Ghent University

An experimental comparison of Reservoir Computing methods

David Verstraeten

Electronics and Information Systems Department

Ghent University - Belgium

December 9 2006 - NIPS 2006

www.elis.UGent.be/SNN/

Outline

- Different reservoir types and readout functions exist
- All implementations of the same or similar ideas (from the standpoint of Machine Learning)
- How do these reservoir methods compare?
  - → Node complexity
- Nodes memory vs. reservoir memory
- Linking between *a priori* stability bounds and actual reservoir dynamics

UNIVERSITEIT GENT

The benchmarks tests

- Tenth order NARMA system identification
  - input u(t) uniform noise
  - $y(t+1) = .3 * y(t) + .5 * y(t) * \sum^{9} y(t-i) + 1.5u(t-9)u(t) + .1$
- Memory capacity (MC)
  - input u(t) uniform noise
  - y(t): successively delayed versions of u(t)
  - ‘Performance’ : MC = Σ (determination coefficients)
- Speech
  - 500 samples, 5 female speakers
  - Trained and tested using ten-fold cross-validation
  - Performance expressed as Word Error Rate (WER)

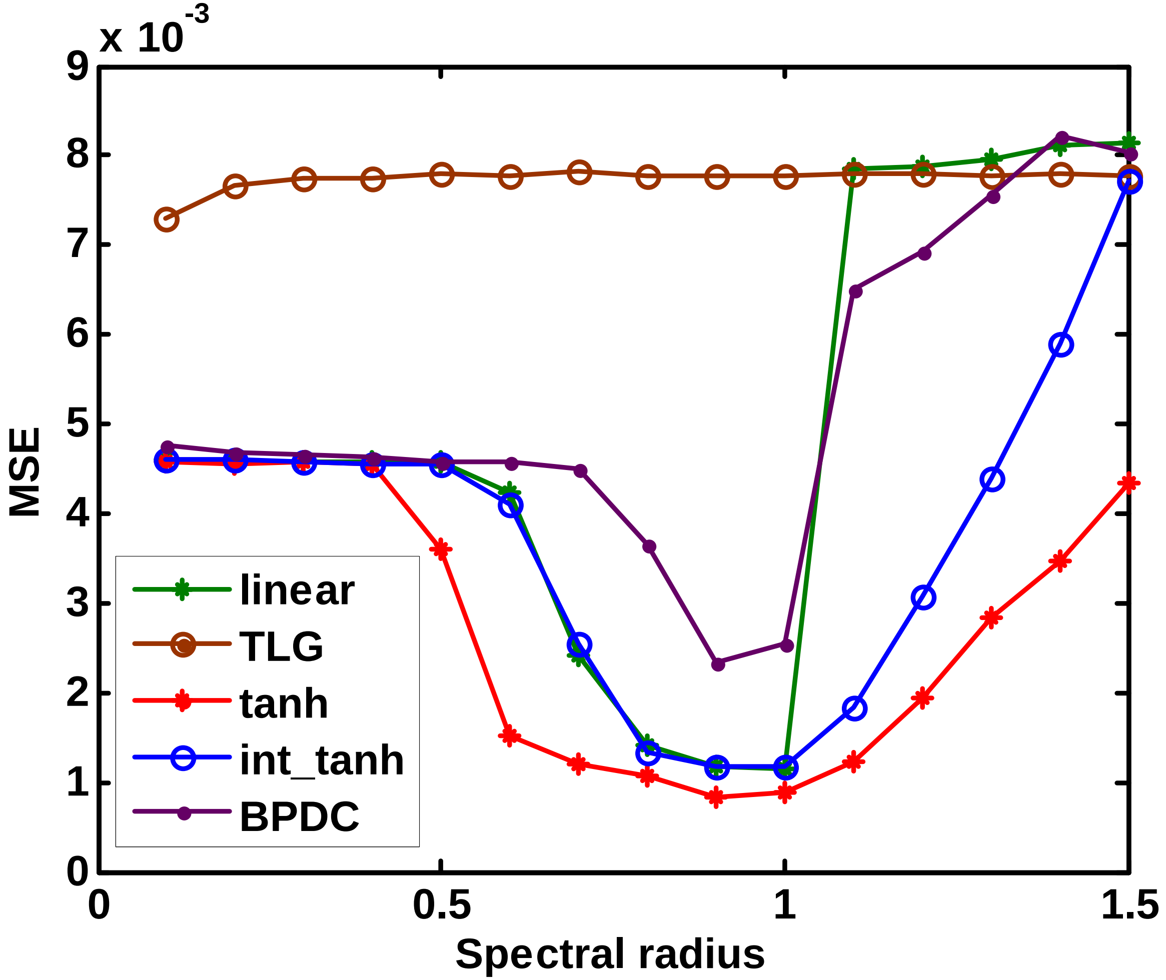
UNIVERSITEIT GENT

Node complexity

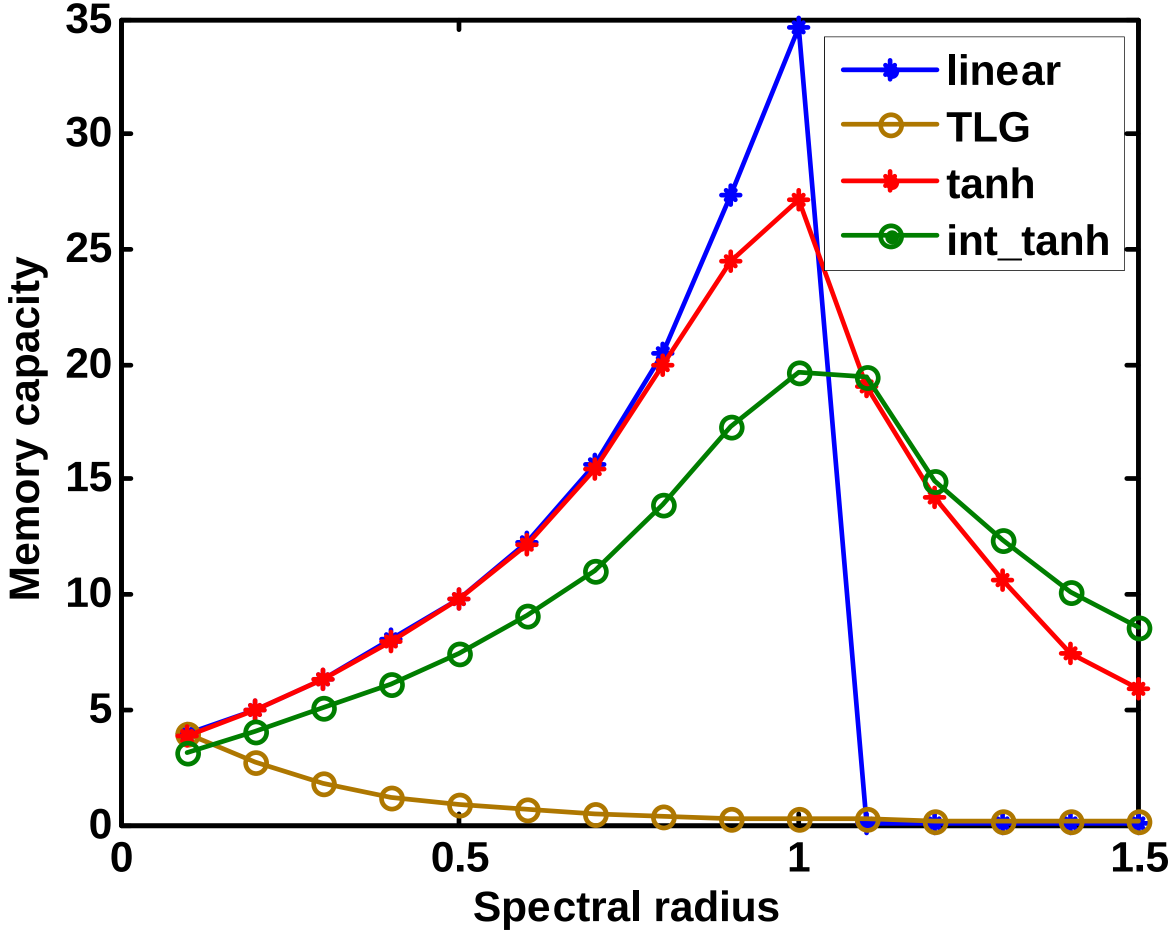
- Reservoir types
  - Analog neurons: linear, TLG, tanh, analog integrator tanh
    - Note: we put the integration timeconstant *inside* the nonlinearity
  - Spiking neurons: Leaky Integrate and Fire, Booij (LIF with synapse model)
- How do these reservoir types compare to each other?
- How does performance scale with the spectral radius?
- How does performance scale with reservoir size?

  → see paper

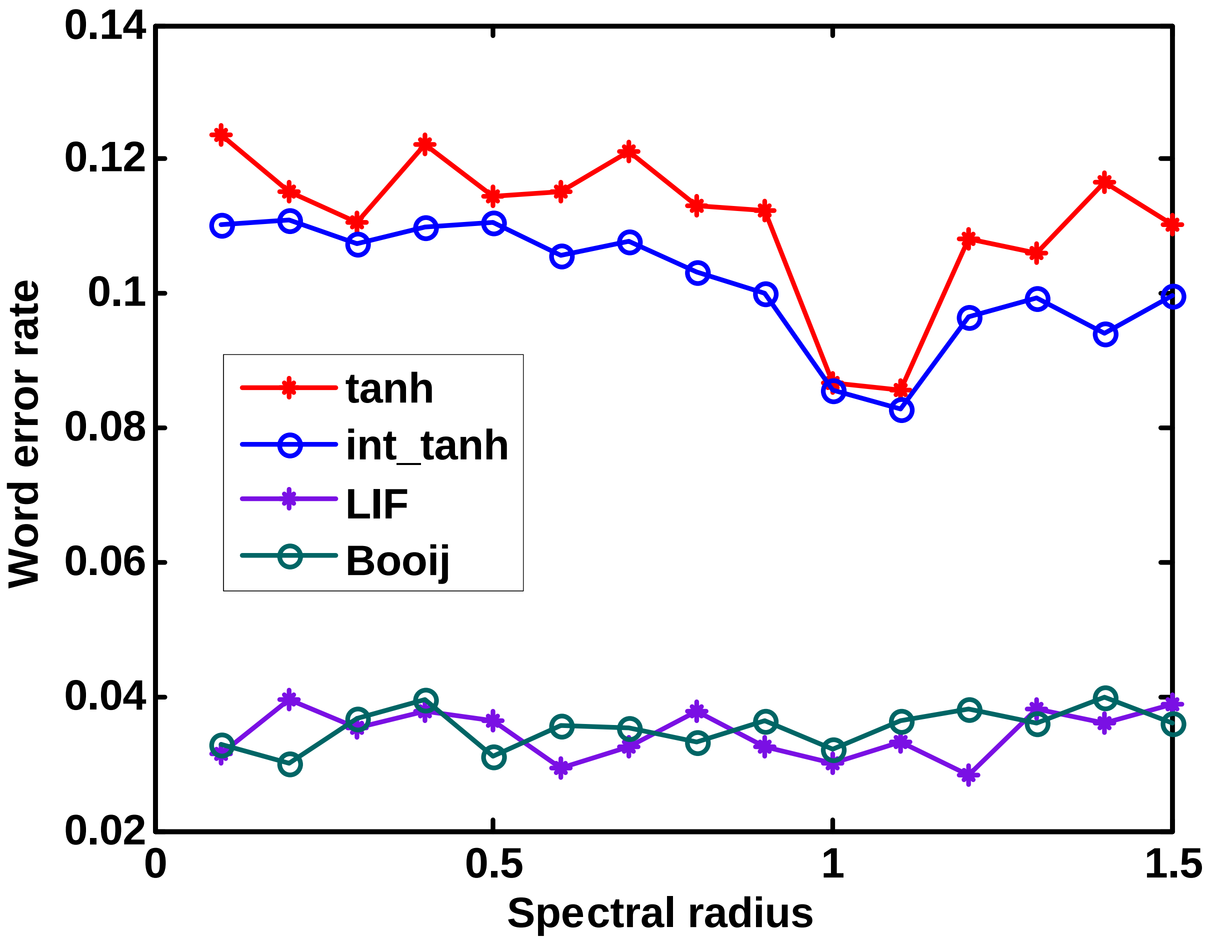
NARMA:
Node complexity vs. spectral radius



