COMMUNICATION

# Reservoir computing using networks of memristors: effects of topology and heterogeneity (Supplementary Information)

J. B. Mallinson [a], Z. E. Heywood [a,b], R. K. Daniels [a], M. D. Arnold [c], P. J. Bones [a,b] and S. A. Brown [a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

## S1 Comparison of Percolating Networks of Nanoparticles (PNNs) and Heterogeneous Arrays (HAs) based on Network Statistics

In order to examine the relationship between reservoir structure and performance, we compare the computational performance of PNNs with that of regular (square) arrays of memristive tunnel gaps (MTGs). In order to facilitate decisions that must be made about which array sizes should be compared with the PNNs, here we examine key statistical properties of an ensemble of 500 randomly constructed PNNs which each have dimensions of 200×200 particle diameters. We then relate these properties to those of square arrays in order to choose the appropriate array sizes that ensure a 'fair' comparison.

The key properties considered are the number of output electrodes $N_{out}$ (Fig. S1a), the mean path length $\langle L_P \rangle$ (Fig. S1b) and the number of MTGs in the network $N_{MTG}$ (Fig. S1c). Finally, we discuss the distribution of MTG sizes (Fig. S1d).

### PNNs

In the RC literature $N_{out}$ typically has a strong impact on task performance at least partly because it determines the number of predictor variables which are used for training (and hence the number of weights to be trained). For PNNs, the most common number of outputs (i.e. number of groups of particles overlapping the right edge of the system) is $N_{out} = 12$ (Fig. S1a), and so, we present data only for PNNs which have this exact value of $N_{out}$. The mean path length $\langle L_P \rangle$ is the average number of MTGs which are connected in series along the conducting pathways within the reservoir. The voltage divider effect means that MTGs in reservoirs with larger (smaller) $\langle L_P \rangle$ experience smaller (larger) local potentials and therefore exhibit weaker (stronger) responses (See ESI S2). The most commonly observed path length in PNNs is $\langle L_P \rangle = 12$ (Fig. S1b). The number of MTGs in a reservoir is also important for RC because it potentially controls the collective response of the MTGs and hence the spatiotemporal dynamics which provide the nonlinearity and memory to the transformation of input signals. A larger number of MTGs likely provides richer dynamics to the reservoir and greater diversity to the reservoir output currents. For PNNs with dimensions of 200×200 particle diameters, $N_{MTG} \approx 2000$ (Fig. S1c).

### Regular arrays

The properties discussed in the previous subsection differ for square arrays and it is not possible to match the values of all three quantities for PNNs with those of single square arrays of a single size. For a square array of size $n \times n$ nodes, as shown in Fig. S1e, there are $n$ right-hand (left-hand) edge nodes which are treated as output (input) electrodes i.e. $N_{out} = n$. Since the arrays are square, the mean path length $\langle L_P \rangle$ from the input to output electrodes is also $n$. There are no MTGs which connect adjacent input (output) nodes and so the number of MTGs within the network is $N_{MTG} = (n-1)(2n-2)$. Given that in the PNNs $N_{out} = \langle L_P \rangle = 12$, an array size of $n = 12$ seems to be an appropriate choice. However, the 12×12 array has $N_{MTG} = 242$, which is roughly an order of magnitude smaller than the corresponding value for the PNNs. We therefore consider a second array size of $n = 36$ which has $N_{MTG} = 2450$. While this is not an exact match to the PNNs ($n = 33$ gives the closest match with $N_{MTG} = 2048$), $n = 36$ allows us

[a] The MacDiarmid Institute for Advanced Materials and Nanotechnology, School of Physical and Chemical Sciences, University of Canterbury, Christchurch, New Zealand. E-mail: simon.brown@canterbury.ac.nz
[b] Electrical and Computer Engineering, University of Canterbury, Private Bag 4800,
[c] Christchurch 8140, New Zealand
[d] School of Mathematical and Physical Sciences, University of Technology Sydney, PO Box 123 Broadway NSW 2007, Australia
† Electronic Supplementary Information (ESI) available: DOI: 10.1039/x0xx00000x

to set $N_{out} = 12$ by dividing the 36 output nodes into groups of 3 (i.e. each 3 adjacent output nodes are shorted together). This is important as the PNNs, the 12×12 arrays and the 36×36 arrays all have $N_{out} = 12$.

### MTG Heterogeneity

Finally, we note that the distribution of MTG sizes $D$ for the PNNs (Fig. S1d) is approximately a random distribution over the range $L = [1\times10^{-6}, 0.173]$ particle diameters. Particles separated by gaps < $1\times10^{-6}$ particle diameters are considered to be connected (i.e. groups are merged) and the upper limit, $L = 0.173$, corresponds to conductance values approaching machine precision, i.e. $G = \alpha e^{-\beta(0.173)} \approx 1\times10^{-15}$. There are of course larger gaps between particle groups but these have negligible conductance and are therefore excluded from the calculations. In the heterogeneous arrays (HAs) we assign random MTG sizes over this same range. A single value of $D_{\mathrm{uniform}} = 0.05$ particle diameters is used for all MTGs in the uniform arrays (UAs) and uniform PNNs (UPNNs).

