METHODOLOGY   **Open Access**

Deep learning-based wave digital modeling of rate-dependent hysteretic nonlinearities for virtual analog applications

Oliviero Massi* , Alessandro Ilic Mezza, Riccardo Giampiccolo and Alberto Bernardini

Abstract

Electromagnetic components greatly contribute to the peculiar timbre of analog audio gear. Indeed, distortion effects due to the nonlinear behavior of magnetic materials are known to play an important role in enriching the harmonic content of an audio signal. However, despite the abundant research that has been devoted to the characterization of nonlinearities in the context of virtual analog modeling over the years, the discrete-time simulation of circuits exhibiting rate-dependent hysteretic phenomena remains an open challenge. In this article, we present a novel data-driven approach for the wave digital modeling of rate-dependent hysteresis using recurrent neural networks (RNNs). Thanks to the modularity of wave digital filters, we are able to locally characterize the wave scattering relations of a hysteretic reluctance by encapsulating an RNN-based model into a single one-port wave digital block. Hence, we successfully apply the proposed methodology to the emulation of the output stage of a vacuum-tube guitar amplifier featuring a nonlinear transformer.

Keywords Hysteresis, Neural networks, Wave digital filters

1 Introduction

The practice of emulating analog circuits and devices in the context of digital audio effects is known as virtual analog (VA) modeling [1–3]. Over the last few years, a lot of research effort has been devoted to deriving efficient and accurate digital implementations of circuit nonlinearities found in analog audio gear, which are well appreciated for their peculiar tonal character by industry professionals. Among such nonlinear phenomena, frequency-dependent saturation effects due to magnetic materials are of particular interest. Indeed, electromagnetic components can be found along the entire analog sound recording chain, which comprises, e.g., guitar pickups, electrodynamic microphones, loudspeaker drivers,

electrical transformers, and magnetic tapes. A distinctive characteristic of magnetic materials is rate-dependent hysteresis [4] which affects the magnetic flux density B in response to a variation in the magnetic field H and its gradient [5, 6]. In general, the output of a system exhibiting hysteresis follows different paths with increasing or decreasing inputs, resulting in various loops depending on the past history. For this reason, modeling hysteresis is known to be a challenging task, especially for what concerns discrete-time circuital simulation. Furthermore, modeling the dynamic ferromagnetic effects that modify the hysteresis characteristics depending on the input frequency adds a further aspect of complexity when trying to tackle the problem with conventional VA methods.

In the literature, VA modeling approaches can be divided into two main categories: *black-box* methods that try to infer the global behavior of a reference circuit from observational input/output data using, e.g., Volterra series [7] or neural networks [8], and *white-box* methods that emulate the reference circuit

*Correspondence:
Oliviero Massi
oliviero.massi@polimi.it
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di
Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy

by solving the underlying system of ordinary differential equations using, e.g., state-space models [9, 10], port-Hamiltonian methods [11], or wave digital filters (WDFs) [3]. In particular, among white-box approaches, WDFs have recently proved to be an efficient framework for deriving digital representations of electromagnetic circuits [12, 13].

First introduced by A. Fettweis in the 70s to derive digital implementations of passive analog filters [14], WDFs are realized describing a reference analog circuit as an interconnection of wave digital (WD) blocks. This is accomplished by substituting Kirchhoff port variables (port voltages and port currents) with linear combinations of *wave variables* (incident waves and reflected waves) through the addition of a free parameter at each port called *port resistance*. Circuit elements and connection networks are dealt with separately as they are described by one-port or multi-port blocks characterized by input/output scattering relations.

Several works in the literature on WDFs are devoted to the modeling of circuit nonlinearities, including diodes [15–21], transistors [22–25], and vacuum tubes [26–29]. Paiva et al. [15] derived an explicit WD description of exponential diodes and diode pairs based on the Lambert function, which was later extended in [30]. A general approach based on the Lambert function to derive closed-form scattering relations for exponential nonlinearities, such as the Shockley diode model or simplifications of the Ebers-Moll model for certain bipolar junction transistor (BJT) amplifier configurations, was later discussed in [17]. D'Angelo et al. [20], in order to reduce computational cost, proposed to reformulate the expressions involving the main branch of the Lambert function in terms of the Wright omega function. A different approach based on one-dimensional Newton-Raphson (NR) solvers was presented in [18] and [19]. Canonical piecewise linear (CPWL) representations of nonlinear functions [31, 32] were also employed to derive explicit WD scalar mappings [12], whereas models of the diode and of triode nonlinearities based on multilayer perceptrons [21, 29] have been recently proposed.

Linear circuits and circuits with up to one nonlinear element, as long as the nonlinearity is characterized by an explicit wave mapping, can be modeled via stable discretization methods without the need of iterative solvers [19, 33]. This is a considerable advantage of WD modeling over VA methods that operate in the Kirchhoff domain, as the latter are typically characterized by systems of implicit equations and entail the use of iterative algorithms [33–35]. Moreover, although the use of iterative solvers is still required if multiple nonlinear elements are present in the circuit [36, 37], the modular structure of WDFs and the freedom in selecting the

port parameters have proven advantageous in terms of efficiency and robustness compared to iterative methods designed in the Kirchhoff domain [12, 18, 19, 38, 39].

Despite a rich literature on WDFs, only few works to date have focused on the modeling of nonlinearities with memory. For instance, [40] presented an approach based on mutators for the implementation of a class of nonlinear dynamic one-port elements. In [41], instead, the modeling of generic memristors in the WD domain was discussed. Unfortunately, however, none of the existing methods can be readily used for the WD implementation of rate-dependent hysteresis.

Thanks to their excellent nonlinear approximation capabilities [42], deep neural networks have been recently employed for the modeling of hysteretic phenomena in various physical domains [43–49]. However, whilst neural networks had been previously introduced in the field of WDFs as an alternative methodology to define explicit wave mappings for static nonlinear components, deep learning methods capable of dealing with nonlinearities with memory and input rate dependency have yet to be integrated in a general WDF framework for discrete-time circuit simulation.

In this article, we bridge this gap by studying the modeling and implementation of electromagnetic audio circuits with rate-dependent hysteretic nonlinearities in the WD domain. We locally model a nonlinear reluctance exhibiting hysteresis as a one-port circuit element making use of a recurrent neural network (RNN) [45]. In fact, RNNs are capable of modeling the long-term memory effects that characterize rate-dependent hysteresis. The resulting WD block is trained using wave variables, and it can be readily inserted into multiphysics WD structures [12, 13] in order to implement nonlinear electromagnetic reference circuits. To the best of our knowledge, this constitutes the first example of using RNNs to model the dynamic behavior of circuital blocks in the WD domain. Ultimately, we pursue a hybrid approach, supplementing the white-box WDFs formalism with purely data-driven modules. Notably, the modular structure of WDFs allows us to limit the scope of black-box modeling only to the characterization of specific circuit elements by learning the respective wave scattering relations from measurement data.

The remainder of this manuscript is organized as follows. In Section 2, we provide the theoretical background on WDFs. Section 3 is devoted to the analysis of rate-dependent hysteresis models. In Section 4, we illustrate how to implement reluctances with hysteresis in the WD domain using an RNN-based model. In Section 5, we present the model training procedure. In Section 6, we utilize the proposed WD circuital block for the emulation of the output stage of a vacuum tube guitar amplifier. Conclusions are
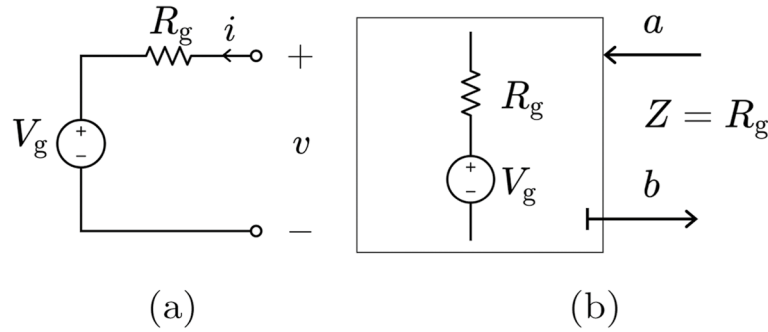
Fig. 1 Generic linear one-port element **a** in the Kirchhoff domain and **b** in the wave digital domain. The T-shaped stub indicates port adaptation

drawn in Section 7, where future work and applications are also discussed.

2 Background on wave digital filters

The design of WDFs is based on a port-wise description of a reference analog circuit. Circuit elements and topological connection networks are modeled using one- or multi-port WD blocks characterized by scattering relations. This is made possible by substituting each pair of Kirchhoff variables, i.e., port voltage v and port current i, with a pair of *wave* variables. Although different types of waves exist in the literature on WDFs [14, 35, 50, 51], the most used definition is that of voltage waves [14]

$$a = v + Zi, \quad b = v - Zi, \tag{1}$$

where a and b are the incident and reflected waves, respectively, whereas Z is a free parameter called *port resistance*. The inverse mapping of (1) is given by

$$v = \frac{a+b}{2}, \quad i = \frac{a-b}{2Z}, \tag{2}$$

which holds true if and only if $Z \neq 0$. Since it can be arbitrarily selected, Z constitutes a powerful degree of freedom in the port description. Indeed, a proper choice of Z is of fundamental importance for the numerical solution of WD structures [14].

2.1 Linear circuit elements

As a representative example, let us consider a generic linear one-port circuit element as the one in Fig. 1a. A large class of linear one-port circuit elements, including resistors, resistive voltage/current sources, and dynamic elements such as capacitors and inductors discretized using stable methods, can be described by means of the discrete-time Thévenin equivalent model [19]

$$v[k] = R_g[k]i[k] + V_g[k], \tag{3}$$

where k is the sampling index, $v[k]$ is the port voltage, $i[k]$ is the port current, $R_g[k]$ is a resistive parameter,

and $V_g[k]$ is a voltage parameter. According to (1), the Thévenin equivalent model can be expressed in the WD domain as follows

$$b[k] = \frac{R_g[k] - Z[k]}{R_g[k] + Z[k]} a[k] + \frac{2Z[k]}{R_g[k] + Z[k]} V_g[k]. \tag{4}$$

The instantaneous dependence between $b[k]$ and $a[k]$ can be eliminated by setting $Z[k] = R_g[k]$; in this case, (4) reduces to $b[k] = V_g[k]$, and the linear one port-element is said to be *adapted* [14] (see Fig. 1b).

