Two-Step Bayesian PINNs for Uncertainty Estimation

Pablo Flores
Pontificia Universidad Católica de Chile
ptflores1@uc.cl

Olga Graf
Technical University of Munich
graf@ma.tum.de

Pavlos Protopapas
Harvard University
pavlos@seas.harvard.edu

Karim Pichara
Pontificia Universidad Católica de Chile
kpb@ing.puc.cl

Abstract

We use a two-step procedure to train Bayesian neural networks that provide uncertainties over the solutions to differential equation systems provided by Physics-Informed Neural Networks (PINNs). We take advantage of available error bounds over PINNs to formulate a heteroscedastic variance that improves the uncertainty estimation. Furthermore, we solve forward problems and utilize the uncertainties obtained to improve parameter estimation in inverse problems in the fields of cosmology and fermentation.

1 Introduction

Physics-Informed Neural Networks (PINNs) are a type of neural network (NN) first proposed by [25] that are trained to solve a differential equation (DE) by incorporating the underlying physics of the problem into the network architecture. This allows the network to learn the DE solution without needing additional data. PINNs present many advantages over traditional numerical solvers; they are continuous, differentiable, and parallelizable, i.e., they do not need to compute previous time steps in order to compute the next one. Although these advantages are noteworthy, the task of computing solution errors or error bounds remains largely unresolved, with ongoing research in this area [28, 29, 15].

Since the first appearance of PINNs, there has been a rapid growth of publications that cover a wide range of applications. Heat transfer [6], wave equations [39] and fluid mechanics [23, 5, 30] are some examples. There has also been interest in identifying and overcoming failure modes and optimization challenges [36, 42, 24, 14]. Bayesian PINNs have also been explored [45, 27, 20].

Given the lack of advances for estimating errors, we leverage Bayesian NNs to obtain uncertainties over the solutions. The contributions of our work are the following:

- We propose an error bound based heteroscedastic variance that improves uncertainty quality.
- We solve forward problems for equations in the fields of cosmology and fermentation processes.
- Finally, we utilize solution bundles [17] to solve inverse problems, i.e., do parameter estimation over the equation parameters.

DLDE-III Workshop in the 37th Conference on Neural Information Processing Systems (NeurIPS 2023).

2 Background

2.1 Problem Formulation

We adopt a slightly different formulation from [38]. The DEs we will work with can be defined as follows:

$$\mathcal{F}_\lambda[u(\mathbf{x})] = f(\mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{1}$$

$$\mathcal{B}_\lambda[u(\mathbf{x})] = b(\mathbf{x}), \quad \mathbf{x} \in \Gamma, \tag{2}$$

where $\mathbf{x}$ is the space-time coordinate, Ω is a bounded domain with boundary Γ, $f(\mathbf{x})$ is the source term (also called forcing function), u is the solution of the system, $\mathcal{F}_\lambda$ is a differential operator, $\mathcal{B}_\lambda$ and $b(\mathbf{x})$ are the initial condition (IC) or boundary condition (BC) operator and term, respectively, and λ denotes the parameters of the system.

In this paper, we will focus on problems where the operators $\mathcal{F}_\lambda$, $\mathcal{B}_\lambda$ and terms $f(\mathbf{x})$, $b(\mathbf{x})$ are known. If λ is assumed to be known, the goal is to find the solution u, which is referred to as the *forward problem*. Conversely, if u is known and the aim is to estimate λ, then this is known as the *inverse problem*.

2.2 Physics-Informed Neural Networks

A *Physics-Informed Neural Network* uses a deep neural network, $u_\theta(\mathbf{x})$, to approximate the true solution $u(\mathbf{x})$. The network is trained to minimize the residual $r_\theta(\mathbf{x}) = \mathcal{F}_\lambda[u_\theta(\mathbf{x})] - f(\mathbf{x})$, aiming for $r_\theta(\mathbf{x}) = 0$.

The mean squared error is commonly used as the loss function. A transformation $\tilde{u}_\theta(\mathbf{x})$ ensures that initial conditions are met (we use the transformation shown in [20, 9]). The optimization problem is: $\min_\theta \sum_i^N \tilde{r}_\theta^2(\mathbf{x}_i)$.

Solution bundles [17] extend PINNs by allowing the network to take system parameters $\lambda \in \Lambda$ as inputs. This eliminates the need for multiple trainings for different λ values, and the optimization objective becomes: $\min_\theta \sum_i^N \sum_j^M \tilde{r}_\theta^2(\mathbf{x}_i, \lambda_j)$.

2.3 Uncertainty Quantification with Bayesian Neural Networks

To quantify uncertainty in NNs, we will employ the Bayesian approach. There are a number of different methods for posterior $p(\theta|\mathcal{D})$ calculation. In this paper, we will focus on the following three:

Bayes By Backpropagation (BBB) [3] employs variational inference to approximate the posterior. It assumes a variational distribution $q(\theta|\rho)$ and minimizes its Kullback-Leibler divergence from the true posterior.

Hamiltonian Monte Carlo (HMC) [16] draws samples from the posterior distribution using a Markov chain. Specifically, it is a Metropolis-Hastings algorithm [31, 21] that generates proposals by simulating the movement of a particle using Hamiltonian dynamics [37]. We will use the No-U-Turn Sampler (NUTS) [22] which is an extension of HMC.

Neural Linear Model (NLM) [41] is a Bayesian linear regression model with learned feature basis. Given specific priors and likelihood choices, analytical forms for both the posterior $p(\theta|\mathcal{D})$ and the posterior predictive $p(y|\mathbf{x})$ can be derived.

2.4 Error Bounds for PINNs

Good quality uncertainties should correlate with the true error of a solution. Since true error is not accessible, in Section 3.2 we use error bounds to improve the uncertainty in the Bayesian NNs.

In [28, 29], the authors present algorithms for calculating error bounds on PINNs. These bounds are applicable to linear ODEs, systems of linear ODEs, non-linear ODEs in the form ϵv^k[1], as well as certain types of PDEs. These algorithms are independent of the NN architecture and depend solely on the structure of the equation as defined in Eq. (1) and the residuals of the DE.

[1] Here, v is a variable and $|\epsilon| \ll 1$.

The network error is denoted as $\eta(\mathbf{x}) := u(\mathbf{x}) - \tilde{u}_\theta(\mathbf{x})$, and the error bound is represented by a scalar function $\mathbb{B}$ such that $\|\eta(\mathbf{x})\| \leq \mathbb{B}(\mathbf{x})$.

3 Two-Step Bayesian PINNs

We use a two-step approach to obtain uncertainties in the solutions of equations. In the first step, we train a PINN as a bundle network (Section 2.2) that we refer to as the *deterministic network*. We denote this network as $u_{\theta_{\text{det}}} : \Omega, \Lambda \to \mathbb{R}$.

3.1 Bayesian Neural Network Training

The second step involves training a Bayesian neural network using the outputs of $u_{\theta_{\text{det}}}$ as targets.

Dataset: We define a dataset by taking the space-time coordinates as independent variables, the equation parameters as additional input variables, and the outputs of the deterministic net as dependent variables $\mathcal{D} = \{(\mathbf{x}_i, \lambda_i, u_{\theta_{\text{det}}}(\mathbf{x}_i, \lambda_i)) \mid \mathbf{x}_i \in \Omega, \lambda_i \in \Lambda, u_{\theta_{\text{det}}}(\mathbf{x}_i, \lambda_i) \in \mathbb{R}\}_{i=1}^{N'}$.

