega v_2 w_3^2\,dx + \widetilde{Re}\delta\int_\Omega v_2 v_3 w_3\,dx \\
&= \widetilde{Re}\,G\int_\Omega w_3\,dx - Re_1\int_\Omega H\nabla\pi\cdot\mathbf{w}\,dx - \widetilde{Re}\int_\Omega H\nabla q\cdot\mathbf{w}\,dx \\
&+m\int_\Omega H(\nabla\times\boldsymbol{\phi})\cdot\mathbf{w}\,dx + m\delta\int_\Omega \phi_3 w_1\,dx \\
&+(1+m)\int_\Omega \nabla\cdot(H\nabla\mathbf{w})\cdot\mathbf{w}\,dx - (1+m)\delta^2\int_\Omega \frac{1}{H}(w_2^2+w_3^2)\,dx \\
&-\sigma^2\int_\Omega H\mathbf{w}\cdot\mathbf{w}\,dx - Re_1\sigma C_F\int_\Omega H(|\mathbf{u}|\mathbf{u}-|\mathbf{v}|\mathbf{v})\cdot\mathbf{w}\,dx \\
&-\widetilde{Re}\sigma C_F\int_\Omega H|\mathbf{v}|\mathbf{v}\cdot\mathbf{w}\,dx. \qquad (17)
\end{aligned}$$

Since $(|\mathbf{u}|u_i - |\mathbf{v}|v_i)w_i = \frac{1}{2}(|\mathbf{u}|+|\mathbf{v}|)w_i^2 + \frac{1}{2}(|\mathbf{u}|-|\mathbf{v}|)^2(|\mathbf{u}|+|\mathbf{v}|)$ and H is a positive function, we have $\int_\Omega H(|\mathbf{u}|\mathbf{u}-|\mathbf{v}|\mathbf{v})\cdot\mathbf{w}\,dx \geq 0$. Also, considering the nonnegativity of $\int_\Omega \frac{1}{H}(w_2^2+w_3^2)\,dx$ together with (13) and the fact that $\int_\Omega H\big((\mathbf{u}\cdot\nabla)\mathbf{w}\big)\cdot\mathbf{w}\,dx = 0$, the equality (17) becomes

$$\begin{aligned}
&(1+m)H_0\|\nabla\mathbf{w}\|_2^2 + \sigma^2 H_0\|\mathbf{w}\|_2^2 \leq \left|Re_1\int_\Omega H\big((\mathbf{w}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx\right| \\
&+\left|Re_1\delta\int_\Omega v_3 w_3 w_2\,dx\right| + \left|Re_1\delta\int_\Omega v_2 w_3^2\,dx\right| + \left|\widetilde{Re}\int_\Omega H\big((\mathbf{v}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx\right| \\
&+\left|\widetilde{Re}\delta\int_\Omega v_3^2 w_2\,dx\right| + \left|\widetilde{Re}\delta\int_\Omega v_2 v_3 w_3\,dx\right| + \left|\widetilde{Re}\,G\int_\Omega w_3\,dx\right|
\end{aligned}$$

$$+\left|\widetilde{Re}\sigma C_F\int_\Omega H|\mathbf{v}|\mathbf{v}\cdot\mathbf{w}\,dx\right|+\left|m\int_\Omega H(\nabla\times\boldsymbol{\phi})\cdot\mathbf{w}\,dx\right|+\left|m\delta\int_\Omega \phi_3 w_1\,dx\right|. \tag{18}$$

Using Hölder's inequality, Sobolev inequality and the bound (5), the terms of the right hand side of (18) can be estimated as follows:

$$\begin{aligned}\left|Re_1\int_\Omega H\big((\mathbf{w}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx\right| &\le Re_1H_1\|\mathbf{w}\|_4\|\nabla\mathbf{v}\|_2\|\mathbf{w}\|_4\\ &\le \mathfrak{C}Re_1E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2^2,\end{aligned}$$

$$\begin{aligned}\left|Re_1\delta\int_\Omega v_3w_3w_2\,dx\right| &\le Re_1\delta\|\mathbf{v}\|_2\|\mathbf{w}\|_4^2\\ &\le \mathfrak{C}Re_1\delta E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2^2,\end{aligned}$$

$$\begin{aligned}\left|\widetilde{Re}\int_\Omega H\big((\mathbf{v}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx\right| &\le |\widetilde{Re}|H_1\|\mathbf{v}\|_4\|\nabla\mathbf{v}\|_2\|\mathbf{w}\|_4\\ &\le \mathfrak{C}|\widetilde{Re}|E_{\mathbf{v}}^2\|\nabla\mathbf{w}\|_2,\end{aligned}$$

$$\begin{aligned}\left|\widetilde{Re}\delta\int_\Omega v_3^2w_2\,dx\right| &\le |\widetilde{Re}|\delta\|\mathbf{v}\|_4^2\|\mathbf{w}\|_2\\ &\le \mathfrak{C}|\widetilde{Re}|\delta E_{\mathbf{v}}^2\|\nabla\mathbf{w}\|_2,\end{aligned}$$

$$\begin{aligned}\left|\widetilde{Re}\,G\int_\Omega w_3\,dx\right| &\le |\widetilde{Re}||G||\Omega|\|\mathbf{w}\|_2\\ &\le \mathfrak{C}|\widetilde{Re}|\|\nabla\mathbf{w}\|_2,\end{aligned}$$

$$\begin{aligned}\left|\widetilde{Re}\sigma C_F\int_\Omega H|\mathbf{v}|\mathbf{v}\cdot\mathbf{w}\,dx\right| &\le |\widetilde{Re}|\sigma C_FH_1\||\mathbf{v}|\mathbf{v}\|_2\|\mathbf{w}\|_2\\ &\le \mathfrak{C}|\widetilde{Re}|\sigma C_F\|\nabla\mathbf{v}\|_2^2\|\nabla\mathbf{w}\|_2 \text{ (see [18])}\\ &\le \mathfrak{C}|\widetilde{Re}|\sigma C_FE_{\mathbf{v}}^2\|\nabla\mathbf{w}\|_2,\end{aligned}$$

$$\begin{aligned}\left|m\int_\Omega H(\nabla\times\boldsymbol{\phi})\cdot\mathbf{w}\,dx\right| &\le mH_1\|\nabla\times\boldsymbol{\phi}\|_2\|\mathbf{w}\|_2\\ &\le \mathfrak{C}m\|\nabla\boldsymbol{\phi}\|_2\|\nabla\mathbf{w}\|_2,\end{aligned}$$

and

$$\left|m\delta\int_\Omega \phi_3 w_1\,dx\right| \le m\delta\|\boldsymbol{\phi}\|_2\|\mathbf{w}\|_2 \le \mathfrak{C}m\delta\|\nabla\boldsymbol{\phi}\|_2\|\nabla\mathbf{w}\|_2.$$

The preceding estimates transform the inequality (18) into

$$(1+m)H_0\|\nabla\mathbf{w}\|_2^2 \le \mathfrak{C}\Big\{Re_1(1+\delta)E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2^2 + |\widetilde{Re}|\Big((1+\delta+\sigma C_F)E_{\mathbf{v}}^2+1\Big)\|\nabla\mathbf{w}\|_2 + m(1+\delta)\|\nabla\boldsymbol{\phi}\|_2\|\nabla\mathbf{w}\|_2\Big\}. \tag{19}$$

Taking the $L^2(\Omega)$-inner product of (15) with $\boldsymbol{\phi}$, we get

$$\begin{aligned}
&Re_1J\int_\Omega H\big((\mathbf{u}\cdot\nabla)\boldsymbol{\phi}\big)\cdot\boldsymbol{\phi}\,dx + Re_1J\int_\Omega H\big((\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\big)\cdot\boldsymbol{\phi}\,dx\\
&+\widetilde{Re}J\int_\Omega H\big((\mathbf{v}\cdot\nabla)\boldsymbol{\eta}\big)\cdot\boldsymbol{\phi}\,dx - Re_1J\delta\int_\Omega w_3\eta_3\phi_2\,dx - \widetilde{Re}J\delta\int_\Omega v_3\eta_3\phi_2\,dx\\
&+Re_1J\delta\int_\Omega w_3\eta_2\phi_3\,dx + \widetilde{Re}J\delta\int_\Omega v_3\eta_2\phi_3\,dx\\
&\le -2m\int_\Omega H\boldsymbol{\phi}\cdot\boldsymbol{\phi}\,dx + m\int_\Omega H(\nabla\times\mathbf{w})\cdot\boldsymbol{\phi}\,dx + m\delta\int_\Omega w_3\phi_1\,dx\\
&+n_1\int_\Omega \nabla\cdot(H\nabla\boldsymbol{\phi})\cdot\boldsymbol{\phi}\,dx - n_1\delta^2\int_\Omega\frac{1}{H}(\phi_2^2+\phi_3^2)\,dx\\
&+(n_2+n_3)\int_\Omega H\nabla(\nabla\cdot\boldsymbol{\phi})\cdot\boldsymbol{\phi}\,dx + (n_2+n_3)\delta\int_\Omega(\nabla\boldsymbol{\phi})\cdot\boldsymbol{\phi}\,dx\\
&-(n_2+n_3)\delta^2\int_\Omega\frac{1}{H}\phi_2^2\,dx.
\end{aligned}\tag{20}$$

Integration by parts allows us to replace the integral $(n_2+n_3)\int_\Omega H\nabla(\nabla\cdot\boldsymbol{\phi})\cdot\boldsymbol{\phi}\,dx$ by $-(n_2+n_3)\int_\Omega H(\nabla\cdot\boldsymbol{\phi})^2\,dx - (n_2+n_3)\delta\int_\Omega\phi_2(\nabla\cdot\boldsymbol{\phi})\,dx$ in the inequality (20); subsequently, using the nonnegativity of $\int_\Omega H(\nabla\cdot\boldsymbol{\phi})^2\,dx$, $\int_\Omega\frac{1}{H}(\phi_2^2+\phi_3^2)\,dx$ and $\int_\Omega\frac{1}{H}\phi_2^2\,dx$ and the fact that $\int_\Omega H\big((\mathbf{u}\cdot\nabla)\boldsymbol{\phi}\big)\cdot\boldsymbol{\phi}\,dx = 0$, we obtain

$$\begin{aligned}
n_1H_0\|\nabla\boldsymbol{\phi}\|_2^2 + 2mH_0\|\boldsymbol{\phi}\|_2^2 &\le \left|Re_1J\int_\Omega H\big((\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\big)\cdot\boldsymbol{\phi}\,dx\right|\\
&+\left|Re_1J\delta\int_\Omega w_3\eta_3\phi_2\,dx\right| + \left|Re_1J\delta\int_\Omega w_3\eta_2\phi_3\,dx\right|\\
&+\left|(n_2+n_3)\delta\int_\Omega\phi_2(\nabla\cdot\boldsymbol{\phi})\,dx\right| + \left|(n_2+n_3)\delta\int_\Omega(\nabla\boldsymbol{\phi})\cdot\boldsymbol{\phi}\,dx\right|\\
&+\left|\widetilde{Re}J\int_\Omega H\big((\mathbf{v}\cdot\nabla)\boldsymbol{\eta}\big)\cdot\boldsymbol{\phi}\,dx\right| + \left|\widetilde{Re}J\delta\int_\Omega v_3\eta_3\phi_2\,dx\right| + \left|\widetilde{Re}J\delta\int_\Omega v_3\eta_2\phi_3\,dx\right|
\end{aligned}$$

$$+\left|m\int_\Omega H(\nabla\times\mathbf{w})\cdot\boldsymbol{\phi}\,dx\right|+\left|m\delta\int_\Omega w_3\phi_1\,dx\right|. \tag{21}$$

