Response to Reviewer XYcj

Comment — The examples selected to demonstrate cd-PINN's performance are limited to relatively simple cases.

Reply: In order to test whether our method is applicable to more complicated cases, here we look into a multiscale model for p53 activation, which is a key gene closely related to cancer development [1]. This model is composed of seven coupled ordinary differential functions,

$$
\begin{aligned}
&\frac{d[MDM2]}{dt} = \frac{k_{MD_S}[S]}{K_{MD_S}+[S]} + \frac{k_{M_p}[p53]^{n_1}}{K_{M_p}^{n_1}+[p53]^{n_1}} + D_{MA}[MA] + \frac{k_{Dp4}[MDM2_p]}{K_{Mp}+[MDM2_p]} \\
&\quad - k_{MA}[MDM2][ARF] - \frac{k_{p_M}[Akt][MDM2]}{K_{Akt_M}+[MDM2]} - d_{Mdm2}[MDM2], \\
&\frac{d[MDM2_p]}{dt} = D_{MpA}[M_pA] + \frac{k_{p_M}[Akt][MDM2]}{K_{Akt_M}+[MDM2]} - k_{MpA}[MDM2_p][ARF] \\
&\quad - \frac{k_{Dp4}[MDM2_p]}{K_{Mp}+[MDM2_p]} - d_{Mp}[MDM2_p], \\
&\frac{d[MA]}{dt} = k_{MA}[MDM2][AFR] - D_{MA}[MA] - d_{MA}[MA], \\
&\frac{d[M_pA]}{dt} = k_{MpA}[MDM2_p][ARF] - D_{MpA}[M_pA] - d_{MpA}[M_pA], \\
&\frac{d[p53]}{dt} = k_{p53} - \frac{k_{M53}[MDM2][p53]}{K_{M53}+[p53]} - \frac{k_{Mp53}[MDM2_p][p53]}{K_{Mp53}+[p53]} - d_{p53}[p53], \\
&\frac{d[PTEN]}{dt} = k_{PTEN} + \frac{k_{P_p}[p53]^{n_2}}{K_{P_p}^{n_2}+[p53]^{n_2}} - d_{PTEN}[PTEN], \\
&\frac{d[Akt]}{dt} = \frac{k_{A_S}[S]}{K_{A_S}+[S]}\frac{[Akt]_t-[Akt]}{K_0+[Akt]_t-[Akt]} - \frac{k_{DP3}[Akt]}{K_{Akt}+[Akt]} - \frac{k_{A_p}[PTEN][Akt]}{K_{AP}+[Akt]},
\end{aligned}
\tag{1}
$$

with the initial values (ranging from 0.005 to 5) and model parameters (ranging from 0.05 to 50) directly taken from the cited paper [1].

In this task, we expect that the cd-PINN can correctly learn the solutions to the above model from time $t = 0$ to $t = 5$ with respect to arbitrary inputs of $[S] \in [0.1, 10.0]$ and $[ARF] \in [0.1, 1.5]$. For this purpose, we uniformly select 41×41 groups of $[S]$ and $[ARF]$ during the interval $[0.1, 10.0] \times [0.1, 1.5]$ to generate the test data. Meanwhile, the real data consists of 51 points corresponding to the solution with $[S] = 2.575$, $[ARF] = 0.45$ and 51 points corresponding to the solution with $[S] = 7.525$, $[ARF] = 0.695$. The training data set includes the real data and $N_t = N_0 = 2^{13}$ residual data points.

The final predictions of cd-PINN on this example are summarized in Figure 1, whose MSE is 3.32×10^{-4}, two order lower than the MSE of the model without C.D. constraints (5.38×10^{-2}). Therefore, it can be concluded that our cd-PINN is capable for handling more challenging situations.

Changes: See Section 3.4 and Appendix B.4 – A Multiscale Model for P53 Activation for details.

