micromachines

MDPI

Article

Fourier Approximation of magMEMS Oscillations: Neural Network Space Handling

Piotr Skrzypacz [1], Arman Bolatov [1], Andrzej Dziedzic [2], Grant Ellis [1], Kaisar Tangirbergen [1], Bartosz Pruchnik [2,*] and Piotr Putek [2]

[1] Department of Mathematics, School of Sciences and Humanities, Nazarbayev University, Astana 010000, Kazakhstan; piotr.skrzypacz@nu.edu.kz (P.S.); arman.bolatov@nu.edu.kz (A.B.)
[2] Faculty of Electronics, Photonics and Microsystems, Wrocław University of Science and Technology, ul. Janiszewskiego 11/17, 50-372 Wrocław, Poland
* Correspondence: bartosz.pruchnik@pwr.edu.pl

Abstract

The initial value problem for a model of magnetic Micro-Electro-Mechanical System (magMEMS) with current-carrying conductors is considered. The model equation is a conservative MEMS oscillator due to the presence of the singular term corresponding to the magnetic force between two current-conducting wires. If the excitation parameter is below a certain threshold, the actuator oscillates. Fourier approximation of periodic solutions is enhanced by neural network space handling for both the limiting case where the geometric parameter approaches zero and the general case with arbitrary geometric parameter values. The method is validated through numerical comparisons with high-precision solutions, and its limitations are identified during the verification of experimental results. The findings can be useful for the design of magMEMS models.

Keywords: magMEMS; conservative singular oscillator; periodic solutions; dynamic pull-in; magMEMS switch; amplitude-frequency relation; neural networks; Fourier approximation

check for updates

Academic Editor: Ha Duong Ngo

Received: 12 September 2025
Revised: 27 October 2025
Accepted: 23 November 2025
Published: 28 November 2025

Citation: Skrzypacz, P.; Bolatov, A.; Dziedzic, A.; Ellis, G.; Tangirbergen, K.; Pruchnik, B.; Putek, P. Fourier Approximation of magMEMS Oscillations: Neural Network Space Handling. *Micromachines* **2025**, *16*, 1355. https://doi.org/10.3390/mi16121355

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The operation of micro-electro-mechanical systems (MEMS) is strongly influenced by the pull-in effect, which occurs when electrostatic or magnetostatic forces overcome the restoring elastic forces, leading to mechanical instability and collapse. This phenomenon has been extensively investigated for electrostatically actuated MEMS devices [1–3]. The classical mass–spring model describing the pull-in effect was first introduced by Nathanson et al. [4], and further analyses clarified that the dynamic pull-in voltage is generally lower than the static pull-in threshold [5,6]. Although the electrostatic pull-in problem is now well understood, the corresponding analysis for magnetic MEMS (magMEMS) actuated by current-carrying conductors remains less developed. Understanding the oscillatory behavior and pull-in thresholds in magMEMS devices is crucial for ensuring reliable operation, particularly because excessive actuation currents can cause structural failure or permanent magnetic saturation [7–11]. Previous studies [12–15] have addressed dynamic pull-in and periodic oscillations for magnetic actuation models, emphasizing the need for accurate and computationally efficient approximation methods to guide device design.

Over the past two decades, various analytical and numerical techniques have been developed to approximate periodic solutions in nonlinear oscillatory systems. Among the

most versatile methods is the Fourier–least squares method (FLSM) [16], which constructs periodic approximations by minimizing the residual of the governing nonlinear differential equation in the Fourier domain. This approach has been successfully applied to classical nonlinear oscillators, including the Duffing and Van der Pol systems [17], yielding accurate results even for strongly nonlinear regimes. Complementary to these classical methods, the rise of physics-informed neural networks (PINNs) [18] has opened a new avenue for integrating machine learning with differential equation modeling. Extensions of this framework, such as Fourier-feature-based PINNs [19], have demonstrated improved convergence for oscillatory solutions by embedding spectral information into the neural network input space. These developments motivate the combination of spectral and learning-based strategies for nonlinear problems with periodic responses.

The novelty of the present work lies in the formulation and application of a Fourier–neural network hybrid approach to compute approximate periodic solutions for magnetostatic MEMS oscillators with current-carrying wires. The proposed methodology combines the Fourier–least squares principle for residual minimization with neural network (NN) space handling, enabling efficient and high-accuracy computation of periodic orbits below the dynamic pull-in threshold. The approach is validated for both the limiting case of vanishing geometric parameter $\xi \to 0^+$ and for the general nonzero ξ configuration, providing quantitative insights into parameter dependencies and dynamic stability. This hybrid Fourier–neural framework enhances the analytical tractability and computational efficiency of nonlinear magMEMS modeling and can be extended to other classes of oscillatory systems with similar nonlinearities.

This paper is organized as follows. Section 2 introduces the mathematical model of the magMEMS device. In Section 3, we present the construction of approximate Fourier periodic solutions and the neural network-based determination of Fourier parameters. Section 4, in turn, extends the method to the general case and discusses validation and parameter dependencies, while Section 5 concludes the paper with remarks on the effectiveness and potential extensions of the proposed approach.

2. Mathematical Model of magMEMS Oscillator

The magnetic force between two current carrying conductors having length L and separated by the distance b (Figure 1) can be expressed as [12]:

$$F_L = \frac{i_1 i_2 \mu_0 \mu_r}{2\pi} \left[\frac{b}{\sqrt{b^2 + L^2}} + \frac{L^2}{b^2 \sqrt{\frac{L^2}{b^2} + 1}} - 1 \right]. \tag{1}$$

Then, the second Newton's law of motion implies

$$m\ddot{\widetilde{x}} + k_s \widetilde{x} - \frac{i_1 i_2 \mu_0 \mu_r}{2\pi} \left[\frac{(b - \widetilde{x})}{\sqrt{(b - \widetilde{x})^2 + L^2}} + \frac{L^2}{(b - \widetilde{x})^2 \sqrt{\frac{L^2}{(b-\widetilde{x})^2} + 1}} - 1 \right] = 0, \tag{2}$$

where the restoring force due to the presence of a spring or springs with constant k_s is

$$F_R = -k_s \widetilde{x},$$

and the attraction force between the current-carrying filaments due to the magnetic fields produced by the currents i_1, i_2 is given as

$$F_L = \frac{i_1 i_2 \mu_0 \mu_r}{2\pi} \left[\frac{(b - \widetilde{x})}{\sqrt{(b - \widetilde{x})^2 + L^2}} + \frac{L^2}{(b - \widetilde{x})^2 \sqrt{\frac{L^2}{(b-\widetilde{x})^2} + 1}} - 1 \right].$$

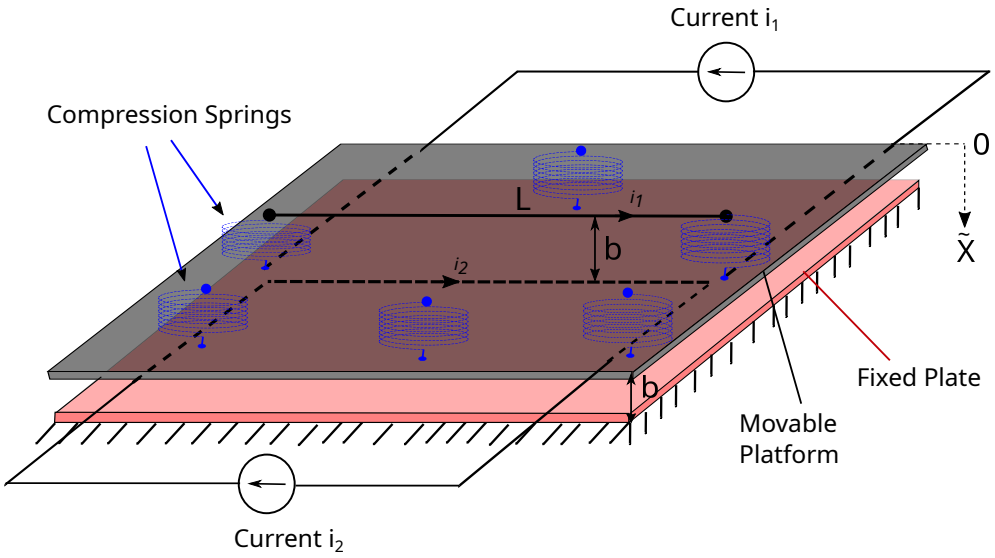

Figure 1. magMEMS with current-carrying filaments. The instantaneous distance between filaments is $b - \tilde{x}(t)$.

The governing differential Equation (2) does not include a damping term with $\dot{\tilde{x}}$ since we assume that the air resistance and the mechanical damping of the flexible platform are negligible. Let us introduce the dimensionless distance and time as follows

$$x = \frac{\tilde{x}}{b} \quad \text{and} \quad t = \tilde{t}\sqrt{\frac{k_s}{m}}, \tag{3}$$

respectively. The dimensionless model equation then becomes

$$\ddot{x} + x - K\left[\frac{\xi^2\,(1-x)}{\sqrt{\xi^2\,(1-x)^2+1}} + \frac{1}{(1-x)\sqrt{\xi^2\,(1-x)^2+1}} - \xi\right] = 0\,, \tag{4}$$

where the scaled excitation parameter K is given by

$$K = \frac{\mu_0 \mu_r i_1 i_2 L}{2\pi\, k_s\, b^2}\,, \tag{5}$$

and

$$\xi = \frac{b}{L}$$

denotes the scaled geometric parameter. We complement Equation (2) with zero initial conditions

$$x(0) = \dot{x}(0) = 0\,,$$

and assume that the currents in both wires are in the same direction or unidirectional, i.e., $K \geq 0$.