Distributions: We define a Bayesian neural network u_θ by assigning distributions to the network parameters and dataset as it is done in [20]. For all three methods described in Section 2.3 we use the following distributions:

$$\theta \sim \mathcal{N}(0, \sigma_{\text{prior}}), \tag{3}$$

$$u_{\theta_{\text{det}}}(\hat{\mathbf{x}}, \hat{\lambda}) \mid \hat{\mathbf{x}}, \hat{\lambda}, \theta \sim \mathcal{N}(u_\theta(\hat{\mathbf{x}}, \hat{\lambda}), \sigma_{\text{Likelihood}}(\hat{\mathbf{x}}, \hat{\lambda})). \tag{4}$$

Given Eq. (4), we define the likelihood as $p(\mathcal{D}|\theta) = \prod_{i=1}^{N'} p(u_{\theta_{\text{det}}}(\mathbf{x}_i, \lambda_i) \mid \mathbf{x}_i, \lambda_i, \theta)$.

For BBB, we also define the variational posterior distribution using the mean field approximation [2].

For NLM, we can obtain the posterior and posterior predictive distributions through analytical derivation:

$$u \mid \mathbf{x}, \lambda, \mathcal{D} \sim \mathcal{N}(\Phi_\theta(\mathbf{x}, \lambda)\mu_{\text{post}}, \sigma^2_{\text{Likelihood}}(\mathbf{x}, \lambda) + \Phi_\theta(\mathbf{x}, \lambda)\Sigma_{\text{post}}\Phi_\theta^T(\mathbf{x}, \lambda)), \tag{5}$$

where μ_{post} and Σ_{post} are the posterior parameters and $\Phi_\theta(\mathbf{x}, \lambda)$ is the learned feature map, we provide details in Section VII.I.I of the supplementary material.

3.2 Heteroscedastic Variance Based on Error Bounds

Bayesian NNs are usually trained using homoscedastic variance, i.e., $\sigma_{\text{Likelihood}}(\mathbf{x}, \lambda) = \texttt{const}$. We take advantage of the error bounds described in Section 2.4 to define a heteroscedastic variance $\sigma_{\text{Likelihood}}(\mathbf{x}, \lambda) = \mathbb{B}(\mathbf{x}, \lambda)$.

Since the error bounds are not available for all equations, we use a slightly modified homoscedastic variance in the remaining equations; it is defined as:

$$\sigma_{\text{Likelihood}}(\mathbf{x}, \lambda) = \begin{cases} 0 & \mathbf{x} \in \Gamma, \\ \text{const} & \mathbf{x} \notin \Gamma, \end{cases} \tag{6}$$

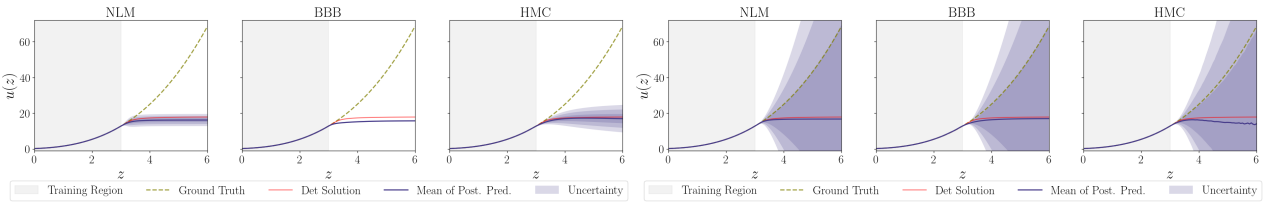
which means that we set to zero the variance at the coordinates of ICs. The reason for this is the enforcement of ICs, thus we are certain that the solution is correct at those coordinates.

3.3 Uncertainty Computation

The calculation of uncertainty is straightforward as it is the standard deviation of the posterior predictive distribution. We can obtain it by applying the law of total variance, its derivation can be found in Section VII.I.II of the supplementary material.

4 Parameter Estimation for Inverse Problems

We can use Bayesian framework to estimate the parameters λ of a DE system. For a given set of observations $\mathcal{O} = \{(\mathbf{x}_i, \mu_i, \sigma_i) \mid \mathbf{x}_i \in \Omega, \mu_i \in \mathbb{R}, \sigma_i \in \mathbb{R}^+\}_{i=1}^{O}$, we seek to find the posterior



(a) Homoscedastic variance (b) Error bounds based heteroscedastic variance

Figure 1: ΛCDM Bayesian solutions

distribution $p(\lambda|\mathcal{O})$ of the parameters. Here μ_i and σ_i are the mean and standard deviation of the true solution, respectively.

In this work, we assume a uniform prior over λ. To define the likelihood $p(\mathcal{O}|\lambda)$, we assume the observations have a normal distribution around the true solution, i.e., $\mu_i \sim \mathcal{N}(u(\mathbf{x}_i), \sigma_i)$. Since in Section 3 we obtain a distribution over all possible solutions, we need to marginalize over them to get the likelihood, the details of this marginalization are provided in Section VII.I.III of the supplementary material. Having defined the prior and likelihood, we can apply a sampling algorithm to approximate $p(\lambda|\mathcal{O})$. We use the emcee Python library [18] which implements the samplers introduced in [19].

5 Experiments and discussion

In this section we present the results we obtained for solving the forward problem using the standard PINN as described in Section 3 and inverse problem using bundle solutions and estimating the parameters as described in Sections 2.2 and 4, respectively.

Cosmology We address the same four cosmological models as in [8]. Details of the equations are given in Section VII.II.I of the Supplementary Material. In Fig. 1 we present the solutions we obtained using the three Bayesian methods described in Section 2.3. To solve this forward problem we select fixed values for the equation parameters.

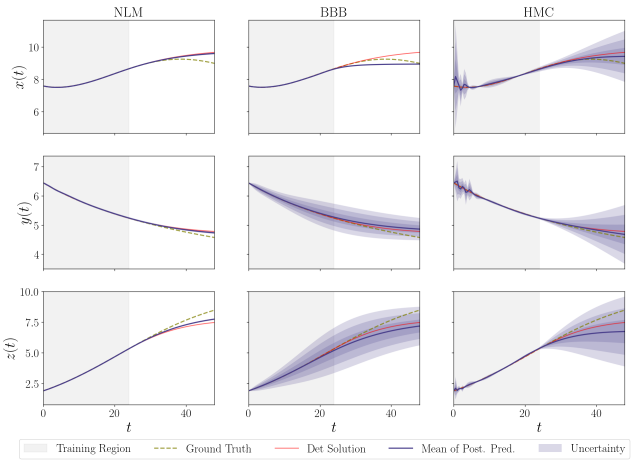


Figure 2: Bayesian solutions for LAB fermentation.

We estimated the parameters of the equation for the ΛCDM model as explained in Section 4. The ΛCDM model has only one parameter for which we obtained 10,000 samples for each Bayesian NN and also for a deterministic net. The means and standard deviations obtained were: 0.25 ± 0.025, 0.22 ± 0.0024, 0.20 ± 0.012, 0.1 ± 0.0001 for the deterministic net, NLM, BBB, and HMC, respectively.

Fermentation: We solve equations for Lactic Acid Bacteria (LAB) fermentation and cacao fermentation proposed in [44, 32], respectively. Details of the equations are given in Section VII.II.II of the Supplementary Material. In Fig. 2 we present the solutions to the forward problem for the LAB system.

6 Conclusions

We successfully applied a two-step procedure to train a Bayesian NN and proposed a heteroscedastic variance based on error bounds. We have shown that the use of this heteroscedastic variance effectively improves uncertainty quality on equations from cosmological and fermentation dynamical models. This emphasizes the need for methods that provide PINNs' error bounds for any DE class.