Memory capacity: Node complexity vs. spectral radius



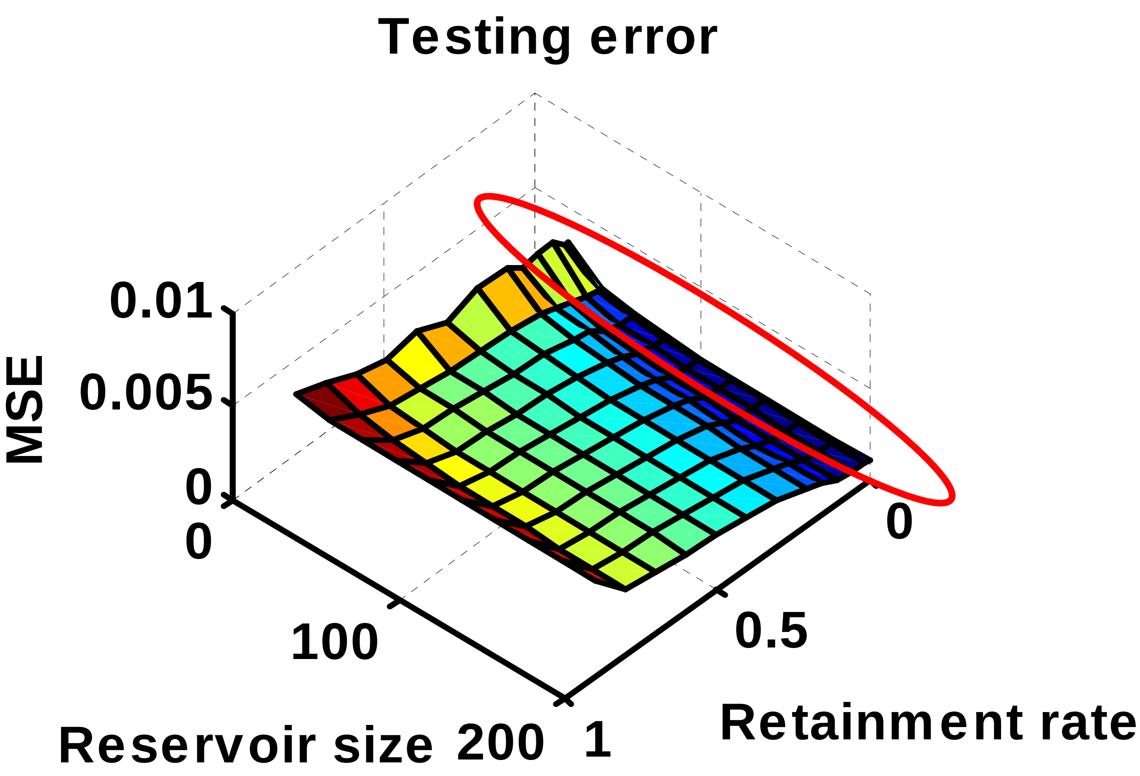
Speech recognition:
Node complexity vs. spectral radius