Multiplying both sides of Equation (4) by $\dot{x}$ and then integrating with respect to t leads to the conservation of the total energy

$$\begin{aligned}\mathcal{E}_{K,\xi}(t) = &\frac{1}{2}(\dot{x}(t))^2 + \frac{1}{2}x^2(t) \\ &- K\left[-\sqrt{\xi^2\,(1-x(t))^2+1} + \operatorname{artanh}\left(\frac{1}{\sqrt{\xi^2\,(1-x(t))^2+1}}\right) - \xi\,x(t)\right] \\ &- K\left[\sqrt{1+\xi^2} - \operatorname{artanh}\left(\frac{1}{\sqrt{1+\xi^2}}\right)\right],\end{aligned} \tag{6}$$

from which it follows that

$$\begin{aligned}(\dot{x}(t))^2 = &-x^2(t) \\ &+ 2K\left[-\sqrt{\xi^2\,(1-x(t))^2+1} + \operatorname{artanh}\left(\frac{1}{\sqrt{\xi^2\,(1-x(t))^2+1}}\right) - \xi\,x(t)\right] \\ &+ 2K\left[\sqrt{1+\xi^2} - \operatorname{artanh}\left(\frac{1}{\sqrt{1+\xi^2}}\right)\right].\end{aligned} \tag{7}$$

Let

$$\begin{aligned}f_{K,\xi}(s) = &-s^2 + 2K\left[-\sqrt{\xi^2\,(1-s)^2+1} + \operatorname{artanh}\left(\frac{1}{\sqrt{\xi^2\,(1-s)^2+1}}\right) - \xi\,s\right] \\ &+ 2K\left[\sqrt{1+\xi^2} - \operatorname{artanh}\left(\frac{1}{\sqrt{1+\xi^2}}\right)\right]\end{aligned} \tag{8}$$

be related to the righ-hand-side of Equation (7). The solution $x(t)$ is periodic if the corresponding phase portrait in the $(x, \dot{x})$ forms a closed curve. This holds true when the function $f_{K,\xi}(s)$ has a root in the interval $(0,1)$. In the critical case $K = K^*_\xi$, the function $f_{K,\xi}(s)$ has a double root at $s = A^*_\xi$. This condition is satisfied when

$$f_{K^*_\xi,\xi}(A^*_\xi) = 0 \quad \text{and} \quad \frac{\partial f_{K^*_\xi,\xi}}{\partial s}(A^*_\xi) = 0\,, \tag{9}$$

which results in the transcendental equation for A^*_ξ:

$$\begin{aligned}&\frac{A^*_\xi}{2}\left[\frac{\xi^2\,(1-A^*_\xi)}{\sqrt{\xi^2\,(1-A^*_\xi)^2+1}} + \frac{1}{(1-A^*_\xi)\sqrt{\xi^2\,(1-A^*_\xi)^2+1}} - \xi\right] \\ &= -\sqrt{\xi^2\,(1-A^*_\xi)^2+1} + \operatorname{artanh}\left(\frac{1}{\sqrt{\xi^2\,(1-A^*_\xi)^2+1}}\right) - \xi\,A^*_\xi \\ &\quad + \sqrt{1+\xi^2} - \operatorname{artanh}\left(\frac{1}{\sqrt{1+\xi^2}}\right).\end{aligned} \tag{10}$$

The value of A^*_ξ, representing the maximum deflection, can be obtained numerically for various values of the parameter ξ by solving Equation (10). Then, the corresponding dynamic pull-in threshold K^*_ξ can be calculated as follows

$$K^*_\xi = \left[\frac{\xi^2\,(1-A^*_\xi)}{\sqrt{\xi^2\,(1-A^*_\xi)^2+1}} + \frac{1}{(1-A^*_\xi)\sqrt{\xi^2\,(1-A^*_\xi)^2+1}} - \xi\right]^{-1} A^*_\xi \tag{11}$$

due to the second condition from Equation (9). In Figure 2, the effect of the geometry parameter ξ on both the maximum deflection A^*_ξ of the flexible platform and the dynamic pull-in threshold K^*_ξ is illustrated. The maximum deflection A^*_ξ decreases as ξ increases due to a weakening Lorentz force. For the same reason, the dynamic pull-in threshold K^*_ξ increases with increasing ξ.

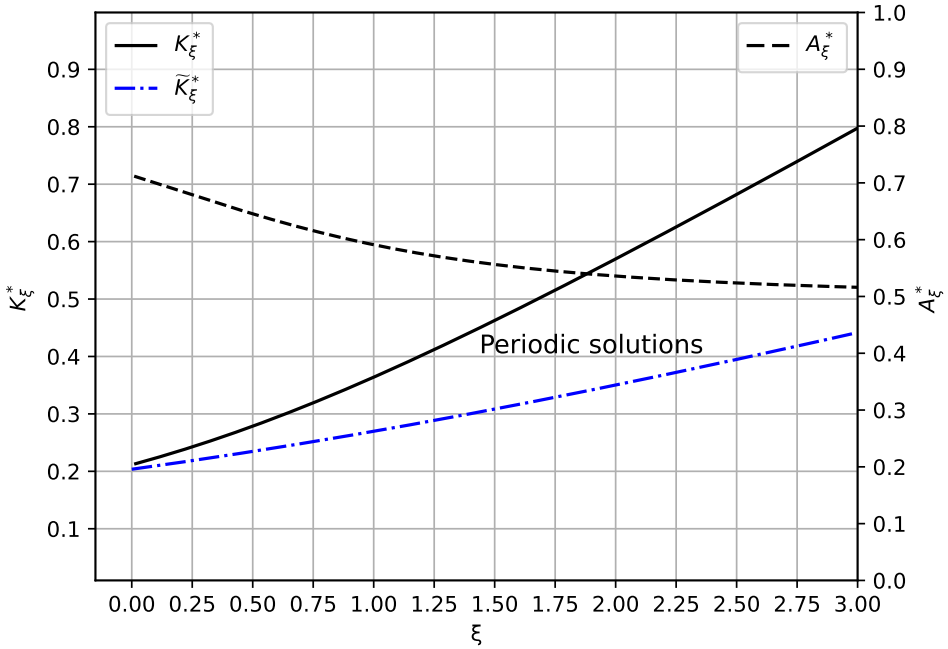

Figure 2. Maximum deflection A^*_ξ of flexible part and dynamic pull-in threshold K^*_ξ along with its lower bound $\widetilde{K}^*_\xi$. Pairs (ξ, K) below the separatrix $K^* = K^*_\xi$ (solid line) lead to periodic solutions.

When the motion of the filament is caused by the magnetic field of an infinite current-carrying conductor, i.e., $\xi \to 0^+$, the zero initial value problem for Equation (4) becomes

$$\begin{aligned} \ddot{x} + x - \frac{K}{1-x} &= 0\,, \\ x(0) = \dot{x}(0) &= 0\,. \end{aligned} \tag{12}$$

Multiplying (12) by $\dot{x}$ and integrating with respect to time, we get the conservation of energy

$$\mathcal{E}_{K,0}(t) = \frac{1}{2}(\dot{x}(t))^2 + \frac{1}{2}x^2(t) + K\ln(1-x(t)) \equiv 0 \tag{13}$$

for all $t \geq 0$, and the corresponding function $f_{K,\xi}(s)$ defined in Equation (8) is given in the limit case $\xi = 0$ by $f_{K,0}(s) = -s^2 - 2K\ln(1-s)$. The initial value problem by Equation (12) has periodic solutions if $K < K^*_0$, otherwise it has pull-in solutions. The dynamic pull-in threshold K^*_0 and the maximum deflection A^*_0 are given by [13,15]

$$K^*_0 = -\frac{2W_{-1}\left(-e^{-1/2}/2\right) + 1}{4W^2_{-1}\left(-e^{-1/2}/2\right)} = 0.2036321888\ldots\,, \tag{14}$$

and

$$A^*_0 = \frac{1+\sqrt{1-4K^*_0}}{2} = 0.71533\ldots\,, \tag{15}$$

which can be deduced from Equation (9). In Equation (14), $W_{-1} : [-1/e, 0) \to [-1, -\infty)$ denotes the second branch of the LambertW function which is defined as a solution

$y(x) \leq -1$ to the transcendental equation $ye^y = x$ for $x \in [-1/e, 0)$. The maximum deflection A satisfies for $0 < K < K_0^*$ the transcendental equation

$$A^2 + 2K\ln(1-A) = 0\,. \tag{16}$$

The lower bound for the separatrix $K(\xi) = K_\xi^*$ can be inferred from Equation (11) as follows

$$\begin{aligned} K_\xi^* &\geq \frac{(1-A_0^*)A_0^*\sqrt{\xi^2(1-A_0^*)^2+1}}{\xi^2(1-A_0^*)^2+1-\xi(1-A_0^*)\sqrt{\xi^2(1-A_0^*)^2+1}} \\ &= (1-A_0^*)A_0^* + (1-A_0^*)^2A_0^*\xi + \frac{(1-A_0^*)((A_0^*)^2-2A_0^*+1)A_0^*}{2}\xi^2 + \mathcal{O}(\xi^4) \end{aligned} \tag{17}$$

due to the fact that $A_\xi^* \geq A_0^* = 0.715331863\ldots$, see Equation (15) and Figure 2. Notice that $K_0^* = (1-A_0^*)A_0^*$ according to the second condition in Equation (9). The lower bound $\widetilde{K}_\xi^*$ from Equation (17) confirms that the dynamic pull-in threshold K_ξ^* increases with increasing geometric parameter ξ as shown in Figure 2.