The results show that when using homoscedastic variance BBB heavily underestimates the uncertainty while NLM and HMC tend to increase their uncertainty as the solution moves away from the training region. We then utilized the solutions to estimate the parameter of the ΛCDM model and observed that Bayesian solutions yield a less disperse estimation than *deterministic nets*.

In future work we intend to do parameter estimation for the remaining DEs and analyze in depth the benefits of using well calibrated uncertainties.

References

[1] C. Armendariz-Picon, V. Mukhanov, and P. J. Steinhardt. Dynamical Solution to the Problem of a Small Cosmological Constant and Late-Time Cosmic Acceleration. *Physical Review Letters*, 85(21):4438–4441, Nov. 2000. doi: 10.1103/PhysRevLett.85.4438. URL https://link.aps.org/doi/10.1103/PhysRevLett.85.4438. Publisher: American Physical Society.

[2] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg, July 2006. ISBN 978-0-387-31073-2.

[3] C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra. Weight uncertainty in neural networks. *32nd International Conference on Machine Learning, ICML 2015*, 2:1613–1622, 2015. arXiv: 1505.05424 ISBN: 9781510810587.

[4] H. A. Buchdahl. Non-Linear Lagrangians and Cosmological Theory. *Monthly Notices of the Royal Astronomical Society*, 150(1):1–8, Sept. 1970. ISSN 0035-8711. doi: 10.1093/mnras/150.1.1. URL https://doi.org/10.1093/mnras/150.1.1.

[5] S. Cai, Z. Mao, Z. Wang, M. Yin, and G. E. Karniadakis. Physics-informed neural networks (PINNs) for fluid mechanics: a review. *Acta Mechanica Sinica*, 37(12):1727–1738, Dec. 2021. ISSN 1614-3116. doi: 10.1007/s10409-021-01148-1. URL https://doi.org/10.1007/s10409-021-01148-1.

[6] S. Cai, Z. Wang, S. Wang, P. Perdikaris, and G. E. Karniadakis. Physics-Informed Neural Networks for Heat Transfer Problems. *Journal of Heat Transfer*, 143(6), Apr. 2021. ISSN 0022-1481. doi: 10.1115/1.4050542. URL https://doi.org/10.1115/1.4050542.

[7] R. R. Caldwell, R. Dave, and P. J. Steinhardt. Cosmological Imprint of an Energy Component with General Equation of State. *Physical Review Letters*, 80(8):1582–1585, Feb. 1998. doi: 10.1103/PhysRevLett.80.1582. URL https://link.aps.org/doi/10.1103/PhysRevLett.80.1582. Publisher: American Physical Society.

[8] A. T. Chantada, S. J. Landau, P. Protopapas, C. G. Scóccola, and C. Garraffo. Cosmological informed neural networks to solve the background dynamics of the Universe, May 2022. URL http://arxiv.org/abs/2205.02945. arXiv:2205.02945 [astro-ph, physics:gr-qc, physics:hep-ph].

[9] F. Chen, D. Sondak, P. Protopapas, M. Mattheakis, S. Liu, D. Agarwal, and M. Di Giovanni. NeuroDiffEq: A Python package for solving differential equations with neural networks. *Journal of Open Source Software*, 5(46):1931, Feb. 2020. ISSN 2475-9066. doi: 10.21105/joss.01931. URL https://joss.theoj.org/papers/10.21105/joss.01931.

[10] M. Chevallier and D. Polarski. Accelerating universes with scaling dark matter. *International Journal of Modern Physics D*, 10(02):213–223, Apr. 2001. ISSN 0218-2718. doi: 10.1142/S0218271801000822. URL https://www.worldscientific.com/doi/abs/10.1142/S0218271801000822. Publisher: World Scientific Publishing Co.

[11] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis. Modified gravity and cosmology. *Physics Reports*, 513(1):1–189, Mar. 2012. ISSN 0370-1573. doi: 10.1016/j.physrep.2012.01.001. URL https://www.sciencedirect.com/science/article/pii/S0370157312000105.

[12] Z. Cong, Z. Han, Y. Shuo, L. Siqi, Z. Tong-Jie, and S. Yan-Chun. Four new observational H(z) data from luminous red galaxies in the Sloan Digital Sky Survey data release seven. *Research in Astronomy and Astrophysics*, 14(10):1221, Oct. 2014. ISSN 1674-4527. doi: 10.1088/1674-4527/14/10/002. URL https://dx.doi.org/10.1088/1674-4527/14/10/002.

[13] E. J. Copeland, A. R. Liddle, and D. Wands. Exponential potentials and cosmological scaling solutions. *Physical Review D*, 57(8):4686–4690, Apr. 1998. doi: 10.1103/PhysRevD.57.4686. URL https://link.aps.org/doi/10.1103/PhysRevD.57.4686. Publisher: American Physical Society.

[14] A. Daw, J. Bu, S. Wang, P. Perdikaris, and A. Karpatne. Mitigating Propagation Failures in PINNs using Evolutionary Sampling, Oct. 2022. URL http://arxiv.org/abs/2207.02338. arXiv:2207.02338 [cs].

[15] T. De Ryck and S. Mishra. Generic bounds on the approximation error for physics-informed (and) operator learning. May 2022. URL https://openreview.net/forum?id=bF4eYy3LTR9.

[16] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Physics Letters B*, 195(2):216–222, Sept. 1987. ISSN 0370-2693. doi: 10.1016/0370-2693(87)91197-X. URL https://www.sciencedirect.com/science/article/pii/037026938791197X.

[17] C. Flamant, P. Protopapas, and D. Sondak. Solving Differential Equations Using Neural Network Solution Bundles, June 2020. URL https://arxiv.org/abs/2006.14372v1.

[18] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306, Feb. 2013. ISSN 1538-3873. doi: 10.1086/670067. URL https://iopscience.iop.org/article/10.1086/670067/meta. Publisher: IOP Publishing.

[19] J. Goodman and J. Weare. Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80, Jan. 2010. ISSN 2157-5452. doi: 10.2140/camcos.2010.5.65. URL https://msp.org/camcos/2010/5-1/p04.xhtml. Publisher: Mathematical Sciences Publishers.

[20] O. Graf, P. Flores, P. Protopapas, and K. Pichara. Uncertainty Quantification in Neural Differential Equations, Nov. 2021. URL http://arxiv.org/abs/2111.04207. arXiv:2111.04207 [cs].

[21] W. K. Hastings. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57(1):97–109, 1970. ISSN 0006-3444. doi: 10.2307/2334940. URL https://www.jstor.org/stable/2334940. Publisher: [Oxford University Press, Biometrika Trust].

[22] M. D. Hoffman and A. Gelman. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.

[23] X. Jin, S. Cai, H. Li, and G. E. Karniadakis. NSFnets (Navier-Stokes flow nets): Physics-informed neural networks for the incompressible Navier-Stokes equations. *Journal of Computational Physics*, 426:109951, Feb. 2021. ISSN 0021-9991. doi: 10.1016/j.jcp.2020.109951. URL https://www.sciencedirect.com/science/article/pii/S0021999120307257.

[24] A. Krishnapriyan, A. Gholami, S. Zhe, R. Kirby, and M. W. Mahoney. Characterizing possible failure modes in physics-informed neural networks. In *Advances in Neural Information Processing Systems*, volume 34, pages 26548–26560. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper/2021/hash/df438e5206f31600e6ae4af72f2725f1-Abstract.html.

