Noise2Noise Image Reconstruction of Lifetime Maps in Halide Perovskite Thin Films

Gabriele Scrivanti
IPVF, UMR 9006, CNRS,
Ecole Polytechnique,
IP Paris, Chimie Paristech, PSL
Palaiseau, France
gabriele.scrivanti@cnrs.fr

Luca Calatroni
MaLGa Center, DIBRIS,
Università di Genova,
MMS, Istituto Italiano di Tecnologia,
Genoa, Italy
luca.calatroni@unige.it

Stefania Cacovich
IPVF, UMR 9006, CNRS,
Ecole Polytechnique,
IP Paris, Chimie Paristech, PSL
Palaiseau, France
stefania.cacovich@cnrs.fr

***Abstract*—We present an unsupervised deep-learning approach for lifetime map reconstruction from noisy time-resolved fluorescence imaging (TR-FLIM) datasets. In the context of semiconductor and photovoltaic device characterisation, this method is critical for accurately predicting solar cell performance and detecting early signs of degradation. More precisely, we consider an unsupervised Noise2Noise (N2N) training framework combined with physics-driven modelling for the quantitative reconstruction of lifetime maps. The proposed approach incorporates a log-linear fit in the N2N loss function and parameterises the unknown maps as outputs of a shallow neural network with a multi-branch architecture. By learning from multiple noisy acquisitions of the same scene, our method effectively allows an accurate estimation with shorter acquisition protocols, which translates into a lower risk of damage for the sample under consideration. Tests on simulated data and comparisons with available model-based approaches show that the proposed approach improves robustness w.r.t. noise levels with limited tuning of the regularisation/algorithmic parameters.**

***Index Terms*—Quantitative image reconstruction, Noise2Noise, perovskite cell characterisation.**

I. Introduction

Photovoltaic (PV) devices play a pivotal role in the global shift toward a low-carbon economy by offering sustainable solutions for light-to-electricity conversion. As the demand for solar energy grows, optimising PV technology performance and stability is crucial [1]. Advanced characterization techniques, such as steady-state photoluminescence imaging and time-resolved fluorescence imaging (TR-FLIM), are fundamental tools for analysing optoelectronic properties such as carrier dynamics, defect densities, and diffusion lengths [2]. Here, TR-FLIM is employed to obtain spatially resolved maps of the surface recombination rate and carrier lifetime in halide perovskite thin films. After excitation at $t = 0$, the long-time ($t \gg 0$) PL intensity $P(t)$ emitted by a semiconductor can be described by a negative exponential time-dependent model, where for any measurement point $n \in \{1, \dots, N\}$ in the detection array, the intensity is given by

$$P_n(t) = \delta_n \exp(-t/\tau_n) \tag{1}$$

with $\delta_n = P_n(0) > 0$ representing the initial intensity coefficient and $\tau_n > 0$ is the material-specific decay time, often called carrier lifetime. The δ parameter can be related to the surface recombination rate of a semiconductor.

Such characterisation could be quite challenging from a practical viewpoint. First, in photoluminescence measurements, a critical trade-off exists between data quality and sample preservation. Performing experiments with a high number of accumulations – that is, many individual measurements combined to form a single data point within a repetition of the experiment – increases the signal-to-noise ratio (SNR), at the price of longer experimental times. This can potentially cause damage to the sample or alter the properties under analysis, especially in the case of laser-sensitive materials such as halide perovskite or organic materials. Shorter experimental times are therefore preferable to minimise sample degradation, but this comes at the cost of having noisier data. A reconstruction pipeline robust to different SNR levels, thus requires effective processing methods specifically tailored to PL data.