3. Approximate Periodic Solutions

3.1. Fourier Approximation

The differential equation in (12) can be rewritten as

$$(1-x)\ddot{x} - x^2 + x = K\,. \tag{18}$$

In order to obtain an approximate periodic solution of this conservative singular oscillator, we employ a Fourier approximation method. The key idea is to project the dynamics onto a subspace spanned by a finite number of basis functions that inherently satisfy the initial conditions. In our case, we consider the following ansatz:

$$\hat{x}(t) = \alpha\sin^2(\omega t) + \beta\sin^2(2\omega t) + \gamma\sin^2(3\omega t)\,, \tag{19}$$

which guarantees that

$$\hat{x}(0) = 0 \quad \text{and} \quad \dot{\hat{x}}(0) = 0\,.$$

In Equation (19), the parameter ω denotes the frequency. The parameter vector with four parameters

$$\boldsymbol{\theta} = [\omega, \alpha, \beta, \gamma]\,,$$

which include both a dominant frequency component and amplitude coefficients for the first three harmonics, are then determined by enforcing that the ansatz approximately satisfies the governing equation. After differentiating the ansatz and substituting into Equation (18), we obtain

$$\begin{aligned} &2\omega^2\Big(\alpha\sin^2(\omega t) + \beta\sin^2(2\omega t) + \gamma\sin^2(3\omega t) - 1\Big) \\ &\quad\times\Big[2\alpha\sin^2(\omega t) - \alpha - 32\beta\sin^4(\omega t) + 32\beta\sin^2(\omega t) - 4\beta + 18\gamma\sin^2(3\omega t) - 9\gamma\Big] \\ &\quad+\alpha\sin^2(\omega t) + \beta\sin^2(2\omega t) + \gamma\sin^2(3\omega t) \\ &\quad-\Big(\alpha\sin^2(\omega t) + \beta\sin^2(2\omega t) + \gamma\sin^2(3\omega t)\Big)^2 = K. \end{aligned} \tag{20}$$

Denoting the left-hand side of Equation (20) by $f_{\boldsymbol{\theta}}(t)$, we define a loss function as the mean squared error

$$G(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N}(f_{\boldsymbol{\theta}}(t_i) - K)^2,$$

where in our implementation, the residual $f_{\boldsymbol{\theta}}(t_i) - K$ is computed at $N = 100$ equidistant points t_i in the time interval $[0, 5\pi]$. This range was chosen because it covers several periods of $\hat{x}(t)$. Our goal is to minimize $G(\boldsymbol{\theta})$ over the parameter vector $\boldsymbol{\theta}$. This is achieved by an unconstrained gradient descent procedure.

3.2. Minimization Strategy

For each fixed value of K, the initial guess for the parameters was chosen as

$$\boldsymbol{\theta} = [1,\ K,\ K/2,\ K/4].$$

This choice is motivated by physical intuition: for small values of the excitation parameter K, the amplitude of the oscillations is expected to be small and to scale approximately linearly with K, while the dominant oscillatory behavior remains near the natural (or fundamental) frequency of the linear part of the system (here, approximately unity). Although simple stochastic gradient descent [20] or SGD with momentum can be used to update $\boldsymbol{\theta}$, our experiments showed that the Adam optimizer [21] converged significantly faster and more robustly. The Fourier residual is computed using PyTorch's (PyTorch Stable (2.9.0) release) automatic differentiation [22], and the parameter vector is updated iteratively over 1000 iterations. Details of the optimization algorithm are as follows:

1. **Initialization:** For a given K (with $K < K^*$), initialize the parameter vector
$$\boldsymbol{\theta} = [1,\ K,\ K/2,\ K/4]$$
and generate N equidistant time points $t_i \in [0, 5\pi]$.
2. **Loss Evaluation:** Compute the Fourier residual $f_{\boldsymbol{\theta}}(t_i)$ at each t_i and evaluate the mean squared error loss
$$G(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N}(f_{\boldsymbol{\theta}}(t_i) - K)^2.$$
3. **Gradient Update:** Use the Adam optimizer to update $\boldsymbol{\theta}$ by backpropagating the gradient of $G(\boldsymbol{\theta})$.
4. **Iteration:** Perform the update for a fixed number of iterations (e.g., 1000 iterations). At each iteration, compare the current loss with the best loss recorded so far and save the corresponding $\boldsymbol{\theta}$ if the current loss is lower.
5. **Output:** After the final iteration, output the parameter vector $\boldsymbol{\theta}$ that achieved the minimum loss along with its corresponding loss value.

3.3. Comparison with High-Precision Runge-Kutta ODE Solver

We validate each approximate solution by comparing it to a high-precision Runge-Kutta (RK4) integration of

$$(1 - x)\,\ddot{x} \;-\; x^2 \;+\; x \;=\; K,$$

with Δt chosen as 10^{-6} and with a precision of 10^{-8}. Figure 3 illustrates these comparisons for three representative values of K. The solid lines show the Fourier-based approximations, whereas the dashed lines represent the RK4 solution sampled at several time steps. We observe excellent agreement.

To further quantify the accuracy, we compute the absolute error between the Fourier approximation and the RK4 solution, sampled on the same temporal grid, and plot this in Figure 4. We find that an error remains small for ω and α, with slightly larger deviations

on β, γ. The Fourier-based method captures the behavior of the larger modes, but it underestimates the amplitudes of the highest-frequency terms somewhat.

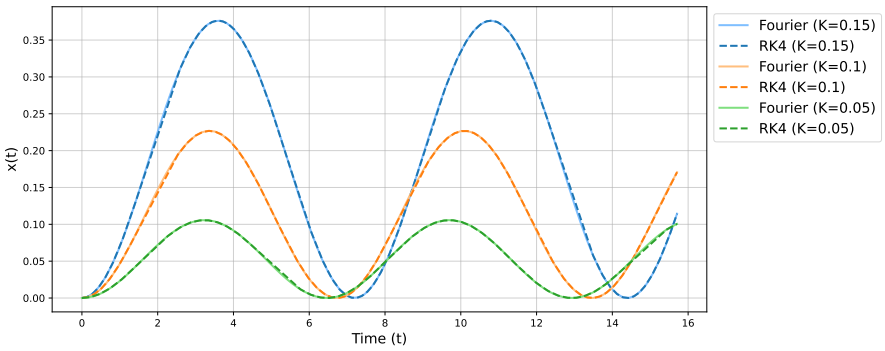

Figure 3. Comparison of the Fourier approximation (solid lines) with a high-precision Runge-Kutta solution (dashed lines) for three different values of K.

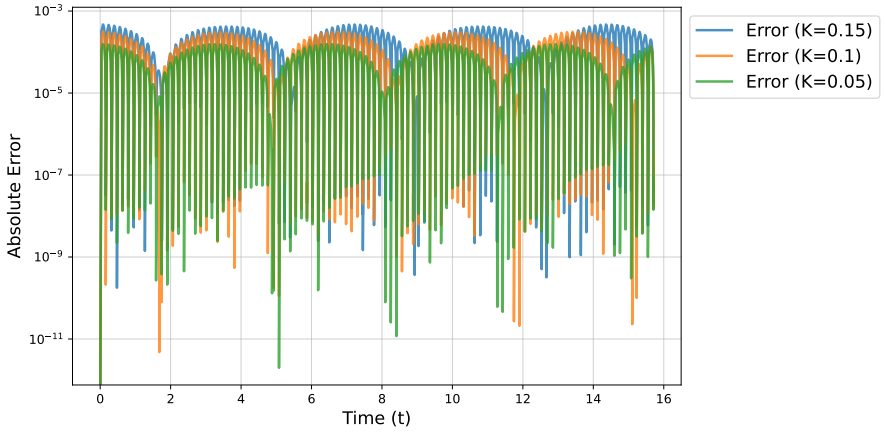

Figure 4. Absolute error between the Fourier-based approximation and the Runge-Kutta solution for different values of K. Note the logarithmic scale on the vertical axis.

3.4. Neural Network Fitting of Amplitude-Frequency Parameters

Although we can solve for $\boldsymbol{\theta} = [\omega, \alpha, \beta, \gamma]$ repeatedly for each individual K black by minimizing $\mathcal{L}(\boldsymbol{\theta})$, it is often convenient to have an explicit (or near-explicit) mapping from K to the parameters $[\omega, \alpha, \beta, \gamma]$. We accomplish this by training a single neural network whose output layer has four nodes, corresponding to $\omega(K)$, $\alpha(K)$, $\beta(K)$, and $\gamma(K)$. Concretely, we sample K at various points in $(0, \tilde{K}]$, with $\tilde{K}$ near the dynamic pull-in threshold K^*, and for each K we find the best-fit $\boldsymbol{\theta}$ that minimizes $\mathcal{L}(\boldsymbol{\theta})$. This yields a training set $\{ [K_i,\ \omega_i,\ \alpha_i,\ \beta_i,\ \gamma_i] \}$.

