Physics-informed neural networks viewpoint for solving the Dyson-Schwinger equations of quantum electrodynamics

Rodrigo Carmo Terin

King Juan Carlos University, Department of Physics,
Av. del Alcalde de Móstoles, 28933, Madrid, Spain

rodrigo.carmo@urjc.es

Abstract

Physics-informed neural networks (PINNs) are employed to solve the Dyson–Schwinger equations of quantum electrodynamics (QED) in Euclidean space, with a focus on the non-perturbative generation of the fermion's dynamical mass function in the Landau gauge. By inserting the integral equation directly into the loss function, our PINN framework enables a single neural network to learn a continuous and differentiable representation of the mass function over a spectrum of momenta. Also, we benchmark our approach against a traditional numerical algorithm showing the main differences among them. Our novel strategy, which is expected to be extended to other quantum field theories, is the first step towards forefront applications of machine learning in high-level theoretical physics.

1 Introduction and motivation

The predictions of the behavior of quantum field theories require an in-depth knowledge of the interactions among fields and particles on different energy scales. One of the theoretical methods to study these interactions is through the famous Dyson-Schwinger equations (DSEs), which are known for being an infinite set of integral equations fundamental to investigate the infrared (IR) region in QFTs, particularly in QED [1–3]. These integral equations are responsible for describing the dynamics of n-point Green's functions, which have an importance in investigating phenomena like dynamical mass generation, confinement, and the nature of phase transitions in quantum systems [4–11].

In terms of QED, the fermion and photon propagators' DSEs are fundamental in understanding how gauge invariance and chiral symmetry breaking manifest in different energy regimes [12–16]. It is important to mention that these integral equations form an infinite tower of equations, where each Green function is connected to higher-order ones, and therefore practical applications in general involve truncating the system. The Rainbow-Ladder approximation, for instance, is a commonly used truncation that simplifies the fermion-boson interaction vertex to its lowest-order term [17–21], a method often employed to study hadronic physics and the high-energy behavior of the quark-gluon vertex in quantum chromodynamics (QCD) [4, 22, 23].

In recent years, machine learning (ML) algorithms have been recognized as important tools for addressing difficult high-dimensional problems in physics. Among these approaches, the PINNs have emerged as particularly promising. These networks integrate the physical laws governing the system directly into the architecture of the neural network, using loss functions informed by the residuals of the differential equations that must be solved [24–26]. This method has been effectively applied to solve forward and inverse problems involving non-linear partial differential equations (PDEs), enabling high-precision solutions even in the presence of noisy and incomplete data [24, 25].

The wide-range applicability of PINNs covers some physical areas, e.g., the fluid dynamics, where they have been used to reconstruct flow fields from partial observations [27]. In cardiovascular flow modeling, these networks have successfully predicted arterial blood pressure from non-invasive 4D flow Magnetic Resonance Imaging (MRI) data [28], and in plasma physics, they have been used to uncover turbulent transport at the edge of magnetic confinement fusion devices [29]. They have also demonstrated utility in quantum chemistry, where they have been designed to handle high-dimensional quantum many-body problems, such as solving the Schrödinger equation [30–32].

Moreover, recent advances have expanded the scope of PINNs; for instance, the introduction of Bayesian physics-informed neural networks (B-PINN) has further improved the ability to quantify uncertainty in predictions, making these networks stronger in scenarios where data may be sparse or noisy [33]. In addition, the development of extended PINNs (XPINNs) has enabled more efficient training in parallel architectures [34]. Other extensions of these networks, such as deep operator networks (DeepONets), were responsible for learning mappings between infinite-dimensional function spaces, allowing the solution of operator learning problems [35]. There are more examples like multi-fidelity PINNs that combine data from different sources to increase model accuracy and to reduce computational costs [36]. Furthermore, applications in molecular simulations have led PINNs to accurately predict potential energy surfaces and simulate molecular dynamics [37].

In this work, we extend the use of PINNs to solve the DSEs for the fermion propagator in QED, particularly in Euclidean space. Unlike their conventional applications, which focus on solving PDEs, we employ these networks to tackle the integral equations that govern quantum field theories, such as those represented by DSEs, which are fundamentally more challeng-

ing due to their non-local nature and the need to account for all possible interactions in the quantum field [38, 39].

Our adaptation involves incorporating the integral equations of the DSEs directly into the loss function used during training. By doing so, we preserve the full structure and non-local nature of the DSEs. The neural networks approximate the wave function renormalization, the dynamical mass function, and are trained to minimize the discrepancy between their outputs and the solutions to the integral equations. This is achieved by first expressing the momentum variable in dimensionless form, and discretizing a desired interval using a logarithmic grid of N points. Then, the integrals are numerically approximated using the trapezoidal rule [40–42], which splits the integration domain into small intervals and approximates the area under the curve by summing the areas of the corresponding trapezoids. In each sub-interval, the quantity of interest is evaluated using our neural network, and these computations are directly incorporated into the loss function.

To confirm numerical stability and physical credibility, we implement positivity of the output through a softplus activation and improve resolution over several momentum scales by adding specific loss terms for intermediate, high, and ultraviolet (UV) regions. This multiscale supervision strategy enables the network to agree with numerical state-of-the-art traditional solutions even when $B(p^2)$ scales many orders of magnitude. In contrast to our previous approach relying only on base residuals, our novel strategy introduces log-scale constraints and perturbative tail matching to capture the full behavior of $B(p^2)$ on IR and UV regimes. As a consequence, the networks learn solutions that are consistent with both the mathematical form of the DSEs and the underlying physical principles they represent, without modifying the network architecture to include integral operators directly.

Finally, our research contributes to a effort to develop physics-informed learning methods capable of solving different problems in distinct scientific domains. By inserting intrinsic physical knowledge within PINNs, this study aims to provide efficient tools for investigating quantum field dynamics. The results could be a significant first step towards novel applications of machine learning in physics, from high-energy particle collisions to condensed matter systems.