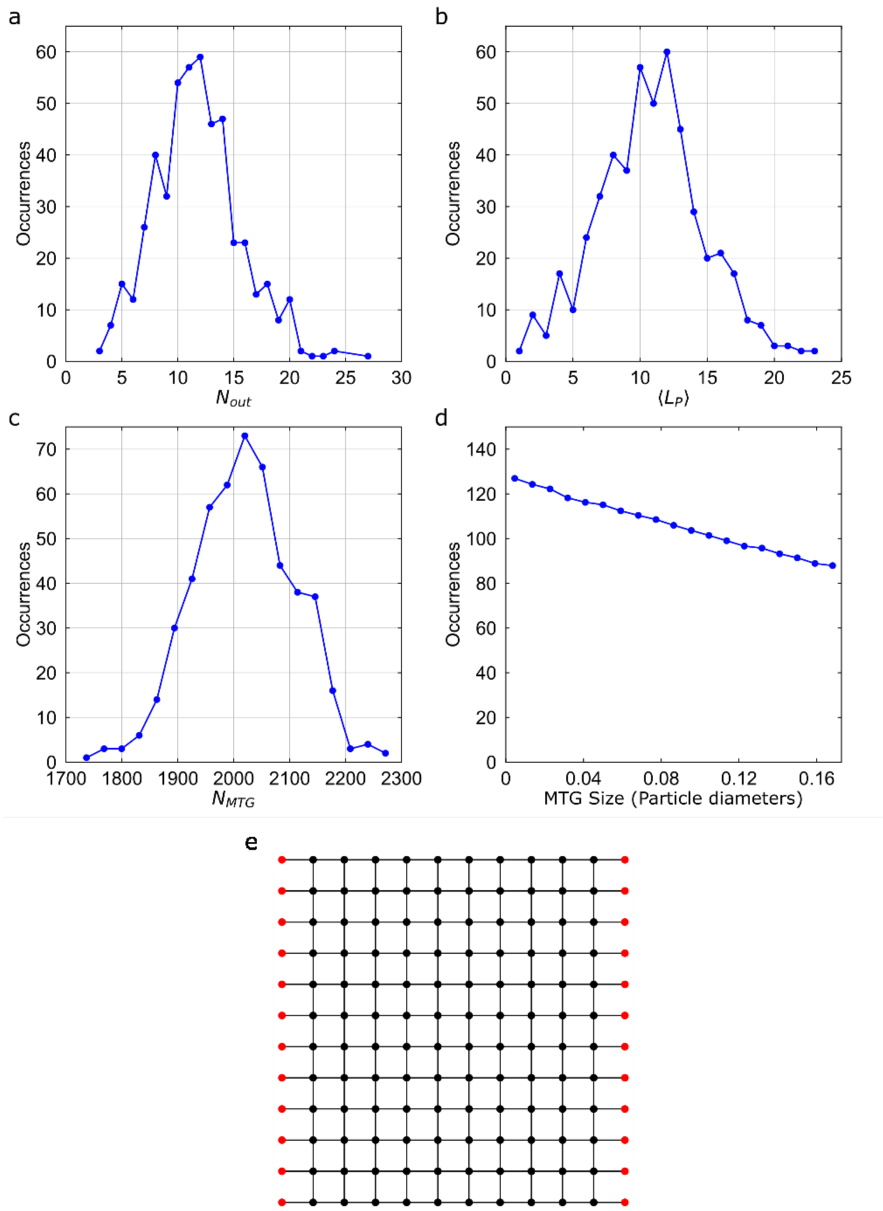


Figure S1: PNN characteristics. All statistics are calculated from 500 realizations of PNNs with size 200×200 particle diameters. (a) Distribution of the number of output electrodes $N_{out}$. Output electrodes are defined as any particle group which overlaps the right-hand edge of the network and which carries a measurable current (≥$1\times10^{-15}$ A). (b) Distribution of the PNN mean path length $\langle L_P \rangle$. $\langle L_P \rangle$ is the mean number of MTGs along the dominant current pathways. Note that these values are calculated when the system has reached dynamical equilibrium after the application of a constant DC voltage (0.5 V). (c) Distribution of the number of MTGs $N_{MTG}$ within the PNN. (d) Distribution of the initial sizes (D) of memristive tunnel gaps (MTG) in PNNs. Note these initial gap sizes are measured in the absence of hillocks. (e) Schematic of a regular array network. Nodes are analogous to particle groups in PNNs while links represent MTGs. Edge nodes (red) are designated as input (left) and output (right) nodes.