The neural network is a fully connected feedforward 4-layer net with widths of 64, 128, 64, and 32 nodes, using GELU [23] activations for smoothness. Specifically, this activation is defined as

$$\mathrm{GELU}(x) = \frac{x}{2}\left[1 + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right],$$

where erf denotes the error function. We optimize for up to 500,000 epochs, employing the Adam optimizer and a cosine-annealing learning rate schedule [24]. An early stopping criterion is used based on a patience threshold to halt training once no further improvement in the loss is observed.

Figure 5 shows a direct comparison of the neural-network-predicted profiles of $\omega(K), \alpha(K), \beta(K)$, and $\gamma(K)$ in the range of $0 \leq K \leq 0.2$ versus the true values obtained from repeated least-squares fits of $\boldsymbol{\theta}$. The network captures the parameter trends very accurately.

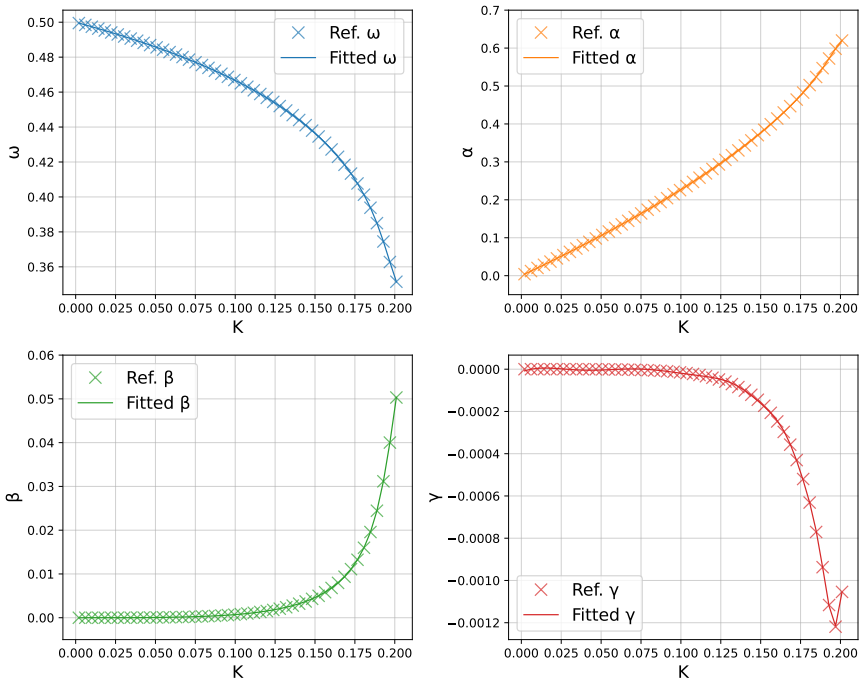

Figure 5. Neural-network fitting of the parameters $\omega, \alpha, \beta, \gamma$ as functions of K. Crosses ($\times$) represent the individual solutions obtained from the Fourier-residual minimization; solid lines show the neural network predictions.

Finally, Figure 6 illustrates the absolute errors

$$|\omega_{\mathrm{NN}}(K) - \omega_{\mathrm{ref}}(K)|, \quad |\alpha_{\mathrm{NN}}(K) - \alpha_{\mathrm{ref}}(K)|,$$
$$|\beta_{\mathrm{NN}}(K) - \beta_{\mathrm{ref}}(K)|, \quad |\gamma_{\mathrm{NN}}(K) - \gamma_{\mathrm{ref}}(K)|$$

across the sampled K-values, plotted in a log scale. These small errors confirm that the neural networks successfully learn the parameter dependencies over the entire range $(0, \tilde{K}]$.

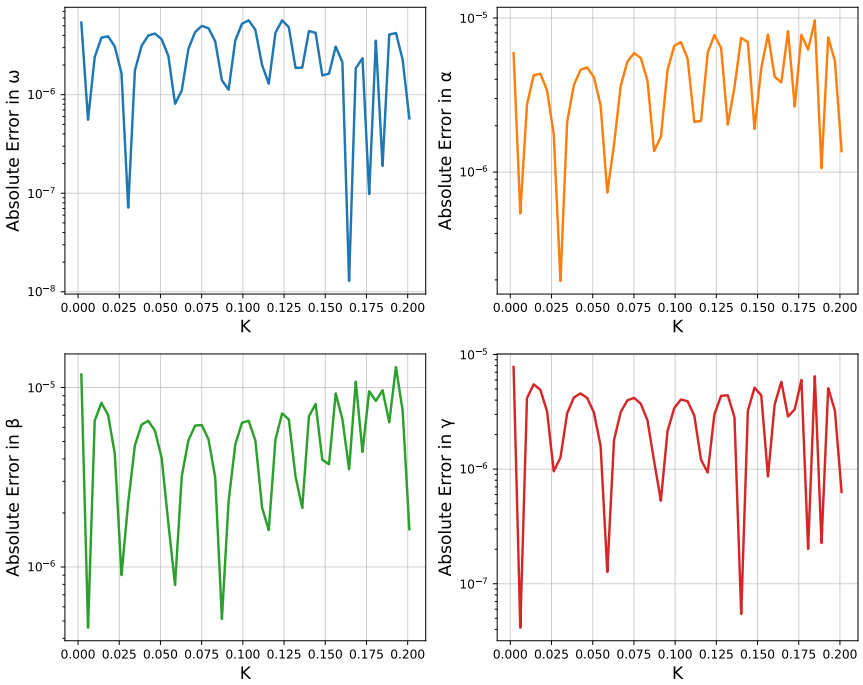

Figure 6. Log-scale plot of absolute errors between the neural network predictions and the individually optimized parameter values.

3.5. Maximum Deflection Comparison with Exact Solution

In order to further assess the accuracy of the approximate Fourier solutions, we compared the maximum (peak) amplitude of each Fourier solution against the exact amplitude obtained from the transcendental equation (see Equation (16))

$$-A^2 - 2K\ln(1-A) = 0\,, \tag{21}$$

where we ignore the trivial solution $A = 0$. For each sampled value of K, we used a standard root-finding method (Brent's method) to solve for the nontrivial maximum deflection A_{ref}. Next, we optimized the Fourier parameters $\boldsymbol{\theta} = [\omega, \alpha, \beta, \gamma]$ with 2000 Adam steps, constructed the corresponding Fourier approximate solution and extracted its maximum deflection A_{Fourier}. The procedure was repeated for 20 values of K in the range $(0, K^*)$.

Figure 7 illustrates the results. The left panel shows both the true amplitude (solid line) and the Fourier-based amplitude (dashed line) as functions of K. The right panel plots the absolute difference $|A_{\text{Fourier}} - A_{\text{ref}}|$ in a logarithmic scale. We observe that the approximation remains accurate over the entire range of K, with the absolute error staying within a small margin for practical purposes.

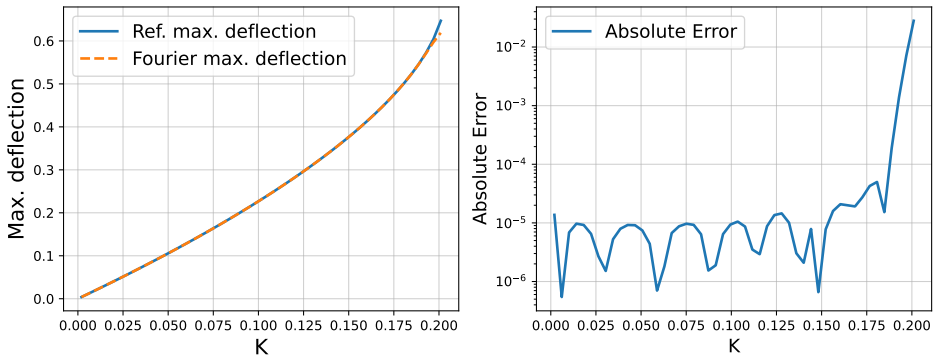

Figure 7. (**Left**) Comparison of the reference maximum deflection (solid line) derived from solving $-A^2 - 2K\ln(1-A) = 0$ with the Fourier-based maximum deflection (dashed line). (**Right**) The absolute error between these two maximum deflections displayed on a logarithmic scale.

Notice that the exact maximum deflection can be expressed in terms of Taylor expansion using Equation (21) and implicit differentation:

$$A = 2K + 2K^2 + \frac{14}{3}K^3 + 14K^4 + \frac{2138}{45}K^5 + \frac{1562}{9}K^6 + \dots\,. \tag{22}$$

This result was obtained in *Mathematica* [25] by employing `AsymptoticSolve`.

4. Extension to Non-Zero Geometric Parameter

The approach described in Section 3 focused on the limiting case $\xi \to 0$, where the governing equation simplifies to Equation (12). In practical magMEMS design, however, the geometric parameter $\xi = b/L$ is typically non-zero and can significantly affect the system dynamics. This section extends our Fourier neural network method to handle the general case with arbitrary $\xi > 0$, using the full governing Equation (4).

