ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-WECD02

COMPUTING SPIN-POLARIZATION IN ELECTRON STORAGE RINGS BY MACHINE LEARNING VIA RANDOMIZED FOURIER NEURAL NETWORKS*

J. M. Agudelo Velez†, K. Heinemann‡, M. Motamed§, J.A. Ellison¶,
The University of New Mexico, Albuquerque, NM, USA

Abstract

Our work addresses the challenge of estimating spin polarization in high-energy electron and positron storage rings, such as the Electron Storage Ring (ESR) of the Electron-Ion Collider (EIC) at Brookhaven National Lab (BNL) and those in the electron/positron Future Circular Collider (FCC-ee) at CERN. We model the spin and orbital motion of particle bunches using the recently introduced spin-orbit Fokker-Planck (SOFP) equation, a linear time-evolution partial differential equation (PDE). In this paper, we propose a novel machine learning (ML) approach leveraging a randomized Fourier neural network (rFNN) framework, specifically designed to solve linear PDEs. We will discuss the SOFP highlight its relevance to spin polarization studies, and share preliminary results demonstrating the network's performance on the Poisson problem.

INTRODUCTION

In this article, we present our work in two parts: The modeling of orbital and spin dynamics of electron beams via the associated Ito stochastic differential equation (SDE) system of the SOFP in lab frame (LF) from [1], and the development of rFNNs for PDEs as applied to the multidimensional Poisson problem. The long-term goal of our project is to apply the rFNN theory to general Fokker-Planck (FP) systems and apply it to our SOFP in the beam frame (BF), where the BF is defined in [2]. It is thus useful to view our project as two independent problems, one in beam dynamics and the other in ML, with the understanding that we will eventually apply the theory from the latter to the former.

With regard to the beam dynamics portion of our project, the history of our method begins in 1975 with the derivation of Debenev and Kondratenko's (DK) orbital FP and Bloch equations in [3] for phase-space and polarization densities, respectively. This comprehensive model takes into account the Thomas-BMT spin precession effect, the Sokolov-Ternov effect along with the Baier-Katkov correction, and the kinetic polarization and radiative depolarization effects. Though comprehensive, the complexity of the orbital FP and Bloch equations makes it so that in the fifty years since the models introduction, it has yet to be fully implemented. A workaround reduced Bloch model has been implemented; it however only accounts for the radiative depolarization effect. Hence a full realization of the DK model is desired, as it will give a more precise estimate for the phase-space and polarization densities for high-energy accelerators. Fortunately, via the work done in [4] and [1], we may replace the problem of solving the two PDEs from [3] by solving the SOFP from [1]. The SOFP has an associated system of SDEs which we transform into the BF defined in [2].

From the standpoint of accelerator physics, the motivation of our project is to solve the SOFP. We aim to extend the theory developed in [5] on rFNNs to linear time evolution PDEs, such as the SOFP. This task is also motivated from the standpoint of applied mathematics where we seek to leverage the qualities of the rFNN approach in [5] to ideally build a serviceable solver for the FP. Specifically, we highlight a training methodology unrelated to traditional global optimization schemes seen in standard feed-forward networks. This is a "block-by-block" training approach that theoretically guarantees convergence depending on the type of target function.

ELECTRON BEAM DYNAMICS

The two PDEs seen in [3] have been transformed to the single 9-D SOFP in LF, as discussed in detail in [6]. It is preferable to analyze the SOFP in a BF based on the Frenet-Serret formalism on a design orbit, where analysis is easier and effects are more clearly seen, as is standard in accelerator physics. In addition, transforming SDE systems like ours is more convenient than transforming the corresponding FP equation. Thus, we move from the LF SOFP to the equivalent LF SDE system (written below) and do the transformations to BF. The BF SOFP follows immediately from the BF SDE and is ready for our analytical and numerical studies. We note that the LF SDE is transformed through several steps, as was outlined in the Handbook article for the orbital dynamics [2]. What we do now is an extension to include the spin dynamics. Further details into these transformations may be found in the slides accompanying this paper in these proceedings.

