



A Heuristic Physics-Informed Neural Network for Unsteady Magnetohydrodynamic Squeezing Flow through a Darcy-Forchheimer Porous Channel with Brownian Motion, Thermophoresis, Thermal Radiation, and Heat Generation

Umar Daraz¹, Khaleel Ahmad^{2,3,*}

¹ Department of Physics, University of Education Lahore, Pakistan

² Department of Mathematics and Statistics, The University of Lahore, Pakistan

³ Center for Theoretical Physics, Khazar University, 41 Mehseti Str., Baku, AZ1096, Azerbaijan

ARTICLE INFO

Article history:

Received 7 June 2026

Received in revised form 19 September 2026

Accepted 23 September 2026

Available online 26 September 2026

Keywords:

Hybrid Nanofluid Squeezing Flow; Darcy-Forchheimer Magnetohydrodynamics; Buongiorno Nanoparticle Transport; Physics-Informed Neural Network; Green Function Fixed Point Analysis; Particle Swarm and Neural Network Algorithm

ABSTRACT

The unsteady two-dimensional magnetohydrodynamic squeezing flow of an electrically conducting hybrid nanofluid through a Darcy-Forchheimer porous channel is studied under the Buongiorno Brownian and thermophoretic fluxes, nonlinear Rosseland radiation, and a temperature-dependent heat source or sink. A symbolically confirmed similarity transformation reduces the governing partial differential equations to a coupled nonlinear ordinary differential system, which is recast as a fixed point problem through explicit Green functions. Existence, uniqueness, and an a posteriori error certificate are established. The fields are then obtained by a heuristic physics-informed neural network trained by a hybrid particle swarm and neural network algorithm with a trust-region refinement. The influence of every physical parameter on the velocity, temperature, and concentration is analysed.

*Corresponding author.

E-mail address: khalil7066616@gmail.com

<https://doi.org/10.66972/iscis202617>

© The Author(s) 2026 | [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

1. Introduction

The transport of energy in a thin film of fluid that is compressed between two approaching surfaces is a recurring configuration in lubrication, polymer processing, hydraulic machinery and biomedical devices. The classical study of such squeezing motion goes back to the work of Stefan [1] and the exact unsteady solutions catalogued by Wang [2]. When the working fluid carries a dilute suspension of metallic or ceramic nanoparticles, its effective thermal conductivity is raised well beyond that of the carrier liquid, a possibility first articulated by Choi & Eastman [3] and since developed into a large body of modelling work. Two complementary descriptions of nanofluids dominate the literature, namely the single phase dispersion model of Tiwari & Das [4] in which the suspension is treated as a homogeneous medium with modified properties, and the two phase model of Buongiorno [5] in which the slip mechanisms of Brownian diffusion and thermophoresis are resolved explicitly. The boundary layer form of the latter model was reduced to a tractable similarity system by Khan & Pop Khan and Pop [6], whose treatment underlies much of the subsequent analysis of nanoparticle volume fraction transport, including the nonlinear stretching sheet computations of Rana & Bhargava Rana and Bhargava [7].

Hybrid nanofluids, which disperse two distinct species of nanoparticle in a single base liquid, extend this idea by combining the favourable attributes of both fillers; the synthesis and modelling of such suspensions was advanced in [8, 9], while the effective property correlations needed to close the governing equations descend from the conduction analysis [10], the shape dependent conductivity model [11], and the suspension viscosity [12]. The present study adopts magnetite (Fe_3O_4) and molybdenum disulphide (MoS_2) as the two solid phases in an ethylene glycol and water base, a pairing that couples a strongly magnetic ferrite with a high conductivity layered material.

The squeezing of a nanofluid film in the presence of a magnetic field has attracted sustained attention because the Lorentz force offers a contactless means of regulating the flow rate and the wall heat flux. Representative contributions include the analytical and numerical treatments [13–15], the slip flow solutions [16], and the two phase radiative analysis [17]. When the channel is filled with a porous matrix the resistance to motion acquires a linear Darcy contribution together with a quadratic inertial correction first identified by Forchheimer [18], a balance that is described in detail in [19]. Radiative transport in optically thick media is customarily linearised through the Rosseland approximation [20], and additional volumetric heating or cooling is represented by a temperature dependent source term. Further studies that inform the modelling choices made here include the porous enclosure computations [21], the Hall current analysis [22] and the generalised slip treatment [23].

The resulting boundary value problems are stiff and strongly coupled. They are most often solved by shooting or by the residual controlled collocation solver `bvp4c` [24], whose theoretical basis is described in [25]. In recent years a complementary solution paradigm has emerged in which the unknown fields are represented by an artificial neural network and the governing equations are enforced directly on the network output. The universal approximation property that justifies this representation was established in [26, 27], and the use of networks as differential equation solvers was pioneered in [28]. The modern physics-informed neural network [29, 30], embeds the residual in a loss functional that is minimised by gradient descent with automatic differentiation [31–33]. The approximation theory of these solvers has matured in parallel, and rigorous bounds on the generalisation and training error are now available through the analyses in [34–36], which relate the attainable accuracy to the residual that the network is able to drive to zero.

For the small shallow networks that arise in similarity reduced problems, a derivative free global search is often more reliable than gradient descent, and the present work therefore couples particle swarm optimisation [37–39] with the neural network algorithm [40], a population based method inspired by the propagation of signals in biological nervous systems. The hybrid search is polished by a

trust region step in the spirit of the classical least squares method of Levenberg [41] and Marquardt [42], which is also used on its own to train a supervised surrogate that yields the familiar regression and error histogram diagnostics. The accuracy of the stochastic training is summarised through the mean squared error and through the inequality coefficient of Theil [43].

Alongside the numerical treatment, the well posedness of the reduced model deserves a rigorous footing, because a neural solver returns a number whether or not the underlying problem possesses a unique solution near the computed state. The classical instruments for this purpose are the contraction principle of Banach [44] and the compactness based fixed point theorem of Schauder [45], both presented in the modern monographs [46, 47] together with the Green function representation of boundary value problems treated in [48]. These tools allow the similarity system to be written as a fixed point equation whose solvability, uniqueness and conditioning can be quantified, and they connect the residual that the network minimises to a certified bound on the distance to the exact solution. The methodology of the present paper extends the heuristic physics-informed framework recently proposed in [49] for squeezing flow, to which several further physical mechanisms are added here, namely Darcy and Forchheimer porous resistance, Brownian and thermophoretic transport, nonlinear radiation and volumetric heat generation, together with a Green function existence, uniqueness and stability analysis, an extensive statistical study and a symbolic verification of the computed solution. The overall structure of the study, from the governing balances through the analysis to the learned and reference solutions, is summarised in Figure 1.

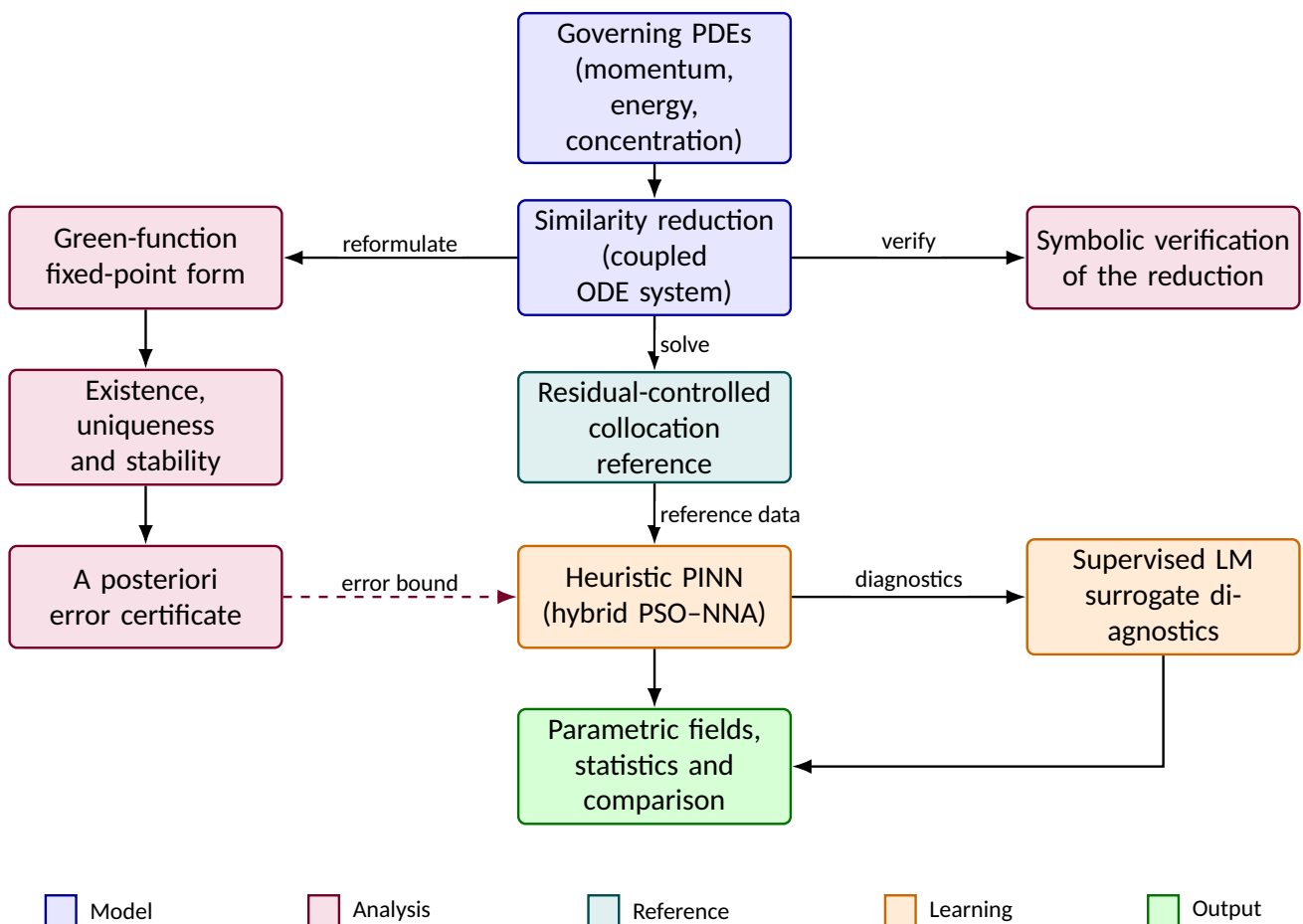


Fig. 1. Overall flow of the study, linking the governing model, the Green function existence, uniqueness and stability analysis, the residual controlled reference, the heuristic physics-informed network and the supervised surrogate to the final parametric and statistical results

2. Mathematical Formulation

Consider the unsteady, laminar and incompressible flow of an electrically conducting hybrid nanofluid confined between two horizontal parallel plates separated by the time dependent gap $2h(t)$, where $h(t) = H\sqrt{1 - \alpha t}$ and $\alpha t < 1$. The lower plate is fixed while the upper plate approaches it, so that the film is squeezed; the geometry and the principal physical mechanisms are sketched in Figure 2. A uniform magnetic field of strength $B_0/\sqrt{1 - \alpha t}$ acts transverse to the plates, the channel is saturated by a homogeneous porous matrix of permeability k_1 with Forchheimer inertial coefficient C_b , and the suspension carries Fe_3O_4 and MoS_2 nanoparticles in an ethylene glycol and water base.

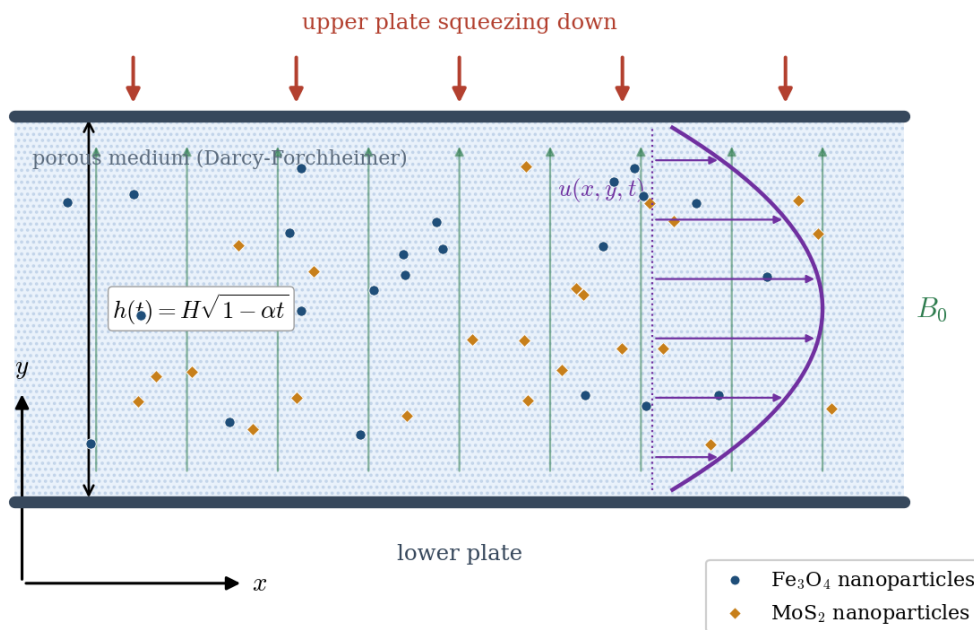


Fig. 2. Geometry of the problem: the hybrid nanofluid is squeezed through a Darcy–Forchheimer porous channel between two parallel plates under a transverse magnetic field, with Brownian and thermophoretic transport of the dispersed Fe_3O_4 and MoS_2 nanoparticles

With u and v the velocity components along the x and y axes, conservation of mass requires:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \quad (1)$$

The streamwise momentum balance, including the Lorentz force, the Darcy drag and the quadratic Forchheimer resistance, reads:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -\frac{1}{\rho_{hnf}} \frac{\partial p}{\partial x} + \frac{\mu_{hnf}}{\rho_{hnf}} \frac{\partial^2 u}{\partial y^2} - \frac{\sigma_{hnf} B_0^2}{\rho_{hnf}(1 - \alpha t)} u - \frac{\mu_{hnf}}{\rho_{hnf} k_1} u - \frac{C_b}{\sqrt{k_1}} u^2 \quad (2)$$

Energy transport with Rosseland radiation, the Buongiorno slip fluxes and a temperature dependent heat source or sink obeys:

$$\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} = \frac{k_{hnf}}{(\rho c_p)_{hnf}} \frac{\partial^2 T}{\partial y^2} - \frac{1}{(\rho c_p)_{hnf}} \frac{\partial q_r}{\partial y} + \tau \left[D_B \frac{\partial C}{\partial y} \frac{\partial T}{\partial y} + \frac{D_T}{T_m} \left(\frac{\partial T}{\partial y} \right)^2 \right] + \frac{Q_0}{(\rho c_p)_{hnf}} (T - T_h) \quad (3)$$

where the radiative flux is linearised as $q_r = -(16\sigma^*T_\infty^3/3k^*) \partial T/\partial y$, and $\tau = (\rho c_p)_p/(\rho c_p)_f$ is the ratio of the effective heat capacity of the nanoparticle material to that of the base fluid. The nanoparticle volume fraction is governed by:

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} + v \frac{\partial C}{\partial y} = D_B \frac{\partial^2 C}{\partial y^2} + \frac{D_T}{T_m} \frac{\partial^2 T}{\partial y^2} \quad (4)$$

The boundary conditions express symmetry at the channel centre and the moving impermeable upper wall:

$$\begin{aligned} \frac{\partial u}{\partial y} = 0, \quad v = 0, \quad \frac{\partial T}{\partial y} = 0, \quad \frac{\partial C}{\partial y} = 0 \quad \text{at } y = 0, \\ u = 0, \quad v = \frac{dh}{dt}, \quad T = T_h, \quad C = C_h \quad \text{at } y = h(t) \end{aligned} \quad (5)$$

To reduce the system we introduce the similarity variable and the dimensionless fields:

$$\eta = \frac{y}{H\sqrt{1-\alpha t}}, \quad u = \frac{\alpha x}{2(1-\alpha t)} f'(\eta), \quad v = -\frac{\alpha H}{2\sqrt{1-\alpha t}} f(\eta), \quad \theta = \frac{T}{T_h}, \quad \phi = \frac{C}{C_h} \quad (6)$$

For brevity write $\beta = 1 - \alpha t$. The similarity variable $\eta = y/(H\sqrt{\beta})$ obeys:

$$\frac{\partial \eta}{\partial y} = \frac{1}{H\sqrt{\beta}}, \quad \frac{\partial \eta}{\partial t} = \frac{\alpha \eta}{2\beta} \quad (7)$$

and differentiating the velocity components of Eq. (6) gives:

$$\frac{\partial u}{\partial x} = \frac{\alpha}{2\beta} f', \quad \frac{\partial u}{\partial y} = \frac{\alpha x}{2\beta H\sqrt{\beta}} f'', \quad \frac{\partial v}{\partial y} = -\frac{\alpha}{2\beta} f' \quad (8)$$

whence $\partial u/\partial x + \partial v/\partial y = 0$, so that continuity (1) is met for every choice of f . Carrying the same rule through the convective derivative of the streamwise velocity, and using $\partial u/\partial t = (\alpha^2 x/2\beta^2)[f' + \frac{1}{2}\eta f'']$, $u \partial u/\partial x = (\alpha^2 x/4\beta^2)(f')^2$ and $v \partial u/\partial y = -(\alpha^2 x/4\beta^2)f f''$, collects the inertia into:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = \frac{\alpha^2 x}{4\beta^2} [2f' + \eta f'' + (f')^2 - f f''] \quad (9)$$

while the viscous, Lorentz, Darcy and Forchheimer terms of Eq. (2) take the forms:

$$\begin{aligned} \frac{\mu_{hnf}}{\rho_{hnf}} \frac{\partial^2 u}{\partial y^2} &= \frac{\nu_{hnf} \alpha x}{2H^2 \beta^2} f''', & \frac{\sigma_{hnf} B_0^2}{\rho_{hnf} \beta} u &= \frac{\sigma_{hnf} B_0^2}{\rho_{hnf}} \frac{\alpha x}{2\beta^2} f', \\ \frac{\mu_{hnf}}{\rho_{hnf} k_1} u &= \frac{\nu_{hnf} \alpha x}{k_1} \frac{f'}{2\beta}, & \frac{C_b}{\sqrt{k_1}} u^2 &= \frac{C_b}{\sqrt{k_1}} \frac{\alpha^2 x^2}{4\beta^2} (f')^2 \end{aligned} \quad (10)$$

Dividing the streamwise balance (2) by the viscous scale $\nu_{hnf} \alpha x/(2H^2 \beta^2)$ and reading off the four coefficient ratios:

$$\frac{\alpha H^2}{2\nu_{hnf}} = SA_1, \quad \frac{\sigma_{hnf} B_0^2 H^2}{\mu_{hnf}} = \text{Ha}^2 A_2, \quad \frac{H^2 \beta}{k_1} = K_p, \quad \frac{C_b x}{\sqrt{k_1} \beta} = \text{Fr} \quad (11)$$

recasts it as:

$$f''' - SA_1 [2f' + \eta f'' + (f')^2 - f f''] - (\text{Ha}^2 A_2 + K_p) f' - \text{Fr} (f')^2 = \frac{2H^2 \beta^2}{\rho_{hnf} \nu_{hnf} \alpha x} \frac{\partial p}{\partial x} \quad (12)$$

The pressure is eliminated by differentiating Eq. (12) once with respect to η , which is the similarity image of the transverse cross differentiation ∂_y of the momentum equation and removes the streamwise gradient $\partial p/\partial x$. With $\partial_\eta[2f' + \eta f'' + (f')^2 - f f'''] = 3f'' + \eta f''' + f' f'' - f f'''$ and $\partial_\eta(f')^2 = 2f' f''$, and the factor of two of the Forchheimer term absorbed into Fr , this yields the fourth order momentum equation of (16).