Lifetime maps can only be indirectly estimated by solving the inverse problem modelled by (1), which, depending on the quality of the data, could be a challenging task. In a recent work [3], a variational approach combining robust noise modelling with Structural Total Variation regularisation [4] is employed to estimate such parameter maps. The method has been shown to be effective, despite the tailored choice of regularisation (prone to artefacts such as, e.g., staircasing) and a heuristic strategy to identify the optimal regularisation parameter, which may be inefficient as it requires the solution of several instances of the underlying minimisation problem. To overcome these limitations, one possibility consists in considering learning-based techniques which enjoy better expressivity and limited parameter-tuning upon suitable training. Note, however, that in the specific PL context described above, having at disposal reliable reference parameter maps to be used as clean samples presents a significant challenge, as they can only be derived from high-quality data acquired under optimal

SC and GS acknowledge funding from the ANR under the JCJC project CLEOPATRA (ANR-23-CE50-0017). LC acknowledges the financial support of the European Research Council (grant MALIN, 101117133). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency.

conditions, a process that demands considerable resources and technical expertise. To avoid such need, we then focus on unsupervised deep learning techniques, a class of approaches that allows training also under limited data.

Among the plethora of existing unsupervised approaches, Noise2Noise (N2N) models [5] and their extensions (see, e.g., [6]–[9]) are nowadays a prominent example which has been widely used in applications. N2N is a fully unsupervised approach solely relying on pairs of noisy observations of the same scene for the training of the network parameters, thus eliminating the need for clean ground truth data. Applications of these approaches to photovoltaic materials can be found in [10] for the restoration of a stack of noisy PL images acquired at different wavelengths. Unsupervised approaches have also been adapted to be used under data scarcity conditions. While standard N2N methods necessitate indeed extensive training on large datasets of image pairs, the works in [11]–[13] limit training to solely multiple noisy realisations of the very same scene. Given the absence of task-specific training examples, such methods are also referred to as *zero-shot* methods in the relevant literature. They are particularly valuable in scientific imaging, where only multiple rapid acquisitions of the same scene acquired under the same experimental conditions can be obtained, while a complete training dataset that is comprehensive of multiple setups and acquisition protocols is often challenging to obtain. Note that a similar paradigm to zero-shot N2N, is the Deep Image Prior [14] where the core idea is to transform a random input into an observed noisy and distorted image by means of a deep neural network, with the idea that it will generate a cleaner version of the image upon early stopping.

Contribution: We propose a hybrid unsupervised N2N deep-learning approach combined with physical log-linear modelling for quantitative reconstruction of lifetime parameter maps in PL applications. Drawing inspiration from zero-shot N2N approaches, we take advantage of the availability of multiple noisy acquisitions of the same sample, which allows us to benefit from mini-batching during the optimisation process. We reparametrise the two unknown parameter maps in terms of a single-input two-branched shallow neural network and consider a suitable data term to be optimised in combination with a mild cross-channel Structured Total Variation not requiring any fine parameter tuning to avoid noise overfitting. Extensive simulations and real-data reconstructions are performed in comparison with plain pixel-wise linear regression and a regularised variational technique proposed in [3].

Notations: We use standard common notation from deep learning (B, C, N), with B representing the batch size, C the number of channels and N the size of the signal (that is, images with dimension $n_1 \times n_2$ are assumed to be vectorised so that $N = n_1 n_2$). We use boldface for vectors and italics for their scalar components. For simplicity, we omit explicit indexing whenever all index values of the corresponding dimension are involved. By $\Theta \subset \mathbb{R}^d, d \gg 1$ we denote the space of network weight configurations so that $\theta \in \Theta$.

II. Modelling & proposed approach

In real-world experimental conditions, data collection occurs over discrete time windows or *gates*, whose width $\omega > 0$ can be adjusted to control SNR. During the c-th measurement window $[t_c, t_c + \omega_c]$, the averaged intensity value

$$P_{c,n} = \frac{1}{\omega_c} \int_{t_c}^{t_c+\omega_c} P_n(t) dt \tag{2}$$

is obtained. We refer to the set of all $P_{c,n}$ for $c \in \{1, \dots, C\}$ and $n \in \{1, \dots, N\}$ as the *photoluminescence cube*. To simplify our analysis, we can now convert the exponential relationship (1) into a linear form through a logarithmic transformation. Then, for every $n \in \{1, \dots, N\}$ and $c \in \{1, \dots, C\}$, given the logarithmically transformed measurements

