



PAPER • OPEN ACCESS

Universal embedding of autonomous dynamical systems into a Lotka-Volterra-like format

To cite this article: Diego Frezzato 2024 *Phys. Scr.* **99** 015235

View the [article online](#) for updates and enhancements.

You may also like

- [Decoding algorithm as a moment problem related to the extended Lotka–Volterra system](#)

Yan Pan, Xiang-Ke Chang and Xing-Biao Hu

- [Passivity of Lotka–Volterra and quasi-polynomial systems](#)

Lrinc Márton, Katalin M Hangos and Attila Magyar

- [Dynamically evolved community size and stability of random Lotka–Volterra ecosystems](#)

Tobias Galla



PAPER

Universal embedding of autonomous dynamical systems into a Lotka-Volterra-like format

OPEN ACCESS

RECEIVED
10 October 2023REVISED
27 November 2023ACCEPTED FOR PUBLICATION
4 December 2023PUBLISHED
21 December 2023

Diego Frezzato

Department of Chemical Sciences, University of Padova, via Marzolo 1, I-35131, Padova, Italy

E-mail: diego.frezzato@unipd.it**Keywords:** ordinary differential equations, autonomous dynamical systems, embedding into Lotka-Volterra, universal approximation

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

**Abstract**

We show that the ordinary differential equations (ODEs) of any deterministic autonomous dynamical system with continuous and bounded rate-field components can be embedded into a quadratic Lotka-Volterra-like form by turning to an augmented set of state variables. The key step consists in expressing the rate equations by employing the Universal Approximation procedure (borrowed from the machine learning context) with logistic sigmoid ‘activation function’. Then, by applying already established methods, the resulting ODEs are first converted into a multivariate polynomial form (also known as generalized Lotka-Volterra), and finally into the quadratic structure. Although the final system of ODEs has a dimension virtually infinite, the feasibility of such a universal embedding opens to speculations and calls for an interpretation at the physical level.

1. Introduction

In 1986, Peschel and Mende published a monograph [1] entitled ‘The predator–prey model’ whose subtitle was the intriguing question: *Do we live in a Volterra World?* Although the focus was on system’s growth and structure-building from the viewpoint of ecology, the findings were general and that work fitted in a series of studies, past and subsequent, concerning the embedding of the evolution laws of several types of deterministic dynamical systems into quadratic Lotka-Volterra-like ordinary differential equations (ODEs). Let us introduce the basic notation to enter the topic. Let \mathbf{x} be the finite set of physical variables of interest that evolve according to

$$\dot{\mathbf{x}} = \mathbf{y}(\mathbf{x}) \quad (1)$$

where $\mathbf{y}(\mathbf{x})$ is a time-independent rate-field vector. In our discussion, it will be assumed that the components of the rate field are bounded and continuous functions of the state variables. The embedding into a quadratic format is generally achieved by turning from \mathbf{x} to an enlarged set of new dynamical variables. In doing this, the dimension of the system’s representation increases, but the big advantage is that the original nonlinearity can be brought down to second order. The feasibility of such a kind of recasting, and especially its potential applicability to a variety of dynamical systems, justifies the above question and still triggers new investigations.

The history of the ODEs ‘quadratization’ has developed in a very branched tree of interconnected studies. Besides the work of Peschel and Mende cited above, some other important contributions have to be mentioned. First, it is worth mentioning Kerner’s work [2] on the embedding of ODEs into ‘elemental Riccati equations’. A milestone is surely the work of Brenig and Goriely [3] who showed how the Lotka-Volterra (LV) format can be achieved through intermediate quasi-polynomial forms [4], also termed generalized Lotka-Volterra (GLV) formats, into which the original ODEs can be exactly or approximately converted. The same kind of transformation was rediscovered by several authors with little variations. For instance, in the (deterministic) chemical kinetics context, in which the mass-action rate equations already have a GLV format, we mention the works of Gouzé [5], of Fairén and Hernández-Bermejo [6], and of the present author and co-workers for both closed [7] and open [8] chemical reaction networks. Quasi-polynomial (GLV) and LV formats have been widely studied in terms of stability of the stationary points [9–16], boundedness of the solutions [5, 12], stabilizing feedback control in process systems [17–19], integrability of ODEs with polynomial nonlinearities [20], connection with stochastic urn processes [21], and connection with abstract Lie algebra [22]. This brief overview

is by no means complete and should only give the idea of the broad interest in such topic across decades of research.

Beyond the mere mathematical aesthetics, the crucial point is that the real dynamics of a variety of systems can be mirrored by a *unique* kind of dynamics in the extended space of the new variables. If some interesting feature emerges at such an abstract level of representation, it might be interesting to inspect how that feature manifests in the physical space. For instance, under this perspective the analysis of the quadratic format enabled to let emerge a definition of the slow manifolds observed in mass-action chemical kinetics [23] and to devise computational routes for their identification [24, 25], to discover the existence of mutually orthogonal attracting subspaces in what was termed the ‘hyper-spherical representation’ of the dynamics [24, 26], and even the existence of an intrinsic timing for Markov jump processes represented in the probability space [27].

Let us now circumscribe the context. The specific universal quadratic format we will deal with is of the following type:

$$\dot{h}_Q = -h_Q \sum_{Q'} M_{QQ'} h_{Q'} \quad (2)$$

where the h_Q are the new dynamical variables enumerated by the index Q , the overdot stands for time-derivative, and the elements $M_{QQ'}$ are dimensionless constant factors. It is assumed that a suitable transformation from the original physical variables \mathbf{x} to the new set \mathbf{h} does exist and can be found to convert the original ODEs equation (1) into the quadratic format of equation (2). The new variables are strictly positive and have physical dimension of inverse-of-time. The dimension of the extended space is supposed to be finite although possibly indefinitely large. The crucial point to be stressed is that the $h_Q(\mathbf{x})$ are interrelated one each other by mutual constrains in the way that the independent number of degrees of freedom remains unaltered. This means that the initial conditions are not arbitrary, i.e., they have to be fixed by setting $h_Q(0) \equiv h_Q(\mathbf{x}(0))$. From such initial condition, the subsequent evolution in the extended space is autonomous and ruled by equation (2). It is finally supposed that a backward transformation allows to retrieve the physical state $\mathbf{x}(t)$, at any future time, from the ensemble of the $h_Q(t)$. As a side note, by turning to the variables $V_{QQ'} := M_{QQ'} h_{Q'}$ we can directly convert equation (2) into an evolution equation devoid of any system-depended parameter: $\dot{V}_{QQ'} = -V_{QQ'} \sum_{Q''} V_{Q'Q''}$. All system’s features bear on the mutual interrelations between these new variables. The $V_{QQ'}$ prove to own interesting properties and constitute the basis on which to build the hyper-spherical representation mentioned above.¹