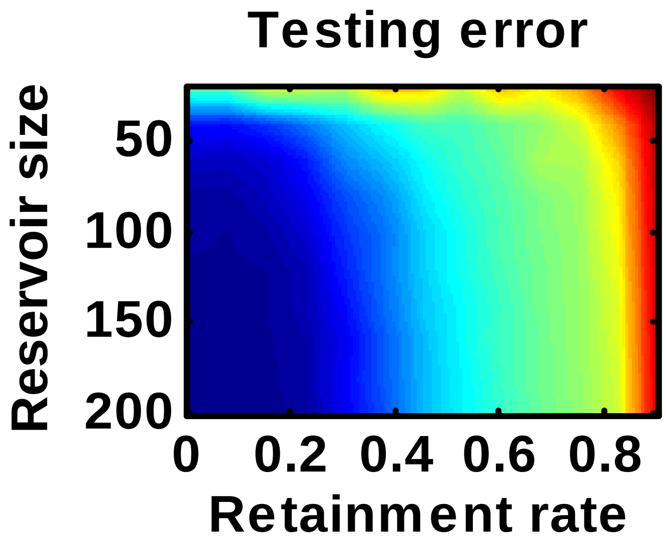


Node memory vs.reservoir memory

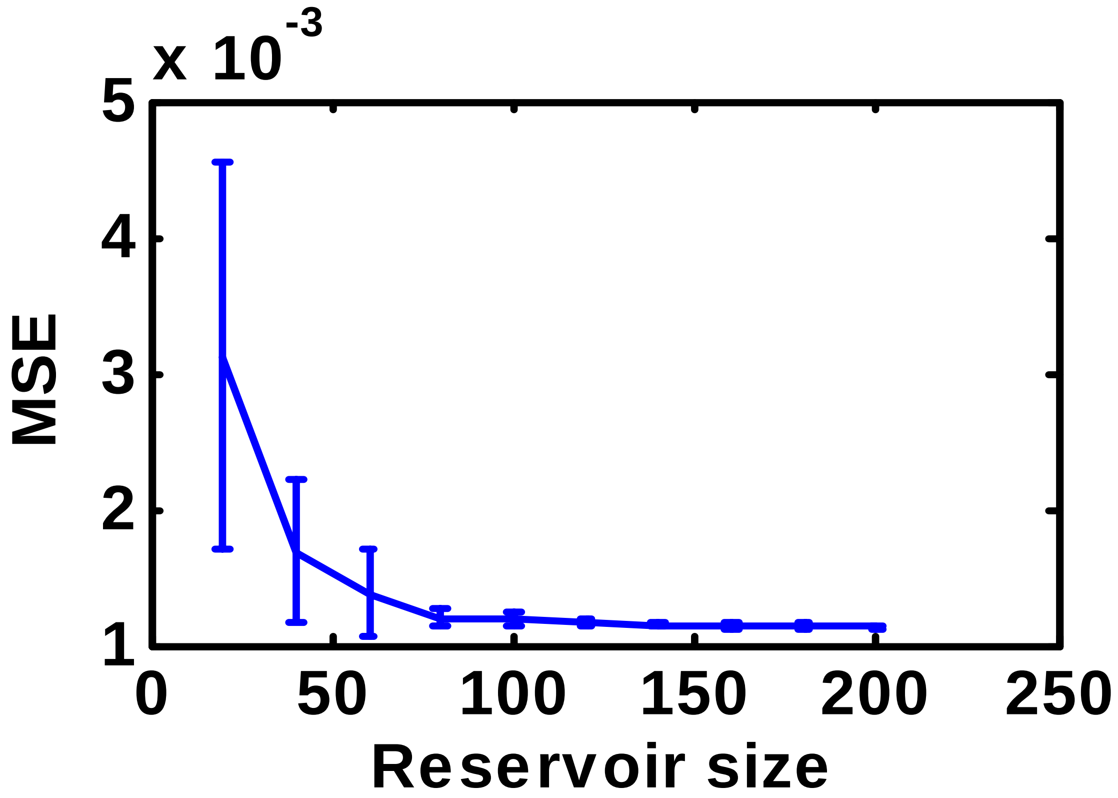
- Some nodes have internal memory (temporal integration):
  - Analog integrator neurons
  - Leaky Integrate and Fire (with or without synapses)
- We focus on analog integrator neurons:
  - Only one parameter controls node memory
  - Expressed as retainment rate r: r=0 ➔ simple tanh nodes
  - In our case, the integration is *inside* the nonlinearity: more chance of stable reservoirs

