Unfolding recurrence by Green's functions for optimized reservoir computing

arxiv: 2010.06247

Sandra Nestler, Christian Keup, David Dahmen, Matthieu Gilson, Holger Rauhut, Moritz Helias

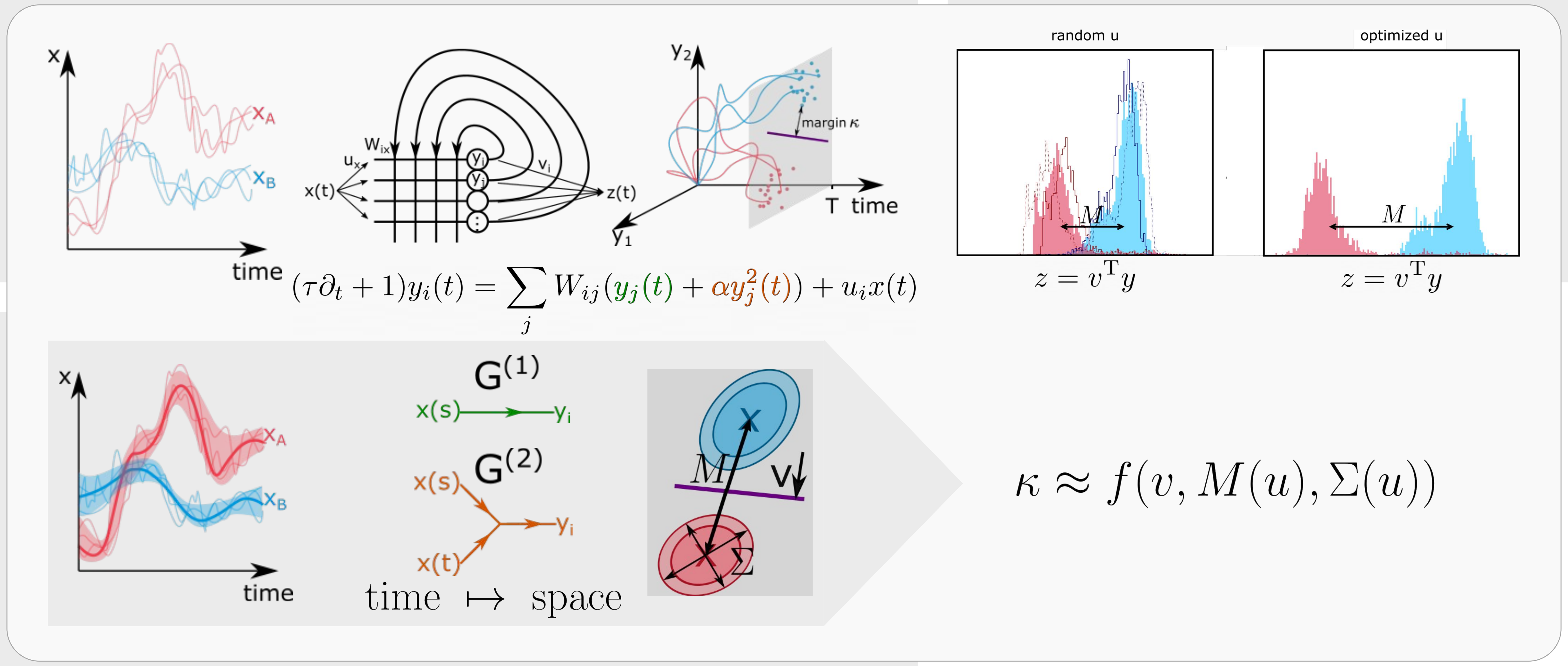
Setup

- Reservoir Computing as computationally efficient machine learning system [1, 2]
- Task: Binary classification of one-dimensional, time-dependent stimuli
- Dynamics governed by random recurrent reservoir with connectivity W and transfer functions $\varphi(y) \in \{y, y + \alpha y^2\}$
- Stimulation via input projection u and classification via hyperplane with readout vector v
- Dependence of the performance on reservoir properties has already been studied [3, 4]

Objective

- Joint optimization of input and readout projections
- Classification quality measure: margin $\kappa(u,v) = \min_\nu (\zeta_\nu v^{\mathrm{T}} y^{u,\nu})$
- Differentiable and less sensitive to exact realizations of stimuli: soft margin $\kappa_\eta(u,v) = -\frac{1}{\eta}\ln\left[\sum_\nu \exp(-\eta\zeta_\nu v^{\mathrm{T}} y^{u,\nu})\right]$
- For large set of sample data: κ_η becomes cumulant generating function
- Gradient can be calculated to desired degree of complexity of the network state distribution using a cumulant expansion

$$\kappa_\eta(u,v) \approx v^{\mathrm{T}} M^u - \frac{1}{2}\eta v^{\mathrm{T}} \Sigma^u v$$



Linear dynamics

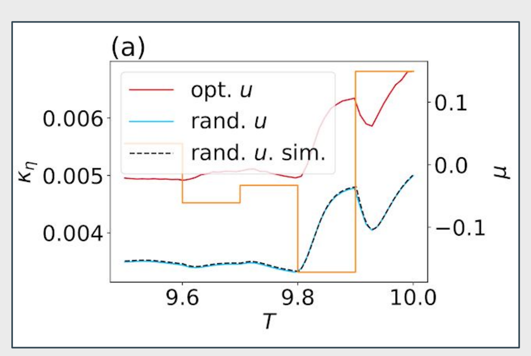
- Linear dynamics: exact solution

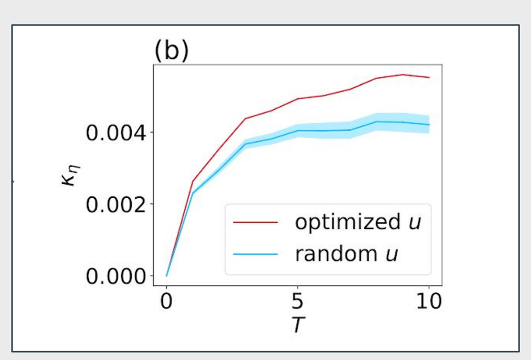
  $y(t) = \sum \int dt'\, G^{(1)}(t-t')\, u\, x(t')$