This paper is organized as follows. In Sec. 2, we briefly review the Dyson–Schwinger equations in Euclidean space, presenting the fundamental expressions for the fermion and photon propagators. Then, in Sec. 3, we introduce the rainbow truncation framework and detail our neural network solution strategy, including the trapezoidal numerical integration method, the network architectures, and the training procedure (hyperparameters, loss function, etc.). Next, in Sec. 4, we present and analyze the results attained from our PINNs, focusing on the wave function renormalization and the dynamical mass function. Also, we compare our framework against a traditional numerical algorithm. Lastly, Sec. 5 concludes the manuscript, summarizing our findings and suggesting potential directions for future research.

2 Background of the Dyson-Schwinger equations in Euclidean space

The Dyson-Schwinger equations provide a non-perturbative framework for determining Green's functions in quantum field theories. In this work, we focus on the DSEs for the renormalized fermion in QED within Euclidean space by following the theoretical formalism developed in our previous work [43]. The transition from Minkowski space-time to Euclidean space is performed through a Wick rotation, which converts the Minkowski time component p_0 into an imaginary component $p_4 = ip_0$. This results in the Euclidean metric $p^2 = p_1^2 + p_2^2 + p_3^2 + p_4^2$, simplifying the analytic structure of the integrals.

In Euclidean space, the inverse renormalized fermion propagator $S_E^{-1}(p)$ can be written in terms of the wavefunction renormalization $A(p^2)$ and the mass function $B(p^2)$:

$$S_E^{-1}(p) = i\slashed{p} A(p^2) + B(p^2), \tag{1}$$

where $\slashed{p} = \gamma_\mu p_\mu$ and the Euclidean gamma matrices γ_μ satisfy the anticommutation relation $\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}$. Furthermore, we also have the ratio $M(p^2) = \frac{B(p^2)}{A(p^2)}$ that defines the momentum-dependent mass function of the fermion. It characterizes the non-perturbative dynamical generation of a momentum-dependent mass for the fermion, even if the bare mass m is zero. In general, both $A(p^2)$ and $B(p^2)$ are determined by solving the coupled system of integral equations derived from the DSEs.

To make the DSEs tractable, we use the rainbow approximation [4,18], which truncates the infinite set of coupled integral equations down to a manageable form while retaining essential non-perturbative features. The bare fermion-photon vertex is given by:

$$\Gamma^\mu(p,k) = \gamma^\mu. \tag{2}$$

The functions $A(p^2)$ and $B(p^2)$ satisfy the following coupled integral equations derived from the DSEs:

$$B(p^2) = m_{ph} + g_{ph}^2 \left[\Sigma_s(p^2) - \Sigma_s(\mu_F^2) \right], \tag{3}$$

$$A(p^2) = 1 - g_{ph}^2 \left[\Sigma_v(p^2) - \Sigma_v(\mu_F^2) \right], \tag{4}$$

Here, Σ_s and Σ_v are the scalar and vector components of the self-energy of the fermion. Moreover, m_{ph} is the renormalized physical mass of the fermion, and the coupling is defined as $g_{ph}^2 = 4\pi\alpha$, where α is the effective fine-structure constant in the theory. These quantities are defined at the renormalization scale μ_F^2, and their relation to the bare parameters is established through the renormalization conditions discussed later in this section.

The fermion self-energy in Euclidean space is expressed as:

$$\Sigma(p) = -g_{ph}^2 \int \frac{d^4k}{(2\pi)^4} \gamma_\mu S(p-k) \gamma_\nu D_{\mu\nu}(k). \tag{5}$$

The fermion propagator is:

$$S(p-k) = \frac{-A((p-k)^2)\left(i(\slashed{p} - \slashed{k})\right) + B((p-k)^2)}{A^2((p-k)^2)(p-k)^2 + B^2((p-k)^2)}, \tag{6}$$

while the photon propagator is decomposed as:

$$D_{\mu\nu}(k) = \left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right) D(k^2) + \xi \frac{k_\mu k_\nu}{k^4}, \tag{7}$$

which ξ is the gauge parameter. Using the gamma matrix identities:

$$\gamma_\mu \gamma_\nu = \delta_{\mu\nu} - i\sigma_{\mu\nu}, \qquad \{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}, \tag{8}$$

where $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$, we can now decompose the fermion self-energy $\Sigma(p)$ into its scalar and vector components. Since $\Sigma_s(p^2)$ and $\Sigma_v(p^2)$ are defined without including the coupling factor, we explicitly factor out g_{ph}^2:

$$\Sigma(p) = g_{ph}^2 \left[(i\slashed{p}) \Sigma_v(p^2) + \Sigma_s(p^2) \right]. \tag{9}$$

To determine $\Sigma_s(p^2)$ and $\Sigma_v(p^2)$, one performs the gamma matrix algebra, applies the integrals to the loop momentum, and identifies the coefficients of the scalar and vector structures. The scalar component $\Sigma_s(p^2)$ is attained from terms that do not involve gamma matrices (after all contractions), whereas $\Sigma_v(p^2)$ is extracted from terms proportional to $i\slashed{p}$.

Although the integral expressions for Σ_s and Σ_v are formally similar, the vector part $\Sigma_v(p^2)$ does not simplify easily due to the non-trivial momentum dependence of A, B, and D. A representative form of these integrals is:

$$\Sigma_s(p^2) = \int \frac{d^4k}{(2\pi)^4} \frac{\left[3D(k^2)B((p-k)^2) + \xi \frac{B((p-k)^2)}{k^2} \right]}{A^2((p-k)^2)(p-k)^2 + B^2((p-k)^2)}, \tag{10}$$

$$\Sigma_v(p^2) = \int \frac{d^4k}{(2\pi)^4} \frac{A((p-k)^2)(p-k)^2 \left[3D(k^2) + \xi \frac{1}{k^2} \right]}{\left[A^2((p-k)^2)(p-k)^2 + B^2((p-k)^2) \right] p^2}. \tag{11}$$

