# Uncertainty-aware Physics-informed Neural Networks for Robust CARS-to-Raman Signal Reconstruction

**Aishwarya Venkataramanan**      **Sai Karthikeya Vemuri**      **Adithya Ashok Chalain Valapil**

**Joachim Denzler**
Computer Vision Group, Friedrich Schiller University Jena, Germany

## Abstract

Coherent anti-Stokes Raman scattering (CARS) spectroscopy is a powerful and rapid technique widely used in medicine, material science, and chemical analyses. However, its effectiveness is hindered by the presence of a non-resonant background that interferes with and distorts the true Raman signal. Deep learning methods have been employed to reconstruct the true Raman spectrum from measured CARS data using labeled datasets. A more recent development integrates the domain knowledge of Kramers-Kronig relationships and smoothness constraints in the form of physics-informed loss functions. However, these deterministic models lack the ability to quantify uncertainty, an essential feature for reliable deployment in high-stakes scientific and biomedical applications. In this work, we evaluate and compare various uncertainty quantification (UQ) techniques within the context of CARS-to-Raman signal reconstruction. Furthermore, we demonstrate that incorporating physics-informed constraints into these models improves their calibration, offering a promising path toward more trustworthy CARS data analysis.

## 1   Introduction

Coherent anti-Stokes Raman scattering (CARS) spectroscopy has emerged as an important technique in chemistry, physics, and biomedical imaging due to its high-speed, label-free detection of molecular vibrations [1, 2]. Its ability to provide rapid chemical contrast makes it particularly attractive for live-cell imaging and real-time diagnostics. However, a fundamental limitation of CARS is the presence of a strong non-resonant background (NRB), which interferes with the resonant Raman response, distorting both the spectral shape and intensity. This reduces the spectral interpretability and imaging contrast [3].

To address this challenge, recent works use deep learning to learn the mappings from the NRB-contaminated CARS spectra to the underlying true Raman spectra. One promising direction is the use of physics-informed neural networks (PINNs) [4, 5, 6, 7], which embed physical constraints directly into the learning process. Recently, a model named RamPINN was developed by integrating known physical relationships between signal components. It is shown to reconstruct Raman spectra that are both data-consistent and physically plausible, improving the separation of resonant and non-resonant components even under challenging measurement conditions [7]. Nevertheless, RamPINN is deterministic in nature and does not account for model uncertainty. This is an increasingly important consideration when deploying machine learning models in domains where decisions have significant consequences, such as in biomedical diagnostics or materials characterization [8, 9].

In this study, we explore the integration of uncertainty quantification (UQ) methods with PINNs for CARS-to-Raman signal reconstruction. To this end, we systematically evaluate the prominent UQ strategies: Monte Carlo Dropout [8], Bayesian Neural Networks (BNNs) [10, 11], Deep En-

sembles [12], and Neural Processes (NPs) [13]. The results indicate that incorporating physical knowledge consistently improves reconstruction accuracy across the UQ methods and yields better confidence calibration.

## 2 Physics-Informed Learning

### 2.1 General Setup

Physics-informed learning integrates neural networks with governing physical laws to embed domain knowledge directly into the training process [5, 14]. The goal is to obtain models that not only fit observed data but also satisfy underlying physical constraints.

Let $x$ be the input, $\hat{y}$ the prediction within the domain $\Omega$, and $f_\theta$ a neural network parameterized by $\theta$. A known operator $\mathcal{D}$ encodes the physical constraint:

$$\mathcal{D}(\hat{y}, x) = 0 \,, \quad x \in \Omega \,, \tag{1}$$

where $\mathcal{D}$ is typically a differential or integral operator [5]. To enforce this constraint during training, a physics loss term is defined to penalize violations:

$$\mathcal{L}_{\text{phy}} = \frac{1}{|\Omega|} \int_\Omega |\mathcal{D}(f_\theta(x), u(x), x)|^2 \, dx \,. \tag{2}$$

where $u(x)$ represents any additional known functions or auxiliary variables required by the constraint. The total training objective combines the physics loss with a standard data-fitting term (e.g., mean squared error):

$$\mathcal{L}_{\text{total}} = \lambda_{\text{data}} \cdot \mathcal{L}_{\text{data}} + \lambda_{\text{phy}} \cdot \mathcal{L}_{\text{phy}}. \tag{3}$$

Here, $\lambda_{\text{data}}$ and $\lambda_{\text{phy}}$ control the relative weights of the data and physics components [15, 6, 16].