2.2 Topological connection networks

In the Kirchhoff domain, a N-port topological connection network [51, 52] is characterized by a vector of port voltages $\mathbf{v} = [v_1, \ldots, v_N]^T$ and a vector of port currents $\mathbf{i} = [i_1, \ldots, i_N]^T$. Let $\mathbf{v}_t \in \mathbb{R}^\chi$ be the vector of independent port voltages and $\mathbf{i}_l \in \mathbb{R}^\psi$ be the vector of independent port currents, where $\chi + \psi = N$. Thus, it is possible to write

$$\mathbf{v} = \mathbf{Q}^T \mathbf{v}_t, \quad \mathbf{i} = \mathbf{B}^T \mathbf{i}_l, \tag{5}$$

where $\mathbf{B}$ is the fundamental loop matrix and $\mathbf{Q}$ is the fundamental cut-set matrix [53]. Given that topological connection networks are lossless and reciprocal, the ortogonality property $\mathbf{Q}\mathbf{B}^T = \mathbf{0}$ holds true [51, 53]. The matrix $\mathbf{Q}$ of size $\chi \times N$ and the matrix $\mathbf{B}$ of size $\psi \times N$ can be derived performing a tree-cotree decomposition of the reference circuit [54]. In the WD domain, topological connection networks are modeled using N-port junctions characterized by the wave variables

$$\mathbf{a} = \mathbf{v} + \mathbf{Zi}, \quad \mathbf{b} = \mathbf{v} - \mathbf{Zi}, \tag{6}$$

where $\mathbf{a} = [a_1, \ldots, a_N]^T$ is the vector of waves incident to the junction and $\mathbf{b} = [b_1, \ldots, b_N]^T$ is the vector of waves reflected by the junction, while $\mathbf{Z} = \text{diag}[Z_1, \ldots, Z_N]$ is a diagonal matrix having port resistances as diagonal entries. The scattering relation between $\mathbf{a}$ and $\mathbf{b}$ is given by $\mathbf{b} = \mathbf{Sa}$, where $\mathbf{S} \in \mathbb{R}^{N \times N}$ is a scattering matrix that
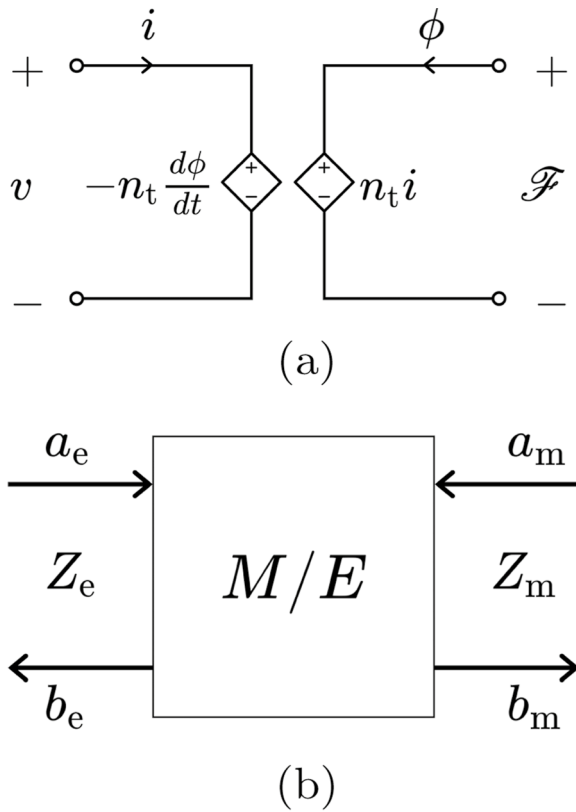
Fig. 2 Magnetic/electric junction **a** in the Kirchhoff domain and **b** in the wave digital domain

can be computed using either of the two following equivalent equations [51]

$$\mathbf{S} = 2\mathbf{Q}^{\mathrm{T}}(\mathbf{Q}\mathbf{Z}^{-1}\mathbf{Q}^{\mathrm{T}})^{-1}\mathbf{Q}\mathbf{Z}^{-1} - \mathbf{I},\tag{7}$$

$$\mathbf{S} = \mathbf{I} - 2\mathbf{Z}\mathbf{B}^{\mathrm{T}}(\mathbf{B}\mathbf{Z}\mathbf{B}^{\mathrm{T}})^{-1}\mathbf{B},\tag{8}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix. If $\psi > \chi$, (7) is computationally cheaper than (8). If $\chi > \psi$, the opposite holds true.

2.3 Magnetic/electric junction

As shown in [55], an analogy can be drawn between the magneto-motive force and the electric voltage and between the magnetic flux and the electric current. Consequently, the coupling between magnetic and electric domains can be realized by means of the magnetic/electric (ME) junction [12, 13], shown in Fig. 2a, whose continuous-time constitutive equations are

$$\begin{cases} v(t) = -n_{\mathrm{t}}\frac{d\phi}{dt} \\ \mathscr{F}(t) = n_{\mathrm{t}}i(t) \end{cases},\tag{9}$$

where $\mathscr{F}$ is the magneto-motive force (m.m.f.), ϕ is the magnetic flux, and n_{t} is the number of winding turns of

an inductive coil. The port facing the electrical subcircuit is called *electric port*, and the relative signals are marked with the subscript "e," whereas the port facing the magnetic subsystem is referred to as *magnetic port*, and the relative signals are marked with the subscript "m." We set $v_{\mathrm{e}} = v$, $v_{\mathrm{m}} = \mathscr{F}$, $i_{\mathrm{e}} = i$, and $i_{\mathrm{m}} = \phi$, and we rewrite (9) to express the electrical variables as functions of the magnetic variables

$$\begin{cases} v_{\mathrm{e}}(t) = -n_{\mathrm{t}}\frac{di_{\mathrm{m}}(t)}{dt} \\ i_{\mathrm{e}}(t) = \frac{1}{n_{\mathrm{t}}}v_{\mathrm{m}}(t) \end{cases}.\tag{10}$$

In order to implement the ME junction in the discrete-time domain, the time derivative in (10) can be discretized using the Backward Euler method obtaining

$$\begin{cases} v_{\mathrm{e}}[k] = -\frac{n_{\mathrm{t}}}{T_{\mathrm{s}}}(i_{\mathrm{m}}[k] - i_{\mathrm{m}}[k-1]) \\ i_{\mathrm{e}}[k] = \frac{v_{\mathrm{m}}[k]}{n_{\mathrm{t}}} \end{cases}\tag{11}$$

where T_{s} is the sampling period, and k is the sampling index. To obtain the WD implementation of the ME junction [12, 13] shown in Fig. 2b, we express the Kirchhoff variables in terms of the wave variables a_{e}, b_{e}, a_{m}, and b_{m} by using (2):

$$v_{\mathrm{e}}[k] = \frac{a_{\mathrm{e}} + b_{\mathrm{e}}}{2}, \quad i_{\mathrm{e}}[k] = \frac{a_{\mathrm{e}} - b_{\mathrm{e}}}{2Z_{\mathrm{e}}[k]},\tag{12}$$

$$v_{\mathrm{m}}[k] = \frac{a_{\mathrm{m}} + b_{\mathrm{m}}}{2}, \quad i_{\mathrm{m}}[k] = \frac{a_{\mathrm{m}} - b_{\mathrm{m}}}{2Z_{\mathrm{m}}[k]},\tag{13}$$

where Z_{e} and Z_{m} are the free parameters of the electric and the magnetic ports, respectively. By substituting (12) and (13) into (11) and solving for the reflected waves, we obtain the following system of equations
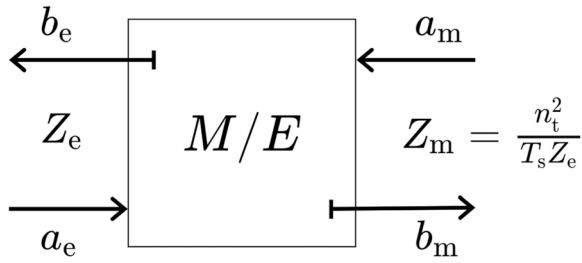
$$\begin{bmatrix} b_{\mathrm{e}}[k] \\ b_{\mathrm{m}}[k] \end{bmatrix} = \mathbf{S}_{\mathrm{ME}}\begin{bmatrix} a_{\mathrm{e}}[k] \\ a_{\mathrm{m}}[k] \end{bmatrix} + \mathbf{M}_{\mathrm{ME}}\begin{bmatrix} a_{\mathrm{m}}[k-1] \\ b_{\mathrm{m}}[k-1] \end{bmatrix},\tag{14}$$

where

$$\mathbf{S}_{\mathrm{ME}} = \begin{bmatrix} -\frac{\frac{2Z_{\mathrm{e}}[k]}{n_{\mathrm{t}}} - \beta[k]}{\beta[k]} & \frac{1}{2}\left[\frac{\left(\frac{2Z_{\mathrm{e}}[k]}{n_{\mathrm{t}}} - \beta[k]\right)^2}{\beta[k]} - \beta[k]\right] \\ \frac{2}{\beta[k]} & -\frac{\frac{2Z_{\mathrm{e}}[k]}{n_{\mathrm{t}}} - \beta[k]}{\beta[k]} \end{bmatrix},\tag{15}$$

$$\mathbf{M}_{\mathrm{ME}} = \begin{bmatrix} \frac{Z_{\mathrm{e}}[k]}{T_{\mathrm{s}}Z_{\mathrm{m}}[k-1]}\frac{1}{\beta[k]} & -\frac{Z_{\mathrm{e}}[k]}{T_{\mathrm{s}}Z_{\mathrm{m}}[k-1]}\frac{1}{\beta[k]} \\ -\frac{n_{\mathrm{t}}}{T_{\mathrm{s}}Z_{\mathrm{m}}[k-1]}\frac{1}{\beta[k]} & \frac{n_{\mathrm{t}}}{T_{\mathrm{s}}Z_{\mathrm{m}}[k-1]}\frac{1}{\beta[k]} \end{bmatrix},\tag{16}$$

and $\beta[k] = \frac{Z_{\mathrm{e}}[k]}{n_{\mathrm{t}}} + \frac{n_{\mathrm{t}}}{T_{\mathrm{s}}Z_{\mathrm{m}}[k]}$. The two matrices $\mathbf{S}_{\mathrm{ME}}$ and $\mathbf{M}_{\mathrm{ME}}$ need to be recomputed whenever a variation of the two port resistances Z_{e} and Z_{m} occurs.

Fig. 3 ME junction showing adapted ports symbolically represented by a T-shaped stub. The magnetic port resistance is set according to the adaptation condition given in (19)

In order to efficiently solve WD structures, it is desirable to remove as many implicit relations as possible. As mentioned in Section 2.1, this can be achieved through the *adaptation process*. Since the diagonal entries of $\mathbf{S}_{\mathrm{ME}}$ are equal, it is possible to make both the electric and the magnetic port reflection-free at the same time by properly setting the free parameters Z_{e} and Z_{m}. By imposing the constraint

$$\frac{\frac{Z_{\mathrm{e}}}{n_{\mathrm{t}}} - \frac{n_{\mathrm{t}}}{T_{\mathrm{s}} Z_{\mathrm{m}}[k]}}{\frac{Z_{\mathrm{e}}}{n_{\mathrm{t}}} + \frac{n_{\mathrm{t}}}{T_{\mathrm{s}} Z_{\mathrm{m}}[k]}} = 0, \tag{17}$$

we can solve for Z_{e} to obtain the adaptation condition for the electric port

$$Z_{\mathrm{e}}[k] = \frac{n_{\mathrm{t}}^2}{T_{\mathrm{s}} Z_{\mathrm{m}}[k]}, \tag{18}$$

or solve for Z_{m} to obtain the adaptation condition for the magnetic port

$$Z_{\mathrm{m}}[k] = \frac{n_{\mathrm{t}}^2}{T_{\mathrm{s}} Z_{\mathrm{e}}[k]}. \tag{19}$$

For example, Fig. 3 shows an ME junction adapted using (19). Finally, it is worth pointing out that other discretization methods can be considered for the implementation of the ME junction, as shown in [13].

