AIMS Mathematics, 11(11): 36026–36064
DOI:10.3934/math.20261395
Received: 06 July 2026
Revised: 31 August 2026
Accepted: 21 October 2026
Published: 03 November 2026

https://www.aimspress.com/journal/Math

Research article

Physics-informed machine learning approach for radiative MHD dusty Maxwell nanofluid flow over a stretching/shrinking surface in a non-Darcy porous medium: A case of dual solutions

Bilal Ahmad[1,*], Muhammad Abubakar[2], Amber Sultan[3], Syed Baqar Hussain[3] and Yazeed Yasin Ghadi[4]

[1] Department of Mathematics and Statistics, The University of Lahore, Lahore, Pakistan
[2] Department of Computer Science, Lahore Garrison University, Lahore, Pakistan
[3] School of Systems and Technology, University of Management and Technology, Lahore, Pakistan
[4] Department of Computer Science, Al Ain University, Al Ain, United Arab Emirates

* **Correspondence:** Email: bilal.math@outlook.com.

Abstract: Wire coating, fiber drawing, and the extrusion of filled thermoplastics all draw a slender continuous filament through a hot, particle-laden viscoelastic melt, often in a packed die and under an imposed magnetic field. This paper formulates that configuration as a steady two-phase boundary layer flow of a radiative magnetohydrodynamic (MHD) dusty Maxwell nanofluid over a stretching or shrinking cylinder in a Darcy–Forchheimer porous medium, retaining a micrometer-scale dust phase with velocity and thermal slip alongside a Buongiorno nanoparticle description, nonlinear Rosseland radiation, and an Arrhenius reaction. Similarity reduction yields a tenth-order system in which transverse curvature enters through the parameter γ, multiplying every diffusion operator by $(1 + 2\gamma\eta)$ and adding a viscoelastic term with no plane-sheet analog. A collocation solver with fold-following continuation, verified against published stretching cylinder data to within 0.004% at matched truncation, resolves both solution branches for the decelerating suction-assisted regime. Curvature is found to govern the problem principally through feasibility: Because the coefficient of the highest derivative is $(1 + 2\gamma\eta) - \beta f^2$, curvature relaxes the condition for a physically admissible solution, enlarging the admissible region of the suction–Deborah plane by about two-thirds at $\gamma = 0.5$ and doubling or tripling the largest attainable Deborah number at fixed suction. It also contracts the range of shrinking rates supporting dual solutions, moving the critical value from $\lambda_c = -1.8336$ on a plane sheet to -1.6840 at $\gamma = 0.5$, with the contraction strengthening as suction increases. A linear temporal stability analysis of the complete disturbance system confirms the upper branch to be stable and the lower to be unstable; the linearized operator is shown to be block lower-triangular, so the customary restriction to the momentum–dust subsystem is exact rather than assumed. Finally, a machine learning surrogate is assessed against classical alternatives on equal data. A Gaussian process built on 10 reference solutions attains 0.18% accuracy for under three seconds of construction and repays itself

after 11 evaluations, against roughly 875 for a network trained on the differential residual. The same Gaussian process degrades by nearly three orders of magnitude approaching the turning point, erring by 12.3% at a distance of 0.036 from the fold, where the branch has a vertical tangent and smooth interpolation is the wrong hypothesis class. The case for embedding the governing equations in a surrogate rests on that neighborhood, and not on the cost.

Keywords: stretching surface; curvature parameter; dusty Maxwell nanofluid; dual solutions; temporal stability; physics-informed neural network
Mathematics Subject Classification: 76A05, 76T15, 76W05, 76S05, 65L10, 68T07

Nomenclature

$\hat{x}, \hat{r}$	Axial, radial coordinate (m)	Pr	Prandtl number
$\hat{u}, \hat{v}$	Fluid velocity components ($\mathrm{m\,s^{-1}}$)	R_d	Radiation parameter
$\hat{u}_p, \hat{v}_p$	Dust's velocity components ($\mathrm{m\,s^{-1}}$)	θ_w	Temperature ratio T_w/T_∞
R	Cylinder radius (m)	N_b	Brownian motion parameter
γ	Curvature parameter $R^{-1}\sqrt{\nu/a}$	N_t	Thermophoresis parameter
f, g	Dimensionless stream functions	Q	Heat source/sink parameter
θ, θ_p	Fluid and dust temperature (dimensionless)	Sc	Schmidt number ν/D_B
φ	Nanoparticle concentration (dimensionless)	σ_c	Reaction rate parameter k_r/a
λ	Stretching/shrinking parameter	E	Activation energy parameter
λ_c	Critical (fold) value of λ	n	Reaction rate exponent
S	Suction parameter	l	Dust's mass concentration ρ_p/ρ
β	Maxwell (Deborah) number	β_v, β_t	Momentum, thermal interaction
M	Magnetic parameter	Γ	Specific-heat ratio c_s/c_p
K_p	Porosity parameter	Re_x	Local Reynolds number
F_r	Local Forchheimer parameter	ω	Temporal growth rate (eigenvalue)
C_f	Skin friction coefficient	η	Similarity variable
Nu_x	Nusselt number	η_∞	Domain truncation
Sh_x	Sherwood number	$\Lambda_1, \Lambda_2, \Lambda_3$	Coded response surface methodology (RSM) factors
T	Fluid temperature (K)	T_w, T_∞	Wall and ambient temperature (K)
C	Nanoparticle concentration ($\mathrm{mol\,m^{-3}}$)	C_w, C_∞	Wall and ambient concentration ($\mathrm{mol\,m^{-3}}$)
a	Reference stretching rate ($\mathrm{s^{-1}}$)	V_0	Suction velocity ($\mathrm{m\,s^{-1}}$)
B_0	Magnetic flux density (T)	k	Permeability ($\mathrm{m^2}$)
C_b	Forchheimer drag coefficient ($\mathrm{m^{-1}}$)	κ	Thermal conductivity ($\mathrm{W\,m^{-1}K^{-1}}$)
σ	Electrical conductivity ($\mathrm{S\,m^{-1}}$)	ρ, ρ_p	Fluid and dust density ($\mathrm{kg\,m^{-3}}$)
μ	Dynamic viscosity (Pa s)	ν	Kinematic viscosity ($\mathrm{m^2\,s^{-1}}$)
c_p, c_s	Specific heats: fluid, dust ($\mathrm{J\,kg^{-1}K^{-1}}$)	D_B	Brownian diffusivity ($\mathrm{m^2\,s^{-1}}$)
D_T	Thermophoretic diffusivity ($\mathrm{m^2\,s^{-1}}$)	τ_m, τ_T	Momentum, thermal relaxation time (s)
λ_M	Maxwell relaxation time (s)	k_r	Reaction rate constant ($\mathrm{s^{-1}}$)
E_a	Molar activation energy ($\mathrm{J\,mol^{-1}}$)	R_g	Universal gas constant ($\mathrm{J\,mol^{-1}K^{-1}}$)

σ^*	Stefan–Boltzmann constant ($\mathrm{W\,m^{-2}K^{-4}}$)	k^*	Mean absorption coefficient ($\mathrm{m^{-1}}$)
Q_0	Heat source/sink coefficient ($\mathrm{W\,m^{-3}K^{-1}}$)	St	Stokes number
α	Thermal diffusivity ($\mathrm{m^2\,s^{-1}}$)	τ_n	Nanoparticle-to-fluid heat capacity ratio

1. Introduction

Training a network to satisfy a differential system itself, rather than to fit samples of its solution, has become a workable alternative to conventional discretization for nonlinear transport problems. The construction is due to Raissi et al. [1]: The residual of the governing equations is evaluated at collocation points scattered through the domain and used as the training objective, so that the physics constrains the fit everywhere and not only where data happen to exist. Two properties recommend the approach for thermal engineering design. The first is that the operating parameters may be taken as inputs, so that one fit answers queries across a design space for which a boundary value solver would otherwise have to be called repeatedly [2]. The second is that a residual objective does not require the solution to be marched from an initial profile, which matters where the boundary value problem has more than one solution, and a shooting method converges to a given one only if it is started close to it. Both properties bear directly on the flow considered here, in which withdrawal of fluid through a shrinking wall produces a finite band of shrinking rates supporting two steady states. Whether they are delivered in practice, and how such a surrogate stands against the classical interpolants that the same reference data would support is one of the questions this paper sets out to answer.

Continuously drawing a slender cylindrical body through a hot, particle-laden polymer melt is the common mechanical core of several manufacturing routes: Wire and cable coatings, optical- and glass-fiber drawing, and the extrusion of filled thermoplastics through annular dies. In each of these, the product is a cylinder rather than a plane sheet, the surrounding melt is viscoelastic rather than Newtonian, and the melt carries a solid dispersed load of pigment, filler, or reinforcing powder whose particles are massive enough to lag the carrier fluid. The quality of the finished product depends on the wall shear stress and the heat and mass transfer rates at the moving surface, and these, in turn, depend on the coupled action of the geometry, the rheology, and the suspended phase.

Several further effects enter once the process conditions are realistic. When the die is packed or the preform fibrous, the flow resistance ceases to be Darcy-linear and acquires a quadratic Forchheimer contribution. When the melt is electrically conductive, an imposed magnetic field is used to damp draw resonance and to stabilize the line. When the line passes through a furnace section, radiative transport competes with conduction, and the temperature ratio across the layer is large enough that a linearized Rosseland approximation is inadequate. When the coating cures, a temperature-activated reaction consumes the diffusing species. A model intended to describe such a line must therefore carry the cylindrical geometry, viscoelasticity, a second particulate phase, non-Darcy drag, nonlinear radiation and an Arrhenius reaction simultaneously, and the present study assembles that combination.

The operating regime of interest is the decelerating one. When the surface is shrunk rather than stretched and mass is withdrawn through it, the boundary layer admits more than one steady state over a finite range of shrinking rates, terminating at a critical value beyond which no similarity solution exists. That critical value delimits the operating window of the process, and locating it, establishing which of the coexisting states is physically realizable, and determining how the geometry moves it are the questions this paper addresses.

1.1. Literature review

Boundary layer flow along a stretching cylinder differs from the plane-sheet problem through the transverse curvature, which enters the similarity reduction as the parameter $\gamma = R^{-1}\sqrt{\nu/a}$ and multiplies every diffusion operator by $(1 + 2\gamma\eta)$. The configuration has been examined for Newtonian and generalized Newtonian fluids: Poply et al. [3] treated the radiative laminar layer, Manjunatha et al. [4] added a dusty phase in a porous medium, and Hashim et al. [5] and Mahdy et al. [6] examined Carreau fluids with melting and with activation energy, respectively. Abbas et al. [7] added Darcy resistance, radiation and a chemical reaction on the same geometry, and comparable transport with a reacting species has been reported for stratified sheets [8]. These studies are confined to the stretching case $\lambda > 0$, where the similarity solution is unique.

Viscoelasticity is represented here by the upper-convected Maxwell model, which introduces a single relaxation time and has been widely applied to polymeric boundary layers since the stagnation point analysis of Sadeghy et al. [9]. Nanoparticle transport follows the two-component formulation of Buongiorno [10], in which Brownian diffusion and thermophoresis appear as drift fluxes; the corresponding boundary layer problem was solved by Khan and Pop [11]. Two-phase dusty suspensions, in which micrometer-scale particles slip relative to the carrier and require their own momentum and thermal balances, have been studied for dusty nanoliquids under quadratic radiation and convection [12] and, on the cylindrical geometry considered here, for magnetohydrodynamic (MHD) dusty Casson nanofluids with a variable heat source and the Arrhenius reaction [13]. The suspension modeled in the present work contains both populations, which are physically distinct and are treated distinctly; the separation of scales that justifies this is set out in Section 2.

Nonuniqueness in boundary layers over shrinking surfaces with wall suction has been recognized since Miklavčič and Wang [14], and dual solutions have since been reported for a wide range of configurations, including Maxwell ternary nanofluids past a stretching/shrinking sheet [15] and, on the stretching/shrinking cylinder considered here, for magnetized nanofluids with velocity slip [16]. Which branch is physically realizable is settled by the temporal stability argument introduced by Merkin [17] and developed by Weidman et al. [18]: A disturbance proportional to $e^{-\omega\tau}$ is superposed on the base state, and the sign of the smallest eigenvalue ω decides the outcome. Such analyses are commonly restricted to the momentum subsystem; whether that restriction is exact for a coupled two-phase system is examined in Section 3.2.

Physics-informed neural networks [1] embed a differential residual in the training objective and have been applied across the physical sciences [19] and to problems admitting multiple solutions [20]. Neural surrogates have likewise been used for nanofluid boundary layer transport on stretching surfaces [21] and in Darcy–Forchheimer media [22]. Their training is not without difficulty: Stiffness in the residual produces unbalanced back-propagated gradients, which is a documented failure mode [23]. Methodological development continues: Enriched formulations have been proposed for coupled transport systems [24], and fundamental solution networks for inverse problems governed by the Laplace and biharmonic operators [25]. Such networks are frequently proposed as parametric surrogates, and the question of whether they outperform classical interpolation on that task, and where they do so, is taken up in Section 4.7.

1.2. Significance of the study

The configuration matters because the quantities it determines are the ones that determine product quality. The wall shear stress fixes the draw tension and therefore the risk of filament breakage; the wall heat flux fixes the rate at which the coating is cooled and hence its final microstructure; and the wall mass flux fixes how far the reacting species penetrates before the cure is complete. All three are governed by the same tenth-order system, and all three respond to the geometry through the single curvature group γ.

The decelerating regime is significant for a further reason. Where two steady states coexist, the state actually realized depends on how the operating point was approached, and a process crossing the critical shrinking rate loses its steady solution entirely. Locating that boundary, and establishing which of the coexisting states is physically realizable, is therefore a prerequisite for operating near it.

1.3. Research gap

Three gaps follow from the literature reviewed above. First, flow over a stretching cylinder has been treated only for the accelerating case $\lambda > 0$, where the similarity solution is unique; the shrinking branch, on which multiplicity arises, has not been mapped for a viscoelastic two-phase suspension on a cylinder. Second, the admissibility of the Maxwell formulation itself, which requires the coefficient of the highest derivative to remain of one sign, is rarely examined, although it bounds the Deborah number that any such computation may legitimately report. Third, machine learning surrogates for this class of problem are commonly proposed without comparison against the classical interpolants that the same reference data would support, so the conditions under which they are preferable remain unclear.

1.4. Novelty and contributions

This paper formulates the two-phase problem described above on a stretching or shrinking cylinder, reduces it to a tenth-order similarity system, resolves both solution branches by fold-following continuation, determines their stability, and assesses a machine learning surrogate against classical alternatives. The specific contributions are as follows.

- An admissibility criterion for the Maxwell coefficient, $(1 + 2\gamma\eta) - \beta f^2 > 0$, which bounds the physically attainable Deborah number, together with a map of the resulting operating envelope in the suction–Deborah plane. Curvature is shown to enlarge that envelope by about two-thirds, so that operating points unreachable on a plane sheet become attainable on a cylinder.
- Identification of the curvature parameter γ as a control on the dual-solution structure, with the critical shrinking rate $\lambda_c(\gamma)$ mapped and its interaction with suction quantified.
- A dimensionally consistent formulation of radiative MHD dusty Maxwell nanofluid flow over a stretching/shrinking cylinder in a Darcy–Forchheimer medium, in which the two solid populations are distinguished by their Stokes number and the Forchheimer group is defined as an explicitly local quantity.
- A complete linear temporal stability analysis of the coupled disturbance system, including a proof that its operator is block lower-triangular and hence that the momentum–dust reduction is exact, with grid convergence and branch-resolved eigenvalues.
- An assessment of a machine learning surrogate against Gaussian process and radial basis function (RBF) interpolation on identical data, with per-variable errors including the dust phase fields,

predictions at held-out shrinking rates near the turning point, and break-even costing.

The remainder of the paper is organized as follows. Section 2 sets out the model and its similarity reduction. Section 3 describes the numerical treatment, the stability formulation, and the verification suite. Section 4.3 presents the curvature control of the dual-solution structure, and Section 4.5 presents the transport fields. Section 4.7 reports the surrogate study, Section 4.8 outlines a compact parametric sensitivity analysis, and Section 5 concludes the paper.

2. Mathematical formulation

Consider a steady, axisymmetric, incompressible boundary-layer flow along a circular cylinder of radius R aligned with the axial coordinate $\hat{x}$, the radial coordinate being $\hat{r}$. The cylinder's surface is stretched or shrunk axially with a velocity $\hat{u}_w(\hat{x}) = \lambda a\hat{x}$, where $a > 0$ is the reference rate and λ the stretching/shrinking parameter ($\lambda > 0$ stretching, $\lambda < 0$ shrinking). A uniform radial suction $-V_0$ and a uniform transverse magnetic field B_0 are applied. The configuration is shown in Figure 1. The model rests on the following assumptions.

