ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

Random Unicycle Network (RUN!): supercharging harmonic oscillator networks via non-holonomic constraints

Mariano Ramírez Montero[1], Andrea Ceni[2], Andrea Cossu[2], Davide Bacciu[2], Claudio Gallicchio[2], and Cosimo Della Santina[1] *

1- Delft University of Technology - Cognitive Robotics
Delft - The Netherlands

2- University of Pisa - Department of Computer Science
Pisa - Italy

Abstract. Motivated by advances in physical reservoir computing, we seek models that retain the modularity of echo state networks while enriching their internal dynamics. Recent studies have demonstrated that oscillator networks can achieve this balance, although their simple harmonic nature may limit their expressiveness. Here, we investigate the idea of augmenting harmonic oscillators with non-holonomic (velocity-level) constraints, known to induce rich, nonlocal behaviors. We implement these constraints intrinsically within each dynamical unit, yielding a model equivalent to the unicycle — the canonical representation of the simplest vehicle. We test the model on three time-series classification benchmarks, achieving competitive or superior accuracy compared to the state of the art, with reservoirs as small as 20 unicycles.

1 Introduction

Time-series data are ubiquitous, and extracting meaningful patterns from them is increasingly important. Yet, practical constraints often limit the use of large deep learning models: cloud-based processing poses privacy and communication challenges, while edge devices such as wearables face tight power budgets. Reservoir computing (RC) offers an efficient alternative, achieving strong temporal modeling performance with far lower computational cost [1].

In physical reservoir computing, a wide range of nonlinear dynamical systems has been exploited as computational substrates, from photonic oscillators to mechanical and spintronic devices [2]. These studies show that rich dynamics can enhance temporal processing capabilities. Yet, within modular network formulations, the use of structured or non-trivial dynamical laws remains limited. Most works still rely on standard echo state networks (ESNs) [3] or spiking-based models [4]. In this vein, recent work on oscillator-based deep learning models has shown substantial promise [5, 6, 7] and motivated researchers to investigate the use of coupled harmonic oscillators in RC [8], demonstrating that this simple dynamical alternative can outperform ESNs.

*This work was supported under the European Union's Horizon Europe Program from Project EMERGE - Grant Agreement No. 101070918.

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

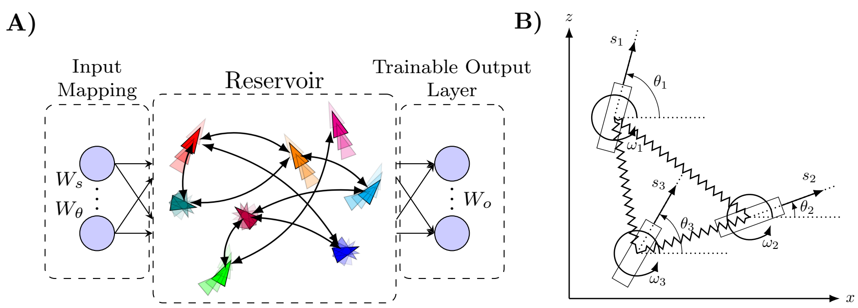

Figure 1: **Panel (A)**: Our model consists of coupled, non-holonomic unicycle units, represented above with coloured triangles. We use this network as a reservoir, training a single linear layer to map the last reservoir state to a desired output label. **Panel (B)**: A detailed example of what a three-unit network could look like, with the positions x, z, orientations θ, linear velocities s, and angular velocities ω shown for each unit.

With this work, we take a step further and ask: *how can we enhance the expressiveness of harmonic oscillator networks without compromising their modular structure and architectural simplicity?* To this end, we draw inspiration from a well-studied property in physics and control theory—*non-holonomy* [9]. These are systems whose state variables can be separated into position and velocity components, $(q, \dot{q})$, constrained to evolve according to velocity-level relations of the form $A(q)\dot{q} = 0$, where $A(q)$ is a matrix-valued function. Despite their apparent simplicity, such systems exhibit rich, nonlocal behaviors even in low-dimensional settings [10], having for example steady state configurations depending on the path taken.

Building on this idea, we propose to augment triplets of harmonic oscillators via a one-dimensional integrable constraint, yielding a compact model formally equivalent to a set of unicycles coupled through linear springs (Fig. 1). Physically, each unit behaves as a two-dimensional oscillator whose axis of oscillation can rotate within the plane. We refer to this augmented architecture as the *Random Unicycle Network (RUN!)*.

We evaluate the proposed RUN! reservoir on three benchmark time-series classification tasks widely used in the RC literature. Across all tests, the model achieves accuracy comparable to or exceeding that of state-of-the-art RC architectures: ESNs and oscillator-based reservoirs [8].