Please do not adjust margins

## S2 I-V characteristics Showing Network Nonlinearity and Hysteresis

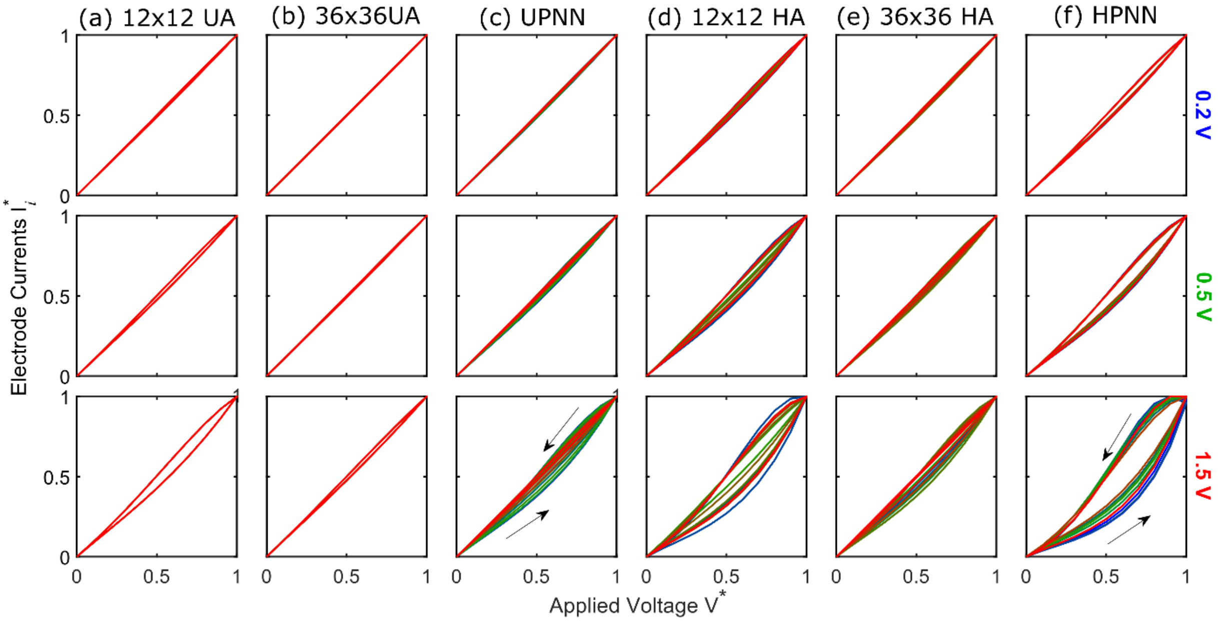


Figure S2: Network nonlinearity at different applied voltages. Normalized I-V characteristics showing how both the nonlinearity and diversity of the different electrode currents increase with input amplitude for representative networks. $\mathrm{I}_i^* = \mathrm{I}_i/\max(\mathrm{I}_i)$ where $\mathrm{I}_i$ is the output current from electrode $i$. $\mathrm{V}^* = \mathrm{V}_{\mathrm{app}}/\mathrm{V}_{\max}$ where $\mathrm{V}_{\mathrm{app}}$ is the applied voltage and $\mathrm{V}_{\max} = \max(\mathrm{V}_{\mathrm{app}})$. The voltage sweep period was chosen to be close to the MTG response time (period = 20, $T = 5$) allowing the hysteresis to be clearly seen. Sweeps were applied with three different $\mathrm{V}_{\max}$: 0.2 V, 0.5 V, 1.5 V. The 12 individual (normalized) electrode currents $\mathrm{I}_i^*$ are shown in different colours. It is clear that all of the electrode outputs are close to linear at $\mathrm{V}_{\max} = 0.2$ V, while for increasing $\mathrm{V}_{\max}$ the levels of both nonlinearity and hysteresis increase.

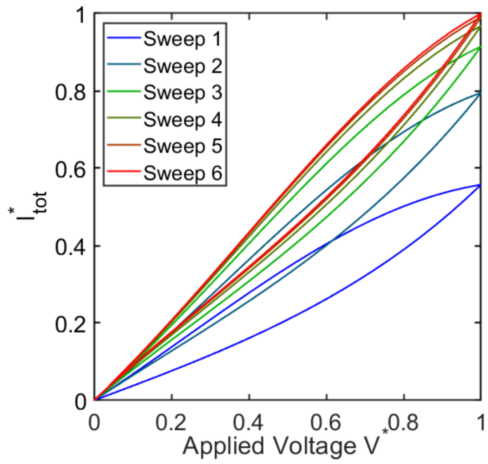


Figure S3: Evolution of hysteresis. Normalized $\mathrm{I}_{\mathrm{tot}}$-V curves from a HPNN are shown for relatively fast voltage sweeps with respect to the network response times (period = 50, $T = 50$) with $\mathrm{V}_{\max} = 1.5$ V. The different coloured lines represent the normalized total current ($\mathrm{I}_{\mathrm{tot}}^* = \mathrm{I}_{\mathrm{tot}}/\max(\mathrm{I}_{\mathrm{tot}})$ where $\mathrm{I}_{\mathrm{tot}}$ is the sum of all output electrode currents. $\mathrm{V}^* = \mathrm{V}_{\mathrm{app}}/\mathrm{V}_{\max}$ where $\mathrm{V}_{\mathrm{app}}$ is the applied voltage and $\mathrm{V}_{\max} = \max(\mathrm{V}_{\mathrm{app}})$) for subsequent sweeps of the applied voltage. The PNN is initially in the ground state (no hillocks exist within the MTGs in the PNN) and the network response evolves from sweep to sweep as the dynamics approach a limit cycle (Sweeps 5 and 6 are almost identical and subsequent sweeps are the same as Sweep 6).

## S3 Examples of Explicit RC Results

This section presents examples of the performance of each reservoir for all five RC tasks (Figures S4-S8). The results shown correspond to the optimal $T$ and $V_{max}$ shown in the tables below for a randomly selected network realization and input sequence. For the UAs, where there is no dependence on $T$ and $V_{max}$, the optimal values from the corresponding HAs were used. Note that for consistency with the labelling of the curves in the main text (Figures 5-6) the same colours are used here to present the responses of the various systems for each task.