[25] I. E. Lagaris, A. Likas, and D. I. Fotiadis. Artificial Neural Networks for Solving Ordinary and Partial Differential Equations. *IEEE Transactions on Neural Networks*, 9(5):987–1000, May 1997. doi: 10.1109/72.712178. URL http://arxiv.org/abs/physics/9705023. arXiv:physics/9705023v1.

[26] E. V. Linder. Exploring the Expansion History of the Universe. *Physical Review Letters*, 90(9):091301, Mar. 2003. doi: 10.1103/PhysRevLett.90.091301. URL https://link.aps.org/doi/10.1103/PhysRevLett.90.091301. Publisher: American Physical Society.

[27] K. Linka, A. Schafer, X. Meng, Z. Zou, G. E. Karniadakis, and E. Kuhl. Bayesian Physics-Informed Neural Networks for real-world nonlinear dynamical systems, May 2022. URL http://arxiv.org/abs/2205.08304. arXiv:2205.08304 [nlin].

[28] S. Liu, X. Huang, and P. Protopapas. Evaluating Error Bound for Physics-Informed Neural Networks on Linear Dynamical Systems, July 2022. URL http://arxiv.org/abs/2207.01114. arXiv:2207.01114 [cs, math].

[29] S. Liu, X. Huang, and P. Protopapas. *Residual-based error bound for physics-informed neural networks*. June 2023.

[30] Z. Mao, A. D. Jagtap, and G. E. Karniadakis. Physics-informed neural networks for high-speed flows. *Computer Methods in Applied Mechanics and Engineering*, 360:112789, Mar. 2020. ISSN 0045-7825. doi: 10.1016/j.cma.2019.112789. URL https://www.sciencedirect.com/science/article/pii/S0045782519306814.

[31] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6): 1087–1092, Dec. 2004. ISSN 0021-9606. doi: 10.1063/1.1699114. URL https://doi.org/10.1063/1.1699114.

[32] M. Moreno-Zambrano, S. Grimbs, M. S. Ullrich, and M.-T. Hütt. A mathematical model of cocoa bean fermentation. *Royal Society Open Science*, 5(10):180964, Oct. 2018. ISSN 2054-5703. doi: 10.1098/rsos.180964. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6227950/.

[33] M. Moresco. Raising the bar: new constraints on the Hubble parameter with cosmic chronometers at z 2. *Monthly Notices of the Royal Astronomical Society: Letters*, 450 (1):L16–L20, June 2015. ISSN 1745-3925. doi: 10.1093/mnrasl/slv037. URL https://doi.org/10.1093/mnrasl/slv037.

[34] M. Moresco, A. Cimatti, R. Jimenez, L. Pozzetti, G. Zamorani, M. Bolzonella, J. Dunlop, F. Lamareille, M. Mignoli, H. Pearce, P. Rosati, D. Stern, L. Verde, E. Zucca, C. M. Carollo, T. Contini, J.-P. Kneib, O. L. Fèvre, S. J. Lilly, V. Mainieri, A. Renzini, M. Scodeggio, I. Balestra, R. Gobat, R. McLure, S. Bardelli, A. Bongiorno, K. Caputi, O. Cucciati, S. d. l. Torre, L. d. Ravel, P. Franzetti, B. Garilli, A. Iovino, P. Kampczyk, C. Knobel, K. Kovač, J.-F. L. Borgne, V. L. Brun, C. Maier, R. Pelló, Y. Peng, E. Perez-Montero, V. Presotto, J. D. Silverman, M. Tanaka, L. A. M. Tasca, L. Tresse, D. Vergani, O. Almaini, L. Barnes, R. Bordoloi, E. Bradshaw, A. Cappi, R. Chuter, M. Cirasuolo, G. Coppa, C. Diener, S. Foucaud, W. Hartley, M. Kamionkowski, A. M. Koekemoer, C. López-Sanjuan, H. J. McCracken, P. Nair, P. Oesch, A. Stanford, and N. Welikala. Improved constraints on the expansion rate of the Universe up to z 1.1 from the spectroscopic evolution of cosmic chronometers. *Journal of Cosmology and Astroparticle Physics*, 2012(08):006, Aug. 2012. ISSN 1475-7516. doi: 10.1088/1475-7516/2012/08/006. URL https://dx.doi.org/10.1088/1475-7516/2012/08/006.

[35] M. Moresco, L. Pozzetti, A. Cimatti, R. Jimenez, C. Maraston, L. Verde, D. Thomas, A. Citro, R. Tojeiro, and D. Wilkinson. A 6% measurement of the Hubble parameter at z0.45: direct evidence of the epoch of cosmic re-acceleration. *Journal of Cosmology and Astroparticle Physics*, 2016(05):014, May 2016. ISSN 1475-7516. doi: 10.1088/1475-7516/2016/05/014. URL https://dx.doi.org/10.1088/1475-7516/2016/05/014.

[36] M. A. Nabian, R. J. Gladstone, and H. Meidani. Efficient training of physics-informed neural networks via importance sampling. *Computer-Aided Civil and Infrastructure Engineering*, 36(8):962–977, 2021. ISSN 1467-8667. doi: 10.1111/mice.12685. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/mice.12685. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mice.12685.

[37] R. M. Neal. Monte Carlo Implementation. In R. M. Neal, editor, *Bayesian Learning for Neural Networks*, Lecture Notes in Statistics, pages 55–98. Springer, New York, NY, 1996. ISBN 978-1-4612-0745-0. doi: 10.1007/978-1-4612-0745-0_3. URL https://doi.org/10.1007/978-1-4612-0745-0_3.

[38] A. F. Psaros, X. Meng, Z. Zou, L. Guo, and G. E. Karniadakis. Uncertainty Quantification in Scientific Machine Learning: Methods, Metrics, and Comparisons. 2022. doi: 10.48550/arxiv.2201.07766. arXiv: 2201.07766.

[39] M. Rasht-Behesht, C. Huber, K. Shukla, and G. E. Karniadakis. Physics-Informed Neural Networks (PINNs) for Wave Propagation and Full Waveform Inversions. *Journal of Geophysical Research: Solid Earth*, 127(5):e2021JB023120, 2022. ISSN 2169-9356. doi: 10.1029/2021JB023120. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2021JB023120. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2021JB023120.

[40] J. Simon, L. Verde, and R. Jimenez. Constraints on the redshift dependence of the dark energy potential. *Physical Review D*, 71(12):123001, June 2005. doi: 10.1103/PhysRevD.71.123001. URL https://link.aps.org/doi/10.1103/PhysRevD.71.123001. Publisher: American Physical Society.

[41] J. Snoek, O. Rippel, K. Swersky, R. Kiros, N. Satish, N. Sundaram, M. M. A. Patwary, Prabhat, and R. P. Adams. Scalable Bayesian Optimization Using Deep Neural Networks. Feb. 2015. URL https://www.semanticscholar.org/paper/Scalable-Bayesian-Optimization-Using-Deep-Neural-Snoek-Rippel/93bc65d2842b8cc5f3cf72ebc5b8f75daeacea35.

[42] S. Steger, F. M. Rohrhofer, and B. C. Geiger. How PINNs cheat: Predicting chaotic motion of a double pendulum. Oct. 2022. URL https://openreview.net/forum?id=shUbBca03f.

[43] D. Stern, R. Jimenez, L. Verde, M. Kamionkowski, and S. A. Stanford. Cosmic chronometers: constraining the equation of state of dark energy. I: H(z) measurements. *Journal of Cosmology and Astroparticle Physics*, 2010(02):008, Feb. 2010. ISSN 1475-7516. doi: 10.1088/1475-7516/2010/02/008. URL https://dx.doi.org/10.1088/1475-7516/2010/02/008.

