

# 具分次跳躍之 EBV ODE 的 PINN

Min-Jhe Lu

September 26, 2025

## Abstract

我們提出一條針對 EBV (Epstein–Barr Virus) 血漿動力學的 PINN 實作路線：在連續 ODE  $(B, R, E)$  外，以 LQ 放療模型給出的分次跳躍作外加事件；將跳躍嵌入解的構造式、在事件附近對 ODE 殘差做平滑遮罩。資料端採左設限對數常態似然以處理 LOD (不偵測) 值。全流程採 *split*：先以資料回歸獲得可微的  $E(t)$ ，再開啟物理殘差與事件條件以微調參數與隱變數。

## 1 基本模型與分次跳躍

連續段：

$$\dot{B} = rB\left(1 - \frac{B}{K}\right), \quad \dot{R} = \sigma_0 B - k_{12}R, \quad \dot{E} = k_{12}R - k_c E. \quad (1.1)$$

第  $j$  次分次於  $t = t_j$  (LQ 生存分率  $SF_j = \exp(-\alpha d_j - \beta d_j^2)$ )：

$$B(t_j^+) = B(t_j^-) SF_j, \quad (1.2)$$

$$R(t_j^+) = R(t_j^-) + \sigma_1 [B(t_j^-) - B(t_j^+)], \quad (1.3)$$

$$E(t_j^+) = E(t_j^-) \quad (\text{連續}) . \quad (1.4)$$

LQ 模型背景見 [1, 2]。

## 2 實作步驟

### Step 1 | 時間縮放與非負「硬約束」

**理由**：將時間縮放到  $[0, 1]$  使自動微分的尺度一致、訓練更穩定；對必須非負的狀態/參數用輸出變換比加懲罰穩定。

**方法**：令

$$t_s = \frac{t - t_0}{t_f - t_0} \in [0, 1], \quad (\underline{B}, \underline{R}, \underline{E}) \leftarrow \text{NN}(t_s), \quad (2.1)$$

並對需非負者在輸出層施加 square 或 softplus 變換。

依據 SIR-PINN 方法節 §2.2 明確以  $t_s$  無因次化 (式 (5)–(6))，DeepXDE 文獻系統性使用輸出變換以「硬」滿足條件並保持穩定 [3, 4, §2.2]。

檢核點。早期訓練不會出現負的  $\underline{B}, \underline{R}, \underline{E}$ ；殘差不因號誌錯誤而震盪。

## Step 2 | 把「分次跳躍」嵌入到解的構造式

理由：直接在解的構造式保證 式 (1.2) and (1.3) ( $\kappa \rightarrow \infty$  時精確)，避免以懲罰近似。

方法：用平滑階躍  $H_\kappa(\tau) = \sigma(\kappa\tau)$ ，設

$$M(t) = \exp\left(\sum_j (\log SF_j) H_\kappa(t - t_j)\right), \quad B(t) = M(t) \underline{B}(t), \quad (2.2)$$

$$\Delta B_j = B(t_j^-) - B(t_j^+) = M(t_j^-) \underline{B}(t_j^-) [1 - SF_j], \quad R(t) = \underline{R}(t) + \sigma_1 \sum_j \Delta B_j H_\kappa(t - t_j), \quad (2.3)$$

並令  $E(t)$  連續近似（不嵌入跳階）。依據 可微事件函數與混合流 - 跳系統（Neural ODEs with events §1.1；Neural Jump-SDE）提供將跳躍以可微觸發函數嵌入的範式；界面跳條件與 cPINN/XPINN 的介面強制思路完全相符 [5, 6, 7, 8]。

檢核點。隨  $\kappa$  退火增大， $t_j$  近旁圖像趨近垂直跳階。

### 平滑階躍函數 $H_\kappa(\tau) = \sigma(\kappa\tau)$ 的解釋

平滑階躍函數  $H_\kappa(\tau) = \sigma(\kappa\tau)$  是用來處理模型中「分次跳躍」(piecewise jumps) 的關鍵工具，特別嵌入解的構造式中（見 Step 2）。