### 2.2 Kramers–Kronig and Smoothness Constraints

Following RamPINN[7], let $x$ denote the measured CARS spectrum, and $\hat{y} = (\hat{y}_{\text{raman}}, \hat{y}_{\text{NRB}})$ the predicted Raman and NRB components from $f_\theta(x)$. These components are related through the Kramers–Kronig (KK) relations [17, 18, 19, 20, 7].

**Kramers–Kronig Regularization.** From the causality principle, the real and imaginary parts of the CARS response are related through the Hilbert transform. Accordingly, the Raman component should correspond to the imaginary part of the Hilbert transform of the CARS signal after removing the non-resonant background (NRB) [7]:

$$\mathcal{L}_{\text{KK}} = |\hat{y}_{\text{raman}} - \Im[\mathcal{H}(x - \hat{y}_{\text{NRB}})]|^2 \,, \tag{4}$$

where $\mathcal{H}(\cdot)$ denotes the differentiable Hilbert transform.

**NRB Smoothness Regularization.** The NRB component is expected to be broad and smooth [21, 22, 23]. To enforce this, we penalize sharp variations:

$$\mathcal{L}_{\text{smooth}} = |\nabla \hat{y}_{\text{NRB}}|^2 \,. \tag{5}$$

**Final Loss.** The final objective combines the data fidelity term with both regularization terms:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{data}} \mathcal{L}_{\text{data}} + \lambda_{\text{KK}} \mathcal{L}_{\text{KK}} + \lambda_{\text{smooth}} \mathcal{L}_{\text{smooth}}, \tag{6}$$

where $\lambda_{\text{data}}$, $\lambda_{\text{KK}}$, and $\lambda_{\text{smooth}}$ control the relative contributions of each component.

## 3 Uncertainty Quantification

Unlike deterministic approaches, UQ enables models to express predictive confidence, providing a principled measure of reliability. In this work, we evaluate six Bayesian and ensemble-based UQ methods: Gaussian Processes (GP) [24], Bayesian Neural Networks (Full BNN), Partial BNNs [11], Monte Carlo Dropout (MC-Dropout)[8], Deep Ensembles[12], and Distance-informed Neural Processes (DNP) [13]. Among these, GP serves as a non-neural baseline and is a probabilistic model

with closed-form uncertainty estimates. The remaining methods are neural network-based. BNNs place Gaussian distributions over the network weights and use variational inference to approximate the full posterior. However, these methods are typically not scalable to large or deep networks due to the high computational cost of posterior approximation. Partial BNNs reduce this complexity by applying the Bayesian treatment to only the final layer of the model, offering a trade-off between scalability and uncertainty quality. MC-Dropout approximates Bayesian inference by performing multiple stochastic forward passes at test time, utilizing dropout, which makes it computationally efficient and easy to implement. Deep Ensembles train multiple independently initialized models and combine their predictions to estimate uncertainty. Finally, DNP extends Neural Processes by incorporating input-dependent uncertainty through distance-aware mechanisms, enabling flexible and structured uncertainty modeling.

To investigate the impact of physics-informed inductive biases on uncertainty estimates, we train each neural network-based model both with and without the physics-based loss described in Equation 6. In this setup, the data loss term $\mathcal{L}_{\text{data}}$ corresponds to the standard objective traditionally used to train the respective UQ model. The physics-informed training extends this baseline by incorporating additional constraints: Kramers-Kronig consistency loss $\mathcal{L}_{\text{KK}}$ and the NRB smoothness loss $\mathcal{L}_{\text{smooth}}$ to embed domain knowledge into the learning process.