The photon propagator satisfies its own DSE:

$$\frac{1}{D(k^2)} = k^2 \left[1 - g_{ph}^2 \left(\Pi\left(k^2\right) - \Pi\left(\mu_B^2\right) \right) \right], \tag{12}$$

where the photon self-energy $\Pi(k^2)$ is given by:

$$\Pi(k^2) = \int \frac{d^4p}{(2\pi)^4} F(p^2)F((p-k)^2) \left[A(p^2)A((p-k)^2)(p^2 - p\cdot k) + 2B(p^2)B((p-k)^2) \right], \tag{13}$$

with

$$F(p^2) = \frac{1}{A^2(p^2)p^2 + B^2(p^2)}. \tag{14}$$

Renormalization in QED involves distinct renormalization constants: Z_2 for the fermion field, Z_m for the mass, and Z_A for the photon field. The physical mass and coupling at the scale μ_F relate to the bare parameters as:

$$m_{ph} = Z_2 m - g^2 \Sigma_s\left(\mu_F^2\right), \tag{15}$$

where $Z_2 = \frac{1}{1 - g^2 \Sigma_v(\mu_F^2)}$. The renormalized coupling is defined through:

$$g_{ph} = \frac{g}{Z_g}, \quad \text{with} \quad Z_g = \frac{1}{\sqrt{Z_A}}. \tag{16}$$

By specifying the appropriate renormalization conditions in μ_F, one fixes Z_2, Z_m, Z_A, and thus determines the physical parameters m_{ph} and g_{ph}. In this way, all renormalized quantities are manifested finite and well defined at the chosen renormalization scale.

3 The neural network approach

Our main focus in this work is to approximate the dynamical mass function of the fermion through a single PINN in the Landau gauge, where the gauge parameter ξ is equal null and compare our results with a traditional (conservative) numerical algorithm and with works [44, 45]. Under these conditions, the wave function renormalization is effectively 1 at leading

order, which simplifies the DSEs considerably. Consequently, the mass function is given directly by

$$M(p^2) = B(p^2). \qquad (17)$$

While more general scenarios could involve PINNs for $A(p^2)$, $B(p^2)$, and $D(k^2)$ simultaneously, our present effort focuses on learning $B(p^2)$, since the photon propagator is ill-defined in the above-mentioned gauge by using the DSEs formalism [43]. To do so we start from the full DSE as given in Eq. (3), we then apply a series of approximations to derive its simplified form. First, we employ the rainbow approximation by replacing the dressed fermion–photon vertex with its bare form Eq. (2), and approximating the photon propagator by its leading-order expression. Although such simplifications are standard in the study of dynamical mass generation, we note that the structure of our PINN architecture is, in principle, flexible enough to incorporate more sophisticated truncations, including gauge-independent quantities, as long as the corresponding integral expressions are properly implemented in the loss function. This potential extension is currently under investigation [46]. We then choose renormalization conditions that remove the subtraction terms, or equivalently, we work in the chiral limit where $m_{ph} = 0$; as a consequence, the scalar part simplifies to $B(p^2) \approx g_{ph}^2 \Sigma_s(p^2)$. Following this, the angular integrations in four-dimensional Euclidean space are performed, which reduce the full loop integral to a one-dimensional integral over k^2 with an integration kernel that naturally splits into two regions. Specifically, for $0 \leq k^2 \leq p^2$ the kernel is $K(p^2, k^2) = \dfrac{k^2}{p^2}$, while for $p^2 \leq k^2 \leq \kappa^2$ it becomes $K(p^2, k^2) = 1$. By incorporating these results, the scalar self-energy can be expressed as

$$\Sigma_s(p^2) \propto \int_0^{\kappa^2} \frac{B(k^2)}{k^2 + B^2(k^2)} K(p^2, k^2) \, dk^2 .$$

Finally, combining all these approximations leads to the simplified integral equation for $B(p^2)$:

$$B(p^2) = \frac{3\alpha}{4\pi} \left[\int_0^{p^2} \frac{B(k^2)}{k^2 + B^2(k^2)} \frac{k^2}{p^2} \, dk^2 + \int_{p^2}^{\kappa^2} \frac{B(k^2)}{k^2 + B^2(k^2)} \, dk^2 \right], \qquad (18)$$

where κ^2 is a fixed infrared scale (often set to 1 for convenience). In what follows, we outline our neural network structure, describe how we numerically approximate the integrals, and detail the training procedure that implements Eq. (18).

3.1 Neural network architecture

We apply a single neural network $model\_B$ that approximates $B(p^2)$. This network has:

1. An input layer of dimension 1 receiving $x = p^2/\kappa^2$ on a logarithmic grid (e.g., $x \in [10^{-12}, 1]$).

2. Three hidden layers, each with 64 neurons, using the hyperbolic tangent (tanh) activation function.

3. An output layer of dimension 1 with a softplus activation to ensure $B_{\text{pred}}(x) > 0$. That is, the predicted value is given by $B_{\text{pred}}(x) = \log\left(1 + \exp\left(\tilde{B}(x)\right)\right)$.

This architecture better approximate $B(p^2)$ over several orders of magnitude, especially in the UV region where the values become very small. The positivity constraint is automatically implemented by the softplus function, which also avoids potential numerical instabilities in the integral denominators $k^2 + B^2(k^2)$.

The network parameters (weights and biases) are initialized by default with Glorot uniform initialization [47], and the bias in the final layer can optionally be initialized to $\log(\exp(10^{-3}) - 1)$ to make $B \approx 10^{-3}$ at the start of training.

