A Use of Gauss Law for the Process of Training Artificial Neural Networks

Bartosz Chaber
Faculty of Electrical Engineering
Warsaw University of Technology
Warsaw, Poland
bartosz.chaber@pw.edu.pl

Helena Jursza
Faculty of Electrical Engineering
Warsaw University of Technology
Warsaw, Poland
helena.jursza.dokt@pw.edu.pl

Jacek Starzyński
Faculty of Electrical Engineering
Warsaw University of Technology
Warsaw, Poland
jacek.starzynski@pw.edu.pl

***Abstract*—In this paper, the authors compare the training cost and predictive accuracy of various neural-network-based surrogate models trained on data generated using the finite-difference method.**

Traditionally, the loss function used to train artificial neural networks measures the discrepancy between the network's output and a reference solution. A neural network trained in this manner can approximate the electric potential; however, the resulting field may be non-physical (e.g., the electric field may violate Gauss's law)

Modern automatic differentiation frameworks enable the incorporation of additional terms into the loss function, such as constraints enforcing zero divergence of the field predicted by the neural network.

The results demonstrate that incorporating such physics-based constraints improves the physical consistency of the predicted fields while maintaining competitive accuracy. Furthermore, the proposed approach enhances the generalization capability of surrogate models beyond the training dataset.

***Index Terms*—Poisson equation, Gauss law, Physics-Informed Neural Networks**

I. Introduction

Optimization during the design process of electrical equipment often requires many evaluations of potential device designs. Surrogate models that sacrifice some accuracy during optimization but enable much faster evaluation are highly valued.

In this research, we present a preliminary surrogate model for electrical insulator chains. They are composed of various dielectric materials (e.g., ceramics, glass, and composite materials) and operate in high-voltage environments. Before an insulator chain is accepted for production, it often needs to be simulated to assess the electric field distribution around the device. Our surrogate model aims to replace the simulation in the optimization process.

In recent years, computer simulation suites have attempted to use *neural networks* to automatically build such surrogate models [4]. However, merely fitting a model to simulation data might not be enough, as this kind of model does not automatically capture physical laws and may extrapolate poorly. In our work, we test whether incorporating additional constraints during the training of such neural networks benefits the model [1], [2]. We focus on training a neural network on two-dimensional simulated data, where two terms drive the training process: a data error and a divergence residual arising from Gauss Law.

Our work deliberately leaves a few aspects unaddressed: we don't yet test the model's generalization properties, nor do we incorporate more complex geometries present in real-world electrical insulators. Our work focuses on the effect of incorporating physics into training.

II. Gauss law and dielectric interface

For two-dimensional electrostatics, the governing equation follows from Gauss law in differential form, together with the constitutive relation and the definition of the scalar electric potential φ:

$$\nabla \cdot \mathbf{D} = \rho_f, \qquad \mathbf{D} = \varepsilon(\mathbf{x})\,\mathbf{E}, \qquad \mathbf{E} = -\nabla\varphi, \tag{1}$$

where $\mathbf{D}$ is electric displacement field, ρ_f is a free charge density, ε is electric permittivity, $\mathbf{x} = (x_1, x_2)$ is the spatial position vector in 2D, and $\mathbf{E}$ is the electric field.

We can combine (1) into the generalized Poisson equation on a two-dimensional domain Ω

$$\nabla \cdot \big(\varepsilon(\mathbf{x})\,\nabla\varphi(\mathbf{x})\big) = -\rho_f(\mathbf{x}), \qquad \mathbf{x} \in \Omega \subset \mathbb{R}^2. \tag{2}$$

Expanding the divergence in (2) yields

$$\varepsilon\,\nabla^2\varphi + \nabla\varepsilon \cdot \nabla\varphi = -\rho_f. \tag{3}$$

A frequent simplification is to drop the second term and consider only $\nabla^2\varphi = -\rho_f/\varepsilon$. This approximation is valid only *inside homogeneous* subdomains.

Physically, the electric displacement field $\mathbf{D} = \varepsilon\mathbf{E}$ is discontinuous at dielectric interface perpendicular to the normal vector $\hat{\mathbf{n}}$ and produces a bound surface charge density σ_b

$$\big(\mathbf{D}_1 - \mathbf{D}_2\big) \cdot \hat{\mathbf{n}} = \sigma_b, \tag{4}$$

so that $\nabla \cdot \mathbf{D} = 0$ while $\nabla \cdot \mathbf{E} = \sigma_b/\varepsilon \neq 0$ on the dielectric boundaries, even in the absence of free charge. Therefore, in the surrogate model, we should enforce $\nabla \cdot \mathbf{D} = 0$, but only inside the homogeneous subdomains. The non-zero divergence on the dielectric boundaries will be addressed later in this paper.