# 4 Experiments and Results

## 4.1 Experimental Setup

**Synthetic spectra generation.** We generate 2000 CARS-Raman spectral pairs for our experiments. The spectral synthesis [25] and specific parameters for generation are as in [26]. The synthetic Raman spectra with $N$ peaks are modeled using a Lorentzian function as shown in Equation 7. Here, $A$ is the amplitude, $\Omega$ is the normalized resonance frequency, $\gamma$ is the linewidth and $\omega$ is the normalized Raman shift:

$$\chi_r^{(3)}(\omega) = \sum_{n=1}^{N} \frac{A_n}{\Omega_n - \omega - i\gamma_n}. \tag{7}$$

A polynomial or sigmoid function is used to replicate the NRB as shown in Equation 8. Here, $b_1, b_2$ control the steepness while $c_1, c_2$ determine the position of inflection points in the sigmoid function. Variables $a, b, c, d, e$ are the coefficients of polynomial function:

$$\chi_{nrb}^{(3)}(\omega) = \begin{cases} \frac{1}{1+e^{-b_1(\omega-c_1)}} \cdot \frac{1}{1+e^{b_2(\omega-c_2)}} & \text{(Sigmoid)} \\ a\omega^4 + b\omega^3 + c\omega^2 + d\omega + e & \text{(Polynomial)} \end{cases} . \tag{8}$$

This background along with some random noise $\epsilon$ is added over the Raman spectra to generate CARS spectra:

$$I_{\text{CARS}}(\omega) = \left| \chi_r^{(3)}(\omega) + \chi_{\text{nrb}}^{(3)}(\omega) \right|^2 + \epsilon. \tag{9}$$

**Real spectra.** We use six homogeneous sample spectra in our experiments. They are measured on broadband CARS and Spontaneous Raman setups, and are publicly available from the work [27, 28].

**Training and Evaluation.** For synthetic experiments, we randomly split the generated dataset into 80% for training and 20% for evaluation. For real-world experiments, we adopt a zero-shot evaluation protocol, where the models trained on the synthetic dataset and directly evaluated on real CARS-Raman measurements without any fine-tuning. The neural network backbone follows a 1D ResNet-style architecture consisting of four residual blocks [29]. These shared features are then passed to two separate $1 \times 1$ convolutional heads for predicting the resonant Raman component and the NRB. We train all models in PyTorch [30] using the Adam optimizer [31] with a learning rate of $10^{-3}$. The weight terms for the loss function are $\lambda_{\text{data}} = 10$, $\lambda_{\text{KK}} = 1$ and $\lambda_{\text{smooth}} = 10$. The MC-Dropout results were computed using 50 stochastic forward passes, while the deep ensemble results were obtained from an ensemble of 5 independently trained models.

For quantitative evaluation, we use the log-likelihood (LL) and expected calibration error (ECE) [32]. LL measures how well the predicted probability distribution explains the observed data, with higher values indicating more reliable uncertainty modeling. ECE quantifies the deviation between predicted confidence intervals and actual coverage. A lower value of ECE indicates a better performance.

Table 1: Comparison of uncertainty quantification methods on a synthetic dataset, evaluated with and without physics. Results are averaged over 10 runs.

| Method | Without Physics | | With Physics | |
|---|---|---|---|---|
| | LL ($\uparrow$) | ECE ($\downarrow$) | LL ($\uparrow$) | ECE ($\downarrow$) |
| GP | 1.115±0.017 | 0.202±0.004 | - | - |
| MC-Dropout | -1.164±0.013 | 0.213±0.015 | -2.089±0.053 | 0.278±0.012 |
| Deep Ensemble | -1.102±0.015 | 0.215±0.005 | 0.119±0.023 | 0.145±0.006 |
| Full BNN | **1.189±0.046** | **0.076±0.003** | **1.274±0.076** | **0.049±0.004** |
| Partial BNN | 0.731±0.012 | 0.280±0.003 | 0.762±0.004 | 0.234±0.007 |
| DNP | 1.175±0.021 | 0.114±0.003 | 1.188±0.017 | 0.104±0.005 |

Table 2: Comparison of uncertainty quantification methods on a real dataset, evaluated with and without physics. Results are averaged over 10 runs.

| Method | Without Physics | | With Physics | |
|---|---|---|---|---|
| | LL ($\uparrow$) | ECE ($\downarrow$) | LL ($\uparrow$) | ECE ($\downarrow$) |
| GP | 0.685±0.164 | **0.087±0.049** | - | - |
| MC-Dropout | -3.089±0.530 | 0.278±0.012 | -3.530 ±0.598 | 0.242±0.009 |
| Deep Ensemble | -2.567±0.436 | 0.255±0.021 | -2.016±0.418 | 0.208±0.012 |
| Full BNN | 0.685±0.231 | 0.172±0.041 | 0.716±0.263 | 0.160±0.076 |
| Partial BNN | 0.497±0.172 | 0.227±0.021 | 0.565±0.188 | 0.215±0.027 |
| DNP | **0.914±0.185** | 0.168±0.062 | **1.014±0.144** | **0.131±0.064** |