[44] E. E. R. Talavera, D. P. Antonio, V. Trujillo, and D. Y. S. Muñoz. Differential equations system to describe bacterial growth, pH variation and lactic acid production in batch fermentation.

[45] L. Yang, X. Meng, and G. E. Karniadakis. B-PINNs: Bayesian Physics-Informed Neural Networks for Forward and Inverse PDE Problems with Noisy Data. Mar. 2020. doi: 10.1016/j.jcp.2020.109913. URL http://arxiv.org/abs/2003.06097. arXiv: 2003.06097.

[46] I. Zlatev, L. Wang, and P. J. Steinhardt. Quintessence, Cosmic Coincidence, and the Cosmological Constant. *Physical Review Letters*, 82(5):896–899, Feb. 1999. doi: 10.1103/PhysRevLett.82.896. URL https://link.aps.org/doi/10.1103/PhysRevLett.82.896. Publisher: American Physical Society.

VII Supplementary Material

VII.I Bayesian Methods Details

VII.I.I NLM Posterior Predictive Parameters

$$\Sigma_{\text{post}} = \left(\Phi_\theta^T(x,\lambda)\Sigma\Phi_\theta(x,\lambda) + \sigma^2_{\text{prior}} I\right) \tag{7}$$

$$\mu_{\text{post}} = \Sigma_{\text{post}}\left(\Sigma^{-1}\Phi_\theta(x,\lambda)\right)^T u_{\theta_{\text{det}}}(x_{\mathcal{D}}, \lambda_{\mathcal{D}}) \tag{8}$$

VII.I.II Law of Total Variance Derivation

$$\begin{aligned}
\text{Var}(u|x,\lambda,\mathcal{D}) &= \mathbb{E}_{\theta|\mathcal{D}}\left[\text{Var}(u|x,\lambda,\theta)\right] + \text{Var}_{\theta|\mathcal{D}}\left[\mathbb{E}(u|x,\lambda,\theta)\right] \\
&= \mathbb{E}_{\theta|\mathcal{D}}\left[\sigma^2_{\text{Likelihood}}(x,\lambda)\right] + \text{Var}_{\theta|\mathcal{D}}[u_\theta(x,\lambda)] \\
&\approx \sigma^2_{\text{Likelihood}}(x,\lambda) + \frac{1}{M}\sum_{i=1}^{M}(u_{\theta_i}(x,\lambda) - \overline{u(x,\lambda)})^2,
\end{aligned} \tag{9}$$

where $\theta_i \sim p(\theta|\mathcal{D})$ and $\overline{u(x,\lambda)} \approx \frac{1}{M}\sum_{i=1}^{M} u_{\theta_i}(x,\lambda)$ is the sample mean of the net outputs.

The latter approximation applies to BBB and NUTS, for BBB the samples are taken from the variational posterior and for NUTS we use the samples of the true posterior the method yields. For NLM we have the analytical expression in Eq. (7).

VII.I.III Marginalization Over Solutions

$$\begin{aligned}
p(x_i,\mu_i,\sigma_i|\lambda) &= \int_{\mathcal{U}} \mathcal{N}(\mu_i; u(x_i,\lambda),\sigma_i)\cdot p(u(x_i,\lambda)|x_i,\lambda,\mathcal{D})du \\
&\approx \frac{1}{M}\sum_{j=1}^{M}\mathcal{N}(\mu_i; u^{(j)}(x_i,\lambda),\sigma_i), \text{ where } u^{(j)} \sim p(u(x_i,\lambda)|x_i,\lambda,\mathcal{D}),
\end{aligned} \tag{10}$$

$$p(\mathcal{O}|\lambda) = \prod_{i=1}^{O} p(x_i,\mu_i,\sigma_i|\lambda). \tag{11}$$

VII.II Equation details

We now describe the equations, initial conditions and parameters of each dynamical model.

VII.II.I Cosmology

In this section we specify the cosmological equations we used. These equations were solved with PINNs in [8].

ΛCDM

$$\textbf{Equation:} \quad \frac{dx_m}{dz} = \frac{3x_m}{1+z} \tag{12}$$

$$\textbf{Initial conditions:} \quad x_m(z) = \Omega_{m,0} \tag{13}$$

$$\textbf{Parameters:} \quad \Omega_{m,0} \tag{14}$$

$$\textbf{Hubble function:} \quad H(z) = H_0\sqrt{x_m(z) + 1 - \Omega_{m,0}} \tag{15}$$

Parametric dark energy [26, 10]

Equation:
$$\frac{dx_{DE}}{dz} = \frac{3x_{DE}}{1+z}\left(1 + w_0 + \frac{w_1 z}{1+z}\right) \tag{16}$$

Initial conditions:
$$x_m(z) = 1 - \Omega_{m,0} \tag{17}$$

Parameters:
$$w_0, w_1, \Omega_{m,0} \tag{18}$$

Hubble function:
$$H(z) = H_0\sqrt{\Omega_{m,0}(1+z)^3 + x_{DE}(z)} \tag{19}$$

Quintessence [7, 1, 13, 46]

Equation:
$$\frac{dx}{dN} = -3x + \frac{\sqrt{6}}{2}\lambda y^2 + \frac{3}{2}x(1 + x^2 - y^2) \tag{20}$$

$$\frac{dy}{dN} = \frac{\sqrt{6}}{2}xy\lambda + \frac{3}{2}y(1 + x^2 - y^2) \tag{21}$$

Initial conditions:
$$x_0 = 0 \tag{22}$$

$$y_0 = \sqrt{\frac{1 - \Omega^{\Lambda}_{m,0}}{\Omega^{\Lambda}_{m,0}(1+z_0)^3 + 1 - \Omega^{\Lambda}_{m,0}}} \tag{23}$$

Parameters:
$$H_0, \Omega^{\Lambda}_{m,0} \tag{24}$$

Hubble function:
$$H(z) = H^{\Lambda}_0\sqrt{\frac{\Omega^{\Lambda}_{m,0}(1+z)^3}{1 - x^2 - y^2}} \tag{25}$$

$f(R)$ **gravity** [11, 4]

Equation:
$$\frac{dx}{dz} = \frac{1}{z+1}(-\Omega - 2v + x + 4y + xv + x^2) \tag{26}$$

$$\frac{dy}{dz} = \frac{-1}{z+1}(vx\Gamma - xy + 4y - 2yv) \tag{27}$$

$$\frac{dv}{dz} = \frac{-v}{z+1}(x\Gamma + 4 - 2v) \tag{28}$$

$$\frac{d\Omega}{dz} = \frac{\Omega}{z+1}(-1 + 2v + x) \tag{29}$$

$$\frac{dr}{dz} = -\frac{r\Gamma x}{z+1} \tag{30}$$

$$\text{where } \Gamma(r) = \frac{(r+b)\left[(r+b)^2 - 2b\right]}{4br} \tag{31}$$

Initial conditions:
$$x_0 = 0 \tag{32}$$

$$y_0 = \frac{\Omega^{\Lambda}_{m,0}(1+z_0)^3 + 2(1 - \Omega^{\Lambda}_{m,0})}{2\left[\Omega^{\Lambda}_{m,0}(1+z_0)^3 + (1 - \Omega^{\Lambda}_{m,0})\right]} \tag{33}$$

$$v_0 = \frac{\Omega^{\Lambda}_{m,0}(1+z_0)^3 + 4(1 - \Omega^{\Lambda}_{m,0})}{2\left[\Omega^{\Lambda}_{m,0}(1+z_0)^3 + (1 - \Omega^{\Lambda}_{m,0})\right]} \tag{34}$$

$$\Omega_0 = \frac{\Omega^{\Lambda}_{m,0}(1+z_0)^3}{\Omega^{\Lambda}_{m,0}(1+z_0)^3 + (1 - \Omega^{\Lambda}_{m,0})} \tag{35}$$

$$r_0 = \frac{\Omega^{\Lambda}_{m,0}(1+z_0)^3 + 4(1 - \Omega^{\Lambda}_{m,0})}{1 - \Omega^{\Lambda}_{m,0}} \tag{36}$$

Parameters:
$$b, \Omega^{\Lambda}_{m,0}, H^{\Lambda}_0 \tag{37}$$

Hubble function:
$$H(z) = H^{\Lambda}_0\sqrt{\frac{r}{2v}(1 - \Omega^{\Lambda}_{m,0})} \tag{38}$$

VII.II.II Fermentation

Lactic acid bacteria This system was proposed in [44].

