

# レザバーコンピューティングを用いた 適応的な参照軌道生成による外乱にロバストな制御

○河合 祐司 (阪大) 熱田 洋史 (阪大)  
浅田 稔 (大阪国際工科専門職大学・阪大・中部大・情報通信研究機構)

本研究は、リカレントニューラルネットワークの一種であるレザバーコンピューティングを用いて、外乱に対して小さなトルクで応答するロボット制御法を提案する。レザバーコンピューティングは現在の関節角度状態に対して望ましい補正を生成するように学習し、比例・微分 (proportional-derivative: PD) 制御器の参照軌道を生成する。それにより、外乱によりロボットの状態が突発的に変化しても、参照軌道が適応的に変化することで生成トルクが抑えられる。2リンクアームシミュレータにより提案手法を評価し、インパルス状の外乱トルクや観測ノイズが与えられた場合でも、提案手法は少ないトルクでロバストにターゲット姿勢に到達できることを示す。

## 1. はじめに

予期せぬ未知の外乱に最小限の力で柔らかく応答することは、人間を含む未知環境ではたらくロボットの安全性や省エネ性の点において重要である。PD 制御器を有するロボットの状態が大きな外乱により突発的に変化した場合、参照軌道との誤差が大きくなるため、特に微分制御の作用によりロボットが非常に大きな力を生成する。これにより、周囲の人に危険が及び、ロボット自体の損傷につながる可能性があり、また、エネルギー効率も悪い。

そこで、外乱により変化したロボットの状態に対して、参照軌道が適応的に変化することで、参照軌道との誤差の大きな変化を避けるアプローチを考える。再計画による新しい参照軌道の生成には計算コストや時間がかかる。そのため、参照軌道を時間のみに依存させるのではなく、ロボットの状態に依存させるアプローチが提案されている。Seto and Sugihara [1] は、ロボットの手先位置と最終的に到達すべきターゲット位置との内分点を参照点とする reference shaping を提案した。これにより、予期せぬ外乱によりロボットの手先位置が変化しても、その参照点との誤差の変化が比較的緩やかになるため、力の生成が抑えられる。しかし、ターゲットに対する応答性も緩やかになるという問題がある。Khansari-Zadeh and Billard [2] は、与えられた手先軌道のデータセットからターゲットに漸近収束する速度ベクトル場 (力学系) をガウス混合モデルを用いて学習する手法を提案した。この力学系では、学習した軌道以外の任意の初期値からの軌道もターゲットに収束する安定性が保証されている。したがって、外乱により手先位置が変化したとしても、この力学系に沿った参照軌道の更新により、その状態から緩やかにターゲットに収束でき、過度に大きな力の生成を避けられる。しかし、この手法はロボットの課題空間 (手先の位置・速度) で力学系を構成しており、その空間上の参照軌道を状態空間 (関節角度) に写す運動学モデルが必要になる。また、ガウス混合モデルの学習コストは高く、より軽量な手法が望まれる。

レザバーコンピューティングは学習コストが非常に低い時系列学習法として注目を集めている [3, 4]。一般に重み固定のランダムリカレントニューラルネットワー

クをレザバー層とし、各素子の活動の線形和をリードアウト出力とする。出力とターゲット時系列の誤差を最小にするように、最小二乗法を用いてその線形和の重みを決定する。レザバー層への入力をロボットの状態とし、リードアウト出力を運動指令とすることで、ロボットの運動軌道を学習するシステムが提案されている [5, 6]。レザバーコンピューティングの汎化能力により、未学習の初期状態からでも学習した運動軌道を実現できる。しかし、これらの研究では、外乱に対する過剰な力生成を避ける仕組みを有していない。

本研究では、レザバーコンピューティングが 1 例の関節角度軌道を学習することで、ロボットの任意の状態 (関節角度) から参照軌道を生成し、予期せぬ外乱に対して参照軌道を適応的に変化させ、過剰に大きな力の生成を避けるフィードバック制御系を提案する。レザバーコンピューティングの出力と現在のロボットの状態の和を PD 制御器の参照軌道とし、その参照軌道がターゲット関節角度系列になるようにレザバーコンピューティングが学習する。これにより、レザバーは現在のロボットの状態に対する望ましい補正を出力する。PD 制御における参照軌道の誤差を計算する際、参照軌道からロボットの状態を減算すると、その望ましい補正のみが残る。したがって、ロボットの状態に含まれる外乱や観測ノイズの直接的な影響が取り除かれる。また、レザバーコンピューティングの汎化能力により、ロボットが外乱により未学習の状態に変位してもターゲット軌道に復帰できる。本論文では、2リンクアームのロボットシミュレータを用いて、提案システムに外乱や観測ノイズを加えた際の参照軌道、手先軌道、および、生成トルクを評価する。