4.1. Generalized Fourier Approximation for Arbitrary ξ

For non-zero ξ, the governing Equation (4) can be rewritten as:

$$\frac{\ddot{x} + x}{\tau(x, \xi)} = K\,, \tag{23}$$

where $\tau(x, \xi)$ is defined as:

$$\tau(x, \xi) = \frac{\xi^2(1-x)}{\sqrt{\xi^2(1-x)^2+1}} + \frac{1}{(1-x)\sqrt{\xi^2(1-x)^2+1}} - \xi\,. \tag{24}$$

We employ the same Fourier ansatz as in Equation (19):

$$\hat{x}(t) = \alpha \sin^2(\omega t) + \beta \sin^2(2\omega t) + \gamma \sin^2(3\omega t)\,, \tag{25}$$

which automatically satisfies the initial conditions $\hat{x}(0) = \dot{\hat{x}}(0) = 0$. The key difference is that now we must optimize the parameter vector $\boldsymbol{\theta} = [\omega, \alpha, \beta, \gamma]$ for each pair (K, ξ) rather than just K alone.

The residual function becomes:

$$R_{\boldsymbol{\theta}}(t; K, \xi) = \frac{\ddot{\hat{x}}(t) + \hat{x}(t)}{\tau(\hat{x}(t), \xi)} - K\,, \tag{26}$$

and the loss function is defined as:

$$\mathcal{L}(\boldsymbol{\theta}; K, \xi) = \frac{1}{N}\sum_{i=1}^{N} [R_{\boldsymbol{\theta}}(t_i; K, \xi)]^2\,. \tag{27}$$

4.2. Two-Input Neural Network Architecture

To efficiently handle the extended parameter space, we train a neural network that maps from the two-dimensional input (K, ξ) directly to the four Fourier parameters $[\omega, \alpha, \beta, \gamma]$ (see Figure 8). This contrasts with the one-dimensional input K used in the $\xi \to 0$ case. The same surrogate rationale from Stage 2 applies here: once trained, the map $(K, \xi) \mapsto [\omega, \alpha, \beta, \gamma]$ provides instantaneous, differentiable evaluations that amortize per-query optimization and enable gradient-based inverse design across the two-parameter space.

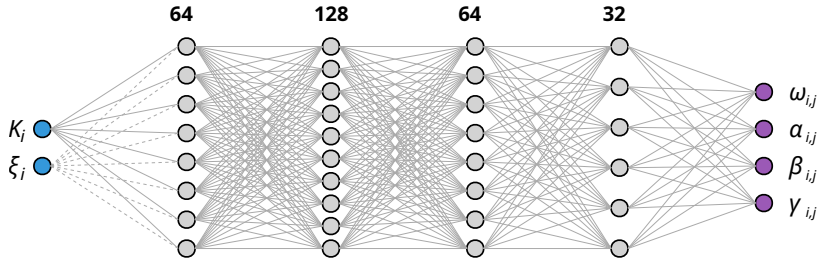

Figure 8. Neural network architecture for mapping (K_i, ξ_j) to Fourier parameters $[\omega, \alpha, \beta, \gamma]$. The network uses two inputs for the general case ($\xi \neq 0$) compared to one input (K only) when $\xi = 0$.

The neural network architecture consists of fully connected layers with widths and GELU activations. We optimize the parameters for a grid of (K, ξ) pairs where $K \in [0.01, 0.19]$ and $\xi \in [0.1, 2.0]$, ensuring that all combinations remain below the dynamic pull-in threshold K^*_ξ.

Workflow Overview and Grid Summary

The procedure follows two concise steps: (i) direct optimization of $\boldsymbol{\theta} = [\omega, \alpha, \beta, \gamma]$ for each (K_i, ξ_j) on a coarse grid (no neural network in this step), producing a dataset $(K, \xi) \mapsto (\omega, \alpha, \beta, \gamma)$; and (ii) training a feedforward neural network to model this mapping. In the extended setting, the NN has two inputs (K, ξ) and four outputs $(\omega, \alpha, \beta, \gamma)$, compared to one input (K only) when $\xi = 0$.

We use 20 points in K and 20 points in ξ (i.e., 400 total pairs) (Table 1).

Table 1. Illustrative 20×20 grid of (K, ξ) pairs. Each entry stores the optimized parameter vector $\theta_{i,j} = [\omega, \alpha, \beta, \gamma]$ obtained in Step 1 of optimization algorithm.

	K_1	K_2	...	K_{20}
ξ_1	$\theta_{1,1}$	$\theta_{1,2}$	...	$\theta_{1,20}$
ξ_2	$\theta_{2,1}$	$\theta_{2,2}$	...	$\theta_{2,20}$
⋮	⋮	⋮	⋱	⋮
ξ_{20}	$\theta_{20,1}$	$\theta_{20,2}$	...	$\theta_{20,20}$

4.3. Validation Against High-Precision Solutions

Figure 9 presents a comprehensive validation of our approach for different values of ξ. For each $\xi \in \{0.1, 0.5, 1.0, 2.0\}$ and various K values, we compare the Fourier approximation against high-precision Runge-Kutta solutions. The left panels show excellent agreement between the methods, while the right panels display the absolute errors on a logarithmic scale, demonstrating the high accuracy of our approach across different geometric configurations.

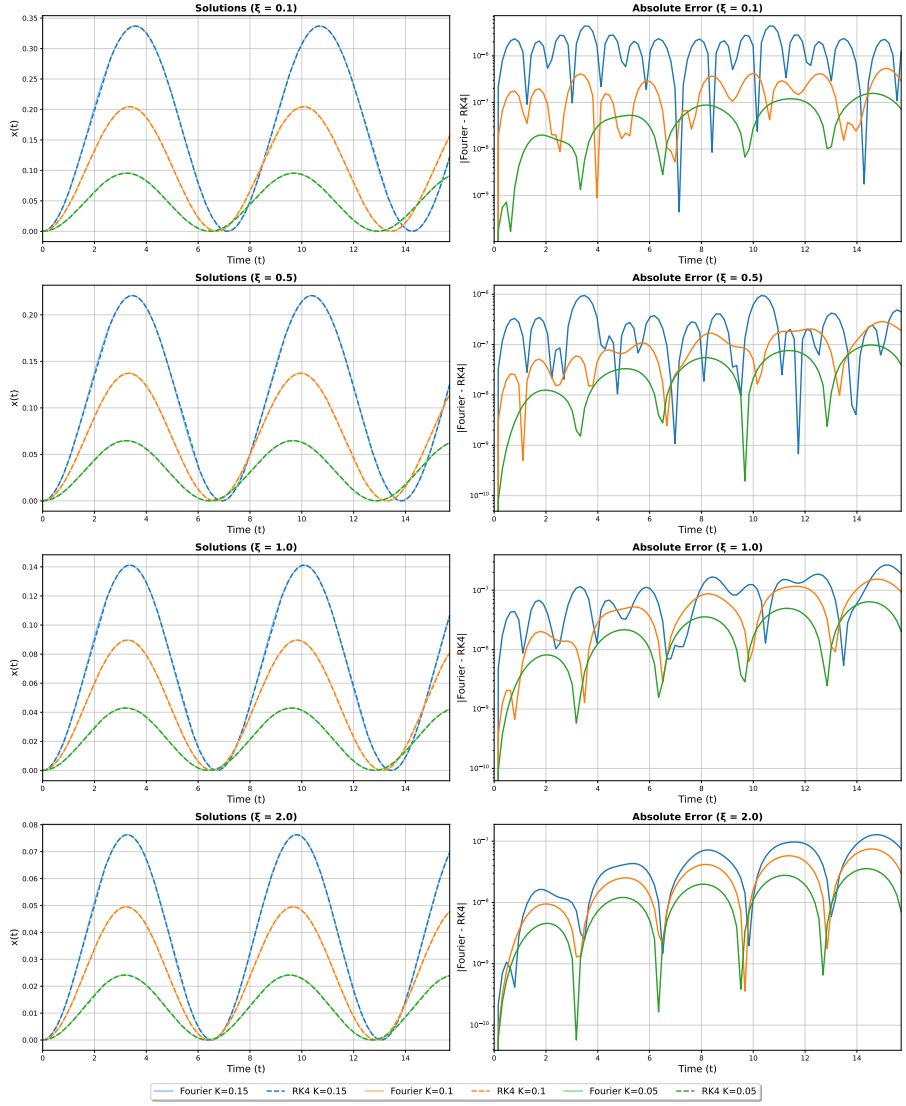

Figure 9. Validation of Fourier approximations for non-zero ξ values. Left panels: Comparison between Fourier solutions (solid lines) and RK4 solutions (dashed lines). Right panels: Absolute errors on logarithmic scale. Each row corresponds to a different ξ value: 0.1, 0.5, 1.0, and 2.0.

4.4. Parameter Dependency Analysis

The trained neural network enables efficient exploration of how the Fourier parameters vary with both K and ξ. Figure 10 illustrates the dependence of each parameter $[\omega, \alpha, \beta, \gamma]$ on K for selected values of ξ. The crosses represent the reference values obtained from individual optimizations, while the solid lines show the neural network predictions.