NARMA:
Reservoir size vs. retainment rate

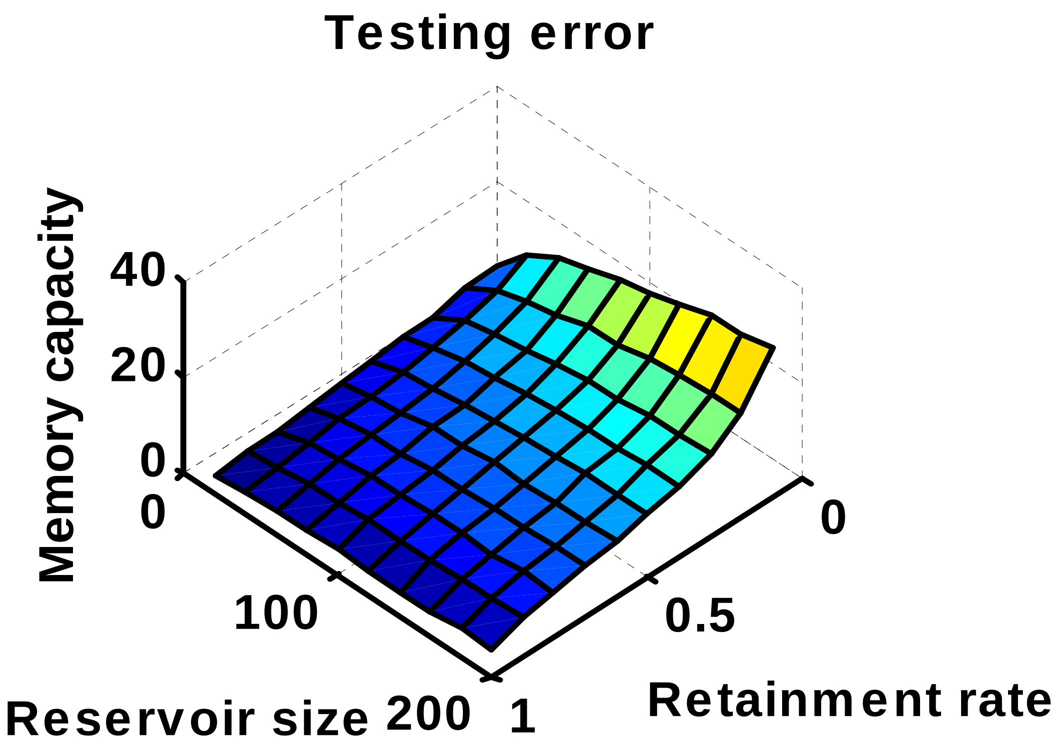


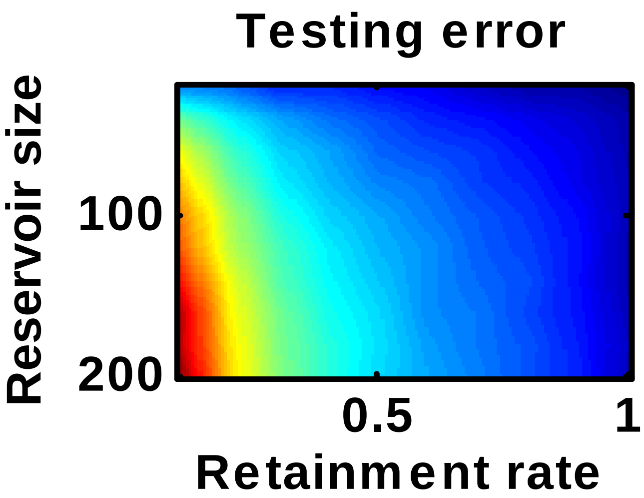


NARMA:
Reservoir size vs. retainment rate

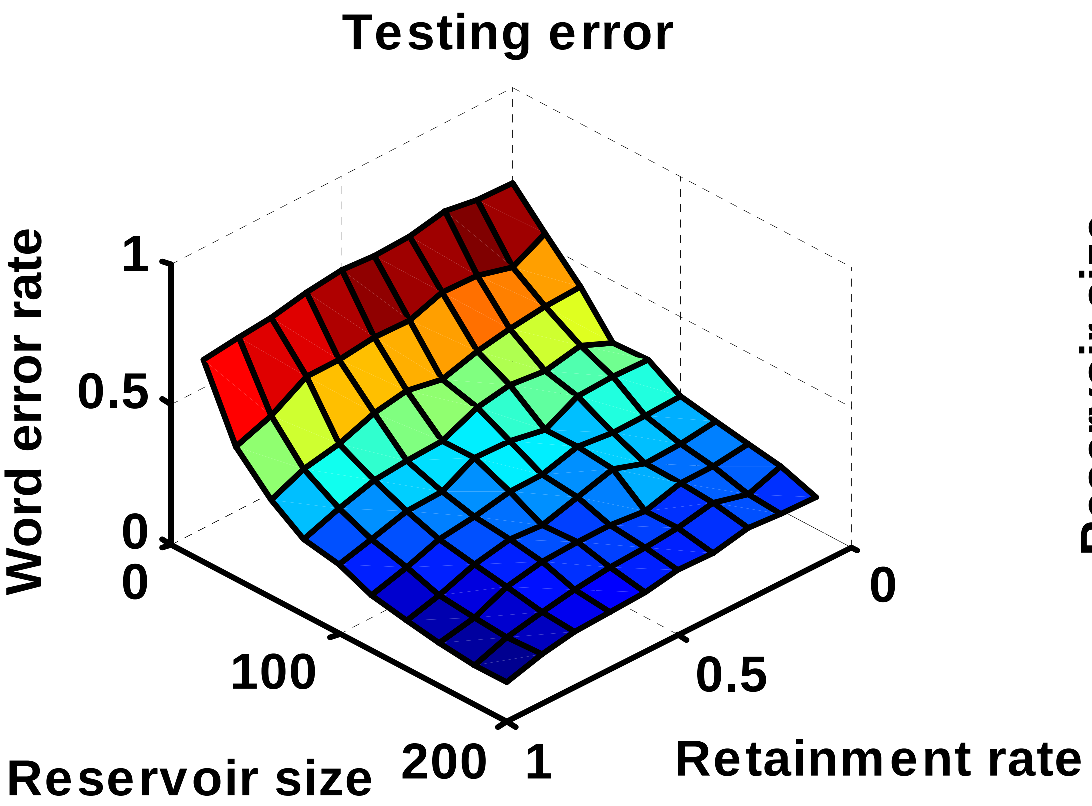


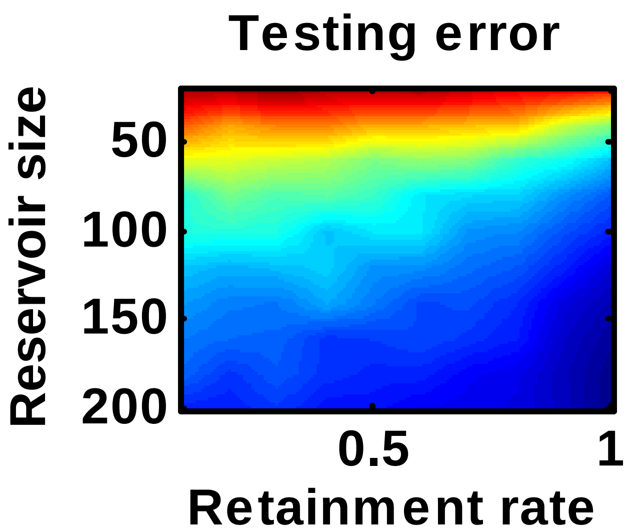
Memory capacity: Reservoir size vs. retainment rate