The LF System of SDEs

In the LF, the SDE system for position $\vec{R}$, momentum $\vec{\Pi}$, and spin $\vec{\Xi}$ of a single particle are given by Eq. (1)-(3) where $Q = (\vec{R}, \vec{\Pi})$. Here e, m denote the charge and mass of the electron, γ is the Lorenz factor, $\vec{E}, \vec{B}$ are the electric and magnetic fields due to the accelerator, $\vec{\mathcal{F}}_{rad}$ encompasses the radiation forces on the electron, $\vec{\mathcal{B}}_{\vec{\pi}}, \vec{\mathcal{B}}_{\vec{\xi}}$ together give kinetic polarization. $\vec{\Omega}$ gives the Thomas-BMT effect and

* Work supported by U.S. Department of Energy, Office of Science, under Award Numbers DE-SC0018008 and DE-SC0025476
† joseagudelo@unm.edu
‡ heineman@math.unm.edu
§ motamed@unm.edu
¶ ellison@math.unm.edu

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-WECD02

$\vec{\mathscr{W}}_{rad}$ gives the Sokolov-Ternov effect with the Baier-Katkov correction. Lastly ν is the white-noise process. From [1]

$$\frac{d\vec{R}}{dt} = \frac{\vec{\Pi}}{m\gamma(\vec{\Pi})}\,, \tag{1}$$

$$\frac{d\vec{\Pi}}{dt} = e\left(\vec{E}(t,\vec{R}) + \frac{\vec{\Pi}\times\vec{B}(t,\vec{R})}{m\gamma(\vec{\Pi})}\right) + \vec{\mathscr{F}}_{rad}(t,Q) + \vec{\mathscr{B}}_{\vec{\pi}}(t,Q)\nu(t)\,, \tag{2}$$

$$\frac{d\vec{\Xi}}{dt} = \vec{\Omega}(t,Q)\times\vec{\Xi} + \vec{\mathscr{W}}_{rad}(t,Q,\vec{\Xi}) + \vec{\mathscr{B}}_{\vec{\xi}}(t,Q)\nu(t) \tag{3}$$

We note that the terms $\vec{\mathscr{W}}_{rad}, \vec{\mathscr{B}}_{\vec{\xi}}$ were introduced in [1], so they are ignored in conventional Monte-Carlo spin tracking. In [1] it was also shown that if either piece of $\vec{\mathscr{B}}_{\vec{\pi}}, \vec{\mathscr{B}}_{\vec{\xi}}$ is neglected, the entire kinetic polarization effect is lost.

The LF Spin-Orbit Fokker-Planck Equation

The SDE system of (1)-(3) can be written as $dY/dt = \mathscr{D}(t,Y) + \mathscr{B}(t,Y)\nu(t)$ where $Y = (\vec{R}, \vec{\Pi}, \vec{\Xi})$ and $\mathscr{D}, \mathscr{B}$ are the "drift" and "diffusion" parts of the SDEs. Then, following 4.3.41 and 4.3.42 p. 95 in [7], the FP equation for the probability density function $\mathscr{P}$ of the random variable Y is

$$\partial_t\mathscr{P} = -\sum_i \partial_i\left[\mathscr{D}_i(t,y)\mathscr{P}\right] + \frac{1}{2}\sum_{i,j}\partial_i\partial_j\left\{\left[\mathscr{B}(t,y)\mathscr{B}^T(t,y)\right]_{i,j}\mathscr{P}\right\}\,,\ y\in\mathbb{R}^9. \tag{4}$$

This equation defines the SOFP.

RANDOMIZED FOURIER NEURAL NETWORKS FOR THE POISSON PROBLEM

The second task of our project is to design a method to solve time-evolution FP type PDEs and apply it to the SOFP. We extend the work done in [5], where the authors approximated some target function $u : \mathbb{R}^d \mapsto \mathbb{R}$ using a rFNN framework over target function data, to the setting where we are given data from a PDE instead. Specifically, [5] introduced a training algorithm for rFNNs which avoids global and gradient-based optimization while maintaining error control. In training, they employ the Metropolis-Hastings (MH) sampling procedure to iteratively train network layers, avoiding the need to commit to a network architecture before training.

In this article we present our current work on applying the theory in [5] to approximate the solution of the d-dimensional Poisson problem given by

$$\begin{cases}\Delta u(x) = f(x)\,,\ x\in D\subset\mathbb{R}^d\,,\\ u(x) = h(x)\,,\ x\in\partial D\subset\mathbb{R}^d\end{cases} \tag{5}$$

where D is an open set and ∂D is the boundary. Though the Poisson problem is simple relative to the FP, developing the theory here provides useful insights on how to do the same for the FP equation. It is also a more manageable setting in which to analyze potential numerical issues during implementation so as to have better foresight into the eventual implementation of an rFNN solver for the FP.