The terms present in right hand side of (21) can be estimated as

$$\begin{aligned}\left|Re_1J\int_\Omega H\big((\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\big)\cdot\boldsymbol{\phi}\,dx\right|&\le Re_1JH_1\|\mathbf{w}\|_4\|\nabla\boldsymbol{\eta}\|_2\|\boldsymbol{\phi}\|_4\\&\le \mathfrak{C}Re_1JE_{\boldsymbol{\eta}}\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2,\end{aligned}$$

$$\begin{aligned}\left|Re_1J\delta\int_\Omega w_3\eta_3\phi_2\,dx\right|&\le Re_1J\delta\|\mathbf{w}\|_4\|\boldsymbol{\eta}\|_2\|\boldsymbol{\phi}\|_4\\&\le \mathfrak{C}Re_1J\delta E_{\boldsymbol{\eta}}\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2,\end{aligned}$$

$$\begin{aligned}\left|(n_2+n_3)\delta\int_\Omega \phi_2(\nabla\cdot\boldsymbol{\phi})\,dx\right|&\le (n_2+n_3)\delta\|\boldsymbol{\phi}\|_2\|\nabla\cdot\boldsymbol{\phi}\|_2\\&\le \mathfrak{C}(n_2+n_3)\delta\|\nabla\boldsymbol{\phi}\|_2^2,\end{aligned}$$

$$\begin{aligned}\left|(n_2+n_3)\delta\int_\Omega (\nabla\boldsymbol{\phi})\cdot\boldsymbol{\phi}\,dx\right|&\le (n_2+n_3)\delta\|\nabla\boldsymbol{\phi}\|_2\|\boldsymbol{\phi}\|_2\\&\le \mathfrak{C}(n_2+n_3)\delta\|\nabla\boldsymbol{\phi}\|_2^2,\end{aligned}$$

$$\begin{aligned}\left|\widetilde{Re}J\int_\Omega H\big((\mathbf{v}\cdot\nabla)\boldsymbol{\eta}\big)\cdot\boldsymbol{\phi}\,dx\right|&\le |\widetilde{Re}|JH_1\|\mathbf{v}\|_4\|\nabla\boldsymbol{\eta}\|_2\|\boldsymbol{\phi}\|_4\\&\le \mathfrak{C}|\widetilde{Re}|JE_{\mathbf{v}}E_{\boldsymbol{\eta}}\|\nabla\boldsymbol{\phi}\|_2,\end{aligned}$$

$$\begin{aligned}\left|\widetilde{Re}J\delta\int_\Omega v_3\eta_3\phi_2\,dx\right|&\le |\widetilde{Re}|J\delta\|\mathbf{v}\|_4\|\boldsymbol{\eta}\|_2\|\boldsymbol{\phi}\|_4\\&\le \mathfrak{C}|\widetilde{Re}|J\delta E_{\mathbf{v}}E_{\boldsymbol{\eta}}\|\nabla\boldsymbol{\phi}\|_2,\end{aligned}$$

$$\begin{aligned}\left|m\int_\Omega H(\nabla\times\mathbf{w})\cdot\boldsymbol{\phi}\,dx\right|&\le mH_1\|\nabla\times\mathbf{w}\|_2\|\boldsymbol{\phi}\|_2\\&\le \mathfrak{C}m\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2,\end{aligned}$$

and

$$\begin{aligned}\left|m\delta\int_\Omega w_3\phi_1\,dx\right|&\le m\delta\|\mathbf{w}\|_2\|\boldsymbol{\phi}\|_2\\&\le \mathfrak{C}m\delta\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2.\end{aligned}$$

Combining all these bounds with (21), we obtain

$$n_1 H_0 \|\nabla \boldsymbol{\phi}\|_2^2 \leq \mathfrak{C}\Big\{(1+\delta)(m+Re_1 JE_{\boldsymbol{\eta}})\|\nabla \mathbf{w}\|_2 \|\nabla \boldsymbol{\phi}\|_2 + (n_2+n_3)\delta \|\nabla \boldsymbol{\phi}\|_2^2 + |\widetilde{Re}| J(1+\delta)E_{\mathbf{v}} E_{\boldsymbol{\eta}} \|\nabla \boldsymbol{\phi}\|_2 \Big\}. \tag{22}$$

Adding (19) and (22), we get

$$\begin{aligned}
&(1+m)H_0\|\nabla \mathbf{w}\|_2^2 + n_1 H_0 \|\nabla \boldsymbol{\phi}\|_2^2 \\
&\leq \mathfrak{C}\Big\{Re_1(1+\delta)E_{\mathbf{v}}\|\nabla \mathbf{w}\|_2^2 + |\widetilde{Re}|\Big((1+\delta+\sigma C_F)E_{\mathbf{v}}^2+1\Big)\|\nabla \mathbf{w}\|_2 \\
&\quad +(1+\delta)(m+Re_1 JE_{\boldsymbol{\eta}})\|\nabla \mathbf{w}\|_2\|\nabla \boldsymbol{\phi}\|_2 + (n_2+n_3)\delta\|\nabla \boldsymbol{\phi}\|_2^2 \\
&\quad +|\widetilde{Re}| J(1+\delta)E_{\mathbf{v}}E_{\boldsymbol{\eta}}\|\nabla \boldsymbol{\phi}\|_2\Big\}.
\end{aligned} \tag{23}$$

By Young's inequality, we obtain

$$\begin{aligned}
&(1+m)H_0\|\nabla \mathbf{w}\|_2^2 + n_1 H_0 \|\nabla \boldsymbol{\phi}\|_2^2 \\
&\leq \frac{(1+m)H_0}{2}\|\nabla \mathbf{w}\|_2^2 + \frac{n_1 H_0}{2}\|\nabla \boldsymbol{\phi}\|_2^2 + \mathfrak{C}(n_2+n_3)\delta\|\nabla \boldsymbol{\phi}\|_2^2 \\
&\quad +\mathfrak{C}\Big\{Re_1(1+\delta)E_{\mathbf{v}} + \frac{(1+\delta)^2(m+Re_1 JE_{\boldsymbol{\eta}})^2}{n_1 H_0}\Big\}\|\nabla \mathbf{w}\|_2^2 \\
&\quad +\mathfrak{C}\left\{\frac{\Big((1+\delta+\sigma C_F)E_{\mathbf{v}}^2+1\Big)^2}{(1+m)H_0} + \frac{J^2(1+\delta)^2 E_{\mathbf{v}}^2 E_{\boldsymbol{\eta}}^2}{n_1 H_0}\right\}\widetilde{Re}^2.
\end{aligned} \tag{24}$$

After simplification, (24) yields

$$\begin{aligned}
&\left\{\frac{(1+m)H_0}{2} - \mathfrak{C}\left(Re_1(1+\delta)E_{\mathbf{v}} + \frac{(1+\delta)^2(m+Re_1 JE_{\boldsymbol{\eta}})^2}{n_1 H_0}\right)\right\}\|\nabla \mathbf{w}\|_2^2 \\
&\quad +\left\{\frac{n_1 H_0}{2} - \mathfrak{C}(n_2+n_3)\delta\right\}\|\nabla \boldsymbol{\phi}\|_2^2 \\
&\leq \mathfrak{C}\left\{\frac{\Big((1+\delta+\sigma C_F)E_{\mathbf{v}}^2+1\Big)^2}{(1+m)H_0} + \frac{J^2(1+\delta)^2 E_{\mathbf{v}}^2 E_{\boldsymbol{\eta}}^2}{n_1 H_0}\right\}\widetilde{Re}^2.
\end{aligned} \tag{25}$$

It follows from inequality (25) that the required result holds whenever $Re_1 \leq \epsilon_1$, $m \leq \epsilon_2$ and $\delta \leq \epsilon_3$ for some positive constants ϵ_1, ϵ_2 and ϵ_3. □

Remark 1. *The convergence of* $\mathbf{v}$ *to* $\mathbf{u}$ *in* $\mathcal{X}$ *and* $\boldsymbol{\eta}$ *to* $\boldsymbol{\nu}$ *in* $\left(H_0^1(\Omega)\right)^3$ *as* $Re_2 \to 0$ *follows from the inequality* (25) *when* $Re_1 = 0$. *Therefore, it can be concluded that the BVP* (1)-(4) *admits a regular perturbation with respect to the Reynolds number* Re.

3.2 Continuous dependence on the curvature ratio (δ)

Suppose that $(\mathbf{u}, p, \boldsymbol{\nu})$ and $(\mathbf{v}, q, \boldsymbol{\eta})$ are the solutions of the BVP (1)-(4) corresponding to the curvature ratios δ_1 and δ_2, respectively. Let $(\mathbf{w}, \pi, \boldsymbol{\phi}) := (\mathbf{u} - \mathbf{v}, p - q, \boldsymbol{\nu} - \boldsymbol{\eta})$. Then $(\mathbf{w}, \pi, \boldsymbol{\phi})$ satisfies the BVP

$$\nabla \cdot (H_{\delta_1}\mathbf{w}) + \tilde{\delta}\nabla \cdot (x_2\mathbf{v}) = 0, \tag{26}$$

$$\begin{aligned}
&ReH_{\delta_1}\Big((\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{v}\Big) + \tilde{\delta}Rex_2(\mathbf{v}\cdot\nabla)\mathbf{v} - Re\delta_1 w_3(u_3+v_3)\hat{a}_2 \\
&\quad -\tilde{\delta}Rev_3^2\hat{a}_2 + Re\delta_1(w_2u_3 + v_2w_3)\hat{a}_3 + \tilde{\delta}Rev_2v_3\hat{a}_3 \\
&= -ReH_{\delta_1}\nabla\pi - \tilde{\delta}Rex_2\nabla q + mH_{\delta_1}\nabla\times\boldsymbol{\phi} + \tilde{\delta}mx_2\nabla\times\boldsymbol{\eta} + m\delta_1\phi_3\hat{a}_1 \\
&\quad +\tilde{\delta}m\eta_3\hat{a}_1 + (1+m)\nabla\cdot(H_{\delta_1}\nabla\mathbf{w}) + \tilde{\delta}(1+m)\nabla\cdot(x_2\nabla\mathbf{v}) \\
&-(1+m)\frac{\delta_1^2}{H_{\delta_1}}(w_2\hat{a}_2 + w_3\hat{a}_3) - \tilde{\delta}(1+m)\left(\frac{\delta_1+\delta_2}{H_{\delta_1}} - \frac{\delta_2^2x_2}{H_{\delta_1}H_{\delta_2}}\right)(v_2\hat{a}_2 + v_3\hat{a}_3) \\
&\quad -\sigma^2H_{\delta_1}\mathbf{w} - \tilde{\delta}\sigma^2x_2\mathbf{v} - Re\sigma C_FH_{\delta_1}(|\mathbf{u}|\mathbf{u} - |\mathbf{v}|\mathbf{v}) - \tilde{\delta}Re\sigma C_Fx_2|\mathbf{v}|\mathbf{v},
\end{aligned} \tag{27}$$