Speech recognition: Reservoir size vs. retainment rate



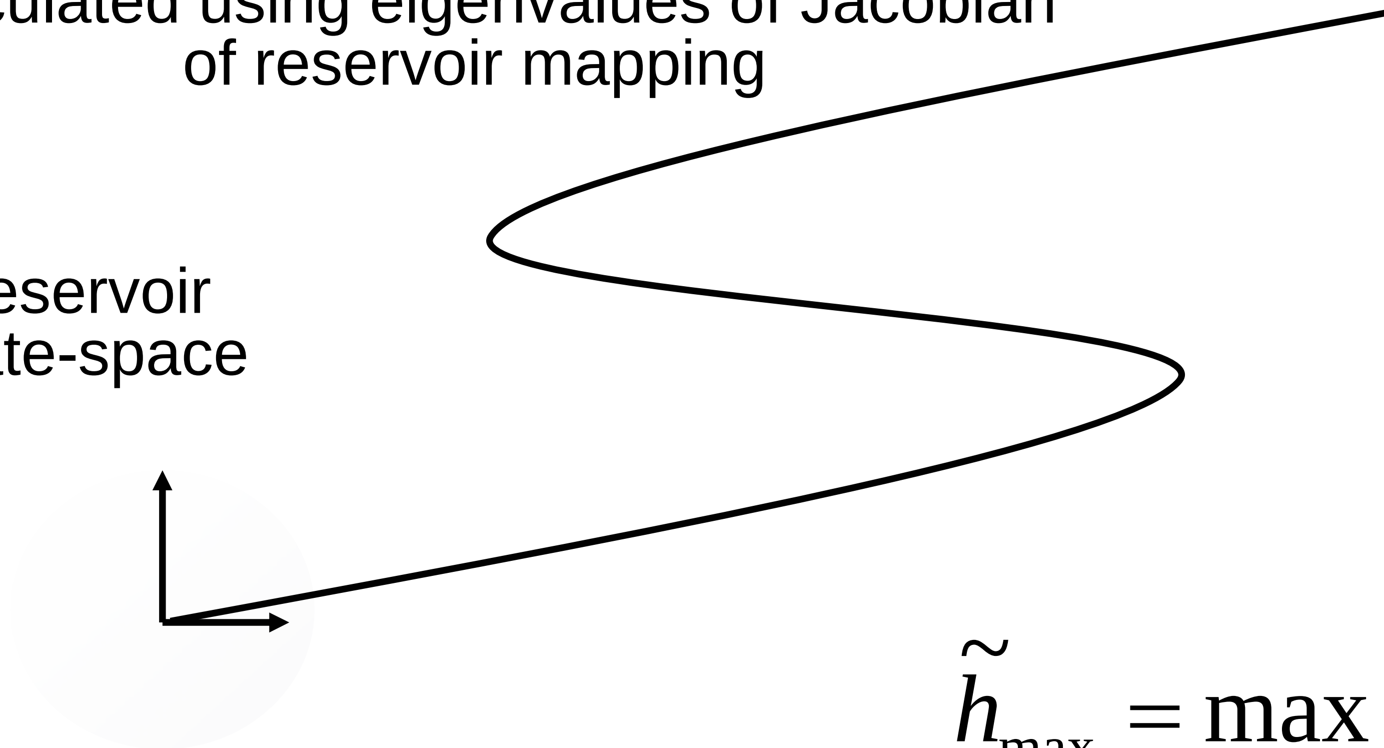


Characterising reservoir dynamics

- Dynamics are of key importance for reservoirs
- Only *a priori* stability bounds based on weight matrix are described:
  - Spectral radius
  - Largest singular value
  - μ_{ssv} : structured singular value (Buehner et al. 2006)
- Actual reservoir dynamics are combination of system properties (weight matrix) and input signals
- How do we characterize the task-dependent dynamics?

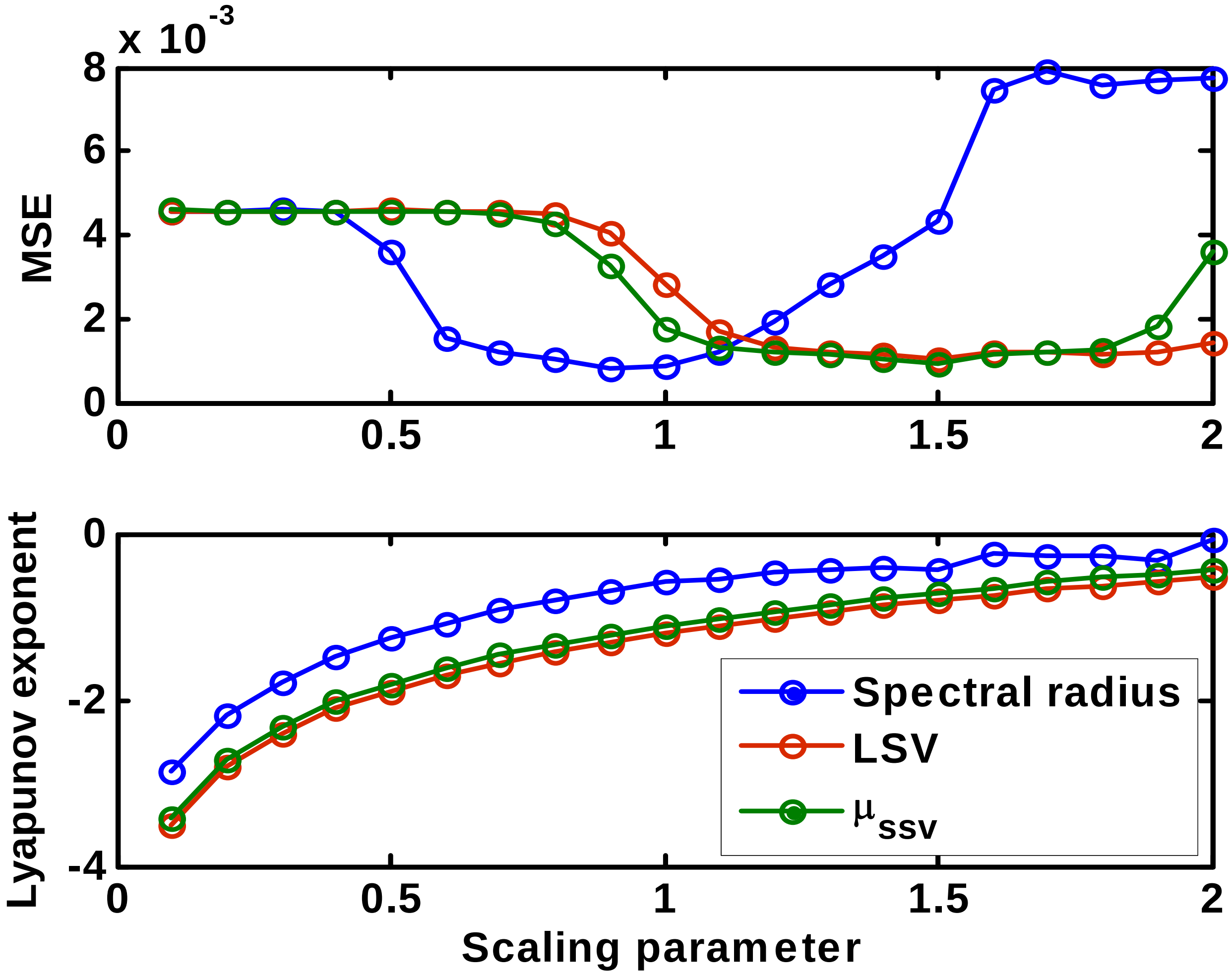
Characterising reservoir dynamics: Local lyapunov exponents

Length of hyperellipsoid axes can be calculated using eigenvalues of Jacobian of reservoir mapping