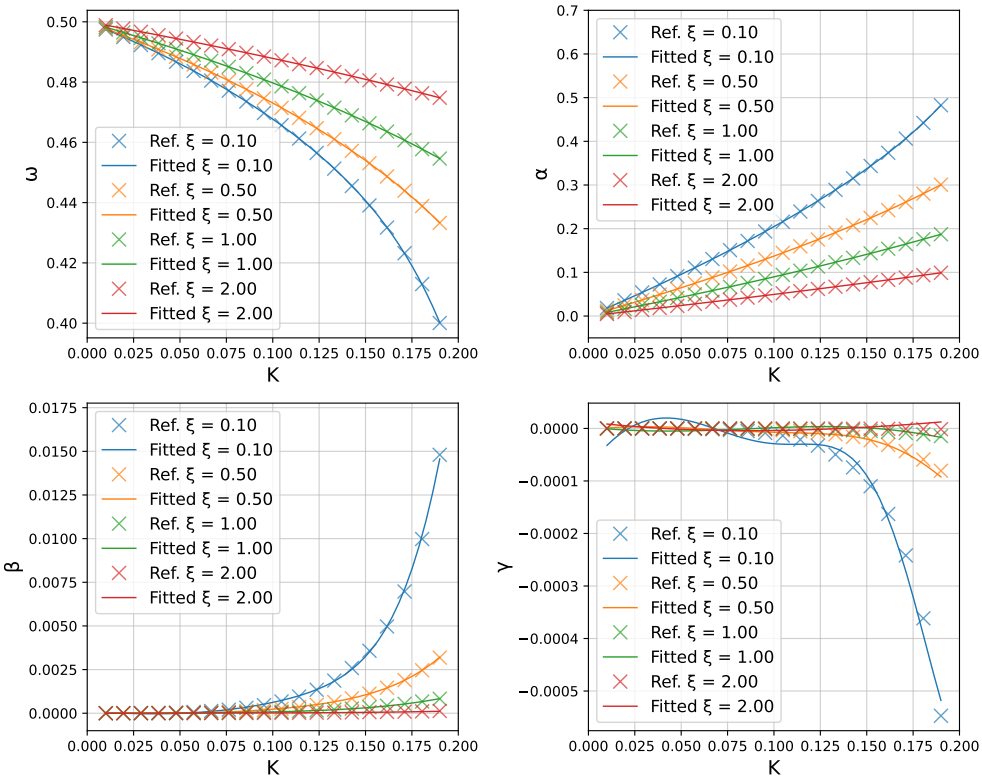

Figure 10. Fourier parameters as functions of K for different ξ values. Crosses: individual optimization results. Solid lines: neural network predictions. The geometric parameter ξ significantly influences the parameter trends, particularly for β and γ.

The results reveal that the geometric parameter ξ has a pronounced effect on the parameter dependencies. In particular, the higher-order harmonic coefficients β and γ show significant variation with ξ, indicating that the geometric ratio between conductor separation and length plays a crucial role in determining the oscillatory behavior.

The fitted trends are consistent with the magnetostatic softening process. The fundamental frequency ω decreases almost linearly with K, with a steeper slope for smaller ξ. The leading amplitude α grows roughly linearly with K and is suppressed as ξ increases, reflecting a weaker effective forcing. Higher harmonics remain small, that is, β increases with K and is most pronounced for small ξ, whereas γ stays close to zero and vanishes as ξ grows. Consequently, for larger ξ the response is nearly single-harmonic, which explains the improved accuracy observed in Figure 9.

Figure 11 quantifies the accuracy of the neural network predictions by showing the absolute errors between predicted and reference values. The errors remain small across the entire parameter space, confirming the reliability of our approach.

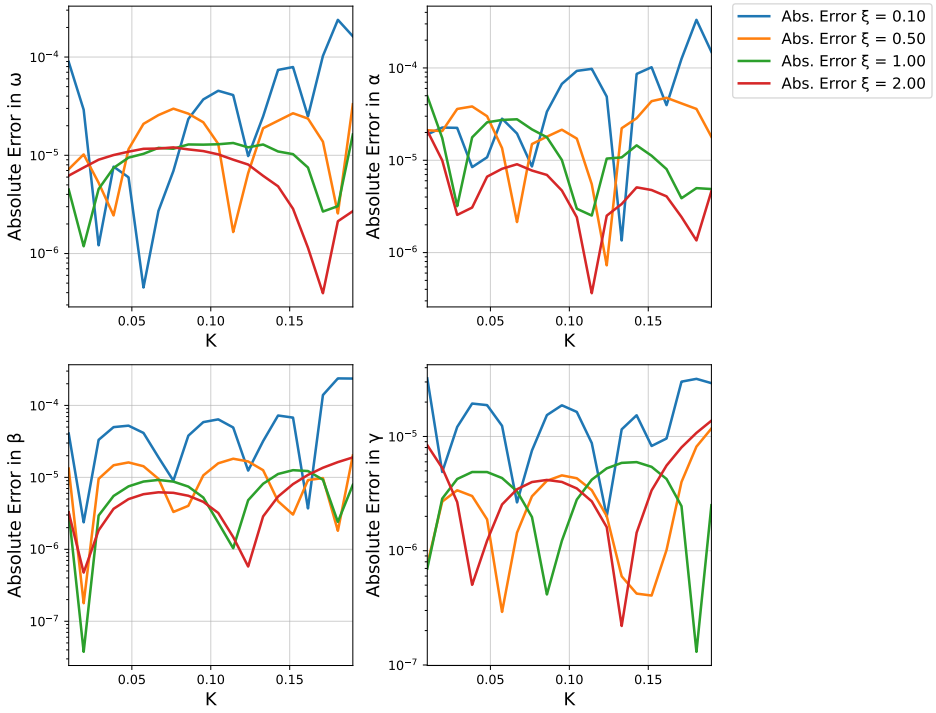

Figure 11. Absolute errors between neural network predictions and reference parameter values for different ξ values, displayed on logarithmic scale. The errors remain small across all parameters and ξ values.

4.5. Three-Dimensional Parameter Surfaces

To provide a comprehensive view of the parameter space, Figure 12 presents three-dimensional surface plots showing how each Fourier parameter varies simultaneously with both K and ξ. These surfaces capture the full parameter dependencies and can serve as visual guides for magMEMS design optimization.

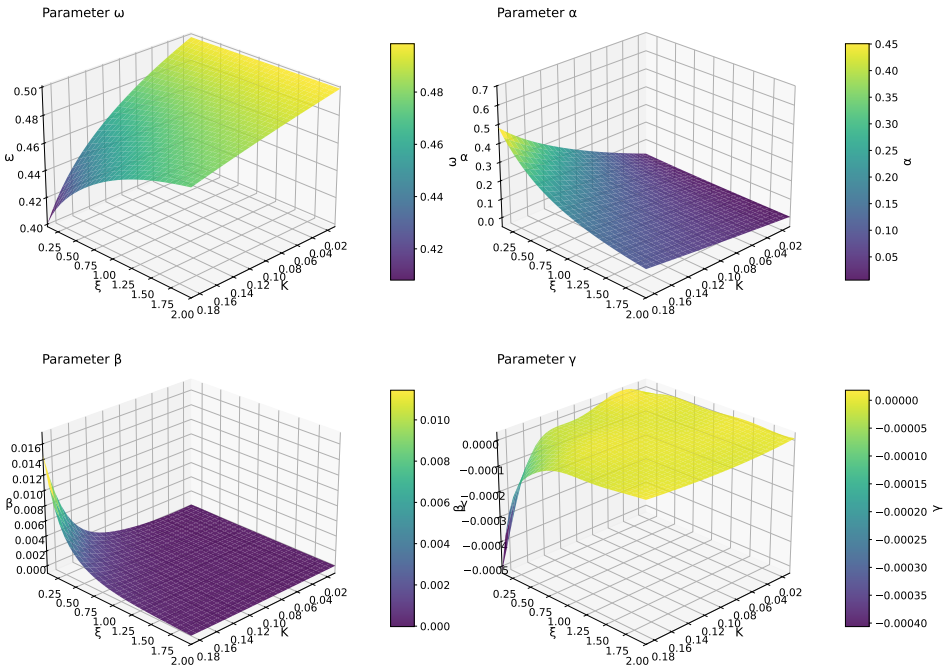

Figure 12. Three-dimensional parameter surfaces showing the variation of Fourier parameters $[\omega, \alpha, \beta, \gamma]$ with both excitation parameter K and geometric parameter ξ. These surfaces provide comprehensive insight into the parameter dependencies for magMEMS design.

The surface plots show that the frequency parameter ω increases smoothly with both K and ξ, exhibiting a nearly linear relationship across the parameter space. The primary amplitude coefficient α displays a more complex behavior, decreasing with increasing ξ while showing nonlinear dependence on K. The higher-order harmonic coefficients β and γ remain relatively small, with β showing the most variation across the parameter space and γ remaining close to zero throughout most of the domain. All parameters demonstrate smooth, well-behaved surfaces without discontinuities, ensuring the numerical stability of the approach.

This extended framework provides magMEMS designers with a comprehensive tool for analyzing devices with arbitrary conductor geometries, enabling rapid optimization across the full design space defined by both electrical and geometric parameters.