## 4.2 Synthetic data experiments

Table 1 presents the quantitative comparison of various UQ methods on the synthetic dataset, evaluated both with and without physics-based constraint. Among all methods, the Full BNN achieves the best performance in both settings. It obtains the highest LL and lowest ECE, indicating accurate and well-calibrated predictive uncertainties. The DNP model performs competitively, closely matching the Full BNN. The incorporation of physics-based loss consistently improves performance across most neural-based models. These results suggest that physics-informed training improves both the predictive performance of the models and calibration of uncertainty estimates for neural network-based approaches. The qualitative results are provided in Figure 1 in Appendix A.

## 4.3 Real-world experiments

Table 2 summarizes the performance of the UQ baselines on the real CARS-Raman dataset under zero-shot setting. DNP achieves the best overall performance, yielding the highest LL and lowest ECE both with and without physics-based regularization. This strong performance can be attributed to DNP's meta-learning architecture, which is explicitly designed to enable fast adaptation and generalization across tasks. Full BNN also performs competitively, showing strong LL scores and improved calibration when physics-informed loss terms are included. Similar to the observation from the experiments on synthetic dataset, incorporating physics improves both metrics for most models. This suggests that domain knowledge can help reduce overconfidence and improve generalization even in zero-shot scenarios. Qualitative results are provided in Figure 2 in Appendix A.

## 5 Conclusion

In this study, we conducted a systematic evaluation of UQ methods for CARS-to-Raman signal reconstruction. Experiments on synthetic and real datasets show that embedding domain knowledge via physics consistently improves reconstruction accuracy and uncertainty calibration. Among the evaluated UQ methods, Full BNNs and Distance-informed Neural Processes achieved the best overall performance, with physics-informed training yielding higher log-likelihood scores and lower calibration errors. More broadly, these results demonstrate that integrating physical priors not only improves signal fidelity but also enhances the reliability and interpretability of uncertainty estimates. Future work will explore scaling these approaches to broader spectral ranges, multi-modal imaging, and active learning frameworks that leverage uncertainty estimates for adaptive data acquisition.

# References

[1] C. Krafft, I. W. Schie, T. Meyer, M. Schmitt, and J. Popp. Developments in spontaneous and coherent raman scattering microscopic imaging for biomedical applications. *Chemical Society Reviews*, 45(7):1819–1849, 2016.

[2] Ji-Xin Cheng and X. Sunney Xie. Coherent anti-stokes raman scattering microscopy: instrumentation, theory, and applications. *The Journal of Physical Chemistry B*, 108(3):827–840, 2004.

[3] Dario Polli, Vikas Kumar, Carlo M. Valensise, Marco Marangoni, and Giulio Cerullo. Broadband Coherent Raman Scattering Microscopy. *Laser & Photonics Reviews*, 12(9):1800020, 2018.

[4] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics informed deep learning (part ii): Data-driven discovery of nonlinear partial differential equations, 2017.

[5] M. Raissi, P. Perdikaris, and G. E. Karniadakis. 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, 2019.

[6] Sai Karthikeya Vemuri, Tim Büchner, Julia Niebling, and Joachim Denzler. Functional tensor decompositions for physics-informed neural networks, 2024.

[7] Sai Karthikeya Vemuri, Adithya Ashok Chalain Valapil, Tim Büchner, and Joachim Denzler. Rampinn: Recovering raman spectra from coherent anti-stokes spectra using embedded physics, 2025.

[8] Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059. PMLR, 2016.

[9] Aishwarya Venkataramanan, Assia Benbihi, Martin Laviale, and Cédric Pradalier. Gaussian latent representations for uncertainty estimation using mahalanobis distance in deep classifiers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4488–4497, 2023.

[10] Charles Blundell, Julien Cornebise, Koray Kavukcuoglu, and Daan Wierstra. Weight uncertainty in neural network. In *International conference on machine learning*, pages 1613–1622. PMLR, 2015.

[11] Mrinank Sharma, Sebastian Farquhar, Eric Nalisnick, and Tom Rainforth. Do bayesian neural networks need to be fully stochastic? In *International Conference on Artificial Intelligence and Statistics*, pages 7694–7722. PMLR, 2023.