3.2 Numerical approximation of the integral

Since $A(p^2) = 1$ in Landau gauge at leading order, the self-energy expression simplifies to Eq. (18). We rewrite p^2 in dimensionless form $x = p^2/\kappa^2$, and similarly k^2/κ^2 for the integration variable. Thus, the integrals become

$$B(x) = \frac{3\alpha}{4\pi}\left[\underbrace{\int_0^x \frac{B(k)}{k+B^2(k)}\frac{k}{x}\,dk}_{\text{IR part}} + \underbrace{\int_x^1 \frac{B(k)}{k+B^2(k)}\,dk}_{\text{UV part}}\right]. \tag{19}$$

Here, $0 \le x \le 1$. In practice, we discretize $[10^{-12}, 1]$ using a logarithmic grid of N points, $\{x_1, x_2, \ldots, x_N\}$, and approximate the integrals using a trapezoidal rule. For the interval $[0, x_i]$, we sum over $\{x_j\}_{j=1}^i$; for $[x_i, 1]$, we sum over $\{x_j\}_{j=i}^N$. Symbolically,

$$\int_0^{x_i} f(k)\,dk \approx \sum_{j=1}^{i-1} \tfrac{1}{2}\big[f(x_j)+f(x_{j+1})\big]\big[x_{j+1}-x_j\big], \tag{20}$$

$$\int_{x_i}^1 g(k)\,dk \approx \sum_{j=i}^{N-1} \tfrac{1}{2}\big[g(x_j)+g(x_{j+1})\big]\big[x_{j+1}-x_j\big]. \tag{21}$$

In each trapezoid sub-interval, we evaluate $B(x_j)$ using our neural network. The coupling α is written as a constant in front of the integral, but it can also be treated as a parameter to be fitted.

To improve precision in specific momentum regions, we define masks over subsets of the domain: a "mid" region ($10^{-7} \le x \le 10^{-3}$), "high" ($10^{-3} \le x \le 10^{-1}$), and "very high" ($10^{-1} \le x \le 1$). These intervals will be used in the loss function to guide the network with which we refer to as regional supervision, this strategy consists of incorporating localized loss contributions into the total loss function, enabling the network to balance its performance over multiple orders of magnitude.

Moreover, a logarithmic comparison between the network prediction and the traditional solution is included in each region to better work out the dynamic range of $B(x)$. This will be fundamental to ensure that the PINN matches the benchmark results both qualitatively and quantitatively on all energy scales.

3.3 Training procedure and loss construction

Let us show now our PINN strategy, i.e., at both each epoch and discrete point x_i, we compute:

$$B_{\text{pred}}(x_i) = \texttt{model\_B}(x_i),$$

and also form

$$B_{\text{target}}(x_i) = \frac{3\alpha}{4\pi}\left[\int_0^{x_i} \frac{B(k)}{k+B^2(k)}\frac{k}{x_i}\,dk + \int_{x_i}^1 \frac{B(k)}{k+B^2(k)}\,dk\right],$$

through the trapezoidal sums. We then store up a mean-squared error (MSE) over all points x_i:

$$\text{Loss}_{\text{base}} = \frac{1}{N}\sum_{i=1}^N \big[B_{\text{pred}}(x_i) - B_{\text{target}}(x_i)\big]^2. \tag{22}$$

To improve learning over several orders of magnitude of $B(p^2)$, especially for values $\lesssim 10^{-6}$, additional loss terms were introduced. These include:

i. A mid-momentum range loss:

$$\text{Loss}_{\text{mid}} = \frac{1}{N_{\text{mid}}} \sum_{x_i \in [10^{-7}, 10^{-3}]} \left[\log_{10} B_{\text{pred}}(x_i) - \log_{10} B_{\text{num}}(x_i) \right]^2. \tag{23}$$

ii. A high-momentum loss in the interval $[10^{-3}, 10^{-1}]$:

$$\text{Loss}_{\text{high}} = \frac{1}{N_{\text{high}}} \sum_{x_i \in [10^{-3}, 10^{-1}]} \left[\log_{10} B_{\text{pred}}(x_i) - \log_{10} B_{\text{num}}(x_i) \right]^2. \tag{24}$$

iii. A perturbative UV tail constraint:

$$\text{Loss}_{\text{tail}} = \frac{1}{N_{\text{uv}}} \sum_{x_i \in [10^{-2}, 1]} \left[x_i B_{\text{pred}}(x_i) - C \right]^2, \tag{25}$$

where C is the mean $C = \langle x_i B_{\text{num}}(x_i) \rangle$ in the UV.

The final loss is a weighted sum of all components:

$$\text{Loss}_{\text{total}} = \text{Loss}_{\text{base}} + \lambda_{\text{mid}} \text{Loss}_{\text{mid}} + \lambda_{\text{high}} \text{Loss}_{\text{high}} + \lambda_{\text{tail}} \text{Loss}_{\text{tail}}, \tag{26}$$

with chosen weights $\lambda_{\text{mid}} = \lambda_{\text{high}} = 10^6$, and $\lambda_{\text{tail}} = 10^2$. Moreover, $N_{\text{mid}}, N_{\text{high}}$, and N_{uv} denote the number of grid points x_i lying in the intervals $[10^{-7}, 10^{-3}]$, $[10^{-3}, 10^{-1}]$, and $[10^{-2}, 1]$, respectively. These quantities correspond to the cardinality of the masked subsets used during loss computation, and serve to normalize the loss terms on different momentum regions, ensuring that no interval dominates due to a larger number of points. Also, this composite loss structure enables the network to reproduce the state-of-the-art traditional (conservative) numerical results. To accomplish that, we use the automatic differentiation capabilities of JAX [48] together with the neural network infrastructure of Equinox [49] and the optimization routines provided by Optax [50]. These libraries let efficient gradient-based optimization of the neural parameters. Furthermore it is worth mentioning that our learning strategy is not a standard supervised learning approach. Although some components of the loss function involve comparisons with results from the traditional numerical algorithm, these are used only as auxiliary regularization terms to stabilize training in specific momentum regions. The core of our method lies in minimizing the residual of the DSEs on the input domain, following the PINN paradigm. In this sense, the training falls under the category of residual-based physics-informed learning, rather than supervised or unsupervised learning in the classical machine learning taxonomy.

4 Discussion of results

In this section, we present our solutions attained from both a traditional numerical iterative solver and our PINN approach for $A(p^2)$ and $B(p^2)$ in Landau gauge.