Comment — Additionally, while the parameter dependence may be continuous, it does not necessarily have to be differentiable. This distinction is crucial, as the lack of differentiability can present challenges in certain contexts, such as in hyperbolic conservation laws, e.g. Burger's

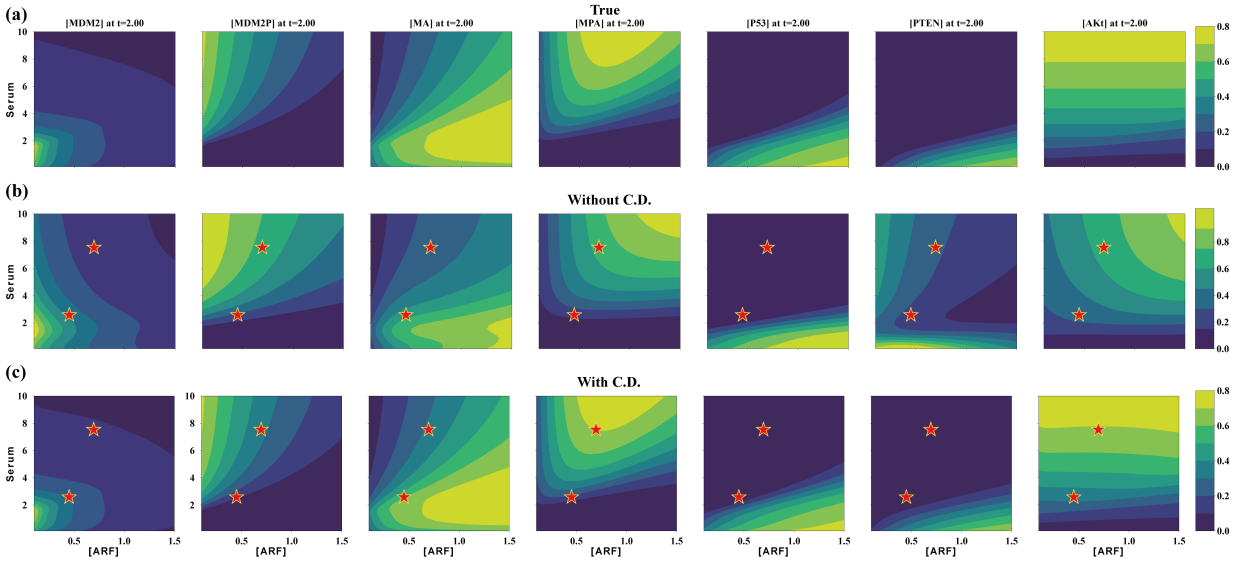

Figure 1: Phase diagram for the solutions of p53 activation model at time $t = 2$. The expression levels for seven genes are calculated by (a) the ODE solver, (b) PINN without C.D. constraints, and (c) cd-PINN separately in comparison.

equation, where discontinuities like shock locations arise. Please discuss how cd-PINN might handle systems with non-differentiable parameter dependence.

Reply: The differentiability is a special case of continuity. In this paper, we focus on the former due to its easy implementation. While for hyperbolic conservation laws, like Burger's equations, we may turn to the more general conditions, like the Rankine-Hugoniot jump condition, to determine the exact locations where the shock structure arises. Furthermore, our preliminary results show that incorporating the condintious dependence could indeed improve generality of cd-PINN on the visocity solutions of Burgers equation, but is not directly applicable to the non-viscous Burgers equation.

Changes: See Page 11 for changes.

"In the current paper, we restrict our study to ordinary differential equations for clarity. Obviously, the same approach is applicable to partial differential equations too, e.g. the viscosity in Burgers equation or the Reynolds number in Navier-Stokes equations. However, it should be noted that the PDE cases are far more complicated in general. For example, in many cases the parameter dependence of PDEs may be continuous but not necessarily differentiable. This subtle distinction is crucial for certain contexts, such as the shock structures in hyperbolic conservation laws. Under these situations, we need to turn to more general conditions, like the Rankine-Hugoniot jump condition, to determine the exact locations where the shock structure arises. The related work is ongoing."

Comment — Including a logarithmic scale in Figure 4 would be beneficial, Please give a better explanations to your figures on how you compare with other methods.

Reply: Thank you for suggestions. We have revised Figure 4 by adding logarithmic scale plots and provided more detailed explanations of our comparison methods.

References

[1] Xinyu Tian, Bo Huang, Xiao-Peng Zhang, Mingyang Lu, Feng Liu, José N Onuchic, and Wei Wang. Modeling the response of a tumor-suppressive network to mitogenic and oncogenic signals. Proceedings of the National Academy of Sciences, 114(21):5337–5342, 2017.