The Randomized Fourier Neural Network

We aim to approximate a target function u with a depth L rFNN of width W. We denote this network by u_L and define it recursively via

$$u_0(x) = \Re\sum_{j=1}^{W}\beta_{0,j}s(i\omega_{0,j}\cdot x) =: g_0(x;\Phi_0)\,, \tag{6}$$

$$\begin{aligned}u_\ell(x) &= u_{\ell-1}(x) + \Re\sum_{j=1}^{W}\beta_{\ell,j}s(i\omega_{\ell,j}\cdot x)\\ &\quad+\Re\sum_{j=1}^{W}\beta'_{\ell,j}s(i\omega'_{\ell,j}u_{\ell-1}(x))\\ &=: u_{\ell-1}(x) + g_\ell(x;\Phi_\ell) + g'_\ell(u_{\ell-1}(x);\Phi'_\ell)\end{aligned} \tag{7}$$

where $s(t) = e^{it}$. We refer to $\omega_{\ell,j}\in\mathbb{R}^d, \omega'_{\ell,j}\in\mathbb{R}$ and $\beta_{\ell,j}, \beta'_{\ell,j}\in\mathbb{C}$ as the frequency and amplitude parameters respectively. We denote $\omega_\ell = \{\omega_{\ell,j}\}_{j=1}^W$, $\omega'_\ell = \{\omega'_{\ell,j}\}_{j=1}^W$, $\beta_\ell = \{\beta_{\ell,j}\}_{j=1}^W$ and $\beta'_\ell = \{\beta'_{\ell,j}\}_{j=1}^W$. Often we refer to u_ℓ as the "ℓ^{th} block" of the rFNN. In the next section we demonstrate how to obtain good $\Phi_\ell = \{\omega_\ell, \beta_\ell\}$ and $\Phi'_\ell = \{\omega'_\ell, \beta'_\ell\}$ in a block-by-block basis for the Poisson problem.

The rFNN Trained with Poisson Data

For the Poisson problem given in Eq. (5), we generate two data sets: $(x_n, f(x_n))_{n=1}^{N_1}$ inside the domain D and $(y_n, h(y_n))_{n=1}^{N_2}$ on the boundary ∂D which we refer to as the "Poisson data". In this article, we consider $\{x_n\}_{n=1}^{N_1}$ and $\{y_n\}_{n=1}^{N_2}$ to be a uniform mesh on D and ∂D respectively.

Finding the amplitudes β_ℓ, β'_ℓ At each each u_ℓ define the loss function $\mathscr{L}_\ell : \mathbb{C}^{2W}\mapsto\mathbb{R}$ over β_ℓ, β'_ℓ by

$$\begin{aligned}\mathscr{L}_\ell(\beta_\ell,\beta'_\ell) &:= N_1^{-1}\sum_{n=1}^{N_1}|\Delta u_\ell(x_n) - f(x_n)|^2\\ &+N_2^{-1}\sum_{n=1}^{N_2}|u_\ell(y_n) - h(y_n)|^2 + \lambda\|\left(\beta_\ell,\beta'_\ell\right)\|_2^2\end{aligned} \tag{8}$$

where λ is a so-called Tikhonov regularization parameter. To find the amplitudes β_ℓ, β'_ℓ given the frequencies, we define a least-squares problem which aims to minimize $\mathscr{L}_\ell$ and returns the minimizers β_ℓ, β'_ℓ. For notational convenience, we take $\beta'_0 = \{0\}_{j=1}^W$ and thus ω'_0 is irrelevant. Theoretically, this problem is simple as it can be reduced to solving a $2W\times 2W$ linear system. However, in implementation, this has been challenging as the condition number of the system can be high, leading to instability in the results. We are currently focused on addressing this issue. Details will be published in [8].