- The flow is steady, axisymmetric, laminar, and incompressible, and the boundary layer approximation applies.
- The carrier obeys the upper-convected Maxwell model with the relaxation time λ_M.
- The suspension carries two dilute solid populations: Micrometer-scale dust particles of uniform size and constant number density, treated as a separate phase with momentum and thermal relaxation times τ_m and τ_T, and nanoparticles, treated by the Buongiorno description. Table 1 states the separation of scales.
- The porous medium follows the Darcy–Forchheimer relation.
- Radiative transport is modeled by the nonlinear Rosseland approximation.
- A first-order reaction with molar Arrhenius activation energy is present.
- The magnetic Reynolds number is small, so the induced field is neglected.

Table 1. Separation of scales between the two solid populations.

	Dust particles	Nanoparticles
Characteristic size	1–100 μm	10–100 nm
Stokes number	$St = O(1)$, slips	$St \ll 1$, no slip
Representation	Separate phase (τ_m, τ_T)	Buongiorno drift
Governed fields	g, θ_p	N_b, N_t via φ
Volume fraction	Dilute; ρ, ν unaffected	Dilute; enters via N_b, N_t

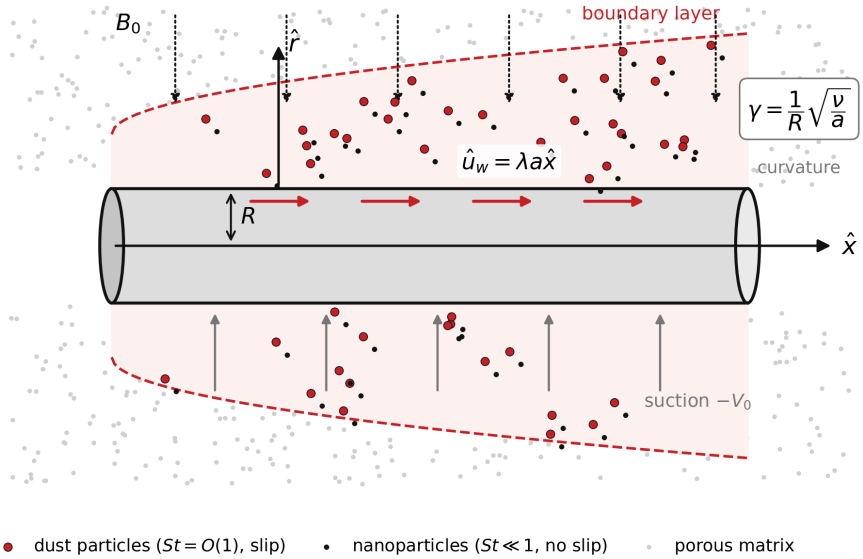


Figure 1. Stretching/shrinking cylinder of radius R in a non-Darcy porous medium, carrying a dusty Maxwell nanofluid under a transverse magnetic field with radial suction.

For the fluid phase, we have

$$\frac{\partial(\hat{r}\hat{u})}{\partial\hat{x}} + \frac{\partial(\hat{r}\hat{v})}{\partial\hat{r}} = 0, \tag{2.1}$$

$$\hat{u}\hat{u}_{\hat{x}} + \hat{v}\hat{u}_{\hat{r}} + \lambda_M\left(\hat{u}^2\hat{u}_{\hat{x}\hat{x}} + \hat{v}^2\hat{u}_{\hat{r}\hat{r}} + 2\hat{u}\hat{v}\,\hat{u}_{\hat{x}\hat{r}}\right) = \frac{\nu}{\hat{r}}\frac{\partial}{\partial\hat{r}}\left(\hat{r}\frac{\partial\hat{u}}{\partial\hat{r}}\right) + \frac{\rho_p}{\rho\tau_m}(\hat{u}_p - \hat{u}) - \frac{\sigma B_0^2}{\rho}\hat{u} - \frac{\nu}{k}\hat{u} - \frac{C_b}{\sqrt{k}}\hat{u}^2, \tag{2.2}$$

$$\hat{u}T_{\hat{x}} + \hat{v}T_{\hat{r}} = \frac{1}{\rho c_p\hat{r}}\frac{\partial}{\partial\hat{r}}\left[\hat{r}\left(\kappa + \frac{16\sigma^* T^3}{3k^*}\right)T_{\hat{r}}\right] + \tau_n\left[D_B C_{\hat{r}}T_{\hat{r}} + \frac{D_T}{T_\infty}T_{\hat{r}}^2\right] + \frac{Q_0}{\rho c_p}(T - T_\infty) + \frac{\rho_p c_s}{\tau_T\rho c_p}(T_p - T), \tag{2.3}$$

$$\hat{u}C_{\hat{x}} + \hat{v}C_{\hat{r}} = \frac{D_B}{\hat{r}}\frac{\partial}{\partial\hat{r}}(\hat{r}C_{\hat{r}}) + \frac{D_T}{T_\infty\hat{r}}\frac{\partial}{\partial\hat{r}}(\hat{r}T_{\hat{r}}) - k_r(C - C_\infty)\left(\frac{T}{T_\infty}\right)^n\exp\left(-\frac{E_a}{R_g T}\right), \tag{2.4}$$

and for the dust phase, we have

$$\frac{\partial(\hat{r}\hat{u}_p)}{\partial\hat{x}} + \frac{\partial(\hat{r}\hat{v}_p)}{\partial\hat{r}} = 0, \quad \hat{u}_p\hat{u}_{p,\hat{x}} + \hat{v}_p\hat{u}_{p,\hat{r}} = \frac{1}{\tau_m}(\hat{u} - \hat{u}_p), \quad \hat{u}_p T_{p,\hat{x}} + \hat{v}_p T_{p,\hat{r}} = -\frac{1}{\tau_T}(T_p - T). \tag{2.5}$$

The boundary conditions are

$$\hat{r} = R:\ \hat{u} = \lambda a\hat{x},\ \hat{v} = -V_0,\ T = T_w,\ C = C_w; \qquad \hat{r}\to\infty:\ \hat{u}, \hat{u}_p \to 0,\ \hat{v}_p \to \hat{v},\ T, T_p \to T_\infty,\ C \to C_\infty. \tag{2.6}$$

Two points of dimensional bookkeeping are noted here. First, k_r is a first-order rate constant with units s^{-1} and enters the reaction term linearly; writing k_r^2 is dimensionally inadmissible. Second, E_a as tabulated for cure and cross-linking reactions is a *molar* quantity, so the Arrhenius exponent must be formed with the universal gas constant R_g rather than the Boltzmann constant.

2.1. Term-wise interpretation

In the momentum equation, the convective inertia is balanced by the viscous term, whose cylindrical form supplies both the variable coefficient $(1 + 2\gamma\eta)$ and the additional first-derivative contribution $2\gamma(\cdot)'$; by the interphase Stokes drag $\rho_p(\hat{u}_p - \hat{u})/(\rho\tau_m)$ that couples the carrier to the dust; by the Lorentz retardation $-\sigma B_0^2\hat{u}/\rho$; by the Darcy resistance $-\nu\hat{u}/k$ and the quadratic Forchheimer drag $-C_b\hat{u}^2/\sqrt{k}$; and by the Maxwell bracket, which represents stress relaxation and is the only term carrying λ_M.

In the energy equation, the divergence term carries Fourier conduction augmented by the nonlinear Rosseland radiative flux, so that the effective conductivity varies with the local temperature through Θ^3 rather than being linearized about the ambient value. It is followed by the Brownian $(D_B C_{\hat{r}} T_{\hat{r}})$ and thermophoretic $(D_T T_{\hat{r}}^2/T_\infty)$ energy fluxes of the Buongiorno description, a volumetric heat source or sink, and the interphase thermal exchange with the dust.

The concentration equation combines Brownian and thermophoretic diffusion with an Arrhenius-type reaction whose rate depends on the local temperature through both the power law $(T/T_\infty)^n$ and the exponential factor. The dust phase equations express the momentum and thermal equilibration of the particles with the carrier through the relaxation times τ_m and τ_T; they contain no diffusion operator and are therefore free of curvature terms.

2.2. Similarity reduction

Introduce

$$\eta = \frac{\hat{r}^2 - R^2}{2R}\sqrt{\frac{a}{\nu}}, \qquad \psi = \sqrt{a\nu}\,\hat{x}Rf(\eta), \qquad \gamma = \frac{1}{R}\sqrt{\frac{\nu}{a}}, \tag{2.7}$$

so that $\hat{u} = \hat{r}^{-1}\psi_{\hat{r}} = a\hat{x}f'(\eta)$ and $\hat{v} = -\hat{r}^{-1}\psi_{\hat{x}} = -(R/\hat{r})\sqrt{a\nu}f(\eta)$, with the corresponding dust phase forms $\hat{u}_p = a\hat{x}g'$, $\hat{v}_p = -(R/\hat{r})\sqrt{a\nu}g$, together with $T = T_\infty[1 + (\theta_w - 1)\theta]$, $T_p = T_\infty[1 + (\theta_w - 1)\theta_p]$, and $\varphi = (C - C_\infty)/(C_w - C_\infty)$. The identity

$$\frac{\hat{r}^2}{R^2} = 1 + 2\gamma\eta \tag{2.8}$$

converts every radial operator into its similarity form, and both continuity equations are satisfied identically. Writing $\Theta = 1 + (\theta_w - 1)\theta$ and $N(\theta) = 1 + R_d\Theta^3$, the system reduces to

$$(1 + 2\gamma\eta)f''' + 2\gamma f'' + ff'' - f'^2 - \beta\left[f^2 f''' - 2ff'f'' + \frac{\gamma f^2 f''}{1 + 2\gamma\eta}\right] - (M + K_p)f' - F_r f'^2 + l\beta_v(g' - f') = 0, \tag{2.9}$$

$$(1 + 2\gamma\eta)\left[N\theta'' + 3R_d(\theta_w - 1)\Theta^2\theta'^2\right] + 2\gamma N\theta' + \Pr f\theta' + \Pr(1 + 2\gamma\eta)\left(N_b\theta'\varphi' + N_t\theta'^2\right) + \Pr Q\theta + \Pr l\Gamma\beta_t(\theta_p - \theta) = 0, \tag{2.10}$$

$$(1 + 2\gamma\eta)\varphi'' + 2\gamma\varphi' + Sc\,f\varphi' + \frac{N_t}{N_b}[(1 + 2\gamma\eta)\theta'' + 2\gamma\theta'] - Sc\,\sigma_c\Theta^n\exp\left(-\frac{E}{\Theta}\right)\varphi = 0, \tag{2.11}$$

$$gg'' - g'^2 + \beta_v(f' - g') = 0, \qquad g\theta_p' + \beta_t(\theta - \theta_p) = 0, \tag{2.12}$$

subject to

$$\begin{gathered} f(0) = S, \quad f'(0) = \lambda, \quad \theta(0) = 1, \quad \varphi(0) = 1; \\ f'(\infty) = g'(\infty) = 0, \quad g(\infty) = f(\infty), \quad \theta(\infty) = \theta_p(\infty) = \varphi(\infty) = 0. \end{gathered} \tag{2.13}$$

Three features of the reduction are specific to the cylinder and are easily lost. The Maxwell bracket in (2.9) acquires the additional term $\gamma f^2 f''/(1 + 2\gamma\eta)$, which originates in $\hat{v}^2\hat{u}_{\hat{r}\hat{r}}$ and has no plane-sheet counterpart. The Brownian and thermophoretic contributions in (2.10) carry the factor $(1 + 2\gamma\eta)$, because each is a product of two radial derivatives and so collects the metric factor twice. The dust phase equations (2.12), by contrast, contain no diffusion operator and are therefore free of curvature terms. Setting $\gamma = 0$ recovers the plane-sheet system exactly.

The dimensionless groups of the reduced system are collected in Table 2.

Table 2. Dimensionless groups appearing in (2.9)–(2.12), with their definitions in terms of the dimensional quantities.

Symbol	Name	Definition
γ	Curvature parameter	$\frac{1}{R}\sqrt{\frac{\nu}{a}}$
β	Maxwell (Deborah) number	$\lambda_M a$
M	Magnetic (Hartmann) parameter	$\frac{\sigma B_0^2}{\rho a}$
K_p	Porosity parameter	$\frac{\nu}{ka}$
F_r	Local Forchheimer parameter	$\frac{C_b\hat{x}}{\sqrt{k}}$
l	Dust's mass concentration	$\frac{\rho_p}{\rho}$
β_v	Momentum interaction parameter	$\frac{1}{a\tau_m}$
β_t	Thermal interaction parameter	$\frac{1}{a\tau_T}$
Γ	Specific-heat ratio	$\frac{c_s}{c_p}$
Pr	Prandtl number	$\frac{\nu}{\alpha}$
R_d	Radiation parameter	$\frac{16\sigma^* T_\infty^3}{3k^*\kappa}$
θ_w	Temperature ratio	$\frac{T_w}{T_\infty}$
N_b	Brownian motion parameter	$\frac{\tau_n D_B(C_w - C_\infty)}{\nu}$
N_t	Thermophoresis parameter	$\frac{\tau_n D_T(T_w - T_\infty)}{\nu T_\infty}$
Q	Heat source/sink parameter	$\frac{Q_0}{a\rho c_p}$
Sc	Schmidt number	$\frac{\nu}{D_B}$
σ_c	Reaction rate parameter	$\frac{k_r}{a}$
E	Activation energy parameter	$\frac{E_a}{R_g T_\infty}$

Three of these require comment. First, M and K_p enter the reduced system only through the combination $-(M + K_p)f'$ in (2.9), and the thermal and species equations are influenced by them only through f. The magnetic and Darcy resistances are therefore degenerate: No solution of this system can distinguish a Hartmann number M from an equal increment in the porosity parameter K_p. The sum is accordingly reported as a single resistance group.

The Forchheimer group $F_r = C_b\hat{x}/\sqrt{k}$ is explicitly *local*: The coefficient $C_b/\sqrt{k}$ carries the dimension (length)$^{-1}$, so a dimensionless group can only be formed with an axial length, and the results are accordingly reported at a fixed station in the local similarity sense. Absorbing the axial dependence into a position-varying permeability is formally possible but physically artificial and is not adopted here.

The specific heat ratio is defined as $\Gamma = c_s/c_p$ and appears in the fluid phase energy equation only. This follows from the structure of the interphase exchange: The dust's thermal balance carries $\rho_p c_s$ on both sides, so the factor cancels there, whereas the corresponding sink in the fluid equation inherits $\rho_p c_s/(\rho c_p)$. Carrying Γ in both equations, as is sometimes done, is inconsistent.

2.3. Quantities of engineering interest

With $\mathrm{Re}_x = a\hat{x}^2/\nu$, we have

$$\mathrm{Re}_x^{1/2} C_f = f''(0), \qquad \mathrm{Re}_x^{-1/2} Nu_x = -\left(1 + R_d\theta_w^3\right)\theta'(0), \qquad \mathrm{Re}_x^{-1/2} Sh_x = -\varphi'(0). \tag{2.14}$$

3. Solution methodology

3.1. Numerical solution and continuation

The tenth-order system (2.9)–(2.12) is written in first-order form with the state vector $(f, f', f'', g, g', \theta, \theta', \varphi, \varphi', \theta_p)$ and solved by a fourth-order collocation method with adaptive mesh refinement and a residual tolerance of 10^{-8}. Because θ'' enters (2.11) and f''' appears inside the Maxwell bracket, the derivatives are recovered sequentially rather than by solving a coupled algebraic system: (2.9) yields f''', (2.10) then yields θ'', and (2.11) uses that value to yield φ''.

Prescribing λ and solving for $f''(0)$ fails near the turning point, where $\mathrm{d}\lambda/\mathrm{d}f''(0) \to 0$. The branch is therefore traced in two stages. In the first, λ is marched downwards from $\lambda = 1$ with warm starts. In the second, the roles are exchanged: $f''(0)$ is prescribed and λ is carried as an unknown parameter of the boundary value problem, giving 11 boundary conditions for 10 states plus one parameter. Under this parameterization, the fold is an ordinary interior point of the solution curve, so both branches are obtained continuously and no pseudo-arclength bordering or step-size heuristic is required.

The complete procedure, from the governing equations through the two-stage continuation and the admissibility screen to the stability eigenproblem, is summarized in Figure 2.