2.4 WD structures
WDFs can be organized into tree structures called *connection trees*. Three types of constitutive blocks can be identified in a WD connection tree: the *root*, which has no upward-facing ports and can have one or more downward-facing ports; *nodes* (typically multi-port topological or ME junctions), which have one upward-facing port and one or more downward-facing ports; *leaves* (typically circuit elements), which have upward-facing ports and no downward-facing ports [12].

A WD structure can be solved without employing iterative solvers only if there are no delay-free loops [14]. Delay-free loops are formed every time instantaneous implicit relations exist among wave variables. Breaking delay-free loops at each upward-facing port makes the structure realizable and computable in the WD domain. This is achieved through *adaptation* of all the elements and upward-facing junctions except for the root, which has no upward-facing ports.

2.5 Solving WD structures with at most one nonlinear element
Contrary to linear elements, nonlinear elements cannot be adapted as described in Section 2.1 [12, 18, 19, 38, 39]. However, in the WD domain, it is possible to implement circuits with up to one nonlinear element without resorting to iterative solvers as long as the nonlinear element is characterized by an explicit WD mapping. This is accomplished by choosing the nonlinear element as the root of the connection tree and adapting all nodes and leaves. As done in [12] for solving electromagnetic circuits characterized by an arbitrary number of WD topological junctions, we divide the WD structure in levels, where level $\ell = 1$ contains only the root and level $\ell = L$ contains only leaves, and we index each node/leaf on the ℓ-th level with the subscript u. At each sampling step k, the computational flow comprises four stages, depicted in Fig. 4:

1. *Leaves scattering stage*: the waves reflected by the leaves are computed using their scattering relations (4).
2. *Forward scattering stage*: the wave reflected by each node of level ℓ is computed and propagated towards level $\ell - 1$, until the root is reached. The generic reflected wave is computed as

$$b_{\ell,u,n}[k] = \mathbf{s}_{\ell,u,n}[k]\mathbf{a}_{\ell,u}[k],$$

where $\mathbf{s}_{\ell,u,n}$ is the nth row of the scattering matrix $\mathbf{S}_{\ell,u}[k]$ corresponding to the port with index n facing level $\ell - 1$. In case the considered node is a ME junction, the reflected wave is computed using one of the scalar scattering relations in (14), depending on which port (electric or magnetic) faces level $\ell - 1$.
3. *Root scattering stage*: the wave reflected by the root is computed according to the WD scattering relation f characterizing the nonlinear element

$$b_{1,1}[k] = f_{1,1}(a_{1,1}[k]).$$

4. *Backward scattering stage*: the waves scattered by the nodes in the ℓth level are computed and propagated towards level $\ell + 1$. The computation starts at level

(a) Leaves Scattering

(b) Forward Scattering

(c) Root Scattering

(d) Backward Scattering
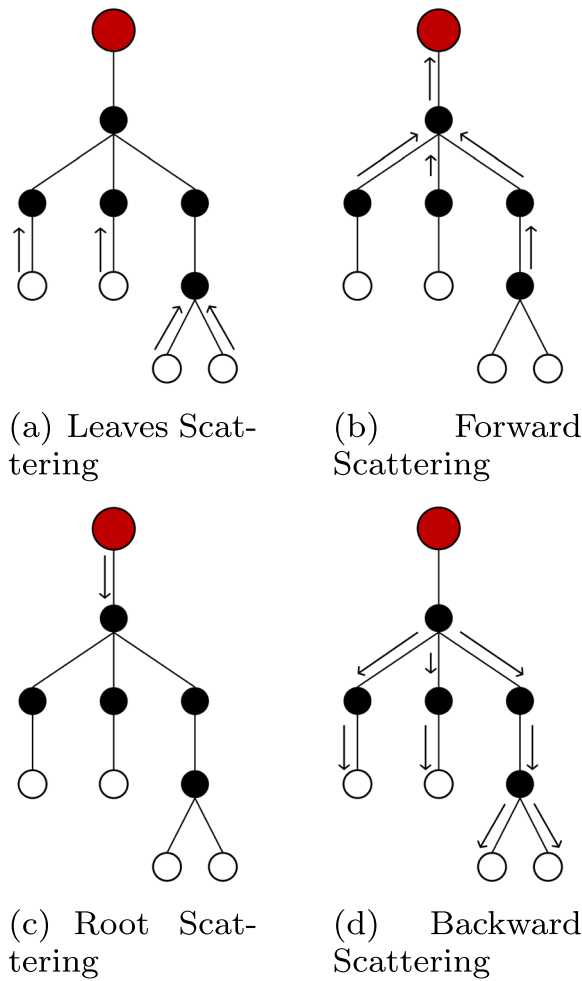
Fig. 4 WD connection tree computational flow. Black elements are nodes, white elements are the leaves. The root is the red element

$\ell = 2$ and ends at $\ell = L - 1$. The vector of reflected waves is evaluated as

$$\mathbf{b}_{\ell,u}[k] = \mathbf{S}_{\ell,u}[k]\mathbf{a}_{\ell,u}[k].$$

In case the considered node is a ME junction, the reflected wave is computed using one of the scalar scattering relations in (14), depending on which port (electric or magnetic) faces level $\ell + 1$.

3 Background on hysteresis modeling

In the literature, two main categories of hysteresis models can be found: physical and phenomenological. Physical models are based on the physics laws governing the target system [56, 57]. Typically, such models are mathematically complex and require detailed knowledge of the physical properties of the materials. Phenomenological models, instead, make use of conventional system identification

techniques and often lack a direct physical interpretation [58, 59]. Most phenomenological models approximate the whole hysteretic nonlinearity by weighting elementary hysteretic operators, known as *hysterons*, characterized by a simple mathematical description [60–62]. Such models are mostly empirical and rely on acquired experimental data [60].

The most popular operator-based phenomenological model is the Preisach model [63], which is used in a wide range of applications. The Preisach model is a *rate-independent* model, which means that the variation rate of the input signal within a given range does not affect the shape of the hysteresis loop. If instead the output hysteresis depends both on the value of the input and on the speed at which it changes, we call it *rate-dependent* hysteresis or *dynamic* hysteresis. Although rate-dependent extensions of the Preisach model exist [64, 65], they entail solving computationally-intensive parameter identification problems, which make their use in practical applications very challenging.

More recently, starting from the definition of the Preisach model, new phenomenological methods based on neural networks have been introduced in the literature. Such methods model rate-dependent hysteretic nonlinearities in a data-driven fashion, relying on physical measurements of ferromagnetic or ferroelectric materials [44–46, 49]. Specifically, Farrokh et al. [44] proposed a multi-layer feedforward neural network called extended Preisach neural network (XPNN), based on newly defined hysteron-like neurons, which proved to be capable of simulating both rate-independent and rate-dependent hysteresis loops. Chen et al. [46] proved that diagonal recurrent neural networks (dRNNs) are able to realize the superposition of a number of rate-dependent hysterons. Moreover, Amodeo et al. [49] employed multilayer nonlinear autoregressive exogenous neural networks (NARX) for both quasi-static and dynamic hysteretic modeling of iron-dominated magnets. In the following, we focus on the recently proposed Preisach-RNN [45] which showed promising results for predicting the behavior of dynamic hysteresis in ARMCO pure iron.

3.1 Preisach-RNN

The Preisach-RNN [45] is a rate-dependent implementation of the Preisach model based on a single-layer RNN [66, 67].

The traditional mathematical formulation of the Preisach model [63, 68] can be written as

$$\hat{y}(t) = \iint_{\alpha \geq \beta} \mu(\alpha, \beta)\gamma_{\alpha\beta}(u(t))d\alpha d\beta, \tag{20}$$
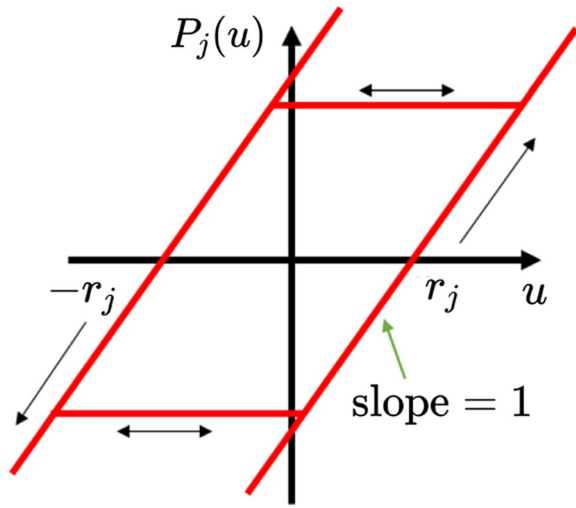
Fig. 5 Play Operator $P_j(u)$

where $\hat{y}(t)$ is the model output at time t, $u(t)$ is the model input at time t, whereas $\mu(\alpha, \beta)$ is the density function that weights the elementary rectangular hysteresis operators $\gamma_{\alpha\beta}$, also called Preisach hysterons [46, 47], with α and β being the ascending and descending switching thresholds, respectively. Applying the following change in coordinates [69]

$$r = \frac{\alpha - \beta}{2}, \quad \nu = \frac{\alpha + \beta}{2}, \quad \hat{\mu} = \mu(\nu + r, \nu - r),$$
(21)

the Preisach half-plane $\{(\alpha, \beta) \,|\, \alpha \geq \beta\}$ is mapped onto the half-plane $\{(r, \nu) \,|\, r > 0, \, \nu \in \mathbb{R}\}$, where the boundary between the $+1$ and -1 regions is described by the curve $\nu = P(u(t))$, known as *Play Operator*. This allows us to rearrange (20) as

$$\hat{y}(t) = \int_0^\infty \rho(r, P(u(t)))dr,$$
(22)

where the function ρ is defined as

$$\rho(r, P(u(t))) = \int_{-\infty}^{P(u(t))} \hat{\mu}(r, \nu)d\nu \; - \int_{P(u(t))}^{\infty} \hat{\mu}(r, \nu)d\nu.$$
(23)

Equation (22) can be approximated by considering only M operators, yielding

$$\hat{y}(t) = \sum_{j=1}^{M} \varphi_j P_j(u(t)),$$
(24)

where $P_j(u(t))$ is the jth Play Operator, and φ_j represents its density function. In the literature, (24) is commonly referred to as Prandtl-Ishlinskii model [44, 70]. Figure 5 shows a Play Operator, which is defined as

$$P_j(u(t)) = \max(u(t) - r_j,$$
$$\min(u(t) + r_j, P_j(u(t - 1)))),$$
(25)

$$P_j(u(0)) = \max(u(0) - r_j,$$
$$\min(u(0) + r_j, \kappa_0)),$$
(26)

where κ_0 is the initial condition for the operator and r_j, which represents the discrete counterpart of r, is defined as

$$r_j = \frac{j - 1}{M}(\max(u(t)) - \min(u(t))),$$
(27)

for $j = 1, 2, \ldots, M$.