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-WECD02

Finding the frequencies $\omega_\ell, \omega'_\ell$ To find the frequencies, we derive two optimal probability densities p_ℓ, q_ℓ from which to sample $\omega_\ell, \omega'_\ell$ respectively. For $\ell > 0$ define $r_\ell(x) = u(x) - u_{\ell-1}(x)$. We want to tune the function parameters to approximate $r_\ell(x) \approx g_\ell(x;\Phi_\ell) + g'_\ell(u_{\ell-1}(x);\Phi'_\ell)$. Here we make the assumption that there exists functions $\bar{r}_\ell, \bar{r}'_\ell$ such that $r_\ell = \bar{r}_\ell + \bar{r}'_\ell$ with $g_\ell(x;\Phi_\ell) \approx \bar{r}_\ell(x)$ and $g'_\ell(u_{\ell-1}(x);\Phi'_\ell) \approx \bar{r}'_\ell(u_{\ell-1}(x))$. This assumption is motivated by the idea that certain features of the target function r_ℓ are more efficiently represented with respect to the variable x, while others are more efficiently represented with respect to the variable $u_{\ell-1}$(x), and that an optimal split between $\bar{r}_\ell$ and $\bar{r}'_\ell$ needs to be learned by the network. For $\ell = 0$, take $\bar{r}_0 = u, \bar{r}'_0 = 0$.

Let $\hat{\cdot}$ denote the Fourier transform. We have used similar techniques as in [9] to prove the following bound in [8]:

$$\mathbb{E}_{\omega_\ell,\omega'_\ell}\left[\min_{\beta_\ell,\beta'_\ell} \mathscr{L}_\ell\right] \leq \frac{1}{W}\mathbb{E}_\omega\left[P_\ell\left(\|\omega\|_2\right)\left|\hat{\bar{r}}_\ell(\omega)\right|^2 / p_\ell^2(\omega)\right] + \frac{1}{W}\mathbb{E}_{\omega'}\left[Q_\ell\left(\omega'\right)\left|\hat{\bar{r}}'_\ell(\omega')\right|^2 / q_\ell^2(\omega')\right] \tag{9}$$

where P_ℓ, Q_ℓ are known polynomials. The goal is to derive optimal densities p_ℓ, q_ℓ which minimize the bound in Eq. (9). In [8] we derived these densities to be

$$p_\ell(\omega) = P_\ell^{1/2}\left(\|\omega\|_2\right)\left|\hat{\bar{r}}_\ell(\omega)\right| , \tag{10}$$

$$q_\ell(\omega) = Q_\ell^{1/2}\left(\omega'\right)\left|\hat{\bar{r}}'_\ell(\omega')\right| . \tag{11}$$

Sampling from these densities minimizes the bound in Eq. (9). This bound is to be interpreted as follows: If we sample $\omega_\ell, \omega'_\ell$ from these p_ℓ, q_ℓ and use the β_ℓ, β'_ℓ which minimize the loss, we expect $\mathscr{L}_\ell$ to be bounded by W^{-1}, though that may not hold. However, if we repeated this sampling over many attempts, the average $\mathscr{L}_\ell$ will be bounded by W^{-1}.

Of course, as stated before, we do not know $\bar{r}_\ell, \bar{r}'_\ell$ so that sampling from these densities is not directly possible. However, we provide a heuristic argument in [8] similar to one found in [5] which addresses this issue within the MH sampling procedure. In the slides of this paper, we provide a sort of "pseudo code" which details how to do the sampling without the exact p_ℓ, q_ℓ, though the full justification is left to [8].

Building a Deep rFNN

With now summarize the training procedure:

(I) Sample $\omega_0 \sim p_0$ with MH and minimize $\mathscr{L}_0$ to find β_0.

(II) Build u_0 using these parameters.

(III) Sample $\omega_1 \sim p_1, \omega'_1 \sim q_1$ with MH and minimize $\mathscr{L}_1$ to find β_1, β'_1.

(IV) Build u_1 using these parameters.

(V) Continue building blocks until within a desired error.

RESULTS

We here present an example of the Poisson problem solved using the rFNN framework. We consider the case where $d = 1, D = (0, \pi)$ and $\partial D = \{0, \pi\}$. The target function is $u(x) = 1.5\cos(4x) + 0.25\cos(10x) + 0.04\cos(25x)$ and so, by the method of manufactured solutions, we get $f(x) = \Delta u(x) = -24\cos(4x) - 25\cos(10x) - 25\cos(25x)$ and $h(0) = u(0) = 1.79, h(\pi) = u(\pi) = 1.71$. Here we