Let us note that equation (2) are rate equations of pure second order. Just by analogy, but without any real physical correspondence, these equations can be collectively seen as a Lotka-Volterra-like structure devoid of linear terms. The structure equation (2) was obtained for mass-action chemical kinetics [4, 5, 7], for some classes of phase-space dynamics under the extension to complex-valued h_Q [26] and, recently, even for classical master equation dynamics in which the embedding into equation (2) implies the increase from linear to quadratic ODEs [27]. In all these situations, the change from \mathbf{x} to \mathbf{h} was devised ad hoc on the basis of the specific type of original ODEs. This allowed to confer, in each case, a specific form to the new variables and, even more importantly, to delimit the dimension of the extended space.

Given the relevance of the quadratic structure equation (2) as an intermediate step to derive other formats of the evolution law (like, for instance, the evolution in the hyper-spherical representation mentioned above), what we ask here is if it exists a *universal* quadratization route, i.e., a general procedure applicable regardless of the specific features of the given dynamical system. This would allow us to establish a direct connection between (potentially) any autonomous dynamical system and the evolution law equation (2) in the corresponding extended space. We will see that such a universal quadratization procedure does indeed exist regardless of the smoothness of the rate-field components $y_i(\mathbf{x})$, on condition that they are continuous and bounded functions in the physical domain of the \mathbf{x} variables.

To devise such universal quadratization route, we will resort to the ‘Universal Approximation’ (UA) method [28, 29] applied to the components of $\mathbf{y}(\mathbf{x})$. The UA is a well-known approach to approximate virtually any bounded and generally multivariate function. It is worth mentioning that the UA provides the theoretical basis for the single-hidden-layer feedforward strategies in regression and classification machine learning tasks [28–31] and supports the efficacy of the multi-hidden-layer deep learning architectures [32]. In section 2 we show that the UA applied to the rate equations allows us to devise a universal recasting of the original ODEs. Then, as

¹ The $V_{QQ'}$ and the $\dot{V}_{QQ'}$ proved useful to construct suitable scalar fields to guide the localization of possible slow manifolds in the physical space [24, 25]. Furthermore, the empirical inspection of the high-order time derivatives of the $V_{QQ'}(t)$ has led to conjecture the existence of an intrinsic ‘timing function’ as long as the trajectory $\mathbf{x}(t)$ is inside an ‘attractiveness region’ of the physical space [24, 27]. Concerning the hyper-spherical representation, let n_{dim} be the total number of h_Q variables. The $V_{QQ'}$ collectively specify a point in the n_{dim}^2 -dimensional space whose orthogonal directions are labelled by the index $J \leftrightarrow (Q, Q')$. Each direction is orthogonal to an associate $(n_{\text{dim}}^2 - 1)$ -dimensional hyper-plane, and all such hyper-planes are orthogonal one each other. In the hyper-spherical representation [24], the representative variables become a ‘radial’ coordinate and a set of ‘angular’ coordinates specifying a point on the $(n_{\text{dim}}^2 - 1)$ -dimensional unit sphere. It was shown that the evolution is sectioned into parts during which the point on the spherical surface is attracted by one specific hyper-plane at a time.

shown in section 3, by employing the logistic sigmoid as ‘activation function’ in the UA we can achieve a generalized Lotka-Volterra format from which equation (2) is eventually obtained. The technical step from GLV to LV is detailed in appendix. Section 5 is devoted to conclusions and perspectives.

To the best of the author’s knowledge, the Lotka-Volterra embedding method here proposed is a novelty. A posteriori, however, the author notes several points of closeness with a remarkable work by Moreau, Brenig and co-workers [33]. In that work the authors showed that several neural networks architectures, if used as generators of continuous-time self-evolution, can be embedded into Lotka-Volterra formats. They treated explicitly the ‘dynamical perceptron’ (in its basic form and with self-connections) and the ‘dynamical multi-layer perceptron’. Since the neural network architecture is potentially a universal approximator of functions, and the specific application to the trajectories of autonomous dynamical systems was in fact developed by Funahashi and Nakamura in a previous work [34], in their concluding remarks the authors claimed that Lotka-Volterra formats can be used to approximate any finite-dimensional dynamical system for any finite time. This is what we elaborate here in explicit way. While in [33] the perceptron units are by themselves the velocities of change of the dynamical variables, in our work they collectively enter into the expansion of the components of $\mathbf{y}(\mathbf{x})$ (the actual velocities according to equation (1)) through the Universal Approximation strategy (see the key equation equation (3) given later). Analogies and differences between this work and [33] will be pointed out where appropriate.

2. Recasting of the evolution law via Universal Approximation

Let us consider the system of ODEs in equation (1), where \mathbf{x} is a finite N -dimensional set of dynamical variables and $\mathbf{y}(\mathbf{x})$ is the associate N -dimensional vector field of rate equations. We assume that the components $y_i(\mathbf{x})$ are continuous (systems confined by reflecting boundaries are not considered here) and bounded in the physical domain where the dynamics are inspected.