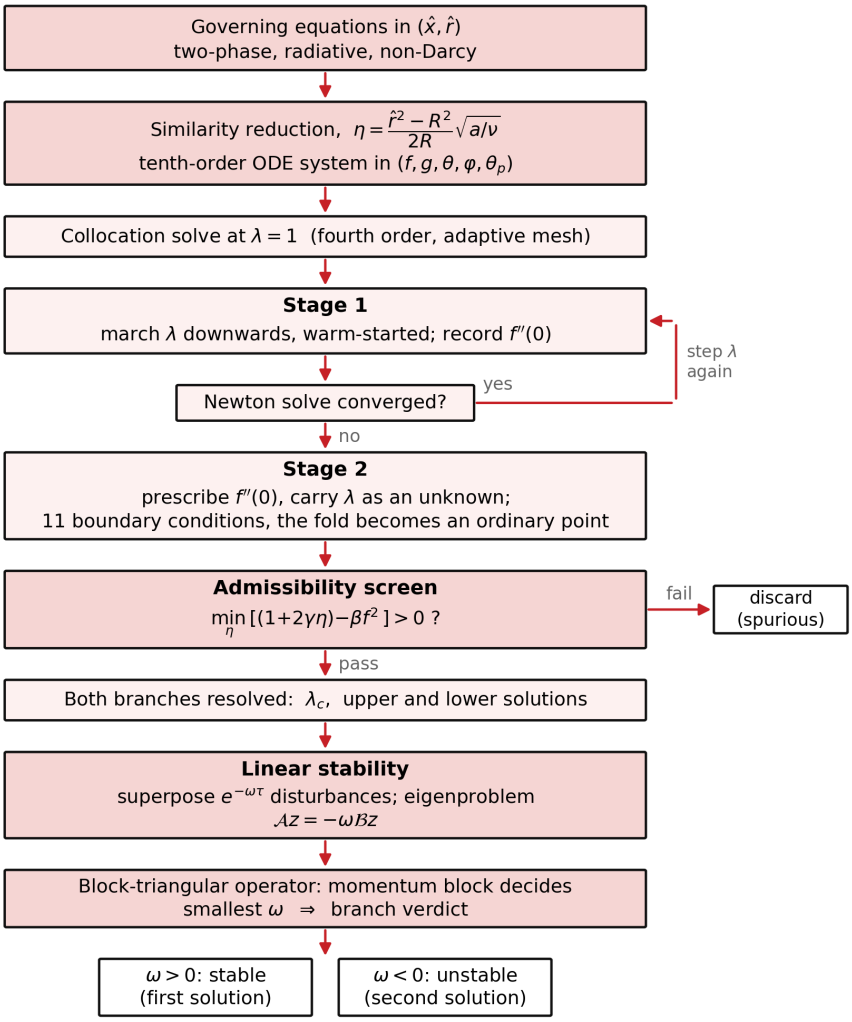


Figure 2. The numerical procedure. Stage 1 marches λ downwards until the Newton solver fails near the fold; Stage 2 exchanges the roles of λ and $f''(0)$ so that the turning point becomes an ordinary point of the solution curve. Every accepted solution is screened on the Maxwell admissibility margin before entering the stability analysis.

Solving (2.9) for the highest derivative gives

$$\left[(1+2\gamma\eta) - \beta f^2\right] f''' = -\Big(2\gamma f'' + f f'' - f'^2 + \beta\big(2 f f' f'' - \tfrac{\gamma f^2 f''}{1+2\gamma\eta}\big) - (M + K_p) f' - F_r f'^2 + l\beta_v (g' - f')\Big), \quad (3.1)$$

so the coefficient $(1 + 2\gamma\eta) - \beta f^2$ must remain strictly positive throughout the domain. On a plane sheet this reduces to the familiar condition $1 - \beta f^2 > 0$. This is not a formality: A collocation solver will report successful convergence on a spurious branch that has passed through the singularity, and such solutions display large nonmonotonic overshoots and approach the far field only at the truncation boundary. Every solution reported in this paper is screened for the margin $\min_\eta \left[(1 + 2\gamma\eta) - \beta f^2\right]$, and solutions failing the test are discarded rather than plotted. Because $1 + 2\gamma\eta \geq 1$, curvature relaxes the condition; the admissible range of β is mapped in Section 4.3.

3.2. Temporal stability analysis

Let $\tau = at$ and superpose the disturbances

$$\begin{aligned} f(\eta,\tau) &= f_0(\eta) + e^{-\omega\tau}F(\eta), \quad & g(\eta,\tau) &= g_0(\eta) + e^{-\omega\tau}P(\eta), \\ \theta(\eta,\tau) &= \theta_0(\eta) + e^{-\omega\tau}G(\eta), \quad & \varphi(\eta,\tau) &= \varphi_0(\eta) + e^{-\omega\tau}H(\eta), \\ \theta_p(\eta,\tau) &= \theta_{p0}(\eta) + e^{-\omega\tau}Q(\eta). \end{aligned} \tag{3.2}$$

Retaining terms that are linear in the disturbance, the momentum and dust-momentum equations give

$$\begin{aligned} &(1+2\gamma\eta)F''' + 2\gamma F'' + f_0F'' + f_0''F - 2f_0'F' \\ &\quad - \beta\Big[f_0^2F''' + 2f_0f_0'''F - 2f_0'f_0''F - 2f_0f_0''F' - 2f_0f_0'F'' + \tfrac{\gamma}{1+2\gamma\eta}(2f_0f_0''F + f_0^2F'')\Big] \\ &\quad - (M+K_p)F' - 2F_rf_0'F' + l\beta_v(P'-F') = -\omega F', \end{aligned} \tag{3.3}$$

$$g_0P'' + g_0''P - 2g_0'P' + \beta_v(F'-P') = -\omega P', \tag{3.4}$$

subject to $F(0) = F'(0) = 0$, $F'(\infty) = P'(\infty) = 0$, and $P(\infty) = F(\infty)$. The thermal disturbances satisfy the corresponding linearizations of (2.10), (2.11), and the dust energy equation, with $G(0) = H(0) = 0$ and $G(\infty) = H(\infty) = Q(\infty) = 0$.

Equations (3.3) and (3.4) involve F and P alone, whereas the thermal equations involve F and P as well as G, H, and Q. The linearized operator is therefore block lower-triangular

$$\mathcal{L} = \begin{pmatrix} \mathcal{A}_{\mathrm{mom}} & 0 \\ C & \mathcal{A}_{\mathrm{th}} \end{pmatrix}, \tag{3.5}$$

and its spectrum is the union of the spectra of the two diagonal blocks. The thermal block is a stable advection–diffusion operator whose eigenvalues remain strictly positive and bounded away from zero, so no eigenvalue of $\mathcal{A}_{\mathrm{th}}$ can change sign. Every stability transition is consequently governed by $\mathcal{A}_{\mathrm{mom}}$. This establishes, rather than assumes, the reduction to the momentum–dust subsystem that is customary in this literature; both blocks are nevertheless computed in Section 4.3 so that the claim is verified numerically.

The blocks are discretized by second-order central differences on a uniform mesh, boundary rows are substituted directly, and the resulting generalized eigenvalue problem $\mathcal{A}z = -\omega\mathcal{B}z$ is solved densely. Grid convergence is reported in Section 4.3.

The $O(\beta)$ unsteady Maxwell relaxation contributions ($\hat{u}_{tt}$, $\hat{u}_{\hat{x}t}$, $\hat{u}_{\hat{r}t}$) are omitted; retaining them would render the disturbance problem a quadratic eigenvalue problem. The omission is consistent with the boundary layer scaling and with established practice, and is recorded among the limitations in Section 5.

3.3. Physics-informed neural network

The network approximates the five primitive fields $(f, g, \theta, \varphi, \theta_p)$ as functions of η, with the derivatives obtained by automatic differentiation so that the residual is evaluated exactly rather than by finite differences. It is important to be precise about what is being minimized: The training objective is the residual of the similarity-reduced ordinary differential system (2.9)–(2.12), not of the governing

partial differential equations. We refer to it throughout as the ordinary differential equation (ODE) residual. The loss is

$$\mathcal{L} = \sum_i \langle r_i^2 \rangle_\eta + w_{\mathrm{bc}} \langle r_{\mathrm{bc}}^2 \rangle, \tag{3.6}$$

where r_i are the five equation residuals evaluated at the collocation nodes and r_{bc} collects the 10 boundary residuals.

The collocation nodes are Chebyshev-clustered toward the wall. The boundary layer occupies a small fraction of the truncated domain, so uniform sampling spends most of its points in the far field where every field is flat; clustering concentrates them where the residual is actually large.

The network, the derivative chain, and the composition of the loss are shown in Figure 3.

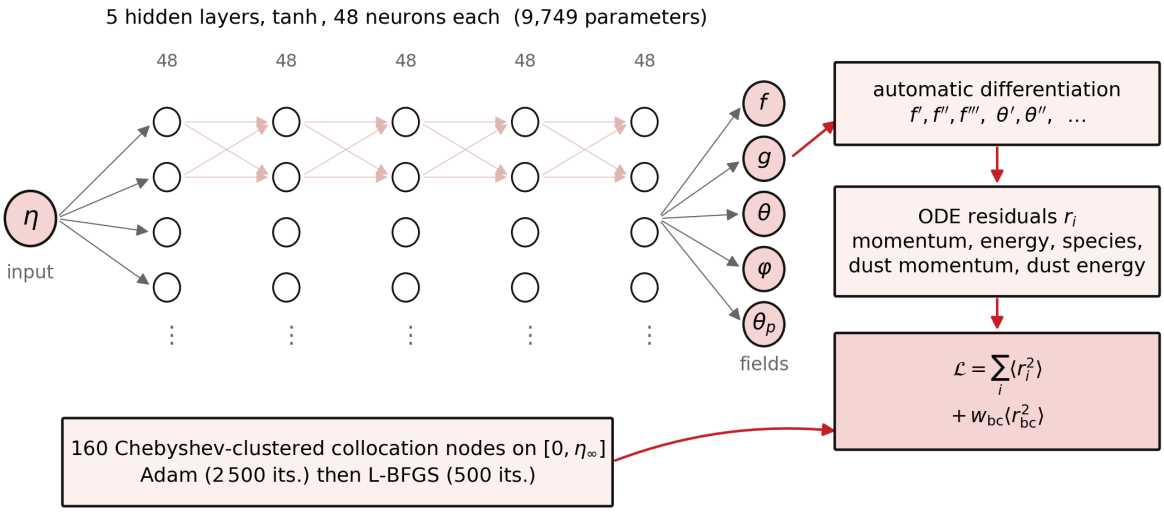


Figure 3. Surrogate architecture. A single input η maps through five tanh layers to the five primitive fields; derivatives are obtained by automatic differentiation, so the five ODE residuals are evaluated exactly rather than by finite differences and combined with the boundary residuals into the training objective.

4. Results and discussion

4.1. Numerical validation

4.1.1. Benchmark against published results

In the limit of the no dust phase with $\beta = M = K_p = F_r = Q = 0$, $\lambda = 1$, and $S = 0$, the momentum equation (2.9) collapses to

$$(1 + 2\gamma\eta)f''' + 2\gamma f'' + ff'' - f'^2 = 0, \qquad f(0) = 0,\ f'(0) = 1,\ f'(\infty) = 0, \tag{4.1}$$

for which skin friction values are available from Hashim et al. [5], Poply et al., [3] and Mahdy et al. [6].

A direct comparison is complicated by domain truncation, and the complication is instructive rather than incidental. Solving (4.1) on progressively larger domains, $f''(0)$ converges to values that differ systematically from the published ones, by as much as 1.4% at $\gamma = 1$. Inverting each published entry for the truncation η_∞ that would reproduce it gives $\eta_\infty = 9.99 \pm 0.02$ across all four curvatures: The published values are mutually consistent with a common truncation at $\eta_\infty = 10$, and the residual discrepancy is a truncation artefact rather than a difference of formulation.

Table 3 therefore reports both comparisons. At matched truncation, the present solver reproduces the reference values to within 0.004%; the converged column shows the truncation error the benchmark itself carries.

Table 3. Skin friction $\mathrm{Re}^{1/2}C_f$ for the reduced problem (4.1). The final column is the error the published values carry through truncation at $\eta_\infty = 10$. Here, — denotes a value not reported in the cited source.

γ	Hashim [5]	Poply [3]	Mahdy [6]	Present ($\eta_\infty = 10$)	Error	Present (converged)
0.0	−1.000000	—	−1.000000	−1.000008	0.001%	−1.000000
0.1	−1.036979	−1.036977	−1.036980	−1.036980	0.000%	−1.036575
0.3	−1.111165	−1.111138	−1.111162	−1.111150	0.001%	−1.108359
0.7	−1.257045	−1.257011	−1.257045	−1.257020	0.002%	−1.246319
1.0	−1.453373	−1.363865	−1.363924	−1.363870	0.004%	−1.345405

We note that the $\gamma = 1$ entry of Hashim et al. departs from the other two sources by about 6.5%; this disagreement is present in the published literature and is not resolved here. The present computation reproduces the Poply and Mahdy values.

4.1.2. Domain and mesh independence

Figure 4 plots both the benchmark comparison and the convergence behavior. For the full tenth-order system at the baseline operating point, $f''(0)$ and Nu are independent of the truncation to eight significant figures by $\eta_\infty = 16$, changing by less than 10^{-8} between $\eta_\infty = 20$ and 30. Because the collocation mesh is refined adaptively to a residual tolerance of 10^{-8}, the results are additionally invariant to the initial mesh over $200 \le n \le 3200$ nodes. All production runs use $\eta_\infty \ge 12$, which is substantially beyond the truncation at which the published benchmark values were obtained.

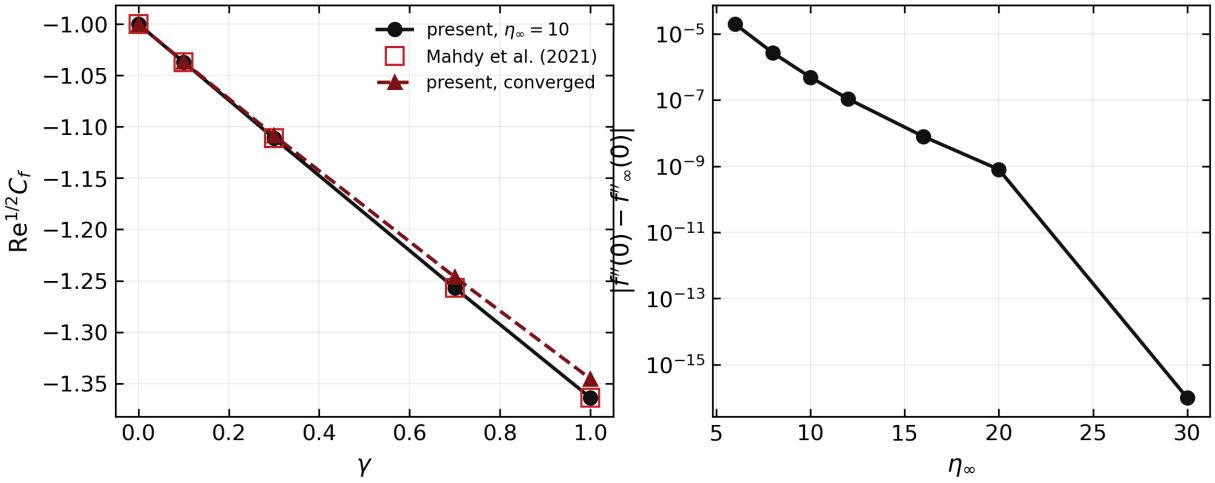


Figure 4. Left: Skin friction against curvature, comparing the present solver at matched truncation and at convergence with the published values. Right: Convergence of $f''(0)$ with the truncation η_∞ for the full tenth-order system.

4.1.3. Independent implementation

The momentum limit (4.1) was additionally solved by a second, independently written collocation routine that shares no code path with the full solver. The two agree to six significant figures at every curvature tested, which excludes an implementation error in the curvature terms as the source of the departure from the published values.

That comparison verifies the base flow but not the turning point, which is the quantity the remainder of the paper rests on. The fold locus was therefore recomputed by a third route: A bordered Newton pseudo-arclength continuation of the momentum and dust subsystem, discretized by trapezoidal collocation on a uniform grid with its own sparse solver, and sharing no numerical path with the routine used elsewhere. Restricting to that subsystem is exact rather than an approximation, because the reduced system is triangular: $f''(0)$ as a function of λ is fixed by the momentum and dust equations alone, the thermal and species fields being driven by f without feeding back. Grid refinement gives an observed order of accuracy of 1.97, against 2 for the trapezoidal scheme, and Richardson extrapolation at $\gamma = 0.3$ gives $\lambda_c = -1.734388$, where the continuation of Section 3 gives -1.734400, a difference of 1.2×10^{-5}. Over $\gamma = 0$ to 0.5, the two methods agree to between 0.002% and 0.077%, which is the size of the discretization error at the grids used. The critical shrinking rates reported below are therefore not the product of a single numerical route.