$$\begin{aligned}
&ReH_{\delta_1}J\Big((\mathbf{u}\cdot\nabla)\boldsymbol{\phi} + (\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\Big) + \tilde{\delta}ReJx_2(\mathbf{v}\cdot\nabla)\boldsymbol{\eta} - ReJ\delta_1(u_3\phi_3 + w_3\eta_3)\hat{a}_2 \\
&\quad -\tilde{\delta}ReJv_3\eta_3\hat{a}_2 + ReJ\delta_1(u_3\phi_2 + w_3\eta_2)\hat{a}_3 + \tilde{\delta}ReJv_3\eta_2\hat{a}_3 \\
&= -2mH_{\delta_1}\boldsymbol{\phi} - 2\tilde{\delta}mx_2\boldsymbol{\eta} + mH_{\delta_1}\nabla\times\mathbf{w} + \tilde{\delta}mx_2\nabla\times\mathbf{v} + m\delta_1w_3\hat{a}_1 + \tilde{\delta}mv_3\hat{a}_1 \\
&\quad +n_1\nabla\cdot(H_{\delta_1}\nabla\boldsymbol{\phi}) + \tilde{\delta}n_1\nabla\cdot(x_2\nabla\boldsymbol{\eta}) - n_1\frac{\delta_1^2}{H_{\delta_1}}(\phi_2\hat{a}_2 + \phi_3\hat{a}_3) \\
&-\tilde{\delta}n_1\left(\frac{\delta_1+\delta_2}{H_{\delta_1}} - \frac{\delta_2^2x_2}{H_{\delta_1}H_{\delta_2}}\right)(\eta_2\hat{a}_2 + \eta_3\hat{a}_3) + (n_2+n_3)H_{\delta_1}\nabla(\nabla\cdot\boldsymbol{\phi}) \\
&\quad +\tilde{\delta}(n_2+n_3)x_2\nabla(\nabla\cdot\boldsymbol{\eta}) + (n_2+n_3)\delta_1\nabla\phi_2 + \tilde{\delta}(n_2+n_3)\nabla\eta_2 \\
&-(n_2+n_3)\frac{\delta_1^2}{H_{\delta_1}}\phi_2\hat{a}_2 - \tilde{\delta}(n_2+n_3)\left(\frac{\delta_1+\delta_2}{H_{\delta_1}} - \frac{\delta_2^2x_2}{H_{\delta_1}H_{\delta_2}}\right)\eta_2\hat{a}_2
\end{aligned} \tag{28}$$

in Ω,

and

$$\mathbf{w} = \mathbf{0} = \boldsymbol{\phi} \text{ on } \partial\Omega, \tag{29}$$

where $\mathbf{w} = (w_1, w_2, w_3)$, $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)$, $w_i = u_i - v_i$, $\phi_i = \nu_i - \eta_i$ $(i = 1, 2, 3)$, $\tilde{\delta} = \delta_1 - \delta_2$ and $H_{\delta_j} = 1 + \delta_j x_2$ $(j = 1, 2)$.

The principal result concerning the continuous dependence on the curvature ratio (δ) is presented below:

Theorem 2. *Let* $(\mathbf{w}, \pi, \boldsymbol{\phi})$ *be the solution of the boundary value problem* (26)-(29). *Then* $\mathbf{w} \to \mathbf{0}$ *in* $\mathcal{X}$ *and* $\boldsymbol{\phi} \to \mathbf{0}$ *in* $(H_0^1(\Omega))^3$ *as* $\delta_2 \to \delta_1$ *for sufficiently small values of* Re, m *and* δ_1.

Proof. Taking the $L^2(\Omega)$-inner product on both sides of (27) with $\mathbf{w}$, we get

$$
\begin{aligned}
&Re \int_\Omega H_{\delta_1}\big((\mathbf{u}\cdot\nabla)\mathbf{w}\big)\cdot\mathbf{w}\,dx + Re\int_\Omega H_{\delta_1}\big((\mathbf{w}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx\\
&+\tilde{\delta}Re\int_\Omega x_2\big(\mathbf{v}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx - Re\delta_1\int_\Omega (u_3+v_3)w_3w_2\,dx\\
&-\tilde{\delta}Re\int_\Omega v_3^2w_2\,dx + Re\delta_1\int_\Omega (w_2u_3+v_2w_3)w_3\,dx + \tilde{\delta}Re\int_\Omega v_2v_3w_3\,dx\\
&= -Re\int_\Omega H_{\delta_1}\nabla\pi\cdot\mathbf{w}\,dx - \tilde{\delta}Re\int_\Omega x_2\nabla q\cdot\mathbf{w}\,dx + m\int_\Omega H_{\delta_1}(\nabla\times\boldsymbol{\phi})\cdot\mathbf{w}\,dx\\
&+\tilde{\delta}m\int_\Omega x_2(\nabla\times\boldsymbol{\eta})\cdot\mathbf{w}\,dx + m\delta_1\int_\Omega \phi_3w_1\,dx + \tilde{\delta}m\int_\Omega \eta_3w_1\,dx\\
&+(1+m)\int_\Omega \nabla\cdot(H_{\delta_1}\nabla\mathbf{w})\cdot\mathbf{w}\,dx + \tilde{\delta}(1+m)\int_\Omega \nabla\cdot(x_2\nabla\mathbf{v})\cdot\mathbf{w}\,dx\\
&-(1+m)\delta_1^2\int_\Omega \frac{1}{H_{\delta_1}}(w_2^2+w_3^2)\,dx - \tilde{\delta}(1+m)(\delta_1+\delta_2)\int_\Omega \frac{1}{H_{\delta_1}}(v_2w_2+v_3w_3)\,dx\\
&+\tilde{\delta}\delta_2^2(1+m)\int_\Omega \frac{x_2}{H_{\delta_1}H_{\delta_2}}(v_2w_2+v_3w_3)\,dx - \sigma^2\int_\Omega H_{\delta_1}\mathbf{w}\cdot\mathbf{w}\,dx\\
&-\tilde{\delta}\sigma^2\int_\Omega x_2\mathbf{v}\cdot\mathbf{w}\,dx - Re\sigma C_F H_{\delta_1}\int_\Omega (|\mathbf{u}|\mathbf{u}-|\mathbf{v}|\mathbf{v})\cdot\mathbf{w}\,dx\\
&-\tilde{\delta}Re\sigma C_F\int_\Omega x_2|\mathbf{v}|\mathbf{v}\cdot\mathbf{w}\,dx. \qquad (30)
\end{aligned}
$$

Using the identity $\int_\Omega H_{\delta_1}\big((\mathbf{u}\cdot\nabla)\mathbf{w}\big)\cdot\mathbf{w}\,dx = 0$, and the nonnegativity of the integrals $\int_\Omega \frac{1}{H_{\delta_1}}(w_2^2+w_3^2)\,dx$ and $\int_\Omega(|\mathbf{u}|\mathbf{u}-|\mathbf{v}|\mathbf{v})\cdot\mathbf{w}\,dx$, the equality (30) gives

$$
\begin{aligned}
&(1+m)\left\|\sqrt{H_{\delta_1}}\nabla\mathbf{w}\right\|_2^2 + \sigma^2\left\|\sqrt{H_{\delta_1}}\mathbf{w}\right\|_2^2 \le \left|Re\int_\Omega H_{\delta_1}\big((\mathbf{w}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx\right|\\
&+\left|\tilde{\delta}Re\int_\Omega x_2\big(\mathbf{v}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx\right| + \left|Re\delta_1\int_\Omega v_3w_3w_2\,dx\right| + \left|\tilde{\delta}Re\int_\Omega v_3^2w_2\,dx\right|\\
&+\left|Re\delta_1\int_\Omega v_2w_3^2\,dx\right| + \left|\tilde{\delta}Re\int_\Omega v_2v_3w_3\,dx\right| + \left|Re\int_\Omega H_{\delta_1}\nabla\pi\cdot\mathbf{w}\,dx\right|\\
&+\left|\tilde{\delta}Re\int_\Omega x_2\nabla q\cdot\mathbf{w}\,dx\right| + \left|m\int_\Omega H_{\delta_1}(\nabla\times\boldsymbol{\phi})\cdot\mathbf{w}\,dx\right| + \left|\tilde{\delta}m\int_\Omega x_2(\nabla\times\boldsymbol{\eta})\cdot\mathbf{w}\,dx\right|
\end{aligned}
$$

$$+\left|m\delta_1\int_\Omega \phi_3 w_1\,dx\right|+\left|\tilde{\delta}m\int_\Omega \eta_3 w_1\,dx\right|+\left|\tilde{\delta}(1+m)\int_\Omega \nabla\cdot(x_2\nabla\mathbf{v})\cdot\mathbf{w}\,dx\right|$$
$$+\left|\tilde{\delta}(1+m)(\delta_1+\delta_2)\int_\Omega \frac{1}{H_{\delta_1}}(v_2w_2+v_3w_3)\,dx\right|$$
$$+\left|\tilde{\delta}\delta_2^2(1+m)\int_\Omega \frac{x_2}{H_{\delta_1}H_{\delta_2}}(v_2w_2+v_3w_3)\,dx\right|+\left|\tilde{\delta}\sigma^2\int_\Omega x_2\mathbf{v}\cdot\mathbf{w}\,dx\right|$$
$$+\left|\tilde{\delta}Re\sigma C_F\int_\Omega x_2|\mathbf{v}|\mathbf{v}\cdot\mathbf{w}\,dx\right|. \tag{31}$$

By Hölder's and Sobolev inequalities together with (26), the terms of the right-hand side of (31) can be estimated as follows:

$$\left|Re\int_\Omega H_{\delta_1}\big((\mathbf{w}\cdot\nabla)\mathbf{v}\big)\cdot\mathbf{w}\,dx\right|\le \mathfrak{C}ReE_{\mathbf{v}}\|\nabla\mathbf{w}\|_2^2,$$

$$\left|\tilde{\delta}Re\int_\Omega x_2\big(\mathbf{v}\cdot\nabla\big)\mathbf{v}\big)\cdot\mathbf{w}\,dx\right|\le \mathfrak{C}|\tilde{\delta}|ReE_{\mathbf{v}}^2\|\nabla\mathbf{w}\|_2,$$

$$\left|Re\delta_1\int_\Omega v_3w_3w_2\,dx\right|\le \mathfrak{C}Re\delta_1E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2^2,$$

$$\left|\tilde{\delta}Re\int_\Omega v_3^2w_2\,dx\right|\le \mathfrak{C}|\tilde{\delta}|ReE_{\mathbf{v}}^2\|\nabla\mathbf{w}\|_2,$$

$$\begin{aligned}\left|Re\int_\Omega H_{\delta_1}\nabla\pi\cdot\mathbf{w}\,dx\right|&=\left|Re\int_\Omega \pi\nabla\cdot(H_{\delta_1}\mathbf{w})\,dx\right|\\&=\left|Re\int_\Omega \pi\big(-\tilde{\delta}\nabla\cdot(x_2\mathbf{v})\big)\,dx\right|\ \text{(by (26))}\\&\le|\tilde{\delta}|Re\|\pi\|_2\|x_2\nabla\cdot\mathbf{v}+v_2\|_2\\&\le\mathfrak{C}|\tilde{\delta}|Re(\|p\|_2+\|q\|_2)\|\nabla\mathbf{v}\|_2\\&\le\mathfrak{C}|\tilde{\delta}|Re(E_p+E_q)E_{\mathbf{v}},\end{aligned}$$

$$\begin{aligned}\left|\tilde{\delta}Re\int_\Omega x_2\nabla q\cdot\mathbf{w}\,dx\right|&=\left|\tilde{\delta}Re\int_\Omega q\nabla\cdot(x_2\mathbf{w})\,dx\right|\\&=\left|\tilde{\delta}Re\int_\Omega q(x_2\nabla\cdot\mathbf{w}+w_2)\,dx\right|\\&\le|\tilde{\delta}|Re\|q\|_2\big(\|x_2\|_\infty\|\nabla\cdot\mathbf{w}\|_2+\|\mathbf{w}\|_2\big)\\&\le\mathfrak{C}|\tilde{\delta}|ReE_q\|\nabla\mathbf{w}\|_2,\end{aligned}$$

$$\left|m\int_\Omega H_{\delta_1}(\nabla\times\boldsymbol{\phi})\cdot\mathbf{w}\,dx\right|\le\mathfrak{C}m\|\nabla\boldsymbol{\phi}\|_2\|\nabla\mathbf{w}\|_2,$$