The thermal and solutal balances reduce by the same substitution. Writing $T = T_h \theta$ and $C = C_h \phi$, both fields depend on x only through η , so the streamwise convection vanishes an:

$$\frac{\partial T}{\partial t} + v \frac{\partial T}{\partial y} = -\frac{\alpha T_h}{2\beta}(f - \eta)\theta', \quad \frac{\partial C}{\partial t} + v \frac{\partial C}{\partial y} = -\frac{\alpha C_h}{2\beta}(f - \eta)\phi' \quad (13)$$

The linearised radiative flux $q_r = -(16\sigma^* T_\infty^3/3k^*) \partial T/\partial y$ enters the energy equation as:

$$-\frac{1}{(\rho c_p)_{hnf}} \frac{\partial q_r}{\partial y} = \frac{16\sigma^* T_\infty^3}{3k^*(\rho c_p)_{hnf}} \frac{\partial^2 T}{\partial y^2} \quad (14)$$

which merges with Fourier conduction to place the factor $1 + R_d$ on θ'' . Dividing the energy balance (3) by $k_{hnf} T_h / [(\rho c_p)_{hnf} H^2 \beta]$ and the concentration balance (4) by $D_B C_h / (H^2 \beta)$, and identifying:

$$\frac{\alpha H^2}{2} \frac{(\rho c_p)_{hnf}}{k_{hnf}} = \text{Pr} S A_3, \quad \frac{\alpha H^2}{2 D_B} = S_c S, \quad \frac{D_T T_h}{T_m D_B C_h} = \frac{N_t}{N_b} \quad (15)$$

reproduces the reported thermal and solutal equations. The wall data transform through $dh/dt = -\alpha H/(2\sqrt{\beta})$ and $\eta(h) = 1$: the kinematic condition $v = dh/dt$ gives $f(1) = 1$, the no slip condition $u = 0$ gives $f'(1) = 0$, and the centreline symmetry gives $f(0) = 0$, $f''(0) = 0$, $\theta'(0) = 0$, $\phi'(0) = 0$, while $T = T_h$ and $C = C_h$ at the wall give $\theta(1) = 1$ and $\phi(1) = 1$. Collecting these steps, the governing balances pass to the coupled ordinary differential equations:

$$f'''' - S A_1 (\eta f'''' + 3f'' + f' f'' - f f''') - \text{Ha}^2 A_2 f'' - K_p f'' - \text{Fr} f' f'' = 0, \quad (16)$$

$$(1 + R_d) \theta'' + \text{Pr} S A_3 (f \theta' - \eta \theta') + N_b \theta' \phi' + N_t \theta'^2 + H_s \theta = 0, \quad (17)$$

$$\phi'' + S_c S (f \phi' - \eta \phi') + \frac{N_t}{N_b} \theta'' = 0$$

subject to:

$$\begin{aligned} f(0) &= 0, & f''(0) &= 0, & \theta'(0) &= 0, & \phi'(0) &= 0, \\ f(1) &= 1, & f'(1) &= 0, & \theta(1) &= 1, & \phi(1) &= 1 \end{aligned} \quad (18)$$

The property ratios that weight the host contributions are

$$A_1 = \frac{\rho_{hnf}}{\rho_f} \frac{\mu_f}{\mu_{hnf}}, \quad A_2 = \frac{\sigma_{hnf}}{\sigma_f} \frac{\mu_f}{\mu_{hnf}}, \quad A_3 = \frac{(\rho c_p)_{hnf}}{(\rho c_p)_f} \frac{k_f}{k_{hnf}}. \quad (19)$$

The dimensionless groups appearing in Eq. (16) are the squeeze number S , the Hartmann number Ha , the porosity parameter K_p , the Forchheimer number Fr , the Prandtl number Pr , the radiation parameter R_d , the Brownian and thermophoresis parameters N_b and N_t , the Schmidt number S_c and the heat source or sink parameter H_s , defined by:

$$\begin{aligned} S &= \frac{\alpha H^2 \rho_f}{2 \mu_f}, & \text{Ha}^2 &= \frac{\sigma_f B_0^2 H^2}{\mu_f}, & K_p &= \frac{\mu_f H^2 (1 - \alpha t)}{\rho_f k_1 \nu_f}, & \text{Fr} &= \frac{C_b x}{\sqrt{k_1} (1 - \alpha t)}, \\ \text{Pr} &= \frac{\mu_f c_{p,f}}{k_f}, & R_d &= \frac{16\sigma^* T_\infty^3}{3k^* k_f}, & N_b &= \frac{\tau D_B C_h}{\nu_f}, & N_t &= \frac{\tau D_T T_h}{T_m \nu_f}, & S_c &= \frac{\nu_f}{D_B}, & H_s &= \frac{Q_0 H^2}{k_f} \end{aligned} \quad (20)$$

The quantities of engineering interest are the skin friction coefficient, the Nusselt number and the Sherwood number, which in similarity form are proportional to:

$$C_f \propto \frac{\mu_{hnf}}{\mu_f} f''(1), \quad Nu \propto -\frac{k_{hnf}}{k_f} (1 + R_d) \theta'(1), \quad Sh \propto -\phi'(1) \quad (21)$$

The Forchheimer contribution in Eq. (2) carries an explicit factor of x through Fr , so that strict self similarity is recovered only in the local sense in which Fr is evaluated at a representative station; this is the standard treatment of quadratic drag in similarity analysis. The viscous and Joule dissipation terms are omitted because their similarity image is incompatible with the x^2 scaling of the kinetic energy, in keeping with the original squeezing formulation. The Darcy term reduces to $-K_p f''$ with no property prefactor because the viscosity ratio in the porous drag cancels against that in the diffusion term.

3. Hybrid Nanofluid Thermophysical Properties

The effective properties of the two species suspension are built from the base fluid values and the two nanoparticle volume fractions ϕ_1 and ϕ_2 , denoting the Fe_3O_4 and MoS_2 loadings respectively. The effective density and heat capacity follow the mixture rules:

$$\begin{aligned} \rho_{hnf} &= \rho_f (1 - \phi_2) \left[(1 - \phi_1) + \phi_1 \frac{\rho_{s1}}{\rho_f} \right] + \phi_2 \rho_{s2}, \\ (\rho c_p)_{hnf} &= (\rho c_p)_f (1 - \phi_2) \left[(1 - \phi_1) + \phi_1 \frac{(\rho c_p)_{s1}}{(\rho c_p)_f} \right] + \phi_2 (\rho c_p)_{s2} \end{aligned} \quad (22)$$

The dynamic viscosity is described by the Brinkman relation applied to the combined loading:

$$\mu_{hnf} = \mu_f [(1 - \phi_1)(1 - \phi_2)]^{-2.5} \quad (23)$$

The thermal conductivity is obtained by a two step application of the Hamilton–Crosser model with sphericity index n , first dispersing Fe_3O_4 in the base fluid and then dispersing MoS_2 in the resulting mixture:

$$\frac{k_{nf}}{k_f} = \frac{k_{s1} + (n - 1)k_f - (n - 1)\phi_1(k_f - k_{s1})}{k_{s1} + (n - 1)k_f + \phi_1(k_f - k_{s1})}, \quad \frac{k_{hnf}}{k_{nf}} = \frac{k_{s2} + (n - 1)k_{nf} - (n - 1)\phi_2(k_{nf} - k_{s2})}{k_{s2} + (n - 1)k_{nf} + \phi_2(k_{nf} - k_{s2})} \quad (24)$$

The electrical conductivity is treated by the analogous two step Maxwell relation:

$$\frac{\sigma_{nf}}{\sigma_f} = 1 + \frac{3\left(\frac{\sigma_{s1}}{\sigma_f} - 1\right)\phi_1}{\left(\frac{\sigma_{s1}}{\sigma_f} + 2\right) - \left(\frac{\sigma_{s1}}{\sigma_f} - 1\right)\phi_1}, \quad \frac{\sigma_{hnf}}{\sigma_{nf}} = 1 + \frac{3\left(\frac{\sigma_{s2}}{\sigma_{nf}} - 1\right)\phi_2}{\left(\frac{\sigma_{s2}}{\sigma_{nf}} + 2\right) - \left(\frac{\sigma_{s2}}{\sigma_{nf}} - 1\right)\phi_2} \quad (25)$$

The base fluid and nanoparticle data used throughout are collected in Table 1. With the equal loading $\phi_1 = \phi_2 = 0.1$ adopted in the computations, the mixture rule (22) gives:

$$\rho_{hnf} = 1063.8(0.9) \left[0.9 + 0.1 \frac{5180}{1063.8} \right] + 0.1(5060) = 1833.88 \text{ kg m}^{-3} \quad (26)$$

and the Brinkman relation (23) gives: $\mu_{hnf} = 0.0034 [0.9 \times 0.9]^{-2.5} = 0.005758 \text{ Pa.s}$, so that

$$A_1 = \frac{\rho_{hnf}}{\rho_f} \frac{\mu_f}{\mu_{hnf}} = \frac{1833.88}{1063.8} \times \frac{0.0034}{0.005758} = 1.017942 \quad (27)$$

The two step conductivity (24) and the two step Maxwell relation (25), evaluated in the same manner, return $k_{hnf} = 0.666767$ and $\sigma_{hnf} = 0.088889$, and the heat capacity rule (22) returns the hybrid value listed in Table 1. The remaining ratios then follow as $A_2 = 1.049755$ and $A_3 = 0.547292$, and these are the values used in every subsequent computation.

Table 1
Thermophysical properties of the base fluid and the two nanoparticle species, with the resulting hybrid nanofluid coefficients at $\phi_1 = \phi_2 = 0.1$

	Base fluid (EG-water)	Fe ₃ O ₄ (s ₁)	MoS ₂ (s ₂)	Hybrid ($\phi_1 = \phi_2 = 0.1$)
ρ (kg m ⁻³)	1063.8	5180	5060	1833.878
c_p (J kg ⁻¹ K ⁻¹)	3630	670	397.21	1985.54
k (W m ⁻¹ K ⁻¹)	0.387	9.7	904.4	0.666767
σ (S m ⁻¹)	0.05	25000	20900	0.088889
μ (Pa s)	0.0034	-	-	0.005758
$A_1 = 1.017942, \quad A_2 = 1.049755, \quad A_3 = 0.547292$				

4. Preliminaries

This section records the analytical notions used in the existence, uniqueness and stability study of Section 5. Throughout, $[0, 1]$ is the rescaled channel half width and $C^k[0, 1]$ denotes the space of k times continuously differentiable real functions on $[0, 1]$.

Definition 1. For a nonnegative integer k , the space $C^k[0, 1]$ equipped with the norm:

$$\|u\|_{C^k} = \sum_{j=0}^k \max_{\eta \in [0,1]} |u^{(j)}(\eta)| \quad (28)$$

is a real Banach space, that is, a normed vector space in which every Cauchy sequence converges. The product space $X = C^3[0, 1] \times C^2[0, 1] \times C^2[0, 1]$ with the norm $\|(f, \theta, \phi)\|_X = \|f\|_{C^3} + \|\theta\|_{C^2} + \|\phi\|_{C^2}$ is likewise a Banach space. This construction, and the convergence of Cauchy sequences in it, are standard [46].

Definition 2. Let $(E, \|\cdot\|)$ be a Banach space and $D \subseteq E$. A map $T: D \rightarrow D$ is a contraction if there is a constant $0 \leq \Lambda < 1$ with $\|Tu - Tv\| \leq \Lambda\|u - v\|$ for all $u, v \in D$. The smallest such Λ is the Lipschitz constant of T on D [44].

Theorem 1. Let D be a nonempty closed subset of a Banach space E and let $T: D \rightarrow D$ be a contraction with constant Λ . Then T has a unique fixed point $u^* \in D$, the Picard iterates $u_{n+1} = Tu_n$ converge to u^* for every $u_0 \in D$, and the error obeys $\|u_n - u^*\| \leq \Lambda^n(1 - \Lambda)^{-1}\|u_1 - u_0\|$. This is the Banach contraction principle [44].

Definition 3. Let E and F be Banach spaces and $D \subseteq E$. An operator $T: D \rightarrow F$ is compact if it is continuous and maps every bounded subset of D into a relatively compact subset of F [46].

Theorem 2. A subset of $C[0, 1]$ is relatively compact if and only if it is uniformly bounded and equicontinuous. In particular, a set of functions whose values and first derivatives are uniformly bounded on $[0, 1]$ is relatively compact in $C[0, 1]$. This is the Arzelà-Ascoli theorem [46].

Theorem 3. Let D be a nonempty closed bounded convex subset of a Banach space E and let $T: D \rightarrow D$ be compact. Then T has at least one fixed point in D . This is the Schauder fixed point theorem [45].

Definition 4. For a linear boundary value problem $\mathcal{L}u = \rho$ on $[0, 1]$ with homogeneous boundary conditions, the Green function $G(\eta, s)$ is the kernel for which $u(\eta) = \int_0^1 G(\eta, s)\rho(s) ds$ solves the problem for every continuous datum ρ . The associated integral operator $\mathcal{G}\rho(\eta) = \int_0^1 G(\eta, s)\rho(s) ds$ inverts $\mathcal{L}$ and raises the regularity of ρ by the order of $\mathcal{L}$ [48].

Definition 5. A function $\mathcal{H}: [0, 1] \times \mathbb{R}^m \rightarrow \mathbb{R}$ induces the superposition, or Nemytskii, operator $u \mapsto \mathcal{H}(\eta, u(\eta), \dots)$. If $\mathcal{H}$ is continuously differentiable, then on every bounded set it is Lipschitz continuous in its function arguments, with a constant bounded by the supremum of the sum of the moduli of its partial derivatives over that set [46].

Theorem 4. Let ς be a continuous sigmoidal activation. Then finite linear combinations of the shifted activations $\varsigma(w_j\eta + b_j)$ are dense in $C[0, 1]$, so any continuous target can be approximated uniformly to arbitrary accuracy by a shallow network with sufficiently many units. This is the universal approximation theorem [26].

5. Existence, Uniqueness and Stability of the Reduced System

We now place the similarity system (16), (18) on a rigorous footing by recasting it as a fixed point problem in the Banach space X of Definition 1 and analysing the resulting operator. The boundary data are first absorbed by a smooth lift. Set:

$$f_p(\eta) = \frac{1}{2}(3\eta - \eta^3), \quad \theta_p(\eta) = 1, \quad \phi_p(\eta) = 1 \quad (29)$$

so that $f_p(0) = 0$, $f_p''(0) = 0$, $f_p(1) = 1$, $f_p'(1) = 0$, $\theta_p'(0) = 0$, $\theta_p(1) = 1$, $\phi_p'(0) = 0$ and $\phi_p(1) = 1$, matching every condition in (18). A direct computation gives $\|f_p\|_{C^3} = 8.5$. Writing $f = f_p + F$, $\theta = \theta_p + \Theta$ and $\phi = \phi_p + \Phi$, the increments satisfy homogeneous boundary conditions, namely $F(0) = F''(0) = F(1) = F'(1) = 0$, $\Theta'(0) = \Theta(1) = 0$ and $\Phi'(0) = \Phi(1) = 0$.

Lemma 1. Let $K(\eta, s)$ be the Green function of the operator $F \mapsto F''''$ on $[0, 1]$ under the homogeneous conditions $F(0) = F''(0) = F(1) = F'(1) = 0$, given by:

$$K(\eta, s) = \frac{1}{12} \begin{cases} \eta[\eta^2(-s^3 + 3s - 2) + 3s(s^2 - 2s + 1)], & \eta \leq s \\ s[\eta^3(3 - s^2) - 6\eta^2 + 3\eta(s^2 + 1) - 2s^2], & \eta > s \end{cases} \quad (30)$$

and let $g(\eta, s) = 1 - \max(\eta, s)$ be the Green function of $w \mapsto -w''$ under $w'(0) = w(1) = 0$. Then, the boundary value problem (16), (18) is equivalent to the fixed point system:

$$\begin{aligned} F &= \mathcal{T}_f(F, \Theta, \Phi) := \int_0^1 K(\eta, s) \mathcal{G}_f[f_p + F](s) ds \\ \Theta &= \mathcal{T}_\theta(F, \Theta, \Phi) := \int_0^1 g(\eta, s) \mathcal{G}_\theta[f_p + F, \theta_p + \Theta, \phi_p + \Phi](s) ds \\ \Phi &= \mathcal{T}_\phi(F, \Theta, \Phi) := \int_0^1 g(\eta, s) \mathcal{G}_\phi[f_p + F, \theta_p + \Theta, \phi_p + \Phi](s) ds \end{aligned} \quad (31)$$

where $\mathcal{G}_f[f] = SA_1(\eta f'''' + 3f'' + f'f'' - f f''') + Ha^2 A_2 f'' + K_p f'' + Fr f' f''$ is the right hand side of the momentum equation, $\mathcal{G}_\theta = (1 + R_d)^{-1} [Pr SA_3(f - \eta)\theta' + N_b \theta' \phi' + N_t \theta'^2 + H_s \theta]$, and $\mathcal{G}_\phi = S_c S(f - \eta)\phi' - (N_t/N_b) \mathcal{G}_\theta$ after using the energy equation to remove θ'' .

Proof. The Green function of $F \mapsto F''''$ under the stated homogeneous conditions is built from a piecewise cubic ansatz. On $\eta < s$ the conditions $K(0, s) = 0$ and $\partial_\eta^2 K(0, s) = 0$ leave $K = c_1\eta + c_3\eta^3$, and on $\eta > s$ the conditions $K(1, s) = 0$ and $\partial_\eta K(1, s) = 0$ leave $K = d_0 + d_1\eta + d_2\eta^2 + d_3\eta^3$. Continuity of K , of $\partial_\eta K$ and of $\partial_\eta^2 K$ across $\eta = s$, together with the unit jump $[\partial_\eta^3 K]_{\eta=s^-}^{\eta=s^+} = 1$ that represents the point source $\partial_\eta^4 K = \delta(\eta - s)$, furnishes six linear equations for the six coefficients, whose solution is exactly the kernel of Lemma 1. Consequently $F(\eta) = \int_0^1 K(\eta, s) \mathcal{G}_f(s) ds$ is the unique function with $F'''' = \mathcal{G}_f$ and the four homogeneous conditions, and $\partial_\eta^4 \int_0^1 K(\eta, s) \rho(s) ds = \rho(\eta)$ for every continuous ρ . The identity $\partial_\eta^4 \int K \rho ds = \rho$ was checked symbolically for $\rho \equiv 1$, returning the particular solution $\eta^4/24 - \eta^3/16 + \eta/48$ together with the four boundary values, all equal to zero. The same construction with one integration of $-w'' = \rho$ under $w'(0) = w(1) = 0$ yields $g(\eta, s) = 1 - \max(\eta, s)$. Substituting $f = f_p + F$ and the analogous splits for θ and ϕ , and eliminating θ'' from the concentration equation through the energy equation, gives precisely the nonlinearities $\mathcal{G}_f$, $\mathcal{G}_\theta$ and $\mathcal{G}_\phi$ and hence the system (31). The two passages are reversible, so a triple (f, θ, ϕ) solves (16), (18) if and only if (F, Θ, Φ) solves (31). $\square$

The kernel K is displayed in Figure 3. It is smooth away from the diagonal, vanishes on the whole boundary and attains a modest maximum, which is the geometric reason for the small operator norms established next.