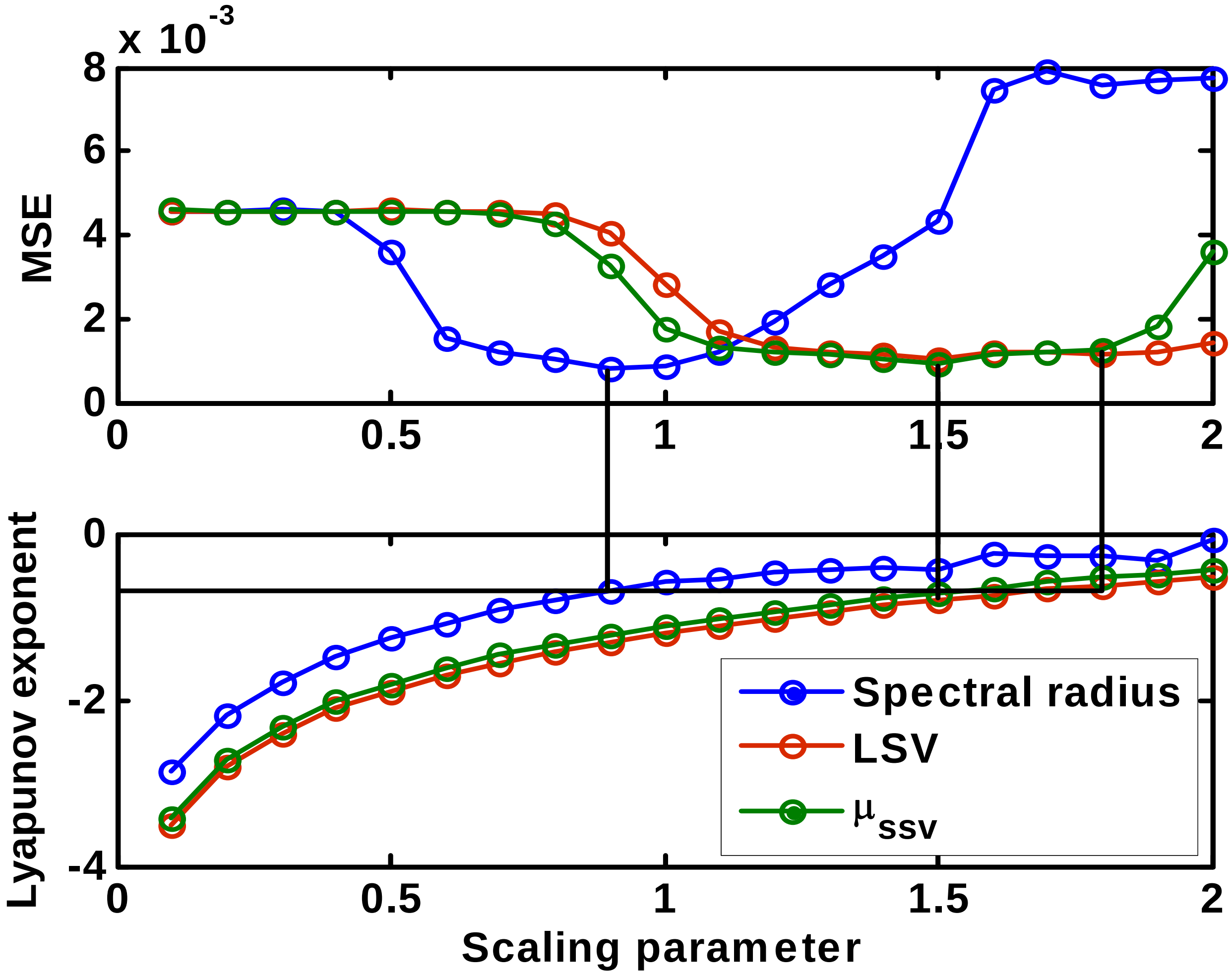
$$\tilde{h}_{\max} = \max_k \prod_{n=1}^{N} \sqrt[N]{r_k}$$