$$p_{c,n} := \ln(P_{c,n}) \tag{3}$$

and a punctual acquisition time t_c (for sufficiently small measurement window $\omega_c \ll 1$), we seek for $x_{1,n} = \log(\delta_n)$ and $x_{2,n} = -1/\tau_n$ using the following linear model

$$p_{c,n} = x_{1,n} + x_{2,n} t_c. \tag{4}$$

The data collection process can be performed repetitively to obtain multiple photoluminescence cubes of the sample under analysis, thus allowing for an averaging of the results which reduces the effect of acquisition noise and enhances the accuracy of the measurements. These multiple acquisitions can also be stacked in an additional dimension, thus considering the set of indices $b = 1, \dots, B$ and the tensors $\mathbf{P} = (P_{b,c,n})_{b,c,n}$ and $\mathbf{p} = (p_{b,c,n})_{b,c,n}$.

A. Noise2Noise (N2N)

The Noise2Noise approach [5] (and the self-supervised denoising approaches inspired by it) relies on two key ideas. The former is that a neural network $\mathsf{N}_\theta : \mathbb{R}^{C \times N} \to \mathbb{R}^{C \times N}$ parametrised by $\theta \in \Theta$ is used to model the reconstruction process mapping noisy observations $\mathbf{y} \to \mathbb{R}^{C \times N}$ into a denoised one $\mathbf{x} \in \mathbb{R}^{C \times N}$. The latter is that, for its training, a loss function $\mathcal{L} : (\mathbb{R}^{C \times N})^2 \to \mathbb{R}$ is used, which involves two different noisy observation $\mathbf{y}^{\text{input}} \in \mathbb{R}^{C \times N}$ and $\mathbf{y}^{\text{target}} \in \mathbb{R}^{C \times N}$ of the same clean image $\mathbf{x} \in \mathbb{R}^{C \times N}$ as a training pair

$$\operatorname*{argmin}_{\theta \in \Theta} \sum_{b=1}^{B} \mathcal{L}(\mathsf{N}_\theta(\mathbf{y}_b^{\text{input}}), \mathbf{y}_b^{\text{target}}), \tag{5}$$

under the assumption that the noise is unstructured and independent. Hereabove, B corresponds to the dimension of the training set. Extension of N2N to proper imaging inverse problems such as, e.g., undersampled MRI data reconstruction were also showed [5].

Note that while bypassing the need for clean reference data, N2N approaches were originally applied using image datasets containing multiple acquisitions with independent noise realisations of the same scene. In scientific imaging and, more in particular, for the type of application targeted here, the construction of a proper dataset would demand considerable time and resources. Note, in particular, that the acquisition

parameters (such as temporal sampling intervals ω_c) are inherently experiment-specific and can vary significantly across different experimental protocols. Dataset-free, or zero-shot methods [11]–[13] are thus essential for feasibility.

B. N2N quantitative estimation of lifetime maps

Recalling (4), we design the N2N approach such that given B acquisitions $\mathbf{p} \in \mathbb{R}^{B\times C\times N}$ and the corresponding time steps $\mathbf{t} \in \mathbb{R}^C_+$, a reconstruction of the desired quantitative maps is obtained. We randomly choose a batch index $\bar{b} \in \{1,\dots,B\}$ to be our *input* data and we process it through a network $\mathsf{N}_\theta\colon \mathbb{R}^{C\times N} \to \mathbb{R}^{2\times N}$ (whose weights $\theta \in \Theta$ are trained using the remaining PL cubes $b \neq \bar{b}$ as *target* data) to output the sought-for pair of parameter maps $([\mathsf{N}_\theta(\mathbf{p}_{\bar{b}})]_1, [\mathsf{N}_\theta(\mathbf{p}_{\bar{b}})]_2)$. The training of the weights is performed as follows. For each batch index $b \in \{1,...,B\}\setminus\bar{b}$ and each channel $c \in \{1,\dots,C\}$, to estimate the linear fits in equation (4) we minimise $\mathcal{J}:\ \Theta \to \mathbb{R}$ defined as