$$\left|\tilde{\delta}m\int_\Omega x_2(\nabla\times\boldsymbol{\eta})\cdot\mathbf{w}\,dx\right|\le\mathfrak{C}|\tilde{\delta}|mE_{\boldsymbol{\eta}}\|\nabla\mathbf{w}\|_2,$$

$$\left|m\delta_1\int_\Omega\phi_3 w_1\,dx\right|\le\mathfrak{C}m\delta_1\|\nabla\boldsymbol{\phi}\|_2\|\nabla\mathbf{w}\|_2,$$

$$\left|\tilde{\delta}m\int_\Omega\eta_3 w_1\,dx\right|\le\mathfrak{C}|\tilde{\delta}|mE_{\boldsymbol{\eta}}\|\nabla\mathbf{w}\|_2,$$

$$\begin{aligned}\left|\tilde{\delta}(1+m)\int_\Omega\nabla\cdot(x_2\nabla\mathbf{v})\cdot\mathbf{w}\,dx\right|&=\left|\tilde{\delta}(1+m)\int_\Omega x_2\nabla\mathbf{v}:\nabla\mathbf{w}\,dx\right|\\&\le\mathfrak{C}|\tilde{\delta}|(1+m)E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2,\end{aligned}$$

$$\left|\tilde{\delta}(1+m)(\delta_1+\delta_2)\int_\Omega\frac{1}{H_{\delta_1}}(v_2w_2+v_3w_3)\,dx\right|\le\mathfrak{C}|\tilde{\delta}|(1+m)\frac{(\delta_1+\delta_2)}{H^0_{\delta_1}}E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2,$$

$$\left|\tilde{\delta}\delta_2^2(1+m)\int_\Omega\frac{x_2}{H_{\delta_1}H_{\delta_2}}(v_2w_2+v_3w_3)\,dx\right|\le\mathfrak{C}|\tilde{\delta}|(1+m)\frac{\delta_2^2}{H^0_{\delta_1}H^0_{\delta_2}}E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2,$$

$$\left|\tilde{\delta}\sigma^2\int_\Omega x_2\mathbf{v}\cdot\mathbf{w}\,dx\right|\le\mathfrak{C}|\tilde{\delta}|\sigma^2E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2,$$

and

$$\left|\tilde{\delta}Re\sigma C_F\int_\Omega x_2|\mathbf{v}|\mathbf{v}\cdot\mathbf{w}\,dx\right|\le\mathfrak{C}|\tilde{\delta}|Re\sigma C_FE_{\mathbf{v}}^2\|\nabla\mathbf{w}\|_2,$$

where $H^0_{\delta_1}:=\inf_\Omega H_{\delta_1}$ and $H^0_{\delta_2}:=\inf_\Omega H_{\delta_2}$.
In view of these, the inequality (31) becomes

$$\begin{aligned}(1+m)H^0_{\delta_1}\|\nabla\mathbf{w}\|_2^2\le\mathfrak{C}\Bigg\{&Re(1+\delta_1)E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2^2+|\tilde{\delta}|\|\nabla\mathbf{w}\|_2\Bigg(Re(1+\sigma C_F)E_{\mathbf{v}}^2\\&+ReE_q+mE_{\boldsymbol{\eta}}+(1+m)\left(1+\frac{\delta_1+\delta_2}{H^0_{\delta_1}}+\frac{\delta_2^2}{H^0_{\delta_1}H^0_{\delta_2}}\right)E_{\mathbf{v}}+\sigma^2E_{\mathbf{v}}\Bigg)\\&+|\tilde{\delta}|Re(E_p+E_q)E_{\mathbf{v}}+m(1+\delta_1)\|\nabla\boldsymbol{\phi}\|_2\|\nabla\mathbf{w}\|_2\Bigg\}.\end{aligned}\tag{32}$$

Taking the $L^2(\Omega)$-inner product on both sides of (28) with $\boldsymbol{\phi}$, we get

$$
\begin{aligned}
&ReJ\int_\Omega H_{\delta_1}\Big((\mathbf{u}\cdot\nabla)\boldsymbol{\phi}\Big)\cdot\boldsymbol{\phi}\,dx + ReJ\int_\Omega H_{\delta_1}\Big((\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\Big)\cdot\boldsymbol{\phi}\,dx\\
&+\tilde{\delta}ReJ\int_\Omega x_2\Big((\mathbf{v}\cdot\nabla)\boldsymbol{\eta}\Big)\cdot\boldsymbol{\phi}\,dx - ReJ\delta_1\int_\Omega w_3\eta_3\phi_2\,dx - \tilde{\delta}ReJ\int_\Omega v_3\eta_3\phi_2\,dx\\
&+ReJ\delta_1\int_\Omega w_3\eta_2\phi_3\,dx + \tilde{\delta}ReJ\int_\Omega v_3\eta_2\phi_3\,dx\\
&= -2m\int_\Omega H_{\delta_1}\boldsymbol{\phi}\cdot\boldsymbol{\phi}\,dx - 2\tilde{\delta}m\int_\Omega x_2\boldsymbol{\eta}\cdot\boldsymbol{\phi}\,dx + m\int_\Omega H_{\delta_1}(\nabla\times\mathbf{w})\cdot\boldsymbol{\phi}\,dx\\
&+\tilde{\delta}m\int_\Omega x_2(\nabla\times\mathbf{v})\cdot\boldsymbol{\phi}\,dx + m\delta_1\int_\Omega w_3\phi_1\,dx + \tilde{\delta}m\int_\Omega v_3\phi_1\,dx\\
&+n_1\int_\Omega \nabla\cdot(H_{\delta_1}\nabla\boldsymbol{\phi})\cdot\boldsymbol{\phi}\,dx + \tilde{\delta}n_1\int_\Omega \nabla\cdot(x_2\nabla\boldsymbol{\eta})\cdot\boldsymbol{\phi}\,dx\\
&-n_1\delta_1^2\int_\Omega\frac{1}{H_{\delta_1}}(\phi_2^2+\phi_3^2)\,dx - \tilde{\delta}n_1(\delta_1+\delta_2)\int_\Omega\frac{1}{H_{\delta_1}}(\eta_2\phi_2+\eta_3\phi_3)\,dx\\
&+\tilde{\delta}n_1\delta_2^2\int_\Omega\frac{x_2}{H_{\delta_1}H_{\delta_2}}(\eta_2\phi_2+\eta_3\phi_3)\,dx + (n_2+n_3)\int_\Omega H_{\delta_1}\nabla(\nabla\cdot\boldsymbol{\phi})\cdot\boldsymbol{\phi}\,dx\\
&+\tilde{\delta}(n_2+n_3)\int_\Omega x_2\nabla(\nabla\cdot\boldsymbol{\eta})\cdot\boldsymbol{\phi}\,dx + (n_2+n_3)\delta_1\int_\Omega\nabla\phi_2\cdot\boldsymbol{\phi}\,dx\\
&+\tilde{\delta}(n_2+n_3)\int_\Omega\nabla\eta_2\cdot\boldsymbol{\phi}\,dx - (n_2+n_3)\delta_1^2\int_\Omega\frac{1}{H_{\delta_1}}\phi_2^2\,dx\\
&-\tilde{\delta}(n_2+n_3)(\delta_1+\delta_2)\int_\Omega\frac{1}{H_{\delta_1}}\eta_2\phi_2\,dx + \tilde{\delta}(n_2+n_3)\delta_2^2\int_\Omega\frac{x_2}{H_{\delta_1}H_{\delta_2}}\eta_2\phi_2\,dx.
\end{aligned}
\tag{33}
$$

After simplification, (33) turns out to be

$$
\begin{aligned}
n_1H^0_{\delta_1}\|\nabla\boldsymbol{\phi}\|_2^2 + 2mH^0_{\delta_1}\|\boldsymbol{\phi}\|_2^2 &\le \left|ReJ\int_\Omega H_{\delta_1}\Big((\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\Big)\cdot\boldsymbol{\phi}\,dx\right|\\
&+\left|\tilde{\delta}ReJ\int_\Omega x_2\Big((\mathbf{v}\cdot\nabla)\boldsymbol{\eta}\Big)\cdot\boldsymbol{\phi}\,dx\right| + \left|ReJ\delta_1\int_\Omega w_3\eta_3\phi_2\,dx\right|\\
&+\left|\tilde{\delta}ReJ\int_\Omega v_3\eta_3\phi_2\,dx\right| + \left|ReJ\delta_1\int_\Omega w_3\eta_2\phi_3\,dx\right| + \left|\tilde{\delta}ReJ\int_\Omega v_3\eta_2\phi_3\,dx\right|\\
&+\left|2\tilde{\delta}m\int_\Omega x_2\boldsymbol{\eta}\cdot\boldsymbol{\phi}\,dx\right| + \left|m\int_\Omega H_{\delta_1}(\nabla\times\mathbf{w})\cdot\boldsymbol{\phi}\,dx\right| + \left|\tilde{\delta}m\int_\Omega x_2(\nabla\times\mathbf{v})\cdot\boldsymbol{\phi}\,dx\right|\\
&+\left|m\delta_1\int_\Omega w_3\phi_1\,dx\right| + \left|\tilde{\delta}m\int_\Omega v_3\phi_1\,dx\right| + \left|\tilde{\delta}n_1\int_\Omega\nabla\cdot(x_2\nabla\boldsymbol{\eta})\cdot\boldsymbol{\phi}\,dx\right|
\end{aligned}
$$

$$+\left|\tilde{\delta}n_1(\delta_1+\delta_2)\int_\Omega \frac{1}{H_{\delta_1}}(\eta_2\phi_2+\eta_3\phi_3)\,dx\right|+\left|\tilde{\delta}n_1\delta_2^2\int_\Omega \frac{x_2}{H_{\delta_1}H_{\delta_2}}(\eta_2\phi_2+\eta_3\phi_3)\,dx\right|$$
$$+\left|(n_2+n_3)\delta_1\int_\Omega \phi_2(\nabla\cdot\boldsymbol{\phi})\,dx\right|+\left|\tilde{\delta}(n_2+n_3)\int_\Omega x_2(\nabla\cdot\boldsymbol{\eta})(\nabla\cdot\boldsymbol{\phi})\,dx\right|$$
$$+\left|\tilde{\delta}(n_2+n_3)\int_\Omega \phi_2(\nabla\cdot\boldsymbol{\eta})\,dx\right|+\left|(n_2+n_3)\delta_1\int_\Omega \nabla\phi_2\cdot\boldsymbol{\phi}\,dx\right|$$
$$+\left|\tilde{\delta}(n_2+n_3)\int_\Omega \nabla\eta_2\cdot\boldsymbol{\phi}\,dx\right|+\left|\tilde{\delta}(n_2+n_3)(\delta_1+\delta_2)\int_\Omega \frac{1}{H_{\delta_1}}\eta_2\phi_2\,dx\right|$$
$$+\left|\tilde{\delta}(n_2+n_3)\delta_2^2\int_\Omega \frac{x_2}{H_{\delta_1}H_{\delta_2}}\eta_2\phi_2\,dx\right|. \tag{34}$$

The terms appearing on the right-hand side of (34) admit the following bounds.