4.2. Admissible operating envelope

The first effect of curvature is not on where the fold sits but on whether a physically admissible solution exists at all. Section 3.1 showed that the coefficient $(1+2\gamma\eta)-\beta f^2$ multiplying f''' must remain positive; since $1 + 2\gamma\eta \geq 1$ with equality only at the wall, curvature can only relax that condition. The question is by how much.

Table 4 reports the largest admissible Maxwell parameter over a grid of suction rates and curvatures, each entry obtained by solving the full system and testing the margin $\min_\eta[(1 + 2\gamma\eta) - \beta f^2]$.

Table 4. Largest admissible Maxwell parameter β as a function of suction and curvature. Entries are the largest tested value of β for which the Maxwell coefficient remains positive throughout the layer.

	Curvature γ			
S	0.00	0.10	0.30	0.50
1.5	0.15	0.20	0.30	0.30
2.0	0.10	0.15	0.20	0.20
2.5	0.05	0.10	0.15	0.15
3.0	0.05	0.05	0.10	0.10
3.5	0.02	0.05	0.05	0.05

The effect is substantial, and Figure 5 shows it directly. Of the 30 (S,β) combinations tested, 12 are admissible on a plane sheet and 20 at $\gamma = 0.5$ (an enlargement of the operating region by two-thirds), and the largest attainable Deborah number doubles or triples at every suction rate. At $S = 2.5$, for instance, a plane sheet admits $\beta \le 0.05$, whereas a cylinder of curvature $\gamma = 0.3$ admits $\beta \le 0.15$.

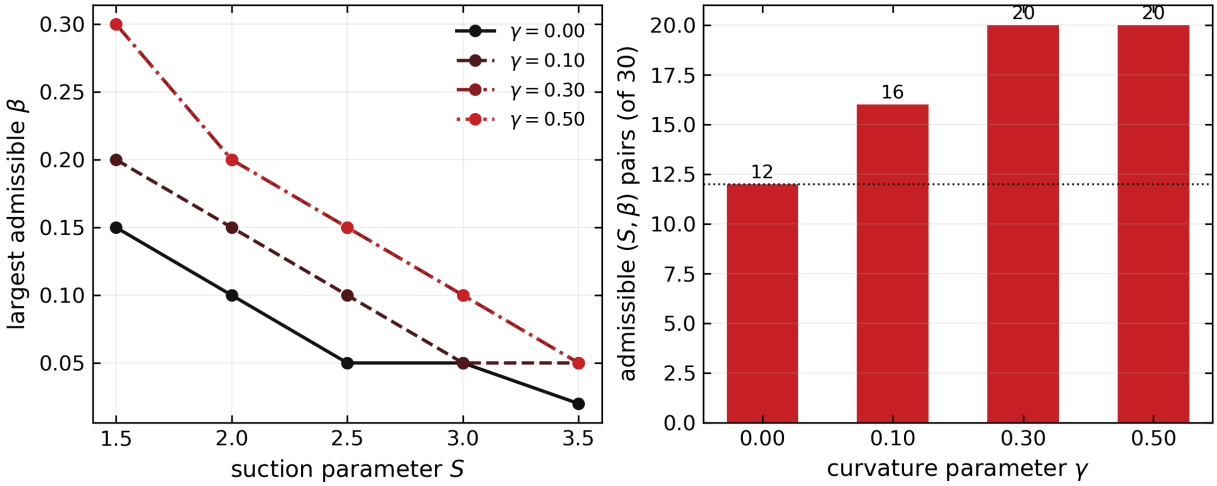


Figure 5. Left: Largest admissible Maxwell parameter against suction, at four curvatures. Right: The number of admissible (S,β) combinations out of the 30 tested; the dotted line marks the plane sheet value.

This matters practically because β and S are not free: The Maxwell number is set by the melt and the draw rate, and suction by the extraction system. Table 4 says that a viscoelastic operating point unreachable on a plane geometry may be reachable on a cylinder of the same nominal conditions, which is a statement about feasibility rather than about the size of a correction.

The condition also has a numerical corollary. Outside the admissible region the collocation solver does not fail: It converges, reports success, and returns a profile whose far-field value has run away to $f(\infty) \approx 80$–100 with a Maxwell coefficient of order -10^3. Such solutions display large nonmonotonic velocity overshoots and meet the outer condition only at the truncation boundary. Every solution reported in this paper has been screened for the margin and those failing it are discarded; without that

screen, they are indistinguishable from converged results.

4.3. *Dual solutions and the critical point*

For a shrinking surface with wall suction, the solution curve folds. Two solutions coexist for $\lambda_c < \lambda < 0$ and none exists for $\lambda < \lambda_c$. The location of that fold is what curvature controls.

Table 5 and Figure 6 report λ_c obtained by the two-stage continuation of Section 3, at $S = 2, \beta = 0.1$, and otherwise baseline parameters. The critical value rises monotonically with γ: From -1.8336 on a plane sheet to -1.6840 at $\gamma = 0.5$, a contraction of the dual-solution window by about 8% over that range. The skin friction at the fold moves in the opposite direction, from 1.9311 to 2.3388.

Table 5. Critical shrinking rate and the skin friction at the fold, as functions of the curvature parameter ($S = 2, \beta = 0.1$).

γ	0.0	0.1	0.2	0.3	0.4	0.5
λ_c	-1.8336	-1.7961	-1.7635	-1.7344	-1.7081	-1.6840
$f''(0)$ at the fold	1.9311	2.0125	2.0972	2.1808	2.2617	2.3388

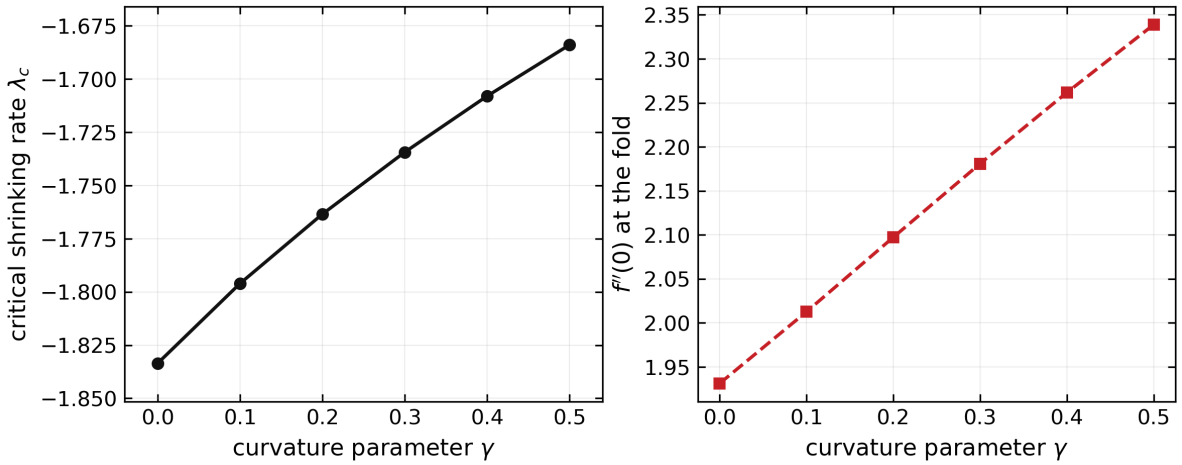


Figure 6. Critical shrinking rate λ_c (left) and the skin friction at the turning point (right) against the curvature parameter.

The mechanism is straightforward. Curvature enters the momentum balance through the diffusion coefficient $(1 + 2\gamma\eta)$, which exceeds unity everywhere in the layer and grows with distance from the wall. Radial transport of axial momentum is therefore more effective on a cylinder than on a plane sheet at the same nominal conditions, and the layer thins. The second solution branch, which is sustained by the balance between the imposed wall motion and a thick, weakly diffusing outer layer, is the more fragile of the two, so enhanced diffusion removes it first. The result is that the fold retreats towards $\lambda = 0$ as γ increases.

Note that λ_c is converged to six significant figures: It is unchanged to that precision across continuation step sizes $\delta f''(0) \in \{0.16, 0.08, 0.04\}$ and truncations $\eta_\infty \in \{10, 15, 20, 25\}$.

The sweep is reported to $\gamma = 0.5$, and the fold ceases to exist not far beyond it. The reason is

internal to the model rather than to the solver, and is worth stating because it bounds the range over which a critical shrinking rate exists at all. Along the first branch, $f(\infty)$ falls monotonically as λ is reduced. The dust phase satisfies $g(\infty) = f(\infty)$, and g is the coefficient of g'' in the dust momentum equation (2.12), so the reduced system degenerates where $f(\infty)$ reaches zero. At $\gamma = 0.3$, the turning point is reached with $f(\infty) = 0.280$ still in hand, at $\gamma = 0.5$ with 0.119, and at $\gamma = 0.6$ with only 0.017. Beyond $\gamma \simeq 0.6$, the degeneracy is encountered first, so the branch terminates instead of folding and no critical shrinking rate exists. This is not a truncation effect: At $\gamma = 0.65$, the branch ends at $\lambda = -1.6565$ for $\eta_\infty = 15$, 20 and 25 alike. Nor is it the Maxwell condition, whose margin remains at 0.600 throughout and never binds.

4.3.1. Branch structure

Figure 7 shows the solution curves in the $(\lambda, f''(0))$ plane at two curvatures, and Figure 8 shows the corresponding profiles on each branch. The second solution is distinguished by a much thicker layer: At a comparable distance from the fold, its velocity deficit extends roughly twice as far into the fluid, and its thermal layer is thicker still.

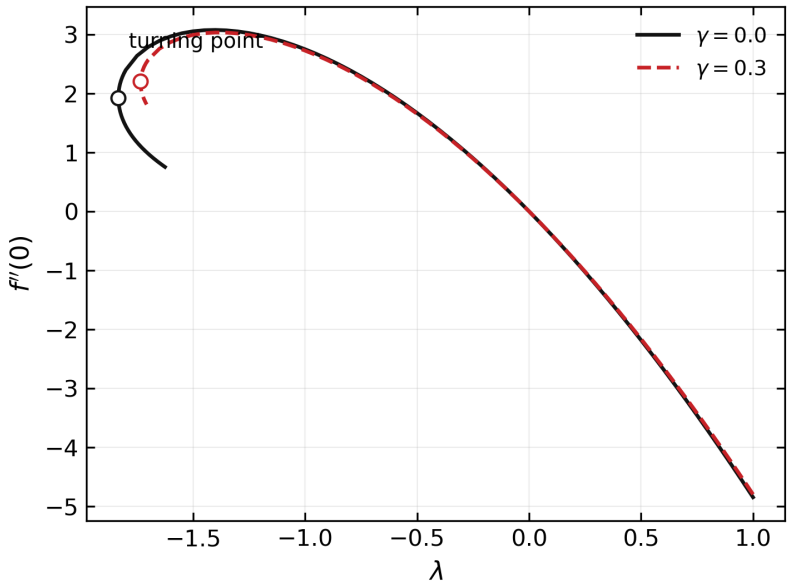


Figure 7. Solution curves $f''(0)$ against λ at two curvatures. Open circles mark the turning points. The upper portion of each curve is the first (stable) solution; the lower portion is the second (unstable) solution.

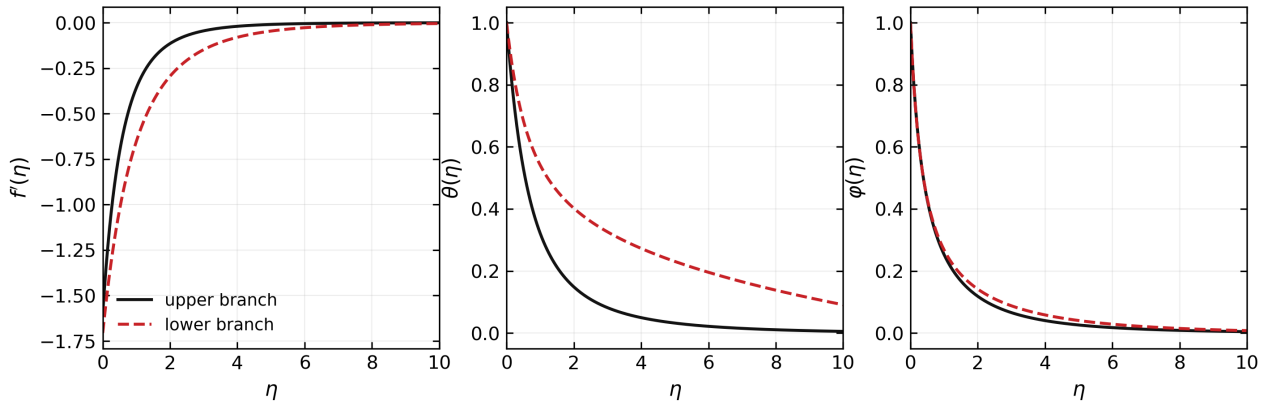


Figure 8. Velocity, temperature, and concentration profiles on both branches at $\gamma = 0.3$. The two solutions are taken at the traced points nearest $0.92\lambda_c$ on their respective branches, with $\lambda = -1.600$ on the upper and $\lambda = -1.707$ on the lower; the lower branch occupies only a narrow interval above λ_c, so an exact match in λ is not available. The second solution carries a markedly thicker velocity and thermal layer.

4.3.2. Interaction with suction

Suction and curvature act on the fold in opposite senses, and the two do not simply add. Table 6 reports λ_c over a range of suction rates at two curvatures, computed at $\beta = 0.05$ so that both geometries remain admissible throughout (Section 4.2).

Table 6. Critical shrinking rate as a function of suction on a plane sheet and on a cylinder ($\beta = 0.05$). The final column is the fractional contraction of the dual-solution window produced by curvature. Here, — denotes a case in which no turning point could be resolved, with the branch reaching the degeneracy of the dust momentum equation first; see the text.

S	λ_c at $\gamma = 0$	λ_c at $\gamma = 0.3$	Contraction
1.5	−1.3713	—	—
2.0	−1.8574	−1.7493	5.8%
2.5	−2.4375	−2.2635	7.1%
3.0	−3.1735	−2.9050	8.5%
3.5	−4.1644	−3.7694	9.5%

Figure 9 plots the same data. Suction is by far the stronger influence on the location of the fold: Raising S from 1.5 to 3.0 moves λ_c from −1.3713 to −3.1735, more than doubling the range of shrinking rates over which two solutions coexist. Wall extraction thins the layer and carries momentum towards the surface, which extends the second branch to much larger shrinking rates.

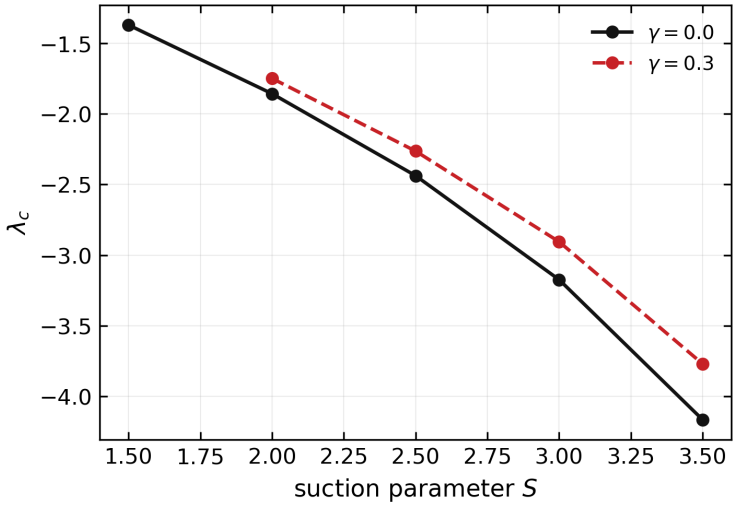


Figure 9. Critical shrinking rate against suction at two curvatures.

Curvature opposes this, and the opposition strengthens as suction increases: The contraction it produces grows from 5.8% at $S = 2.0$ to 9.5% at $S = 3.5$, monotonically and without sign of saturating at the strongest extraction rate considered. At $S = 3.5$, both geometries still retain a well-resolved turning point, namely at $\lambda_c = -4.1644$ on the plane sheet and -3.7694 on the cylinder. The two effects are therefore not independent, and a design in which both are large cannot be assessed by treating either alone.