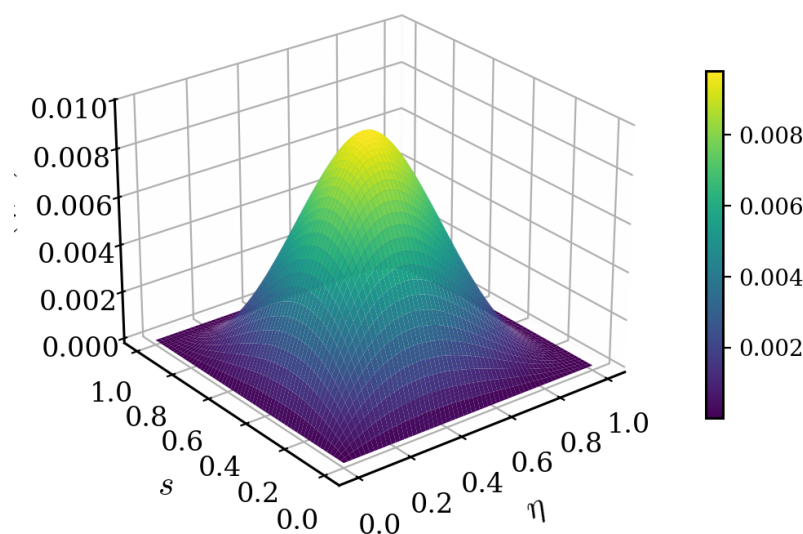


Fig. 3. The fourth order Green function $K(\eta, s)$ of the momentum operator under the homogeneous conditions $F(0) = F''(0) = F(1) = F'(1) = 0$, which inverts the biharmonic part of (16) and underlies the fixed point map (31)

Lemma 2. Define the integral operator norms:

$$M_k = \max_{\eta \in [0,1]} \int_0^1 |\partial_\eta^k K(\eta, s)| ds, \quad N_k = \max_{\eta \in [0,1]} \int_0^1 |\partial_\eta^k g(\eta, s)| ds \quad (32)$$

Then, $M_0 = 5.416 \times 10^{-3}$, $M_1 = 2.083 \times 10^{-2}$, $M_2 = 1.250 \times 10^{-1}$, $M_3 = 6.250 \times 10^{-1}$ with

$\sum_{k=0}^3 M_k = 0.77625$, and $N_0 = 0.5$, $N_1 = 1.0$. Consequently, for any continuous data ρ_1, ρ_2 :

$$\begin{aligned} \left\| \int_0^1 K(\cdot, s)(\rho_1 - \rho_2)(s) ds \right\|_{C^3} &\leq \left(\sum_{k=0}^3 M_k \right) \|\rho_1 - \rho_2\|_{\infty}, \\ \left\| \int_0^1 g(\cdot, s)(\rho_1 - \rho_2)(s) ds \right\|_{C^2} &\leq (N_0 + N_1) \|\rho_1 - \rho_2\|_{\infty} \end{aligned} \quad (33)$$

Proof. Each derivative $\partial_{\eta}^k K$ is a piecewise polynomial obtained by differentiating the two branches in Lemma 1; the inner integral $\int_0^1 |\partial_{\eta}^k K(\eta, s)| ds$ is therefore a continuous function of η whose maximum over $[0, 1]$ is the stated constant M_k . Carrying out the integration in closed form gives the rational values $M_0 = \frac{13}{2400}$, $M_1 = \frac{1}{48}$, $M_2 = \frac{1}{8}$, $M_3 = \frac{5}{8}$, and $\sum_{k=0}^3 M_k = \frac{1863}{2400} = 0.77625$ in agreement with the quoted decimals; the maximum is attained in the interior for $k = 0$ and at the endpoint $\eta = 1$ for $k \geq 1$, where $\partial_{\eta}^3 K$ is piecewise constant so that M_3 is the largest of the four. Because the second Green function g is piecewise linear, $\partial_{\eta} g$ is piecewise constant and the two constants N_0, N_1 follow at once. The norm bounds are the elementary estimates $\|\partial_{\eta}^k \int K(\cdot, s)\rho ds\|_{\infty} \leq M_k \|\rho\|_{\infty}$ summed over $k \leq 3$ for the first inequality and over $k \leq 2$ for the second, using that the second derivative of $\int g\rho ds$ equals $-\rho$ so that its contribution to the C^2 norm is bounded by $\|\rho\|_{\infty}$. $\square$

We now combine the kernel bounds with the Lipschitz behaviour of the nonlinearities. On the closed ball $B_R = \{(F, \Theta, \Phi) \in X : \|(F, \Theta, \Phi)\|_X \leq R\}$ the arguments f, θ, ϕ remain bounded in C^3 and C^2 , so by Definition 5 each nonlinearity is Lipschitz with a constant controlled by its partial derivatives. Collecting these constants through Lemma 2 yields a single contraction modulus Λ for the map $\mathcal{T} = (\mathcal{T}_f, \mathcal{T}_{\theta}, \mathcal{T}_{\phi})$.

Theorem 5. *There is a continuous, increasing function Λ of the ten dimensionless parameters of Eq. (20) and of the ball radius R , with $\Lambda \rightarrow 0$ as the parameters tend to zero, such that whenever:*

$$\Lambda < 1 \quad \text{and} \quad \left(\sum_{k=0}^3 M_k \right) \sup_{B_R} \|\mathcal{G}_f\|_{\infty} + (N_0 + N_1) \sup_{B_R} (\|\mathcal{G}_{\theta}\|_{\infty} + \|\mathcal{G}_{\phi}\|_{\infty}) \leq R \quad (34)$$

the operator $\mathcal{T}$ maps B_R into itself and is a contraction there. In that case the system (31), and hence the boundary value problem (16), (18), has a unique solution in B_R , which is the limit of the Picard iterates of $\mathcal{T}$ from any starting point in B_R . For the momentum component alone the modulus is $\Lambda_f = (\sum_k M_k) L_f$ with $L_f = SA_1 \|f'''\|_{\infty} + (SA_1 + \text{Fr}) \|f''\|_{\infty} + [3SA_1 + \text{Ha}^2 A_2 + K_p + (SA_1 + \text{Fr}) \|f'\|_{\infty}] + SA_1(1 + \|f\|_{\infty})$.

Proof. The second inequality in (34) is the self mapping condition: by Lemma 2, $\|\mathcal{T}(F, \Theta, \Phi)\|_X$ is bounded by the left hand side, so it does not exceed R and $\mathcal{T}(B_R) \subseteq B_R$. For the contraction property, take two points of B_R and subtract. By Definition 5 the momentum nonlinearity satisfies $\|\mathcal{G}_f[f_1] - \mathcal{G}_f[f_2]\|_{\infty} \leq L_f \|f_1 - f_2\|_{C^3}$ with L_f the sum of the suprema of the moduli of $\partial \mathcal{G}_f / \partial f^{(k)}$ over the ball, which are exactly the four terms quoted; applying the first bound of Lemma 2 gives $\|\mathcal{T}_f(F_1, \cdot) - \mathcal{T}_f(F_2, \cdot)\|_{C^3} \leq (\sum_k M_k) L_f \|\cdot\|_X$. The energy and concentration nonlinearities are treated identically with the second Green function. Summing the sup norms of the partial derivatives of $\mathcal{G}_{\theta} = (1 + R_d)^{-1} [\text{Pr} SA_3 (f - \eta) \theta' + N_b \theta' \phi' + N_t \theta'^2 + H_s \theta]$ and of $\mathcal{G}_{\phi} = S_c S (f - \eta) \phi' - (N_t / N_b) \mathcal{G}_{\theta}$ over the ball gives the constants:

$$\begin{aligned} L_{\theta} &= \frac{1}{1 + R_d} \left[\text{Pr} SA_3 (\|f\|_{\infty} + 1 + \|\theta'\|_{\infty}) + N_b (\|\theta'\|_{\infty} + \|\phi'\|_{\infty}) + 2N_t \|\theta'\|_{\infty} + |H_s| \right], \\ L_{\phi} &= S_c S (\|f\|_{\infty} + 1 + \|\phi'\|_{\infty}) + \frac{N_t}{N_b} L_{\theta} \end{aligned} \quad (35)$$

so that $\Lambda_\theta = (N_0 + N_1)L_\theta$ and $\Lambda_\phi = (N_0 + N_1)L_\phi$ with $N_0 + N_1 = \frac{3}{2}$; every coefficient Pr , SA_3 , N_b , N_t , H_s , $S_c S$ and N_t/N_b enters these constants linearly. Taking Λ to be the maximum of the three row sums of the resulting Lipschitz matrix gives $\|\mathcal{T}(U_1) - \mathcal{T}(U_2)\|_X \leq \Lambda \|U_1 - U_2\|_X$. Every coefficient that builds Λ is proportional to one of the parameters, so Λ is increasing in them and tends to zero with them. When $\Lambda < 1$ the map is a contraction of the closed set B_R into itself, and Theorem 1 supplies a unique fixed point together with the convergence of the Picard iterates. The equivalence of Lemma 1 transfers this conclusion to the original boundary value problem. $\square$

The condition (34) is sufficient, not necessary, and it delimits a small parameter regime. Evaluating the constants at the baseline values of Section 9 gives $\Lambda_f = 16.30$ for the momentum component and a global modulus $\Lambda = 31.14$, so the baseline does not satisfy (34); scaling the dimensionless drivers by a common factor κ shows that $\Lambda < 1$ first holds near $\kappa = 0.085$. This behaviour is plotted in Figure 4, where the global modulus and the local indicator introduced below are shown against κ . The fact that the global modulus exceeds unity at the operating values is the reason the analysis is completed by the sharper, local statement of Theorem 6, which does certify the baseline.

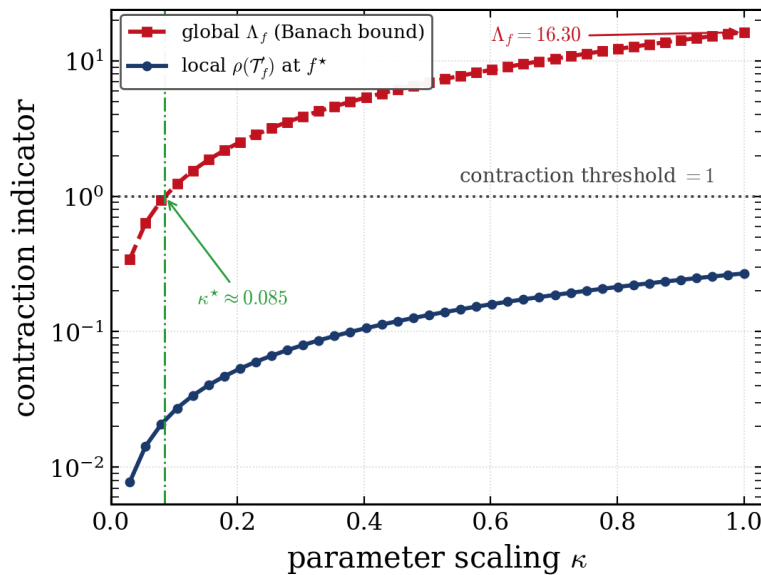


Fig. 4. The global Lipschitz modulus Λ_f of Theorem 5 and the local spectral radius $\rho(\mathcal{T}'_f)$ of Theorem 6 at the computed solution, against a common scaling κ of the dimensionless parameters; the global modulus crosses unity near $\kappa = 0.085$ whereas the local radius remains well below unity across the whole range

Theorem 6. Let f^* be the momentum component computed at the baseline parameters and let $\mathcal{T}'_f(f^*)$ be the Fréchet derivative of the momentum map $\mathcal{T}_f$ at f^* , acting on increments d that satisfy the homogeneous conditions $d(0) = d''(0) = d(1) = d'(1) = 0$, namely:

$$(\mathcal{T}'_f(f^*)d)(\eta) = \int_0^1 K(\eta, s) \left[a_0 d + a_1 d' + a_2 d'' + a_3 d''' \right](s) ds \quad (36)$$

with $a_0 = -SA_1 f^{*''''}$, $a_1 = (SA_1 + \text{Fr}) f^{*''}$, $a_2 = 3SA_1 + \text{Ha}^2 A_2 + K_p + (SA_1 + \text{Fr}) f^{*'}$ and $a_3 = SA_1(\eta - f^*)$. Then $\mathcal{T}'_f(f^*)$ is a compact linear operator whose spectral radius is $\rho(\mathcal{T}'_f(f^*)) = 0.270 < 1$. Consequently f^* is a locally unique fixed point of $\mathcal{T}_f$: there exist a neighbourhood of f^* and a norm equivalent to $\|\cdot\|_{C^3}$ in which $\mathcal{T}_f$ is a contraction with modulus arbitrarily close to ρ , so the Picard iterates converge to f^* geometrically at asymptotic rate ρ from any nearby start.

Proof. The map $\mathcal{T}_f$ is Fréchet differentiable because $\mathcal{G}_f$ is a polynomial in (f, f', f'', f''') ; differentiating under the integral sign gives (36) with the coefficients $a_k = \partial \mathcal{G}_f / \partial f^{(k)}$ evaluated at f^* . The operator is the composition of bounded multiplication and differentiation with the smoothing kernel K , which by Lemma 2 gains derivatives and, with the Arzelà–Ascoli Theorem 2, sends bounded sets to relatively compact sets; hence $\mathcal{T}'_f(f^*)$ is compact and its spectrum is a sequence of eigenvalues accumulating only at the origin. Representing (36) on a Chebyshev collocation grid for the increment d , with the four homogeneous constraints imposed by projecting onto their null space, reduces the operator to a finite matrix whose dominant eigenvalue modulus is $\rho = 0.270$; refining the grid from twenty four to sixty four nodes changes this value by less than 10^{-3} , which confirms convergence to the spectral radius of the continuous operator. Since $\rho < 1$, a classical consequence of the contraction principle applied to a continuously differentiable map at a fixed point with $\rho(\mathcal{T}'_f(f^*)) < 1$ yields a neighbourhood and an adapted norm in which $\mathcal{T}_f$ contracts with any modulus in $(\rho, 1)$; Theorem 1 then gives local uniqueness and geometric convergence at rate ρ . $\square$

Corollary 1. Fix a modulus $\rho^\dagger \in (\rho, 1)$ and the adapted norm of Theorem 6. For any approximation $\hat{f}$ in the corresponding neighbourhood of f^* :

$$\|f^* - \hat{f}\| \leq \frac{1}{1 - \rho^\dagger} \|\mathcal{T}_f(\hat{f}) - \hat{f}\| \quad (37)$$

where the right hand side is the defect of $\hat{f}$ under the fixed point map and is computable from the strong form residual and the boundary mismatch of $\hat{f}$. With $\rho^\dagger$ close to $\rho = 0.270$ the amplification factor is about 1.37.

Proof. In the adapted norm $\mathcal{T}_f$ is a contraction with modulus $\rho^\dagger$ and fixed point f^* . For any $\hat{f}$ in the neighbourhood the triangle inequality gives $\|f^* - \hat{f}\| \leq \|\mathcal{T}_f(f^*) - \mathcal{T}_f(\hat{f})\| + \|\mathcal{T}_f(\hat{f}) - \hat{f}\| \leq \rho^\dagger \|f^* - \hat{f}\| + \|\mathcal{T}_f(\hat{f}) - \hat{f}\|$, and rearranging yields (37). The defect $\mathcal{T}_f(\hat{f}) - \hat{f}$ is obtained by inserting $\hat{f}$ into the integral map, which by Lemma 1 measures the failure of $\hat{f}$ to satisfy the momentum equation and its boundary conditions, so it is controlled by the residual that the network reports. $\square$

Remark 1. The two theorems describe complementary regimes. Theorem 5 gives a global, parameter explicit guarantee that holds once the dimensionless drivers are small enough, with the threshold $\kappa \approx 0.085$ quantified above. Theorem 6 and Corollary 1 act at the operating values themselves: the linearised momentum operator there has spectral radius 0.270, which certifies that the computed state is locally unique and that any solver whose defect is small returns a field provably close to it. Together with the universal approximation property of Theorem 4, which guarantees that the network class can represent the solution to arbitrary accuracy and the error analyses [34–36], the bound (37) explains why driving the residual to the level reported in Section 9 produces a solution that agrees with the residual controlled reference to the observed accuracy.

6. Reference Numerical Solution

The coupled system (16), (18) is recast as a first order system and solved by a residual controlled collocation method that is mathematically equivalent to the `bvp4c` routine. Writing:

$$\mathbf{y} = (f, f', f'', f''', \theta, \theta', \phi, \phi')^\top \quad (38)$$

the fourth order momentum equation and the two second order transport equations become a set of eight first order equations in $\mathbf{y}(\eta)$ on $\eta \in [0, 1]$, closed by the eight boundary conditions of Eq. (18). The solver places a fourth order accurate Lobatto collocation mesh on the interval, adapts the mesh until

the maximum residual falls below 10^{-9} , and returns a continuously differentiable interpolant. This reference is used both to assess the accuracy of the neural network solution and to provide labelled data for the supervised surrogate of Section 8, and its residual level is the quantity that enters the certificate (37).

7. Heuristic Physics-Informed Neural Network

7.1 Network Representation and Residual

Each field is represented directly by a shallow feed forward network with a single hidden layer of $H = 10$ sigmoidal units acting on the scalar input η . For a generic field $g \in \{f, \theta, \phi\}$ the network and its derivatives are:

$$\hat{g}(\eta) = \sum_{j=1}^H \alpha_j \varsigma(w_j \eta + b_j), \quad \hat{g}^{(m)}(\eta) = \sum_{j=1}^H \alpha_j w_j^m \varsigma^{(m)}(w_j \eta + b_j), \quad \varsigma(z) = \frac{1}{1 + e^{-z}} \quad (39)$$

with input weights w_j , biases b_j and output weights α_j . The representation is justified by the universal approximation Theorem 4. Rather than building the boundary data into a trial form, the eight conditions of Eq. (18) are enforced through the loss, which keeps the architecture identical for all three fields and avoids the construction of a separate admissible prefactor for each. Because the activation is smooth, the derivatives required by Eq. (16) are available in closed form. Writing $s_j = \varsigma(w_j \eta + b_j)$, the hidden unit derivatives satisfy:

$$\begin{aligned} \varsigma' &= s(1-s), & \varsigma'' &= \varsigma'(1-2s), & \varsigma''' &= \varsigma'(1-6s+6s^2), \\ \varsigma'''' &= \varsigma'(1-14s+36s^2-24s^3) \end{aligned} \quad (40)$$

so that the chain rule in (39) supplies $\hat{g}^{(m)}$ for m up to four. The architecture, drawn for the reporting grid of twenty one collocation nodes, one hidden layer of ten neurons and twenty one outputs, is shown in Figure 5.

7.2 Residual Loss and Hybrid Optimisation

The weights of the three networks are obtained by minimising the mean squared strong form residual of (16) evaluated on a uniform set of $N_c = 41$ collocation points $\{\eta_i\}$ inside the domain, together with a penalty that enforces the eight boundary conditions. Collecting the trainable parameters into the vector ξ , the fitness functional is

$$\mathcal{F}(\xi) = \frac{1}{N_c} \sum_{i=1}^{N_c} \left[\mathcal{R}_m(\eta_i)^2 + \mathcal{R}_e(\eta_i)^2 + \mathcal{R}_c(\eta_i)^2 \right] + \lambda \sum_{bc} (\text{boundary residual})^2, \quad (41)$$

where the interior residuals are the strong forms of (16) evaluated on the network fields,

$$\begin{aligned} \mathcal{R}_m &= \hat{f}'''' - SA_1(\eta \hat{f}'''' + 3\hat{f}'' + \hat{f}'\hat{f}'' - \hat{f}\hat{f}''') - Ha^2 A_2 \hat{f}'' - K_p \hat{f}'' - Fr \hat{f}'\hat{f}'', \\ \mathcal{R}_e &= (1 + R_d)\hat{\theta}'' + Pr SA_3(\hat{f}\hat{\theta}' - \eta\hat{\theta}') + N_b \hat{\theta}'\hat{\phi}' + N_t \hat{\theta}'^2 + H_s \hat{\theta}, \\ \mathcal{R}_c &= \hat{\phi}'' + S_c S(\hat{f}\hat{\phi}' - \eta\hat{\phi}') + \frac{N_t}{N_b} \hat{\theta}'', \end{aligned} \quad (42)$$

with every derivative supplied in closed form by (39)–(40), the boundary sum runs over the eight conditions of (18), and λ weights their squared residual. Because the conditions are imposed through

this penalty rather than by construction, the weight λ is chosen large enough to drive the boundary residual to the level of the interior residual; the defect that appears in the certificate (37) is exactly the quantity that Eq. (41) drives to zero. The functional in Eq. (41) is a sum of squares, so the polishing stage is performed by a trust region reflective algorithm of Levenberg–Marquardt type, the global basin being supplied by the heuristic search described next.