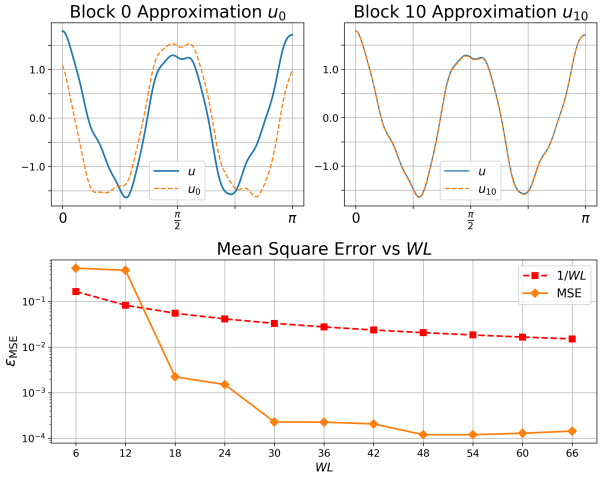

Figure 1: A rFNN approximation of $u(x) = 1.5\cos(4x) + 0.25\cos(10x) + 0.04\cos(25x)$ given Poisson data and the mean squared error (MSE) over a uniform grid of 10000 test points. The MSE is compared with a bound derived on the generalization error of rFNN approximations found in [10]

.

generated $N_1 = 2000$ uniformly spaced points between, but not including, the boundary $\{0, \pi\}$. We took a width of $W = 6$.

Figure 1 shows the MSE for the rFNN approximation of u, with width $W = 6$, trained on Poisson data and measured over 10000 test points in $[0, \pi]$. The rapid initial decay in error aligns with behavior we anticipate for smooth target functions [11]. The saturation beyond $WL = 30$ warrants further investigation, as it may stem from numerical limitations such as effects of regularization and sensitivity to conditioning, rather than from any fundamental barrier to approximation.

NEXT STEPS

The next step in our work is to stabilize the least-squares solver and adjust our implementation so that it may be run with GPU computing hardware as we will need large amounts of training data for higher dimensional problems. We will then turn our attention to the FP, derive optimal densities and the appropriate least-squares problem, and apply the rFNN framework. Simultaneous to all this, we complete the transformation of our SOFP into BF.

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-WECD02

REFERENCES

[1] K. Heinemann, D. Appelö, D. P. Barber, O. Beznosov, and J. A. Ellison, "The Bloch equation for spin dynamics in electron storage rings: Computational and theoretical aspects,"*Int. J. Mod. Phys. A*, vol. 34, no. 36, p. 1942032, Dec. 2019. doi:10.1142/s0217751x19420326

[2] J. A. Ellison, K. Heinemann, H. Mais, "Orbital Eigenanalysis for Electron Storage Rings," in Handbook of Accelerator Physics and Engineering, A. W. Chao, M. Tigner, H. Weise, F. Zimmermann (Eds.),3rd Edition, World Scientific, 2023, pp. 109-113.

[3] Y. S. Derbenev and A. M. Kondratenko, "Relaxation and equilibrium state of electrons in storage rings," *Sov. Phys. Dokl.*, vol. 19, pp. 438–441, 1975.

[4] K. Heinemann and D. P. Barber, "Spin transport, spin diffusion and Bloch equations in electron storage rings," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 463, no. 1-2, pp. 62–67, 2001. doi:10.1016/S0168-9002(01)00276-5

[5] O. Davis, G. Geraci, and M. Motamed, "Deep learning without global optimization by random Fourier neural networks," *SIAM J. Sci. Comput.*, vol. 47, no. 2, pp. C265–C290, 2025. doi:10.1137/24M1661777

[6] K. Heinemann, "Further inquiries into the kinetic approach of Derbenev and Kondratenko to spin-polarized beams in electron storage rings," research slides, University of New Mexico, 2025. To be published.

[7] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, 4th ed., Springer, Berlin, 2009.

[8] J. Agudelo Velez and M. Motamed, "Randomized Fourier networks for solving Poisson's problem," unpublished manuscript, University of New Mexico, 2025.

[9] A. Kammonen, J. Kiessling, P. Plechač, M. Sandberg, and A. Szepessy, "Adaptive Random Fourier Features with Metropolis Sampling," 2020,arXiv:2007.10683 [math.NA]. doi:10.48550/arXiv.2007.10683

[10] A. Kammonen, J. Kiessling, P. Plecháč, M. Sandberg, A. Szepessy, and R. Tempone, "Smaller generalization error derived for a deep residual neural network compared with shallow networks,"*IMA J. Numer. Anal.*, vol. 43, no. 5, pp. 2585–2632, Sep. 2022. doi:10.1093/imanum/drac049

[11] O. Davis, M. Motamed, O. Runborg, "Deep Fourier networks approximate jump discontinuities with spectral accuracy," 2025 preprint.