The idea behind the Preisach-RNN, shown in Fig. 6, is to model the density function φ_j in (24) in the discrete-time domain using an RNN. Furthermore, RNNs allow us to model rate-dependent hysteresis, extending the traditional Preisach model to this case. The hidden state of a U-node RNN at sample k is computed as

$$\mathbf{h}[k] = f_h(\mathbf{W}_{xh}\mathbf{x}[k] + \mathbf{W}_{hh}\mathbf{h}[k - 1] + \mathbf{b}_h),$$
(28)

where $\mathbf{h}[k]$, $\mathbf{h}[k - 1] \in \mathbb{R}^U$ are the current and previous hidden states. The scalar output at sample k is computed as

$$\hat{y}[k] = f_o\left(\mathbf{w}_{hy}^T \mathbf{h}[k] + b_y\right).$$
(29)

The input vector $\mathbf{x}[k] \in \mathbb{R}^{M+2}$ is built concatenating the input signal $u[k]$, the input derivative $\dot{u}[k]$, and M Play Operators $P_j(u[k])$ with $j = 1, \ldots, M$.
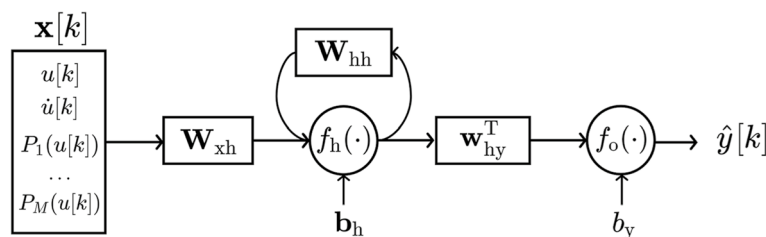
Fig. 6 Diagram of a Preisach-RNN architecture

$\mathbf{W}_{xh} \in \mathbb{R}^{U \times (M+2)}$, $\mathbf{W}_{hh} \in \mathbb{R}^{U \times U}$, $\mathbf{w}_{hy} \in \mathbb{R}^{U}$, $\quad \mathbf{b}_h \in \mathbb{R}^{U}$, and $b_y \in \mathbb{R}$ are the network weights and biases. Finally, the hidden state activation function f_h, which is applied element-wise to the output of each hidden neuron, is the hyperbolic tangent, and the output activation function f_o is a linear activation, i.e., $f_o(x) = x$.

4 Wave digital hysteretic nonlinearities

The conventional approach to develop a magnetic equivalent circuit, employed, for example, in [71], is based on the analogy between magneto-motive force $\mathscr{F}$ and electric voltage v and between the magnetic flux ϕ and electric current i [55]. In the rest of this manuscript, with an abuse of nomenclature, we will generally refer to both pairs of v-i variables and ϕ-$\mathscr{F}$ variables at the circuit ports as variables in the *Kirchhoff domain*, in order to distinguish them from the corresponding variables in the *WD domain*. The electrical variables v and i in linear resistors are related by the Ohm's law $v = Ri$, whereas the two magnetic variables are related by an equivalent linear and instantaneous Ohm-like law

$$\mathscr{F} = \mathscr{R}\,\phi, \tag{30}$$

known as Hopkinson's law, which relates $\mathscr{F}$ and ϕ through the *reluctance* parameter $\mathscr{R}$, which is analogous to the electrical resistance. Since, as a first approximation, higher-order magnetic effects can be considered negligible in the audio frequency band, we can assume ϕ to be uniform along cross sections, and thus that the topology of the magnetic equivalent circuit can be directly derived from the magnetic structure: each winding of n_t turns is represented by an ideal magnetic voltage generator $\mathscr{F} = n_t i$, whose polarity is given by the sign convention depicted in Fig. 7, where i is the winding current. The magnetic path is represented by an equivalent reluctance, possibly nonlinear, whose value depends on the geometry and physical properties of the magnetic material. Although (30) refers to the linear case, in the

following we will consider a nonlinear mapping between ϕ and $\mathscr{F}$ to address rate-dependent hysteresis.

While the reluctance defines the constitutive relation between the magnetic variables ϕ-$\mathscr{F}$, the pair B-H, where B is the flux density and H is the magnetic field, is related by the permeability. In fact, as a first approximation, the relation between B and H can be defined by the formula $B = \mu_0 \mu_r H$, where μ_0 is the vacuum permeability and μ_r is the relative permeability. Moreover, we can convert the B-H curve of the material into its ϕ-$\mathscr{F}$ representation as follows

$$\phi = B\Lambda, \quad \mathscr{F} = H\Gamma, \tag{31}$$

where Λ and Γ are the cross-section and the length of the magnetic path expressed in meters, respectively [12, 13, 55]. Therefore, if we consider the magnetic structure to be made of a magnetic material with a B-H characteristic exhibiting rate-dependent hysteresis, the equivalent nonlinear reluctances that model the different magnetic paths across the magnetic material are going to exhibit hysteresis. The dynamic nonlinearity is thus confined into the constitutive equation of the reluctance $\mathscr{R}$, i.e., the relation between ϕ and $\mathscr{F}$.

In the Kirchhoff discrete-time domain, we can recast the problem of modeling the constitutive equation of a reluctance with rate-dependent hysteresis into a nonlinear regression problem, i.e.,

$$\hat{\phi}[k] = g(\mathscr{F}[k], \mathscr{F}[k-1], \ldots, \mathscr{F}[0]\,; \theta_{KD}), \tag{32}$$

where the mapping g is modeled by a suitable RNN. The set of parameters θ_{KD} is obtained by training the network in the Kirchhoff domain to predict the current value of the magnetic flux $\phi[k]$ given the input time-series $\mathscr{F}[k], \mathscr{F}[k-1], \ldots, \mathscr{F}[0]$. RNNs are particularly suitable for modeling non-instantaneous nonlinearities, because they are fed with input time-series data and use recurrent connections to implement an infinite dynamic response.

(a)   (b)
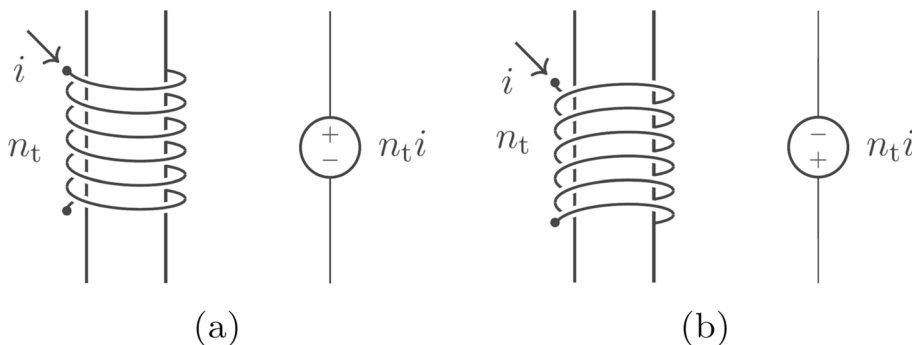
Fig. 7 Sign convention of magnetic voltage generators. **a** Coil wound counter-clockwise. **b** Coil wound clockwise. The figure is taken from [12]

The strategy employed to avoid the shortcomings of learning long-term dependencies from excessively long time-series [72] is to limit the length of the input sequence. Each time-series is thus split into sequences of length K, which are then sequentially fed to the RNN. To ensure long-term memory, however, we propagate the hidden states of the recurrent layers between consecutive sequences, according to a cross-batch statefulness paradigm. Hence, the rate-dependent hysteresis nonlinear regression problem in (32) can be rewritten as

$$\hat{\phi}[k] = g(\mathscr{F}[k], \ldots, \mathscr{F}[k - K + 1]; \theta'_{\text{KD}}), \qquad (33)$$

where θ'_{KD} are obtained by training a stateful RNN.

A similar approach can be adopted in the WD domain by properly converting Kirchhoff variables into wave variables according to (1). This yields

$$\hat{b}[k] = g(a[k], \ldots, a[k - K + 1]; \theta'_{\text{WD}}), \qquad (34)$$

where θ'_{WD} is the set of parameters obtained by training the neural network to predict the reflected wave $b[k]$ given an input sequence composed of the incident wave $a[k]$ and its $K - 1$ previous values, thus obtaining an explicit scattering relation. This relation can be then used to implement the one-port WD realization of a nonlinear reluctance with rate-dependent hysteresis.

It is interesting to note that the newly defined WD block shares some common characteristics with WD models of linear dynamic elements such as capacitors and inductors [19], as their behavior depends on (buffers of) past samples of wave variables.

In the following, we implement the rate-dependent hysteretic mapping g using a Preisach-RNN architecture (see Section 3.1) with $U = 32$ hidden units and $M = 8$ Play Operators.

5 Model training and evaluation

We hereby present the training procedure of the WD hysteretic block for the specific application scenario described in Section 6. In particular, we will assume that a single nonlinear reluctance component is sufficient to model the characteristics of the constitutive magnetic material. Let us consider the magnetic components to be made of Pearlitic Steel R260, whose measurements are used to determine the *B-H* hysteretic behavior under consideration.