$$\left|ReJ\int_\Omega H_{\delta_1}\Big((\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\Big)\cdot\boldsymbol{\phi}\,dx\right|\le \mathfrak{C}ReJE_{\boldsymbol{\eta}}\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2,$$

$$\left|\tilde{\delta}ReJ\int_\Omega x_2\Big((\mathbf{v}\cdot\nabla)\boldsymbol{\eta}\Big)\cdot\boldsymbol{\phi}\,dx\right|\le \mathfrak{C}|\tilde{\delta}|ReJE_{\mathbf{v}}E_{\boldsymbol{\eta}}\|\nabla\boldsymbol{\phi}\|_2,$$

$$\left|ReJ\delta_1\int_\Omega w_3\eta_3\phi_2\,dx\right|\le \mathfrak{C}ReJ\delta_1E_{\boldsymbol{\eta}}\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2,$$

$$\left|\tilde{\delta}ReJ\int_\Omega v_3\eta_3\phi_2\,dx\right|\le \mathfrak{C}|\tilde{\delta}|ReJE_{\mathbf{v}}E_{\boldsymbol{\eta}}\|\nabla\boldsymbol{\phi}\|_2,$$

$$\left|2\tilde{\delta}m\int_\Omega x_2\boldsymbol{\eta}\cdot\boldsymbol{\phi}\,dx\right|\le \mathfrak{C}|\tilde{\delta}|mE_{\boldsymbol{\eta}}\|\nabla\boldsymbol{\phi}\|_2,$$

$$\left|m\int_\Omega H_{\delta_1}(\nabla\times\mathbf{w})\cdot\boldsymbol{\phi}\,dx\right|\le \mathfrak{C}m\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2,$$

$$\left|\tilde{\delta}m\int_\Omega x_2(\nabla\times\mathbf{v})\cdot\boldsymbol{\phi}\,dx\right|\le \mathfrak{C}|\tilde{\delta}|mE_{\mathbf{v}}\|\nabla\boldsymbol{\phi}\|_2,$$

$$\left|m\delta_1\int_\Omega w_3\phi_1\,dx\right|\le \mathfrak{C}m\delta_1\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2,$$

$$\left|\tilde{\delta} m \int_\Omega v_3 \phi_1 \, dx\right| \le \mathfrak{C}|\tilde{\delta}| m E_{\mathbf{v}} \|\nabla \boldsymbol{\phi}\|_2,$$

$$\left|\tilde{\delta} n_1 \int_\Omega \nabla \cdot (x_2 \nabla \boldsymbol{\eta}) \cdot \boldsymbol{\phi} \, dx\right| = \left|\tilde{\delta} n_1 \int_\Omega x_2 \nabla \boldsymbol{\eta} : \nabla \boldsymbol{\phi} \, dx\right| \le \mathfrak{C}|\tilde{\delta}| n_1 E_{\boldsymbol{\eta}} \|\nabla \boldsymbol{\phi}\|_2,$$

$$\left|\tilde{\delta} n_1 (\delta_1 + \delta_2) \int_\Omega \frac{1}{H_{\delta_1}} (\eta_2 \phi_2 + \eta_3 \phi_3) \, dx\right| \le \mathfrak{C}|\tilde{\delta}| \frac{n_1 (\delta_1 + \delta_2)}{H^0_{\delta_1}} E_{\boldsymbol{\eta}} \|\nabla \boldsymbol{\phi}\|_2,$$

$$\left|\tilde{\delta} n_1 \delta_2^2 \int_\Omega \frac{x_2}{H_{\delta_1} H_{\delta_2}} (\eta_2 \phi_2 + \eta_3 \phi_3) \, dx\right| \le \mathfrak{C}|\tilde{\delta}| \frac{n_1 \delta_2^2}{H^0_{\delta_1} H^0_{\delta_2}} E_{\boldsymbol{\eta}} \|\nabla \boldsymbol{\phi}\|_2,$$

$$\left|(n_2 + n_3) \delta_1 \int_\Omega \phi_2 (\nabla \cdot \boldsymbol{\phi}) \, dx\right| \le \mathfrak{C}(n_2 + n_3) \delta_1 \|\nabla \boldsymbol{\phi}\|_2^2,$$

$$\left|\tilde{\delta} (n_2 + n_3) \int_\Omega x_2 (\nabla \cdot \boldsymbol{\eta})(\nabla \cdot \boldsymbol{\phi}) \, dx\right| \le |\tilde{\delta}| (n_2 + n_3) E_{\boldsymbol{\eta}} \|\nabla \boldsymbol{\phi}\|_2,$$

$$\left|\tilde{\delta} (n_2 + n_3) \int_\Omega \phi_2 (\nabla \cdot \boldsymbol{\eta}) \, dx\right| \le |\tilde{\delta}| (n_2 + n_3) E_{\boldsymbol{\eta}} \|\nabla \boldsymbol{\phi}\|_2,$$

$$\left|(n_2 + n_3) \delta_1 \int_\Omega \nabla \phi_2 \cdot \boldsymbol{\phi} \, dx\right| \le \mathfrak{C}(n_2 + n_3) \delta_1 \|\nabla \boldsymbol{\phi}\|_2^2,$$

$$\left|\tilde{\delta} (n_2 + n_3) \int_\Omega \nabla \eta_2 \cdot \boldsymbol{\phi} \, dx\right| \le \mathfrak{C}|\tilde{\delta}| (n_2 + n_3) E_{\boldsymbol{\eta}} \|\nabla \boldsymbol{\phi}\|_2,$$

$$\left|\tilde{\delta} (n_2 + n_3)(\delta_1 + \delta_2) \int_\Omega \frac{1}{H_{\delta_1}} \eta_2 \phi_2 \, dx\right| \le \mathfrak{C}|\tilde{\delta}| \frac{(n_2 + n_3)(\delta_1 + \delta_2)}{H^0_{\delta_1}} E_{\boldsymbol{\eta}} \|\nabla \boldsymbol{\phi}\|_2,$$

and

$$\left|\tilde{\delta} (n_2 + n_3) \delta_2^2 \int_\Omega \frac{x_2}{H_{\delta_1} H_{\delta_2}} \eta_2 \phi_2 \, dx\right| \le \mathfrak{C}|\tilde{\delta}| \frac{(n_2 + n_3) \delta_2^2}{H^0_{\delta_1} H^0_{\delta_2}} E_{\boldsymbol{\eta}} \|\nabla \boldsymbol{\phi}\|_2.$$

Taking into account the preceding estimates, (34) gives

$$n_1 H^0_{\delta_1}\|\nabla\boldsymbol{\phi}\|_2^2 \le \mathfrak{C}\Bigg\{\delta_1(n_2+n_3)\|\nabla\boldsymbol{\phi}\|_2^2 + (1+\delta_1)\big(m+ReJE_{\boldsymbol{\eta}}\big)\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2 + |\tilde{\delta}|\|\nabla\boldsymbol{\phi}\|_2\Bigg((n_1+n_2+n_3)\Bigg(1+\frac{\delta_1+\delta_2}{H^0_{\delta_1}}+\frac{\delta_2^2}{H^0_{\delta_1}H^0_{\delta_2}}\Bigg)E_{\boldsymbol{\eta}} + mE_{\boldsymbol{\eta}} + \big(m+ReJE_{\boldsymbol{\eta}}\big)E_{\mathbf{v}}\Bigg)\Bigg\}. \tag{35}$$

Adding (32) and (35), we get

$$(1+m)H^0_{\delta_1}\|\nabla\mathbf{w}\|_2^2 + n_1H^0_{\delta_1}\|\nabla\boldsymbol{\phi}\|_2^2 \le \mathfrak{C}\big\{Re(1+\delta_1)E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2^2 + \delta_1(n_2+n_3)\|\nabla\boldsymbol{\phi}\|_2^2 + (1+\delta_1)\big(m+ReJE_{\boldsymbol{\eta}}\big)\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2 + |\tilde{\delta}|\mathfrak{D}_1\|\nabla\mathbf{w}\|_2 + |\tilde{\delta}|\mathfrak{D}_2\|\nabla\boldsymbol{\phi}\|_2 + |\tilde{\delta}|Re(E_p+E_q)E_{\mathbf{v}}\big\}, \tag{36}$$

where

$$\mathfrak{D}_1 = Re(1+\sigma C_F)E_{\mathbf{v}}^2 + ReE_q + mE_{\boldsymbol{\eta}} + (1+m)\left(1+\frac{\delta_1+\delta_2}{H^0_{\delta_1}}+\frac{\delta_2^2}{H^0_{\delta_1}H^0_{\delta_2}}\right)E_{\mathbf{v}} + \sigma^2E_{\mathbf{v}}$$

and

$$\mathfrak{D}_2 = (n_1+n_2+n_3)\left(1+\frac{\delta_1+\delta_2}{H^0_{\delta_1}}+\frac{\delta_2^2}{H^0_{\delta_1}H^0_{\delta_2}}\right)E_{\boldsymbol{\eta}} + mE_{\boldsymbol{\eta}} + \big(m+ReJE_{\boldsymbol{\eta}}\big)E_{\mathbf{v}}.$$

Applying Young's inequality, (36) yields

$$\frac{(1+m)H^0_{\delta_1}}{2}\|\nabla\mathbf{w}\|_2^2 + \frac{n_1H^0_{\delta_1}}{2}\|\nabla\boldsymbol{\phi}\|_2^2 \le \mathfrak{C}\Big\{Re(1+\delta_1)E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2^2 + \delta_1(n_2+n_3)\|\nabla\boldsymbol{\phi}\|_2^2 + \frac{(1+\delta_1)^2\big(m+ReJE_{\boldsymbol{\eta}}\big)^2}{(1+m)H^0_{\delta_1}}\|\nabla\boldsymbol{\phi}\|_2^2 + \left(\frac{\mathfrak{D}_1^2}{(1+m)H^0_{\delta_1}}+\frac{\mathfrak{D}_2^2}{n_1H^0_{\delta_1}}\right)\tilde{\delta}^2 + Re(E_p+E_q)E_{\mathbf{v}}|\tilde{\delta}|\Big\}$$

which implies

$$\left(\frac{(1+m)H^0_{\delta_1}}{2} - \mathfrak{C}Re(1+\delta_1)E_{\mathbf{v}}\right)\|\nabla\mathbf{w}\|_2^2$$
$$+\left(\frac{n_1 H^0_{\delta_1}}{2} - \mathfrak{C}\delta_1(n_2+n_3) - \mathfrak{C}\frac{(1+\delta_1)^2\left(m+ReJE_{\boldsymbol{\eta}}\right)^2}{(1+m)H^0_{\delta_1}}\right)\|\nabla\boldsymbol{\phi}\|_2^2$$
$$\leq \mathfrak{C}\left(\frac{\mathfrak{D}_1^2}{(1+m)H^0_{\delta_1}} + \frac{\mathfrak{D}_2^2}{n_1 H^0_{\delta_1}}\right)\tilde{\delta}^2 + \mathfrak{C}Re(E_p+E_q)E_{\mathbf{v}}|\tilde{\delta}|. \tag{37}$$

The desired result follows from (37) for sufficiently small values of Re, m, and δ_1. □

Remark 2. *In the particular case when $\delta_1 = 0$ and $\delta_2 \to 0$, inequality (37) implies that the solution $\mathbf{v}$ converges to $\mathbf{u}$ in $\mathcal{X}$ and $\boldsymbol{\eta}$ converges to $\boldsymbol{\nu}$ in $\left(H_0^1(\Omega)\right)^3$. Hence, we conclude that the solution of the curved pipe flow converges to that of the straight pipe flow when $\delta \to 0$.*