NARMA:
Stability bounds vs. local lyapunov exponent

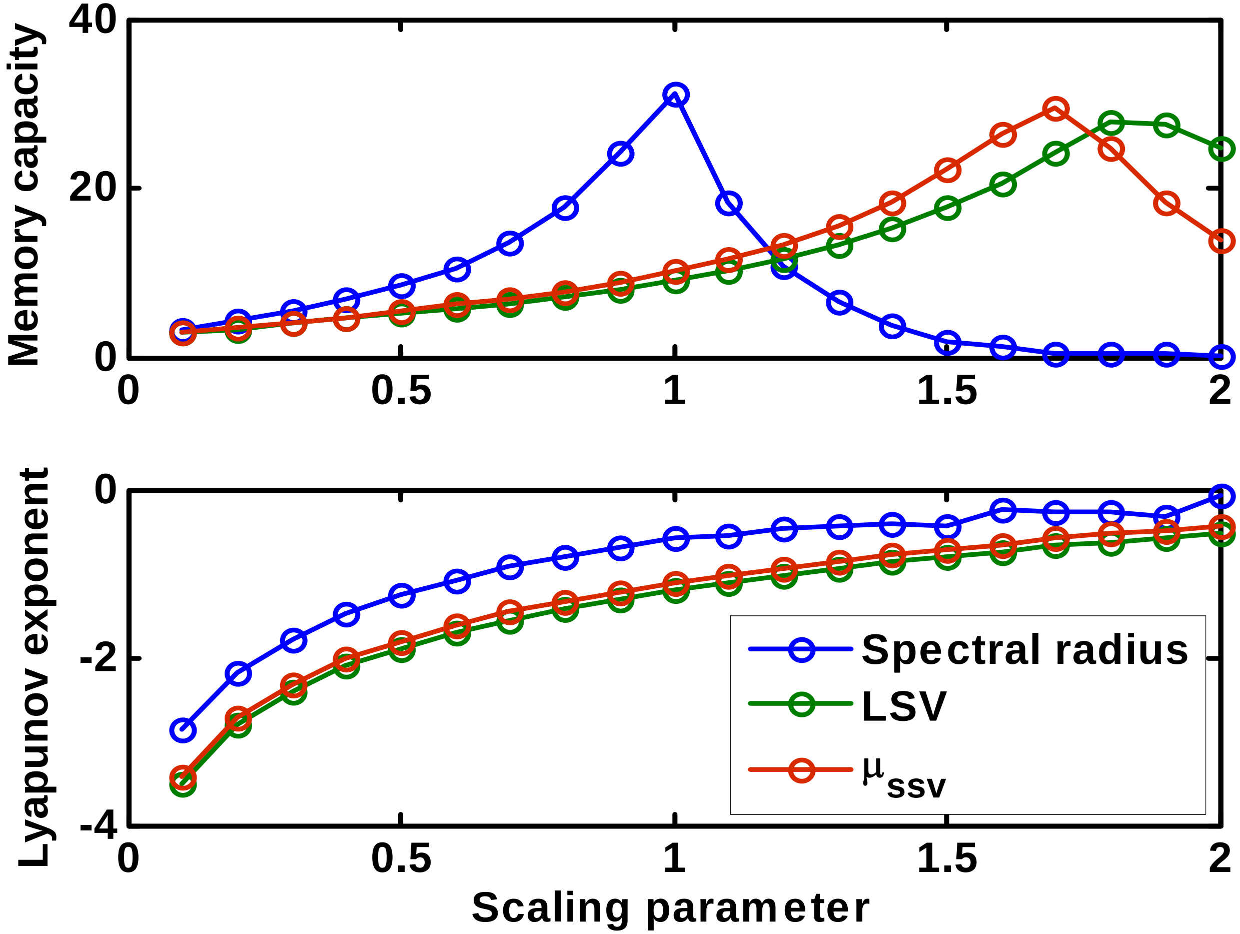


UNIVERSITEIT GENT

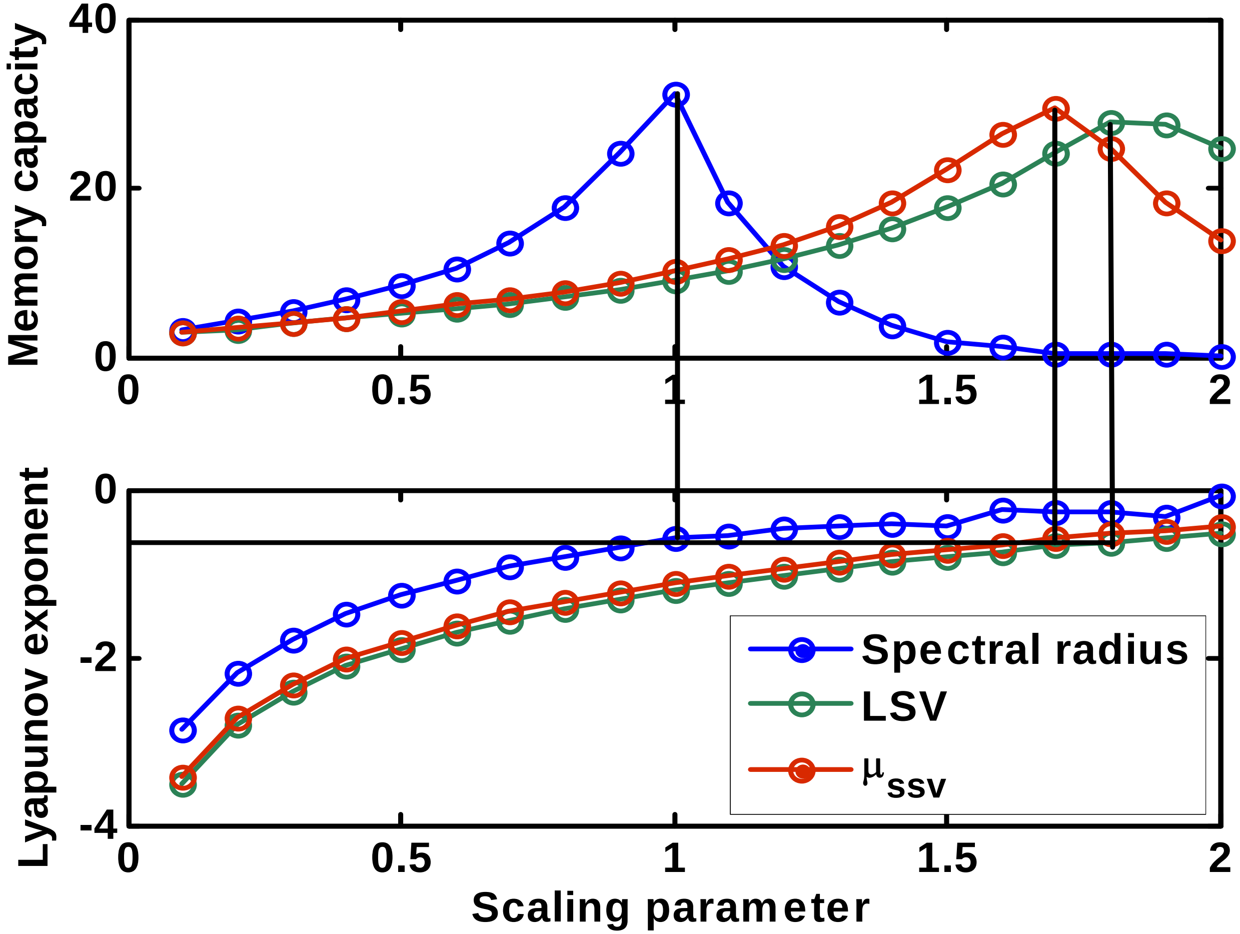
NARMA:
Stability bounds vs. local lyapunov exponent



Memory capacity:
Stability bounds vs. local lyapunov exponent



Memory capacity:
Stability bounds vs. local lyapunov exponent



Conclusions

- We have presented an overview of different reservoir types and compared their performance on three benchmark tests with very different temporal and computational requirements.
- Node complexity allows a trade-off between computational requirements and computational power.
- Different temporal characteristics of the problem can be captured using reservoir and node memory
- Local lyapunov exponents offer a (tentative) measure of reservoir dynamics and are a good (and quick) predictor for performance.