The first row is reported as unresolved, and the reason is worth stating rather than concealing. At $S = 1.5$, the plane sheet folds at $\lambda_c = -1.3713$ with $f(\infty) = 0.182$ still in hand. At $\gamma = 0.3$, the same conditions carry the first branch to the edge of the degeneracy described earlier in this section: $f(\infty)$ falls below 10^{-2} near $\lambda \simeq -1.309$, and a turning point, if one exists there, lies within about 10^{-3} in λ of the degeneracy itself. Neither continuation resolves a turn at that separation which survives grid refinement, and the far-field value is itself not grid-converged in that interval. We therefore report no critical value for this case rather than one the computation does not support. The Maxwell admissibility margin is 0.89 throughout and is not the constraint in play. What the row does establish is the direction of the effect: At low suction, curvature drives the first branch to the edge of the two-phase degeneracy, so the dual-solution structure is lost, or so nearly lost as to be unresolvable, rather than merely displaced.

4.4. Temporal stability

The disturbance problem of Section 3.2 was solved on both branches at several distances from the turning point. Table 7 reports the smallest eigenvalue in each case.

Table 7. Smallest disturbance eigenvalue on each branch. A positive ω denotes a decaying disturbance. Here, $\lambda_c = -1.8336$ at $\gamma = 0$ and -1.7344 at $\gamma = 0.3$.

γ	Branch	λ	$f''(0)$	ω
0.0	Upper	−1.5500	3.0185	+1.234724
0.0	Upper	−1.7500	2.6378	+0.674624
0.0	Upper	−1.8110	2.3186	+0.358111
0.0	Lower	−1.8180	1.5986	−0.228352
0.0	Lower	−1.7728	1.2786	−0.459050
0.3	Upper	−1.4500	3.0195	+0.707925
0.3	Upper	−1.6500	2.7573	+0.454825
0.3	Upper	−1.7134	2.4897	+0.253153
0.3	Lower	−1.7124	1.8497	−0.205833

The pattern is the one a saddle-node requires. On the upper branch, ω is positive and falls monotonically towards zero as $\lambda \to \lambda_c$; on the lower branch, it is negative and rises towards zero from below. The eigenvalue crosses zero at the turning point, so the first solution is physically realizable and the second is not. Curvature lowers the decay rate at a comparable distance from the fold (at $\gamma = 0.3$, the upper branch returns $\omega = +0.708$ where the plane sheet gives $+1.235$), so disturbances on a cylinder relax more slowly, although the sign, and therefore the verdict, is unchanged.

Two points on the method should be noted. First, the eigenvalue quoted is the algebraically smallest, not the one nearest zero. On the lower branch, a small positive eigenvalue coexists with the larger negative one that governs the instability, and selecting by magnitude would return the positive value and misreport the branch as stable.

The second point is the verification of the block argument. At $\lambda = -1.0$ and $\gamma = 0.3$, the momentum block returns $\{+0.956, +1.449, +1.449, +1.666, \dots\}$ and the thermal block $\{+0.501, +0.501, +0.502, +0.502, \dots\}$, the latter clustering near β_t, which is the dust's thermal relaxation rate. The thermal eigenvalues are therefore smaller than the momentum ones at this operating point, but they are strictly positive and bounded away from zero, and they remain so along both branches. Since no thermal eigenvalue changes sign, the stability transition is governed entirely by the momentum block, exactly as the block-triangular structure of Section 3.2 requires. Where the momentum block is comfortably stable, the overall smallest eigenvalue may belong to the thermal block, which does not alter the verdict.

Figure 10 plots both of these: The eigenvalue against λ on each branch, and the leading spectra of the two blocks.

Grid convergence is reported in Table 8: ω varies by less than 0.3% between $N = 81$ and $N = 321$ and converges to three decimal places by $N = 161$, which is the resolution used throughout.

Table 8. Grid convergence of the smallest momentum eigenvalue ($\gamma = 0.3$, $\lambda = -1.0$).

N	81	121	161	241	321
ω	0.958058	0.956740	0.956144	0.955650	0.955459

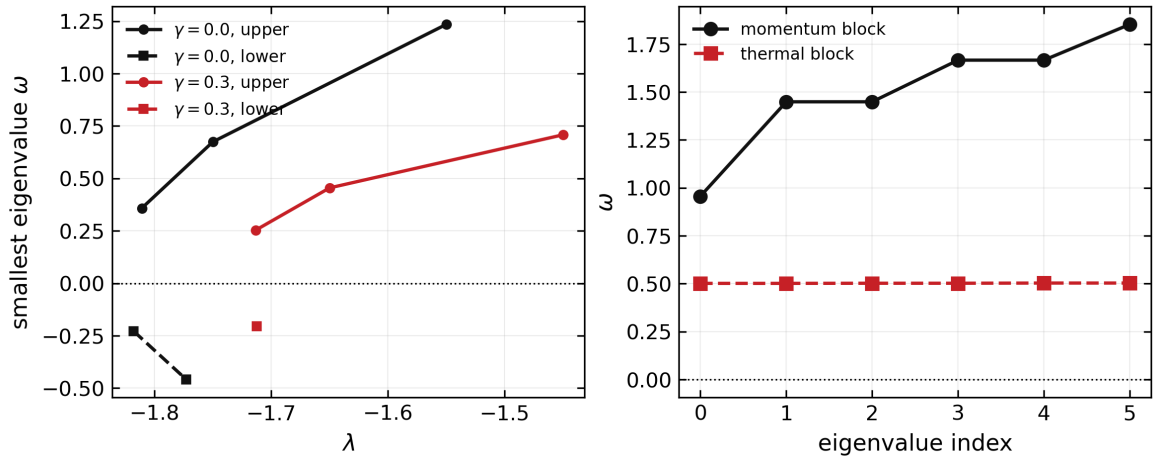


Figure 10. Left: Smallest eigenvalue against λ on both branches at two curvatures; the dotted line is $\omega = 0$. Right: The leading eigenvalues of the momentum and thermal blocks at $\gamma = 0.3$, $\lambda = -1.0$, showing that the thermal block is bounded away from zero.

4.5. Velocity, temperature, and concentration fields

This section reports the stretching branch fields at the baseline operating point ($\gamma = 0.3$, $S = 2$, $\lambda = 1, \beta = 0.1$) as each governing group is varied in turn. Only admissible values of β are included, in the sense of Section 3.1. Table 9 collects the wall quantities.

Table 9. Wall quantities as each parameter is varied about the baseline. $C_f \equiv \mathrm{Re}^{1/2} C_f$, $Nu \equiv \mathrm{Re}^{-1/2} Nu_x$, and $Sh \equiv \mathrm{Re}^{-1/2} Sh_x$.

Parameter	Value	C_f	Nu	Sh
γ	0.0	−4.8469	3.2569	3.1069
	0.3	−4.7975	3.2991	3.1540
	0.5	−4.7936	3.3288	3.1916
F_r	0.0	−4.7463	3.2999	3.1549
	0.9	−4.8978	3.2974	3.1525
β	0.0	−2.9331	3.3561	3.2100
	0.15	−7.0941	3.2569	3.1140
l	0.0	−4.7180	3.1272	3.2252
	0.9	−4.8593	3.4265	3.1012
R_d	0.0	−4.7975	2.8733	2.5671
	1.5	−4.7975	3.8299	3.6176
N_b	0.1	−4.7975	3.8499	0.1222
	0.7	−4.7975	2.3930	3.9875
N_t	0.1	−4.7975	3.4296	3.7227
	0.4	−4.7975	3.0566	2.1559
Sc	1.0	−4.7975	3.5653	0.9658
	4.0	−4.7975	3.0884	7.3314
E	0.0	−4.7975	3.2759	3.4124
	3.0	−4.7975	3.3135	2.9981

4.5.1. One-way coupling in the wall data

The most immediate feature of Table 9 is that C_f takes the identical value -4.7975 for every value of R_d, N_b, N_t, Q, Sc, σ_c, and E. This is not a coincidence of rounding. The momentum equation (2.9) contains no thermal or species variable, so the velocity field is determined without reference to them, while the energy and concentration equations both depend on f. The coupling runs one way only, and the invariance of the skin friction is that structure appearing directly in the results. It is the same property that makes the stability operator block lower-triangular (Section 3.2), and it provides a useful internal consistency check on the solver: Any dependence of C_f on a thermal parameter would indicate an error.

4.5.2. Velocity field

Figure 11 shows the velocity field across the momentum groups. Curvature raises the wall velocity gradient slightly, moving C_f from -4.8469 to -4.7936 between $\gamma = 0$ and 0.5. The enhanced radial transport thins the layer, and the profile relaxes to the free stream over a shorter span of η. The Forchheimer group acts in the opposite sense, its quadratic drag steepening the profile and driving C_f from -4.7463 to -4.8978 as F_r goes from 0 to 0.9.

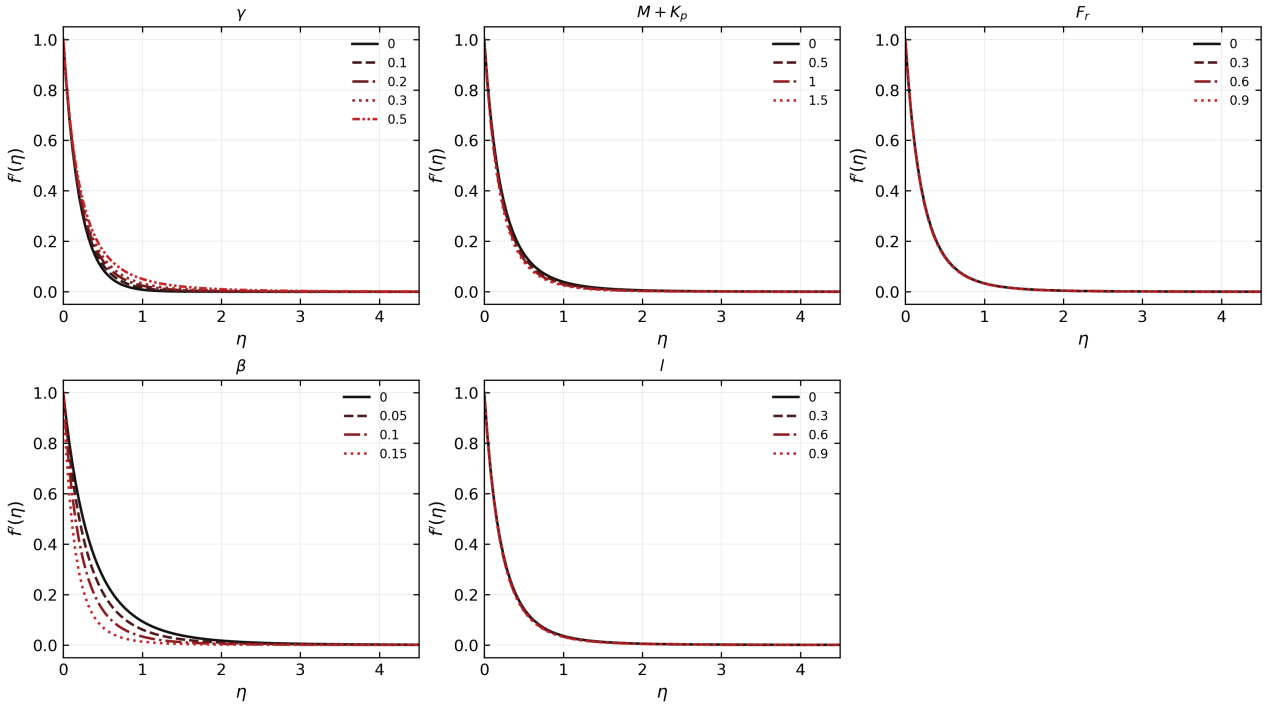


Figure 11. Fluid axial velocity for the curvature, resistance, Forchheimer, Maxwell, and dust loading groups.

The Maxwell parameter is by far the strongest influence on the wall stress, taking C_f from -2.9331 at $\beta = 0$ to -7.0941 at $\beta = 0.15$, a factor of 2.4 across a range that is itself bounded by the admissibility condition. This sensitivity is the reason the admissibility bound matters: The quantity that responds most strongly to β is the one whose governing coefficient β threatens to annihilate.

Dust loading steepens the fluid profile modestly (C_f goes from −4.7180 to −4.8593 as l runs from 0 to 0.9) because the lagging particulate phase extracts momentum through the interphase drag.

Figure 12 shows the slip directly. The dust's velocity g' remains far below the carrier velocity f' throughout the layer, and the dust's temperature θ_p reaches only about a tenth of the carrier value at the wall. This is the expected behavior for $St = O(1)$ particles: They respond to the carrier over the relaxation lengths set by β_v and β_t, rather than following it.

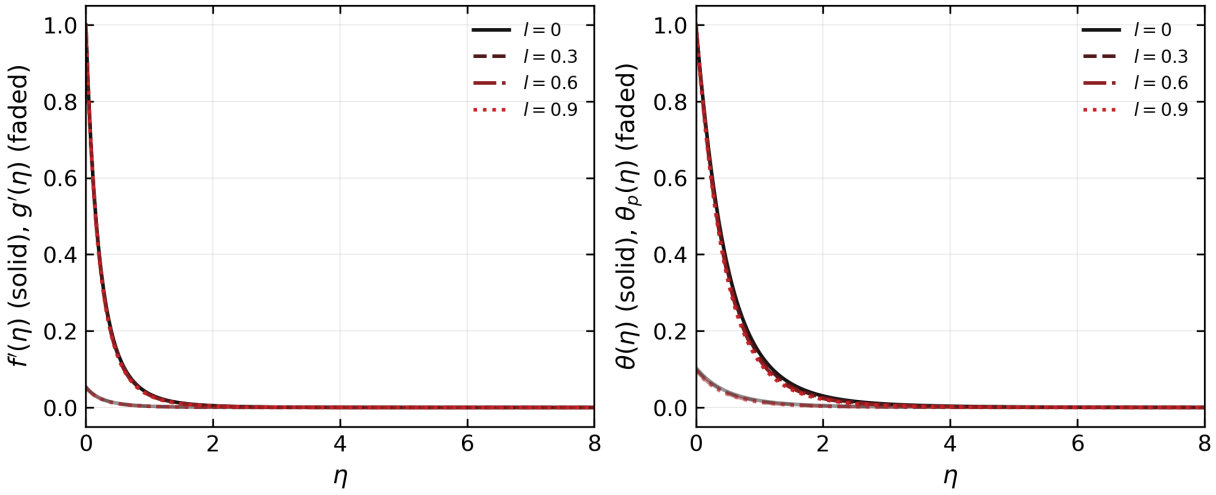


Figure 12. Fluid and dust fields at several loadings; dust quantities are shown faded. The particulate phase lags the carrier in both velocity and temperature.

4.5.3. Temperature field

The temperature field is shown in Figure 13. Thermal radiation is the dominant control on heat transfer: Nu rises from 2.8733 to 3.8299 (a one-third increase) as R_d goes from 0 to 1.5, the additional radiative conductivity thickening the thermal layer while steepening the wall gradient.

The two nanoparticle groups act as the Buongiorno model requires. Brownian motion redistributes energy away from the wall and Nu falls from 3.8499 to 2.3930 as N_b increases from 0.1 to 0.7; thermophoresis drives particles down the temperature gradient and reduces Nu more mildly, from 3.4296 to 3.0566.

Dust loading raises the heat transfer, with Nu climbing from 3.1272 to 3.4265 as l goes from 0 to 0.9. The interphase thermal exchange gives the carrier a second route for depositing energy, and the lagging particles act as a distributed sink within the layer. Note that the loading raises Nu while lowering Sh (from 3.2252 to 3.1012). The two transfer coefficients move in opposite directions, so a dusty carrier cannot be optimized for both simultaneously.