Alternatives such as low-order polynomials, splines, radial-basis interpolation, or Kriging/Gaussian process (GP) can approximate smooth maps very well, but in our two-parameter input (K, ξ) we observed: (i) nonuniform sensitivity and mild nonlinearity near the pull-in boundary; (ii) cross-coupling between K and ξ; and (iii) the need for stable analytic gradients. A small fully connected NN handled these without hand-crafted basis selection, provided smooth, differentiable predictions, and generalized well between grid points. Interpolation-based surrogates typically require denser, more regular grids and may introduce non-smooth derivatives at knots, while GP/Kriging adds kernel selection and $\mathcal{O}(n^3)$ scaling. The compact NN achieved comparable or better accuracy with straightforward training and $\mathcal{O}(1)$ inference.

5. Proof-of-Concept Measurement and Error Budget

5.1. Prototype Experimental Set-Up

The coil wire was driven with a DC current of 4 A, while the magMEMS cantilever carried a DC bias on the order of milliamperes (Figure 13). Displacement was recorded using a SIOS^{TM} Nano Vibration Analyzer (Ilmenau, Germany), a laser Doppler vibrometer with a noise floor of 5 pm/$\sqrt{\text{Hz}}$. Since the instrument's DC path is suppressed, purely static deflection could not be measured directly. Instead, the quasi-DC response was inferred under sub-resonant harmonic actuation. The calibrated quasi-static sensitivity is $S_{\text{stat}} = \frac{\mu_0 \mu_r i_1 L}{2\pi b k_s} \approx 6.46 \times 10^{-2}$ nm/mA ($\approx$65 pm/mA).

In the present large-gap regime, the nondimensional groups obtained from Table 2 satisfy $K \simeq 4.01 \times 10^{-7}$ and $\xi \simeq 3.10 \times 10^{2}$, from which it follows that $R'(x_0) \leq K/\xi \approx 1.3 \times 10^{-9} \ll 1$. The linearization of (4) about the DC equilibrium x_0 yields

$$\delta\ddot{x} + \big(1 - R'(x_0)\big)\,\delta x = f(t), \qquad R'(x) = \frac{K}{(1-x)^2\sqrt{\xi^2(1-x)^2+1}},$$

so $k_{\text{eff}} = k_s\big(1 - R'(x_0)\big) \approx k_s$. Consequently, the quasi-static gain is independent of the DC bias:

$$X_{\text{rms}} = \frac{1}{\sqrt{2}} \frac{\mu_0 \mu_r\, i_1\, i_2^{\text{ac}}\, L_{\text{eff}}}{2\pi\, b\, k_s} \frac{1}{1 - R'(x_0)} \approx \frac{1}{\sqrt{2}} \frac{\mu_0 \mu_r\, i_1\, i_2^{\text{ac}}\, L_{\text{eff}}}{2\pi\, b\, k_s}.$$

Thus any observable DC trend is attributed to a weak DC dependence of parasitic AC ripple and the readout chain rather than elastic/magnetic softening.

Measured RMS contains signal and a DC-independent baseline. Using the $i_2 = 0$ mA point as baseline, $X_{\text{noise}} = 5.54$ pm, the signal-only RMS is computed in the power domain as

$$X_{\text{sig}} = \sqrt{\max\big(X_{\text{meas}}^2 - X_{\text{noise}}^2,\, 0\big)}.$$

With the quasi-static gain $G = S_{\text{stat}}/\sqrt{2} \approx 46$ pm/mA, the equivalent ripple current is $i_{2,\text{ripple}} = X_{\text{sig}}/G$.

Table 2. Variables and parameters, updated with the correct gap and active length.

Variables and Parameters	Value	Units	Description
i_1	4	[A]	current in coil wire
i_2	10×10^{-3}	[A]	current in cantilever, used for K below
b	1.61×10^{-3}	[m]	effective coil–cantilever separation
L	5.20×10^{-6}	[m]	effective active length
k_s	0.04	[N/m]	spring constant of cantilever
μ_0	$4\pi \times 10^{-7}$	[N/A^2]	vacuum permeability
μ_r	1	[−]	relative permeability
$K = \frac{\mu_0 \mu_r i_1 i_2 L}{2\pi k_s b^2}$	4.01×10^{-7}	[−]	excitation at $i_2 = 10$ mA
$\xi = \frac{b}{L}$	3.096×10^2	[−]	geometric parameter
$S_{\text{stat}} = \frac{\mu_0 \mu_r i_1 L}{2\pi b k_s}$	6.46×10^{-2}	[nm/mA]	quasi-static sensitivity (≈65 pm/mA)

The experimental parameters were extracted from Table 2, and the excitation parameter K was computed via Equation (5). For completeness, the nondimensional groups read $\xi = b/L \approx 3.10 \times 10^2$ and $K = \mu_0 \mu_r i_1 i_2 L/(2\pi k_s b^2) \propto i_2$, so $K \approx 4.01 \times 10^{-7}$ at $i_2 = 10$ mA. Since $R'(0) \approx K/\xi \sim 10^{-9}$, the operation is overwhelmingly far from pull-in, and the quasi-static RMS is flat versus DC bias to first order. Thus, any residual variation reflects unintended AC ripple or environmental excitation.

We then analyzed the vibration spectrum. With a DC current applied to the magMEMS cantilever, the first-harmonic RMS amplitude varied by at least a factor of five, from 1.6 pm under repulsive bias to 7 pm under attractive bias; see Table 3. For clarity, Table 3 reports the measured RMS values together with the estimated uncertainty Δ.

Table 3. RMS vibration under DC bias with power-domain baseline removal and implied equivalent ripple.

i_2 DC (mA)	X_{meas} (pm)	$\pm\Delta$ (pm)	X_{sig} (pm)	$i_{2,\text{ripple}}$ (mA)
0	5.54	1.11	0.00	0.000
2	4.29	0.86	0.00	0.000
4	6.95	1.39	4.20	0.092
6	8.46	1.69	6.40	0.140
8	6.74	1.35	3.80	0.083
10	7.19	1.44	4.60	0.100

Baseline $X_{\text{noise}} = X_{\text{meas}}(0\text{ mA}) = 5.54$ pm. Signal-only $X_{\text{sig}} = \sqrt{\max(X_{\text{meas}}^2 - X_{\text{noise}}^2, 0)}$. Gain $G = S_{\text{stat}}/\sqrt{2} \approx 46$ pm/mA. Equivalent ripple $i_{2,\text{ripple}} = X_{\text{sig}}/G$.

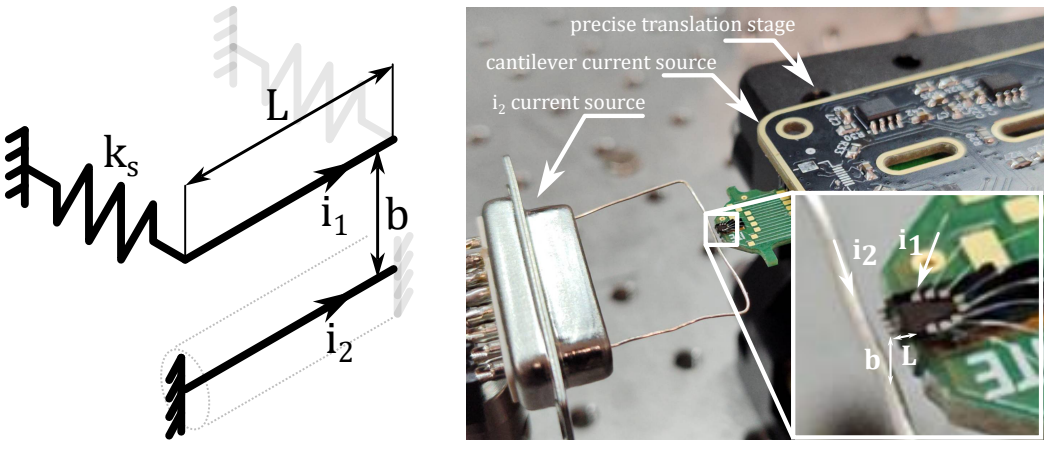

Figure 13. Simplified representation and photograph of the experimental setup with an elastic microcantilever carrying a current-leading filament.

This example illustrates a limitation of the Fourier approach at very small excitation: for the present range of the excitation parameter, the relative error of higher Fourier coefficients deteriorates the approximation. Nevertheless, one can compute the maximum deflection numerically as the root of $f_{K,\xi}(s)$ in Equation (8). For the case considered we obtain $s = 6.476 \cdot 10^{-10}$, which yields $1.043 \cdot 10^{-12}$ m as the maximum deflection.

5.2. Experimental Perspective

To validate the findings obtained with the test Fourier neural network, a model magMEMS device is required. The simplest configuration, both in form and function, consists of two parallel conductor wires positioned opposite one another. These wires are self-supporting, serving simultaneously as conductors and as mechanical scaffolding. To regulate their length and stiffness, the conductors are fabricated in the form of ribbons and placed on an insulating substrate, with certain sections extending partially across a gap. Furthermore, the wires may be curved or bent, enabling them to contract or expand in response to external stimuli. A schematic illustration of the structure is presented in Figure 14.

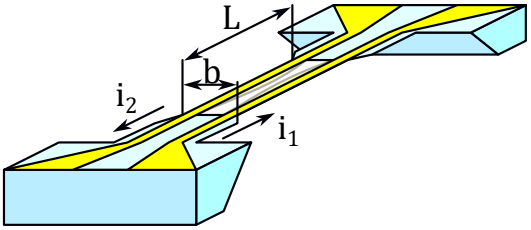

Figure 14. Idea of a MagMEMS—model device, with mechanical and electric structure simplified to the basic case—two conductors.