  $G^{(1)}(t-t') = \frac{1}{\tau}\exp\left[-(1-W)\frac{t-t'}{\tau}\right]$
- Green's function as propagator from stimulus to network space
- Mapping: stimulus statistics → network state statistics

  $M \propto u\langle\zeta_\nu x^\nu\rangle$

  $\Sigma \propto u^2(\langle {x^\nu}^2\rangle - \langle\zeta_\nu x^\nu\rangle^2)$
- Optimization of soft margin: quadratic problem in both u and v
- For fixed reservoir, stimulus and readout time: considerable increase in classification performance
- Optimal input vector composed of modes with various time constants





Simulations carried out using NEST simulator [5]

Non-linear dynamics

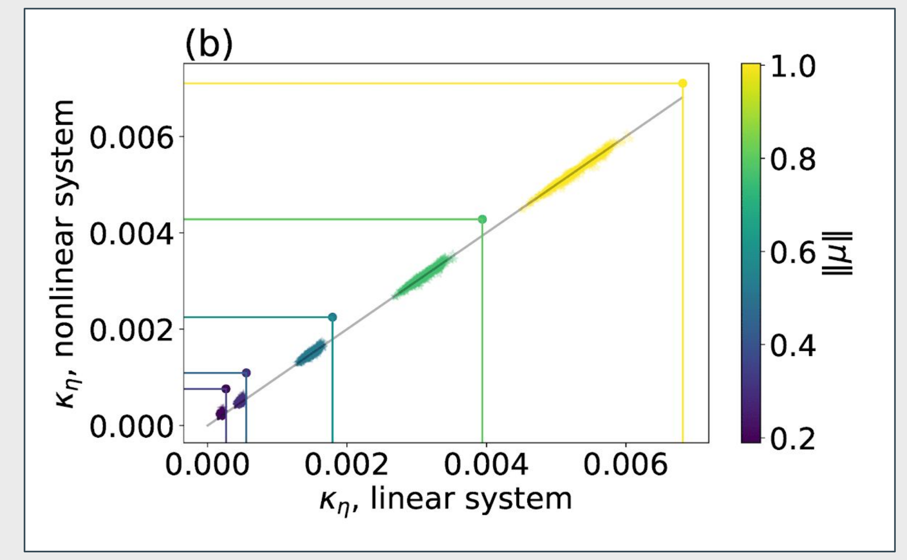
- Non-linear dynamics can be approximated as perturbation series for small α
- Consider only first order correction to linear dynamics: Green's function $G^{(2)}$

  $G^{(2)} = \sum \int G^{(1)}\, \alpha W\, (G^{(1)})^2$
- M becomes sensitive to second order stimulus statistics, Σ becomes sensitive to fourth order statistics

  $M = \sum G^{(1)}\, u\, \langle\zeta_\nu x^\nu\rangle$

  $\quad + \sum G^{(2)}\, u^2(\langle {x^\nu}^2\rangle - \langle\zeta_\nu x^\nu\rangle^2)$

Optimization: Non-linear system



- Maximize closed-form expressions for κ_η
- Clear benefit compared to random u
- Gain from non-linearity varies with linear separability of stimuli
- Significant performance increase for low linear separabilities

Application to ECG5000

- Discriminate between healthy and diseased heartbeats [6]
- Increased separability
  - Strongly increased mean separation M
  - Only moderate increase of fluctuations in readout direction
- Performance increase clearly reflected in both soft margin and test set accuracies

Conclusion

- Unroll recurrent dynamics via Green's functions
- Soft margin yields closed-form expressions for optimization
  - First- and second-order stimulus statistics have strongest influence on performance
  - Effect of higher-order stimulus statistics suppressed by powers of the perturbation parameter α
- Trade-off between separation and variability in readout direction
- Significant gain from non-linearity for weakly linear separable data
- Clear absolute performance gain also in linearly well separable ECG5000 dataset

Analytically unrolling recurrent dynamics into Green's functions is a versatile approach that may be used as a general purpose scheme to analyze recurrent networks.

References

[1] Jäger, H. (2001) The echo state approach to analysing and training recurrent neural networks. Tech. Rep. GMD Report 148

[2] Maass, H., Natschläger, T. (2002) Real-time computing without stable states: A new framework for neural computation based on perturbations. Neural Computation, 14.11, 2531-2560.

[3] Bertschinger N, Natschläger T, Legenstein R. (2004) At the Edge of Chaos: Real-time Computations and Self-Organized Criticality in Recurrent Neural Networks. Conference: Advances in Neural Information Processing Systems 17 NIPS

[4] Toyoizumi T, Abbott L. (2004) Beyond the edge of chaos: Amplification and temporal integration by recurrent networks in the chaotic regime. Phys. Rev. E., 84, 051908

[5] Peyser, Alexander et al. (2017) NEST 2.14.0. Zenodo. 10.5281/zenodo.882971

[6] Chen, Yanping, et al. (2015) The ucr time series classification archive. URL https://www.cs.ucr.edu/%7Eeamonn/time_series_data_2018/