Equations:

$$\frac{dx}{dt} = \alpha x - \beta x^2 - \gamma x y^2 - \tau x z^2 \tag{39}$$

$$\frac{dy}{dt} = \delta y - \theta y^2 - \rho y z \tag{40}$$

$$\frac{dz}{dt} = \sigma z - \phi z^2 - \omega x z \tag{41}$$

Initial conditions:

$$(x,y,z) = (7.5797, 6.44, 1.9) \tag{42}$$

Parameters:

$$\alpha,\beta,\gamma,\tau,\delta,\theta,\rho,\sigma,\phi,\omega \tag{43}$$

(44)

Cocoa This system was proposed in [32]. We present the meaning and initial conditions of each variable in Table 1 and the specification of the system's parameters in Table 2.

Equations:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\mathcal{Y}_{\mathrm{Glc|Y}} \frac{\mu_{\max}^{Y_{\mathrm{Glc}}} x}{x + K_{\mathrm{Glc}}^{\mathrm{Y}}} w - \mathcal{Y}_{\mathrm{Glc|LAB}} \frac{\mu_{\max}^{\mathrm{LAB}} x}{x + K_{\mathrm{Glc}}^{\mathrm{LAB}}} r \tag{45}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -\mathcal{Y}_{\mathrm{Fru|Y}} \frac{\mu_{\max}^{Y_{\mathrm{Fru}}} y}{y + K_{\mathrm{Fru}}^{Y}} w \tag{46}$$

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \mathcal{Y}_{\mathrm{EtOH|Y}}^{\mathrm{Glc}} \frac{\mu_{\max}^{Y_{\mathrm{Glc}}} x}{x + K_{\mathrm{Glc}}^{\mathrm{Y}}} w + \mathcal{Y}_{\mathrm{EtOH|Y}}^{\mathrm{Fru}} \frac{\mu_{\max}^{Y_{\mathrm{Fru}}} y}{y + K_{\mathrm{Fru}}^{Y}} w - \mathcal{Y}_{\mathrm{EtOH|AAB}} \frac{\mu_{\max}^{\mathrm{AAB_{EtOH}}} z}{z + K_{\mathrm{EtOH}}^{\mathrm{AAB}}} s \tag{47}$$

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{Y}_{\mathrm{LA|LAB}} \frac{\mu_{\max}^{\mathrm{LAB}} x}{x + K_{\mathrm{Glc}}^{\mathrm{LAB}}} r - \mathcal{Y}_{\mathrm{LA|AAB}} \frac{\mu_{\max}^{\mathrm{AAB_{LA}}} u}{u + K_{\mathrm{LA}}^{\mathrm{AAB}} s} s \tag{48}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \mathcal{Y}_{\mathrm{Ac|LAB}} \frac{\mu_{\max}^{\mathrm{LAB}} x}{x + K_{\mathrm{Glc}}^{\mathrm{LAB}}} r + \mathcal{Y}_{\mathrm{Ac|AAB}}^{\mathrm{EtOH}} \frac{\mu_{\max}^{\mathrm{AAB_{EtOH}}} z}{z + K_{\mathrm{EtOH}}^{\mathrm{AAB}}} s + \mathcal{Y}_{\mathrm{Ac|AAB}}^{\mathrm{LA}} \frac{\mu_{\max}^{\mathrm{AAB_{LA}}} u}{u + K_{\mathrm{LA}}^{\mathrm{AAB}} s} s \tag{49}$$

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\mu_{\max}^{Y_{\mathrm{Glc}}} x}{x + K_{\mathrm{Glc}}^{\mathrm{Y}}} w + \frac{\mu_{\max}^{Y_{\mathrm{Fru}}} y}{y + K_{\mathrm{Fru}}^{Y}} w - k_Y w z \tag{50}$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{\mu_{\max}^{\mathrm{LAB}} x}{x + K_{\mathrm{Glc}}^{\mathrm{LAB}}} r - k_{LAB} r u \tag{51}$$

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{\mu_{\max}^{\mathrm{AAB_{EtOH}}} z}{z + K_{\mathrm{EtOH}}^{\mathrm{AAB}}} s + \frac{\mu_{\max}^{\mathrm{AAB_{LA}}} u}{u + K_{\mathrm{LA}}^{\mathrm{AAB}} s} s - k_{\mathrm{AAB}}\, s v^2 \tag{52}$$

Table 1: Cocoa fermentation process variables meaning and initial conditions

Variable	Meaning	Initial Condition
x	[Glc]	51.963
y	[Fru]	57.741
z	[EtOH]	0
u	[LA]	0
v	[Ac]	0
w	[Y]	0.029180401
r	[LAB]	0.007868827
s	[AAB]	3.36634e-6

Table 2: Cocoa fermentation process parameters

Parameter	Unit	Interpretation
$\mu_{\max}^{Y_{Glc}}$	h^{-1}	Maximum specific growth rate of Y on Glc
$\mu_{\max}^{Y_{Fru}}$	h^{-1}	Maximum specific growth rate of Y on Fru
$\mu_{\max}^{LAB}$	h^{-1}	Maximum specific growth rate of LAB on Glc
$\mu_{\max}^{AAB_{EtOH}}$	h^{-1}	Maximum specific growth rate of AAB on EtOH
$\mu_{\max}^{AAB_{LA}}$	h^{-1}	Maximum specific growth rate of AAB on LA
K_{Glc}^{Y}	mg(Glc) g(pulp)$^{-1}$	Substrate saturation constant of Y growth on Glc
K_{Fru}^{Y}	mg(Fru) g(pulp)$^{-1}$	Substrate saturation constant of Y growth on Fru
K_{Glc}^{LAB}	mg(Glc) g(pulp)$^{-1}$	Substrate saturation constant of LAB growth on Glc
K_{EtOH}^{AAB}	mg(EtOH) g(pulp)$^{-1}$	Substrate saturation constant of AAB growth on EtOH
K_{LA}^{AAB}	mg(LA) g(pulp)$^{-1}$	Substrate saturation constant of AAB growth on LA
k_Y	mg(EtOH)$^{-1}$ h^{-1}	Mortality rate constant of Y
k_{LAB}	mg(LA)$^{-1}$ h^{-1}	Mortality rate constant of LAB
k_{AAB}	mg(Ac)$^{-2}$ h^{-1}	Mortality rate constant of AAB
$Y_{Glc\|Y}$	mg(Glc) mg(Y)$^{-1}$	Y-to-Glc yield coefficient
$Y_{Glc\|LAB}$	mg(Glc) mg(LAB)$^{-1}$	LAB-to-Glc yield coefficient
$Y_{Fru\|Y}$	mg(Fru) mg(Y)$^{-1}$	Y-to-Fru yield coefficient
$Y_{EtOH\|Y}^{Glc}$	mg(EtOH) mg(Y)$^{-1}$	Y-to-EtOH from Glc yield coefficient
$Y_{EtOH\|Y}^{Fru}$	mg(EtOH) mg(Y)$^{-1}$	Y-to-EtOH from Fru yield coefficient
$Y_{EtOH\|AAB}$	mg(EtOH) mg(AAB)$^{-1}$	AAB-to-EtOH yield coefficient
$Y_{LA\|LAB}$	mg(LA) mg(LAB)$^{-1}$	LAB-to-LA yield coefficient
$Y_{LA\|AAB}$	mg(LA) mg(AAB)$^{-1}$	AAB-to-LA yield coefficient
$Y_{Ac\|LAB}$	mg(Ac) mg(LAB)$^{-1}$	LAB-to-Ac yield coefficient
$Y_{Ac\|AAB}^{EtOH}$	mg(Ac) mg(AAB)$^{-1}$	AAB-to-Ac from EtOH yield coefficient
$Y_{Ac\|AAB}^{LA}$	mg(Ac) mg(AAB)$^{-1}$	AAB-to-Ac from LA yield coefficient