The primary advantage of the proposed structure lies in its relative simplicity of fabrication and operation. Detection of the desired wire deflection induced by the self-generated magnetic field must be carried out using a high-resolution microscopic technique such as scanning electron microscopy (SEM), for example.

The initial test structures were fabricated with $L = 40$ µm and $b = 5$ µm, as illustrated in Figure 15. To ensure mechanical stability and to prevent wedging, the structure may be produced with more conductive lines than strictly required; however, in the verification experiment, only two lines were employed. The stiffness of the structure is governed by the line thickness, which in this case was set to 3 µm, yielding a stiffness of approximately 10 N/m. Based on the general model, the design parameters can subsequently be optimized to satisfy the Fourier parameters defined by the three-term ansatz in Equation (19).

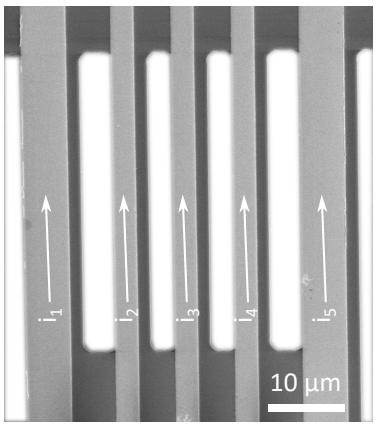

Figure 15. SEM image of an realized, but unoptimized exemplary structure for in-principle observation and validation of the calculated parameters.

6. Conclusions

Approximate oscillatory solutions for a magMEMS switch with current-carrying conductors are obtained using the Fourier Neural Network Method for both the limiting case $\xi \to 0^+$ and the general case with arbitrary geometric parameter $\xi > 0$. Approximate periodic solutions for the range of the excitation parameter values less than the pull-in value exhibit accuracy comparable to that of classical ODE solvers. Numerical illustrations and validations of the effect of both the scaled excitation parameter K and geometric parameter ξ on the oscillation characteristics were presented. This allows for rapid evaluation of magMEMS switch device performance for various excitation currents, filament lengths, maximum separation distances, and spring parameters. The extension to non-zero ξ significantly broadens the applicability of the method, enabling comprehensive design space exploration. These results are especially useful in system optimization and can be easily implemented in a time domain circuit simulator like Spice.

MagMEMS devices under consideration are uniform in material composition, which is not the universal case. Inclusion of auxiliary materials (e.g., support, insulating, forming, etc.) might prove model less applicable by shear change to wire-wire b dimension. However, the most ubiquitous form of RF switches, nanowire-based circuits, or active MEMS devices are formed of a single material only [8]. As such, they are most accurately described by the introduced model. We believe that the nanofield is one of the main application areas.

The results obtained in this work can be useful for the design of magMEMS based on current-carrying filaments. Applications include design of magMEMS switches, oscillators, and resonators. The next research phase will involve a comprehensive experimental study of oscillating current-carrying filaments, transitioning from scaled-up prototypes to application-ready magMEMS devices. To develop a more complete model, future studies should explore damping and nonlinear resonances in harmonically excited systems.

Author Contributions: Conceptualization, P.S., B.P., P.P. and G.E.; methodology, A.B. and P.S.; software, A.B.; validation, P.S., A.B., A.D., P.P., B.P. and G.E.; formal analysis, A.B., A.D., P.S. and K.T.; investigation, A.B.; resources, P.S.; data curation, A.B.; writing—original draft preparation, A.B., P.S., P.P., B.P. and G.E.; writing—review and editing, P.S., A.B., A.D., P.P., B.P., K.T. and G.E.; visualization, A.B.; supervision, P.S., A.D. and G.E.; project administration, P.S.; funding acquisition, P.S., B.P., P.P. and A.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been supported by the Ministry of Education and Science of the Republic of Kazakhstan within the framework of Project AP19676969 "Modeling, Analysis and Optimization of MEMS and magMEMS" and by National Science Centre of Poland project in OPUS 27 contest "Tip quantum dot transistor picometrology: new routes to explore the quantum systems using scanning probe microscopy—SQTMet".

Data Availability Statement: The code for this project is available at: https://github.com/armanbolatov/magmems_optimization (accessed on 25 November 2025). The data supporting the reported results can be reproduced using the provided code.

Acknowledgments: The authors thank the reviewers for their valuable comments and suggestions that helped improve the quality of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Younis, M.I. *MEMS Linear and Nonlinear Statics and Dynamics*; Springer: Berlin/Heidelberg, Germany, 2011.
2. Omarov, D.; Nurakhmetov, D.; Wei, D.; Skrzypacz, P. On the application of Sturm's theorem to analysis of dynamic pull-in for a graphene-based MEMS model. *Appl. Comput. Mech.* **2018**, 12, 59–72. [CrossRef]

3. Skrzypacz, P.; Dongming, W.; Nurakhmetov, D.; Kostsov, E.G.; Sokolov, A.A.; Begzhigitov, M.; Ellis, G. Analysis of dynamic pull-in voltage and response time for a micro-electro-mechanical oscillator made of power-law materials. *Nonlinear Dyn.* **2021**, *105*, 227–240. [CrossRef]
4. Nathanson, H.; Newell, W.; Wickstrom, R.; Davis, J. The resonant gate transistor. *IEEE Trans. Electron Devices* **1967**, *14*, 117–133. [CrossRef]
5. Zhang, W.; Yan, H.; Peng, Z.-K.; Meng, G. Electrostatic pull-in instability in mems/nems: A review. *Sens. Actuators A Phys.* **2014**, *214*, 187–218. [CrossRef]
6. Flores, G. On the dynamic pull-in instability in a mass-spring model of electrostatically actuated MEMS devices. *J. Differ. Equ.* **2017**, *262*, 3597–3609. [CrossRef]
7. Park, B.; Li, M.; Liyanage, S.; Shafai, C. Lorentz force based resonant MEMS magnetic-field sensor with optical readout. *Sens. Actuators A Phys.* **2016**, *241*, 12–18. [CrossRef]
8. Feng, X.L.; He, R.; Yang, P.; Roukes, M.L. Very high frequency silicon nanowire electromechanical resonators. *Nano Lett.* **2007**, *7*, 1953–1959. [CrossRef]
9. Loh, O.Y.; Espinosa, H.D. Nanoelectromechanical contact switches. *Nat. Nanotechnol.* **2012**, *7*, 283–295. [CrossRef] [PubMed]
10. Sun, W.P.; Lim, C.W.; Wu, B.S.; Wang, C. Analytical approximate solutions to oscillation of a current-carrying wire in a magnetic field. *Nonlinear Anal. Real World Appl.* **2009**, *10*, 1882–1890. [CrossRef]
11. Senturia, S.D. *Microsystem Design*; Springer: Boston, MA, USA, 2001.
12. Skrzypacz, P.; Ellis, G.; He, J.-H.; He, C.-H. Dynamic pull-in and oscillations of current-carrying filaments in magnetic micro-electro-mechanical system. *Commun. Nonlinear Sci. Numer. Simul.* **2022**, *109*, 106350. [CrossRef]
13. He, J.-H.; Nurakhmetov, D.; Skrzypacz, P.; Wei, D. Dynamic Pull-in for Micro-Electro-Mechanical Device with a Current-Carrying Conductor. *J. Low Freq. Noise Vib. Act. Control* **2019**, *40*, 1059–1066. [CrossRef]
14. He, J.-H.; Skrzypacz, P.S.; Zhang, Y.; Pang, J. Approximate periodic solutions to microelectromechanicalem oscillator subject to magnetostatic excitation. *Math. Methods Appl. Sci.* **2020**. [CrossRef]
15. Skrzypacz, P.; Ellis, G.; Pruchnik, B.; Putek, P. Generalized analysis of dynamic pull-in for singular magMEMS and MEMS oscillators. *Sci. Rep.* **2025**, *15*, 23691. [CrossRef] [PubMed]
16. Bota, C.; Caruntu, B.; Bundau, O. Approximate periodic solutions for oscillatory phenomena modelled by nonlinear differential equations. *Math. Probl. Eng.* **2014**, *2014*, 513473. [CrossRef]
17. Nayfeh, A.H.; Mook, D.T. *Nonlinear Oscillations*; John Wiley & Sons: Hoboken, NJ, USA, 2024.
18. Raissi, M., Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]
19. Wang, S.; Teng, Y.; Perdikaris, P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM J. Sci. Comput.* **2021**, *43*, A3055-81. [CrossRef]
20. Bottou, L. Large-Scale Machine Learning with Stochastic Gradient Descent. In *Proceedings of COMPSTAT'2010*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 177–186.
21. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
22. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 8024–8035.
23. Hendrycks, D.; Gimpel, K. Gaussian Error Linear Units (GELUs). *arXiv* **2016**, arXiv:1606.08415.
24. Loshchilov, I.; Hutter, F. SGDR: Stochastic Gradient Descent with Warm Restarts. In Proceedings of the International Conference on Learning Representations (ICLR), Toulon, France, 24–26 April 2017.
25. Wolfram Research, Inc. *Mathematica Version 14.3*; Wolfram Research, Inc.: Champaign, IL, USA, 2025. Available online: https://www.wolfram.com/mathematica (accessed on 25 November 2025).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.