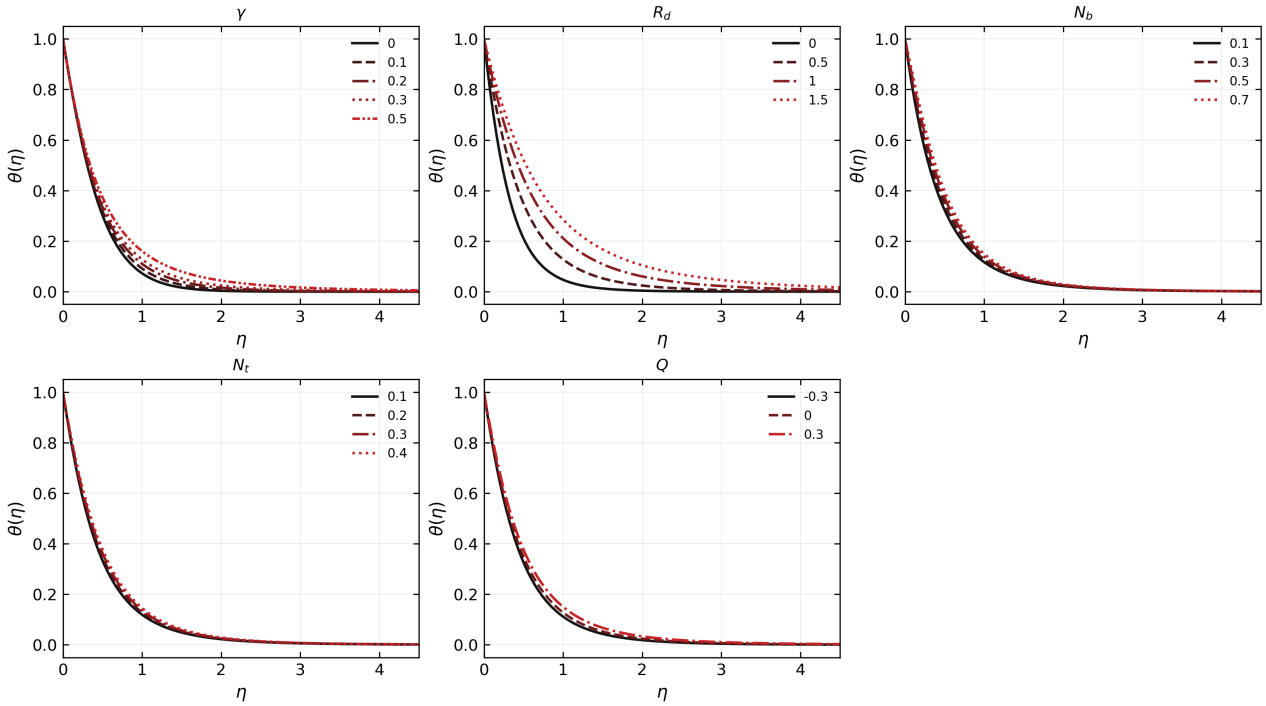


Figure 13. Fluid temperature for the curvature, radiation, Brownian, thermophoresis, and heat source groups.

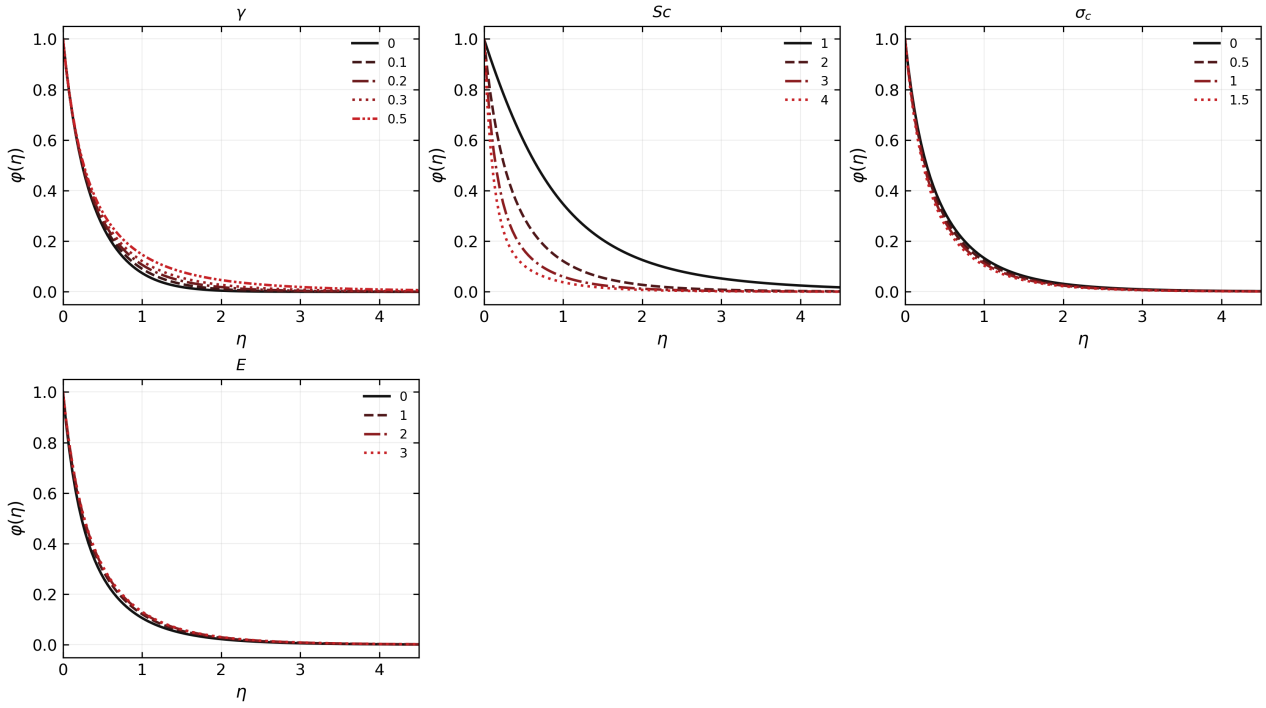


Figure 14. Nanoparticle concentration for the curvature, Schmidt, reaction rate, and activation energy groups.

4.5.4. Concentration field

Figure 14 shows the species field. The Schmidt number governs it, with Sh rising steeply from 0.9658 at $Sc = 1$ to 7.3314 at $Sc = 4$ as weaker Brownian diffusivity confines the concentration boundary layer to an increasingly thin region near the wall.

The reaction terms behave as their physical interpretation demands, and in opposite senses. Increasing the reaction rate σ_c consumes species within the layer and steepens the wall gradient, raising Sh from 2.9694 to 3.4746 across $0 \le \sigma_c \le 1.5$. Increasing the activation energy E suppresses the reaction (fewer collisions carry sufficient energy), so less of the species is consumed, the profile flattens, and Sh falls from 3.4124 to 2.9981 as E runs from 0 to 3. The two therefore provide independent control of the wall mass flux in opposing directions.

Curvature raises both transfer coefficients modestly: Nu by 2.2% and Sh by 2.7% between $\gamma = 0$ and 0.5, consistent with the thinner layer it produces.

4.6. Streamlines and isotherms

The similarity reduction gives the stream function directly as $\psi = \sqrt{a\nu}\,\hat{x}Rf(\eta)$, so streamlines in the physical plane are contours of $\hat{x}f(\eta)$. Figure 15 shows these together with the isotherms for a plane sheet and for a cylinder at $\gamma = 0.3$, at the baseline operating point.

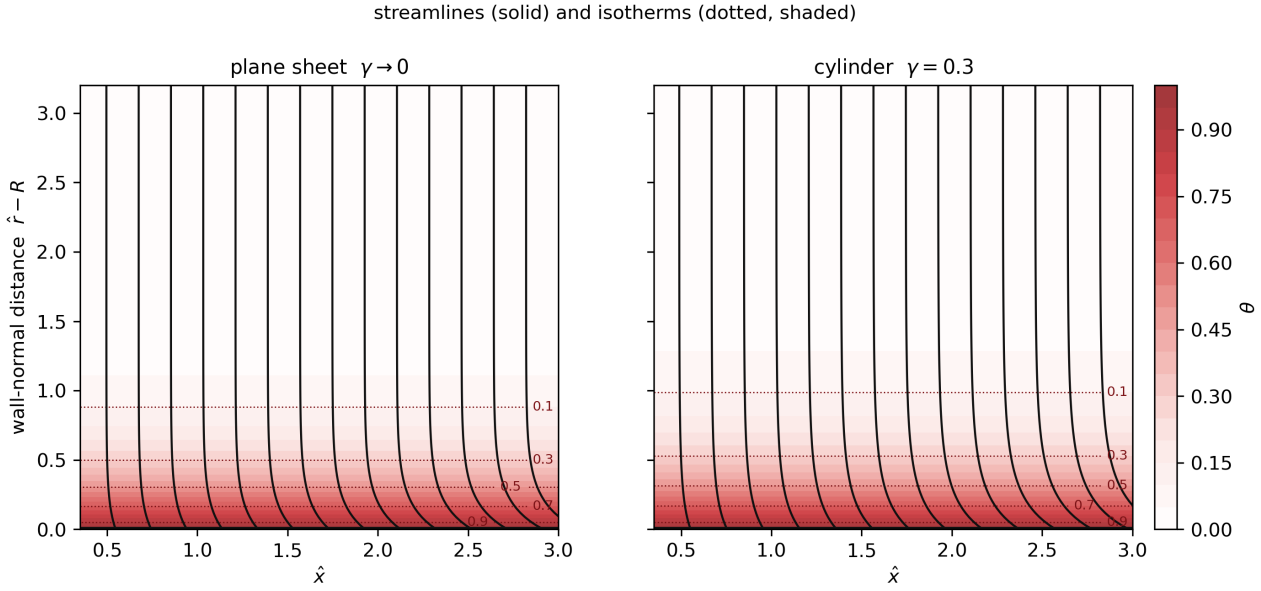


Figure 15. Streamlines (solid) and isotherms (dotted, with red shading) in the physical plane at $\lambda = 1$, $S = 2$, and $\beta = 0.1$. Left: Plane sheet. Right: Cylinder at $\gamma = 0.3$. The vertical coordinate is the wall-normal distance $\hat{r} - R$.

Because the temperature is self-similar, θ depends on η alone and the isotherms are surfaces of the constant η, and hence of the constant $\hat{r}$. They appear as horizontal bands in this plane. That is, a property of the similarity reduction rather than of the particular solution, and the figure is read for the spacing of those bands, which is the thermal layer thickness and varies.

The streamlines are drawn towards the surface by the suction and are compressed axially as $\hat{x}$

increases, with the wall velocity $\lambda a\hat{x}$ growing linearly with distance along the cylinder. Comparing the two panels shows the effect of curvature on the thermal field directly. The wall temperature gradient steepens, with $-\theta'(0)$ rising from 1.9555 at $\gamma = 0$ to 1.9808 at $\gamma = 0.3$, while the outer edge of the thermal layer moves outwards: The station at which $\theta = 0.1$ lies at $\eta = 0.891$ on the plane sheet and at $\eta = 1.133$ at $\gamma = 0.3$, which in physical units, are at $\hat{r} - R = 0.891$ and 0.987, respectively. The trend continues beyond the two curvatures drawn here: Across the sweep of Section 4.5, the gradient reaches 1.9987 at $\gamma = 0.5$ and the $\theta = 0.1$ station moves out to $\eta = 1.327$, or $\hat{r} - R = 1.051$. Curvature therefore produces a more sharply peaked temperature profile with a longer tail, which is the profile shape consistent with the simultaneous rise in Nu reported in Table 9.

4.7. Physics-informed neural network

4.7.1. Forward validation and error metrics

Table 10 reports the accuracy of the trained network against the collocation reference for every state variable, over 10 independent random initializations. Reporting all five matters: The dust phase fields are the least accurate and averaging them into a single figure of merit would conceal that.

Table 10. Relative L_2 error of the trained network against the collocation reference, over 10 independent seeds. Architecture and training protocol as in Appendix A.

Field	Mean	Std. dev.	Min	Max
f	1.48×10^{-4}	6.19×10^{-5}	4.15×10^{-5}	2.37×10^{-4}
g	3.71×10^{-4}	1.70×10^{-4}	1.78×10^{-4}	6.80×10^{-4}
θ	4.97×10^{-4}	3.32×10^{-4}	1.37×10^{-4}	1.29×10^{-3}
φ	5.19×10^{-4}	2.75×10^{-4}	2.33×10^{-4}	1.06×10^{-3}
θ_p	1.21×10^{-2}	8.18×10^{-3}	2.66×10^{-3}	2.56×10^{-2}
Final ODE residual loss	7.90×10^{-7}	2.60×10^{-7}	3.35×10^{-7}	1.12×10^{-6}
training time (s)	231.0	87.6	145.6	459.4

Figure 16 shows the same data by seed. The carrier-phase fields are resolved to between one and five parts in 10^4, and the run-to-run spread is a factor of 4–10. The dust temperature θ_p is the clear outlier: Its mean error is roughly 25 times that of f, and it varies by an order of magnitude between seeds.

The reason is structural rather than incidental, and it is worth stating because it bears on where such networks can be trusted. The dust energy equation $g\theta_p' + \beta_t(\theta - \theta_p) = 0$ is first-order and carries no diffusive term, so it supplies no smoothing of its own; θ_p is governed by g, which is itself constrained only weakly near the wall where g is small. The residual is therefore least informative exactly where the field is hardest to resolve. For the dust's temperature, the collocation solver, not the network, should be regarded as the reference.

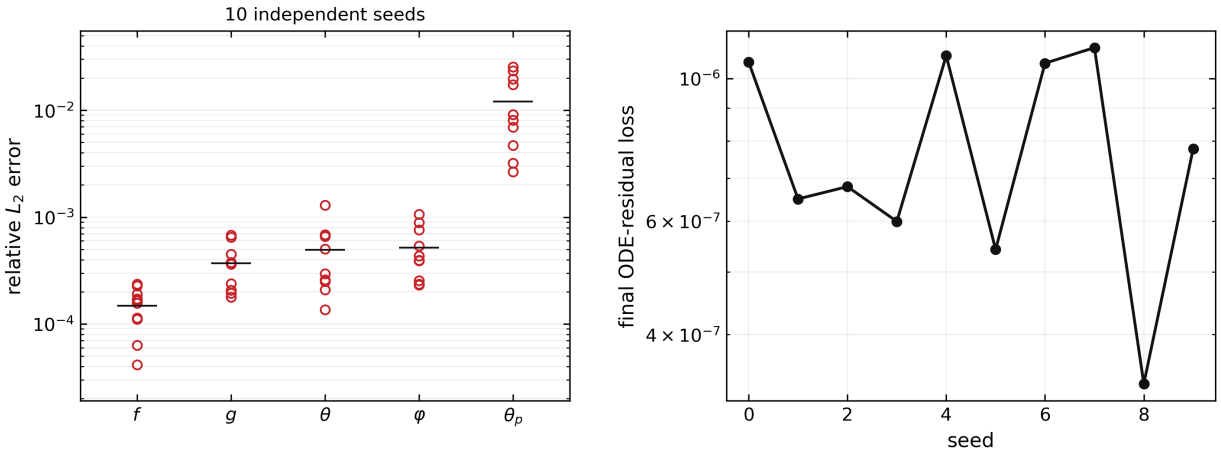


Figure 16. Left: Relative L_2 error for each state variable over 10 seeds; horizontal bars mark the means. Right: The final ODE residual loss by seed.

4.7.2. Comparison with classical surrogates

If the purpose of a surrogate is to return wall quantities across a parameter range, then the relevant question is not whether a network can do it but whether it is the cheapest tool that can. We therefore compare against two classical surrogates fitted to identical reference data: A Gaussian process with a Matérn-5/2 kernel, and a thin-plate-spline radial basis function interpolant.

A single direct solution of the reduced system costs 264 ms. Table 11 reports the accuracy of each surrogate in predicting $f''(0)$ across $\lambda \in [-1.6, 1.0]$ at $\gamma = 0.3$, as a function of the number of training solutions.

Table 11. Classical surrogates for $f''(0)$, fitted to n direct solutions. Errors are relative L_2 over 39 test points. The query cost is below measurement resolution for both. GP, Gaussian process; RBF, radial basis function.

n	Gaussian process	RBF	Data cost (s)	GP fit (s)
5	1.40×10^{-2}	3.37×10^{-2}	1.3	0.14
8	3.49×10^{-3}	1.19×10^{-2}	2.1	0.10
10	1.78×10^{-3}	7.26×10^{-3}	2.6	0.15
15	5.06×10^{-4}	2.96×10^{-3}	4.0	0.19
20	2.13×10^{-4}	1.52×10^{-3}	5.3	0.17

Figure 17 plots the accuracy against training set size and the cumulative cost of each route. The outcome is unambiguous and we state it plainly. Ten direct solutions, costing 2.6 s to generate, support a Gaussian process that predicts $f''(0)$ to 0.18% over the whole interval, fits in 0.15 s, and answers queries at no measurable cost. Twenty solutions reduce the error to 0.021%. Expressed as a break-even against direct solution, the Gaussian process repays its construction after roughly 11 evaluations, whereas a network whose training is two to three orders of magnitude more expensive requires several hundred. For smooth scalar outputs over a parameter interval, the classical surrogate is the better instrument, and no appeal to the network's physics content changes that arithmetic. The radial basis

interpolant is consistently about four times poorer than the Gaussian process at equal data cost, but remains far cheaper than the network.