4.1 Traditional and PINNs methods

First, we have implemented a traditional (conservative) numerical method to attain solutions of the DSEs in the above mentioned gauge, serving as a benchmark for our PINN-based approach. Using as references the works [44, 45], our simulations show how variations in the

coupling (instead of an explicit UV cutoff parameter) can affect the non-perturbative features of the theory. In this iterative method, the equation

$$B(p^2) \,=\, m + \frac{3\,\alpha}{4\pi}\left[\int_0^{p^2} \frac{B(k^2)}{k^2 + B^2(k^2)}\frac{k^2}{p^2}\,dk^2 + \int_{p^2}^{\kappa^2} \frac{B(k^2)}{k^2 + B^2(k^2)}\,dk^2\right] \tag{27}$$

is discretized over a logarithmic grid, with $x = p^2/\kappa^2$ ranging over the interval $[10^{-12}, 1]$. The iterative scheme initializes the function with $B = 10^{-3}$, and then updates it using the trapezoidal rule until convergence (with tolerance 10^{-7}). In parallel, we train our PINN using the same input domain and reference values for the coupling α, while introducing a loss function with additional terms targeting specific momentum intervals as written before.

Figure 1 displays the results for the mass function $B(p^2)$ obtained from both approaches for three values of the coupling $\alpha = 1.13$, 1.15, and 1.18, over the interval $p^2/\kappa^2 \in [10^{-12}, 1]$. The traditional results are shown with dashed lines, while the corresponding PINN predictions are shown with solid lines of matching color. This combined plot shows that our PINN reproduces the non-perturbative behavior on all momentum scales, from deep IR to UV, including features down to $B \sim 10^{-7}$ in the UV.

We also observe that the agreement is especially accurate in the intermediate and UV regimes, where our earlier PINN implementation plateaued. This is a direct consequence of introducing log-scale losses and regional supervision in the training.

As expected, increasing the coupling α increases the non-perturbative effects in the infrared region, yielding a larger value of $B(p^2)$ at low p^2. For all cases, the wave function renormalization remains fixed at $A(p^2) = 1$ in Landau gauge [44, 45], shown in Fig. 2, and used consistently in both methods.

We emphasize that the results shown in Figs. 1 and 2 were attained using the updated PINN framework presented in this work, which includes explicit loss weighting on logarithmic scales and targeted supervision in the mid, high, and very high momentum intervals. These additions were fundamental to suppress the PINN's tendency to plateau and to ensure alignment with the expected numerical behavior.

It is worth commenting that the references [44, 45] focus on the so-called rainbow approximation of the fermion DSE in 4D quantum electrodynamics, in which the photon propagator is kept quenched and the vertex is taken at its bare form γ^μ. Under this approximation, one
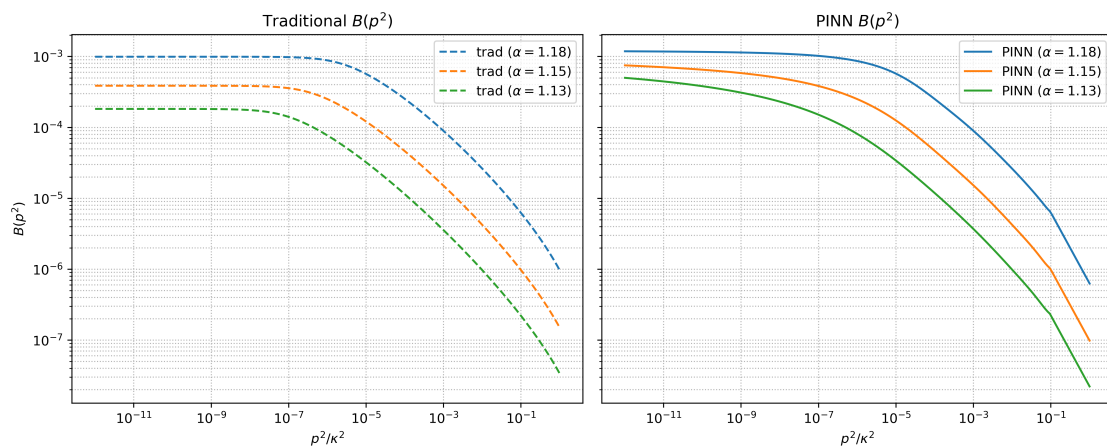
Figure 1: Comparison between traditional numerical solution (dashed) and PINN results (solid) for the fermion dynamical mass function $B(p^2)$ at $\alpha = 1.13$, 1.15, 1.18.
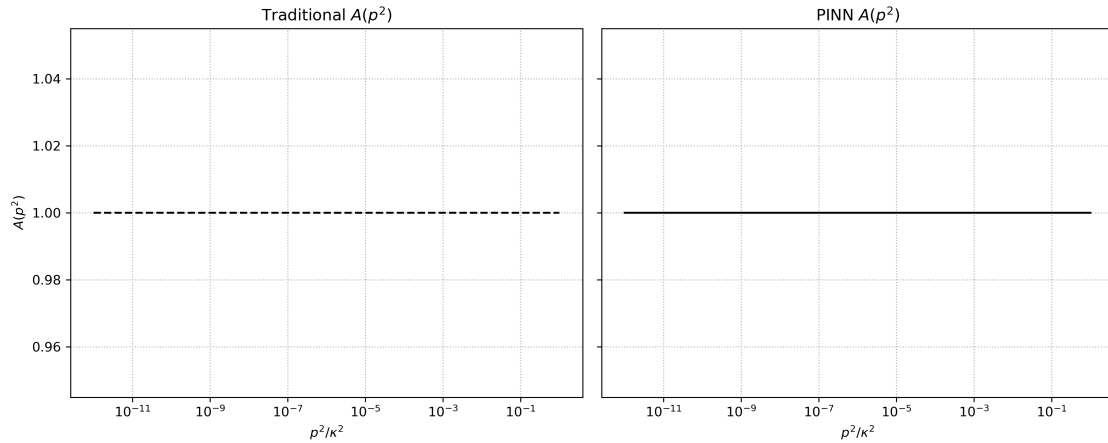
Figure 2: Fermion wave function renormalization $A(p^2)$, fixed to 1 in Landau gauge.

investigates the conditions under which dynamical mass generation emerges, typically identifying a critical coupling α_c above which $M(p^2)$ becomes nonzero. Additionally, the ultraviolet cutoff is implicitly fixed in those analyses, and is not treated as a free parameter; rather, the central interest lies in the gauge parameter ξ and whether $\alpha > \alpha_c$ leads to a spontaneously generated mass (obeying Miransky scaling near α_c).