3.3 Continuous dependence on the Forchheimer coefficient (C_F)

Let $(\mathbf{u}, p, \boldsymbol{\nu})$ and $(\mathbf{v}, q, \boldsymbol{\eta})$ be the solutions of the BVP (1)-(4) corresponding to the Forchheimer coefficients C_{F_1} and C_{F_2}, respectively. Then the differences $(\mathbf{w}, \pi, \boldsymbol{\phi}) := (\mathbf{u}-\mathbf{v}, p-q, \boldsymbol{\nu}-\boldsymbol{\eta})$ solve the problem

$$\nabla\cdot(H\mathbf{w}) = 0, \tag{38}$$

$$ReH\Big((\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{v}\Big) - Re\delta u_3 w_3\hat{a}_2 - Re\delta v_3 w_3\hat{a}_2 + Re\delta u_3 w_2\hat{a}_3$$
$$+Re\delta w_3 v_2\hat{a}_3$$
$$= -ReH\nabla\pi + mH\nabla\times\boldsymbol{\phi} + m\delta\phi_3\hat{a}_1 + (1+m)\nabla\cdot(H\nabla\mathbf{w})$$
$$-(1+m)\frac{\delta^2}{H}(w_2\hat{a}_2 + w_3\hat{a}_3) - \sigma^2 H\mathbf{w} - Re\sigma C_{F_1}H(|\mathbf{u}|\mathbf{u} - |\mathbf{v}|\mathbf{v})$$
$$-\widetilde{C_F}Re\sigma H|\mathbf{v}|\mathbf{v}, \tag{39}$$

$$ReHJ\Big((\mathbf{u}\cdot\nabla)\boldsymbol{\phi}+(\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\Big)-ReJ\delta u_3\phi_3\hat{a}_2-ReJ\delta w_3\eta_3\hat{a}_2+ReJ\delta u_3\phi_2\hat{a}_3$$
$$+ReJ\delta w_3\eta_2\hat{a}_3$$
$$=-2mH\boldsymbol{\phi}+mH\nabla\times\mathbf{w}+m\delta w_3\hat{a}_1+n_1\nabla\cdot(H\nabla\boldsymbol{\phi})-n_1\frac{\delta^2}{H}(\phi_2\hat{a}_2$$
$$+\phi_3\hat{a}_3)+(n_2+n_3)H\nabla(\nabla\cdot\boldsymbol{\phi})+(n_2+n_3)\delta\nabla\phi_2-(n_2+n_3)\frac{\delta^2}{H}\phi_2\hat{a}_2 \quad (40)$$
$$\text{in } \Omega,$$

and

$$\mathbf{w}=\mathbf{0}=\boldsymbol{\phi} \text{ on } \partial\Omega, \quad (41)$$

where $\mathbf{w}=(w_1,w_2,w_3)$, $\boldsymbol{\phi}=(\phi_1,\phi_2,\phi_3)$, $w_i=u_i-v_i$, $\phi_i=\nu_i-\eta_i$ $(i=1,2,3)$ and $\widetilde{C_F}=C_{F_1}-C_{F_2}$.

The following theorem establishes the continuous dependence of the solution on the Forchheimer coefficient (C_F).

Theorem 3. *Suppose that* $(\mathbf{w},\pi,\boldsymbol{\phi})$ *solves the BVP* (38)-(41). *Then, for sufficiently small values of* Re, m *and* δ, $\mathbf{w}\to\mathbf{0}$ *in* $\mathcal{X}$ *and* $\boldsymbol{\phi}\to\mathbf{0}$ *in* $(H_0^1(\Omega))^3$ *as* $C_{F_2}\to C_{F_1}$.

Proof. Taking the $L^2(\Omega)$-inner product of (39) with $\mathbf{w}$ yields the estimate

$$(1+m)H_0\|\nabla\mathbf{w}\|_2^2$$
$$\le\mathfrak{C}\Big(Re(1+\delta)E_{\mathbf{v}}\|\nabla\mathbf{w}\|_2^2+m(1+\delta)\|\nabla\boldsymbol{\phi}\|_2\|\nabla\mathbf{w}\|_2+\widetilde{C_F}Re\sigma E_{\mathbf{v}}^2\|\nabla\mathbf{w}\|_2\Big). \quad (42)$$

Taking the $L^2(\Omega)$-inner product of (40) with $\boldsymbol{\phi}$, we obtain the estimate

$$n_1H_0\|\nabla\boldsymbol{\phi}\|_2^2\le\mathfrak{C}\Big((1+\delta)\big(m+ReJE_{\boldsymbol{\eta}}\big)\|\nabla\mathbf{w}\|_2\|\nabla\boldsymbol{\phi}\|_2+\delta(n_2+n_3)\|\nabla\boldsymbol{\phi}\|_2^2\Big). \quad (43)$$

Adding (42) and (43) and using the Young's inequality, we get

$$\left(\frac{(1+m)H_0}{2}-\mathfrak{C}Re(1+\delta)E_{\mathbf{v}}\right)\|\nabla\mathbf{w}\|_2^2$$
$$+\left(n_1H_0-\mathfrak{C}\delta(n_2+n_3)-\mathfrak{C}\frac{(1+\delta)^2\big(m+ReJE_{\boldsymbol{\eta}}\big)^2}{(1+m)H_0}\right)\|\nabla\boldsymbol{\phi}\|_2^2$$
$$\le\mathfrak{C}\frac{Re^2\sigma^2E_{\mathbf{v}}^4}{(1+m)H_0}\widetilde{C_F}^2. \quad (44)$$

The inequality (44) shows that $\|\nabla \mathbf{w}\|_2 \to 0$ and $\|\nabla \boldsymbol{\phi}\|_2 \to 0$ as $\widetilde{C_F} \to 0$ for sufficiently small values of Re, m and δ. Hence, the required result is proved. □

Remark 3. *For a particular case $C_{F_1} = 0$, when $C_{F_2} \to 0$, the inequality (44) implies that the solution to the Darcy–Lapwood –Forchheimer–Brinkman model for micropolar fluid flow in a curved pipe converges to that of the Darcy–Lapwood–Brinkman model.*

3.4 Continuous dependence on the micropolarity parameter (m)

Let $(\mathbf{u}, p, \boldsymbol{\nu})$ and $(\mathbf{v}, q, \boldsymbol{\eta})$ be the solutions of the BVP (1)-(4) with micropolarity parameters m_1 and m_2, respectively. The difference fields $(\mathbf{w}, \pi, \boldsymbol{\phi}) := (\mathbf{u} - \mathbf{v}, p - q, \boldsymbol{\nu} - \boldsymbol{\eta})$, then satisfy the problem

$$\nabla \cdot (H\mathbf{w}) = 0, \tag{45}$$

$$\begin{aligned} &ReH\Big((\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{v}\Big) - Re\delta u_3 w_3 \hat{a}_2 - Re\delta v_3 w_3 \hat{a}_2 + Re\delta u_3 w_2 \hat{a}_3 \\ &\qquad + Re\delta w_3 v_2 \hat{a}_3 \\ &= -ReH\nabla\pi + m_1 H\nabla\times\boldsymbol{\phi} + \widetilde{m}H\nabla\times\boldsymbol{\eta} + m_1\delta\phi_3\hat{a}_1 + \widetilde{m}\delta\eta_3\hat{a}_1 \\ &+(1+m_1)\nabla\cdot(H\nabla\mathbf{w}) + \widetilde{m}\nabla\cdot(H\nabla\mathbf{v}) - (1+m_1)\frac{\delta^2}{H}(w_2\hat{a}_2 + w_3\hat{a}_3) \\ &\qquad -\widetilde{m}\frac{\delta^2}{H}(v_2\hat{a}_2 + v_3\hat{a}_3) - \sigma^2 H\mathbf{w} - Re\sigma C_F H(|\mathbf{u}|\mathbf{u} - |\mathbf{v}|\mathbf{v}), \end{aligned} \tag{46}$$

$$\begin{aligned} &ReHJ\Big((\mathbf{u}\cdot\nabla)\boldsymbol{\phi} + (\mathbf{w}\cdot\nabla)\boldsymbol{\eta}\Big) - ReJ\delta u_3\phi_3\hat{a}_2 - ReJ\delta w_3\eta_3\hat{a}_2 + ReJ\delta u_3\phi_2\hat{a}_3 \\ &\qquad + ReJ\delta w_3\eta_2\hat{a}_3 \\ &= -2m_1 H\boldsymbol{\phi} - 2\widetilde{m}H\boldsymbol{\eta} + m_1 H\nabla\times\mathbf{w} + \widetilde{m}H\nabla\times\mathbf{v} + m_1\delta w_3\hat{a}_1 + \widetilde{m}\delta v_3\hat{a}_1 \\ &+n_1\nabla\cdot(H\nabla\boldsymbol{\phi}) - n_1\frac{\delta^2}{H}(\phi_2\hat{a}_2 + \phi_3\hat{a}_3) + (n_2+n_3)H\nabla(\nabla\cdot\boldsymbol{\phi}) \\ &\qquad +(n_2+n_3)\delta\nabla\phi_2 - (n_2+n_3)\frac{\delta^2}{H}\phi_2\hat{a}_2 \end{aligned} \tag{47}$$

in Ω,

and

$$\mathbf{w} = \mathbf{0} = \boldsymbol{\phi} \text{ on } \partial\Omega, \tag{48}$$

where $\mathbf{w} = (w_1, w_2, w_3)$, $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)$, $w_i = u_i - v_i$, $\phi_i = \nu_i - \eta_i$ $(i = 1, 2, 3)$ and $\widetilde{m} = m_1 - m_2$.

Our main result on the continuous dependence on the micropolarity parameter (m) is formulated below.

Theorem 4. *If $(\mathbf{w}, \pi, \boldsymbol{\phi})$ is the solution of (45)-(48), then $\mathbf{w} \to \mathbf{0}$ in $\mathcal{X}$ and $\boldsymbol{\phi} \to \mathbf{0}$ in $(H_0^1(\Omega))^3$ as $m_2 \to m_1$, provided the values of Re, m_1 and δ are sufficiently small.*

Proof. Adopting the same line of reasoning from the preceding results, we infer that

$$\begin{aligned}
&\left(\frac{(1+m_1)H_0}{2} - \mathfrak{C}Re(1+\delta)E_{\mathbf{v}}\right)\|\nabla\mathbf{w}\|_2^2 \\
&+\left(\frac{n_1H_0}{2} - \mathfrak{C}\delta(n_2+n_3) - \mathfrak{C}\frac{(1+\delta)^2\left(m_1 + ReJE_{\boldsymbol{\eta}}\right)^2}{(1+m_1)H_0}\right)\|\nabla\boldsymbol{\phi}\|_2^2 \\
&\leq \mathfrak{C}\left(\frac{\left(E_{\boldsymbol{\eta}}(1+\delta) + E_{\mathbf{v}}\left(1+\frac{\delta^2}{H_0}\right)\right)^2}{(1+m_1)H_0} + \frac{\left(E_{\boldsymbol{\eta}} + E_{\mathbf{v}}(1+\delta)\right)^2}{n_1H_0}\right)\widetilde{m}^2. \qquad (49)
\end{aligned}$$

It is clear from (49) that, for sufficiently small values of Re, m_1 and δ, the required result holds. □

Remark 4. *It is well known that in the absence of microrotational effects the system of micropolar equations reduced to the system of Navier-Stokes equations; it is also known that when the microrotation viscosity equals to zero, the linear velocity vector and the microrotation vector are uncoupled and the global motion is unaffected by the microrotation vector [8]. The same can be achieved from (49) by taking $m_1 = 0$ and allowing $m_2 \to 0$.*

References

[1] K.A. Ames and B. Straughan, *Non-Standard and Improperly Posed Problems*, Academic Press, San Diego, 1997.

[2] K.A. Ames and L.E. Payne, On stabilizing against modeling errors in a penetrative convection problem for a porous medium, *Math. Models Methods Appl. Sci.* 4 (1994), 733-740.

[3] N. Arada, On generalized Newtonian fluids in curved pipes, *SIAM J. Math. Anal.* 48 (2016), 1210-1249.

[4] F. Boyer and P. Fabrie, *Mathematical Tools for the Study of the Incompressible Navier-Stokes Equations and Related Models*, Springer, New York, 2012.

[5] H. Brezis, *Functional Analysis, Sobolev Space and Partial Differential Equations*, Springer, New York, 2011.