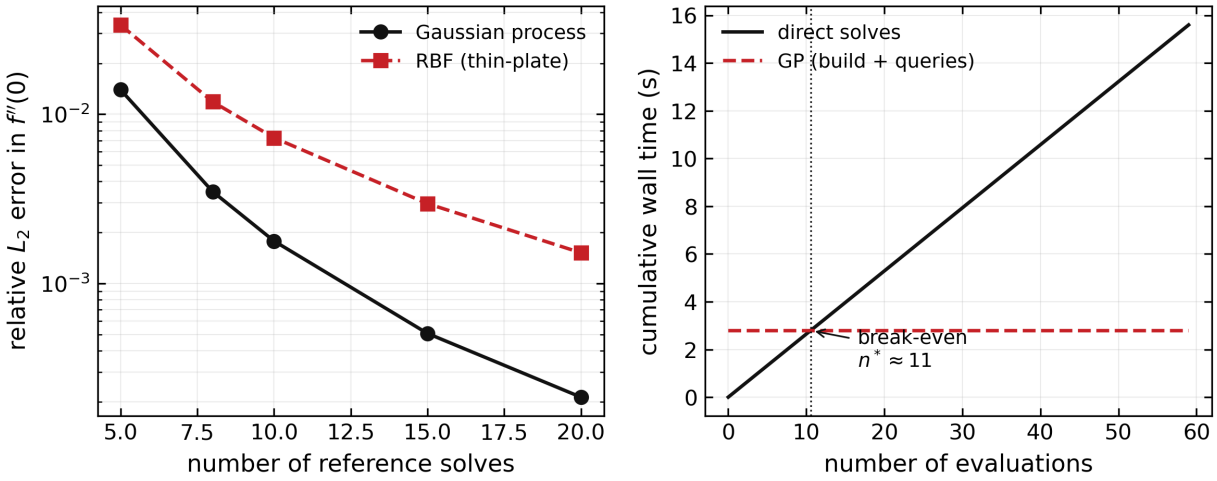


Figure 17. Left: Accuracy of the classical surrogates for $f''(0)$ against the number of reference solutions. Right: Cumulative cost of repeated direct solution compared with building a Gaussian process on 10 solutions and querying it; the surrogate repays its construction after about 11 evaluations.

4.7.3. Prediction near the turning point

The comparison of Section 4.7.2 is not, however, the whole story, and the qualification matters more than the headline. A Gaussian process was fitted to 10 solutions spanning $\lambda \in [-1.4, 1.0]$ and then queried at values withheld from training, chosen to approach the turning point $\lambda_c = -1.736$ from above. Table 12 reports the result.

Table 12. The Gaussian process's (GP) prediction of $f''(0)$ at a held-out λ, ordered by distance from the turning point. Training spanned $\lambda \in [-1.4, 1.0]$; entries below that range are extrapolation. Reference values are taken from the traced first branch, so training and test data lie on the same branch throughout.

λ	$\lambda - \lambda_c$	Reference	GP prediction	Abs. error	
0.60	2.336	−2.6522	−2.6526	4.4×10^{-4}	
0.00	1.736	0.0000	−0.0003	2.8×10^{-4}	
−0.70	1.036	2.1580	2.1573	6.4×10^{-4}	
−1.20	0.536	2.9542	2.9513	3.0×10^{-3}	
−1.45	0.286	3.0200	3.0293	9.2×10^{-3}	Extrapolation
−1.55	0.186	2.9446	2.9988	5.4×10^{-2}	Extrapolation
−1.60	0.136	2.8712	2.9712	1.00×10^{-1}	Extrapolation
−1.65	0.086	2.7597	2.9356	1.76×10^{-1}	Extrapolation
−1.70	0.036	2.5748	2.8922	3.17×10^{-1}	Extrapolation

The error grows by nearly three orders of magnitude as the fold is approached, from 4×10^{-4} at

$\lambda - \lambda_c = 2.3$ to 3.2×10^{-1} at 0.036, where the prediction is in error by 12.3%: It reports 2.89 where the true value is 2.57. The degradation is monotone in the distance to the fold. This is not a deficiency of tuning, and refitting with more training points in the interior does not repair it. At a fold, the solution branch has a vertical tangent in the $(\lambda, f''(0))$ plane, so $f''(0)$ behaves, as λ_c is approached, like a square root of the distance to the turning point and its derivative diverges. A Gaussian process with a smooth Matérn kernel, or any interpolant whose hypothesis class consists of smooth functions of λ, cannot represent that behavior. It is the wrong function class, not a poorly fitted member of the right one.

A natural response is that a residual-trained network should be preferred here, since its objective is defined wherever the equations are and it can therefore be trained in the near-fold region without a single converged solution there, while the Gaussian process must extrapolate into it. We tested that proposition rather than asserting it. A parametric network mapping (η, λ) to the five fields was trained on the differential residual alone over $\lambda \in [\lambda_c + 0.005, 1]$, and a second copy was trained with six additional anchors at $\lambda \geq -1.4$; both were then queried at the held-out values of Table 12. Table 13 reports the comparison.

Table 13. Prediction of $f''(0)$ near the turning point by a Gaussian process fitted to the solutions obtainable at $\lambda \geq -1.55$, and by parametric networks trained on the differential residual over a range reaching to λ_c. Reference values are taken from the traced first branch.

λ	$\lambda - \lambda_c$	Reference	Gaussian process	Network (residual)	Network (anchored)
−1.70	0.036	2.5748	2.7715	2.8311	2.9380
−1.65	0.086	2.7597	2.8394	2.8785	2.9696
−1.60	0.136	2.8712	2.8973	2.9177	2.9954
−1.55	0.186	2.9446	2.9446	2.9485	3.0150
Mean |error| within 0.2 of the fold			0.0756	0.1064	0.1919

The proposition is not supported. Embedding the equations in the objective does not recover the accuracy lost near the fold: The residual-trained network is about 40% less accurate there than the Gaussian process it was meant to improve upon, and adding anchors makes it worse still, with the anchor loss pulling the fit towards the interior where the response is smooth. The advantage the network was expected to hold, namely that it needs no converged solutions in the region of interest, does not translate into a better prediction in that region.

Two practical conclusions follow. In the interior of the parameter range, where the response is smooth, a classical surrogate is accurate, cheap and entirely sufficient: The network earns nothing there. In the neighborhood of the turning point, which is precisely the region of engineering interest because it marks the operating limit, neither route is reliable, and the system should be solved directly. The difficulty is a property of the fold rather than of the hypothesis class: A square-root singularity in the response defeats a smooth kernel and a tanh network alike. This bounds the claim rather than the method, and a surrogate built on the local structure of a turning point, rather than on generic smoothness, remains the appropriate line of attack.

4.7.4. Architecture and collocation sensitivity

The settings of Appendix A were selected rather than inherited. Two studies were run separately so that the effects do not confound: A depth-by-width grid at fixed collocation count, and a collocation-density sweep at the selected architecture.

The best cell has a depth of 3 and a width of 48 (5 045 parameters), and a depth of 6 is worse than a depth of 3 at every width. The grid supports no finer conclusion than this. Each cell is a single training run, and Section 4.7.1 shows that repeating a fixed architecture across random initializations already moves the error in f by a factor of about six, which is comparable with much of the cell-to-cell variation in Table 14. Resolving an architecture trend would require several initializations per cell.

Table 14. Relative L_2 error in f over the architecture grid, at 160 collocation points and a single seed per cell. Parameter counts range from 1 373 to 21 253.

	Width			
Depth	24	32	48	64
3	5.65×10^{-4}	5.96×10^{-4}	$\mathbf{5.94 \times 10^{-5}}$	2.86×10^{-4}
4	8.58×10^{-4}	3.09×10^{-4}	5.62×10^{-4}	8.21×10^{-5}
5	5.14×10^{-4}	2.57×10^{-4}	5.22×10^{-4}	4.32×10^{-4}
6	2.65×10^{-4}	1.44×10^{-3}	3.89×10^{-4}	1.50×10^{-3}

The collocation study is the cleaner of the two, and it speaks to a specific concern about whether a few hundred points can resolve a tenth-order system.

Both studies are summarized in Figure 18. According to Table 15, the error in f falls by a factor of five between 40 and 160 points and then stops improving: 320 and 640 points give the same error as 160 to three significant figures, while the training time doubles. The count is therefore at the plateau rather than below it. The reason a modest number suffices is that the nodes are Chebyshev-clustered toward the wall, so they are concentrated where the residual is large; the order of the system sets how many derivatives must be evaluated at each node, not how many nodes are needed to resolve a boundary layer occupying a small part of the domain.

Table 15. Collocation density study at a depth of 3 and a width of 48.

n_c	40	80	160	320	640
L_2 error in f	2.84×10^{-4}	1.15×10^{-4}	5.94×10^{-5}	7.35×10^{-5}	7.35×10^{-5}
L_2 error in θ_p	7.14×10^{-3}	1.26×10^{-2}	7.43×10^{-3}	8.24×10^{-3}	8.50×10^{-3}
Training time (s)	58	64	105	166	223

The dust's temperature does not improve with density at all, varying without trend between 7×10^{-3} and 1.3×10^{-2} across a 16-fold increase in nodes. This is consistent with the structural explanation of Section 4.7.1: The limitation on θ_p is the absence of a diffusive term in its governing equation, and no amount of sampling repairs that.

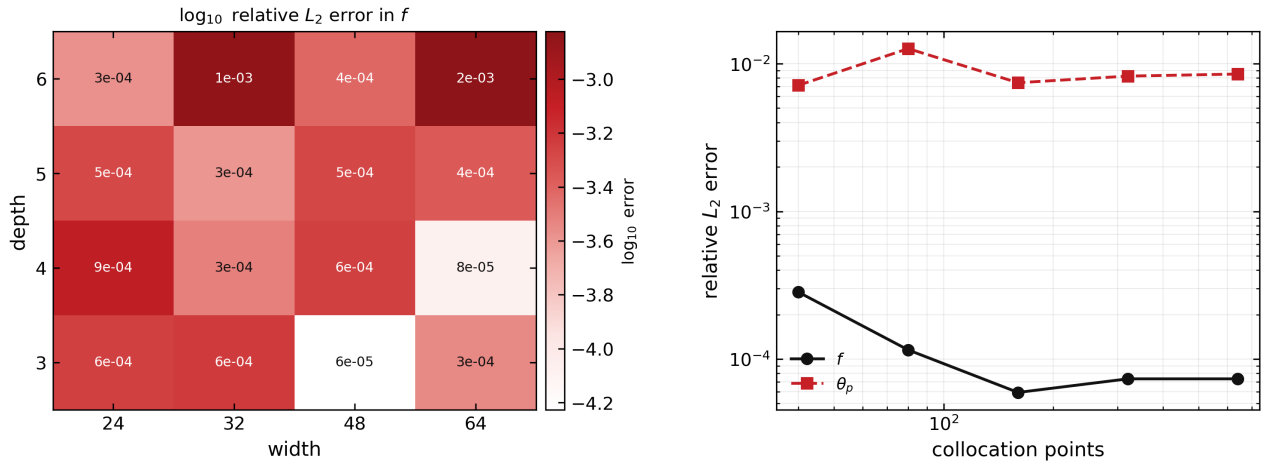


Figure 18. Left: Relative L_2 error in f over the depth-by-width grid. Right: Collocation density convergence for f and θ_p.

4.7.5. Computational cost and break-even

Speed-up figures quoted per query, with the construction cost set aside, are not meaningful for a surrogate that has to be built before it can be used. We therefore report the total cost. If we write T for construction, q for the cost of one query, and s for one direct solution, a surrogate is cheaper than the repeated direct solution only after

$$n^\star = \frac{T}{s - q} \tag{4.2}$$

evaluations. Table 16 collects the three routes on the same basis. All timings are single-machine central processing units (CPUs); no graphics processing units (GPUs) were used.

Table 16. Total cost of each route to $f''(0)$ over a parameter range. Query costs for both surrogates are below the measurement resolution. Here, — denotes not applicable: Direct solution has no construction stage and therefore no break-even point.

	Construction T	Per query q	Break-even $n^\star$
Direct collocation solution	—	0.264 s	—
Gaussian process (10 solutions)	2.8 s	$< 10^{-5}$ s	≈ 11
Network (ODE residual)	231 s	$< 10^{-4}$ s	≈ 875

The Gaussian process repays its construction after about 11 evaluations and the network after about 875, a factor of roughly 80 between them. For a study of the size reported in this paper (a few dozen operating points), the Gaussian process is the economical choice and the direct solver is entirely competitive, and the network is not justified on cost at this scale. Only for campaigns of many hundreds of evaluations does the network's construction cost amortize, and even then Section 4.7.2 shows it must also beat the Gaussian process on accuracy to be worth choosing.

4.7.6. Branch reconstruction

A network trained on the ODE residual alone has no mechanism for selecting between coexisting solutions: It converges to whichever branch its initialization favors. Reconstructing a chosen branch requires a supervised pre-fit to a reference profile for that branch, together with an anchor on $f''(0)$ to prevent the optimizer from drifting back to the other. The reference solution is therefore a prerequisite of the procedure, not a product of it, and the capability is properly described as branch reconstruction under supervision rather than as branch discovery.

4.8. Sensitivity analysis (response surface methodology, RSM)

A quadratic response surface was fitted over the three groups that Section 4.5 identifies as most influential: The curvature γ, the radiation parameter R_d, and the Brownian parameter N_b. The design is a face-centered central composite in these three factors, giving 15 points, all of which converged and satisfied the admissibility test.

A word on what such a fit can and cannot support. The responses here are deterministic outputs of a boundary value solver. Repeating a design point reproduces it to solver tolerance, and there is no replication, no measurement error, and no sampling variability. The residual of a fitted polynomial therefore measures how well a quadratic approximates a smooth deterministic function, and nothing else. Analysis of variance p-values and F-statistics computed from such residuals do not carry their usual inferential meaning, because the null hypothesis they test (namely that an effect is indistinguishable from noise) has no counterpart when there is no noise. We consequently report the goodness-of-fit and cross-validated predictive error, and no significance tests, and we treat the coefficients as a sensitivity ranking rather than as estimates with confidence intervals.

Table 17 collects the fitted surfaces, and Figure 19 gives the corresponding rankings. The Brownian parameter dominates both the transfer coefficients and acts on them in opposite directions, lowering Nu and raising Sh; radiation is second and raises both; curvature is a distant third within this design, which is consistent with Section 4.5 and is the reason why the curvature contribution of this paper is framed in terms of admissibility rather than of transfer enhancement.

Table 17. Quadratic response surfaces over (γ, R_d, N_b), with 15 design points. Leave-one-out (LOO) error is the honest measure of predictive quality for deterministic data. Linear coefficients are in coded units and are reported as a ranking only. RMS, root mean square.

Response	R^2	Adjusted R^2	RMS residual	LOO RMS	Ranking (linear coefficient)
Nu	0.99947	0.99853	1.68×10^{-2}	5.06×10^{-2}	N_b (−0.711), R_d (+0.518), γ (+0.047)
Sh	0.98014	0.94439	2.99×10^{-1}	8.88×10^{-1}	N_b (+2.010), R_d (+0.952), γ (+0.049)
C_f	1.00000	1.00000	2.40×10^{-13}	9.39×10^{-13}	γ (−0.041), R_d (0), N_b (0)

The skin-friction row is worth a moment. Its surface is exact to machine precision, with a residual of 2.4×10^{-13} and coefficients on R_d and N_b that are identically zero. This is not a fitting artefact: C_f depends on none of the thermal or species groups, so within this design, it is a function of γ alone and a quadratic reproduces it exactly. It is the one-way coupling of Section 4.5 appearing a third time, and serves as an independent check that the design and the fitting procedure behave as intended.

The gap between the root-mean-square (RMS) residual and the leave-one-out (LOO) error (roughly a factor of three for both Nu and Sh) is the usual signature of a quadratic fitted close to its capacity on 15 points. It quantifies how much of the apparent fit quality would not survive at an unseen design point, which is precisely the quantity a deterministic study should report in place of a significance test.