Therefore, in order to adapt our forefront method to the traditional DSE literature, our current perspective is to vary the coupling α directly in the integral equation, thus tracking how the infrared behavior of $B(p^2)$ evolves with different α values. Our viewpoint lines up with the idea of exploring dynamical symmetry breaking through a machine learning approach, while remaining consistent with earlier numerical results of the rainbow approximation.

The agreement over all regions, i.e., IR, intermediate, and UV shows that PINNs, when properly guided with physical constraints and loss shaping, can match traditional numerical methods even for non-local integral equations. This is the first step towards the confirmation of the viability of PINNs as a robust method to tackle DSEs. Finally, although the figures do not explicitly incorporate a variable ultraviolet cutoff, one could indeed combine this framework with a regulator function $R(k^2)$ to examine the sensitivity of the solutions to changing Λ, which was done in our very first version of this work.

4.2 Overwiew by comparing the different perspectives

As it is already well-known we employed two distinct numerical viewpoints for solving the DSE for the dynamical mass function $B(p^2)$ in Landau gauge. The summary of both of them is given in tables (1) and (2) below:

Table 1: Characteristics of the traditional numerical method.

Characteristic	Traditional numerical method
Representation	Discrete values of $B(p^2)$ stored on a logarithmic grid from 10^{-12} to 1.
Architecture/Initialization	Started with a small constant for $B(p^2)$ on the grid.
Update mechanism	Iterative, point-by-point updates using trapezoidal integration until convergence.
Convergence criteria	Iterations stop when the maximum difference between successive updates is below 10^{-7} or after a maximum number of iterations.
Computational cost	Can become computationally expensive with a refined momentum grid, requiring careful tuning of initial guesses and iteration parameters.
Output	Produces a discrete solution defined only at the grid points.

Table 2: Features of our PINN framework.

Characteristic	PINN framework
Representation	Continuous function $B(x)$ with $x = p^2/\kappa^2$ represented by a neural network model.
Architecture/Initialization	Three hidden layers with 64 neurons each (using tanh activation), and a positive output required through softplus activation.
Update mechanism	Global gradient descent (Adam) minimizes an MSE loss that compares $B_{\text{pred}}(x_i)$ with a target $B_{\text{target}}(x_i)$ obtained through numerical integration.
Convergence criteria	Training stops when the optimizer stabilizes or a preset epoch limit is reached.
Computational cost	Loss function includes interval-based supervision and logarithmic penalties, enabling multi-scale resolution at all p^2 regions.
Output	Yields a smooth, continuous approximation of $B(p^2)$ valid over the entire domain.

5 Conclusion

In this work, we have used PINNs to tackle the DSEs for QED defined within the bounds of Euclidean space. In summary, we have shown how a single PINN can be applied to the integral equation for the dynamical mass function in Landau gauge. By discretizing the interval $[10^{-12}, 1]$ and applying a trapezoidal approximation, we trained a neural network to match $B(p^2)$ against its own integral expression. Our results confirm that this approach is capable of capturing the main non-perturbative features, namely dynamical mass generation at small momenta and a decreasing behavior for larger p^2.

Here, we introduced targeted loss components in specific momentum intervals, e.g., mid, high, and ultraviolet regimes combined with logarithmic constraints and perturbative tail matching. This multi-scale approach guarantees that the predicted mass function $B(p^2)$ agrees with traditional (conservative) numerical methods over the entire domain, even down to values $\sim 10^{-7}$. These refinements close the previous performance gap and allow the PINN to serve both as qualitative exploratory tool and as a quantitatively accurate method.

This work represents a first step towards applying modern neural architectures to faithfully reproduce traditional numerical results in non-perturbative QFTs, while retaining differentiability and flexibility for further generalizations. Future research may extend these algorithms within the quantum theoretical physics viewpoint to investigate the same problem in Minkowski space-time, explore more general truncation schemes and other gauges to study gauge (in)dependence explicitly, and employ our approach to different coupled systems, e.g., QCD or scalar Yukawa models, as well as to investigate other n-point Green functions. From a computer science perspective other neural network algorithms such as the neuroevolutionary framework [51,52], deep operator networks [35], and transformer-based architectures [53] could share the same ultimate goal of incorporating physical structure into machine learning models to approximate non-local integral equations like DSEs.

Finally, our work intends to extend the use of PINNs to solve difficult Dyson-Schwinger integral equations in QFTs. This approach addresses the missing links between machine learning and theoretical physics and offers a new window into the study of quantum field effects and, we hope, also contributes to the larger efforts to develop computational methods in the era of problems relevant to modern physics.

Acknowledgments

R. C. Terin thanks the useful suggestions and constructive comments from the SciPost Physics reviewers. The author is currently a lecturer in Physics at King Juan Carlos University (Madrid, Spain).

References

[1] F. J. Dyson, *The S matrix in quantum electrodynamics*, Phys. Rev. **75**, 1736 (1949), doi:10.1103/PhysRev.75.1736.

[2] J. Schwinger, *On the Green's functions of quantized fields. I*, Proc. Natl. Acad. Sci. **37**, 452 (1951), doi:10.1073/pnas.37.7.452.

[3] J. Schwinger, *On the Green's functions of quantized fields. II*, Proc. Natl. Acad. Sci. **37**, 455 (1951), doi:10.1073/pnas.37.7.455.

[4] C. D. Roberts and A. G. Williams, *Dyson-Schwinger equations and their application to hadronic physics*, Prog. Part. Nucl. Phys. **33**, 477 (1994), doi:10.1016/0146-6410(94)90049-3 [preprint doi:10.48550/arXiv.hep-ph/9403224].