$$\mathcal{J}(\theta) := \sum_{\substack{b=1\\ b\neq\bar{b}}}^{B}\sum_{c=1}^{C} \mathcal{L}\left(\mathbf{p}_{b,c},\, [\mathsf{N}_\theta(\mathbf{p}_{\bar{b}})]_1 + t_c[\mathsf{N}_\theta(\mathbf{p}_{\bar{b}})]_2\right) + \lambda\mathcal{R}([\mathsf{N}_\theta(\mathbf{p}_{\bar{b}})]_1, [\mathsf{N}_\theta(\mathbf{p}_{\bar{b}})]_2)$$

where $\mathcal{L}:\ (\mathbb{R}^N)^2 \to \mathbb{R}$ is a suitable loss function and $\mathcal{R}:\ \Theta \to \mathbb{R}$ is a regularisation function that reduces the risk of overfitting and obviates the need for defining an early stopping criterion, which is generally a challenging task. The parameter $\lambda > 0$ controls the regularisation strength.

Following [3], as far as $\mathcal{L}$ is concerned, we consider the Huber function defined by:

$$\mathcal{L}(\mathbf{x}_1,\mathbf{x}_2) = H_\epsilon(\mathbf{x}_1,\mathbf{x}_2) = \sum_{n=1}^{N} h_\epsilon(x_{1,n}-x_{2,n}) \tag{6}$$

– where, for $x \in \mathbb{R}$,

$$h_\epsilon(x) = \begin{cases} \frac{1}{2}x^2 & \text{for } |x| \le \epsilon \\ \epsilon(|x| - \frac{\epsilon}{2}) & \text{for } |x| > \epsilon \end{cases}$$

– which enforces robust regression while preserving the regularity of the objective function. Note that such a choice describes a mixture of Gaussian and impulsive noise in the data, see, e.g. [15]. To promote spatial correlation between the estimated maps $[\mathsf{N}_\theta(\mathbf{p}_{\bar{b}})]_1$ and $[\mathsf{N}_\theta(\mathbf{p}_{\bar{b}})]_2$, along with spatial consistency within neighbouring values, we considered the Structure Tensor Total Variation (STV) regularisation [4] which, for a two-channel image $\mathbf{x} = (\mathbf{x}_1,\mathbf{x}_2) \in \mathbb{R}^{2\times N}$, is defined in terms of the image gradient $\mathsf{D}:\ \mathbb{R}^{2\times N} \to \mathbb{R}^{2\times 2\times N}$ which for each $i \in \{1,\dots,N\}$ reads

$$[\mathsf{D}\mathbf{x}]_i = \begin{bmatrix} x_{1,i}-x_{1,v_i} & x_{2,i}-x_{2,v_i} \\ x_{1,i}-x_{1,h_i} & x_{2,i}-x_{2,h_i} \end{bmatrix}, \tag{7}$$

where the pair $(v_i,h_i) \in \{1,\dots,N\}^2$ denotes the position of the vertical and horizontal nearest neighbours of the i-th pixel. For $\mathbf{x} \in \mathbb{R}^{2\times N}$, the STV operator is thus defined as:

$$\mathrm{STV}(\mathbf{x}) = \sum_{i=1}^{N} \|[\mathsf{D}\mathbf{x}]_i\|_p \tag{8}$$

where $\|\cdot\|_p$ denotes the Schatten p-norm [16]. In this work, we adopt a smoothed version of the case $p=2$ (corresponding to the Frobenius norm) defined as

$$\mathrm{STV}_\varepsilon(\mathbf{x}) = \sum_{i=1}^{N} \sqrt{\varepsilon^2 + (x_{1,i}-x_{1,v_i})^2 + (x_{2,i}-x_{2,v_i})^2 + (x_{1,i}-x_{1,h_i})^2 + (x_{2,i}-x_{2,h_i})^2}$$

where $\varepsilon \ll 1$ is set a priori.