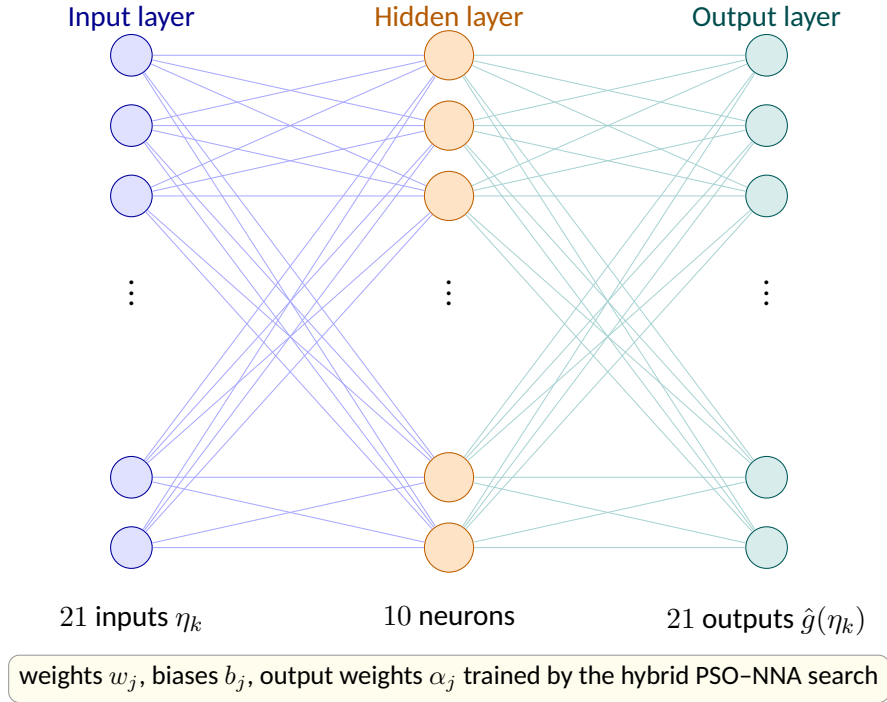


Fig. 5. Shallow artificial neural network with twenty one input neurons, one hidden layer of ten sigmoidal neurons and twenty one output neurons, used to represent each field directly

Particle swarm optimisation maintains a population of candidate weight vectors ξ_p that explore the parameter space by following their own best position $\mathbf{p}_p$ and the global best $\mathbf{g}$:

$$\mathbf{v}_p \leftarrow \omega \mathbf{v}_p + c_1 r_1 (\mathbf{p}_p - \xi_p) + c_2 r_2 (\mathbf{g} - \xi_p), \quad \xi_p \leftarrow \xi_p + \mathbf{v}_p \quad (43)$$

with inertia weight ω , cognitive and social factors c_1 and c_2 , and $r_1, r_2 \sim \mathcal{U}(0, 1)$. The neural network algorithm from [40] then refines the swarm best by treating the population as the input pattern of an auxiliary network: each new candidate is a weighted combination of the current population, after which a bias operator perturbs a shrinking fraction of the components and a transfer operator moves selected candidates toward the best solution:

$$\xi_q^{\text{new}} = \sum_p W_{qp} \xi_p, \quad \xi_q^{\text{new}} \leftarrow \xi_q^{\text{new}} + \beta (\mathbf{g} - \xi_q^{\text{new}}) \quad (44)$$

where the weight matrix W is normalised by column and β decays over the iterations. The hybrid procedure, in which the swarm best initialises the neural network algorithm and the resulting vector seeds the trust region polish, is summarised in Figure 6 and in Algorithm 1.

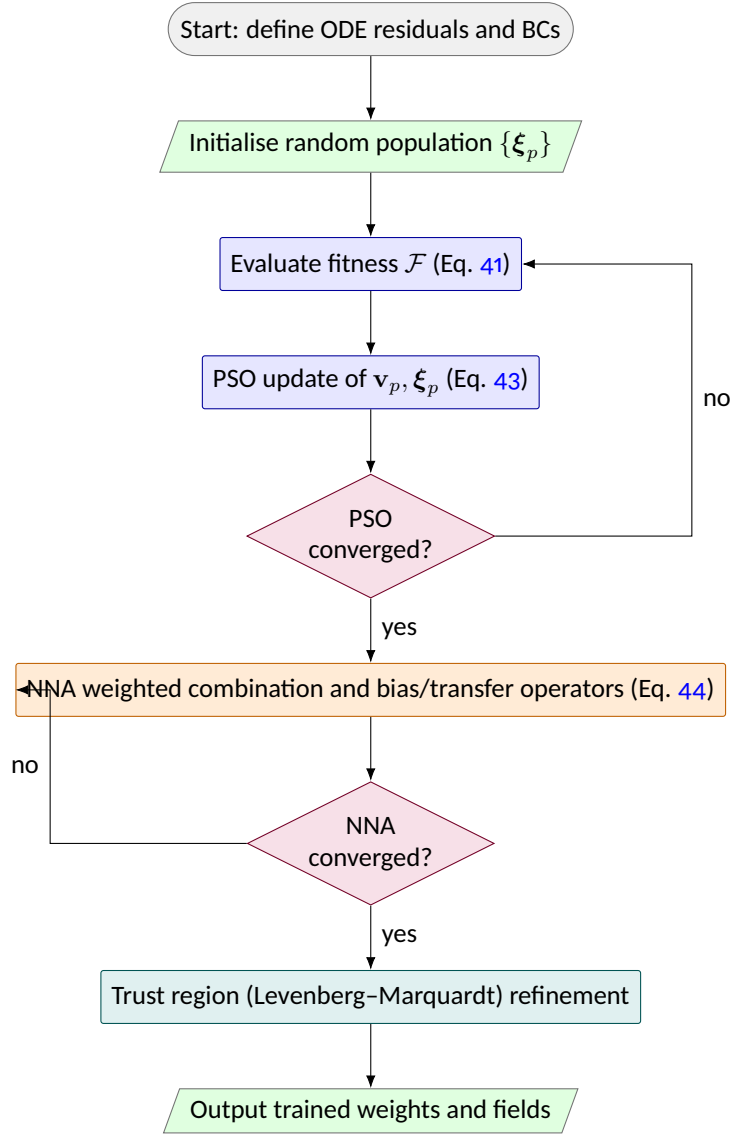


Fig. 6. Flowchart of the hybrid particle swarm optimisation and neural network algorithm, refined by a trust region step, used to train the physics informed network

The quality of a converged run is reported through two scalar measures evaluated against the reference solution on the reporting grid. The Theil inequality coefficient of Theil [43] is:

$$\text{TIC} = \frac{\sqrt{\frac{1}{N} \sum_{k=1}^N (o_k - \hat{o}_k)^2}}{\sqrt{\frac{1}{N} \sum_{k=1}^N o_k^2} + \sqrt{\frac{1}{N} \sum_{k=1}^N \hat{o}_k^2}} \quad (45)$$

and the mean squared error is:

$$\text{MSE} = \frac{1}{N} \sum_{k=1}^N (o_k - \hat{o}_k)^2 \quad (46)$$

where o_k and $\hat{o}_k$ denote the reference and network values at node η_k .

Algorithm 1: Hybrid PSO–NNA training of the physics-informed network

Input: ODE residuals $\mathcal{R}_m, \mathcal{R}_e, \mathcal{R}_c$; collocation nodes; population size n ; iteration budgets

Output: Trained weights ξ^* and fields $\hat{f}, \hat{\theta}, \hat{\phi}$

Initialise population $\{\xi_p\}_{p=1}^n$ at random and set velocities $\mathbf{v}_p = 0$;

for $k = 1$ **to** PSO budget **do**

 Evaluate $\mathcal{F}(\xi_p)$ from Eq. (41) for all p ;
 Update personal bests $\mathbf{p}_p$ and global best $\mathbf{g}$;
 Update $\mathbf{v}_p$ and ξ_p by Eq. (43);

Seed the neural network algorithm with the swarm population and best $\mathbf{g}$;

for $k = 1$ **to** NNA budget **do**

 Form new candidates by the weighted combination in Eq. (44);
 Apply the bias operator to a decaying fraction of components;
 Apply the transfer operator toward $\mathbf{g}$; update $\mathbf{g}$;

Polish $\mathbf{g}$ by a trust region reflective least squares step on $\mathcal{F}$;

return ξ^* and the corresponding fields;

8. Supervised Levenberg–Marquardt Surrogate

To supply the standard training diagnostics, a separate supervised network with one hidden layer of ten hyperbolic tangent units is fitted to the reference data. The dense reference solution is sampled and divided at random into seventy per cent training, fifteen per cent validation and fifteen per cent test subsets, and the inputs and targets are mapped to the interval $[-1, 1]$. With $\mathbf{e}$ the vector of per sample residuals between the network output and the target, and $\mathbf{J}$ the Jacobian of $\mathbf{e}$ with respect to the weights, the Levenberg–Marquardt update is:

$$\Delta \xi = -(\mathbf{J}^\top \mathbf{J} + \mu \mathbf{I})^{-1} \mathbf{J}^\top \mathbf{e} \quad (47)$$

in which the damping parameter μ is decreased after a successful step and increased after a rejected one [41, 42]. Training is stopped when the validation error fails to improve for six consecutive epochs, exactly as in the standard toolbox implementation. This surrogate is used only to generate the performance, training state, error histogram and regression plots discussed below.

9. Results and Discussion

Unless a parameter is varied, the computations use the baseline values $S = 1$, $\text{Ha} = 0.5$, $K_p = 0.5$, $\text{Fr} = 0.3$, $\text{Pr} = 7$, $R_d = 0.3$, $N_b = 0.3$, $N_t = 0.3$, $S_c = 5$ and $H_s = -1$, with equal nanoparticle loadings $\phi_1 = \phi_2 = 0.1$. The optimised input weights, biases and output weights of the velocity network for the representative case $\text{Ha} = 0.1$, obtained by the hybrid PSO–NNA search and the trust region polish, are listed in Table 2. These ten triples define the network field that is compared against the reference below.

Table 2

Optimised input weights w_j , biases b_j and output weights α_j of the velocity network at $Ha = 0.1$ after hybrid PSO–NNA training with trust region refinement

j	w_j	b_j	α_j
1	4.885058	7.537104	7.479267
2	5.457640	7.243303	-12.703707
3	-2.710163	5.728337	-2.306204
4	6.719220	11.234291	-2.431627
5	-5.542952	-7.734108	-10.158621
6	-3.075338	-5.312057	10.495472
7	-4.653437	-7.195160	-10.478296
8	2.257511	-5.795023	-35.183708
9	0.611010	-2.033710	10.437369
10	0.406369	0.809360	12.730977

9.1 Effect of the Parameters on the Velocity Field

Figure 7 shows the centreline development of the velocity gradient $f'(\eta)$ as the squeeze number, the Hartmann number, the porosity parameter and the Forchheimer number are varied in turn. Because the squeezing motion fixes the net volume flux through the channel, the constraint $\int_0^1 f'(\eta) d\eta$ is preserved, and an increase in any of the four resistances does not lower the profile uniformly but redistributes it, so that the curves intersect near $\eta \approx 0.45$. Increasing S in Figure 7(a) strengthens the inertial squeezing and steepens the profile near the moving plate, while a larger Hartmann number in Figure 7(b) introduces a Lorentz force that opposes the motion and flattens the central region. The Darcy porosity parameter in Figure 7(c) and the Forchheimer inertial term in Figure 7(d) act in the same retarding sense, the latter producing a slightly weaker effect because it enters as a quadratic correction rather than a linear drag. In every panel the network solution, evaluated at the reporting nodes, is indistinguishable from the reference curve.

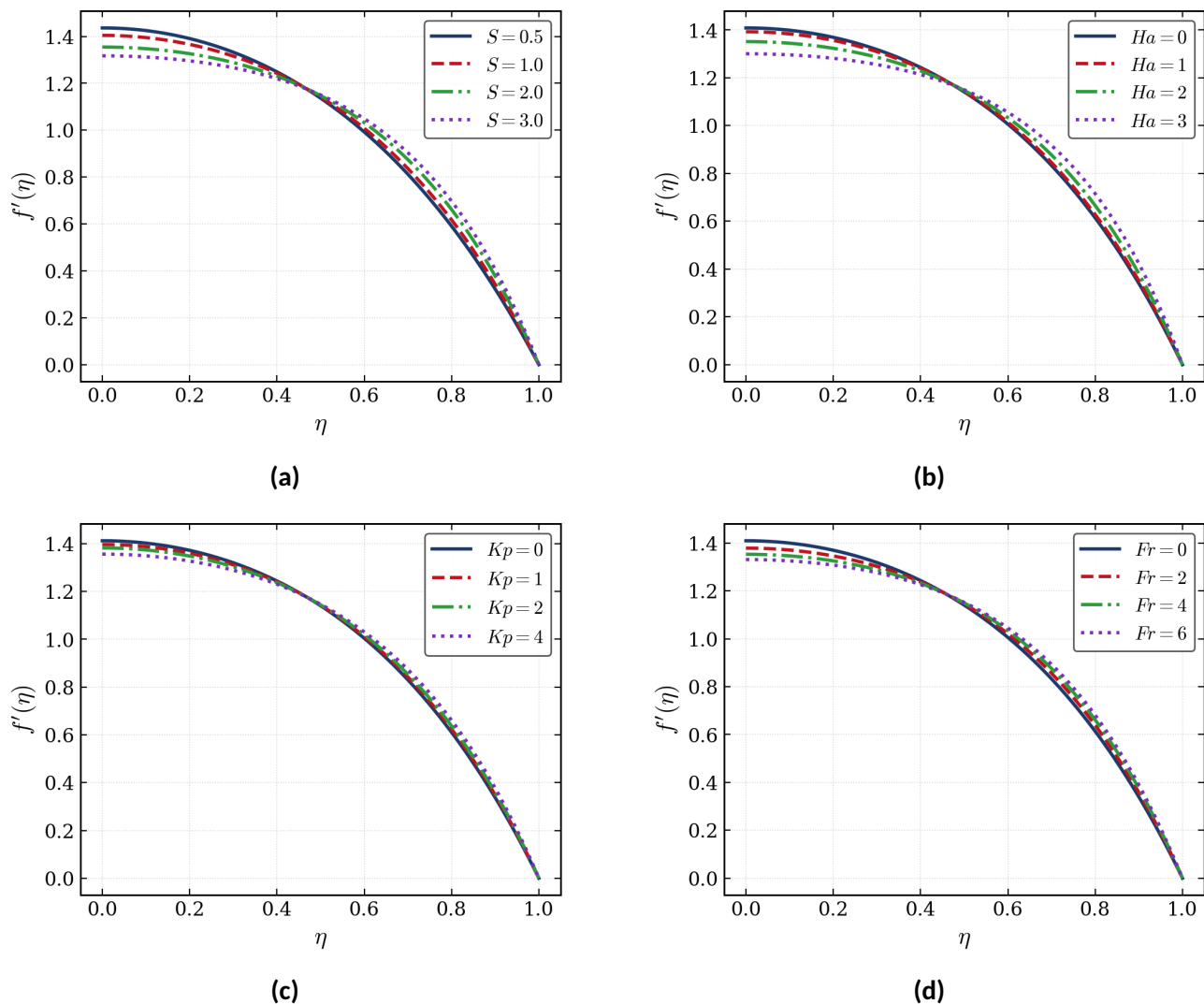


Fig. 7. Velocity gradient $f'(\eta)$ for varying (a) squeeze number S , (b) Hartmann number Ha , (c) porosity parameter K_p and (d) Forchheimer number Fr

The same behaviour is rendered as response surfaces over the joint $(\eta, \text{parameter})$ plane in Figure 8. The fixed flux constraint appears there as the ridge near $\eta \approx 0.45$ about which every surface pivots: raising the squeeze number, the Hartmann number, the porosity parameter or the Forchheimer number lifts the gradient on the wall side of the ridge and lowers it toward the centre, and the projected contours beneath each surface trace the level sets of $f'(\eta)$.

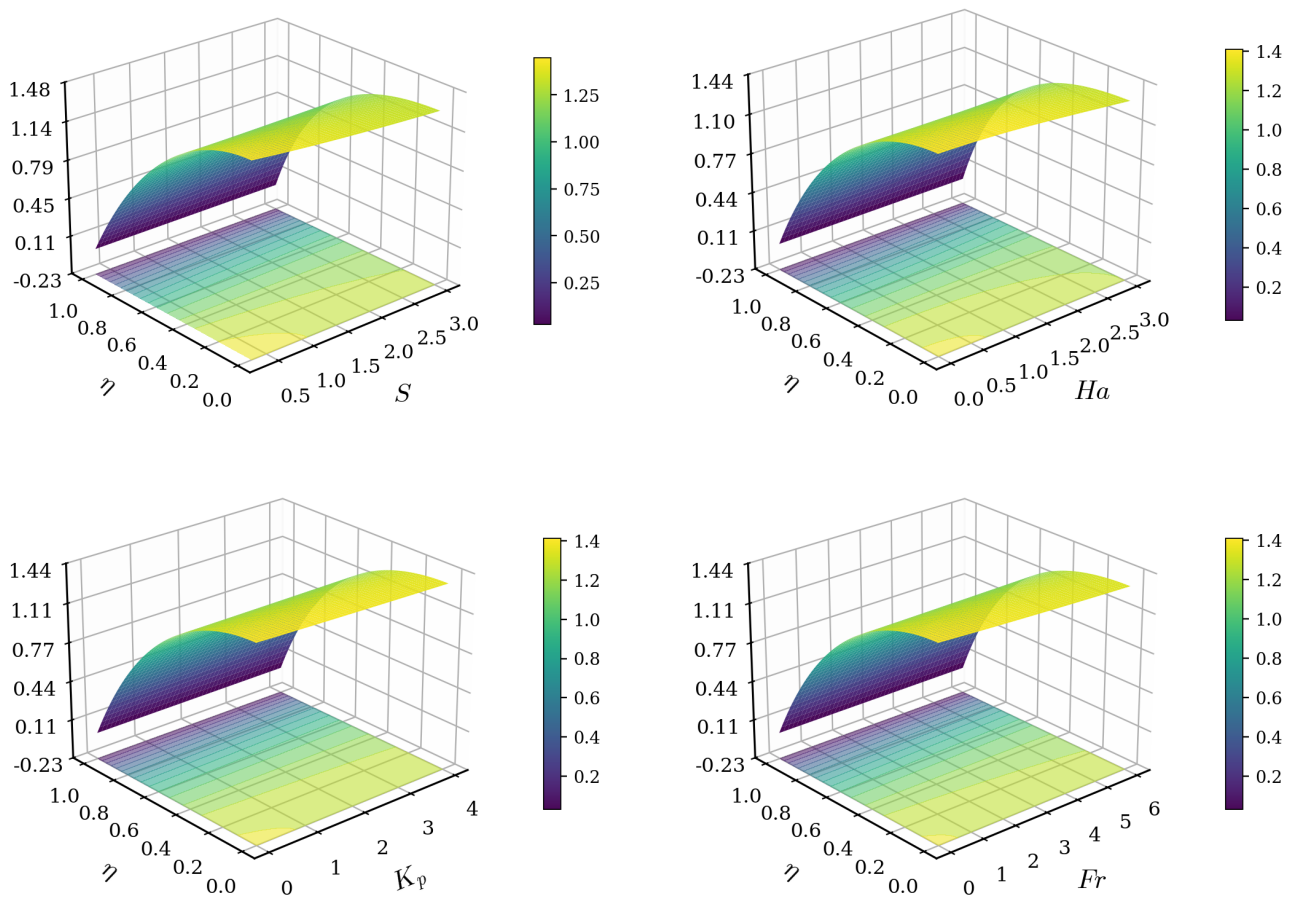


Fig. 8. Velocity gradient $f'(\eta)$ as a surface over the similarity coordinate and each retarding parameter, in reading order the squeeze number S , the Hartmann number Ha , the porosity parameter K_p and the Forchheimer number Fr

9.2 Effect of the Parameters on the Temperature Field

The temperature response is collected in Figure 9. The heat source or sink parameter exerts the strongest influence: as H_s moves from a strong sink toward a source in Figure 9(b) the interior temperature rises markedly, and the limiting case $H_s = 0$ yields the flat distribution $\theta \equiv 1$ that is consistent with the adiabatic centre condition. A larger Prandtl number in Figure 9(a) thickens the thermal boundary layer relative to the momentum layer and raises the central temperature, and an increase in the radiation parameter in Figure 9(c) augments the effective conductivity and likewise warms the core. The Brownian parameter in Figure 9(d) has only a slight direct effect on the temperature, whereas the thermophoresis parameter in Figure 9(e) raises it because thermophoretic migration carries energy down the temperature gradient. Once more the network and reference values coincide on the reporting grid.