Presented in [73], the available dataset [74] consists of magnetic measurements obtained driving the system with a triangular input current sampled at 100 kHz having 11 different input frequencies, i.e., 0.5, 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000 Hz. For each frequency, the curves include measurements of two periods of the triangular input signal *H* and the corresponding flux density

B: the first period contains the first magnetization curve, whereas the second period fully describes the main hysteresis loop. The first magnetization curve is defined as the *B-H* curve branch that describes the material magnetization process starting from a state of no magnetization ($B = 0$). A suitable measurement procedure would involve a complete de-magnetization of the magnetic material between two consecutive acquisitions, leading to the same initial condition $H[0] = B[0] = 0$, regardless of the input frequency. However, each hysteresis loop in the available dataset is characterized by a different initial magnetization value $B[0]$, probably due to incomplete de-magnetization. In fact, in order to obtain a coherent dataset, it would be desirable to have each hysteresis curve characterized by the same $B[0]$. With the aim of reducing possible inconsistencies during training, we preprocess the dataset by excluding the first period of each curve and circularly shifting the result by a quarter of a period. As shown in Fig. 8, this ensures that all training examples start from the same rate-independent value, i.e., the magnetic saturation value. Furthermore, we discard the 0.5 Hz measurements due to discontinuities in the resulting *B-H* curve. We repeat periods to let all remaining curves match the duration of two periods of the lowest frequency curve (1 Hz); this corresponds to fixing the duration of the measurement session to two seconds for all input rates and allows us to balance the training set across frequencies. To transform the *B-H* pair of variables into WD variables *a-b*, it is first necessary to convert the Pearlitic Steel R260 curves into the corresponding ϕ-$\mathscr{F}$ curves, by using the equalities in (31). Then, since $\mathscr{F}$ is the magnetic equivalent of voltage and ϕ is the equivalent of current, the dataset can be expressed in the WD domain via a transformation similar to (1), i.e., [12]

$$a = \mathscr{F} + Z\phi, \quad b = \mathscr{F} - Z\phi, \qquad (35)$$

where the free parameter Z is fixed and can be set according to the adaptation condition of the junction port to which the reluctance is connected. The inverse transformation of (35) is given by

$$\mathscr{F} = \frac{a + b}{2}, \quad \phi = \frac{a - b}{2Z}. \qquad (36)$$

The wave variables are then rescaled in $[-1, 1]$. In order to accomplish this, the values $a_{\min}$, $a_{\max}$, $b_{\min}$, $b_{\max}$ are estimated from the WD data, and used to scale the wave variables according to

$$\begin{aligned} \tilde{a}[k] &= 2 \cdot \frac{a[k] - a_{\min}}{a_{\max} - a_{\min}} - 1, \\ \tilde{b}[k] &= 2 \cdot \frac{b[k] - b_{\min}}{b_{\max} - b_{\min}} - 1, \end{aligned} \qquad (37)$$
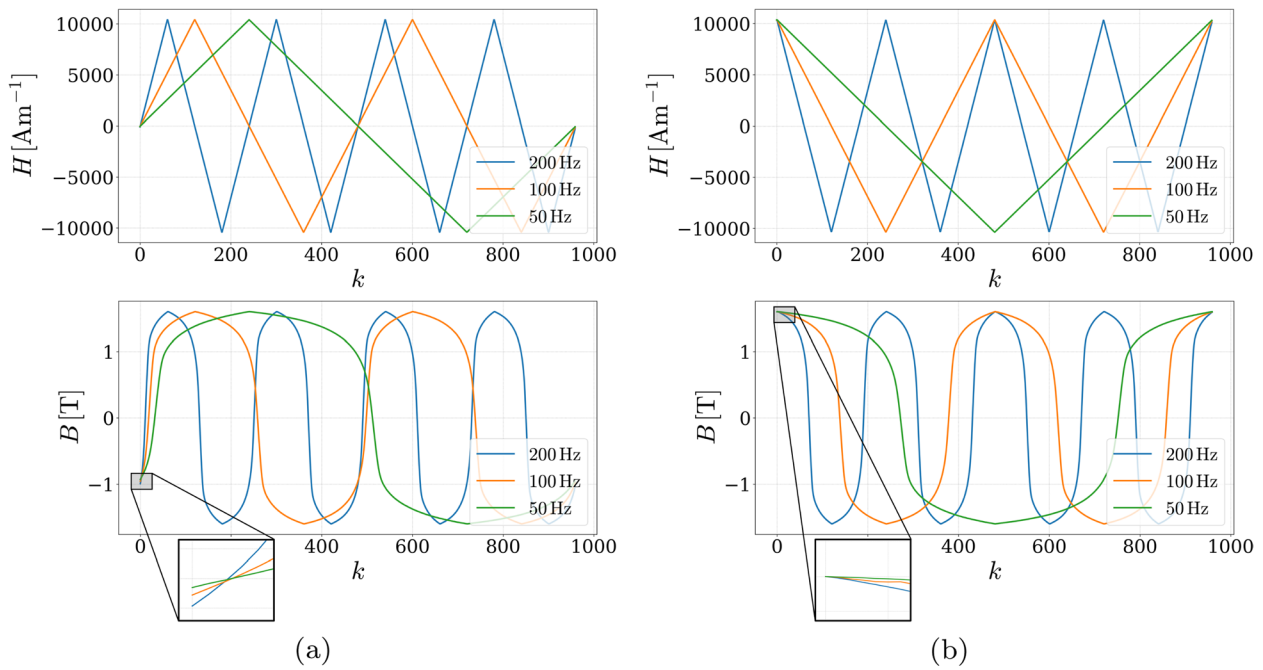
Fig. 8 **a** Exemplificative *H* and *B* curves contained in the measurement dataset before time shift. **b** The corresponding *H* and *B* curves after circular time shift. The curves are starting from positive saturation values

where $\tilde{a}[k]$ is the scaled incident wave, and $\tilde{b}[k]$ is the scaled reflected wave. It is then possible to scale back $\tilde{a}[k]$ and $\tilde{b}[k]$ into their original range using the following equations

$$
\begin{aligned}
a[k] &= a_{\min} + \frac{\tilde{a}[k] + 1}{2} \cdot (a_{\max} - a_{\min}), \\
b[k] &= b_{\min} + \frac{\tilde{b}[k] + 1}{2} \cdot (b_{\max} - b_{\min}).
\end{aligned} \tag{38}
$$

Given the audio application scenario, data are resampled at a common audio sampling frequency, i.e., $f_s = 48$ kHz. Input sequences $\tilde{\boldsymbol{a}}[k] = \tilde{a}[k], \dots, \tilde{a}[k - K + 1]$ are obtained by shifting a rectangular window of length $K = 20$ over the input waves with unitary hop-size and assigning the corresponding ground truth value $\tilde{b}[k]$ to each of them. The pairs $(\tilde{\boldsymbol{a}}[k], \tilde{b}[k])$ are assembled in batches containing one sequence for each input rate included in the training set; this is done to let all training frequencies contribute to each optimization step. The Preisach-RNN described in Section 4 is implemented in Python using Pytorch [75] and comprises 1441 trainable parameters. It is trained to minimize the following loss function defined in the Kirchhoff domain:

$$
\begin{aligned}
\mathcal{L} &= \mathcal{E}(\mathcal{F}, \hat{\mathcal{F}}) + \mathcal{E}(\phi, \hat{\phi}) \\
&= \mathcal{E}\left(\frac{a + b}{2}, \frac{a + \hat{b}}{2} \right) + \mathcal{E}\left(\frac{a - b}{2Z}, \frac{a - \hat{b}}{2Z} \right),
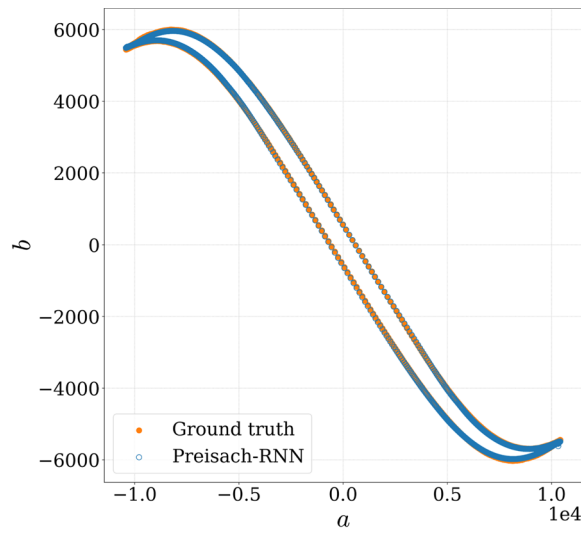\end{aligned} \tag{39}
$$

where

$$
\mathcal{E}(y, \hat{y}) = \frac{\sum_k (y[k] - \hat{y}[k])^2}{\sum_k y^2[k]} \tag{40}
$$

is the normalized mean squared error (NMSE), whereas $a[k]$ and $b[k]$ are obtained from the scaled network inputs $\tilde{a}[k]$ and outputs $\tilde{b}[k]$ through (38). Notably, the loss function in (39) comprises two NMSE terms, one for $\hat{\mathcal{F}}$ and one for $\hat{\phi}$, as both depend on the predicted wave $\hat{b}$.
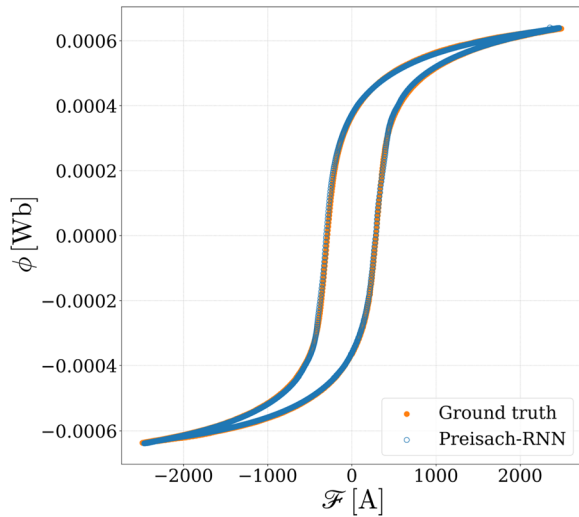
To evaluate the model, we perform leave-one-out cross-validation (LOOCV). Namely, we train ten different Preisach-RNNs, each time selecting nine of the ten frequencies as training set and using the remaining one for evaluation. Each training consists of ten epochs using Adam [76] and a learning rate of 10^{-4}, and it is run on a single NVIDIA TITAN V with 12 GB of RAM. The results are reported in Table 1. Figure 9a shows the model predictions for the test curve at 20 Hz in the WD domain, whereas Fig. 9b shows the corresponding Kirchhoff

Table 1 Leave-one-out cross-validation results

	1Hz	2Hz	5Hz	10Hz	20Hz	50Hz	100Hz	200Hz	500Hz	1000Hz	Average
$\mathcal{E}(b, \hat{b})$	$1.66 \cdot 10^{-4}$	$4.10 \cdot 10^{-5}$	$3.67 \cdot 10^{-5}$	$2.24 \cdot 10^{-5}$	$2.45 \cdot 10^{-5}$	$1.55 \cdot 10^{-5}$	$3.73 \cdot 10^{-4}$	$1.05 \cdot 10^{-4}$	$2.90 \cdot 10^{-4}$	$4.73 \cdot 10^{-4}$	$1.55 \cdot 10^{-4}$
$\mathcal{E}(\mathscr{F}, \hat{\mathscr{F}})$	$5.96 \cdot 10^{-4}$	$1.46 \cdot 10^{-4}$	$1.31 \cdot 10^{-4}$	$7.99 \cdot 10^{-5}$	$8.75 \cdot 10^{-5}$	$5.60 \cdot 10^{-5}$	$1.39 \cdot 10^{-3}$	$3.90 \cdot 10^{-3}$	$1.10 \cdot 10^{-3}$	$1.62 \cdot 10^{-3}$	$9.11 \cdot 10^{-4}$
$\mathcal{E}(\phi, \hat{\phi})$	$2.67 \cdot 10^{-5}$	$6.65 \cdot 10^{-6}$	$5.93 \cdot 10^{-6}$	$3.64 \cdot 10^{-6}$	$3.97 \cdot 10^{-6}$	$2.52 \cdot 10^{-6}$	$6.25 \cdot 10^{-5}$	$1.65 \cdot 10^{-4}$	$4.87 \cdot 10^{-5}$	$7.80 \cdot 10^{-5}$	$4.04 \cdot 10^{-5}$

Fig. 10 A possible output stage of vacuum tube guitar amplifier

(a)

(b)

Fig. 9 a Predictions of a Preisach-RNN in the WD domain with $U = 32$ hidden units and $M = 8$ Play Operators (blue) vs. the ground truth (orange). Input rate: $f = 20$Hz. **b** WD predictions transformed into Kirchhoff domain variables by means of (36)

variables obtained applying the inverse wave transformation in (36). Despite the limited amount of available data, LOOCV shows an average NMSE in the order of magnitude of 10^{-4} for both the WD variables and the Kirchhoff domain variables. These results suggest that the network may exhibit good generalization properties when used to predict waves with an input rate that was not included in the training set. In turn, this gives us confidence that the proposed rate-dependent hysteresis model, trained on the entire measurement dataset, could be successfully applied in a discrete-time circuital simulation scenario.