Table 3: Measurements of the Hubble parameter H using the Cosmic Chronometers technique

z	$H(z) \pm \sigma_H \left[\frac{\text{km/s}}{\text{Mpc}}\right]$	Ref.
0.09	69 ± 12	[40]
0.17	83 ± 8	
0.27	77 ± 14	
0.4	95 ± 17	
0.9	117 ± 23	
1.3	168 ± 17	
1.43	177 ± 18	
1.53	140 ± 14	
1.75	202 ± 40	
0.48	97 ± 62	[43]
0.88	90 ± 40	
0.1791	75 ± 4	[34]
0.1993	75 ± 5	
0.3519	83 ± 14	
0.5929	104 ± 13	
0.6797	92 ± 8	
0.7812	105 ± 12	
0.8754	125 ± 17	
1.037	154 ± 20	
0.07	69 ± 19.6	[12]
0.12	68.6 ± 26.2	
0.2	72.9 ± 29.6	
0.28	88.8 ± 36.6	
1.363	160 ± 33.6	[33]
1.965	186.5 ± 50.4	
0.3802	83 ± 13.5	[35]
0.4004	77 ± 10.2	
0.4247	87.1 ± 11.2	
0.4497	92.8 ± 12.9	
0.4783	80.9 ± 9	

VII.III Results

We present the results for the rest of the equations in Fig. 3, Fig. 4,Fig. 5, Fig. 6 for the cosmological models Dark Parametric Energy with homscedastic and heteroscedastic variance, Quintessence and Hu-Sawicky, respectively. And for the Cocoa fermentation model in Fig. 7