Let us adopt the Universal Approximation [28, 29] to express each of the rate-field components as follows:

$$y_i(\mathbf{x}) = \sum_{k=1}^{\tilde{N}} c_i^{(k)} \sigma \left(b_i^{(k)} + \sum_j w_{ij}^{(k)} x_j \right) \quad (3)$$

where \tilde{N} is the number of terms, $\sigma(\cdot)$ is a function $\mathfrak{R} \rightarrow \mathfrak{R}$, and $\{w_{ij}^{(k)}\}$, $\{c_i^{(k)}\}$ and $\{b_i^{(k)}\}$ are adjustable parameters. A sufficient condition for $\sigma(z)$ to be a valid function in the UA, i.e., capable of producing a dense set in the space of continuous functions, is of being non-constant, bounded² and with different limits at $z \rightarrow +\infty$ and $z \rightarrow -\infty$; continuity is not generally required although it is demanded in our specific application. In the machine learning context, the UA finds its embodiment in the basic single-hidden-layer feedforward network with \tilde{N} ‘nodes’ (or ‘neurons’), $\sigma(\cdot)$ as ‘activation function’, $\{w_{ij}^{(k)}\}$ and $\{c_i^{(k)}\}$ as ‘weights’, and $\{b_i^{(k)}\}$ as ‘additive biases’ [32]. For this reason, in what follows we will freely use such terminology although, of course, there is no evident correspondence between the two contexts.

For the sake of notation, let us write

$$\sigma_i^{(k)}(\mathbf{x}) \equiv \sigma \left(b_i^{(k)} + \sum_j w_{ij}^{(k)} x_j \right) \quad (4)$$

Let us now assume that the activation function is such that

$$\frac{d\sigma(z)}{dz} = f(\sigma(z)) \quad (5)$$

with $f(\cdot)$ some function. Having in mind the achievement of quadratic ODEs as final goal, it is important that $f(\cdot)$ be a polynomial of finite order. This was a key point also in [33] where the first target was to devise a Lotka-Volterra embedding of the continuous-time ‘dynamical perceptron’.³ The problem that we are facing here is different since, in our case, the ‘perceptron’ units (i.e., the $\sigma_i^{(k)}(\mathbf{x})$ components) collectively enter into the expansion equation (3) of the rate-field components. By employing equation (5) we have that $\dot{\sigma}_i^{(k)} = f(\sigma_i^{(k)}) \sum_j w_{ij}^{(k)} \dot{x}_j$ and hence, by using the expansion equation (3) for $\dot{x}_j = y_j(\mathbf{x})$, we obtain

² The boundedness condition is not strictly necessary [29, 35] but the admissibility of unbounded activation functions have to be assessed case by case. For instance, functions like the ‘rectified linear unit’ (ReLU), which is unbounded on one side, are normally employed in the deep learning architecture and prove to be highly efficient at the computational level.

³ The ‘dynamical perceptron’ treated in [33] corresponds to the system of differential equations $\dot{\mathbf{x}} = \sigma(\mathbf{A}\mathbf{x} + \mathbf{a}_0)$, where \mathbf{x} is the n -dimensional column vector of dynamical variables, \mathbf{a}_0 is a column vector of same dimension, \mathbf{A} is a $n \times n$ matrix, and $\sigma(\cdot)$ is the column vector whose entries are $\sigma_i(\mathbf{a}) \equiv \sigma(a_i)$ with $\sigma(\cdot)$ the activation function. We see that the dynamical perceptron of [33] and the evolution law equation (1) are different dynamical systems. While the units of the dynamical perceptron are by themselves the rates of evolution of the dynamical variables, in our case such units are collectively employed to approximate the rate-field components of the actual dynamical system through the UA equation (3).

$$\dot{\sigma}_i^{(k)} = f(\sigma_i^{(k)}) \sum_{j,k'} w_{ij}^{(k)} c_j^{(k')} \sigma_j^{(k')} \quad (6)$$

Overall, once the form of $\sigma(\cdot)$ is chosen, weights and biases are in principle fixed. Thus, equation (6) is a new autonomous set of ODEs whose structure and parameters depend on the choice of the activation function.

Let us note that each rate-field component $y_i(\mathbf{x})$ has its own independent expansion. The physical information needed to determine the associate biases and weights is constituted by the values of $y_i(\mathbf{x})$ at a number N_p of sampling points. We can think of a discretization of the physical space into finite-size cells, taking the centres of the cells as sampling points. By following the reasoning of [31], we can say that if the activation function is infinitely differentiable, then the required number of nodes is $\tilde{N} \leq N_p$. Such number can be however indefinitely large and it increases as the required accuracy on the rate-field components, and of course the required closeness between real and approximate trajectories (those generated by the approximate rate field) in a given time-window, are more and more stringent. An example of rate-field approximation is provided later in section 4. The important point is that we are interested only in the *existence* of weights and biases which allow to get an approximation of the rate field with preset accuracy; the practical way to determine such parameters, and even the criteria to discern between different alternative sets that yield comparable quality of the approximation, are not relevant here.

Let us introduce the further cumulative index

$$\alpha \leftrightarrow (i, k) \quad (7)$$

in which i labels the physical variables and k the ‘nodes’ of the UA structure. With this notation, equation (6) takes on

$$\dot{\sigma}_\alpha = f(\sigma_\alpha) \sum_{\alpha'} A_{\alpha\alpha'} \sigma_{\alpha'} \quad (8)$$

with

$$A_{\alpha \leftrightarrow (i,k), \alpha' \leftrightarrow (j,k')} = w_{ij}^{(k)} c_j^{(k')} \quad (9)$$

3. Quadratization

So far, the form of $\sigma(\cdot)$ was generic. Let us now choose the logistic sigmoid function

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (10)$$

This function is strictly positive-valued, bounded between 0 (for $z \rightarrow -\infty$) and 1 (for $z \rightarrow +\infty$), and monotonically increasing as z increases. With this choice, the derivative in equation (5) is simply the following second-order polynomial:

$$f(\sigma) = \sigma(1 - \sigma) \quad (11)$$

The use of this form in equation (8) yields a system of cubic ODEs:

$$\dot{\sigma}_\alpha = -\sigma_\alpha^2 \sum_{\alpha'} A_{\alpha\alpha'} \sigma_{\alpha'} + \sigma_\alpha \sum_{\alpha'} A_{\alpha\alpha'} \sigma_{\alpha'} \quad (12)$$

The ODEs of equation (12) belong to the GLV class and can be converted into a Lotka-Volterra-like quadratic format by employing the strategy proposed by Brenig and Goriely [3] which, as mentioned in the introduction, was later re-discovered by others.