In the next section, we will describe the use of the proposed WD hysteretic block for the emulation of the output stage of a vacuum tube guitar amplifier.

6 Example of application

As a reference circuit, let us consider the push-pull output stage of a vacuum tube guitar amplifier shown in Fig. 10. Let us assume that the output stage consists only of the nonlinear three-winding audio transformer directly driving the loudspeaker.[1] The secondary side of the transformer is connected to the speaker, modeled by means of the series between the resistor R_{L} and the inductor L_{L}. Without loss of generality, we consider two identical input signals V_{in1} and V_{in2} and, thus, two
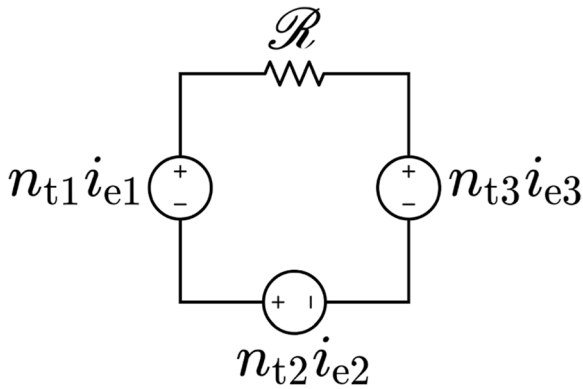
---

[1] If the push-pull power amplifier was to be taken into account, one could resort to the modeling approach discussed in [28] and drive the audio transformer with the output of such an additional stage.
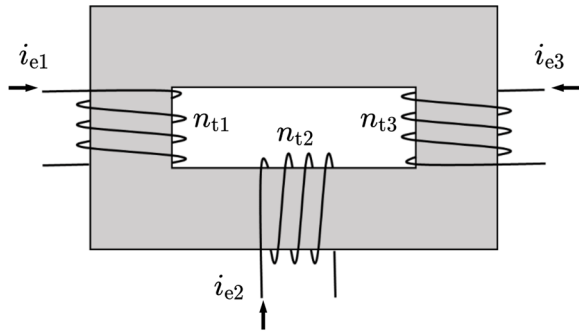
Fig. 11 Core magnetic structure under consideration
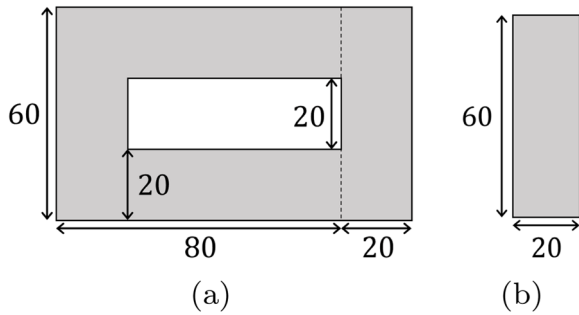
Fig. 12 Core geometry of the output transformer. All dimensions are in millimeters (10^{-3} m). **a** Front view. **b** Side view

Fig. 13 Equivalent circuit model of the magnetic structure shown in Fig. 11

primary windings with n_{t1} and n_{t2} turns, respectively. The secondary side of the transformer presents a single winding with n_{t3} turns. R_{p1} and R_{p2} are the primary coil resistances, whereas R_s corresponds to the secondary coil. Let us assume that the transformer in Fig. 10 has a UI geometry. We also assume that the three windings are positioned on the core structure, as depicted in Fig. 11.

The core geometry is taken from one of GRAU GmbH datasheets [77], and all of its dimensions are reported in Fig. 12.

The magnetic equivalent circuit of the UI core structure shown in Fig. 11 is derived according to Section 4. The result is shown in Fig. 13, where we notice the m.m.f. sources modeling the three windings with n_{t1}, n_{t2}, and n_{t3} turns, respectively, as well as the nonlinear reluctance $\mathscr{R}$ modeling the magnetic material. Given its specific topology and disregarding the effect of eddy currents or other higher-order effects, a single nonlinear reluctance is, in fact, enough to model the whole magnetic core [13, 71]. The geometric parameters of the magnetic path of the considered UI core are $\Gamma = 240$ mm and $\Lambda = 400$ mm^2. The circuit parameters are summarized in Table 2.

Once we derived the magnetic subcircuit, we connect it to the electrical subcircuits by means of the ME junctions (introduced in Section 2.3), thus obtaining a modular multiphysics model [12, 13]. The multiphysics model of the reference circuit in Fig. 10 is shown in Fig. 14. The magnetic domain is represented by the central subcircuit, and it is coupled to the electrical subcircuits by means of three ME junctions.

The WD realization of the circuital model in Fig. 14 is shown in Fig. 15. The scattering matrix $\mathbf{S}_{s1}$ of the WD 4-port junction that embeds the topological information related to the circuit at the secondary side of the transformer can be computed substituting the fundamental loop matrix $\mathbf{B}_{s1} = \begin{bmatrix} -1 & 1 & 1 & 1 \end{bmatrix}$ into (8). Port 1 is connected to the electric port of junction M_1/E_3, resistor R_s is connected to port 2, while resistor R_L and inductor L_L are connected to port 3 and 4, respectively. All these circuit elements are linear and can be thus adapted as described in Section 2.1. The scattering matrix $\mathbf{S}_{s2}$ of the WD 4-port junction related to the magnetic subcircuit is again obtained from (8) but considering $\mathbf{B}_{s2} = \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}$. Ports 2, 3, and 4 are connected to the magnetic port of junction M_1/E_1, M_1/E_2, and M_1/E_3, respectively. The resistive voltage sources V_{in1} and V_{in2} at the primary side of the transformer are also linear, and are connected to the electric ports of junction M_1/E_1 and M_1/E_2, respectively. Finally, the WD one-port block related to the nonlinear

Table 2 Values of the parameters of the circuit in Fig. 10

R_{p1} [Ω]	n_{t1}	R_{p2} [Ω]	n_{t2}	R_s [Ω]	n_{t3}	R_L [Ω]	L_L [mH]
5	25	5	25	5	12	8	0.05
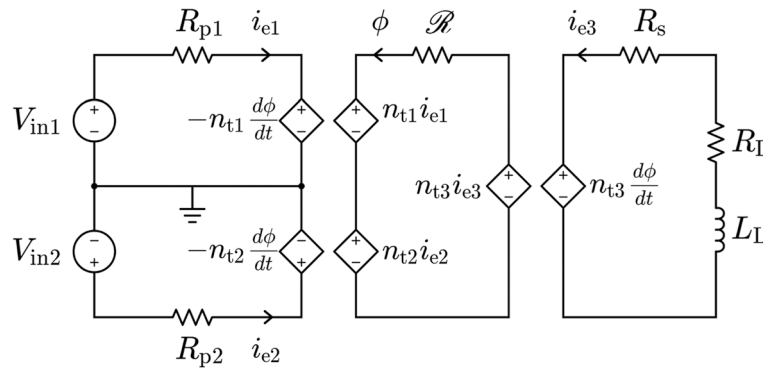
Fig. 14 Output stage of a vacuum tube guitar amplifier including a multiphysics transformer model
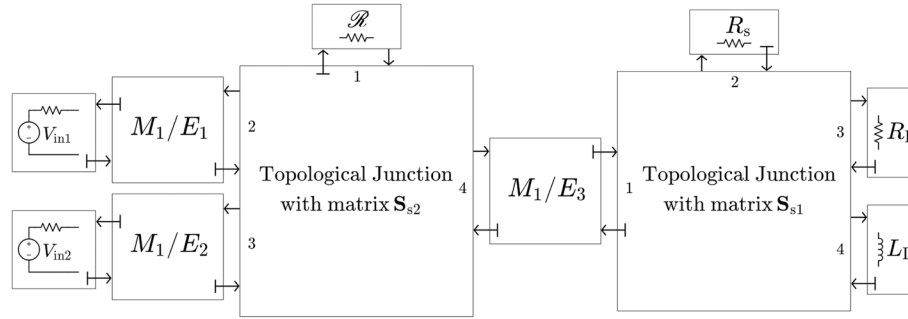
Fig. 15 WD structure implementing the output stage of a vacuum tube guitar amplifier in a multiphysics fashion. The T-shaped stubs indicate port adaptation

reluctance $\mathscr{R}$ is connected to port 1 of the magnetic topological junction as shown in Fig. 15.

The reference circuit contains a single nonlinear one-port element, which means that it is possible to solve the WD structure by employing the algorithm described in Section 2.5 and illustrated in Fig. 4. Since, as a first approximation, the two generators are assumed to be identical, we set $V_{in1} = V_{in2} = A \cos(2\pi k f_0 / f_s)$, where k is the sampling index and f_s is the sampling frequency. We choose the frequency $f_0 = 50$ Hz. We set $A = 250$ V, a value high enough to force the transformer to reach core saturation, in accordance to the available measurement dataset [74].

Before the main WD simulation loop, we run an initialization phase to define the initial hidden state of the WD Preisach-RNN. Namely, we set the signals $V_{in1} = V_{in2} = A$, i.e., the first value of the sinusoidal input signal, and we run the discrete-time simulation for two seconds. Both the input buffer and the hidden states of the WD Preisach-RNN are populated using the values obtained at the end of the initialization loop. The initialization of incident and reflected wave variables at the ports of ME junctions is also performed in the same way.