[12] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems*, 30, 2017.

[13] Aishwarya Venkataramanan and Joachim Denzler. Distance-informed neural processes. *arXiv preprint arXiv:2508.18903*, 2025.

[14] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

[15] Sai Karthikeya Vemuri and Joachim Denzler. Gradient statistics-based multi-objective optimization in physics-informed neural networks. *Sensors*, 23(21), 2023.

[16] Levi D. McClenny and Ulisses M. Braga-Neto. Self-adaptive physics-informed neural networks. *Journal of Computational Physics*, 474:111722, 2023.

[17] Valerio Lucarini, Jarkko J Saarinen, Kai-Erik Peiponen, and Erik M Vartiainen. *Kramers-kronig relations in optical materials research*. Springer Series in Optical Sciences. Springer, 2005 edition, 2005.

[18] Hendrik Anthony Kramers. *La diffusion de la lumière par les atomes*. Como, 1928.

[19] Ralph de Laer Kronig. On the theory of dispersion of x-rays. *Journal of The Optical Society of America*, 12(6):547–557, 1926.

[20] Robert D. Guenther, Duncan G. Steel, and Leopold Bayvel. *Encyclopedia of Modern Optics*. Elsevier / Academic Press, 2004.

[21] Stanislav O Konorov, Michael W Blades, and Robin F B Turner. Non-resonant background suppression by destructive interference in coherent anti-stokes raman scattering spectroscopy. *Opt. Express*, 19(27):25925–25934, 2011.

[22] Ramesh Junjuri, Sina Saghi, Zhen Wang, Xue Li, Yufeng Liu, Wenjun Zhang, and Xiaojun Xu. Bidirectional lstm for coherent anti-stokes raman spectroscopy signal processing. *Journal of Raman Spectroscopy*, 52(4):813–823, 2021.

[23] Zhen Wang, Xue Li, Yufeng Liu, Wenjun Zhang, and Xiaojun Xu. Vector: Very deep convolutional autoencoder for cars spectrum denoising and background removal. *Optics Express*, 29(4):6196–6210, 2021.

[24] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

[25] C. M. Valensise, A. Giuseppi, F. Vernuccio, A. De la Cadena, G. Cerullo, and D. Polli. Removing non-resonant background from cars spectra via deep learning. *APL Photonics*, 5(6):061305, 2020.

[26] Adithya Ashok Chalain Valapil, Carl Messerschmidt, Maha Shadaydeh, Michael Schmitt, Jürgen Popp, and Joachim Denzler. Deep learning-assisted dynamic mode decomposition for non-resonant background removal in cars spectroscopy. In *DAGM German Conference on Pattern Recognition (DAGM-GCPR)*, 2025. (accepted).

[27] Federico Vernuccio, Arianna Bresci, Benedetta Talone, Alejandro de la Cadena, Chiara Ceconello, Stefano Mantero, Cristina Sobacchi, Renzo Vanna, Giulio Cerullo, and Dario Polli. Fingerprint multiplex cars at high speed based on supercontinuum generation in bulk media and deep learning spectral denoising. *Optics Express*, 30(17):30135, 2022.

[28] Federico Vernuccio, Arianna Bresci, Benedetta Talone, Alejandro De la Cadena Pérez-Gallardo, Chiara Ceconello, Stefano Mantero, Cristina Sobacchi, Renzo Vanna, Giulio Cerullo, and Dario Polli. Supplementary material - Fingerprint multiplex CARS at high speed based on supercontinuum generation in bulk media and deep learning spectral denoising, 2022.

[29] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[30] A Paszke. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703*, 2019.

[31] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[32] Volodymyr Kuleshov, Nathan Fenner, and Stefano Ermon. Accurate uncertainties for deep learning using calibrated regression. In *International conference on machine learning*, pages 2796–2804. PMLR, 2018.

# A  Qualitative Evaluation

We qualitatively assess the impact of physics-informed constraints on the Full BNN and DNP models. Figure 1 illustrates their performance on a representative sample from the synthetic dataset, while Figure 2 presents results on the real dataset.