Let us introduce the following quantities built by using the variables σ_α :

$$\Pi_{\gamma\alpha\alpha'}(\boldsymbol{\sigma}) := \prod_{\alpha''} \sigma_{\alpha''}^{e_{\alpha''}(\gamma\alpha\alpha')} \quad (13)$$

with exponents

$$e_{\alpha''}(\gamma\alpha\alpha') = \delta_{\gamma,1}(\delta_{\alpha,\alpha''} + \delta_{\alpha',\alpha''}) + \delta_{\gamma,2}(\delta_{\alpha,\alpha''} + 2\delta_{\alpha',\alpha''}) \quad (14)$$

where γ is a binary index; we assign value 1 or 2 to such index. Here and below, δ stands for the Kronecker’s delta. With these positions, the ODEs in equation (12) can be rewritten in compact form as

$$\dot{\sigma}_\alpha = \sum_{\gamma, \alpha', \alpha''} K_{\gamma\alpha'\alpha'',\alpha} \Pi_{\gamma\alpha'\alpha''}(\boldsymbol{\sigma}) \quad (15)$$

where

$$K_{\gamma\alpha'\alpha'',\alpha} = (\delta_{\gamma,1} - \delta_{\gamma,2}) A_{\alpha\alpha'} \delta_{\alpha,\alpha''} \quad (16)$$

It can be checked by direct substitution that equation (15) is equivalent to equation (12). Let us now consider the new variables

$$h_{\gamma\alpha\alpha'\alpha''}(\boldsymbol{\sigma}) := \prod_q \sigma_q^{e_q(\gamma\alpha'\alpha'') - \delta_{q,\alpha}} \quad (17)$$

which are positive-valued since the logistic sigmoid is strictly positive for any finite value of the argument. As we see, these variables are labelled by four indexes, namely, the binary label γ and the three indexes α , α' and α'' which run over the numbering of the pairs variable-node. By employing the route described in [appendix](#), it is found that the evolution of such variables is ruled by

$$\dot{h}_{\gamma\alpha\alpha'\alpha''} = -h_{\gamma\alpha\alpha'\alpha''} \sum_{\tilde{\gamma}, \tilde{\alpha}, \tilde{\alpha}', \tilde{\alpha}''} M_{\gamma\alpha\alpha'\alpha'', \tilde{\gamma}\tilde{\alpha}\tilde{\alpha}'\tilde{\alpha}''} h_{\tilde{\gamma}\tilde{\alpha}\tilde{\alpha}'\tilde{\alpha}''} \quad (18)$$

in which

$$M_{\gamma\alpha\alpha'\alpha'', \tilde{\gamma}\tilde{\alpha}\tilde{\alpha}'\tilde{\alpha}''} = [\delta_{\alpha, \tilde{\alpha}} - e_{\tilde{\alpha}}(\gamma\alpha'\alpha'')] (\delta_{\tilde{\gamma}, 1} - \delta_{\tilde{\gamma}, 2}) A_{\tilde{\alpha}\tilde{\alpha}'} \delta_{\tilde{\alpha}, \tilde{\alpha}''} \quad (19)$$

Finally, the quadratic form equation (2) is obtained by introducing the following cumulative index Q for the labelling of the quadruplets:

$$Q \leftrightarrow (\gamma, \alpha, \alpha', \alpha'') \quad (20)$$

Totally, the number of h_Q variables is equal to

$$n_{\text{dim}} = 2 (N\tilde{N})^3 \quad (21)$$

The whole quadratization procedure is summarized in [figure 1](#). The universal route described above allows us to obtain a quadratic format of evolution law at the price of working in a very extended space of variables. As anticipated in the Introduction, the $h_Q(\mathbf{x})$ variables are mutually interrelated, as emerges by retracing the global transformation $\mathbf{x} \rightarrow \mathbf{h}(\mathbf{x})$ specified by equations (4), (13) and (17). This implies that the components of the initial array $\mathbf{h}(0)$ cannot be set arbitrarily, but have to be assigned according to the given initial state $\mathbf{x}(0)$ in the physical space. The backward transformation $\mathbf{h}(t) \rightarrow \mathbf{x}(t)$ (see [appendix](#) and remark (iii) below) then allows us to retrieve the physical state at a generic future time t .

3.1. Remarks

(i) *The truncation issue.* While the quadratization procedure from equation (8) to equation (18) is exact at the algebraic level once equation (10) is employed, an approximation is inevitably introduced when equation (3) is adopted with a finite set of addends in the summation.

(ii) *Dimension of the extended space in relation to accuracy.* The dimension of the extended space is determined by the number of nodes in the Universal Approximation equation (3). Such number is determined by the required accuracy, i.e., the largest accepted deviation between approximate trajectories (the ones generated by the approximate rate field) and true trajectories, in a given time-window, starting from the points in a given domain of the physical space. As the domain is taken wider and/or the rate field is more and more featured and/or the time-window is taken longer, the dimension of the extended space increases and can become very rapidly virtually infinite. Let us make an example. Let us suppose dealing with a dynamical system having $N = 10$ physical variables $x_1(t)$ to $x_{10}(t)$, and that the form of the rate equations $y_i(\mathbf{x})$ allows to get a good UA representation with, quite optimistically, only $\tilde{N} = 10$ nodes. The number of extended variables $h_{\gamma\alpha\alpha'\alpha''}(t)$ would be 2×10^6 , which goes against any advantage at the computational level. On the other hand, what we are seeking here is *not* an efficient computational route. In fact, the original ODEs are reasonably easy to solve numerically, whereas the quadratization would bring an unacceptable (and apparently unreasonable) redundant augmentation of the space. Rather, what we wanted to show is that, in principle, *any* dynamical system with continuous and bounded rate-field components is potentially embeddable in the format of equation (18).

(iii) *Invertibility.* As discussed in [appendix](#), the backward transformation from the set of variables $h_{\gamma\alpha\alpha'\alpha''}(\boldsymbol{\sigma})$ to the σ_α is feasible. The monotonicity of the sigmoid function $\sigma(\cdot)$ then allows to get the value of the argument $b_i^{(k)} + \sum_j w_{ij}^{(k)} x_j(t)$ from the value of $\sigma_i^{(k)}(t)$ (see equation (4)). Finally, the physical state $\mathbf{x}(t)$ can be retrieved if at least one of the $N \times N$ matrices $\mathbf{w}^{(k)}$ (or a linear combination of them) is invertible.