[6] A.O. Çelebi, V.K. Kalantarov and D. Uğurlu, Continuous dependence for the convective Brinkman–Forchheimer equations, *Appl. Anal.* 84 (2005), 877-888.

[7] A.O. Çelebi, V.K. Kalantarov and D. Uğurlu, On continuous dependence on coefficients of the Brinkman–Forchheimer equations, *Appl. Math. Lett.* 19 (2006), 801-807.

[8] A.C. Eringen, Theory of micropolar fluids, *J. Math. Mech.* 16 (1966), 1-18.

[9] M.H. Hamdan, Single-phase flow through porous channels a review of flow models and channel entry conditions, *Appl. Math. Comput.* 62 (1994), 203-222.

[10] C.T. Hsu and P. Cheng, Thermal dispersion in a porous medium, *Int. J. Heat Mass Transfer* 33 (1990), 1587-1597.

[11] Y. Liu, Convergence and continuous dependence for the Brinkman–Forchheimer equations, *Math. Comput. Modell.* 49 (2009), 1401-1415.

[12] G. Łukaszewicz, Long time behavior of 2D micropolar fluid flows, *Math. Comput. Modell.* 34 (2001), 487-509.

[13] A. Morro and B. Straughan, Continuous dependence on the source parameters for convective motion in porous media, *Nonlinear Anal.* 18 (1992), 307-315.

[14] L.E. Payne and B. Straughan, Structural stability for the Darcy equations of flow in porous media, *Proc. R. Soc. A* 454 (1974), 1691-1698.

[15] L.E. Payne and B. Straughan, Stability in the initial-time geometry problem for the Brinkman and Darcy equations of flow in porous media, *J. Math. Pures Appl.* 75 (1996), 225-271.

[16] L.E. Payne and B. Straughan, Analysis of the boundary condition at the interface between a viscous fluid and a porous medium and related modelling questions, *J. Math. Pures Appl.* 77 (1998), 317-354.

[17] L.E. Payne and B. Straughan, Convergence and continuous dependence for the Brinkman–Forchheimer equations, *Stud. Appl. Math.* 102 (1999), 419-439.

[18] A. Prabal and M. Devakar, On the existence, uniqueness and regularity of solutions of micropolar fluid flow through porous medium in a curved pipe, *Nonlinear Anal. Real World Appl.* 91 (2026), 104616.

[19] B. Straughan, *Stability and Wave Motion in Porous Media*, Applied Mathematical Sciences, Springer, New York, 2008.

Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 18, No. 4/2026

ON THE FALL SESSION OF THE ROMANIAN ITINERANT SEMINAR ON MATHEMATICS AND ITS APPLICATIONS (RISMA), 10 September 2026, Târgovişte, Romania

Gheorghe Moroşanu, founder and coordinator of RISMA*

Dedicated to Professor Constantin Fetecau on his 80th anniversary

DOI 10.56082/annalsarscimath.2026.4.411

1 Introduction

It is well known within the Romanian mathematical community that I launched in 2017 the Romanian Itinerant Seminar on Mathematical Analysis and Its Applications: www.cs.ubbcluj.ro/rismaa/. Since then, six successful annual sessions have taken place in various cities: Cluj-Napoca (2018), Constanţa (2019), Alba Iulia (2021), Braşov (2022), Craiova (2023), and Cluj-Napoca again in 2024.

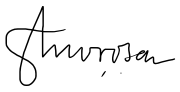
In 2024, I proposed renaming the seminar as *Romanian Itinerant Seminar on Mathematics and Its Applications* (RISMA): http://www.aosr.ro/en/romanian-itinerant-seminar-on-mathematics-and-applications-sirma/, expan ding its scope to include all areas of mathematics and its applications. Under this new format, the seminar sessions are held twice a year, as parts of the Spring and Fall Scientific Conferences organized by the Academy of Romanian Scientists (ARS), and hosted in the same cities where the ARS Spring and Fall Conferences of all ARS sections take place. More precisely, the Spring and Fall Conferences of the ARS Mathematics Section are replaced by extensions called *RISMA Spring and Fall Sessions*. The idea is to

*gheorghe.morosanu@ubbcluj.ro, *Babeş-Bolyai* University, Cluj-Napoca, Romania & Academy of Romanian Scientists, Bucharest

expand the ARS mathematics conferences to a wider audience. The papers resulting from this seminar are usually published in our journal, *Annals of the Academy of Romanian Scientists, Series on Mathematics and Its Applications*.

The participants to this RISMA Fall Session who presented their scientific results on this occasion are glad and honored to dedicate their corresponding abstracts below to Professor Emeritus Constantin Fetecau who turned 80 this year.

September 10, 2026 Gheorghe Moroșanu



2 Abstracts of the Presentations

The abstracts of the presentations are attached below, listed in alphabetical order by the presenters' names. Full papers may be obtained by contacting the authors directly at the email addresses provided.

• **Luminița BARBU**, *Ovidius* University of Constanța, Romania, Associate Member of the Academy of Romanian Scientists; email: lbarbu@univ-ovidius.ro, and **Gheorghe MOROŞANU**, *Babeş-Bolyai* University, Cluj-Napoca, Romania, Full Member of the Academy of Romanian Scientists, email: gheorghe.morosanu@ubbcluj.ro

Title: **On a Delay Second Order Integro-Differential Equation in a General Banach Space**

Abstract: Consider in a real Banach space X the following second order integro-differential equation

$$u''(t) = Au(t) + \int_{-r}^{0} g(s, u_t(s))\, ds + f(t),\ t \in [0, T], \quad (E)$$

$$u(0) = \phi(0),\ u'(0) = b,\ u_0 = \phi, \quad (IC)$$

where $A : D(A) \subset X \to X$ is a densely defined, closed linear operator generating a strongly continuous cosine family, g is a possibly unbounded

nonlinear function which is dominated in some sense by A, and for each $t \in [0,T]$, u_t denotes the function defined by $u_t(s) = u(t+s)$, $-r \leq s \leq 0$.

Existence and uniquennes of solutions for problem $(E)-(IC)$ are derived by using a fixed point approach combined with properties of the cosine family generated by A.

• **Vasile BERINDE**, Technical University of Cluj-Napoca, North University Center at Baia Mare, Romania, Full Member of the Academy of Romanian Scientists; email: vasile.berinde@gmail.com

Title: **Existence and Approximation of Solutions of Boundary Value Problems for Second Order Iterative Differential Equations**

Abstract: We study the existence and approximation of solutions to the second order iterative boundary-value problem

$$\ddot{x}(t) = f(t, x(t), x[2](t)), \quad 0 \leq t \leq 1,$$

where $x[2](t) := x(x(t))$, with solutions satisfying one of the two sets of boundary conditions:

$$(i) \quad x(0) = 0,\ x(1) = 1;$$

$$(ii) \quad x(0) = 1,\ x(1) = 0.$$

• **Aurelian CERNEA**, Faculty of Mathematics and Computer Science, University of Bucharest, Academiei 14, 010014 Bucharest, Corresponding Member of the Academy of Romanian Scientists; email: acernea@fmi.unibuc.ro

Title: **Existence of Solutions for Some Fractional Differential Inclusions**

Abstract: The talk was concerned with two classes of differential inclusions. The first problem is

$$D_C^q x(t) \in F(t, x(t), V(x)(t)) \quad a.e.\ ([0,1]) \qquad (a)$$

$$D_C^\sigma x(\xi) = \sum_{i=1}^{n} \alpha_i D_C^\nu x(\eta_i), \quad x(1) = \sum_{i=1}^{n} \beta_i \int_0^{\eta_i} x(s)ds + \sum_{i=1}^{n} \gamma_i x(\eta_i), \qquad (b)$$

where D_C^q is the Caputo fractional derivative of order $q \in (1,2]$, $0 < \sigma, \nu \leq 1$, $0 < \xi < \eta_1 < \eta_2 < ... < \eta_n < 1$, $\alpha_i, \beta_i, \gamma_i \in \mathbf{R}$, $i = \overline{1,n}$, $F : [0,1] \times \mathbf{R} \times \mathbf{R} \to \mathcal{P}(\mathbf{R})$ is a set-valued map and $V : C([0,1],\mathbf{R}) \to C([0,1],\mathbf{R})$ is a nonlinear Volterra integral operator defined by $V(x)(t) = \int_0^t k(t,s,x(s))ds$ with $k(.,.,.) : [0,1] \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ a given function. The second problem is

$$\begin{cases} D_C^{q_1} x_1(t) \in F_1(t, x_1(t), x_2(t)), & a.e.\ t \in I, \\ D_C^{q_2} x_2(t) \in F_2(t, x_1(t), x_2(t)), & a.e.\ t \in I, \end{cases} \tag{c}$$

$$\begin{cases} x_1(0) = 0, \quad x_1'(T) = 0, \quad x_1(T) = p_1 \int_0^\xi x_2(s)ds + Tp_2 x_2(\xi), \\ x_2(0) = 0, \quad x_2'(T) = 0, \quad x_2(T) = r_1 \int_0^\xi x_1(s)ds + Tr_2 x_1(\xi), \end{cases} \tag{d}$$

where $I = [0,T]$, $T > 0$, $q_1, q_2 \in (2,3)$, $\xi \in (0;T)$, $p_1, p_2, r_1, r_2 \in \mathbf{R}$, $F_i(.,.,.) : I \times \mathbf{R}^2 \to \mathcal{P}(\mathbf{R})$, $i = 1,2$ are given set-valued maps.

Existence results for problem $(a) - (b)$ and $(c) - (d)$ are provided when the values of the set-valued maps are not convex; instead these set-valued maps are assumed to be Lipschitz in state variables.

• **Eduard-Marius CRĂCIUN**, *Ovidius* University of Constanţa, Romania, Associate Member of the Academy of Romanian Scientists; email: emcraciun@yahoo.com

Title: **Cracks Interaction in Pre-Stressed Composites**

Abstract: We consider a pre-stressed, orthotropic, linear, elastic material representing a fiber reinforced composite. We suppose that the material is unbounded and contains two collinear and unequal cracks situated in the same plane, parallel to the reinforcing fibers and to the initial applied stresses. We assume that the upper and lower faces of the cracks are acted by symmetrically distributed incremental stresses. Our first aim is to determine the incremental elastic state produced in the body. To do this we use Guz's representation theorem. Using the theory of the Riemann-Hilbert problem we give the solution of our mathematical problem, assuming that the applied incremental stresses have a given constant value. Our second aim is to determine the critical values of the given incremental stress for which the tips of the cracks start to propagate. Our third aim is to study the interaction of two unequal cracks. According to the obtained results this interaction is strong only if the distance between the cracks is much more smaller as their length. In such a situation, the inner tips start to propagate first and the cracks tend to unify. If the distance is much grater as the length of

the cracks, the interaction is weak and the tips of the longest crack start to propagate first.

• **Vasile DRĂGAN**, *Simion Stoilow* Institute of Mathematics, Romania, Full Member of the Academy of Romanian Scientists; email: vasile.dragan@imar.ro, and **Samir ABERKANE**, Universite de Lorraine, CRAN, UMR Vandoeuvre-les-Nancy Cedex, France, and CNRS, CRAN, UMR 7039, France; email: samir.aberkane@univ-lorraine.fr

Title: **Coupled Matrix Riccati Differential Equations with Applications to the Solution of some Control Problems for Systems Modelled by Mean-field Stochastic Linear Differential Equations**

Abstract: We consider a class of coupled matrix Riccati differential equations involved in the solution of some linear quadratic optimal control problems for systems described by McKean-Vlasov linear stochastic differential equations also known as mean-field stochastic linear differential equations. For this kind of Riccati differential equations several special solutions defined on the whole real axis are studied. We refer to maximal and bounded solution and stabilizing and bounded solution. We shall provide necessary and sufficient conditions which guarantee the existence of that solutions. We emphasize the role of these solutions in the proof of the existence of an optimal control for an optimal control problem consisting of the minimization of a quadratic functional defined on the set of the trajectories of a system of McKean-Vlasov linear stochastic differential equations.