#### 定義與數學形式

- 基本形式： $H_\kappa(\tau)$  是 sigmoid 激活函數  $\sigma(x) = \frac{1}{1+e^{-x}}$  的縮放版本，其中  $x = \kappa\tau$ 。
  - $\tau = t - t_j$ ：相對於第  $j$  次事件時間  $t_j$  的時間偏移（事件觸發後的「延遲」）。
  - $\kappa > 0$ ：陡峭度參數（steepness parameter），控制階躍的「銳利度」。
- 行為：
  - 當  $\tau < 0$ （事件前）， $H_\kappa(\tau) \approx 0$ （平滑過渡）。
  - 當  $\tau > 0$ （事件後）， $H_\kappa(\tau) \approx 1$ （平滑過渡）。
  - 轉換點在  $\tau = 0$ ，轉換寬度約  $\sim 1/\kappa$ 。
- 極限情況：當  $\kappa \rightarrow \infty$ ， $H_\kappa(\tau)$  逼近 Heaviside 階躍函數  $H(\tau)$ （理想步階： $H(\tau) = 0$  若  $\tau < 0$ ， $H(\tau) = 1$  若  $\tau > 0$ ），但保持可微分性，避免梯度中斷。

這是神經網路中常見的「軟階躍」(soft step) 逼近，借鑑自可微事件函數（Neural ODEs with events）與混合流-跳系統（Neural Jump-SDE）[5, 6]。

#### 為何需要平滑階躍？

- 問題：EBV 模型包含連續 ODE（式 (2.1)）與不連續事件（LQ 放療模型的分次跳躍，式 (2.2)–(2.4)），如  $B$  在  $t_j$  處突然衰減  $SF_j$  倍（生存分率）， $R$  增加  $\sigma_1 \Delta B_j$ 。
  - 直接使用理想階躍  $H(\tau)$  會在  $t_j$  產生不連續，導致 PINN 訓練時自動微分（autodiff）梯度爆炸或殘差尖峰。
  - 若用懲罰項（penalty）近似，會引入多損失失衡（multi-loss imbalance）。
- 解法：將跳躍「內建」到解的構造式中，使用平滑  $H_\kappa$  保證可微分性，允許端到端優化。依據 cPINN/XPINN 的介面強制思路 [7, 8]，這能精確滿足跳躍條件（ $\kappa \rightarrow \infty$  時），同時穩定訓練。

## 在模型中的應用

- 構造式嵌入 (式 (1.2) and (1.3)) :

$$M(t) = \exp \left( \sum_j (\log SF_j) H_\kappa(t - t_j) \right), \quad (2.4)$$

$$B(t) = M(t) \underline{B}(t), \quad (2.5)$$

$$\Delta B_j = B(t_j^-) - B(t_j^+) = M(t_j^-) \underline{B}(t_j^-) [1 - SF_j], \quad (2.6)$$

$$R(t) = \underline{R}(t) + \sigma_1 \sum_j \Delta B_j H_\kappa(t - t_j), \quad (2.7)$$

其中  $B(t)$  與  $R(t)$  是神經網路輸出， $M$  提供跳躍疊加， $E$  保持連續。

- 導數計算 (Step 3) : 使用鏈式法則確保可微：

$$\dot{H}_\kappa(\tau) = \kappa \sigma(\kappa \tau) [1 - \sigma(\kappa \tau)] = \kappa H_\kappa(\tau) (1 - H_\kappa(\tau)), \quad (2.8)$$

這是 sigmoid 的導數，峰值在轉換點，避免脈衝誤當 ODE 殘差。

- $\kappa$  退火 (annealing) (Step 7) : 訓練初期  $\kappa$  小 (平滑，易優化)；後期逐步增大 (每  $500\text{--}10^3$  epochs 乘 1.5)，逼近真跳躍，降低優化難度 [9, 10]。