2 Random Unicycle Network (RUN!)

We start from a collection of coupled harmonic oscillators with unit mass as usually considered in the literature [5, 6, 8]: $\ddot{q} + \frac{\partial U}{\partial q} + D\dot{q} = f_e$, with configuration $q \in \mathbb{R}^{3N}$, potential $U : \mathbb{R}^{3N} \to \mathbb{R}$ (containing both the coupling terms and the self oscillatory ones), diagonal damping matrix $D \in \mathbb{R}^{3N \times 3N}$, and external forcing $f_e \in \mathbb{R}^{3N}$. Without loss of generality, we group the oscillators in triplets, partitioning the configuration as $q = (x, z, \theta)$, and similarly split D

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

into D_x, D_z, and $D_\theta \in \mathbb{R}^{N\times N}$, defining $\operatorname{diag}(D_x) = \beta_x$, $\operatorname{diag}(D_z) = \beta_z$, and $\operatorname{diag}(D_\theta) = \beta_\theta \in \mathbb{R}^N$. We also partition f_e into f_x, f_z, and $f_\theta \in \mathbb{R}^N$, yielding

$$\begin{bmatrix}\ddot{x}\\ \ddot{z}\\ \ddot{\theta}\end{bmatrix} + \begin{bmatrix}\frac{\partial U}{\partial x}\\ \frac{\partial U}{\partial z}\\ \frac{\partial U}{\partial \theta}\end{bmatrix} + \begin{bmatrix}D_x & 0 & 0\\ 0 & D_z & 0\\ 0 & 0 & D_\theta\end{bmatrix}\begin{bmatrix}\dot{x}\\ \dot{z}\\ \dot{\theta}\end{bmatrix} = \begin{bmatrix}f_x\\ f_z\\ f_\theta\end{bmatrix}. \tag{1}$$

Although formally equivalent to the original, it is useful to underline that an alternative interpretation of this dynamics is a collection of N oscillators moving in 3D space.

We now enrich this oscillator network by introducing the non-holonomic constraint $\dot{x}_i \sin\theta_i = \dot{z}_i \cos\theta_i$. Define $s \in \mathbb{R}^N$ by $s^2 = \dot{x}^2 + \dot{z}^2$ (element-wise), which allows the velocities of the first two sets of oscillators to be written as $(\dot{x}, \dot{z}) = (\cos\theta \odot s, \sin\theta \odot s)$. This satisfies the constraint for any s and θ, as can be verified by direct inspections. Differentiating s^2 yields

$$\dot{s} \odot s = \ddot{x} \odot \dot{x} + \ddot{z} \odot \dot{z}. \tag{2}$$

Substituting $\ddot{x}$ and $\ddot{z}$ gives

$$\dot{s} \odot s = \left(\frac{\partial U}{\partial x} + D_x \dot{x} + f_x\right)(\cos\theta \odot s) + \left(\frac{\partial U}{\partial z} + D_z \dot{z} + f_z\right)(\sin\theta \odot s). \tag{3}$$

Using $\operatorname{diag}(D_x) = \beta_x$ and $\operatorname{diag}(D_z) = \beta_z$ then yields

$$\dot{s} = \frac{\partial U}{\partial x} \odot \cos\theta + \frac{\partial U}{\partial z} \odot \sin\theta + \underbrace{\left(\beta_x \odot \cos^2\theta + \beta_z \odot \sin^2\theta\right)}_{\triangleq \beta_s(\theta)} s + \underbrace{f_x \odot \cos\theta + f_z \odot \sin\theta}_{\triangleq f_s}. \tag{4}$$

Next, defining $\omega = \dot{\theta}$ and using $\operatorname{diag}(D_\theta) = \beta_\theta$ gives the dynamics of our random unicycle network (RUN!):

$$\begin{bmatrix}\dot{x}\\ \dot{z}\\ \dot{\theta}\\ \dot{s}\\ \dot{\omega}\end{bmatrix} = \begin{bmatrix}\cos(\theta) \odot s\\ \sin(\theta) \odot s\\ \omega\\ \left(\frac{\partial U}{\partial x}\right) \odot \cos(\theta) + \left(\frac{\partial U}{\partial z}\right) \odot \sin(\theta) + f_s - \beta_s \odot s\\ \frac{\partial U}{\partial \theta} + f_\theta - \beta_\theta \odot \omega\end{bmatrix}, \tag{5}$$

with $x, z, \theta, s, \omega, \beta_s, \beta_\theta \in \mathbb{R}^N$.

This structure admits a natural geometric interpretation. The embedding of the constraint has transformed a network of N oscillators evolving in the full 3D space (x_i, z_i, θ_i) into N oscillators whose motion in the plane (x_i, z_i) is restricted to a single *instantaneous* direction $(\cos(\theta)_i, \sin(\theta)_i)$, determined by the current value of θ_i at time t. This mirrors and generalizes the behavior of classical unicycle models from the non-holonomic mechanics literature, and Equation (5) extends this principle to a swarm of unicycles—thus motivating the name of our model.