6.1 Results
In this subsection, we discuss the numerical results obtained from the simulation of the WD structure shown in Fig. 15. The simulation of a single input period takes on average 784 ms on a laptop-mounted Intel Core i5-1240P 1.70 GHz CPU. Figure 16 shows the voltage v_{Z_L} across the series between resistor R_L and inductor L_L, which models the loudspeaker connected to the secondary side of the transformer, whereas Fig. 17 shows the operation points on the nonlinear reluctance curve visited during the simulation. Being proportional to the derivative of the magnetic flux, the voltage v_{Z_L} exhibits sharp peaks associated to magnetic core transitions from a positive magnetic flux saturation region to a negative saturation region, and vice versa. It is difficult to quantify the accuracy of the simulation results shown in Fig. 16, due to the fact that there is no easy way to simulate such a rate-dependent nonlinearity within existing circuit simulation software such as LTspice or Mathworks Simscape. However, referring to Fig. 17, we may state that the hysteretic curve is correctly visited throughout the discrete-time simulation, raising our confidence as far as the accuracy of the proposed method is concerned. The predicted curve reacts
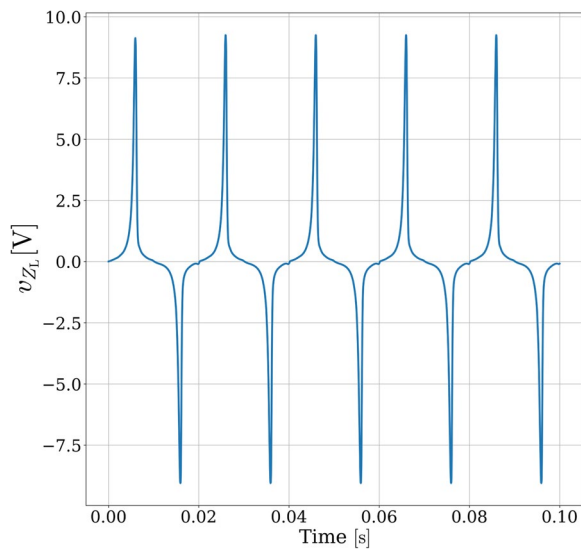
Fig. 16 Voltage across the series of resistor R_L and inductor L_L
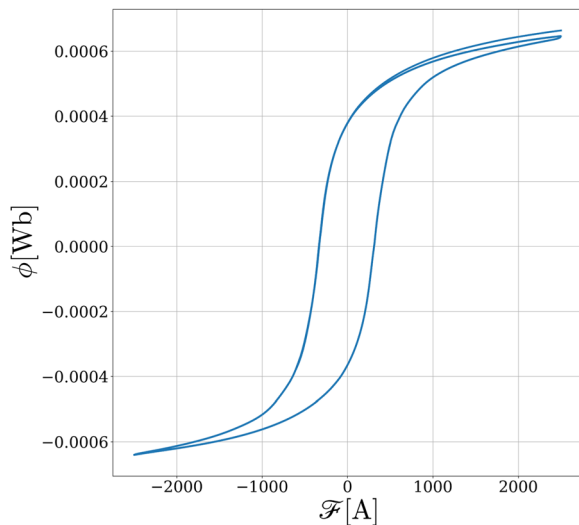
Fig. 17 Hysteresis characteristics of reluctance $\mathcal{R}$ visited during the WD simulation

to different input frequencies f_0 with an hysteresis that is in fact comparable to the physical measurements contained in the dataset presented in Section 5.

7 Conclusions

The modeling and discrete-time circuit simulation of magnetic hysteresis is a notoriously challenging task, especially due to its rate-dependent nature. For this reason, despite the pervasive presence of magnetic components in analog audio gear, circuits with hysteretic elements are usually not tabled for virtual analog

applications. In this manuscript, we explored, for the first time, the possibility of using an RNN-based architecture to model hysteretic nonlinear elements in the WD domain. By properly converting the training data expressed as Kirchhoff variables into wave variables, we defined a data-driven WD circuital block that encapsulates a neural network capable of modeling reluctances with rate-dependent hysteresis. Thus, we successfully employed the proposed WD block for the emulation of the output stage of a vacuum tube guitar amplifier, where the nonlinear transformer is modeled in a multiphysics fashion. More in general, this work does not only constitute the first example of using RNNs to model rate-dependent hysteresis behaviors in the WD domain but also a first step into the exploration of deep learning-based solutions for the WD modeling of nonlinearities with memory for virtual analog applications.

Future work may concern refining the proposed model by considering a dataset of magnetic measurements at input rates spanning the entire audio bandwidth that includes minor hysteresis loops. A noteworthy extension would also be integrating the WD hysteresis block into audio circuits with multiple nonlinearities, which can be then efficiently emulated in the WD domain by exploiting iterative techniques, such as the hierarchical scattering iterative method introduced in [12, 13]. We envision scenarios where data-driven methods could be further developed to supplement the WDF framework in the characterization of circuit nonlinearities directly from experimental measurements. This is done with the full awareness that, in the future, the availability of more efficient simulation algorithms and of more computational power will lead the way towards the real-time implementation of increasingly complex audio circuits.

Abbreviations

BJT	Bipolar junction transistor
CPWL	Canonical piecewise linear
dRNN	Diagonal recurrent neural network
LOOCV	Leave-one-out cross-validation
ME	Magnetic/electric
m.m.f	Magneto-motive force
NARX	Nonlinear autoregressive exogenous
NMSE	Normalized mean squared error
NR	Newton-Raphson
RNN	Recurrent neural network
VA	Virtual analog
WD	Wave digital
WDF	Wave digital filter
XPNN	Extended Preisach neural network

Authors' contributions
A.I.M., R.G., and A.B. conceptualized the study and the method. O.M. implemented the codebase, run the experiments, and wrote the initial draft

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

1. J. Pakarinen, V. Välimäki, F. Fontana, V. Lazzarini, J. Abel, Recent advances in real-time musical effects, synthesis, and virtual analog models. EURASIP J. Adv. Sig. Proc. **2011** (2011). https://doi.org/10.1155/2011/940784
2. V. Valimaki, F. Fontana, J.O. Smith, U. Zolzer, Introduction to the special issue on virtual analog audio effects and musical instruments. IEEE Trans. Audio Speech Lang. Process. **18**(4), 713–714 (2010). https://doi.org/10.1109/TASL.2010.2046449
3. G. De Sanctis, A. Sarti, Virtual analog modeling in the wave-digital domain. IEEE Trans. Audio Speech Lang. Process. **18**(4), 715–727 (2009)
4. J.A. Ewing, W. Thomson, X. Experimental researches in magnetism. Philos. Trans. R. Soc. Lond. **176**, 523–640 (1885). https://doi.org/10.1098/rstl.1885.0010
5. S. Chikazumi, C.D. Graham, *Physics of ferromagnetism* (Oxford University Press, Oxford, 1997)
6. G. Bertotti, *Hysteresis in magnetism: for physicists, materials scientists, and engineers* (Gulf Professional Publishing, Houston, 1998)
7. D. Bouvier, T. Hélie, D. Roze, Phase-based order separation for volterra series identification. Int. J. Control. **94**(8), 2104–2114 (2021)
8. A. Wright, E.P. Damskägg, V. Välimäki et al., Real-time black-box modelling with recurrent neural networks, in *Proc. 22nd Int. Conf. Digital Audio Effects (DAFx-19)*. (University of Birmingham, Birmingham, 2019)
9. D.T. Yeh, J.S. Abel, J.O. Smith, Automated physical modeling of nonlinear audio circuits for real-time audio effects-part i: Theoretical development. IEEE Trans. Audio Speech Lang. Process. **18**(4), 728–737 (2010). https://doi.org/10.1109/TASL.2009.2033978
10. G. Borin, G. De Poli, D. Rocchesso, Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems. IEEE Trans. Speech Audio Process. **8**(5), 597–605 (2000). https://doi.org/10.1109/89.861380
11. A. Falaize-Skrzek, T. Hélie, Simulation of an analog circuit of a wah pedal: a port-hamiltonian approach, in *Audio Engineering Society Convention 135*. (Audio Engineering Society, New York, 2013)
12. R. Giampiccolo, A. Bernardini, G. Gruosso, P. Maffezzoni, A. Sarti, Multiphysics modeling of audio circuits with nonlinear transformers. J. Audio Eng. Soc **69**(6), 374–388 (2021)
13. R. Giampiccolo, A. Bernardini, G. Gruosso, P. Maffezzoni, A. Sarti, Multidomain modeling of nonlinear electromagnetic circuits using wave digital filters. Int J. Circ. Theory Appl. **50**(2), 539–561 (2022). https://doi.org/10.1002/cta.3146
14. A. Fettweis, Wave digital filters: Theory and practice. Proc. IEEE **74**(2), 270–327 (1986). https://doi.org/10.1109/PROC.1986.13458
15. R.C.D. Paiva, S. D'Angelo, J. Pakarinen, V. Valimaki, Emulation of operational amplifiers and diodes in audio distortion circuits. EEE Trans. Circ. Syst. II Express Briefs **59**(10), 688–692 (2012). https://doi.org/10.1109/TCSII.2012.2213358
16. A. Bernardini, K.J. Werner, A. Sarti, J.O. Smith, Multi-port nonlinearities in wave digital structures, in *2015 International Symposium on Signals, Circuits and Systems (ISSCS)*. (IEEE, Iasi, 2015), pp.1–4. https://doi.org/10.1109/ISSCS.2015.7203989
17. A. Bernardini, K.J. Werner, A. Sarti, J.O. Smith III, Modeling nonlinear wave digital elements using the Lambert function. IEEE Trans. Circ. Syst. I Regular Pap. **63**(8), 1231–1242 (2016). https://doi.org/10.1109/TCSI.2016.2573119
18. A. Bernardini, K.J. Werner, P. Maffezzoni, A. Sarti, Wave digital modeling of the diode-based ring modulator, in *Proc. 144th Audio Engineering Society Convention*, Milan 2018. (Audio Engineering Society, New York, 2018), Convention Paper #10015
19. A. Bernardini, P. Maffezzoni, A. Sarti, Linear multistep discretization methods with variable step-size in nonlinear wave digital structures for virtual analog modeling. IEEE/ACM Trans. Audio Speech Lang. Process. **27**(11), 1763–1776 (2019). https://doi.org/10.1109/TASLP.2019.2931759
20. S. D'Angelo, L. Gabrielli, L. Turchet, Fast approximation of the Lambert w function for virtual analog modelling. Practice **100**, 8 (2019)
21. J. Chowdhury, C.J. Clarke, in *19th Sound and Music Computing Conference*. Emulating diode circuits with differentiable wave digital filters (SMC Network, Saint-Étienne, 2022), pp. 2–9. https://zenodo.org/record/6566846
22. D.T. Yeh, J.O. Smith, Simulating guitar distortion circuits using wave digital and nonlinear state-space formulations, in *Proc. 11th Int. Conf. Digital Audio Effects (DAFx-08)*. (Helsinki University of Technology, Espoo, 2008), pp.19–26
23. K.J. Werner, V. Nangia, J.O. Smith III, J.S. Abel, Resolving wave digital filters with multiple/multiport nonlinearities, in *Proc. 18th Int. Conf. Digital Audio Effects (DAFx-15)*. (Norwegian University of Science and Technology, Trondheim, 2015), pp.387–394
24. A. Bernardini, A.E. Vergani, A. Sarti, Wave digital modeling of nonlinear 3-terminal devices for virtual analog applications. Circ. Syst. Signal Process. **39**(7), 3289–3319 (2020)
25. L. Kolonko, B. Musiol, J. Velten, A. Kummert, A split-modular approach to wave digital filters containing bipolar junction transistors, in *2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*. (Lansing, IEEE, 2021), pp.840–843
26. J. Pakarinen, M. Karjalainen, Enhanced wave digital triode model for real-time tube amplifier emulation. IEEE Trans. Audio Speech Lang. Process. **18**(4), 738–746 (2009)
27. R. Cauduro Dias de Paiva, J. Pakarinen, V. Välimäki, M. Tikander, Real-time audio transformer emulation for virtual tube amplifiers. J. Adv. Signal Process. **2011**, 1–15 (2011)
28. J. Zhang, J.O. Smith III, Real-time wave digital simulation of cascaded vacuum tube amplifiers using modified blockwise method, in *Proc. 21st Int. Conf. Digital Audio Effects (DAFx-18)*. (University of Aveiro, Aveiro, 2018)
29. C.C. Darabundit, D. Roosenburg, J.O. Smith, Neural net tube models for wave digital filters, in *Proc. 25th Int. Conf. Digital Audio Effects (DAFx20in22)*. (Vienna University of Music and Performing Arts, Vienna, 2022), pp.153–160
30. K.J. Werner, V. Nangia, A. Bernardini, J.O. Smith III., A. Sarti, An improved and generalized diode clipper model for wave digital filters, in *Proc. 139th Audio Engineering Society Convention*. (Audio Engineering Society, New York, 2015)
31. L. Chua, S.M. Kang, Section-wise piecewise-linear functions: canonical representation, properties, and applications. Proc. IEEE **65**(6), 915–929 (1977). https://doi.org/10.1109/PROC.1977.10589
32. A. Bernardini, A. Sarti, Canonical piecewise-linear representation of curves in the wave digital domain, in *2017 25th European Signal Processing Conference (EUSIPCO)*. (IEEE, Kos, Greece, 2017), pp.1125–1129. https://doi.org/10.23919/EUSIPCO.2017.8081383
33. K. Meerkotter, Digital simulation of nonlinear circuits by wave digital filter principles, vol. 1, in *1989 IEEE International Symposium on Circuits and Systems (ISCAS)*. (IEEE, Portland, 1989), pp.720–723. https://doi.org/10.1109/ISCAS.1989.100452
34. A. Sarti, G. De Sanctis, Systematic methods for the implementation of nonlinear wave-digital structures. EEE Trans Circ. Syst. I Regular Pap. **56**(2), 460–472 (2009). https://doi.org/10.1109/TCSI.2008.2001801
35. A. Bernardini, A. Sarti, Biparametric wave digital filters. IEEE Trans. Circ. Syst. I Regular Pap. **PP**, 1–13 (2017). https://doi.org/10.1109/TCSI.2017.2679007