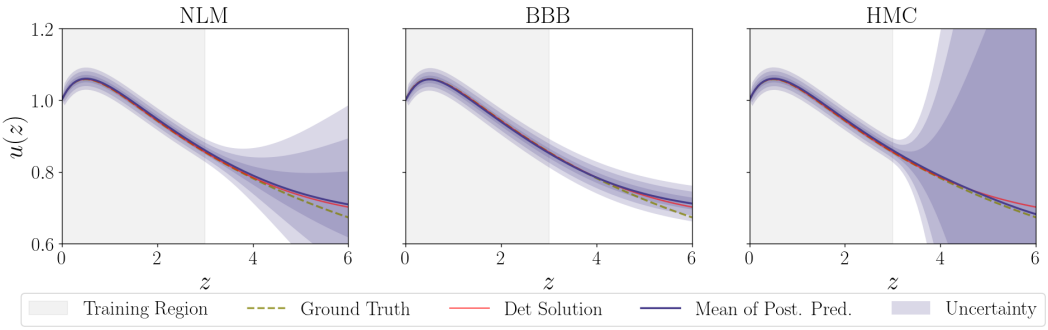


Figure 3: Dark Parametric Energy with homoscedastic variance solutions

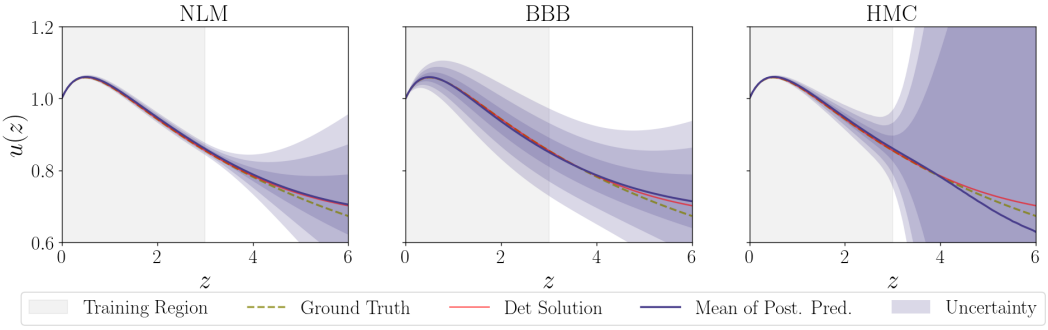


Figure 4: Dark Parametric Energy with heteroscedastic variance solutions

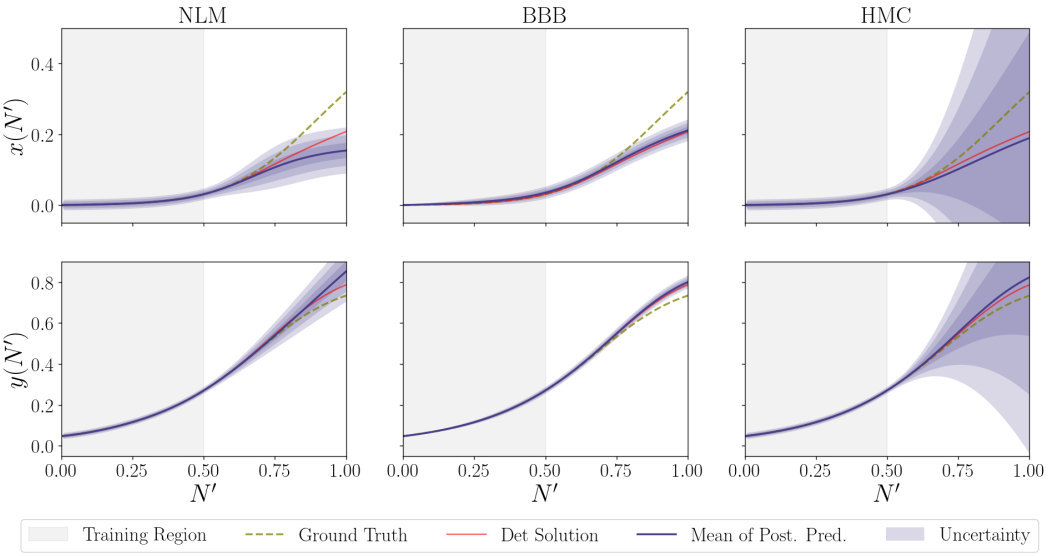


Figure 5: Quintessence solutions

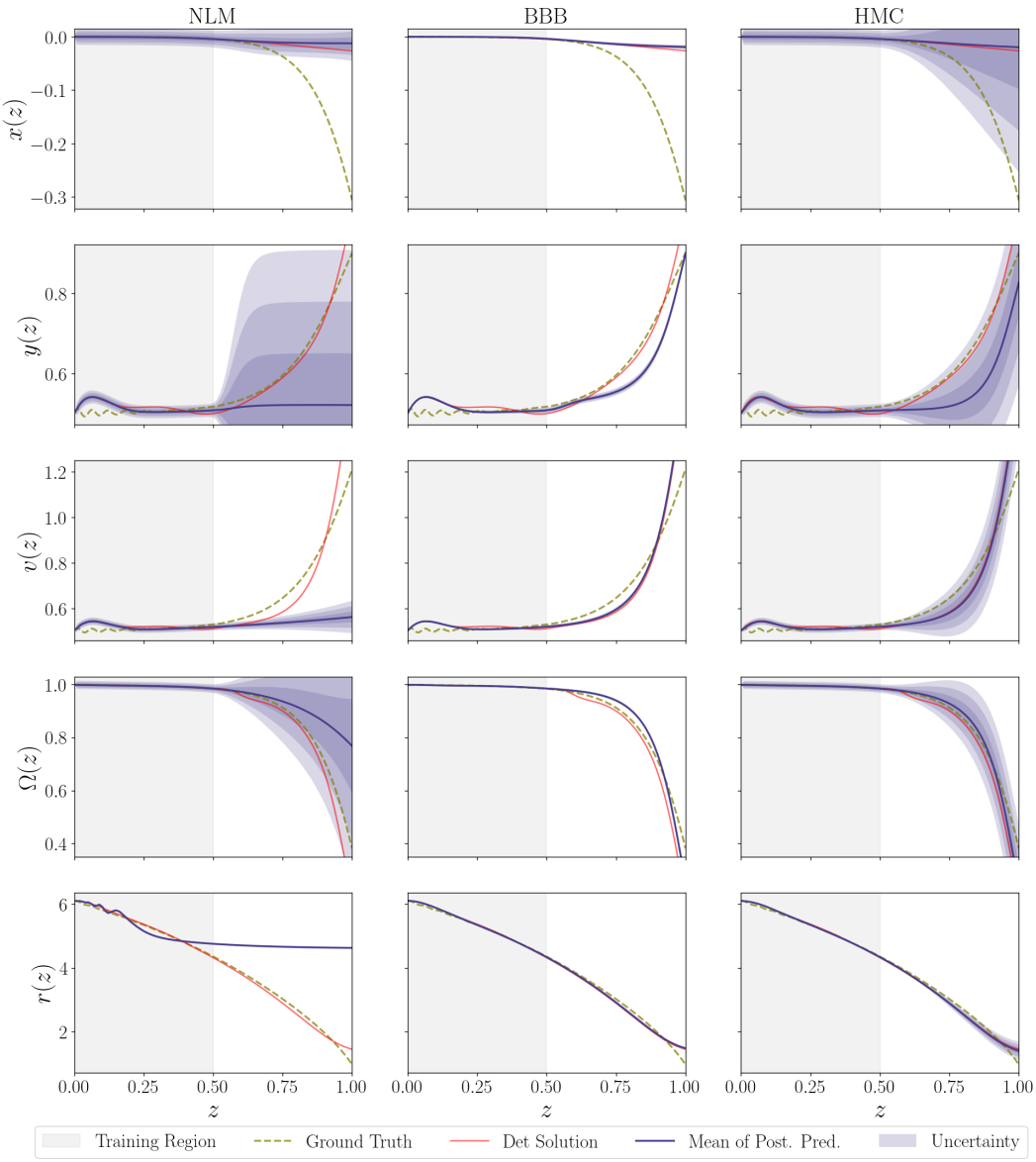


Figure 6: Hu-Sawicky solutions

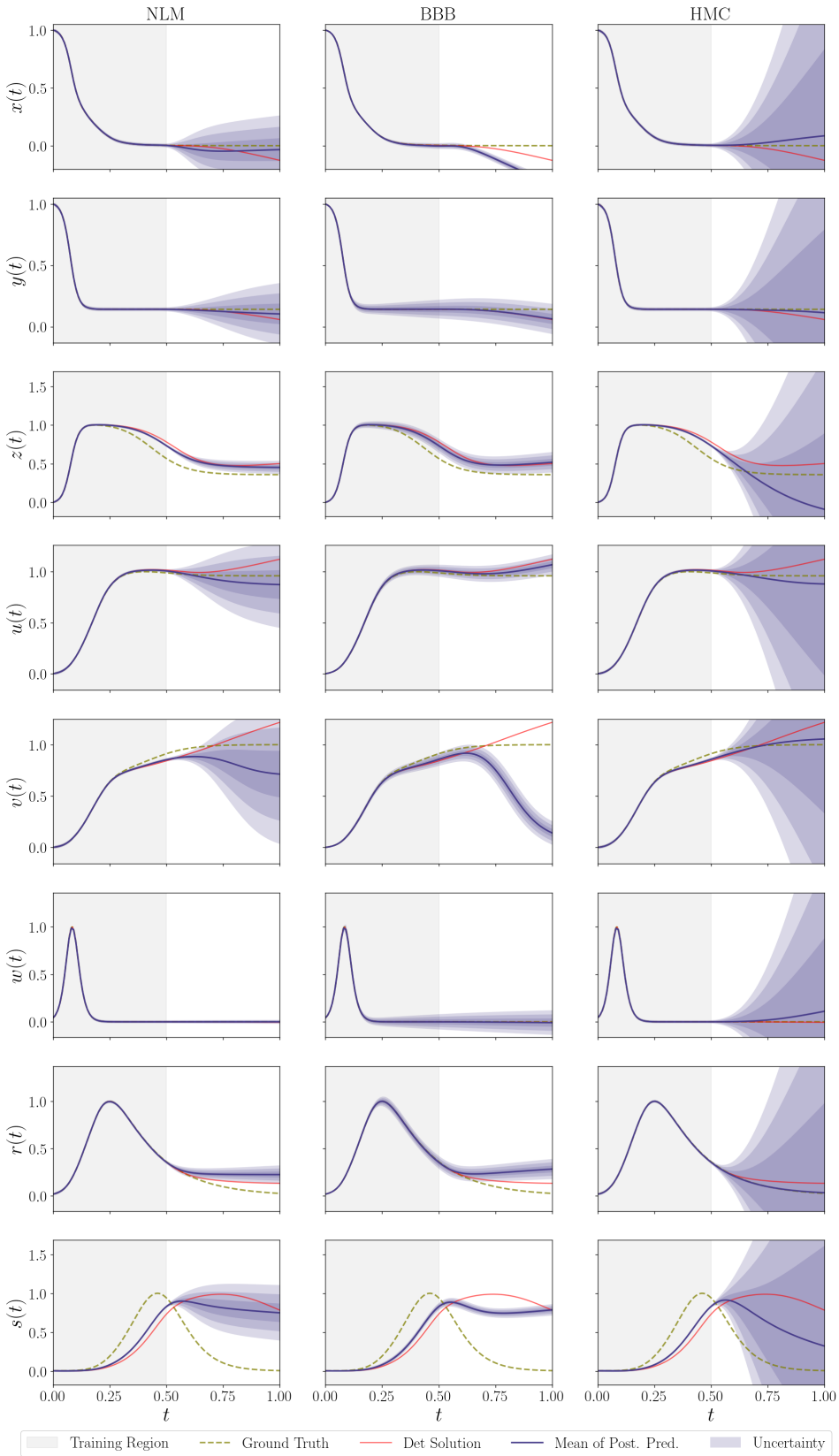


Figure 7: Cocoa fermentation solutions