## 優點與檢核

- **優點**：使不連續系統可微，穩定 PINN 收斂；保留 LQ 物理可解釋性；適用多事件 ( $30+$  次)。
- **檢核點**：隨  $\kappa$  增大， $t_j$  附近圖像趨近垂直跳階 (垂直線)。
- **文獻依據**：源自可微事件與界面條件 [5, 6, 7, 8]，在不連續 PINN 中加權平滑策略 [11]。

此設計是為讓 split 訓練 (Stage-1 資料、Stage-2 物理) 更穩定，尤其在 LOD 雜訊下。

## Step 3 | 事件遮罩與導數鏈式法則

**理由**：避免把事件脈衝誤當 ODE 殘差。

**方法**：每個  $t_j$  設平滑遮罩  $w(t) = \prod_j [1 - \exp(-\frac{(t-t_j)^2}{2\varepsilon^2})]$ ；自動微分時用

$$\dot{B} = \dot{M} \underline{B} + M \dot{\underline{B}}, \quad \dot{R} = \dot{\underline{R}} + \sigma_1 \sum_j \Delta B_j \dot{H}_\kappa(t - t_j) + \sigma_1 \sum_j \Delta \dot{B}_j H_\kappa(t - t_j),$$

其中  $\dot{H}_\kappa(\tau) = \kappa \sigma(\kappa \tau) (1 - \sigma(\kappa \tau))$ 。

**依據** 對不連續/銳層的 PINN 加權與平滑策略 (*Discontinuity Computing using PINNs*) 以及課程式學習可顯著穩定殘差 [11]。

檢核點。關掉遮罩  $w \equiv 1$  時， $t_j$  附近殘差尖峰；打開遮罩恢復平穩。

## Step 4 | 三塊 Loss (核心)

### 4.1 ODE 殘差 (事件外)

$$\mathcal{L}_{\text{ODE}} = \frac{1}{N_c} \sum_i w(t_i) \left[ (\dot{B} - rB(1 - \frac{B}{K}))^2 + (\dot{R} - (\sigma_0 B - k_{12} R))^2 + (\dot{E} - (k_{12} R - k_c E))^2 \right], \quad t_i \notin \cup_j (t_j \pm \varepsilon). \quad (2.9)$$

$K \gg B$  時可改第一項為  $(\dot{B} - rB)^2$ 。

### 4.2 事件點條件 (雙側取樣)

$$\mathcal{L}_{\text{jump}} = \sum_j |B(t_j^+) - SF_j B(t_j^-)|^2 + |R(t_j^+) - R(t_j^-) - \sigma_1 \Delta B_j|^2 + |E(t_j^+) - E(t_j^-)|^2. \quad (2.10)$$

**依據** 將界面 (事件點) 條件作為額外約束明確加入目標函數，與 cPINN/XPINN 於子域介面的作法一致 [7, 8, 介面條件]。

### 4.3 觀測似然 (左設限對數常態)

$$\mathcal{L}_{\text{data}} = \sum_{E_i > \text{LOD}} \left[ \frac{(\ln E_i - \ln E(t_i))^2}{2\sigma_{\text{obs}}^2} + \ln \sigma_{\text{obs}} + \ln E_i \right] + \sum_{E_i \leq \text{LOD}} \left[ -\ln \Phi \left( \frac{\ln \text{LOD} - \ln E(t_i)}{\sigma_{\text{obs}}} \right) \right], \quad (2.11)$$

其中  $\Phi$  為標準常態 CDF、 $\sigma_{\text{obs}} > 0$  可聯合學習。為何非 MSE? LOD 下的 N/D 並非 0，而是「未知且  $\leq \text{LOD}$ 」；以左設限似然能無偏地利用訊息 [12]。