NARMA2

| | Optimal T | Optimal $V_{max}$ |
|---|---|---|
| 12x12 UA | 2 | 0.2 |
| 36x36 UA | 1 | 0.5 |
| HPNN | 1 | 0.2 |
| 12x12 HA | 2 | 0.2 |
| 36x36 HA | 1 | 0.5 |
| UPNN | 1 | 0.2 |

NARMA10

| | Optimal T | Optimal $V_{max}$ |
|---|---|---|
| 12x12 UA | 6 | 0.2 |
| 36x36 UA | 13 | 0.2 |
| HPNN | 5 | 0.2 |
| 12x12 HA | 6 | 0.2 |
| 36x36 HA | 13 | 0.2 |
| UPNN | 4 | 0.2 |

NLT

| | Optimal T | Optimal $V_{max}$ |
|---|---|---|
| 12x12 UA | 1 | 1.5 |
| 36x36 UA | 1 | 1.5 |
| HPNN | 1 | 1.5 |
| 12x12 HA | 1 | 1.5 |
| 36x36 HA | 1 | 1.5 |
| UPNN | 1 | 1.5 |

WD

| | Optimal T | Optimal $V_{max}$ |
|---|---|---|
| 12x12 UA | 1 | 1.5 |
| 36x36 UA | 1 | 1.5 |
| HPNN | 1 | 0.5 |
| 12x12 HA | 1 | 1.5 |
| 36x36 HA | 1 | 1.5 |
| UPNN | 1 | 0.5 |

MC

| | Optimal T | Optimal $V_{max}$ |
|---|---|---|
| 12x12 UA | 14 | 0.5 |
| 36x36 UA | 4 | 1.5 |
| HPNN | 13 | 0.2 |
| 12x12 HA | 14 | 0.5 |
| 36x36 HA | 4 | 1.5 |
| UPNN | 12 | 0.5 |

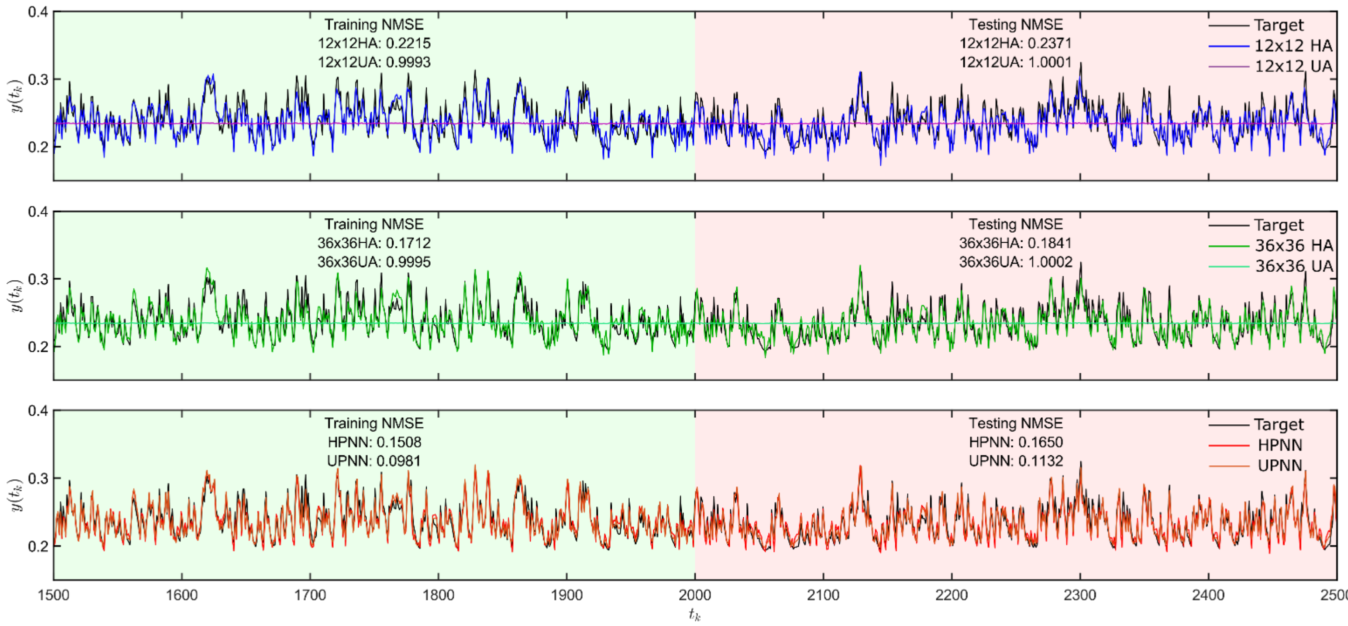


Figure S4: NARMA-2 Task Performance. The NARMA-2 target function $y$ (black) is overlaid in each panel with constructed outputs $\hat{y}$ from the 12x12 arrays (top), 36x36 arrays (middle) and PNNs (bottom). The NMSE performance metric is calculated from $y$ and $\hat{y}$. Note that for consistency with the labelling of the curves in the main text (Figures 5-6) the same colours are used here to present the responses of the various systems for each task. The responses for the UAs (top two panels) are essentially horizontal lines, due to the poor performance of the networks. The responses for the UPNNs and HPNNs are almost indistinguishable on this scale.