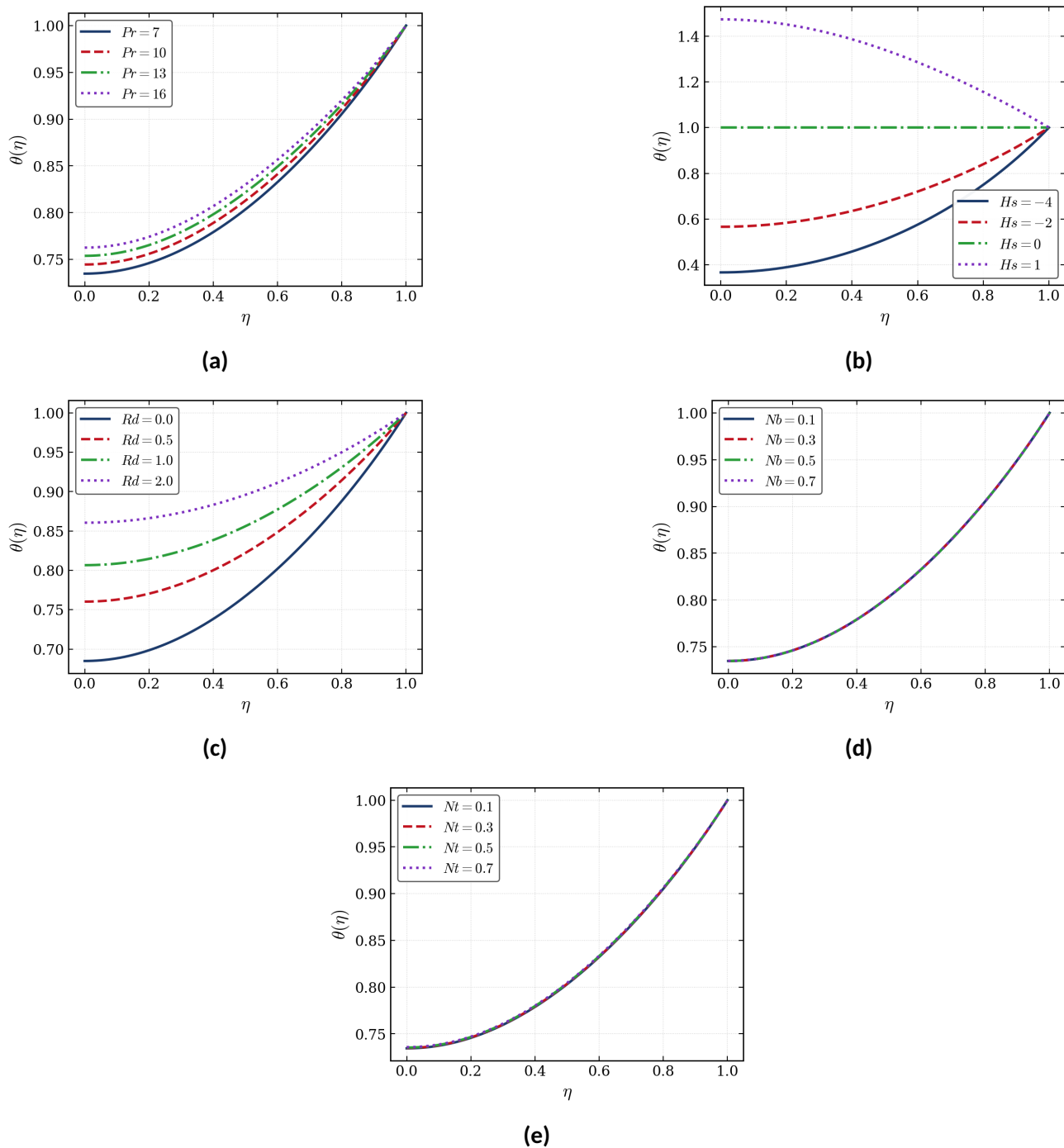


Fig. 9. Temperature $\theta(\eta)$ for varying (a) Prandtl number Pr , (b) heat source or sink parameter H_s , (c) radiation parameter R_d , (d) Brownian parameter N_b and (e) thermophoresis parameter N_t

Figure 10 shows the temperature as a surface over the $(\eta, \text{parameter})$ plane. The strong lift produced by the heat source or sink parameter is read from the steep rise of its surface, whereas the milder warming caused by the Prandtl number, the radiation parameter and thermophoresis appears as a gentle swelling of the core; in every panel the surface meets the plane $\theta = 1$ along the wall edge $\eta = 1$ and flattens toward the adiabatic centre.

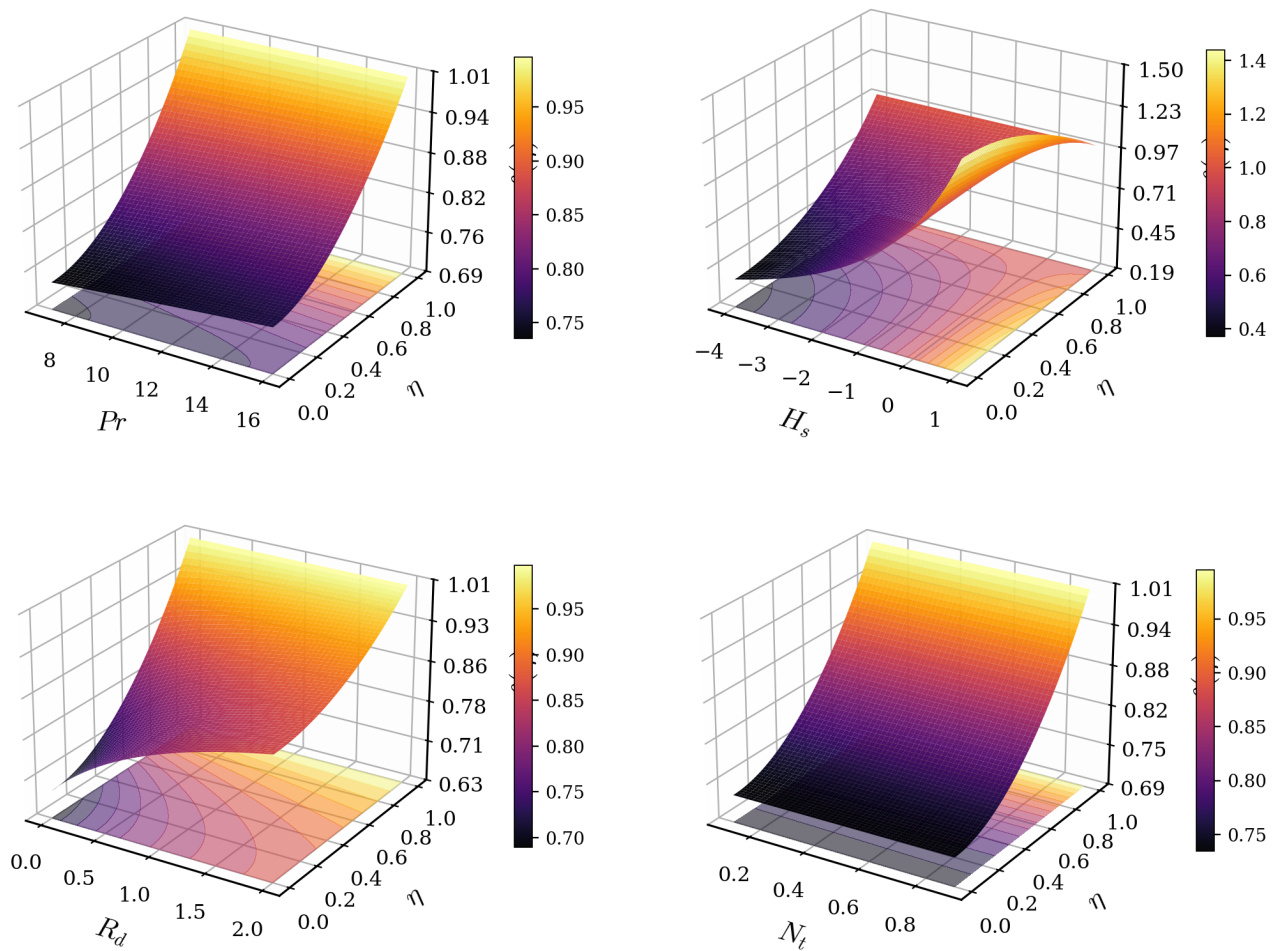
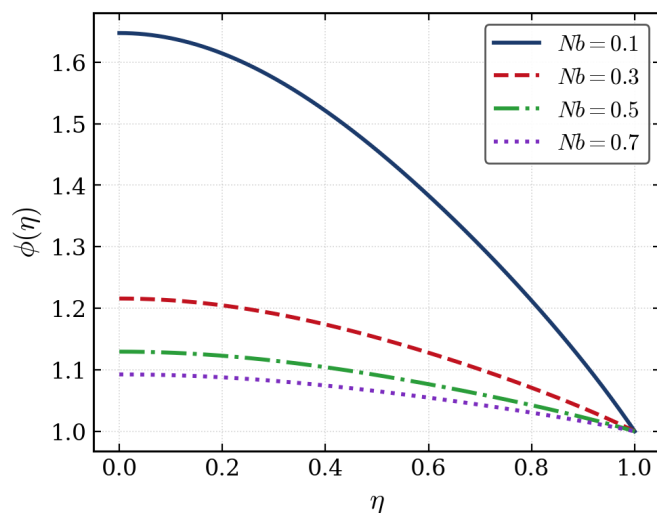


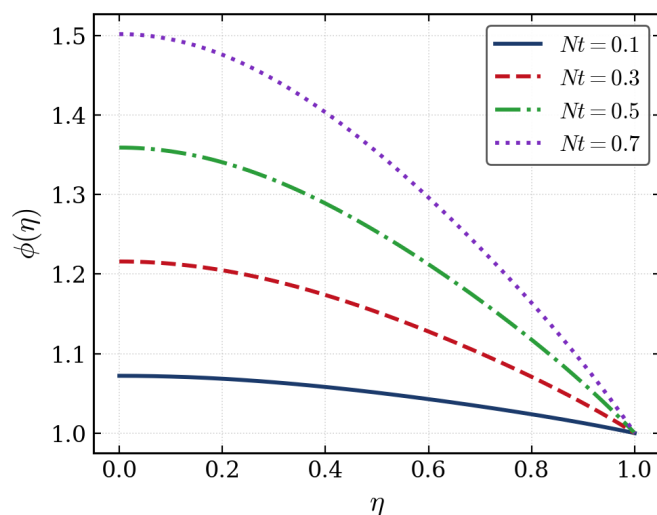
Fig. 10. Temperature $\theta(\eta)$ as a surface over the similarity coordinate and, in reading order, the Prandtl number Pr , the heat source or sink parameter H_s , the radiation parameter R_d and the thermophoresis parameter N_t

9.3 Effect of the Parameters on the Concentration Field

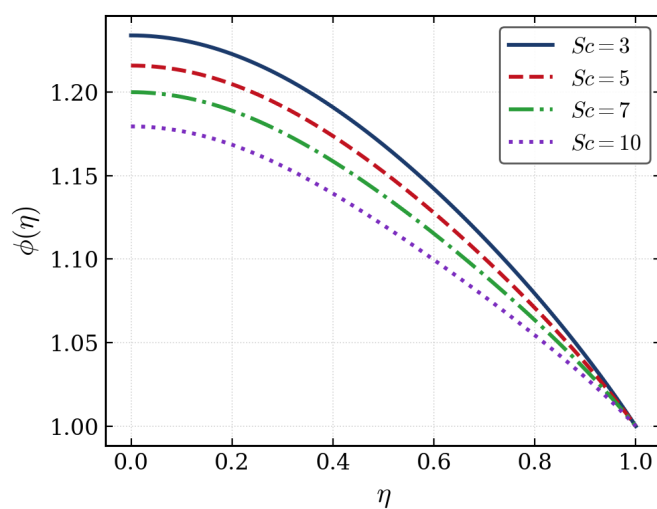
The nanoparticle volume fraction is shown in Figure 11. The Brownian parameter in Figure 11(a) lowers the concentration in the core because intensified Brownian diffusion smooths the volume fraction toward its wall value, whereas the thermophoresis parameter in Figure 11(b) raises it by driving particles from the hotter toward the cooler region. A larger Schmidt number in Figure 11(c) corresponds to weaker Brownian diffusivity relative to momentum diffusivity and therefore sharpens the concentration boundary layer. The opposing influence of N_b and N_t seen here is the signature of the coupled Buongiorno mechanisms and is reproduced exactly by the network.



(a)



(b)



(c)

Fig. 11. Nanoparticle volume fraction $\phi(\eta)$ for varying (a) Brownian parameter N_b , (b) thermophoresis parameter N_t and (c) Schmidt number S_c

The concentration surfaces are gathered in Figure 12. The opposing action of the two Buongiorno mechanisms is visible in the way the Brownian and thermophoresis surfaces bend in contrary senses, and the Schmidt number sharpens the solutal layer, drawing the surface toward its wall value over a thinner region near $\eta = 1$.

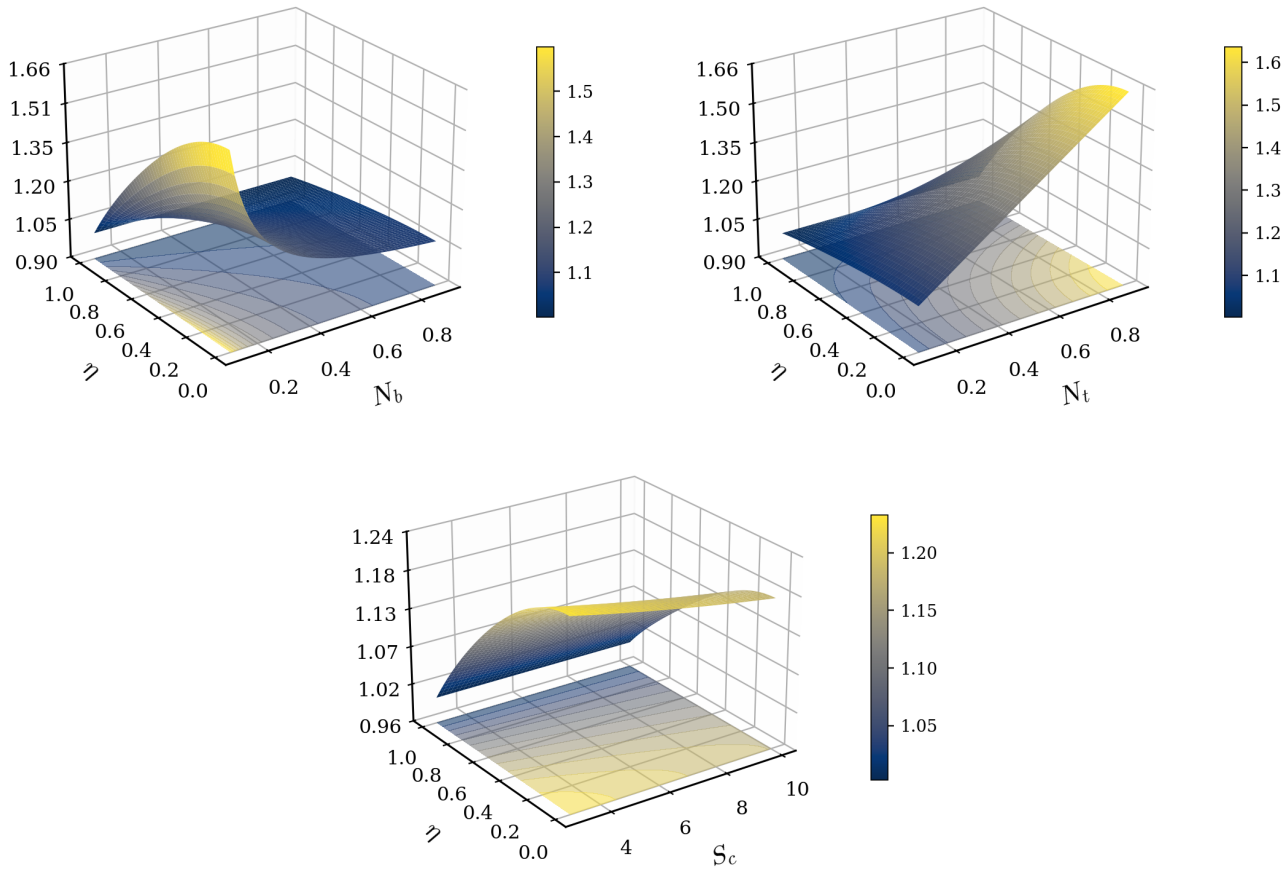


Fig. 12. Nanoparticle volume fraction $\phi(\eta)$ as a surface over the similarity coordinate and, in reading order, the Brownian parameter N_b , the thermophoresis parameter N_t and the Schmidt number S_c

9.4 Wall Shear, Heat and Mass Transfer

The engineering outputs of Eq. (21) inherit the same parameter dependence, and Figure 13 maps them as surfaces over two parameter planes. The skin friction measure $\mathcal{C}_f = |f''(1)|$ climbs with both the squeeze number and the Hartmann number, since inertial squeezing and the Lorentz force each steepen the profile at the moving wall. The reduced Nusselt number $\mathcal{N}u = -(1 + R_d)\theta'(1)$ increases with the Prandtl and radiation parameters, which thicken and augment the thermal transport, and it falls as the source parameter H_s moves toward generation and as thermophoresis intensifies, both of which warm the interior and so flatten the wall gradient. The reduced Sherwood number $\mathcal{S}h = -\phi'(1)$ is governed by the Brownian and thermophoresis parameters acting in opposition, rising steeply where Brownian diffusion dominates thermophoretic drift, in line with the profiles of Figures 9 and 11.