## EBV 模型中 E (細胞外 EBV DNA 濃度) 存在 LOD 問題的參考文獻

在 EBV (Epstein-Barr Virus) 血漿動力學模型中，E 代表細胞外 EBV DNA 濃度，常見於 qPCR (定量聚合酶鏈反應) 測量，其數據經常受限於偵測極限 (LOD, Limit of Detection) 與量化下限 (LLoQ, Lower Limit of Quantification)，導致低濃度樣本為左設限 (left-censored) 數據，即標記為 N/D (Not Detected)，無法精確量化。此問題在臨床與研究中普遍存在，因為 EBV DNA 病毒載量在無症狀或低活性期常低於 LOD，需使用半 LLoQ 值或特殊似然處理以避免偏差。以下為相關參考文獻，證實此現象：

1. **Lazzarotto et al. (2018)**：在異基因造血幹細胞移植受者中，細胞巨病毒與 Epstein-Barr Virus DNA 在全血與血漿樣本的動態分析顯示，低於 LLoQ 的陽性樣本被設限 (censored) 為 LLoQ 的一半值，以處理量化不確定性。這直接證實 EBV DNA 血漿測量常有 LOD 問題，影響病毒載量估計 [13].
2. **Kanakry et al. (2016)**：在有或無 EBV 疾病的患者中，血漿與外周血單核細胞 EBV DNA 的臨床意義研究，強調 qPCR LOD 為 50 copies/mL，低於此值的樣本為不偵測，需統計調整以避免低估病毒活性 [14].
3. **Mayo Clinic Laboratories (EBVQN Assay)**：EBV DNA 血漿檢測的量化範圍為 35–100,000,000 IU/mL，LOD 為 19 IU/mL，低於此值無法可靠量化，常用於移植受者監測，凸顯 LOD 在 EBV 臨床應用中的普遍性 [15].

這些參考支持 EBV PDF 中使用左設限對數常態似然 (式 (2.11)) 處理 E 數據的必要性，確保無偏估計，尤其在含 N/D 值的 split 訓練中。

## 觀測似然（左設限對數常態）的解釋

在 PINN (Physics-Informed Neural Network) 模型中，資料損失  $L_{data}$  用來擬合觀測數據，特別是針對 EBV (Epstein-Barr Virus) 血漿濃度  $E(t_i)$  的測量值。EBV 數據常見的挑戰是「偵測極限」(Limit of Detection, LOD)，即當真實濃度低於儀器可靠偵測的閾值時，測量結果會被記錄為「不偵測」(N/D)，而非精確的 0 值。這類數據屬於**左設限 (left-censored) 數據**：我們知道真實值  $E_i \leq \text{LOD}$ ，但無法得知確切數值。若忽略此特性，直接用 MSE (均方誤差) 或將 N/D 視為 0 擬合，會導致嚴重偏差 (bias)，因為 N/D 值實際上攜帶「低於 LOD」的部分資訊。

### 為何使用左設限對數常態似然？

- **對數常態分佈 (Log-Normal Distribution) 的假設**：EBV 等生物濃度數據通常呈現對數常態分佈，即  $\ln E \sim \mathcal{N}(\mu, \sigma^2)$ ，其中  $\mu = \ln E(t_i)$  (模型預測的對數濃度)， $\sigma = \sigma_{obs}$  (觀測噪音標準差，可聯合學習)。這適合正偏態 (right-skewed) 的濃度數據，避免負值問題。
- **左設限處理**：標準似然無法直接用於設限數據，因此分為兩部分：
  - **偵測值 ( $E_i > \text{LOD}$ )**：使用完整對數似然 (negative log-likelihood, NLL)，充分利用精確測量。
  - **不偵測值 ( $E_i \leq \text{LOD}$ )**：使用累積分佈函數 (CDF) 計算設限機率，無偏地融入「低於 LOD」的資訊。
- **優點**：相較 MSE，這能無偏 (unbiased) 利用所有數據點，提高估計準確性與效率。依據環境統計標準 [12]，這是處理 LOD 的推薦方法，尤其在醫學/環境數據中。