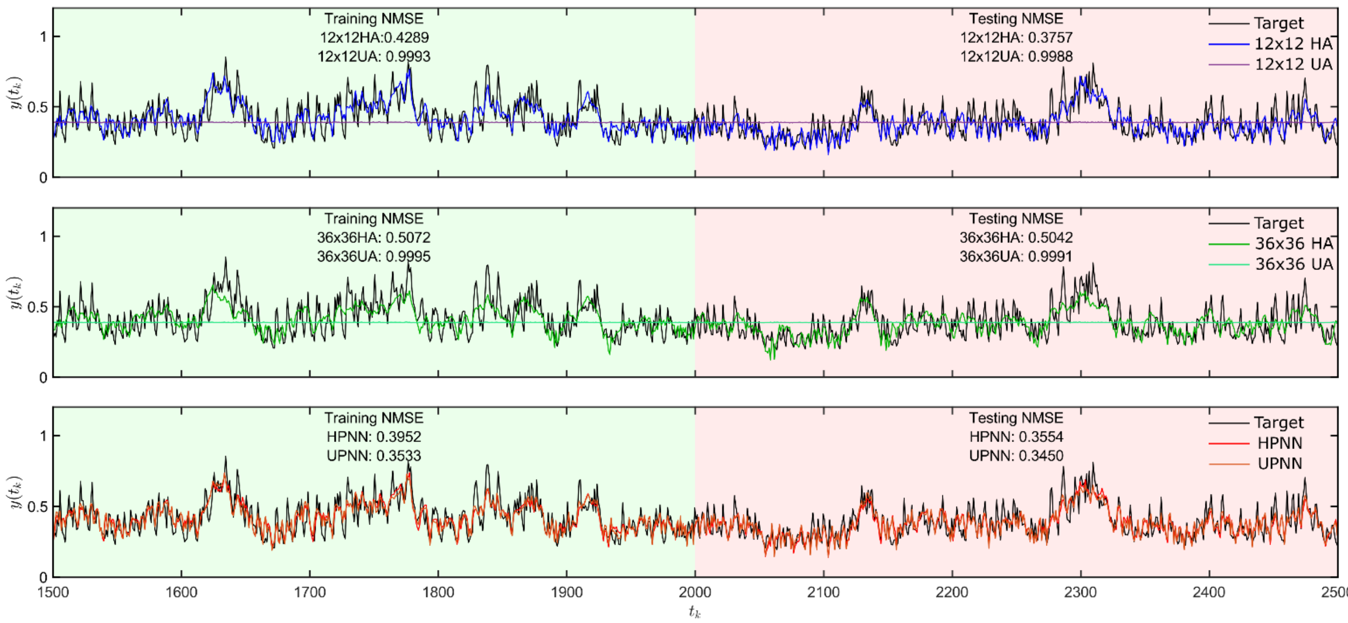


Figure S5: NARMA-10 Task Performance. The NARMA-10 target function $y$ (black) is overlaid in each panel with constructed outputs $\hat{y}$ from the 12x12 arrays (top), 36x36 arrays (middle) and PNNs (bottom). The NMSE performance metric is calculated from $y$ and $\hat{y}$. Note that for consistency with the labelling of the curves in the main text (Figures 5-6) the same colours are used here to present the responses of the various systems for each task. The responses for the UAs (top two panels) are essentially horizontal lines, due to the poor performance of the networks. The responses for the UPNNs and HPNNs are almost indistinguishable on this scale.

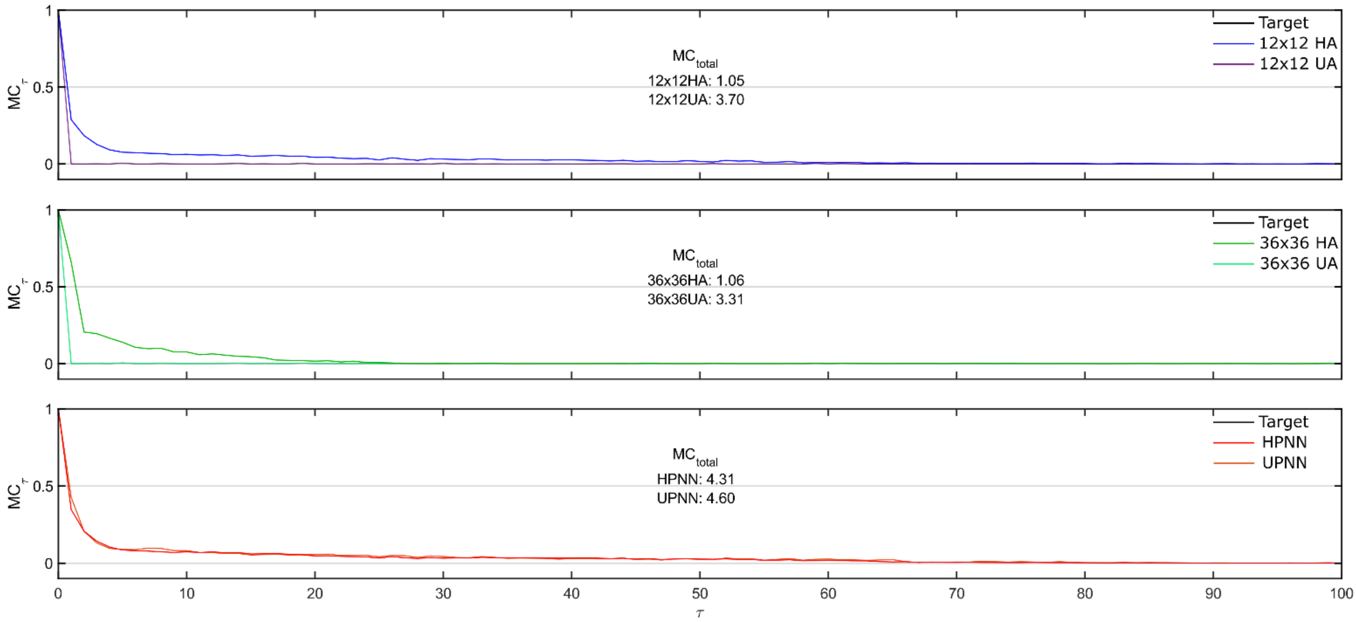


Figure S6: MC Task Forgetting Functions. $\mathrm{MC}_\tau$ as a function of delay $\tau$ for the 12x12 arrays (top), 36x36 arrays (middle) and PNNs (bottom). $\mathrm{MC}_{\mathrm{total}}$ is the sum of $\mathrm{MC}_\tau$ over all $\tau$. Note that for consistency with the labelling of the curves in the main text (Figures 5-6) the same colours are used here to present the responses of the various systems for each task. The responses for the UPNNs and HPNNs are almost indistinguishable on this scale.