In the synthetic case (Figure 1), all models closely reproduce the ground-truth Raman spectra and effectively suppress the non-resonant background. Incorporating the physics constraint further reduces reconstruction errors. The models, however, exhibit distinct uncertainty behaviors: the Full BNN shows broadly distributed uncertainty across the spectrum, whereas the DNP yields smoother, more localized estimates concentrated in low-signal-to-noise regions.

For the real dataset (Figure 2), both methods generalize well in a zero-shot setting, with the DNP maintaining more consistent uncertainty calibration. Consistent with the synthetic results, applying the physics constraint again leads to lower reconstruction errors and improved reliability.
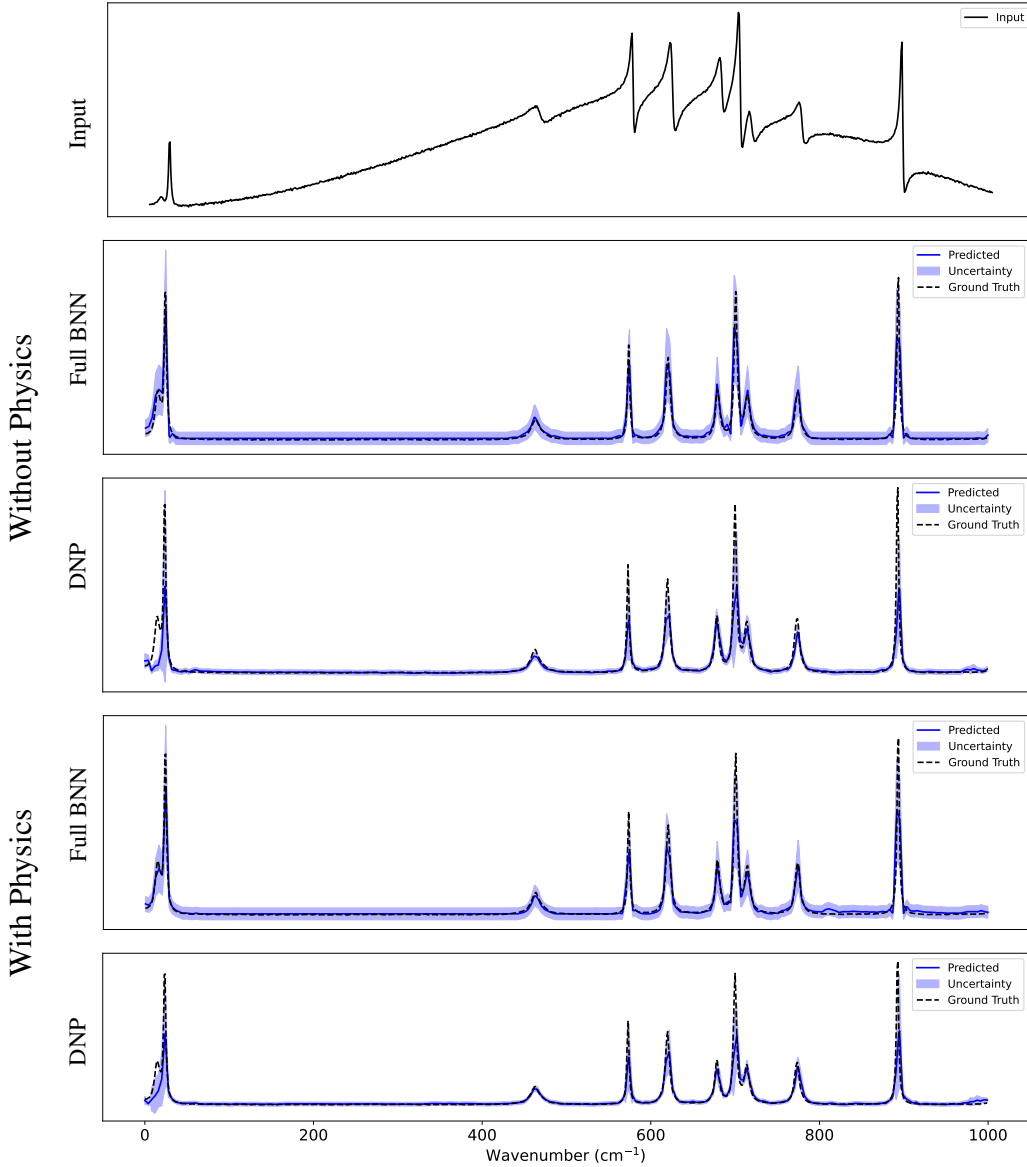