### 公式拆解

損失函數  $L_{data}$  是負對數似然 (NLL)，最小化它等同最大化似然。公式為：

$$L_{data} = \sum_{E_i > \text{LOD}} \left[ \frac{(\ln E_i - \ln E(t_i))^2}{2\sigma_{obs}^2} + \ln \sigma_{obs} + \ln E_i \right] + \sum_{E_i \leq \text{LOD}} \left[ -\ln \Phi \left( \frac{\ln \text{LOD} - \ln E(t_i)}{\sigma_{obs}} \right) \right], \quad (2.12)$$

其中  $\Phi$  是標準常態累積分佈函數 (CDF)， $\sigma_{obs} > 0$  可作為可學習參數。

#### 1. 偵測值部分 ( $E_i > \text{LOD}$ )：

$$\frac{(\ln E_i - \ln E(t_i))^2}{2\sigma_{obs}^2} + \ln \sigma_{obs} + \ln E_i$$

- **解釋**：對數常態密度函數為  $f(y) = \frac{1}{y\sigma_{obs}\sqrt{2\pi}} \exp \left( -\frac{(\ln y - \mu)^2}{2\sigma_{obs}^2} \right)$ ，其負對數為  $\ln y + \ln \sigma_{obs} + \frac{1}{2} \ln(2\pi) + \frac{(\ln y - \mu)^2}{2\sigma_{obs}^2}$ 。

- $(\ln E_i - \ln E(t_i))^2 / (2\sigma_{obs}^2)$ ：測量與預測的對數差異平方，衡量擬合誤差。
- $\ln \sigma_{obs}$ ：懲罰過大噪音 (正則化)。
- $\ln E_i$ ：來自雅可比項 (Jacobian)，因為變數轉換到對數尺度。

- 效果：鼓勵模型  $E(t_i)$  接近觀測  $E_i$ ，同時考慮變異性。
2. 不偵測值部分 ( $E_i \leq \text{LOD}$ )：
- $$-\ln \Phi \left( \frac{\ln \text{LOD} - \ln E(t_i)}{\sigma_{obs}} \right)$$
- 解釋：對數常態的 CDF 為  $P(Y \leq \text{LOD}) = \Phi \left( \frac{\ln \text{LOD} - \mu}{\sigma_{obs}} \right)$ ，其中  $\mu = \ln E(t_i)$ 。
    - 這是設限似然的 NLL：最小化它等同最大化「真實值  $\leq \text{LOD}$ 」的機率。
    - 若模型預測  $E(t_i)$  遠大於 LOD，則  $\Phi$  接近 1， $-\ln \Phi \approx 0$ （低懲罰）；若  $E(t_i)$  過小， $\Phi$  接近 0，懲罰增大。
  - 效果：迫使模型在 N/D 點預測低濃度，但不強制為 0，避免偏差。標準化  $z = (\ln \text{LOD} - \ln E(t_i)) / \sigma_{obs}$  使之無因次。

### 實作注意

- $\sigma_{obs}$  的學習：初始設為小值（如 0.1），聯合優化以捕捉數據變異。
- 數值穩定：確保  $E(t_i) > 0$ （用 softplus 激活），避免  $\ln E(t_i)$  發散。若 LOD 固定，可視為常數。
- 為何優於 MSE：MSE 假設 N/D = 0，忽略「 $\leq \text{LOD}$ 」的分佈資訊，導致低估濃度變異；此似然則借鑑貝氏/最大似然估計，保留所有數據的統計力量。
- 文獻依據：[12] 詳細討論環境數據的設限統計，強調左設限似然在 R/Minitab 中的應用。

此設計使 PINN 在 EBV 數據的 split 訓練中（Stage-1 只用  $L_{data}$  擬合  $E(t)$ ），能產生平滑、可微的初始解，後續引入物理約束更穩定。

### Step 6 | 兩階段 (split) 訓練

理由：資料含 LOD 與事件不連續時，先用資料使  $E(t)$  平滑可微，再引入物理殘差與事件條件。

方法：

- Stage-1：只訓  $E(t)$  與  $\sigma_{obs}$ ，最小化  $\mathcal{L}_{data}$ 。
- Stage-2：解鎖  $B, R$  與參數  $r, K, \sigma_0, k_{12}, k_c, \sigma_1$ ，打開  $\mathcal{L}_{ODE} + \mathcal{L}_{jump}$  並保小權重的  $\mathcal{L}_{data}$ 。

依據 SIR-PINN 方法節 §2.2 明確討論 joint vs. split；課程式/分段時間訓練實證能縮時並穩定收斂 [3, 16]。

## References

- [1] J.F. Fowler. The linear-quadratic formula and progress in fractionated radiotherapy. *British Journal of Radiology*, 62(740):679–694, 1989.