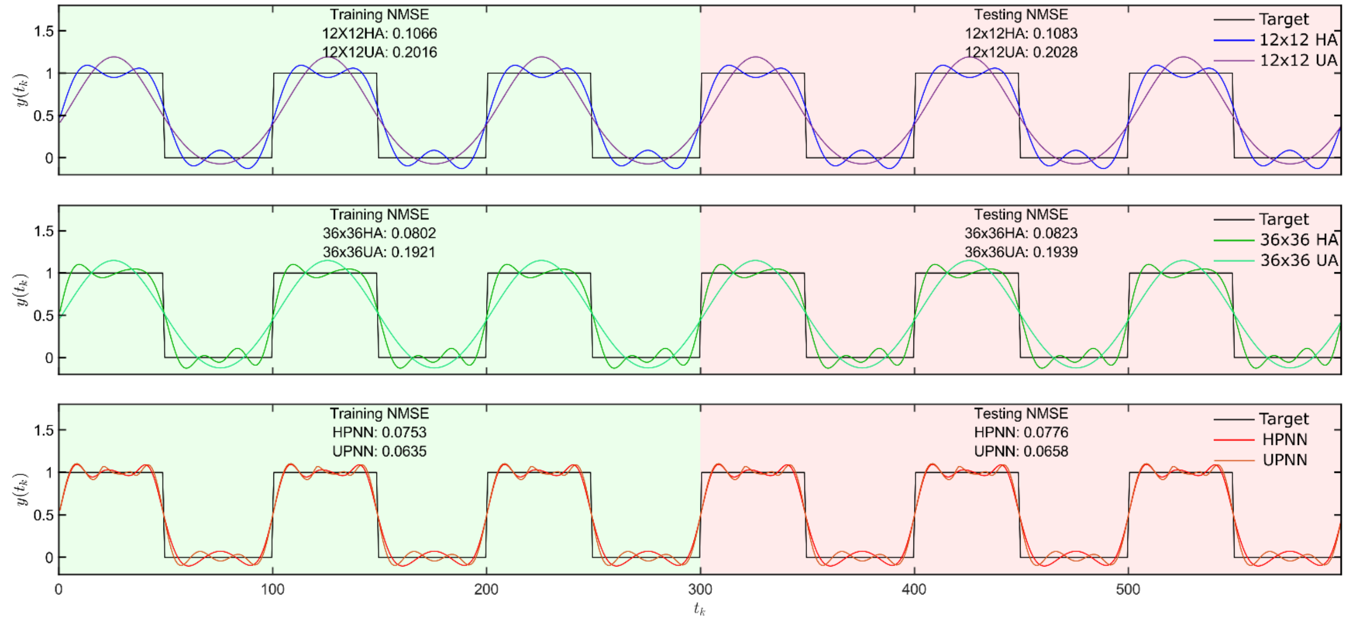


Figure S7: NLT Task Performance. The NLT target function $y$ (black) is overlaid in each panel with constructed outputs $\hat{y}$ from the 12x12 arrays (top), 36x36 arrays (middle) and PNNs (bottom). The NMSE performance metric is calculated from $y$ and $\hat{y}$. The NMSE for the UAs ~ 0.2 is similar to the NMSE value obtained for the error between a sine wave and a square wave, indicating that the UAs do not successfully transform the input. Note that for consistency with the labelling of the curves in the main text (Figures 5-6) the same colours are used here to present the responses of the various systems for each task.