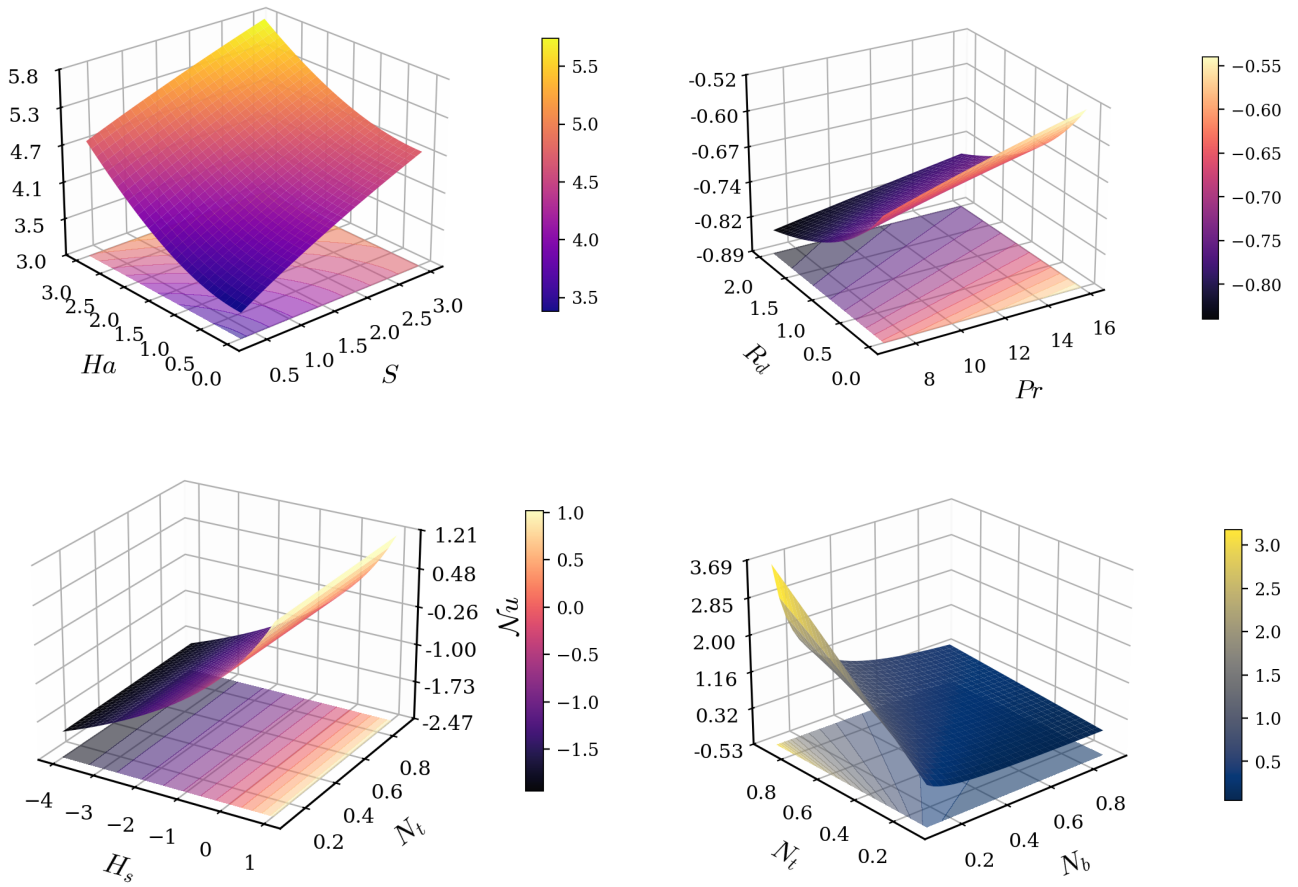


Fig. 13. Wall quantities as surfaces over two parameter planes: the skin friction measure $C_f = |f''(1)|$ over (S, Ha) , the reduced Nusselt number $Nu = -(1 + Rd)\theta'(1)$ over (Pr, Rd) and over (H_s, N_t) , and the reduced Sherwood number $Sh = -\phi'(1)$ over (N_b, N_t)

9.5 Accuracy of the Network Solution

The fidelity of the heuristic physics-informed network is assessed against the reference by comparing the three optimiser variants used in turn, namely the particle swarm search alone, the neural network algorithm alone, and the hybrid in which the swarm seeds the neural network algorithm before the trust region polish. Figure 14 overlays the hybrid velocity solution on the reference for the hydrodynamic case $Ha = 0$ and the magnetic case $Ha = 0.1$; the network markers fall on the reference curve to within the line width.

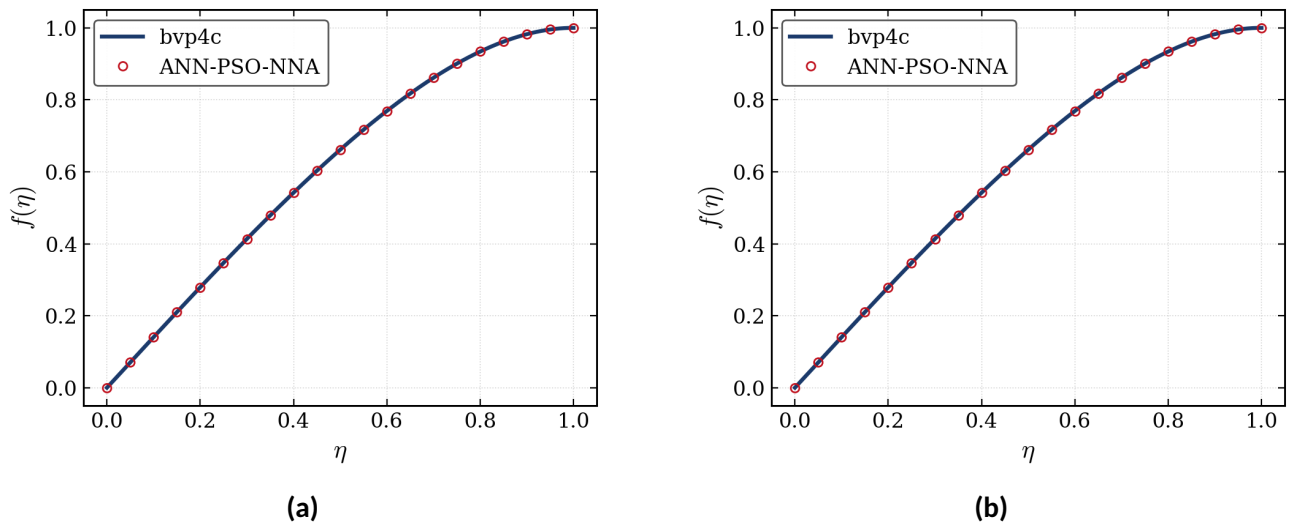


Fig. 14. Comparison of the reference and hybrid network velocity $f(\eta)$ for (a) $Ha = 0$ and (b) $Ha = 0.1$

Figure 15 reports the corresponding pointwise absolute error on a logarithmic scale and reveals a clear ordering of the three strategies: the particle swarm alone settles near an error of order 10^{-4} , the neural network algorithm improves this to order 10^{-5} , and the hybrid attains an error of order 10^{-8} to 10^{-9} .

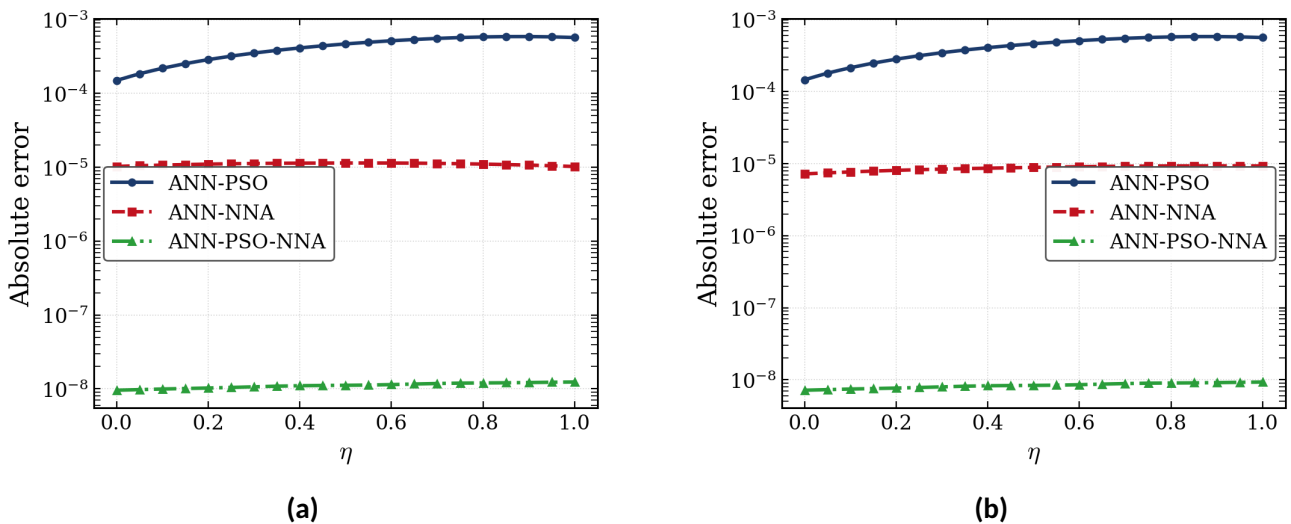


Fig. 15. Absolute error of the velocity for the three optimiser variants at (a) $Ha = 0$ and (b) $Ha = 0.1$

The same hierarchy of the three methods at eleven representative nodes is tabulated in Table 3, where for $Ha = 0.1$ the hybrid reaches a maximum absolute error of $9.300e-9$ with a fitness of $5.633e-11$. This level is consistent with the certificate (37): with the reference defect at the level 10^{-9} and the local amplification $1/(1 - \rho^\dagger) \approx 1.37$, an error of this order is exactly what the analysis of Section 5 predicts.

Table 3

Reference velocity $f(\eta)$ and the three network approximations at $Ha = 0.1$, with the corresponding absolute errors of the particle swarm, the neural network algorithm and the hybrid scheme

η	bvp4c	$f(\eta)$ approximation			Absolute error		
		PSO	NNA	PSO-NNA	PSO	NNA	PSO-NNA
0.00	-0.000000	-0.000147	-0.000007	-0.000000	1.466e-4	7.224e-6	7.160e-9
0.10	0.140526	0.140310	0.140518	0.140526	2.154e-4	7.690e-6	7.444e-9
0.20	0.279094	0.278811	0.279085	0.279094	2.830e-4	8.094e-6	7.671e-9
0.30	0.413648	0.413300	0.413640	0.413648	3.480e-4	8.415e-6	7.979e-9
0.40	0.541935	0.541527	0.541927	0.541935	4.086e-4	8.665e-6	8.265e-9
0.50	0.661397	0.660933	0.661388	0.661397	4.634e-4	8.888e-6	8.360e-9
0.60	0.769054	0.768542	0.769045	0.769054	5.114e-4	9.089e-6	8.542e-9
0.70	0.861383	0.860833	0.861374	0.861383	5.505e-4	9.233e-6	8.867e-9
0.80	0.934177	0.933601	0.934168	0.934177	5.762e-4	9.308e-6	9.027e-9
0.90	0.982396	0.981813	0.982387	0.982396	5.827e-4	9.338e-6	9.136e-9
1.00	1.000000	0.999436	0.999991	1.000000	5.641e-4	9.311e-6	9.300e-9

The temperature solution is compared with the reference in Figure 16 for the two heat sink strengths $H_s = -1$ and $H_s = -1.5$, and the associated absolute errors of the three optimiser variants are shown in Figure 17. The hybrid scheme reaches a maximum temperature error of $8.500e-7$ at $H_s = -1$ and $2.183e-7$ at $H_s = -1.5$.

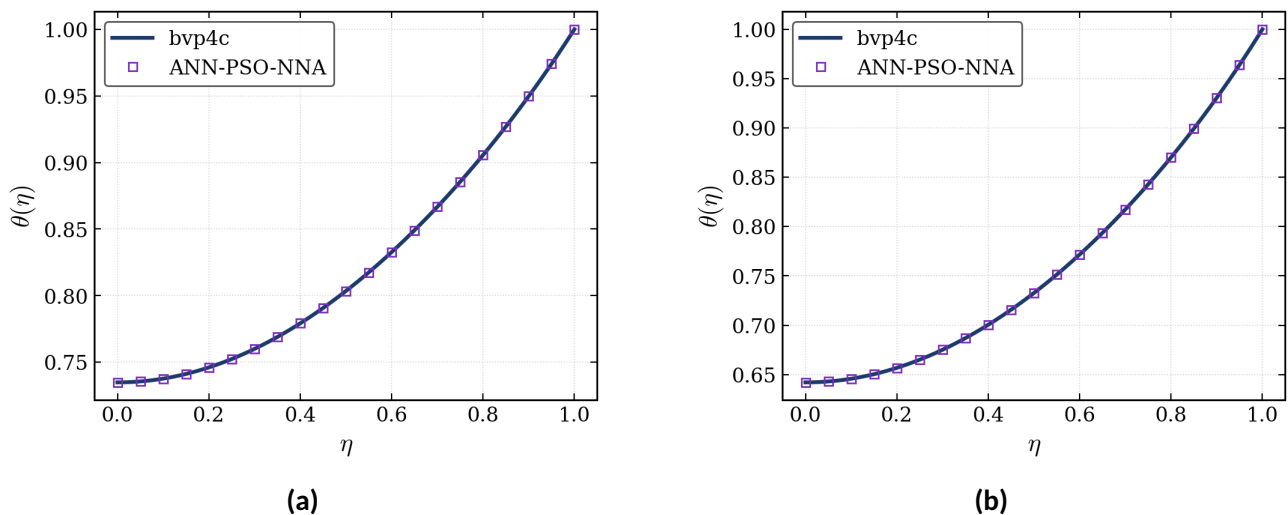


Fig. 16. Comparison of the reference and hybrid network temperature $\theta(\eta)$ for (a) $H_s = -1$ and (b) $H_s = -1.5$

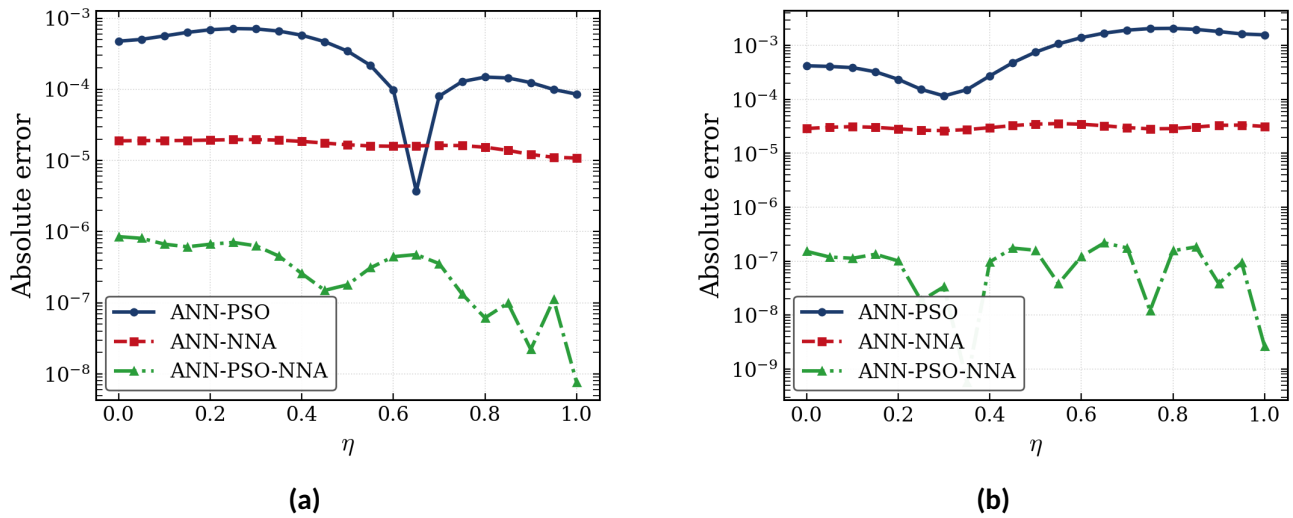


Fig. 17. Absolute error of the temperature for the three optimiser variants at (a) $H_s = -1$ and (b) $H_s = -1.5$

The concentration solution is displayed in Figure 18 and its absolute error in Figure 19, with a maximum hybrid error of $1.041\text{e-}6$ and $1.918\text{e-}6$ for the two cases. In each comparison the dips visible in the error of the better optimisers mark the nodes at which the signed error changes sign, and the overall ordering of the three strategies is preserved across all fields.

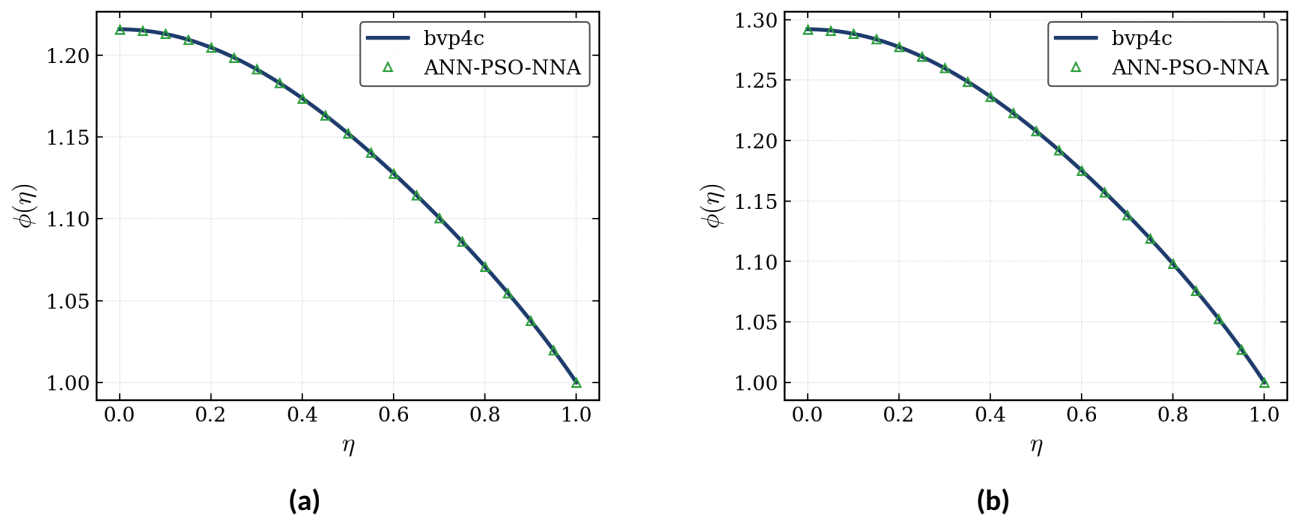


Fig. 18. Comparison of the reference and hybrid network concentration $\phi(\eta)$ for (a) $H_s = -1$ and (b) $H_s = -1.5$

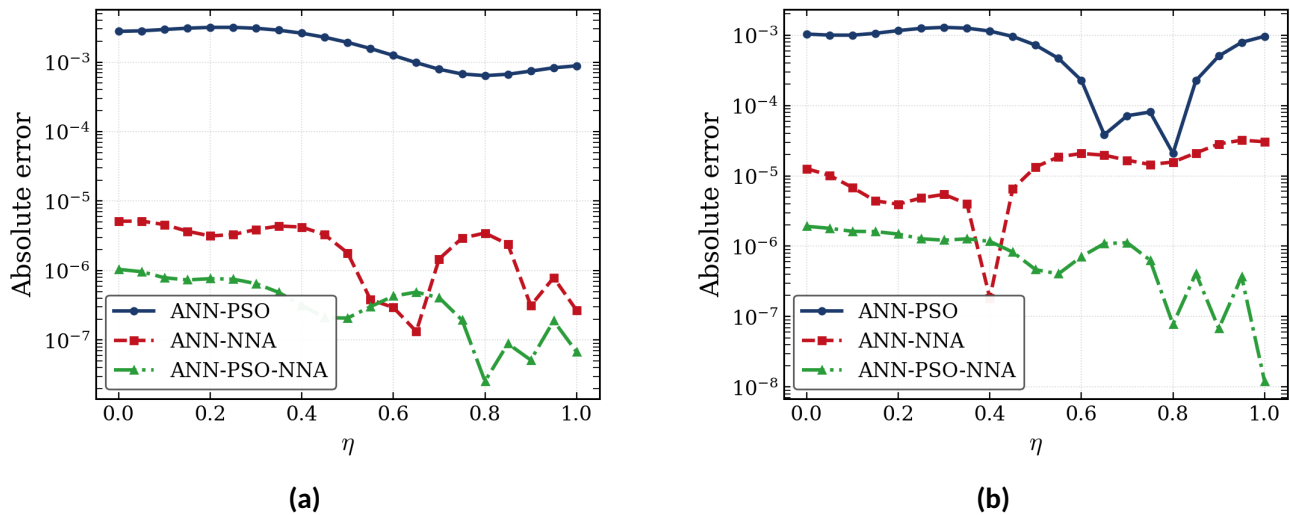


Fig. 19. Absolute error of the concentration for the three optimiser variants at (a) $H_s = -1$ and (b) $H_s = -1.5$

9.6 Statistical Analysis over Independent Runs

Because the hybrid optimiser is stochastic, the robustness of the training is assessed by repeating it one hundred times from independent random populations for each of four scenarios, namely the velocity at $Ha = 0$ and at $Ha = 0.1$ and the temperature at $H_s = -1$ and at $H_s = -1.5$. Figure 20 shows the converged fitness of every run for the velocity and the temperature scenarios, Figure 21 the Theil inequality coefficient of Eq. (45), and Figure 22 the mean squared error of Eq. (46). The scatter of these measures over the runs is narrow and stays at the levels reported for the best run, which confirms that the hybrid scheme converges reliably rather than occasionally. Averaged over the hundred runs, the velocity attains a mean fitness of $1.406e-7$ at $Ha = 0$ and a mean Theil coefficient of $2.096e-6$, while the temperature attains an average mean squared error of $5.272e-12$ at $H_s = -1$.