C. Network architecture

Following [13], we exploit a simple architecture with two parallel branches emerging from a shared input layer, each processing the sought-for map independently. Each branch consists of two sequential blocks, where each block contains a 3×3 convolutional layer, followed by a batch normalisation layer and a Leaky ReLU activation function. The branches conclude with a 1×1 convolutional layer. This double-branched shallow architecture involves approximately 90k parameters. We report a graphical illustration of the network in Figure 1-(c).

III. Experimental Results and Discussion

We validate our proposed approach on both simulated and real TR-FLIM data, following the simulation/data analysis procedure described in [3].

A. Data Simulation

Two synthetic PL datasets (*Data1* and *Data2*) with an image size of $n_1 = n_2 = 512$ and including different geometries were generated using a drift-diffusion physical model starting from different physical values (see [3]). We considered $C = 300$ time steps, uniformly distributed over an interval of 1000 ns. As shown in Figure 1-(a), we simulated data characterised by two and three distinct regions, each defined by specific values of $\log(\delta)$ and τ, which correspond to different drift-diffusion parameters. The plots in Figure 1-(b) represent the mean PL decay, that is, at each time step $c \in \{1,\dots,300\}$ the mean value of $\mathbf{P}_c$ over time. The simulated PL data were corrupted with additive Gaussian noise with a standard deviation of 10^{-2} and 10^{-3}. To simulate multiple noisy acquisitions of the same scene for our N2N approach, we generated $B = 15$ independent noise realisations. For the analysis, we focused on fitting the measurements between 495 ns and 700 ns, as the mono-exponential model is assumed to hold within this range and the signal-to-noise ratio is sufficiently high to provide meaningful estimations.

B. Hyper-parameters setting

For minimising (6), we need to set three hyperparameters: the Huber function threshold $\epsilon > 0$, the smoothing parameter for STV $\varepsilon > 0$ and the regularisation parameter $\lambda > 0$. In the spirit of defining an approach with minimal fine-tuning, we set $\epsilon = 1$ (as in [3]), $\varepsilon = 10^{-3}$ and $\lambda = 10^{-5}$ and keep this configuration in all our experiments. For optimisation, we used Adam optimiser. We let the optimisation run for up to 5000 epochs and set a threshold to 10^{-5} for the relative change between two epochs for the estimated maps and the

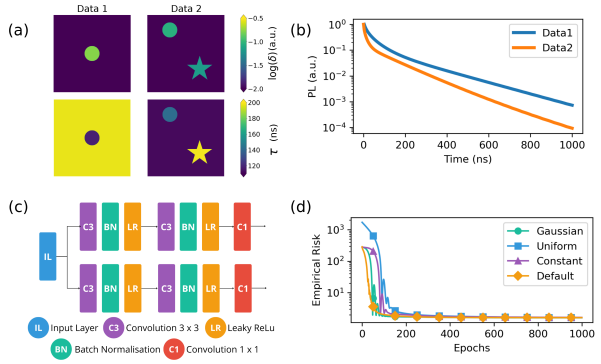

Fig. 1. (a) Parameter maps related to our synthetic datasets Data1 and Data2; (b) mean PL decay (mean value of P_c over time) for Data1 and Data2; (c) illustration of our shallow two-branched NN N_θ; (d) evolution of the loss function for different initialisations of the parameters in N_θ.

loss as stopping criteria. The network weights are initialised using the default settings provided by PyTorch: Kaiming Uniform Initialisation for the convolutional layers and identity initialisations for the batch normalisation layers. We also tested alternative initialisations to analyse the method's robustness: a Gaussian initialisation with mean 0.01 and std 0.01, a constant initialisation to 0.01 and a uniform initialisation over the interval $[0, 1]$. The evolution of the loss over 1000 epochs in the case of Data1 for the four initialisations is reported in Figure 1-(d), which shows that the performances are comparable. All experiments and simulations are run on an RTX 2000 Ada Generation Laptop GPU.