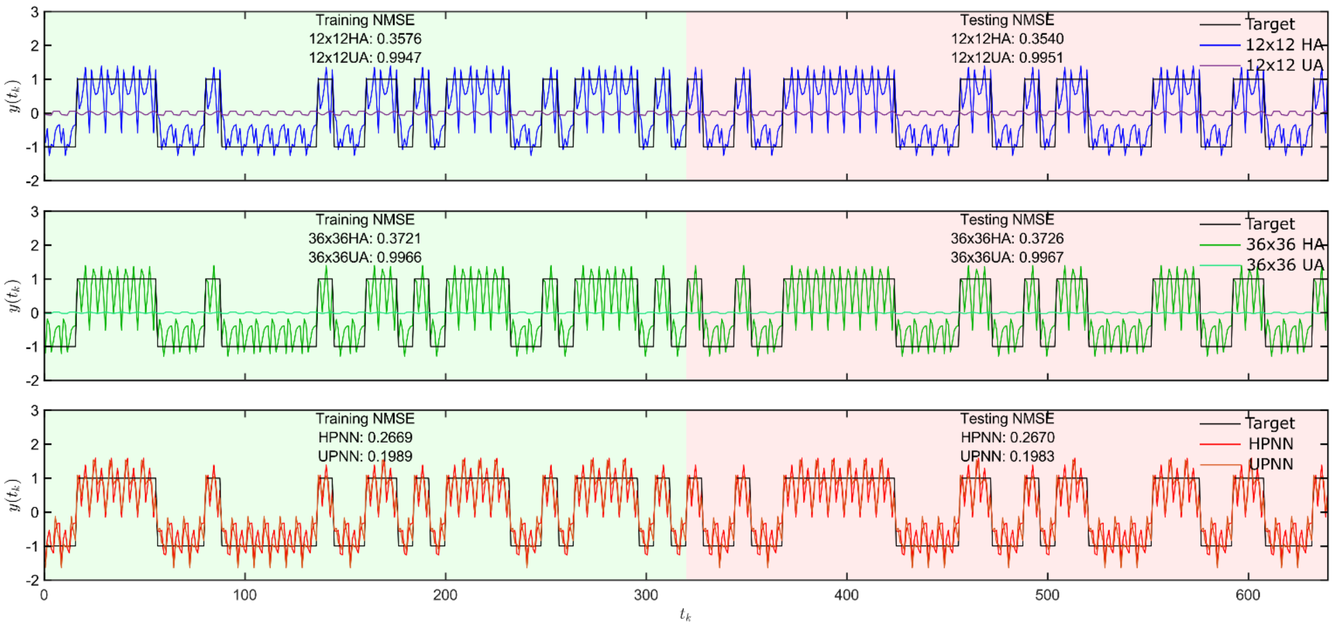


Figure S8: WD Task Performance. The WD target function $y$ (black) is overlaid in each panel with constructed outputs $\hat{y}$ from the 12x12 arrays (top), 36x36 arrays (middle) and PNNs (bottom). The NMSE performance metric is calculated from $y$ and $\hat{y}$. Note that for consistency with the labelling of the curves in the main text (Figures 5-6) the same colours are used here to present the responses of the various systems for each task. The responses for the UAs (top two panels) are essentially horizontal lines, due to the poor performance of the networks. The responses for the UPNNs and PNNs are almost indistinguishable on this scale.

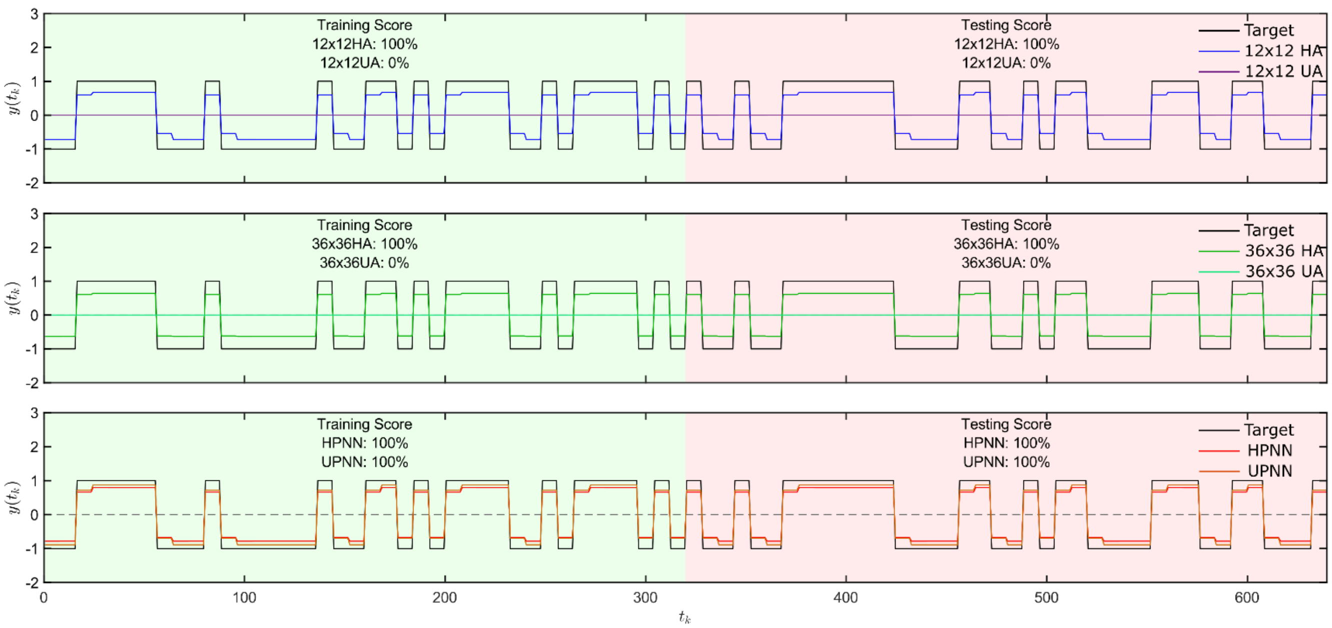


Figure S9: WD Task Performance. The WD target function $y$ (black) is overlaid in each panel with the mean of the output ($\hat{y}$, shown in Fig. S8) over each input waveform $\langle\hat{y}(t)\rangle_{\mathrm{wf}}$: 12x12 arrays (top), 36x36 arrays (middle) and PNNs (bottom). The classification of each input waveform is calculated by comparing $\langle\hat{y}(t)\rangle_{\mathrm{wf}}$ with the decision boundary at zero (dashed). Positive (negative) values greater than a threshold of 0.01 are classified as sine (square). The classification score is the number of correctly classified waveforms divided by the total number of waveforms in the input sequence. Note that for consistency with the labelling of the curves in the main text (Figures 5-6) the same colours are used here to present the responses of the various systems for each task. The responses for the UAs (top two panels) are essentially horizontal lines, due to the poor performance of the networks. The responses for the UPNNs and HPNNs are almost indistinguishable on this scale.

## S4 Symmetry Breaking in Uniform MTG Arrays (UAs)

The uniform arrays (UAs) perform poorly at RC tasks (Fig. 5 in the main text) because the uniform MTG size leads to symmetrically distributed currents and temporally synchronized MTG dynamics. The resulting reservoir dynamics lack richness and the reservoir outputs are linearly dependent on each other ($\mathrm{rank}(X) = 2$). It was shown in Fig. 6 that heterogeneous MTG sizes break this symmetry and allow the arrays to perform as well as the PNNs. Here we demonstrate that it is also possible to break the symmetry of the UAs by randomly weighting the inputs, or by using a single input electrode.