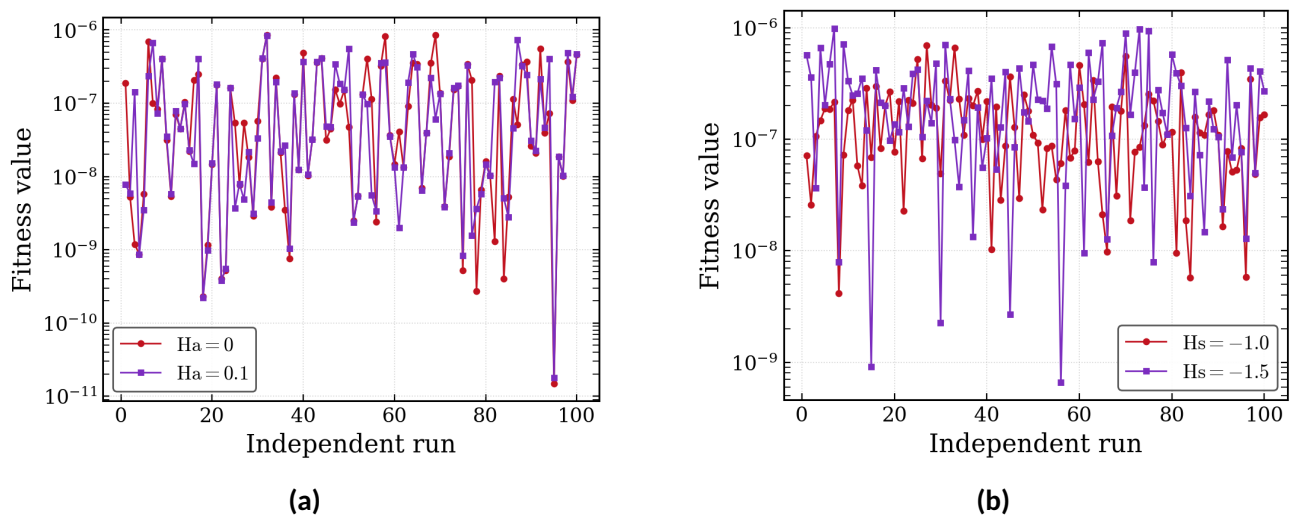


Fig. 20. Converged fitness over one hundred independent runs for (a) the velocity scenarios and (b) the temperature scenarios

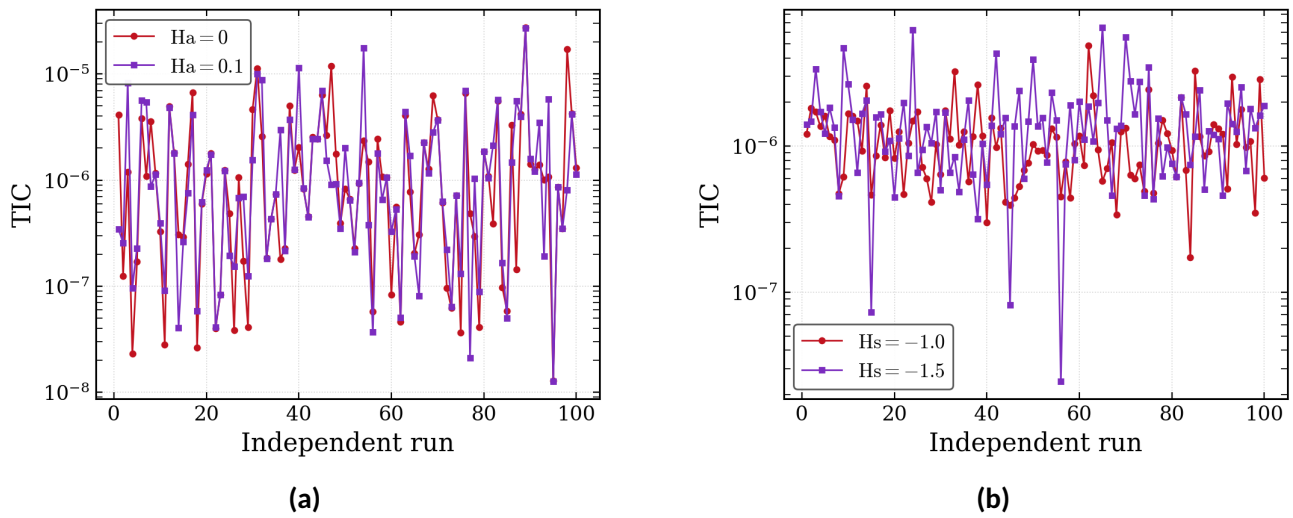


Fig. 21. Theil inequality coefficient over one hundred independent runs for (a) the velocity scenarios and (b) the temperature scenarios.

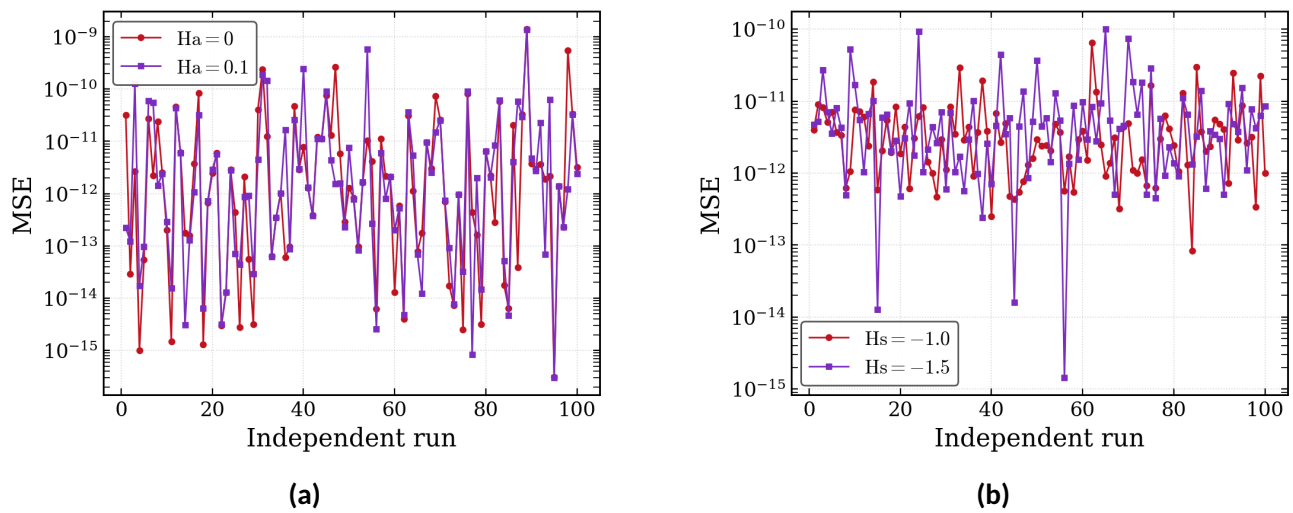


Fig. 22. Mean squared error over one hundred independent runs for (a) the velocity scenarios and (b) the temperature scenarios.

The node by node statistics of the absolute error over the hundred runs are reported in Table 4 and Table 5 for the two velocity scenarios and in Table 6 and Table 7 for the two temperature scenarios. At every reporting node the minimum, maximum, average and standard deviation of the error remain small, the average error staying near 10^{-6} to 10^{-5} throughout the domain, which quantifies the run to run stability already visible in the preceding figures.

Table 4
Statistics of the velocity absolute error over one
hundred independent runs at $Ha = 0$

η	Minimum	Maximum	Average	Std. dev.
0.00	6.215e-9	2.801e-5	2.525e-6	4.426e-6
0.05	9.849e-9	2.913e-5	2.542e-6	4.486e-6
0.10	5.741e-9	3.023e-5	2.560e-6	4.549e-6
0.15	6.212e-10	3.130e-5	2.585e-6	4.614e-6
0.20	1.925e-9	3.236e-5	2.607e-6	4.683e-6
0.25	8.954e-10	3.338e-5	2.629e-6	4.756e-6
0.30	3.659e-9	3.436e-5	2.659e-6	4.825e-6
0.35	6.363e-9	3.530e-5	2.689e-6	4.893e-6
0.40	5.050e-9	3.619e-5	2.718e-6	4.962e-6
0.45	2.001e-9	3.704e-5	2.747e-6	5.030e-6
0.50	6.880e-10	3.786e-5	2.776e-6	5.097e-6
0.55	1.229e-9	3.865e-5	2.807e-6	5.162e-6
0.60	6.700e-9	3.941e-5	2.840e-6	5.226e-6
0.65	1.430e-8	4.016e-5	2.873e-6	5.291e-6
0.70	1.576e-8	4.090e-5	2.904e-6	5.357e-6
0.75	1.302e-8	4.162e-5	2.933e-6	5.424e-6
0.80	1.455e-10	4.231e-5	2.962e-6	5.491e-6
0.85	7.949e-9	4.297e-5	2.995e-6	5.554e-6
0.90	5.282e-9	4.360e-5	3.028e-6	5.615e-6
0.95	2.560e-9	4.419e-5	3.058e-6	5.674e-6
1.00	5.784e-11	4.476e-5	3.087e-6	5.733e-6

Table 5

Statistics of the velocity absolute error over one hundred independent runs at $Ha = 0.1$

η	Minimum	Maximum	Average	Std. dev.
0.00	1.750e-8	2.675e-5	2.495e-6	4.389e-6
0.05	6.304e-9	2.794e-5	2.515e-6	4.449e-6
0.10	1.055e-8	2.910e-5	2.536e-6	4.518e-6
0.15	4.926e-9	3.024e-5	2.558e-6	4.595e-6
0.20	1.497e-10	3.135e-5	2.595e-6	4.670e-6
0.25	4.238e-9	3.242e-5	2.643e-6	4.745e-6
0.30	6.012e-9	3.346e-5	2.698e-6	4.819e-6
0.35	1.830e-9	3.446e-5	2.750e-6	4.897e-6
0.40	3.284e-9	3.541e-5	2.802e-6	4.978e-6
0.45	1.364e-9	3.631e-5	2.851e-6	5.061e-6
0.50	4.546e-9	3.718e-5	2.900e-6	5.146e-6
0.55	3.220e-9	3.802e-5	2.951e-6	5.232e-6
0.60	4.006e-9	3.883e-5	3.002e-6	5.319e-6
0.65	1.067e-8	3.961e-5	3.052e-6	5.407e-6
0.70	1.636e-8	4.038e-5	3.100e-6	5.498e-6
0.75	1.640e-8	4.113e-5	3.146e-6	5.590e-6
0.80	1.515e-8	4.185e-5	3.190e-6	5.684e-6
0.85	1.363e-8	4.253e-5	3.233e-6	5.778e-6
0.90	1.184e-8	4.318e-5	3.275e-6	5.870e-6
0.95	5.617e-9	4.380e-5	3.314e-6	5.962e-6
1.00	2.708e-9	4.439e-5	3.352e-6	6.051e-6

Table 6

Statistics of the temperature absolute error over one hundred independent runs at $H_s = -1$

η	Minimum	Maximum	Average	Std. dev.
0.00	5.631e-8	1.357e-5	2.814e-6	2.126e-6
0.05	5.460e-8	1.327e-5	2.832e-6	2.138e-6
0.10	1.933e-8	1.247e-5	2.567e-6	2.043e-6
0.15	3.255e-8	1.150e-5	2.165e-6	1.810e-6
0.20	5.682e-8	1.084e-5	1.924e-6	1.588e-6
0.25	6.838e-8	1.057e-5	1.939e-6	1.521e-6
0.30	3.025e-9	1.044e-5	2.072e-6	1.550e-6
0.35	3.394e-9	1.009e-5	2.136e-6	1.521e-6
0.40	2.466e-8	9.289e-6	2.065e-6	1.376e-6
0.45	6.444e-9	8.037e-6	1.832e-6	1.310e-6
0.50	8.725e-9	6.546e-6	1.539e-6	1.363e-6
0.55	1.064e-8	6.122e-6	1.283e-6	1.368e-6
0.60	1.091e-8	4.920e-6	1.101e-6	1.196e-6
0.65	1.050e-8	4.921e-6	9.549e-7	1.057e-6
0.70	6.830e-9	6.527e-6	1.092e-6	1.114e-6
0.75	5.973e-8	7.601e-6	1.404e-6	1.208e-6
0.80	6.421e-8	7.782e-6	1.439e-6	1.192e-6
0.85	5.095e-8	7.164e-6	1.106e-6	1.084e-6
0.90	4.464e-9	6.413e-6	8.118e-7	1.056e-6
0.95	3.132e-8	6.487e-6	8.711e-7	1.051e-6
1.00	2.457e-10	7.507e-6	7.966e-7	1.185e-6

Table 7

Statistics of the temperature absolute error over one hundred independent runs at $H_s = -1.5$

η	Minimum	Maximum	Average	Std. dev.
0.00	3.361e-8	1.653e-5	3.440e-6	2.736e-6
0.05	1.467e-8	1.641e-5	3.475e-6	2.741e-6
0.10	1.366e-8	1.537e-5	3.176e-6	2.622e-6
0.15	2.799e-8	1.265e-5	2.679e-6	2.338e-6
0.20	5.659e-8	1.186e-5	2.340e-6	2.097e-6
0.25	7.591e-9	1.206e-5	2.271e-6	2.098e-6
0.30	1.699e-8	1.224e-5	2.392e-6	2.157e-6
0.35	2.267e-10	1.200e-5	2.479e-6	2.111e-6
0.40	3.562e-8	1.116e-5	2.428e-6	1.977e-6
0.45	1.500e-8	9.869e-6	2.237e-6	1.959e-6
0.50	2.806e-9	1.053e-5	2.034e-6	2.028e-6
0.55	1.035e-8	1.008e-5	1.780e-6	2.062e-6
0.60	6.681e-10	8.009e-6	1.576e-6	1.827e-6
0.65	1.269e-8	8.717e-6	1.350e-6	1.641e-6
0.70	7.534e-9	9.640e-6	1.418e-6	1.757e-6
0.75	4.394e-8	1.028e-5	1.866e-6	1.852e-6
0.80	2.065e-8	1.021e-5	1.965e-6	1.809e-6
0.85	1.774e-8	9.324e-6	1.535e-6	1.674e-6
0.90	4.733e-9	8.154e-6	1.117e-6	1.608e-6
0.95	8.680e-9	7.659e-6	1.151e-6	1.559e-6
1.00	1.024e-9	8.141e-6	1.040e-6	1.677e-6

9.7 Supervised Training Diagnostics

The supervised Levenberg–Marquardt surrogate of Section 8 provides the conventional view of the learning process. Figure 23 collects the diagnostics for the velocity: panel (a) shows the training, validation and test mean squared errors falling together over the epochs to the best validation point, panel (b) shows the gradient magnitude, the adaptive damping parameter μ and the validation check counter, panel (c) is the histogram of the fit errors and panel (d) is the regression of the network output against the target, for which the correlation reaches $R = 1.000000$.

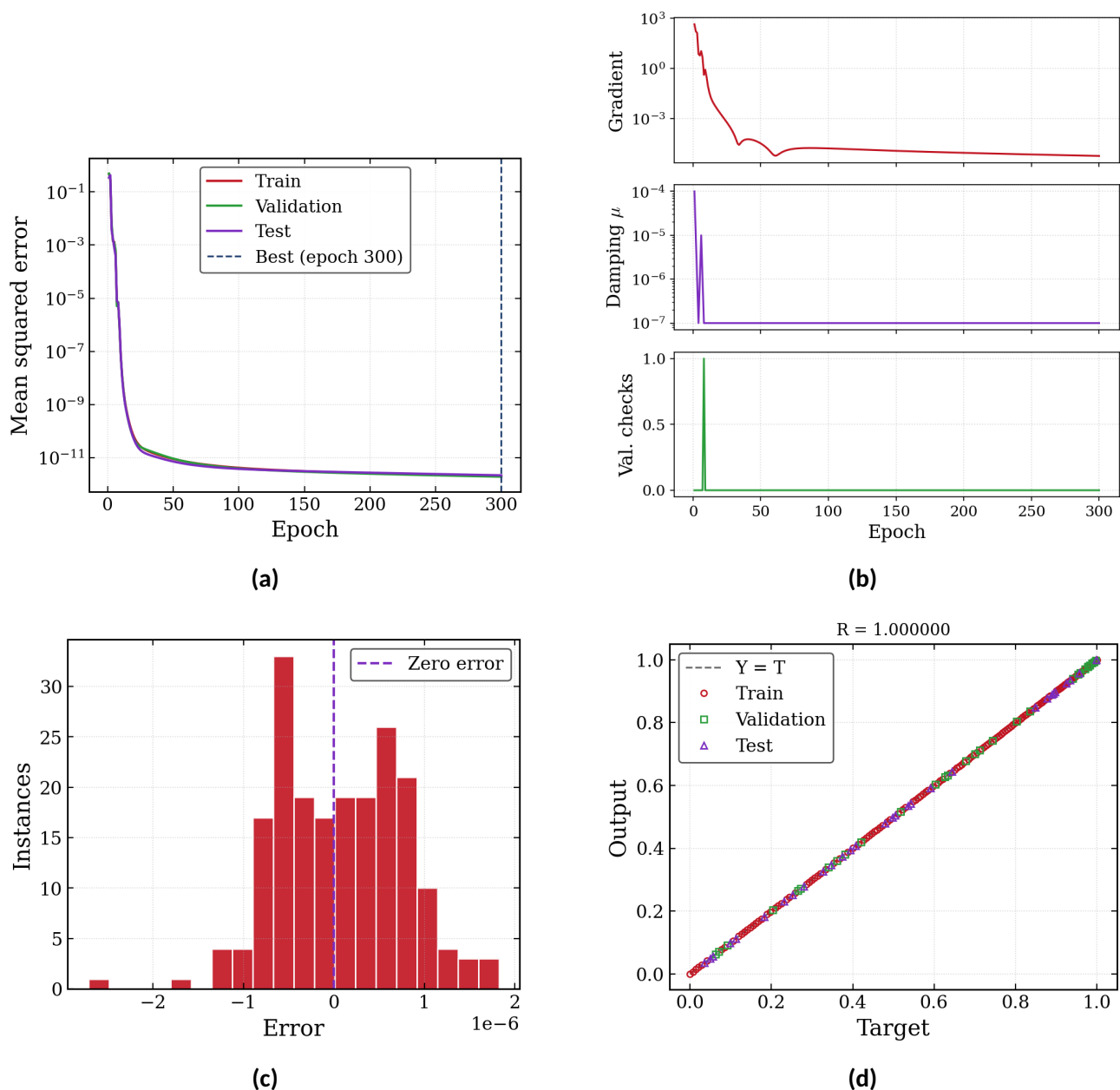


Fig. 23. Supervised Levenberg–Marquardt diagnostics for the velocity: (a) performance, (b) training state, (c) error histogram and (d) regression

The corresponding diagnostics for the temperature are gathered in Figure 24. Here, the validation check mechanism halts the training after 24 epochs once the validation error ceases to improve, and the regression correlation is $R = 0.999999$. The near unit correlation and the symmetric error histograms confirm that the supervised surrogate reproduces the reference fields to high accuracy and corroborate the unsupervised results.