36. S. Petrausch, R. Rabenstein, Wave digital filters with multiple nonlinearities, in *2004 12th European Signal Processing Conference*. (IEEE, Vienna, 2004), pp.77–80
37. M.J. Olsen, K.J. Werner, J.O. Smith III, Resolving grouped nonlinearities in wave digital filters using iterative techniques, in *Proc. 19th Int. Conf. Digital Audio Effects (DAFX-16)*. (Brno University of Technology, Brno, 2016), pp.279–289
38. A. Bernardini, P. Maffezzoni, L. Daniel, A. Sarti, Wave-based analysis of large nonlinear photovoltaic arrays. IEEE Trans. Circ. Syst. I Regular Pap. **65**(4), 1363–1376 (2018). https://doi.org/10.1109/TCSI.2017.275691
39. A. Bernardini, E. Bozzo, F. Fontana, A. Sarti, A wave digital Newton-Raphson method for virtual analog modeling of audio circuits with multiple one-port nonlinearities. IEEE/ACM Trans. Audio Speech Lang. Process. **29**, 2162–2173 (2021). https://doi.org/10.1109/TASLP.2021.3084337
40. A. Sarti, G. De Poli, Toward nonlinear wave digital filters. IEEE Trans. Signal Process. **47**(6), 1654–1668 (1999). https://doi.org/10.1109/78.765137
41. E. Solan, K. Ochs, Wave digital emulation of general memristors. Int. J. Circ. Theory Appl. **46**(11), 2011–2027 (2018)
42. K. Hornik, Approximation capabilities of multilayer feedforward networks. Neural Netw. **4**(2), 251–257 (1991)
43. A.S. Veeramani, J.H. Crews, G.D. Buckner, Hysteretic recurrent neural networks: a tool for modeling hysteretic materials and systems. Smart Mater. Struct. **18**(7), 075004 (2009)
44. M. Farrokh, M.S. Dizaji, F.S. Dizaji, N. Moradinasab, Universal hysteresis identification using extended Preisach neural network. (2019). arXiv preprint arXiv:2001.01559
45. C. Grech, M. Buzio, M. Pentella, N. Sammut, Dynamic ferromagnetic hysteresis modelling using a Preisach-recurrent neural network model. Materials **13**(11) (2020). https://doi.org/10.3390/ma13112561
46. G. Chen, G. Chen, Y. Lou, Diagonal recurrent neural network-based hysteresis modeling. IEEE Trans. Neural Netw. Learn. Syst. 1–11 (2021). https://doi.org/10.1109/TNNLS.2021.3085321
47. G. Chen, Y. Lou, Recurrent-neural-network-based rate-dependent hysteresis modeling and feedforward torque control of the magnetorheological clutch. IEEE/ASME Trans. Mechatron. 1–12 (2021). https://doi.org/10.1109/TMECH.2021.3121498
48. M.P. Soares Barbosa, M. Rakotondrabe, H.V. Hultmann Ayala, Deep learning applied to data-driven dynamic characterization of hysteretic piezoelectric micromanipulators. IFAC-PapersOnLine **53**(2), 8559–8564 (2020). https://doi.org/10.1016/j.ifacol.2020.12.566. 21st IFAC World Congress
49. M. Amodeo, P. Arpaia, M. Buzio, V. Di Capua, F. Donnarumma, Hysteresis modeling in iron-dominated magnets based on a multi-layered NARX neural network approach. International Journal of Neural Systems **31**(09), 2150033 (2021)
50. A. Bernardini, P. Maffezzoni, A. Sarti, Vector wave digital filters and their application to circuits with two-port elements. IEEE Trans. Circ. Syst. I Regular Pap. **68**(3), 1269–1282 (2021). https://doi.org/10.1109/TCSI.2020.3044002
51. A. Bernardini, K.J. Werner, J.O. Smith, A. Sarti, Generalized wave digital filter realizations of arbitrary reciprocal connection networks. IEEE Trans. Circ. Syst. I Regular Pap. **66**(2), 694–707 (2019). https://doi.org/10.1109/TCSI.2018.2867508
52. G. Martens, H. Le, Wave digital adapters for reciprocal second-order sections. IEEE Trans. Circ. Syst. **25**(12), 1077–1083 (1978). https://doi.org/10.1109/TCS.1978.1084433
53. L. Chua, C. Desoer, E. Kuh, *Linear and nonlinear circuits* (McGraw-Hill, New York, 1987)
54. S. Seshu, M. Reed, *Linear graphs and electrical networks* (Addison Wesley Publishing Company, Boston, 1961)
55. E. Laithwaite, Magnetic equivalent circuits for electrical machines. Proc. Inst. Electr. Eng. **114**(11), 1805–1809 (1967). https://doi.org/10.1049/piee.1967.0344
56. D. Atherton, J. Beattie, A mean field Stoner-Wohlfarth hysteresis model. IEEE Trans. Magn. **26**(6), 3059–3063 (1990). https://doi.org/10.1109/20.102891
57. S.E. Zirka, Y.I. Moroz, R.G. Harrison, K. Chwastek, On physical aspects of the Jiles-Atherton hysteresis models. J. Appl. Phys. **112**(4), 043916 (2012). https://doi.org/10.1063/1.4747915
58. J.W. Macki, P. Nistri, P. Zecca, Mathematical models for hysteresis. SIAM Rev. **35**(1), 94–123 (1993)
59. A. Visintin, in *Modelling and optimization of distributed parameter systems applications to engineering*. Mathematical models of hysteresis (Springer, Boston, 1996), p.71–80
60. I.D. Mayergoyz, *Mathematical models of hysteresis and their applications* (Elsevier Science, New York, 2003). https://doi.org/10.1016/B978-0-12-480873-7.X5000-2
61. M.A. Krasnosel'skii, A.V. Pokrovskii, *Systems with hysteresis* (Springer-Verlag, Berlin Heidelberg, 2012)
62. M. Brokate, J. Sprekels, in *Applied Mathematical Sciences*, vol. 121, Hysteresis and phase transitions(Springer, New York, 1996)
63. F. Preisach, Über die magnetische nachwirkung. Z. Phys. **94**(5–6), 277–302 (1935)
64. I.D. Mayergoyz, Dynamic Preisach models of hysteresis. IEEE Trans. Magn. **24**(6), 2925–2927 (1988). https://doi.org/10.1109/20.92290
65. R. Mrad, H. Hu, Dynamic modeling of hysteresis in piezoceramics, vol. 1, in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. (IEEE, Como, 2001), pp.510–515. https://doi.org/10.1109/AIM.2001.936515
66. D.E. Rumelhart, J.L. McClelland, in *Parallel distributed processing: explorations in the microstructure of cognition: foundations*. Learning internal representations by error propagation (MIT Press, Cambridge, 1987), p. 318-362
67. J.L. Elman, Finding structure in time. Cogn. Sci. **14**(2), 179–211 (1990)
68. I.D. Mayergoyz, Mathematical models of hysteresis. IEEE Trans. Magn. **22**(5–6), 603–608 (1986)
69. M. Brokate, Some mathematical properties of the Preisach model for hysteresis. IEEE Trans. Magn. **25**(4), 2922–2924 (1989). https://doi.org/10.1109/20.34325
70. Q. Yanding, Z. Xin, Z. Lu, Modeling and identification of the rate-dependent hysteresis of piezoelectric actuator using a modified Prandtl-Ishlinskii model. Micromachines **8**(4) (2017). https://doi.org/10.3390/mi8040114
71. G. Gruosso, A. Brambilla, Magnetic core model for circuit simulations including losses and hysteresis. Int. J. Numer. Model. Electron. Netw. Devices Fields **21**(5), 309–334 (2008)
72. Y. Bengio, P. Simard, P. Frasconi, Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Netw. **5**(2), 157–166 (1994). https://doi.org/10.1109/72.279181
73. N. Mehboob, Hysteresis properties of soft magnetic materials. Ph.D. thesis, Universität Wien (2012)
74. N. Mehboob. Private communication (2021)
75. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., Pytorch: an imperative style, high-performance deep learning library, in *Proc. Adv. Neural Inf. Process. Syst.*, Vancouver 2019. vol. 32, (Curran Associates Inc., Red Hook, 2019), pp.8026-8037
76. D.P. Kingma, J. Ba, Adam: A method for stochastic optimization. (2014). arXiv preprint arXiv:1412.6980
77. G. GmbH. https://www.grau-stanzwerk.de/d-wAssets/docs/produkte/ui-kernbleche/Zg_Abmessung.pdf. UI-laminations, according DIN EN 60740-1, Accessed 2023-01-30

Publisher's Note