(iv) *Choice of the 'activation function'.* The choice of the function $\sigma(\cdot)$ is subjective. On condition that the derivative function $f(\sigma)$ be a polynomial in σ , the quadratization procedure is always feasible. However, strictly positive-valued functions are *a priori* preferred to avoid possible divergences of the $h_{\gamma\alpha\alpha'\alpha''}(\boldsymbol{\sigma})$. In addition, it proves convenient to opt for the logistic sigmoid because the polynomial is limited to second order (see equation (11)). Also the hyperbolic tangent function (which, by the way, was the activation function adopted in [33]) yields a second-order polynomial, but its employment in our specific quadratization route could cause problems since it can vanish. Other choices would generate an extended space of larger dimension. Thus, the logistic sigmoid appears to be the best choice.

(v) *Alternative embedding schemes.* It is important to recall that other embedding schemes have been proposed for general nonlinear ODEs. For instance, Kerner showed [2] that by introducing additional variables one can convert the original ODEs into a polynomial format, which is a starting point to achieve quadratic

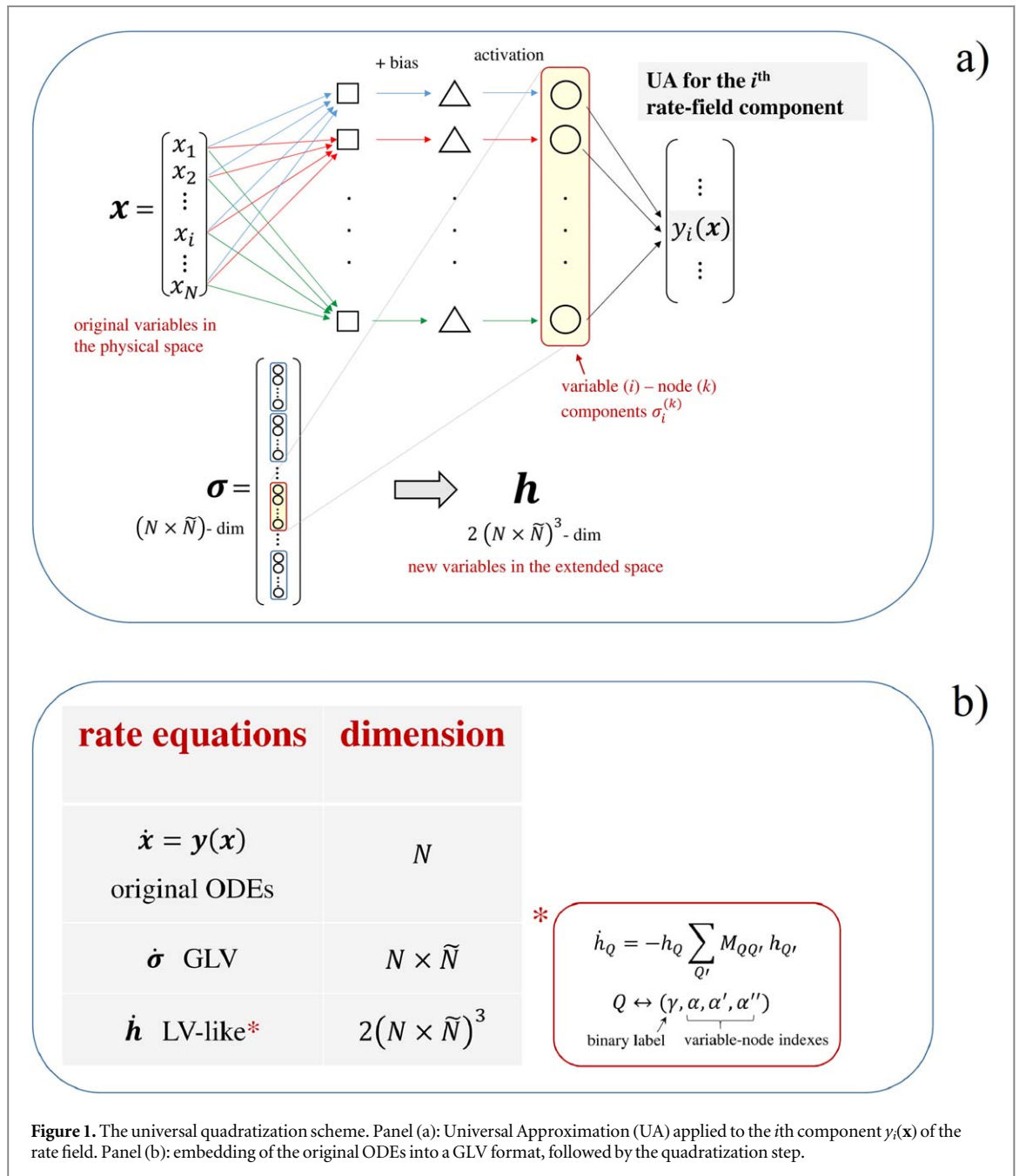
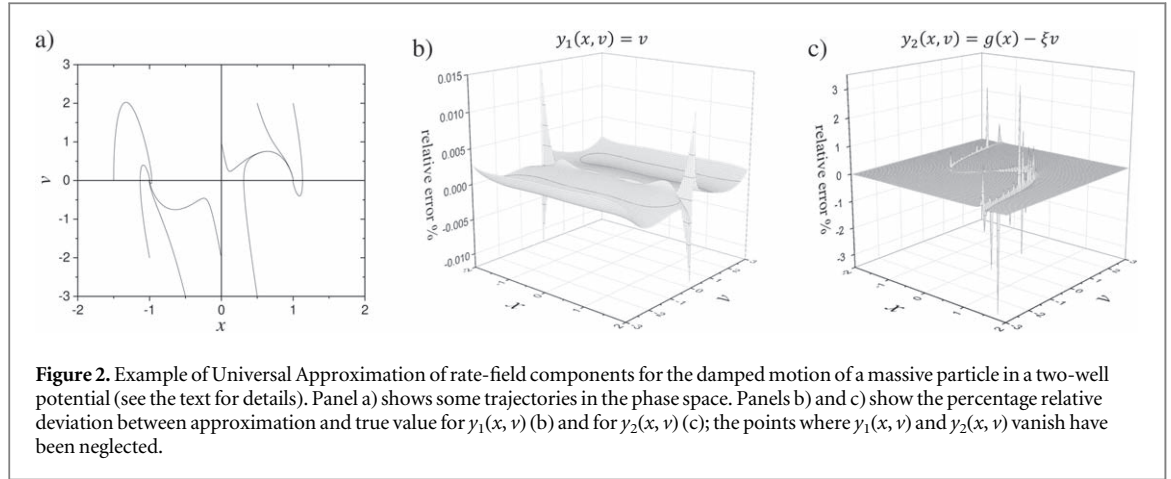