• **Teodor HAVÂRNEANU**, *Octav Mayer* Institute of Mathematics, Romanian Academy, Iaşi, Romania; email: havi@uaic.ro, and **Cătălin POPA**, *Octav Mayer* Institute of Mathematics, Romanian Academy, Iaşi, Romania

Title: **A Splitting Approximation Scheme for the Two-Dimensional Magnetohydrodynamic Equations**

Abstract: We consider the following two-dimensional magnetohydrodynamic equations:

$$\frac{\partial y}{\partial t} - \Delta y + (y \cdot \nabla) y + \nabla p + \nabla\left(\frac{1}{2}B^2\right) - (B \cdot \nabla)B = 0 \quad \text{in } Q = \Omega \times (0,T), \tag{1}$$

$$\frac{\partial B}{\partial t} + \widetilde{\operatorname{curl}}(\operatorname{curl} B) + (y \cdot \nabla)B - (B \cdot \nabla)y = 0 \quad \text{in } Q, \tag{2}$$

$$\operatorname{div} y = 0, \quad \operatorname{div} B = 0 \quad \text{in } Q, \tag{3}$$

$$y = 0, \qquad B \cdot N = 0, \quad \operatorname{curl} B = 0 \quad \text{on } \partial\Omega \times (0,T), \tag{4}$$

$$y(\cdot,0) = y_0, \quad B(\cdot,0) = B_0 \quad \text{in } \Omega, \tag{5}$$

where $y = (y_1, y_2) : \Omega \times [0,T] \to \mathbb{R}^2$ is the fluid velocity, $p : \Omega \times [0,T] \to \mathbb{R}$ is the scalar pressure, $B = (B_1, B_2) : \Omega \times [0,T] \to \mathbb{R}^2$ is the magnetic field, and $\Omega \subset \mathbb{R}^2$ is a bounded domain with a smooth boundary. We proposed a splitting approximation scheme of $(1) - (5)$ which consists of decoupling the Navier–Stokes equations and Maxwell equations which form the system $(1) - (5)$, on small time intervals. To compute the approximate solutions we start with a pair of initial velocity and magnetic fields and alternatively solve Navier–Stokes equations and dynamo equations. The main result states that for sufficiently smooth initial data the scheme converges in L^p with $p > 2$.

• **Liviu IGNAT**, *Simion Stoilow* Institute of Mathematics, Romania, Associate Member of the Academy of Romanian Scientists; email: liviu.ignat@gmail.com, and **Nicola de NITTI**, Polytechnic University of Bari, and **Fernando QUIROS**, Universidad Autonoma de Madrid

Title: **Convection Diffusion Problems with Non-Integrable Initial Data**

Abstract: We describe the long time behaviour of the solutions to

$$u_t - u_{xx} + \left(a|u|^{q-1}u\right)_x = 0 \ \text{ in } (0,\infty) \times \mathbb{R}, \qquad u(0) = u_0 \ \text{ on } \mathbb{R},$$

where $a \in \{\pm 1\}$, $q > 1$ and $u_0 \in L^\infty(\mathbb{R})$ need not be integrable, the model case being $u_0(x) \simeq k_\pm |x|^{-\gamma}$ as $x \to \pm\infty$, $0 < \gamma \le 1$. For integrable data the diffusive and the convective regimes are separated by $q = 2$; here the mass is infinite and the critical exponent becomes $q_c = 1 + 1/\gamma$.

Since $|x|^{-\gamma}$ lies in the weak (Marcinkiewicz) space $L^{1/\gamma,\infty}(\mathbb{R})$ but in no Lebesgue space, we work in $L^{p_0,\infty}(\mathbb{R}) \cap L^\infty(\mathbb{R})$, where we prove well-posedness and the sharp decay estimates

$$\|u(t)\|_{L^\infty(\mathbb{R})} \lesssim \begin{cases} \|u_0\|_{L^{p_0,\infty}(\mathbb{R})}\, t^{-\frac{1}{2p_0}}, & q \ge p_0 + 1, \\ \|u_0\|^{\frac{p_0}{p_0+q-1}}_{L^{p_0,\infty}(\mathbb{R})}\, t^{-\frac{1}{p_0+q-1}}, & 1 < q \le p_0 + 1, \end{cases}$$

parabolic in the first case and hyperbolic in the second, the latter following from Oleinik-type gradient bounds and multi-point estimates.

We then obtain the first term of the asymptotic expansion: for $\gamma \in (0,1)$, $k_\pm \neq 0$, $R > 0$,

$$\lim_{t\to\infty} \int_{|\xi|\le R} \left| t^\alpha u(t, t^\beta \xi) - U_\gamma(\xi) \right| d\xi = 0, \qquad (\alpha, \beta) = \left(\tfrac{\gamma}{1+(q-1)\gamma_*}, \tfrac{1}{1+(q-1)\gamma_*} \right),$$

with $\gamma_* = \gamma$ for $q < q_c$ and $\gamma_* = 0$ otherwise; here U_γ is the profile of the unique self-similar solution with datum $k\,\mathrm{sgn}x|x|^{-\gamma}$ of the heat equation, of the complete equation or of the pure conservation law, according to the regime, its uniqueness in the convective range being the delicate point.

• **Aurelian ISAR**, Faculty of Physics, University of Bucharest; *Horia Hulubei* National Institute of Physics and Nuclear Engineering, Măgurele; Full Member of the Academy of Romanian Scientists; email: isar@theory.nipne.ro, and **Alina STOICA**, Faculty of Physics, University of Bucharest; *Horia Hulubei* National Institute of Physics and Nuclear Engineering, Măgurele; email: dobre.alina@theory.nipne.ro

Title: **Generation of Quantum Correlations in Gaussian Quantum Systems**

Abstract: Hellinger geometric quantum discord and interferometric power are investigated for an open system consisting of two bosonic modes interacting with four different environments: vacuum, squeezed vacuum, thermal and squeezed thermal, taking the initial state of the system either a single-mode squeezed state, which presents no correlations, or a squeezed vacuum state. The time evolution of the correlations is investigated in the framework of the theory of open systems, based on completely positive quantum dynamical semigroups, using the Gorini-Kossakowski-Lindblad-Sudarshan master equation. We show that both quantum correlations can be generated from an initial factorized state and even amplified. For the squeezed vacuum environment, the competition between the influences exerted by the squeezing of the initial state and the squeezing of the environment leads to the decreasing, and even destruction of the correlations, or to their enhancement. In the case of an initial squeezed vacuum state in interaction with the squeezed vacuum environment, for specific conditions one can achieve decoherence-free states.

• **Liviu IXARU**, *Horia Hulubei* National Institute of Physics and Nuclear Engineering, Department of Theoretical Physics, Bucharest, Romania, Full

Member of the Academy of Romanian Scientists; email: ixaru@theory.nipne.ro

Title: **On the Visualization of an Atomic Nucleus**

Abstract: The image of an atom, based on the well-known planetary model of Bohr, is so familiar to the general public that we find it even just on the logo of various institutions in the country or abroad. The situation is different for the atomic nucleus since this is typically based on a simple but naive model such that a modification is needed for a more realistic base. We opt for the nuclear shell model, a standard tool in the current nuclear physics. We examine this approach to present its results by a series of graphs and animations. A series of novelties and surprises will arise. Our presentation is intended to be of use for both specialists and nonspecialists.

• **Mircea LUPU**, *Transilvania* University of Braşov, Romania, Corresponding Member of the Academy of Romanian Scientists; email: emlupu 2006@yahoo.com

Title: **New Methods of Optimization Airfoils, Aerodynamics**

Abstract: The paper addresses direct and inverse boundary value problems, effectively deriving solutions for nonlinear functionals or integral equations. It examines the dynamics of symmetric aerodynamic profiles within an ideal fluid, subjected to the action of either an unbounded or a physically bounded jet. Constitutive equations are formulated in an auxiliary plane, and profile designs are determined based on criteria for resistance or extreme lift. Solutions are obtained using analytical methods, inequalities, or extremal analysis.

• **Constantin P. NICULESCU**, Department of Mathematics, University of Craiova, Romania, Full Member of the Academy of Romanian Scientists; email: constantin.p.niculescu@gmail.com

Title: **In Memoriam, Sorin G. Gal**

Abstract: This talk was dedicated to the memory of my dear friend and collaborator Sorin G. Gal (August 23, 1953 - October 18, 2024). This summarizes the dramatic story behind our latest publication—the book on the Choquet integral—which he initiated as a project in 2023.

• **Adrian PETRUŞEL**, *Babeş-Bolyai* University, Cluj-Napoca, Romania, Full Member of the Academy of Romanian Scientists; email: adrian.petrusel @ubbcluj.ro, and **Gabriela PETRUŞEL**, *Babeş-Bolyai* University, Cluj-Napoca, Romania; email: gabriela.petrusel@ubbcluj.ro

Title: **Some Fixed Point Theorems for Nonself Operators in Metric and b-Metric Spaces**

Abstract: In this paper, we will present a synthesis on some fixed point results for nonself operators. The outline of the paper is the following one:
- Local fixed point theorems;
- Graph contractions;
- Contractions outside a bounded set;
- Fixed point results in b-metric spaces Applications.

• **Savin TREANŢĂ**, National University of Science and Technology Politehnica Bucharest, Romania, Corresponding Member of the Academy of Romanian Scientists; email: savin.treanta@upb.ro

Title: **On Some Constrained Optimization Problems via Gradient Type Variations**

Abstract: The aim of this presentation is the optimization of constrained integral-type functionals by considering gradient-type variation functions. Concretely, the associated necessary conditions of optimality are stated by applying appropriate mathematical pipelines. The main results highlight the nonhomogeneous Euler-Lagrange-type partial differential equations. Compared to classical variations, in this context we are led to controlled right-hand member equations, creating a novel perspective in the field of variational calculus. More precisely, we want to determine the pair (φ^*, u^*) that extremizes the following functional given by a simple integral

$$M\left(\varphi(\cdot), u(\cdot)\right) = \int_{\omega_0}^{\omega_1} G\left(\varphi(\omega), \dot{\varphi}(\omega), \ddot{\varphi}(\omega), u(\omega), \omega\right) d\omega$$

between all functions φ of C^3-class that meet the conditions

$$\varphi(\omega_0) = \varphi_0, \quad \varphi(\omega_1) = \varphi_1, \quad \dot{\varphi}(\omega_0) = \tilde{\varphi}_0, \quad \dot{\varphi}(\omega_1) = \tilde{\varphi}_1,$$

subject to the following ordinary differential inequalities (ODIs) and ordinary differential equations (ODEs)

$$g^{\alpha}\left(\varphi(\omega), \dot{\varphi}(\omega), \ddot{\varphi}(\omega), u(\omega), \omega\right) \leq 0, \quad \alpha = \overline{1, o},\ o \leq n,$$

$$\chi^{\mu}\left(\varphi(\omega),\dot{\varphi}(\omega),\ddot{\varphi}(\omega),\dddot{\varphi}(\omega),u(\omega),\omega\right)$$
$$:=\frac{d^{3}\varphi^{\mu}}{d\omega^{3}}-Z^{\mu}\left(\varphi(\omega),\dot{\varphi}(\omega),\ddot{\varphi}(\omega),u(\omega),\omega\right)=0,\quad \mu=\overline{1,n}$$

(see α as the index set labelling the ODI constraints and μ the index corresponding to each ODE constraint), by employing gradient-type variation functions subject to the boundary conditions.