### Symmetry Breaking using Random Input Weights

In Fig. 5, all of the nodes at the left-hand edge are used as input electrodes and each input electrode receives the same input signal. By multiplying the input to each electrode by a random weight in the range [0, 1], the symmetry of the reservoir dynamics is broken and Fig. S10 shows that the performance becomes similar to that of the PNNs. The HPNN and UPNN performance from Fig. 6 are included for reference.

### Symmetry Breaking using a Single Input Electrode

The symmetry of the UAs can also be broken by applying the input to a single electrode. Figure S6 shows the task performance of 12x12 and 36x36 UAs with the input applied to electrode 6 only (electrodes numbered 1-12 from top to bottom along the left edge of the array). Again, the HPNN and UPNN performance from Fig. 6 is included for reference.

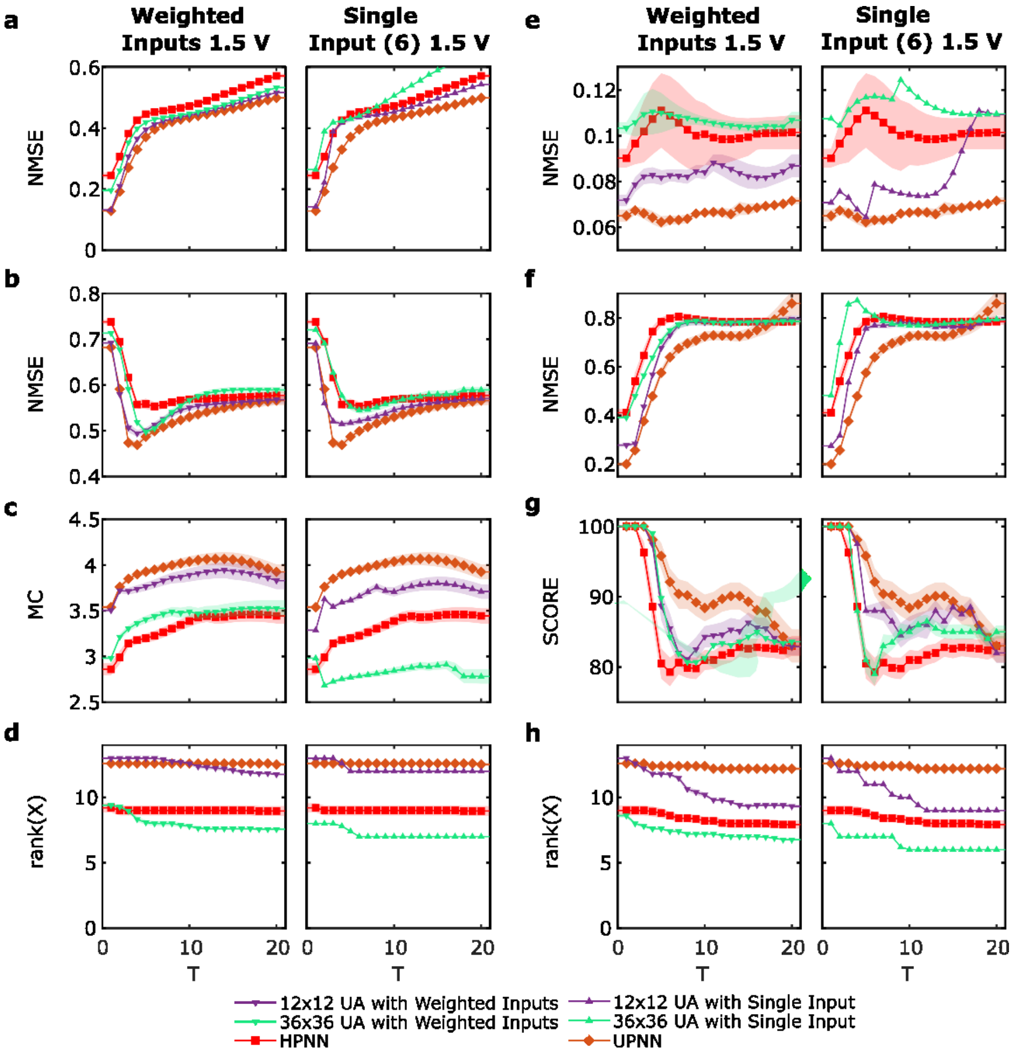


Figure S10: Symmetry breaking using different input configurations. For both (a-d) and (e-h), the left column shows the task performance of the 12x12 and 36x36 UAs with randomly weighted inputs while the right column shows the task performance of the 12x12 and 36x36 UAs with a single (unweighted) input. $V_{max} = 1.5$ V with $V_{min} = 0.1$ V for all tasks. All results are the mean of five network realizations and five input sequences (25 trials total) except for NLT (e) which has only one input sequence. Shaded areas correspond to the mean standard error. All results are expressed as a function of the network response time, $T$. (a) NARMA-2 and (b) NARMA-10 task performance expressed as the NMSE. (c) Linear memory capacity. (d) The rank of the predictor matrix $X$ corresponding to (a-d). (e) NLT and (f) waveform discrimination task performance expressed as the NMSE. (g) Waveform discrimination task performance expressed as the % of correctly classified waveforms. (h) The rank of the predictor matrix $X$ (during waveform discrimination), corresponding to the number of linearly independent outputs from the reservoir. The maximum possible rank is 13, corresponding to 12 electrode outputs and one constant bias term.