To turn (5) into a reservoir, the input $u(t)$ is mapped into f_s and f_θ via

$$f_s = W_s\, u(t), \tag{6}$$

$$f_\theta = W_\theta\, u(t), \tag{7}$$

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

Finally, after the reservoir processes the input, the final state at t_f is passed through a linear readout,

$$\hat{y} = W_o \left[x^\top(t_f)\;\; z^\top(t_f)\;\; \theta^\top(t_f)\;\; s^\top(t_f)\;\; \omega^\top(t_f)\right]^\top + b, \tag{8}$$

where $W_o \in \mathbb{R}^{n_o \times N}$ and $b \in \mathbb{R}^{n_o}$ are trained to predict $\hat{y} \in \mathbb{R}^{n_o}$.

2.1 Tuning the reservoir

In this work, we consider the potential

$$U(y) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{2} K_{ij} \left(A_{ij} - d(i,j)\right)^2 + \frac{1}{2} \kappa_{ij} \left(B_{ij} - (\theta_i - \theta_j)\right)^2 ,$$

where $K, \kappa \succ 0$ are symmetric stiffness matrices, $A, B \in \mathbb{R}^{N \times N}$ are antisymmetric matrices defining the rest lengths, and $d(i,j) = \sqrt{(x_i - x_j)^2 + (z_i - z_j)^2}$ is the Euclidean distance. The first term induces a spatial elastic coupling, while the second produces a rotational one, corresponding to unicycle units linked through virtual springs. To avoid numerical drift, we anchor a single unit by imposing the non-holonomic constraints $\dot{s}_0 = 0$ and $\dot{\omega}_0 = 0$. For simplicity, we also set $\beta_x = \beta_y$. A schematic example of a three-unit reservoir and its state variables is shown in Fig. 1B.

Following the standard RC paradigm, reservoir parameters are randomly initialized and left untrained. Their sampling ranges, however, strongly influence performance. We therefore use Bayesian optimization on a validation set to tune the bounds of the uniform distributions from which the parameters are drawn. The optimized quantities include the linear and angular stiffnesses K and κ, the linear and angular damping coefficients β_s and β_θ, the rest-length matrices A and B, and the input maps W_s and W_θ.

We also tune the sparsity levels of W_s and W_θ, as well as the number of linear and angular couplings between units (i.e., the nonzero entries of K and κ). Finally, we optimize the integration timestep, an input bias term, and a washup parameter defining how many autonomous steps the reservoir evolves before receiving external input.

3 Experiments

RUN! is tested on the sMNIST, FordA, and npCIFAR10 classification benchmarks. The sMNIST task involves classifying 28×28 pixel images of hand-drawn digits, which are flattened into a 1D sequence of 784 time steps. In the FordA dataset, a one-dimensional signal of 500 timesteps is classified as originating from either an anomalous or a normal engine. The npCIFAR10 task uses the CIFAR10 image classification dataset, which contains 32×32-pixel color images of 10 classes of objects, such as airplanes, cats, ships, dogs, and others. The images' RGB channels are reshaped into time series with 32 timesteps and 96 features. Moreover, in this specific task, the time series is padded with 968 timesteps of noise, and thus this task tests the reservoir's capacity to remember the information at the beginning of the signal.

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

Table 1: Test accuracy for the sMNIST, FordA, and npCIFAR-10 benchmarks on test data. We report the mean andobject classes, such as airplanes, cats, ships, and dogsth the best performance for each benchmark in bold and the second is underlined. RON is taken from [8], and is also based on oscillator network.

	Leaky ESN	**RON**	**RUN! (Ours)**
sMNIST ↑	0.8510 $_{0.0020}$	**0.94652** $_{\mathbf{0.0049}}$	<u>0.9439</u> $_{0.0023}$
FordA ↑	0.5461 $_{0.0320}$	<u>0.6885</u> $_{0.0385}$	**0.8522** $_{\mathbf{0.0312}}$
npCIFAR10 ↑	0.2285 $_{0.0010}$	<u>0.3893</u> $_{0.0053}$	**0.4509** $_{\mathbf{0.0056}}$

3.1 Benchmark results

The results for these benchmarks are shown in Table 1. We compare our model with two other RC models: the Leaky ESN [3] and RON [8]. The Leaky ESN is a popular RC model used for time series classification, while RON (Random Oscillators Network) is a more recent model that uses coupled harmonic oscillators as its reservoir. For a given benchmark, we use the same number of trainable parameters for all models, and for the Leaky ESN and RON models we use the hyperparameters reported in [8]. As the results in Table 1 show, our model has comparable or better performance in all of the classification benchmarks.