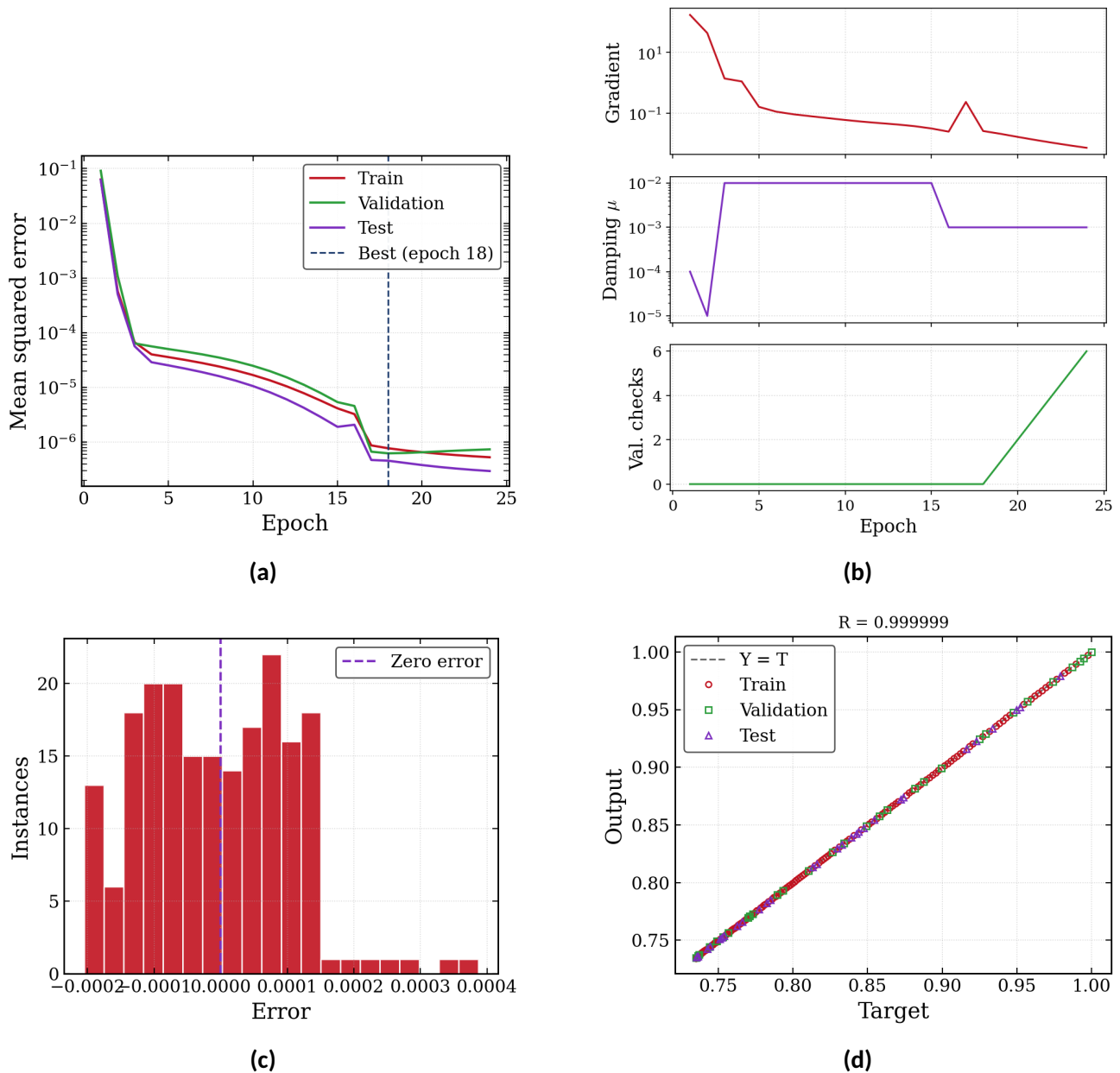


Fig. 24. Supervised Levenberg–Marquardt diagnostics for the temperature: (a) performance, (b) training state, (c) error histogram and (d) regression

9.8 Comparison with Existing Approaches

Table 8 places the present study against a representative selection of earlier work on squeezing and channel flows of nanofluids and on neural solvers for boundary value problems. The comparison is drawn along the physical mechanisms that the model resolves, the type of solver employed, and whether the study supplies a verification of the reduction and an analysis of well posedness. The squeezing analyses [14, 15] and the porous medium treatments [19] resolve the magnetic and porous effects but rely on a single phase description and a purely numerical or perturbation solution. The two phase radiative study [17] adds the Buongiorno fluxes, and the heuristic physics-informed framework [49] introduces a population trained network for squeezing flow, yet neither joins the Darcy–Forchheimer porous resistance to the two phase transport, and neither pairs the learned solution with a Green function existence, uniqueness and stability analysis. The classical physics-informed network

[29], with the error theory of [35, 36], supplies the gradient based counterpart but is not specialised to the present hybrid model. The present work combines all of these mechanisms, attaches the residual controlled reference of accuracy below 10^{-9} , trains the network by the hybrid PSO–NNA search whose hybrid error of $9.300\text{e-}9$ improves on the particle swarm and neural network algorithm alone by four to five orders of magnitude, and certifies the operating point through the spectral radius 0.270 and the bound (37).

Table 8

Comparison of the present study with representative earlier approaches in terms of the resolved physical mechanisms, the solver, and the verification and well posedness analysis provided

Study	MHD	Darcy–Forchheimer porous	Two phase Buongiorno	Nonlinear radiation	Solver type	Symbolic check	Well posedness analysis
Mustafa et al. [14]	yes	no	no	no	analytic/numeric	no	no
Hayat et al. [15]	yes	no	no	yes	numeric	no	no
Sheikholeslami et al. [17]	yes	no	yes	yes	numeric	no	no
Raissi et al. [29]	–	–	–	–	gradient PINN	no	no
Aslam et al. [49]	yes	no	no	no	heuristic PINN	no	no
Present study	yes	yes	yes	yes	heuristic PINN	yes	yes

10. Symbolic Verification of the Reduction

To guarantee that the similarity reduction is free of algebraic error, the passage from the governing partial differential equations to the ordinary differential system was reproduced with a computer algebra system. The similarity ansatz of (6) was substituted into the continuity equation and shown to cancel identically. The momentum operator was assembled term by term and its difference from the reported Eq. (16) simplified to zero, and the same check was applied to the energy and concentration equations. Finally the coefficient of θ'' in the energy equation was confirmed to equal $1 + R_d$, which validates the Rosseland linearisation of the radiative flux. Every check returned a zero residual, so the differential operators enforced by the network in (41) are identical to those of the manuscript, of the reference solver and of the Green function reformulation of Lemma 1. The residual norm reported by the reference collocation solver was below 10^{-9} for all cases considered, which provides an independent numerical confirmation of the same equations.

11. Conclusions

The unsteady magnetohydrodynamic squeezing flow of a Fe_3O_4 – MoS_2 hybrid nanofluid through a Darcy–Forchheimer porous channel was modelled with the Buongiorno two phase description, nonlinear thermal radiation and a temperature dependent heat source or sink, analysed as a fixed point problem and solved by a heuristic physics-informed neural network whose accuracy was verified against a residual controlled reference solution. On the analytical side, the similarity system was written through explicit Green functions as a fixed point equation in a Banach space, for which a global contraction was established under an explicit smallness condition on the parameters and, at the operating values, the linearised momentum operator was shown to have spectral radius 0.270; this yields local uniqueness and an a posteriori error certificate whose amplification factor is about 1.37, so the residual that the network minimises bounds the distance to the exact solution. On the physical side, the squeeze number, the Hartmann number, the porosity parameter and the Forchheimer number each retard the flow and, under the fixed flux constraint of the squeezing geometry, redistribute the velocity profile about a common crossing point. The temperature is governed most strongly by the heat source or sink parameter and is raised further by the Prandtl number, the radiation parameter

and thermophoresis, while the concentration is lowered by Brownian diffusion and raised by thermophoresis, the two Buongiorno mechanisms acting in opposition. The hybrid optimiser that couples particle swarm optimisation with the neural network algorithm and a trust region polish outperformed either metaheuristic alone by several orders of magnitude, reaching a maximum absolute error of $9.300\text{e-}9$ for the velocity, while the supervised Levenberg–Marquardt surrogate returned a regression correlation of $R = 1.000000$. Across one hundred independent runs the training proved stable, with average errors held near 10^{-6} throughout the domain, and the attained accuracy agreed with the certified bound. The framework is general and can be applied to other strongly coupled boundary value problems in transport theory, and a natural extension is the treatment of inverse problems in which unknown physical parameters are recovered jointly with the fields.

Nomenclature

u, v	velocity components	Ha	Hartmann number
x, y	Cartesian coordinates	K_p	porosity parameter
t	time	Fr	Forchheimer number
T	temperature	S	squeeze number
C	nanoparticle volume fraction	Pr	Prandtl number
$h(t)$	channel half width	R_d	radiation parameter
B_0	magnetic field strength	N_b	Brownian parameter
k_1	permeability	N_t	thermophoresis parameter
C_b	Forchheimer coefficient	Sc	Schmidt number
D_B	Brownian diffusion coefficient	H_s	heat source or sink parameter
D_T	thermophoretic diffusion coefficient	η	similarity variable
Q_0	heat generation coefficient	θ	dimensionless temperature
q_r	radiative heat flux	ϕ	dimensionless concentration
f	dimensionless stream function	ρ_{hnf}	hybrid nanofluid density
$\mathcal{F}$	fitness functional	μ_{hnf}	hybrid nanofluid viscosity
K, g	Green functions	σ_{hnf}	hybrid nanofluid conductivity
M_k, N_k	integral operator norms	k_{hnf}	hybrid thermal conductivity
Λ, ρ	contraction modulus and spectral radius	ϕ_1, ϕ_2	nanoparticle volume fractions
R	regression correlation	TIC	Theil inequality coefficient

Acknowledgement

This research was not funded by any grant.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Stefan, J. (1874). Versuche über die scheinbare adhäsion. *Sitzungsberichte der Akademie der Wissenschaften in Wien*, 69, 713–721.
- [2] Wang, C. Y. (1989). Exact solutions of the unsteady Navier-Stokes equations. *Applied Mechanics Reviews*, 42(11S), S269–S282. <https://doi.org/10.1115/1.3152400>
- [3] Choi, S. U. S., & Eastman, J. A. (1995). Enhancing thermal conductivity of fluids with nanoparticles. *ASME International Mechanical Engineering Congress and Exposition*, 66, 99–105.

- [4] Tiwari, R. K., & Das, M. K. (2007). Heat transfer augmentation in a two-sided lid-driven differentially heated square cavity utilizing nanofluids. *International Journal of Heat and Mass Transfer*, 50(9-10), 2002–2018. <https://doi.org/10.1016/j.ijheatmasstransfer.2006.09.034>
- [5] Buongiorno, J. (2006). Convective transport in nanofluids. *Journal of Heat Transfer*, 128(3), 240–250. <https://doi.org/10.1115/1.2150834>
- [6] Khan, W. A., & Pop, I. (2010). Boundary-layer flow of a nanofluid past a stretching sheet. *International Journal of Heat and Mass Transfer*, 53(11-12), 2477–2483. <https://doi.org/10.1016/j.ijheatmasstransfer.2010.01.032>
- [7] Rana, P., & Bhargava, R. (2012). Flow and heat transfer of a nanofluid over a nonlinearly stretching sheet: A numerical study. *Communications in Nonlinear Science and Numerical Simulation*, 17(1), 212–226. <https://doi.org/10.1016/j.cnsns.2011.05.009>
- [8] Devi, S. P. A., & Devi, S. S. U. (2016). Numerical investigation of hydromagnetic hybrid Cu-Al₂O₃/water nanofluid flow over a permeable stretching sheet with suction. *International Journal of Nonlinear Sciences and Numerical Simulation*, 17(5), 249–257. <https://doi.org/10.1515/ijnsns-2016-0037>
- [9] Waini, I., Ishak, A., & Pop, I. (2019). Hybrid nanofluid flow and heat transfer over a nonlinear permeable stretching/shrinking surface. *International Journal of Numerical Methods for Heat and Fluid Flow*, 29(9), 3110–3127. <https://doi.org/10.1108/HFF-01-2019-0057>
- [10] Maxwell, J. C. (1881). *A treatise on electricity and magnetism*. Clarendon Press, Oxford.
- [11] Hamilton, R. L., & Crosser, O. K. (1962). Thermal conductivity of heterogeneous two-component systems. *Industrial and Engineering Chemistry Fundamentals*, 1(3), 187–191. <https://doi.org/10.1021/i160003a005>
- [12] Brinkman, H. C. (1952). The viscosity of concentrated suspensions and solutions. *The Journal of Chemical Physics*, 20(4), 571–581. <https://doi.org/10.1063/1.1700493>
- [13] Sheikholeslami, M., & Ganji, D. D. (2013). Heat transfer of Cu-water nanofluid flow between parallel plates. *Powder Technology*, 235, 873–879. <https://doi.org/10.1016/j.powtec.2012.12.025>
- [14] Mustafa, M., Hayat, T., & Obaidat, S. (2012). On heat and mass transfer in the unsteady squeezing flow between parallel plates. *Meccanica*, 47(7), 1581–1589. <https://doi.org/10.1007/s11012-011-9507-8>
- [15] Hayat, T., Muhammad, T., Qayyum, A., Alsaedi, A., & Mustafa, M. (2016). On squeezing flow of nanofluid in the presence of magnetic field effects. *Journal of Molecular Liquids*, 213, 179–185. <https://doi.org/10.1016/j.molliq.2015.09.004>
- [16] Turkyilmazoglu, M. (2012). Exact analytical solutions for heat and mass transfer of MHD slip flow in nanofluids. *Chemical Engineering Science*, 84, 182–187. <https://doi.org/10.1016/j.ces.2012.07.004>
- [17] Sheikholeslami, M., Ganji, D. D., Javed, M. Y., & Ellahi, R. (2015). Effect of thermal radiation on magnetohydrodynamics nanofluid flow and heat transfer by means of two phase model. *Journal of Magnetism and Magnetic Materials*, 374, 36–43. <https://doi.org/10.1016/j.jmmm.2014.10.013>
- [18] Forchheimer, P. (1901). Wasserbewegung durch boden. *Zeitschrift des Vereines Deutscher Ingenieure*, 45, 1782–1788.
- [19] Nield, D. A., & Bejan, A. (2017). *Convection in porous media* (5th ed.). Springer, Cham. <https://doi.org/10.1007/978-3-319-49562-0>
- [20] Rosseland, S. (1936). *Astrophysik und atom-theoretische grundlagen*. Springer, Berlin.
- [21] Sheikholeslami, M. (2018). Numerical investigation of nanofluid free convection under the influence of electric field in a porous enclosure. *Journal of Molecular Liquids*, 249, 1212–1221. <https://doi.org/10.1016/j.molliq.2017.11.141>

- [22] Gireesha, B. J., Mahanthesh, B., Thammanna, G. T., & Sampathkumar, P. B. (2018). Hall effects on dusty nanofluid two-phase transient flow past a stretching sheet using KVL model. *Journal of Molecular Liquids*, 256, 139–147. <https://doi.org/10.1016/j.molliq.2018.01.186>
- [23] Animasaun, I. L., Mahanthesh, B., & Koriko, O. K. (2018). On the motion of non-Newtonian Eyring-Powell fluid conveying tiny gold particles due to generalized surface slip velocity and buoyancy. *International Journal of Applied and Computational Mathematics*, 4(6), 137. <https://doi.org/10.1007/s40819-018-0550-1>
- [24] Kierzenka, J., & Shampine, L. F. (2001). A BVP solver based on residual control and the MATLAB PSE. *ACM Transactions on Mathematical Software*, 27(3), 299–316. <https://doi.org/10.1145/502800.502801>
- [25] Shampine, L. F., Gladwell, I., & Thompson, S. (2003). *Solving odes with matlab*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511615542>
- [26] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303–314. <https://doi.org/10.1007/BF02551274>
- [27] Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
- [28] Lagaris, I. E., Likas, A., & Fotiadis, D. I. (1998). Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks*, 9(5), 987–1000. <https://doi.org/10.1109/72.712178>
- [29] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- [30] Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., & Piccialli, F. (2022). Scientific machine learning through physics-informed neural networks: Where we are and what is next. *Journal of Scientific Computing*, 92(3), 88. <https://doi.org/10.1007/s10915-022-01939-z>
- [31] Baydin, A. G., Pearlmutter, B. A., Radul, A. A., & Siskind, J. M. (2018). Automatic differentiation in machine learning: A survey. *Journal of Machine Learning Research*, 18(153), 1–43. <https://doi.org/10.5555/3296976.3296981>
- [32] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *International Conference on Learning Representations*.
- [33] Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 249–256.
- [34] Shin, Y., Darbon, J., & Karniadakis, G. E. (2020). On the convergence of physics-informed neural networks for linear second-order elliptic and parabolic type PDEs. *Communications in Computational Physics*, 28(5), 2042–2074. <https://doi.org/10.4208/cicp.OA-2020-0156>
- [35] Mishra, S., & Molinaro, R. (2022). Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs. *IMA Journal of Numerical Analysis*, 42(2), 981–1022. <https://doi.org/10.1093/imanum/drab032>
- [36] De Ryck, T., & Mishra, S. (2022). Generic bounds on the approximation error for physics-informed and related learning algorithms. *Advances in Computational Mathematics*, 48(6), 79. <https://doi.org/10.1007/s10444-022-09980-0>
- [37] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks*, 4, 1942–1948. <https://doi.org/10.1109/ICNN.1995.488968>

- [38] Shi, Y., & Eberhart, R. (1998). A modified particle swarm optimizer. *Proceedings of the IEEE International Conference on Evolutionary Computation*, 69–73. <https://doi.org/10.1109/ICEC.1998.699146>
- [39] Eberhart, R. C., Shi, Y., & Kennedy, J. (2001). *Swarm intelligence*. Morgan Kaufmann, San Francisco.
- [40] Sadollah, A., Sayyaadi, H., & Yadav, A. (2018). A dynamic metaheuristic optimization model inspired by biological nervous systems: Neural network algorithm. *Applied Soft Computing*, 71, 747–782. <https://doi.org/10.1016/j.asoc.2018.06.030>
- [41] Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2(2), 164–168. <https://doi.org/10.1090/qam/10666>
- [42] Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431–441. <https://doi.org/10.1137/0111030>
- [43] Theil, H. (1961). *Economic forecasts and policy*. North-Holland, Amsterdam.
- [44] Banach, S. (1922). Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae*, 3, 133–181.
- [45] Schauder, J. (1930). Der fixpunktsatz in funktionalräumen. *Studia Mathematica*, 2, 171–180.
- [46] Granas, A., & Dugundji, J. (2003). *Fixed point theory*. Springer. <https://doi.org/10.1007/978-0-387-21593-8>
- [47] Agarwal, R. P., Meehan, M., & O'Regan, D. (2001). *Fixed point theory and applications* (Vol. 141). Cambridge University Press. <https://doi.org/10.1017/CBO9780511542633>
- [48] Stakgold, I., & Holst, M. (2011). *Green's functions and boundary value problems* (3rd ed.). John Wiley; Sons. <https://doi.org/10.1002/9781118031000>
- [49] Aslam, M. N., Shaukat, A., & Riaz, M. B. (2024). Heuristic based physics informed neural network approach for the unsteady magnetohydrodynamic viscous flow of a hybrid nanofluid between parallel sheets. *International Communications in Heat and Mass Transfer*, 159, 108320. <https://doi.org/10.1016/j.icheatmasstransfer.2024.108320>