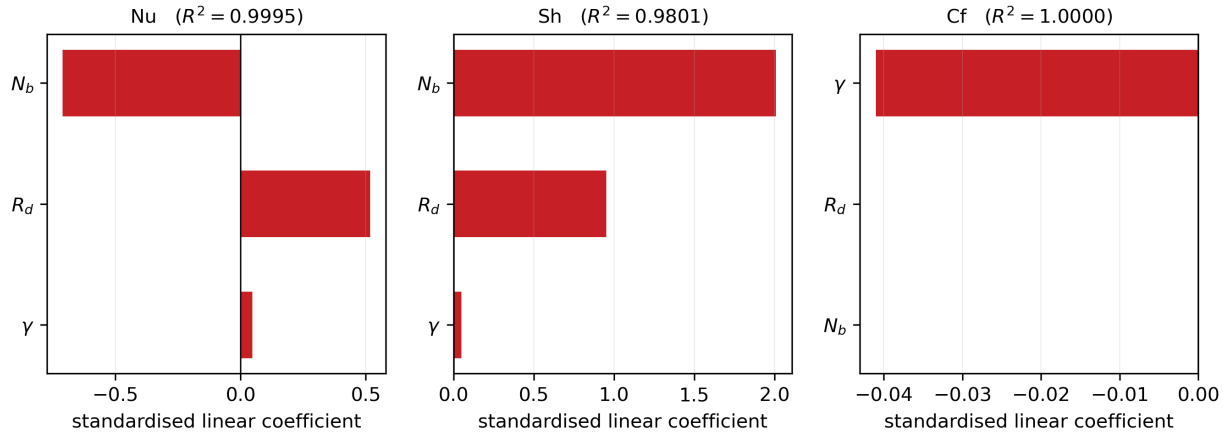


Figure 19. Standardized linear coefficients for each response, ordered by magnitude. Bars are a sensitivity ranking, not effect estimates with confidence intervals.

4.9. Physical insights and applications

The results bear on the drawing configuration described in Section 1 in three ways.

First, the admissibility envelope of Section 4.2 is an operating constraint rather than a numerical technicality. The Maxwell number β is fixed by the melt and the draw rate, and the suction parameter S by the extraction system, and Table 4 states which combinations of the two admit a steady boundary layer at all. A line operating at $S = 2.5$ on a plane geometry is restricted to $\beta \leq 0.05$, whereas the same conditions on a cylinder of curvature $\gamma = 0.3$ admit $\beta \leq 0.15$. For a given melt, this is the difference between a feasible and an infeasible draw.

Second, the critical shrinking rate bounds the decelerating regime. Beyond λ_c, no similarity solution exists, and the stability analysis of Section 4.4 shows that only the upper of the two coexisting states is realizable. A process operated on the lower branch would not persist, so the useful window is the interval between λ_c and zero on the first branch alone. Suction widens that window substantially and curvature narrows it, so the two must be traded against one another rather than chosen independently.

Third, the transport results identify which parameters are worth manipulating. Radiation is the dominant control on the wall heat flux, raising Nu by a third across the range considered, whereas the magnetic and porosity groups act on the momentum field alone and leave Nu and Sh essentially unchanged. Dust loading raises the heat transfer while lowering the mass transfer, so a particle-laden carrier cannot be optimized for both simultaneously, and the choice depends on whether the limiting process is cooling of the coating or diffusion of the reacting species into it.

5. Conclusions

Steady two-phase flow of a radiative, MHD dusty Maxwell nanofluid over a stretching/shrinking cylinder in a Darcy–Forchheimer medium has been formulated, reduced, and solved, with attention to the shrinking regime in which the boundary layer admits coexisting states. The principal findings are as follows.

1. **Curvature controls the multiplicity structure.** The critical shrinking rate rises monotonically with the curvature parameter, from $\lambda_c = -1.8336$ on a plane sheet to $\lambda_c = -1.6840$ at $\gamma = 0.5$ under otherwise identical conditions. Transverse curvature therefore narrows the window of shrinking rates over which dual solutions persist, and does so smoothly enough that γ can be read as a design variable rather than a perturbation. Suction acts in the opposite sense, widening the window, so the two effects can be traded against one another.
2. **The upper branch is stable and the lower unstable.** A linear temporal stability analysis of the complete disturbance system returns a positive smallest eigenvalue on the upper branch and a negative one on the lower, with the eigenvalue crossing zero at the turning point. The linearized operator is block lower-triangular, so the customary restriction of such analyses to the momentum–dust subsystem is exact rather than approximate, a point that is usually assumed and here established and verified by computing both blocks.
3. **The Maxwell parameter is bounded by an admissibility condition.** The coefficient $(1 + 2\gamma\eta) - \beta f^2$ must remain positive, failing which, a collocation solver converges to a spurious state with an unphysical far field. On a plane sheet with $S = 2$, this bounds β below roughly 0.2. Curvature relaxes the bound, so operating points that are inadmissible on a plane sheet become attainable on a cylinder.
4. **Published benchmark values carry a truncation error.** The present solver reproduces the standard stretching cylinder skin friction data to within 0.004% when the domain truncation is matched, but the domain-converged values differ from them by up to 1.4%. The discrepancy is attributable to a common truncation at $\eta_\infty \approx 10$ rather than to any difference in the formulation.
5. **A physics-embedded surrogate is not justified by cost, but by location.** For smooth scalar outputs over a parameter interval, a Gaussian process fitted to 10 direct solutions attains 0.18% accuracy for a total construction cost of under 3 s and repays itself after roughly 11 evaluations, which is far cheaper than training a network. Approaching the turning point, however, that same surrogate degrades by nearly three orders of magnitude and is in error by 12.3% at a distance of 0.036 from the fold, because the branch has a vertical tangent there, and smooth interpolation of solver outputs is the wrong hypothesis class. The case for carrying the governing equations in the surrogate rests on this neighborhood, which is also the region of engineering interest.
6. **Branch reconstruction requires supervision.** A network trained on the ODE residual alone cannot select between coexisting solutions; a reference profile and an $f''(0)$ anchor are prerequisites. The capability is reconstruction, not discovery.

5.1. Limitations

The model is steady, laminar, and axisymmetric, and rests on a local similarity reduction; the Forchheimer group is consequently local, and the results are reported at a fixed axial station. No

experimental validation is offered, and the value of the study is comparative and mechanistic rather than predictive for any particular device. The two solid populations are dilute, so neither modifies the carrier density or viscosity, and no direct dust–nanoparticle exchange is represented.

The two-phase closure also bounds the curvature range over which the dual-solution structure can be examined at all. Because g multiplies g'' in the dust momentum equation and $g(\infty) = f(\infty)$, the reduced system degenerates once $f(\infty)$ falls to zero, and for $\gamma \gtrsim 0.6$, this occurs before the turning point is reached (Section 4.3). Whether this represents genuine dust phase behavior or is an artefact of the Saffman closure combined with that far-field condition is not settled here, and the results are reported for the range in which the fold exists.

In the stability analysis, the $O(\beta)$ unsteady Maxwell relaxation contributions are omitted; retaining them would convert the disturbance problem into a quadratic eigenvalue problem. The omission is consistent with the boundary layer scaling but has not been quantified here. The response surface analysis of Section 4.8 approximates deterministic solvers' output, so its residual statistics measure the quality of a polynomial approximation and support no statistical inference.

5.2. Future work

Three directions follow directly. The unsteady Maxwell terms should be retained so that the disturbance spectrum can be computed without approximation. The behavior of surrogate models in the neighborhood of a fold deserves separate treatment: An interpolant built on the square-root local structure of a turning point, rather than on smooth kernels, would likely outperform both approaches compared here. Finally, the curvature–suction trade-off identified in Section 4.3 invites an optimization study over the admissible region rather than the parameter sweep presentation adopted here.

Appendix.

A. Reproducibility protocol for the network

The settings below are stated in full so that the results of Section 4.7 can be reproduced exactly. The collocation count is the outcome of the density study of Section 4.7.4, which places 160 nodes at the plateau.

The results of Section 4.7.1 are computed at a depth of 5 and a width of 48. The architecture scan of Section 4.7.4 shows the attainable accuracy to be insensitive to depth and width over the range tested, with the spread between cells being comparable with the spread between random initializations of a fixed architecture, so this choice is not critical.

The complete settings are collected in Table 18.

No normalization of the input or outputs was applied: η is $O(1)$–$O(10)$ and every output field is $O(1)$ by construction of the similarity variables, so rescaling changes nothing measurable. There is no train/validation split in the usual sense, because the objective is a differential residual rather than a data-fitting loss; the collocation nodes are the training set and the held-out assessment is the comparison against the collocation solver on a uniform grid of 400 points, reported in Table 10. Where a reference profile is used for branch reconstruction (Section 4.7.6), it enters as an additional supervised term and is stated explicitly at that point.

Table 18. Network, discretization and training settings. IEEE, Institute of Electrical and Electronics Engineers; L-BFGS, limited-memory Broyden–Fletcher–Goldfarb–Shanno.

Architecture	Fully connected, tanh activation
Hidden layers	5
Neurons per hidden layer	48, 48, 48, 48, 48
Inputs/outputs	η / $(f, g, \theta, \varphi, \theta_p)$
Trainable parameters	9 749
Precision	IEEE double (float64)
Collocation nodes	160
Node distribution	Chebyshev, clustered toward the wall
Domain	$\eta \in [0, \eta_\infty]$, $\eta_\infty = 8$
Derivatives	Automatic differentiation (exact, not finite difference)
Loss	$\sum_i \langle r_i^2 \rangle + w_{\mathrm{bc}} \langle r_{\mathrm{bc}}^2 \rangle$
Boundary weight w_{bc}	50
Residuals r_i	The five equations (2.9)–(2.12)
Boundary residuals	The 10 conditions, at $\eta = 0$ and $\eta = \eta_\infty$
First-stage optimizer	Adam, learning rate 10^{-3}, 2 500 iterations
Second-stage optimizer	L-BFGS, up to 500 iterations
Line search	Strong Wolfe
History size	60
Gradient tolerance	10^{-12}
Change tolerance	10^{-14}
Random seeds	$0, 1, \ldots, 9$ (10 independent runs)
Hardware	CPU only; no GPU was available or used
Mean training time	231 s per run

Author contributions

Bilal Ahmad and Muhammad Abubakar: Write original draft; Amber Sulatan and Baqir Hussain: Formal analysis, software; Yazeed Yasin Ghadi: Supervision, review and editing. All authors have read and approved the final version of the manuscript for publication.

Use of generative AI tools declaration

The authors declare they have not used artificial intelligence (AI) tools in the creation of this article.

Data availability

The solver, the stability eigenvalue code, the surrogate training and comparison scripts, the continuation data, and all configuration files required to reproduce every figure and table in this paper are available in a public repository.

Conflict of interest

The authors declare no conflict of interest.

References

1. M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.,* **378** (2019), 686–707. https://doi.org/10.1016/j.jcp.2018.10.045
2. S. Cai, Z. Mao, Z. Wang, M. Yin, G. E. Karniadakis, Physics-informed neural networks (PINNs) for fluid mechanics: A review, *Acta Mech. Sinica,* **37** (2021), 1727–1738. https://doi.org/10.1007/s10409-021-01148-1
3. V. Poply, P. Singh, K. K. Chaudhary, Analysis of laminar boundary layer flow along a stretching cylinder in the presence of thermal radiation, *WSEAS Tran. Fluid Mech.,* **8** (2013), 159–164.
4. P. T. Manjunatha, B. J. Gireesha, B. C. Prasannakumara, Thermal analysis of conducting dusty fluid flow in a porous medium over a stretching cylinder in the presence of non-uniform source/sink, *Int. J. Mech. Mater. Eng.,* **9** (2014), 13.
5. Hashim, M. Khan, A. S. Alshomrani, Characteristics of melting heat transfer during flow of Carreau fluid induced by a stretching cylinder, *Eur. Phys. J. E,* **40** (2017), 1–9.
6. A. Mahdy, F. M. Hady, R. A. Mohamed, O. A. Abo-zaid, Activation energy effectiveness in dusty Carreau fluid flow along a stretched cylinder due to nonuniform thermal conductivity property and temperature-dependent heat source/sink, *Heat Transfer,* **50** (2021), 5760–5778. https://doi.org/10.1002/htj.22147
7. N. Abbas, W. Shatanawi, F. Hasan, T. A. M. Shatnawi, Numerical analysis of Darcy resistant Sutterby nanofluid flow with effect of radiation and chemical reaction over stretching cylinder: Induced magnetic field, *AIMS Math.,* **8** (2023), 11202–11220. https://doi.org/10.3934/math.2023567
8. R. K. Alhefthi, I. Shahzadi, H. A. Khan, N. Khan, M. S. Hashmi, M. Inc, Convective boundary layer flow of MHD tangent hyperbolic nanofluid over stratified sheet with chemical reaction, *AIMS Math.,* **9** (2024), 16901–16923. https://doi.org/10.3934/math.2024821
9. K. Sadeghy, H. Hajibeygi, S. M. Taghavi, Stagnation-point flow of upper-convected Maxwell fluids, *Int. J. Non-Lin. Mech.,* **41** (2006), 1242–1247. https://doi.org/10.1016/j.ijnonlinmec.2006.08.005
10. J. Buongiorno, Convective transport in nanofluids, *J. Heat Tran.,* **128** (2006), 240–250. https://doi.org/10.1115/1.2150834
11. W. A. Khan, I. Pop, Boundary-layer flow of a nanofluid past a stretching sheet, *Int. J. Heat Mass Tran.,* **53** (2010), 2477–2483. https://doi.org/10.1016/j.ijheatmasstransfer.2010.01.032
12. B. Mahanthesh, J. Mackolil, M. Radhika, Significance of quadratic thermal radiation and quadratic convection on boundary layer two-phase flow of a dusty nanoliquid past a vertical plate, *Int. Commun. Heat Mass Tran.,* **120** (2021), 105029. https://doi.org/10.1016/j.icheatmasstransfer.2020.105029

13. M. Ramzan, N. Shaheen, J. D. Chung, S. Kadry, Y. M. Chu, F. Howari, Impact of Newtonian heating and Fourier and Fick's laws on a magnetohydrodynamic dusty Casson nanofluid flow with variable heat source/sink over a stretching cylinder, *Sci. Rep.,* **11** (2021), 2357. https://doi.org/10.1038/s41598-021-81747-x
14. M. Miklavčič, C. Y. Wang, Viscous flow due to a shrinking sheet, *Q. Appl. Math.,* **64** (2006), 283–290. https://doi.org/10.1090/s0033-569x-06-01002-5
15. Y. Ouyang, M. F. Md Basir, K. Naganthran, I. Pop, Dual solutions in Maxwell ternary nanofluid flow with viscous dissipation and velocity slip past a stretching/shrinking sheet, *Alex. Eng. J.,* **105** (2024), 437–448. https://doi.org/10.1016/j.aej.2024.07.093
16. U. Yashkun, L. A. Lund, M. A. Fadhel, N. A. Shah, Velocity slip effect on magnetized Casson nanofluid over shrinking/stretching cylinder: Duality and stability analysis, *Phys. Scripta,* **99** (2024), 025203. https://doi.org/10.1088/1402-4896/ad1795
17. J. H. Merkin, On dual solutions occurring in mixed convection in a porous medium, *J. Eng. Math.,* **20** (1986), 171–179. https://doi.org/10.1007/bf00042775
18. P. D. Weidman, D. G. Kubitschek, A. M. J. Davis, The effect of transpiration on self-similar boundary layer flow over moving surfaces, *Int. J. Eng. Sci.,* **44** (2006), 730–737. https://doi.org/10.1016/j.ijengsci.2006.04.005
19. G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, *Nat. Rev. Phys.,* **3** (2021), 422–440. https://doi.org/10.1038/s42254-021-00314-5
20. Z. Zou, Z. Wang, G. E. Karniadakis, Learning and discovering multiple solutions using physics-informed neural networks with random initialization and deep ensemble, *P. Roy. Soc. A- Math. Phys.,* **481** (2025). https://doi.org/10.1098/rspa.2025.0205
21. G. Ramasekhar, S. Alkarni, N. A. Shah, Machine learning approach of Casson hybrid nanofluid flow over a heated stretching surface, *AIMS Math.,* **9** (2024), 18746–18762. https://doi.org/10.3934/math.2024912
22. A. H. Alharbi, E. H. Boussaha, A. M. Alhartomi, H. Salhi, R. Alrdadi, M. Kezzar, M. R. Sari, AI-driven analysis of Darcy–Forchheimer EMHD boundary layer flow in thermal hybrid nanofluids, *AIMS Math.,* **10** (2025), 27129–27151. https://doi.org/10.3934/math.20251192
23. S. Wang, Y. Teng, P. Perdikaris, Understanding and mitigating gradient flow pathologies in physics-informed neural networks, *SIAM J. Sci. Comput.,* **43** (2021), A3055–A3081. https://doi.org/10.1137/20M1318043
24. X. Huang, F. Wang, B. Zhang, H. Liu, Enriched physics-informed neural networks for dynamic Poisson–Nernst–Planck systems, *Math. Comput. Simulat.,* **237** (2025), 231–246. https://doi.org/10.1016/j.matcom.2025.04.037
25. X. Li, F. Wang, R. Wang, S. Zhao, D. Yang, Fundamental solution neural networks for solving inverse Cauchy problems for the Laplace and biharmonic equations, *Comput. Math. Appl.,* **201** (2026), 1–17. https://doi.org/10.1016/j.camwa.2025.09.032



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)