3.2 Ablations

We perform several ablations on the FordA binary classification task, focusing on the importance of the angular dynamics of the unicycle units. First, we remove any angular input forcing (see (7)), $W_\theta = 0$. In the second case, we only remove angular coupling, $\kappa = 0$. Finally in the last case, we remove both of these: $W_\theta = \kappa = 0$, but run the Bayesian optimization again to find a new set of hyperparameters.

Table 2: FordA classification test accuracy for the nominal RUN! and for three ablations: no angular input ($W_\theta = 0$), no angular coupling ($\kappa = 0$), and no angular input or coupling ($W_\theta = \kappa = 0$) but with re-optimized hyperparameters, with the best result in bold. Note the drop in performance for $W_\theta = 0$.

Model	**FordA test accuracy ↑**
RUN!	**0.8522** $_{\mathbf{0.0312}}$
$W_\theta = 0$	0.5090 $_{0.0199}$
$\kappa = 0$	0.8228 $_{0.0324}$
$T_\theta = \kappa = 0$, re-optimized	0.6027 $_{0.0370}$

The first two ablations make it clear that the angular dynamics are important for the reservoir's performance. Specifically, the angular input forcing is the most important: as the second row shows, performance drops to random chance without it. Finally, while the last ablation shows that some performance can be recovered by re-tuning the reservoir hyperparameters, the full RUN! model is still considerably better.

4 Conclusions

We introduced RUN!, a dynamical architecture obtained by adding nonholonomic constraints to coupled harmonic oscillators, and showed that it matches

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

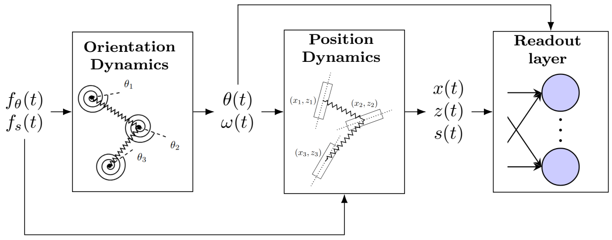

Figure 2: Since the orientation dynamics of the unicycles are independent of their position dynamics, our reservoir can also be seen as a two-layer reservoir.

or exceeds state-of-the-art performance on three time-series classification benchmarks. These results suggest that non-holonomicity can generate non-local state evolution suited for long-term memory, though the underlying mechanism still warrants a more careful analysis. A structural reading is nonetheless available: RUN! can be interpreted, in the spirit of Deep RC, as a two-layer reservoir (Fig. 2). The first reservoir layer corresponds to the orientation dynamics, which already behaves as a full network of N standard coupled harmonic oscillators with configuration $\theta \in \mathbb{R}^N$. The output of this first layer then excites the position dynamics, effectively forming the second reservoir. Looking ahead, we will explore *nonholonomic networks* beyond the reservoir setting—i.e., fully trained architectures—and assess their performance on regression tasks.

References

[1] M. Yan et al. "Emerging opportunities and challenges for the future of reservoir computing". In: *Nature Communications* 15.1 (2024), p. 2056.

[2] E. Picco, A. Lupo, and S. Massar. "Deep Photonic Reservoir Computer for Speech Recognition". In: *IEEE Transactions on Neural Networks and Learning Systems* 36.4 (2025), pp. 7606–7614.

[3] H. Jaeger et al. "Optimization and applications of echo state networks with leaky-integrator neurons". In: *Neural networks* 20.3 (2007), pp. 335–352.

[4] H. Peng et al. "Reservoir computing models based on spiking neural P systems for time series classification". In: *Neural Networks* 169 (2024), pp. 274–281.

[5] T. K. Rusch and S. Mishra. "Coupled Oscillatory Recurrent Neural Network (coRNN): An accurate and (gradient) stable architecture for learning long time dependencies". In: *International Conference on Learning Representations (ICLR 2021)*. OpenReview. 2021.

[6] T. K. Rusch and S. Mishra. "Unicornn: A recurrent model for learning very long time dependencies". In: *International Conference on Machine Learning*. PMLR. 2021, pp. 9168–9178.

[7] M. Stölzle and C. Della Santina. "Input-to-State Stable Coupled Oscillator Networks for Closed-form Model-based Control in Latent Space". In: *Advances in Neural Information Processing Systems*. Ed. by A. Globerson et al. Vol. 37. Curran Associates, Inc., 2024, pp. 82010–82059.

[8] A. Ceni et al. "Random oscillators network for time series processing". In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 2024, pp. 4807–4815.

[9] A. M. Bloch and B. Brogliato. "Nonholonomic mechanics and control". In: *Appl. Mech. Rev.* 57.1 (2004), B3–B3.

[10] R. W. Brockett and L. Dai. "Non-holonomic kinematics and the role of elliptic functions in constructive controllability". In: *Kluwer International Series in Engineering and Computer Science* (1992), pp. 1–1.