[5] P. Maris and P. C. Tandy, *Bethe-Salpeter study of vector meson masses and decay constants*, Phys. Rev. C **60**, 055214 (1999), doi:10.1103/PhysRevC.60.055214 [preprint doi:10.48550/arXiv.nucl-th/9905056].

[6] R. Alkofer, *The infrared behaviour of QCD Green's functions confinement, dynamical symmetry breaking, and hadrons as relativistic bound states*, Phys. Rep. **353**, 281 (2001), doi:10.1016/S0370-1573(01)00010-2 [preprint doi:10.48550/arXiv.hep-ph/0007355].

[7] M. Q. Huber, *Nonperturbative properties of Yang-Mills theories*, Phys. Rep. **879**, 1 (2020), doi:10.1016/j.physrep.2020.04.004 [preprint doi:10.48550/arXiv.1808.05227].

[8] C. D. Roberts and S. M. Schmidt, *Dyson-Schwinger equations: Density, temperature and continuum strong QCD*, Prog. Part. Nucl. Phys. **45**, S1 (2000), doi:10.1016/S0146-6410(00)90011-5 [preprint doi:10.48550/arXiv.nucl-th/0005064].

[9] C. S. Fischer, *QCD at finite temperature and chemical potential from Dyson-Schwinger equations*, Prog. Part. Nucl. Phys. **105**, 1 (2019), doi:10.1016/j.ppnp.2019.01.002 [preprint doi:10.48550/arXiv.1810.12938].

[10] A. C. Aguilar, A. A. Natale and P. S. Rodrigues da Silva, *Relating a gluon mass scale to an infrared fixed point in pure gauge QCD*, Phys. Rev. Lett. **90**, 152001 (2003), doi:10.1103/PhysRevLett.90.152001 [preprint doi:10.48550/arXiv.hep-ph/0212105].

[11] O. Oliveira, T. Frederico and W. de Paula, *On the conformal limit of a QED-inspired model*, Eur. Phys. J. C **84**, 851 (2024), doi:10.1140/epjc/s10052-024-13231-5.

[12] J. M. Cornwall, *Dynamical mass generation in continuum quantum chromodynamics*, Phys. Rev. D **26**, 1453 (1982), doi:10.1103/PhysRevD.26.1453.

[13] D. C. Curtis, M. R. Pennington and D. A. Walsh, *On the gauge dependence of dynamical fermion masses*, Phys. Lett. B **249**, 528 (1990), doi:10.1016/0370-2693(90)91029-B.

[14] D. C. Curtis and M. R. Pennington, *Nonperturbative fermion propagator in massless quenched QED*, Phys. Rev. D **44**, 536 (1991), doi:10.1103/PhysRevD.44.536.

[15] W. Marciano and H. Pagels, *Quantum chromodynamics*, Phys. Rep. **36**, 137 (1978), doi:10.1016/0370-1573(78)90208-9.

[16] C. M. Bender, C. Karapoulitidis and S. P. Klevansky, *Underdetermined Dyson-Schwinger equations*, Phys. Rev. Lett. **130**, 101602 (2023), doi:10.1103/PhysRevLett.130.101602 [preprint doi:10.48550/arXiv.2211.13026].

[17] D. Atkinson, P. W. Johnson and P. Maris, *Dynamical mass generation in three-dimensional QED: Improved vertex function*, Phys. Rev. D **42**, 602 (1990), doi:10.1103/PhysRevD.42.602.

[18] P. Maris, C. D. Roberts and P. C. Tandy, *Pion mass and decay constant*, Phys. Lett. B **420**, 267 (1998), doi:10.1016/S0370-2693(97)01535-9 [preprint doi:10.48550/arXiv.nucl-th/9707003].

[19] C. S. Fischer, R. Alkofer, T. Dahm and P. Maris, *Dynamical chiral symmetry breaking in unquenched QED$_3$*, Phys. Rev. D **70**, 073007 (2004), doi:10.1103/PhysRevD.70.073007 [preprint doi:10.48550/arXiv.hep-ph/0407104].

[20] E. Rojas, A. Ayala, A. Bashir and A. Raya, *Dynamical mass generation in QED with magnetic fields: Arbitrary field strength and coupling constant*, Phys. Rev. D **77**, 093004 (2008), doi:10.1103/PhysRevD.77.093004 [preprint doi:10.48550/arXiv.0803.4173].

[21] A. Kızılersü, T. Sizer, M. R. Pennington, A. G. Williams and R. Williams, *Dynamical mass generation in unquenched QED using the Dyson-Schwinger equations*, Phys. Rev. D **91**, 065015 (2015), doi:10.1103/PhysRevD.91.065015 [preprint doi:10.48550/arXiv.1409.5979].

[22] G. Eichmann, H. Sanchis-Alepuz, R. Williams, R. Alkofer and C. S. Fischer, *Baryons as relativistic three-quark bound states*, Prog. Part. Nucl. Phys. **91**, 1 (2016), doi:10.1016/j.ppnp.2016.07.001 [preprint doi:10.48550/arXiv.1606.09602].

[23] M. Ding, C. Roberts and S. Schmidt, *Emergence of hadron mass and structure*, Particles **6**, 57 (2023), doi:10.3390/particles6010004 [preprint doi:10.48550/arXiv.2211.07763].

[24] 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*, J. Comput. Phys. **378**, 686 (2019), doi:10.1016/j.jcp.2018.10.045.

[25] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang and L. Yang, *Physics-informed machine learning*, Nat. Rev. Phys. **3**, 422 (2021), doi:10.1038/s42254-021-00314-5.

[26] I. E. Lagaris, A. Likas and D. I. Fotiadis, *Artificial neural networks for solving ordinary and partial differential equations*, IEEE Trans. Neural Netw. **9**, 987 (1998), doi:10.1109/72.712178 [preprint doi:10.48550/arXiv.physics/9705023].

[27] M. Raissi, A. Yazdani and G. Em Karniadakis, *Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations*, Science **367**, 1026 (2020), doi:10.1126/science.aaw4741.