Figure 1. The universal quadratization scheme. Panel (a): Universal Approximation (UA) applied to the i th component $y_i(\mathbf{x})$ of the rate field. Panel (b): embedding of the original ODEs into a GLV format, followed by the quadratization step.

forms. A notable approach was proposed by Hernández-Bermejo and Fairén [4] who showed that the original ODEs can be converted into a GLV format (from which LV forms can then be obtained) of the kind $\dot{u}_i = u_i[\lambda_i + \sum_{j=1}^{M^*} A_{ij} \prod_{k=1}^{N^*} u_k^{B_{jk}}]$ where the dynamical variables u_i ($i = 1, \dots, N^*$) comprise the N physical ones (possibly under appropriate shifts to ensure positivity of all the variables) plus additional variables unless the original ODEs have already a GLV format; the $\{\lambda_i\}$, $\{A_{ij}\}$, and $\{B_{jk}\}$ are real-valued parameters. These procedures are however cumbersome and require an ad hoc algebraic elaboration for each given set of ODEs. Moreover, while the conversion into polynomial or GLV formats of finite (and even low) dimension is exact in many cases, a truncation issue bears also on these procedures since an exact closure on the additional variables is generally lacking. The entering coefficients therefore become adjustable parameters and should be optimized, e.g. by means of some kind of fitting procedure inside a given domain of the physical variables. Because of the power-law dependence on the u_i , the quality of the approximation could rapidly degrade when applied outside such domain. The embedding procedure presented here is instead a sort of system-independent ‘black box’ since the UA in equation (3) is a general numerical scheme of approximation which does not require to inspect the specific form of the rate equations; this is the main point of strength. Furthermore, the degradation of the approximation could be mitigated by the fact that the $y_i(\mathbf{x})$ are approximated by the superposition of bounded



functions. Of course, a detailed comparison between the various alternative embedding schemes requires a case-by-case analysis.

4. Example of rate-field approximation

As a simple example, let us consider the dynamics of a damped particle in one dimension. The dynamical variables are $\mathbf{x} \equiv (x, v)$, where x is the position and v the velocity. The rate equations are $\dot{x} = y_1(x, v) = v$ and $\dot{v} = y_2(x, v) = g(x) - \xi v$ where $g(x) = -dU(x)/dx$ being $U(x)$ a potential energy divided by the mass of the particle, and ξ a damping friction coefficient per unit of mass. Let us consider the case of the symmetric two-well potential already studied in [26]. The potential energy is $U(x) = \Delta[(x/c)^2 - 1]^2$ and two stationary points $\mathbf{x} = (\pm c, 0)$ are present.

The computations were done for $\Delta = 5$, $c = 1$ and $\xi = 10$, the same values employed in [26]. Panel (a) of figure 2 shows some trajectories in the phase-space, while panels (b) and (c) show the percentage relative error that bears on the rate-field components approximated by equation (3) with the parameters obtained from the specific route adopted here.⁴ The input information consisted in the values of $y_1(x, v)$ and $y_2(x, v)$ at the centres of the cells of a regular 5×5 grid. The considered portion of phase space was $-2 \leq x \leq 2$ and $-3 \leq v \leq 3$. For each rate-field component, the number \tilde{N} of nodes was therefore equal to 25. We can see that the largest error, which bears on the $y_2(\mathbf{x})$ component, is confined within about 3%. Assuming the achieved accuracy is good enough to describe the trajectories in a certain time-window, according to equation (21) the dimension n_{dim} of the extended space of the variables h_Q would be 2.5×10^5 , which is a huge number considering that the natural variables are only two.

5. Conclusions: do we really live in a Volterra world?

In this work we have shown that virtually any dynamical system of autonomous ODEs with bounded and continuous evolution rates can be converted into the quadratic Lotka-Volterra-like format equation (18). To achieve such recasting we have applied the Universal Approximation (UA) method to the original multivariate rate equations. By specifically employing the logistic sigmoid activation function in the UA procedure, we could achieve the GLV format equation (12), and finally get the quadratic format by means of the quadratization route illustrated in appendix. The notable fact is that this procedure is really universal and system-independent. The new variables, all having physical dimension of inverse-of-time, are the $h_{\gamma\alpha\alpha'\alpha''}$ defined in equation (17). The global transformation is invertible, which means that the physical state of the system can be retrieved from the

⁴ Biases and weights have been determined by resorting to a basic implementation of the ‘extreme machine learning’ approach [31]. Let us adopt the index $c = 1, 2, \dots, \tilde{N}$ to enumerate the cells into which the portion of phase space is subdivided. Note that the numbers of nodes and cells are equal. Let $\mathbf{x}(c)$ be the central points of the cells. For each component $i = 1, 2$, equation (3) can be rewritten in matrix form as $\mathbf{H}^{(i)} \mathbf{g}^{(i)} = \mathbf{z}^{(i)}$ where $\mathbf{H}^{(i)}$ is the $\tilde{N} \times \tilde{N}$ matrix with entries $H_{ck}^{(i)} \equiv \sigma(b_i^{(k)} + \sum_{j=1,2} w_{ij}^{(k)} x_j(c))$, $\mathbf{g}^{(i)}$ is the $\tilde{N} \times 1$ column vector with elements $g_k^{(i)} \equiv c_i^{(k)}$, and $\mathbf{z}^{(i)}$ is the $\tilde{N} \times 1$ column vector with elements $z_c^{(i)} \equiv y_i(\mathbf{x}(c))$. Once the biases $\{b_i^{(k)}\}$ and the weights $\{w_{ij}^{(k)}\}$ are subjectively assigned, the matrix $\mathbf{H}^{(i)}$ is set. The coefficients $\{c_i^{(k)}\}$ are then determined by means of $c_i^{(k)} = \sum_c [\mathbf{H}^{(i)-1}]_{kc} y_i(\mathbf{x}(c))$ under the assumption that the matrix is invertible. In the present computations, the $\{b_i^{(k)}\}$ and the $\{w_{ij}^{(k)}\}$ were randomly generated with uniform distribution between -0.5 and $+0.5$. The same sets were employed for both $i = 1$ and $i = 2$. With such sets, the $\{c_1^{(k)}\}$ and $\{c_2^{(k)}\}$ were determined as described above. It was observed that the largest error bears on the component $y_2(\mathbf{x})$. The relative error $e_2(\mathbf{x}) = (y_2^{\text{appr}}(\mathbf{x}) - y_2(\mathbf{x})) / y_2(\mathbf{x})$ was evaluated over the regular grid 100×100 and its maximum value in module, $e_{2,\text{max}}$, was determined. The calculations were repeated 10^5 times and the set of parameters that yielded the lowest value of $e_{2,\text{max}}$ was stored and used to generate the plots of figure 2. Note that comparable outcomes could have been attained with different sets of parameters.