C. Parameter Maps estimation results

We evaluate our results on synthetic data considering the sum of Normalised Mean Squared Error (NMSE) w.r.t. the ground-truth across both maps. This global metric is specifically chosen to account for the different scaling factors between the two maps, thus providing a balanced global assessment of our model performance. Given an estimated solution $\widehat{\mathbf{x}} \in \mathbb{R}^{2\times N}$ and the GT $\mathbf{x}^{\mathrm{GT}} \in \mathbb{R}^{2\times N}$, such metric is defined as:

$$\mathrm{ERR}(\widehat{\mathbf{x}}, \mathbf{x}^{\mathrm{GT}}) = \frac{1}{N|\overline{\mathbf{x}}_1^{\mathrm{GT}}|}\sum_{i=1}^{N}(\widehat{\mathbf{x}}_{1,i} - \mathbf{x}_{1,i}^{\mathrm{GT}})^2 + \frac{1}{N|\overline{\mathbf{x}}_2^{\mathrm{GT}}|}\sum_{i=1}^{N}(\widehat{\mathbf{x}}_{2,i} - \mathbf{x}_{2,i}^{\mathrm{GT}})^2,$$

where $\overline{\mathbf{x}}_1^{\mathrm{GT}} \in \mathbb{R}$ and $\overline{\mathbf{x}}_2^{\mathrm{GT}} \in \mathbb{R}$ represent the mean values of the reference maps $\mathbf{x}_1^{\mathrm{GT}}$ and $\mathbf{x}_2^{\mathrm{GT}}$. We compare the results of our N2N approach with a pixel-wise (PW) least-squares regression and the regularised method (REG) from [3]. The optimal choice of the regularisation parameter for REG was obtained for each dataset with the L-curve principle by evaluating the solutions for 30 candidate values spanning several reasonable orders of magnitude. The choice of an adequate regularisation parameter is crucial for the proper behaviour of such an approach but could be tedious and time-consuming.

Table I presents the ERR values attained by the three considered methods. For REG, we also report in brackets the minimal ERR value among the 30 candidate solutions used

	Noise level	PW	REG	N2N
Data1	10^{-2}	14.223	**0.033** (0.016)	0.039
	10^{-3}	0.137	**0.006**	0.007
Data2	10^{-2}	33.85	0.176	**0.033**
	10^{-3}	8.473	0.012 (0.006)	**0.002**

TABLE I
ERR VALUES OF THE ESTIMATED MAPS

for the L-curve in case this method identified a solution where REG performed significantly worse. This illustrates how such a heuristic principle does not always recognise the optimal parameter. For Data1, N2N performed slightly worse than REG but the performances are rather comparable, whereas for Data2 the N2N approach significantly outperformed REG.

In Figure 2 we report (in colour map `viridis`) the reconstructions of the maps $\log(\delta)$ – first column – and τ – third column – obtained with the different methods in the case of a noise standard deviation of 10^{-2}, along with their error maps w.r.t. the GT (in colour map `red-white-blue`) which we denote as $\Delta\log(\delta)$ and $\Delta\tau$ – second and fourth column. The fifth column illustrates through a 2D histogram the correlation between the values of the two variables, along with their distributions in the form of 1D histograms placed on the top/right axes. We highlight with red circles the three $(\log(\delta), \tau)$ pairs corresponding to the three regions in the image. The colour map illustrates the number of occurrences (pixels) of a certain $(\log(\delta), \tau)$ pair in the map. While estimating the τ value for the star-shaped region represented a challenge for all the methods, N2N managed to accurately estimate the remaining values. In particular, the background value of $\log(\delta)$ is an example of the superior performance of our approach. Moreover, the error maps of REG highlight how the star-shaped region is surrounded by a staircase-like artefact, which is typical of TV-based regularisation. Note indeed that the solution estimated by N2N best maximises the correlation between the regions.