[28] G. Kissas, Y. Yang, E. Hwuang, W. R. Witschey, J. A. Detre and P. Perdikaris, *Machine learning in cardiovascular flows modeling: Predicting arterial blood pressure from non-invasive 4D flow MRI data using physics-informed neural networks*, Comput. Methods Appl. Mech. Eng. **358**, 112623 (2020), doi:10.1016/j.cma.2019.112623 [preprint doi:10.48550/arXiv.1905.04817].

[29] A. Mathews, M. Francisquez, J. W. Hughes, D. R. Hatch, B. Zhu and B. N. Rogers, *Uncovering turbulent plasma dynamics via deep learning from partial observations*, Phys. Rev. E **104**, 025205 (2021), doi:10.1103/PhysRevE.104.025205 [preprint doi:10.48550/arXiv.2009.05005].

[30] D. Pfau, J. S. Spencer, A. G. D. G. Matthews and W. M. C. Foulkes, *Ab initio solution of the many-electron Schrödinger equation with deep neural networks*, Phys. Rev. Res. **2**, 033429 (2020), doi:10.1103/PhysRevResearch.2.033429 [preprint doi:10.48550/arXiv.1909.02487].

[31] L. Brevi, A. Mandarino and E. Prati, *Addressing the non-perturbative regime of the quantum anharmonic oscillator by physics-informed neural networks*, New J. Phys. **26**, 103015 (2024), doi:10.1088/1367-2630/ad8302.

[32] L. Brevi, A. Mandarino and E. Prati, *A tutorial on the use of physics-informed neural networks to compute the spectrum of quantum systems*, Technologies **12**, 174 (2024), doi:10.3390/technologies12100174 [preprint doi:10.48550/arXiv.2407.20669].

[33] L. Yang, X. Meng and G. Em Karniadakis, *B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data*, J. Comput. Phys. **425**, 109913 (2021), doi:10.1016/j.jcp.2020.109913 [preprint doi:10.48550/arXiv.2003.06097].

[34] A. D. Jagtap and G. Em Karniadakis, *Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations*, Commun. Comput. Phys. **28**, 2002 (2020), doi:10.4208/cicp.OA-2020-0164.

[35] L. Lu, P. Jin, G. Pang, Z. Zhang and G. Em Karniadakis, *Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators*, Nat. Mach. Intell. **3**, 218 (2021), doi:10.1038/s42256-021-00302-5.

[36] X. Meng and G. Em Karniadakis, *A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse PDE problems*, J. Comput. Phys. **401**, 109020 (2020), doi:10.1016/j.jcp.2019.109020 [preprint doi:10.48550/arXiv.1903.00104].

[37] L. Zhang, J. Han, H. Wang, R. Car and W. E., *Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics*, Phys. Rev. Lett. **120**, 143001 (2018), doi:10.1103/PhysRevLett.120.143001 [preprint doi:10.48550/arXiv.1707.09571].

[38] R. Fukuda and T. Kugo, *Schwinger-Dyson equation for massless vector theory and the absence of a fermion pole*, Nucl. Phys. B **117**, 250 (1976), doi:10.1016/0550-3213(76)90572-1.

[39] G. V. Efimov, *Non-local quantum theory of the scalar field*, Commun. Math. Phys. **5**, 42 (1967), doi:10.1007/BF01646357.

[40] R. L. Burden and J. D. Faires, *Numerical analysis*, Brooks/Cole, Pacific Grove, USA, ISBN 9780538733519 (2010).

[41] K. E. Atkinson, *An introduction to numerical analysis*, Wiley & Sons, Hoboken, USA, ISBN 9780471624899 (1989).

[42] E. Süli and D. F. Mayers, *An introduction to numerical analysis*, Cambridge University Press, Cambridge, UK, ISBN 9780511801181 (2003), doi:10.1017/CBO9780511801181.

[43] O. Oliveira, H. Lessa Macedo and R. Carmo Terin, *Looking at QED with Dyson-Schwinger equations: Basic equations, Ward-Takahashi identities and the two-photon-two-fermion irreducible vertex*, Few-Body Syst. **64**, 67 (2023), doi:10.1007/s00601-023-01846-5 [preprint doi:10.48550/arXiv.2204.04197].

[44] A. Kızılersü, T. Sizer, M. R. Pennington, A. G. Williams and R. Williams, *Dynamical mass generation in unquenched QED using the Dyson-Schwinger equations*, Phys. Rev. D **91**, 065015 (2015), doi:10.1103/PhysRevD.91.065015 [preprint doi:10.48550/arXiv.1409.5979].

[45] R. Williams, *Schwinger-Dyson equations in QED and QCD: The Calculation of fermion-antifermion condensates*, PhD thesis, University of Durham, Durham, UK (2007).

[46] R. Carmo Terin, *Physics-informed neural networks learning of gauge (in)dependent Dyson-Schwinger equations*, in preparation.

[47] X. Glorot and Y. Bengio, *Understanding the difficulty of training deep feedforward neural networks*, in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, Proc. Mach. Learn. Res. **9**, 249 (2010).

[48] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin and S. Wanderman-Milne, *JAX: Composable transformations of python+numpy programs* (2018), http://github.com/google/jax.

[49] P. Kidger and C. Garcia, *Equinox: Neural networks in JAX via callable PyTrees and filtered transformations*, (arXiv preprint) doi:10.48550/arXiv.2111.00254.

[50] *Optax: Gradient processing and optimization, in jax* (2020), https://github.com/deepmind/optax.

[51] K. O. Stanley, J. Clune, J. Lehman and R. Miikkulainen, *Designing neural networks through neuroevolution*, Nat. Mach. Intell. **1**, 24 (2019), doi:10.1038/s42256-018-0006-z.

[52] R. Carmo Terin, Z. González Arenas and R. Santana, *Identifying phase transitions in physical systems with neural networks: A neural architecture search perspective*, (arXiv preprint) doi:10.48550/arXiv.2404.15118.

[53] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, *Attention is all you need*, in *Advances in neural information processing systems 30*, Curran Associates, Red Hook, USA, ISBN 9781510860964 (2018) [preprint doi:10.48550/arXiv.1706.03762].