actual point in the extended space of the mutually interrelated new variables. As noted in the Introduction and commented in specific passages, there are several points of closeness with the work of [33]. That study was however devoted to the continuous-time ‘dynamical perceptron’ and to ‘dynamical recurrent neural networks’, while here the ‘perceptron units’ enter collectively into the UA strategy to approximate the actual rate-field components of the system under consideration.

The quadratization procedure presented here could be easily implemented in computer codes. If we had at our disposal a very great computational power and a very great memory-storage capability, we could exploit the quadratic format to explore the real dynamics by looking at their mirrored version in the extended space where interesting features can emerge (see footnote 1). For instance, at the practical level one could devise techniques for the automatic localization of the possible low-dimensional ‘slow manifolds’ toward which the trajectories converge and the system’s evolution slows down; this would be useful to build model-reduction methods based on the fast-slow timescale separation. The problematic aspect is that the dimension of the extended space is huge and rapidly grows as the rate field is more and more featured and as the time-window of interest becomes wider (see remark (ii) in section 3.1). Hence, the procedure here described apparently has little practical utility at the computational level. On the other hand, the scope here was more fundamental: demonstrating the existence of an *universal* route to achieve an *unique* underlying quadratic format of the evolution law regardless of the features of the specific dynamical system.

Finally, let us note that what has been devised here, in essence, is nothing but a way to convert autonomous ODEs into a quadratic extended format: no physics is invoked and even the time could be replaced by a generic progression variable. But if the system is truly *physical*, then we are in the position of going back to the provocative question raised by Peschel and Mende: *Do we live in a Volterra world?* Given the universality of the quadratization method, the answer should be affirmative, at least if we accept to deal with an ensemble of dynamical variables of virtually infinite dimension! Of course, the mapping of the real evolution law (as we write it on the basis of a physico-mathematical model) into the extended Lotka-Volterra counterpart is actually only a matter of ‘mathematical translation’. Going beyond this sensible level of interpretation would lead us to inconvenient metaphysical arguments that should be avoided.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix

A.1. GLV \rightarrow LV transformation

Let us consider a set of positive-valued dynamical variables v_i that evolve according to the following ODEs having a multivariate GLV format:

$$\dot{v}_i = \sum_m c_{mi} \Pi_m(\mathbf{v}), \quad \Pi_m(\mathbf{v}) := \prod_j v_j^{\eta_j^{(m)}} \quad (\text{A1})$$

Let us introduce

$$G_{i\mu}(\mathbf{v}) := \prod_j v_j^{\rho_j^{(i\mu)}} \quad (\text{A2})$$

where μ is an auxiliary index and the exponents $\rho_j^{(i\mu)}$ are, at this level, freely chosen. The time derivative of the functions $G_{i\mu}(\mathbf{v})$ is

$$\begin{aligned} \dot{G}_{i\mu} &= G_{i\mu} \sum_j \rho_j^{(i\mu)} \frac{\dot{v}_j}{v_j} \\ &= G_{i\mu} \sum_{j,m} \rho_j^{(i\mu)} c_{mj} \prod_{j'} v_{j'}^{\eta_{j'}^{(m)} - \delta_{j',j}} \end{aligned} \quad (\text{A3})$$

Now, let us set that the index μ runs over the same entries of m , and choose $\rho_j^{(i\mu)} \equiv \eta_j(\mu) - \delta_{i,j}$. With this position the new dynamical variables $G_{i\mu}(t) \equiv G_{i\mu}(\mathbf{v}(t))$ evolve according to

$$\dot{G}_{i\mu} = -G_{i\mu} \sum_{j,m} M_{i\mu,jm} G_{jm} \quad (\text{A4})$$

with

$$M_{i\mu,jm} = [\delta_{i,j} - \eta_j(\mu)] c_{mj} \quad (\text{A5})$$

Equation (15) has exactly the same structure of equation (A1). This implies that the above $GLV \rightarrow LV$ general route can be straightforwardly applied to the specific case of interest. By making the correct identification of the indexes, equations (A4)–(A5) directly lead to the quadratic format equation (18) with the matrix given in equation (19).

A.2. Backward transformation

Let us introduce the column array \mathbf{s} with elements $s_i := \mathcal{M}^{-1} \sum_{\mu} \ln G_{i\mu}$, where \mathcal{M} is the total number of μ terms. From equation (A2) it follows that $s_i = \sum_j K_{ij} \ln v_j$ where we have introduced the square matrix \mathbf{K} with elements

$$K_{ij} = \frac{1}{\mathcal{M}} \left(\sum_{\mu} \eta_j(\mu) \right) - \delta_{ij} \quad (\text{A6})$$

This sort of matrix is invertible [8], unless the original ODEs equation (A1) are linear (but even in such case the inversion route would be feasible by exploiting specific constraints). Thus we can write

$$v_i = e^{[\mathbf{K}^{-1}\mathbf{s}]_i} \quad (\text{A7})$$

which enables us to obtain the variables $\mathbf{v}(t)$ from the G 's at the time t .