D. Test on Experimental Data

To assess the stability of our approach when dealing with real data, we follow the methodology outlined in [3]. Specifically, we apply it to a dataset obtained with three distinct acquisition protocols, each capturing experimental data from the same sample location. This setup leads to varying Experimental Times (ET) – approximately 4, 7 and 15 minutes – across the different data sets, allowing us to evaluate the robustness of our method under different conditions. Figure 3 shows the estimated lifetime maps obtained using PW, REG, and N2N. The maps generated with PW exhibit significant variations, while the estimates from REG and N2N are visually very similar. This consistency aligns with the notion that the parameter map being analyzed is intrinsic to the sample and should remain unaffected by the acquisition process. The absence of a reference map prevents us from directly quantifying the accuracy of our results. However, several observations can

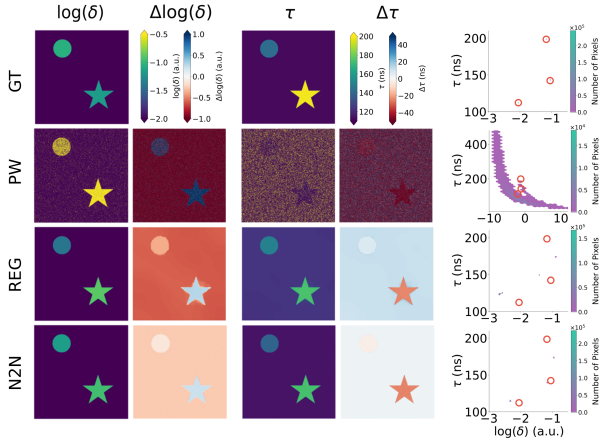

Fig. 2. Results on Data2 for noise standard deviation level of 10^{-2}. First column: $\log(\delta)$ maps. Second column: difference from the GT for the estimated $\log(\delta)$ maps. Third column: τ maps. Fourth column: difference from the GT for the estimated τ maps. Fifth column: correlation plots between the maps expressed in the form of 2D histogram.

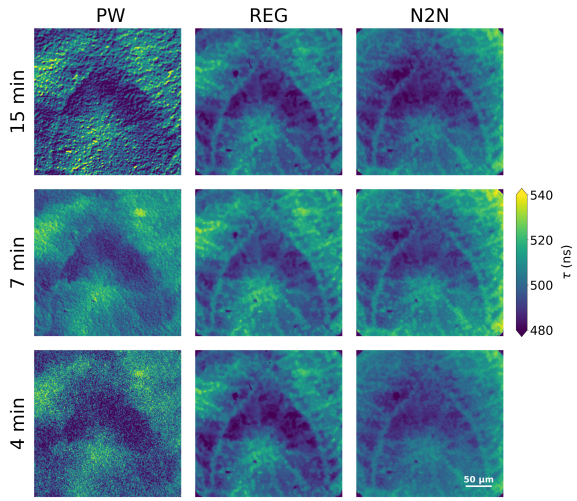

Fig. 3. Estimated τ maps on experimental data acquired on the same sample location with different acquisition protocols having different total experimental time – 15 minutes (first line), 7 minutes (second line) and 4 minutes (third line) – yielding varying SNR and possibly introducing acquisition artefacts.

be made. N2N produces more textured results, and as demonstrated in the previous section, it would have been capable of reconstructing piecewise constant distributions had they been present in the data. In contrast, the REG reconstruction may be affected by oversmoothing, a common issue with total variation regularisation when the regularisation parameter is too strong. Additionally, as shown in the previous section, the heuristic method used to set the regularisation parameter does not always identify the optimal value, which further supports the concern regarding over-smoothing with REG.