- [2] M.C. Joiner and A.J. van der Kogel, editors. *Basic Clinical Radiobiology*. CRC Press, 5 edition, 2018.
- [3] C. Millevoi, D. Pasetto, and M. Ferronato. A physics-informed neural network approach for compartmental epidemiological models. *PLOS Computational Biology*, 20(9):e1012387, 2024.
- [4] L. Lu, X. Meng, Z. Mao, and G.E. Karniadakis. Deepxde: A deep learning library for solving differential equations. *SIAM Review*, 63(1):208–228, 2021.
- [5] R.T.Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud. Learning neural event functions for ordinary differential equations. In *Advances in Neural Information Processing Systems (NeurIPS) Workshop*, 2020. arXiv:2011.03902.
- [6] J. Jia and A.R. Benson. Neural jump stochastic differential equations. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.
- [7] A.D. Jagtap, E. Kharazmi, and G.E. Karniadakis. Conservative physics-informed neural networks on discrete domains for conservation laws. *Computer Methods in Applied Mechanics and Engineering*, 365:113028, 2020.
- [8] A.D. Jagtap and G.E. Karniadakis. Extended physics-informed neural networks (xpinns): A generalized space-time domain decomposition framework. *Communications in Computational Physics*, 28(5):2002–2041, 2020.
- [9] Y. Huang, H. Zheng, Y. Zhang, and H. Zhang. Homotopy physics-informed neural networks for inverse problems of nonlinear differential equations. *Journal of Computational Physics*, 491:112401, 2024. also arXiv:2304.02811.
- [10] C. Gulcehre, M. Moczulski, F. Visin, and Y. Bengio. Mollifying networks. In *ICLR Workshop*, 2017.
- [11] L. Liu, S. Chen, and L. Li. Discontinuity computing using physics-informed neural networks. *arXiv preprint*, 2022.
- [12] D.R. Helsel. *Statistics for Censored Environmental Data Using Minitab and R*. Wiley, 2 edition, 2012.
- [13] Tiziana Lazzarotto, Angela Chiereghin, Antonio Piralla, Giulia Piccirilli, Alessia Girello, Giulia Campanini, Liliana Gabrielli, Cristina Costa, Arcangelo Prete, Francesca Bonifazi, Alessandro Busca, Roberto Cairoli, Anna Amelia Colombo, Marco Zecca, Francesca Sidoti, Gabriele Bianco, Pierpaolo Paba, Carlo Federico Perno, Rossana Cavallo, Fausto Baldanti, and AMCLI-GLaIT working group. Cytomegalovirus and epstein-barr virus dna kinetics in whole blood and plasma of allogeneic hematopoietic stem cell transplantation recipients. *Biology of Blood and Marrow Transplantation*, 24(8):1699–1706, 2018.
- [14] Jennifer A. Kanakry, Aparna M. Hegde, Christine M. Durand, Allan B. Massie, Amy E. Greer, Richard F. Ambinder, and Alexandra Valsamakis. The clinical significance of ebv dna in the plasma and peripheral blood mononuclear cells of patients with or without ebv diseases. *Blood*, 127(16):2007–2017, 2016.

- [15] Epstein-barr virus dna detection and quantification, plasma (ebvqn). Quantification range 35–100,000,000 IU/mL; limit of detection 19 IU/mL.
- [16] Y.W. Bekele, S. Barthelmé, and M.E. Rognes. Physics-informed neural networks with curriculum training. *arXiv preprint*, 2024.