ORCID iDs

Diego Frezzato  <https://orcid.org/0000-0002-9107-2899>

References

- [1] Peschel M and Mende W 1986 *The Predator-Prey Model: Do We Live in A Volterra World?* (Springer)
- [2] Kerner E H 1981 Universal formats for nonlinear ordinary differential systems *J. Math. Phys.* **22** 1366–71
- [3] Brenig L and Goriely A 1989 Universal canonical forms for time-continuous dynamical systems *Phys. Rev. A* **40** 4119–22
- [4] Hernández-Bermejo B and Fairén V 1995 Nonpolynomial vector fields under the Lotka-Volterra normal form *Phys. Lett. A* **206** 31–7
- [5] Gouzé J L 1990 *Transformation of Polynomial Differential Systems in the Positive Orthant Report RR-1308 INRIA p 10* (<https://inria.hal.science/inria-00075251/en>)
- [6] Fairén V and Hernández-Bermejo B 1996 Mass action law conjugate representation for general chemical mechanisms *J. Phys. Chem.* **100** 19023–8
- [7] Nicolini P and Frezzato D 2013 Features in chemical kinetics. I. Signatures of self-emerging dimensional reduction from a general formal of the evolution law *J. Chem. Phys.* **138** 234101
- [8] Ceccato A, Nicolini P and Frezzato D 2019 Recasting the mass-action rate equations of open chemical reaction networks into a universal quadratic format *J. Math. Chem.* **57** 1001–8
- [9] Hernández-Bermejo B 2002 Stability conditions and Liapunov functions for quasi-polynomial systems *Applied Math. Lett.* **15** 25–8
- [10] Rocha Filho T M, Gléria I M, Figueiredo A and Brenig L 2005 The Lotka-Volterra canonical format *Ecol. Model.* **183** 95–106
- [11] Gléria I, Brenig L, Rocha Filho T M and Figueiredo A 2017 Stability properties of nonlinear dynamical systems and evolutionary stable states *Phys. Lett. A* **381** 954–7
- [12] Gléria I, Brenig L, Rocha Filho T M and Figueiredo A 2017 Permanence and boundedness of solutions of quasi-polynomial systems *Phys. Lett. A* **381** 2149–52
- [13] Figueiredo A, Gléria I M and Rocha Filho T M 2000 Boundedness of solutions and Lyapunov functions in quasi-polynomial systems *Phys. Lett. A* **268** 335–41
- [14] Gléria I M, Figueiredo A and Rocha Filho T M 2001 Stability properties of a general class of nonlinear dynamical systems *J. Phys. A Math. Gen.* **34** 3561–75
- [15] Gléria I M, Figueiredo A and Rocha Filho T M 2003 A numerical method for the stability analysis of quasi-polynomial vector fields *Nonlinear Anal.* **52** 329–42
- [16] Motee N, Bahmieh B and Khammash M 2012 Stability analysis of quasi-polynomial dynamical systems with applications to biological network models *Automatica* **48** 2945–50
- [17] Magyar A, Szederkényi G and Hangos K M 2008 Globally stabilizing feedback control of process systems in generalized Lotka-Volterra form *J. Process Control* **18** 80–91
- [18] Magyar A and Hangos K M 2015 Globally stabilizing state feedback control design for Lotka-Volterra systems based on underlying linear dynamics *IFAC-PapersOnLine* **48** 1000–5
- [19] Márton L, Hangos K M and Magyar A 2021 Passivity of Lotka-Volterra and quasi-polynomial systems *Nonlinearity* **34** 1880–9
- [20] Bountis T and Brenig L 2022 Comparison between the QP formalism and the Painlevé property in integrable dynamical systems *Theor. Math. Phys.* **212** 1033–43
- [21] Brenig L, Gléria I, Rocha Filho T M, Figueiredo A and Hernández-Bermejo B 2018 Equivalence between nonlinear dynamical systems and urn processes *J. Phys. A: Math. Theor.* **51** 485101
- [22] Brenig L 2018 Reducing nonlinear dynamical systems to canonical forms *Phil. Trans. R. Soc. A* **376** 20170384
- [23] Nicolini P and Frezzato D 2013 Features in chemical kinetics. II. A self-emerging definition of slow manifolds *J. Chem. Phys.* **138** 234102
- [24] Ceccato A, Nicolini P and Frezzato D 2015 Features in chemical kinetics. III. Attracting subspaces in a hyper-spherical representation of the reactive system *J. Chem. Phys.* **143** 224109
- [25] Ceccato A, Nicolini P and Frezzato D 2017 A low-computational-cost strategy to localize points in the slow manifold proximity for isothermal chemical kinetics *Int. J. Chem. Kinet.* **49** 477–93
- [26] Ceccato A, Nicolini P and Frezzato D 2017 Attracting subspaces in a hyper-spherical representation of autonomous dynamical systems *J. Math. Phys.* **58** 092701

- [27] Frezzato D 2023 Intrinsic timing in classical master equation dynamics from an extended quadratic format of the evolution law *J. Math. Chem.* **61** 806–34
- [28] Cybenko G 1989 Approximation by superposition of a sigmoidal function *Math. Control Signals Systems* **2** 303–14
- [29] Hornik K 1991 Approximation capabilities of multilayer feedforward networks *Neural Netw.* **4** 251–7
- [30] Huang G-B, Chen Y-Q and Babri H A 2000 Classification ability of single hidden layer feedforward neural networks *IEEE Trans, Neural Networks* **11** 799–801
- [31] Huang G-B, Zhu Q-Y and Siew C-K 2006 Extreme learning machine: theory and applications *Neurocomputing* **70** 489–501
- [32] Tikk D, Kóczy L T and Gedeon T D 2003 A survey on Universal Approximation and its limits in soft computing techniques *Int. J. Approx. Reason.* **33** 185–202
- [33] Moreau Y, Louiès S, Vandewalle J and Brenig L 1999 Embedding recurrent neural networks into predator-prey models *Neural Netw.* **12** 237–45
- [34] Funahashi K-I and Nakamura Y 1993 Approximation of dynamical systems by continuous time recurrent neural networks *Neural Netw.* **6** 801–6
- [35] Sonoda S and Murata N 2017 Neural network with unbounded activation functions is universal approximator *Appl. Comput. Harmon. Anal.* **43** 233–68