IV. Conclusions

We have presented a Noise2Noise approach for the quantitative reconstruction of lifetime parameter maps from noisy photoluminescence data in perovskite solar cells. By combining a physical log-linear modelling with an unsupervised training of a multi-branch shallow neural network, our method effectively addresses key challenges in PV characterisation: balancing estimation quality with minimising sample degradation while requiring limited parameter tuning. Our experimental results demonstrate that the proposed approach achieves comparable or superior performance to model-based methods without requiring extensive parameter tuning or clean reference data. The method provides robust performance across different data configurations and noise levels, preserving both global structures and fine details in the parameter maps as evidenced by quantitative metrics and more accurate distributions of the parameter values. Future work will test more expressive network architectures and tackle directly the original non-linear model. Our codes are available at `https://github.com/gbrscr/N2N4LifetimeMaps`.

References

[1] F. J. M. M. Nijsse, J.-F. Mercure, N. Ameli, F. Larosa, S. Kothari, J. Rickman, P. Vercoulen, and H. Pollitt, "The momentum of the solar energy transition," *Nat. Commun.*, vol. 14, no. 1, Oct. 2023.

[2] S. Cacovich, G. Vidon, M. Degani, M. Legrand, L. Gouda, J.-B. Puel, Y. Vaynzof, J.-F. Guillemoles, D. Ory, and G. Grancini, "Imaging and quantifying non-radiative losses at 23% efficient inverted perovskite solar cells interfaces," *Nat. Commun.*, vol. 13, no. 1, May 2022.

[3] G. Vidon, G. Scrivanti, E. Soret, N. Harada, E. Chouzenoux, J.-C. Pesquet, J.-F. Guillemoles, and S. Cacovich, "Rapid and noise-resilient mapping of photogenerated carrier lifetime in halide perovskite thin films," *Adv. Funct. Mater.*, vol. 34, no. 37, 2024.

[4] S. Lefkimmiatis, A. Roussos, P. Maragos, and M. Unser, "Structure tensor total variation," *SIAM J. Imaging Sci.*, vol. 8, no. 2, 2015.

[5] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2noise: Learning image restoration without clean data," in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80. PMLR, 2018.

[6] N. Moran, D. Schmidt, Y. Zhong, and P. Coady, "Noisier2noise: Learning to denoise from unpaired noisy data," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[7] A. Krull, T.-O. Buchholz, and F. Jug, "Noise2void - learning denoising from single noisy images," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[8] J. Batson and L. Royer, "Noise2self: Blind denoising by self-supervision," in *International Conference on Machine Learning*. PMLR, 2019.

[9] T. Huang, S. Li, X. Jia, H. Lu, and J. Liu, "Neighbor2neighbor: Self-supervised denoising from single noisy images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[10] K. Ji, W. Lin, Y. Sun, L.-S. Cui, J. Shamsi, Y.-H. Chiang, J. Chen, E. M. Tennyson, L. Dai, Q. Li, K. Frohna, M. Anaya, N. C. Greenham, and S. D. Stranks, "Self-supervised deep learning for tracking degradation of perovskite light-emitting diodes with multispectral imaging," *Nat. Mach. Intell.*, vol. 5, no. 11, 2023.

[11] Y. Quan, M. Chen, T. Pang, and H. Ji, "Self2self with dropout: Learning self-supervised denoising from single image," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[12] J. Lequyer, R. Philip, A. Sharma, W.-H. Hsu, and L. Pelletier, "A fast blind zero-shot denoiser," *Nat. Mach. Intell.*, vol. 4, no. 11, 2022.

[13] Y. Mansour and R. Heckel, "Zero-shot noise2noise: Efficient image denoising without any data," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2023.

[14] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 5, 2018.

[15] L. Calatroni, J. C. De Los Reyes, and C.-B. Schönlieb, "Infimal convolution of data discrepancies for mixed noise removal," *SIAM J. Imaging Sci.*, vol. 10, no. 3, pp. 1196–1233, 2017.

[16] G. Chierchia, N. Pustelnik, B. Pesquet-Popescu, and J.-C. Pesquet, "A nonlocal structure tensor-based approach for multicomponent image recovery problems," *IEEE Trans. Image Process.*, vol. 23, no. 12, 2014.