Figure 1: **Synthetic dataset results.** Reconstructions from the Full BNN and DNP models without (top) and with (bottom) the physics constraint.
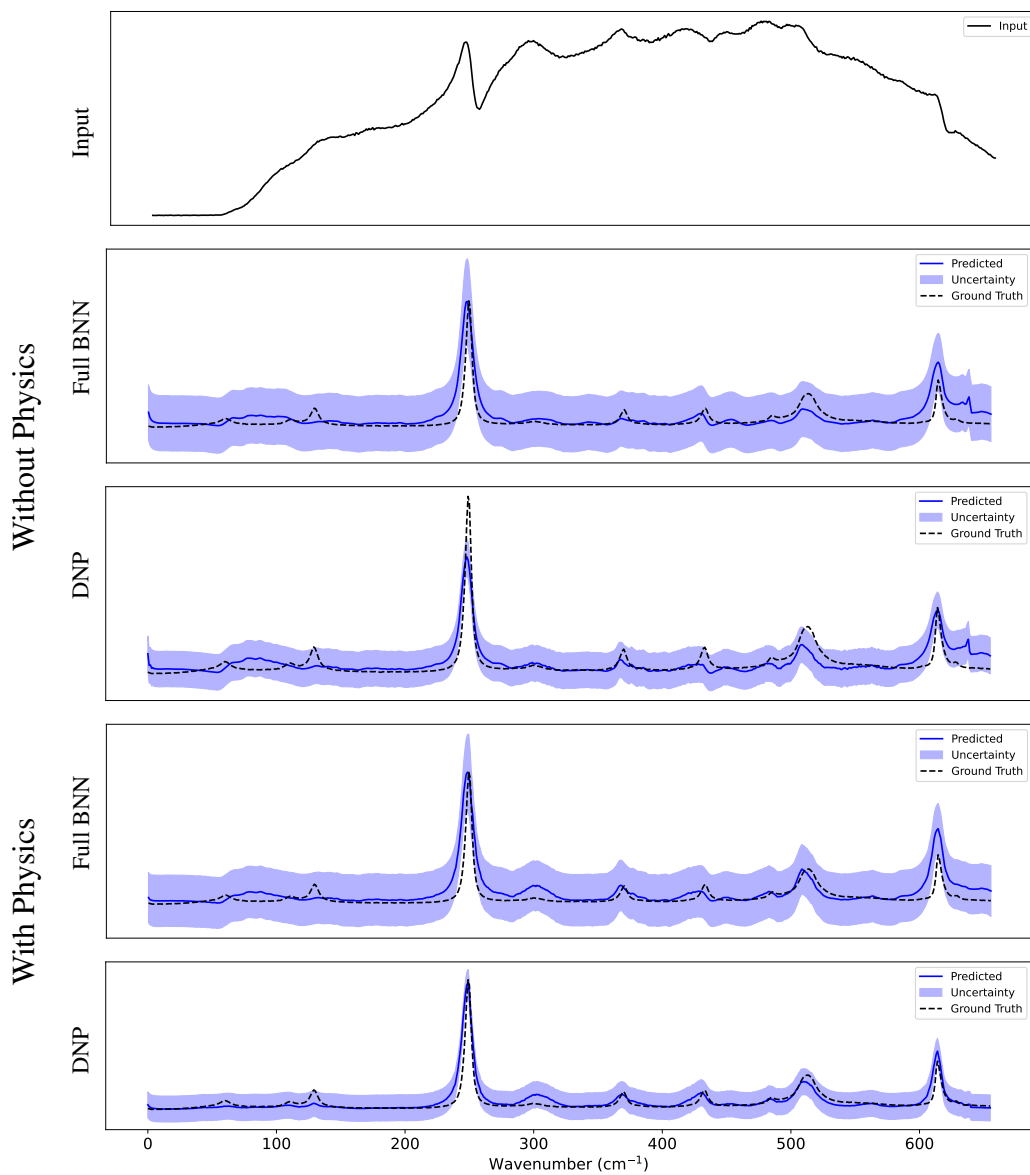
Figure 2: **Real dataset results.** Zero-shot reconstructions from the Full BNN and DNP models without (top) and with (bottom) the physics constraint.