## 2. 提案手法

### 2.1 適応的な参照軌道生成に基づく PD 制御

図 1 に提案システムのブロック線図を示す。簡単のため、以降では 1 自由度のロボットについて説明する。図中左のレザバー層は 1 時刻前の PD 制御器の参照軌道  $r(t - \Delta t)$  を入力とし、 $o(t)$  を出力する。時刻  $t$  における参照軌道  $r(t)$  は次式で与えられる。

$$r(t) = o(t) + \alpha q(t) \quad (1)$$

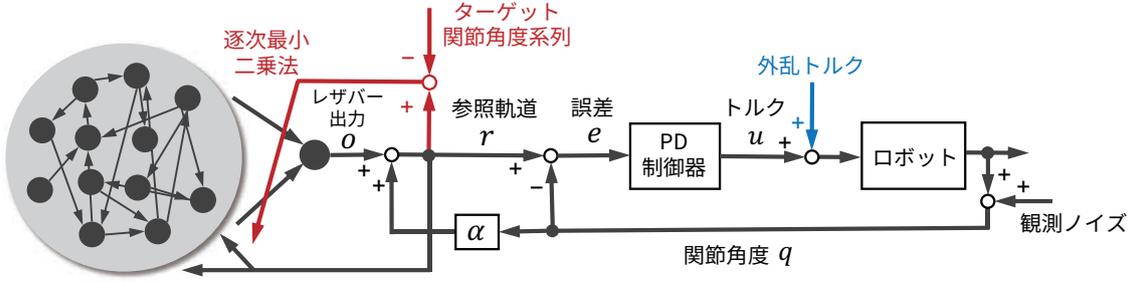


図1 提案システムのブロック線図. 赤い要素は学習時のみ. 青い要素はテスト時のみ.

ここで,  $\alpha$  は  $0 \leq \alpha \leq 1$  の定数である. この  $r(t)$  とターゲット関節角度との誤差が最小になるように逐次最小二乗法によりリードアウト重みを訓練する. そして, PD 制御器が用いる誤差  $e(t)$  は次式で計算される.

$$e(t) = r(t) - q(t) \quad (2)$$

$$= o(t) + (\alpha - 1)q(t) \quad (3)$$

したがって,  $\alpha = 1$  のとき  $e(t) = o(t)$  となり,  $q(t)$  の直接的な影響は取り除かれる. この誤差  $e(t)$  に基づいて PD 制御器がトルク  $u(t)$  を生成し, ロボットを駆動する.

$$u(t) = K_P e(t) + K_D \frac{de(t)}{dt} \quad (4)$$

ここで,  $K_P$  は比例ゲイン定数,  $K_D$  は微分ゲイン定数である.

テスト時には, レザバーコンピューティングの学習を止め (図1の赤い部分を削除), 外乱トルク (図1の青い矢印) を加える.  $\alpha = 0$  の場合, 外乱トルクや観測ノイズの影響を含む  $q(t)$  が  $e(t)$  に直接作用し, PD 制御器が大きなトルクを生成する. 一方,  $\alpha = 1$  の場合,  $q(t)$  の影響が除去されるため, 外乱トルクや観測ノイズの直接的な影響が  $e(t)$  に作用せず, 過剰に大きなトルクの生成が避けられる.

## 2.2 レザバーコンピューティング

レザバー層は  $N$  個の素子のランダム結合のリカレントニューラルネットワークで構成される. レザバー層の素子  $i$  の時刻  $t$  における状態  $x_i(t)$  のダイナミクスは,  $\tanh$  を活性化関数とした発火率モデルで与えられる.

$$\tau \frac{dx_i(t)}{dt} = -x_i(t) + \sum_{j=1}^N W_{ij}^{\text{Rec}} y_j(t) + W_i^{\text{In}} r(t) \quad (5)$$

$$y_i(t) = \tanh(x_i(t)) \quad (6)$$

ここで,  $\tau$  は時定数,  $W_{ij}^{\text{Rec}}$  は素子  $j$  から  $i$  へのリカレント結合重み,  $W_i^{\text{In}}$  は素子  $i$  の入力重みである.  $W_{ij}^{\text{Rec}}$  は平均 0, 分散  $g_{\text{Rec}}/\sqrt{pN}$  とした正規分布からサンプリングされ, ここで  $g_{\text{Rec}}$  はリカレント重みのゲイン,  $p$  は結合確率である.  $W_i^{\text{In}}$  は平均 0, 分散  $g_{\text{In}}$  とした正規分布からサンプリングされる. 今回,  $g_{\text{Rec}} > 1$  とするが, 自身の出力  $o(t)$  を含む参照軌道  $r(t)$  をレザバー層に入力しているため, ネットワーク活動がカオス状態にならず, FORCE 学習が実現される [7].

各素子の活動の重み付き和によりリードアウト出力  $o(t)$  を得る.

$$o(t) = \sum_{i=1}^N W_i^{\text{Out}}(t) y_i(t) \quad (7)$$

ここで,  $W_i^{\text{Out}}(t)$  はリードアウト重みであり, 初期値 0 から逐次最小二乗法により調整される. 時刻  $t$  において,  $W_i^{\text{Out}}(t)$  は  $r(t)$  とターゲット時系列  $d(t)$  の差  $e_d(t)$  を用いて次式で更新される.

$$W_i^{\text{Out}}(t) = W_i^{\text{Out}}(t - \Delta t) - e_d(t) \sum_{j=1}^N P_{ij}(t) y_j(t) \quad (8)$$

$$e_d(t) = r(t) - d(t) \quad (9)$$

$$= \sum_{i=1}^N W_i^{\text{Out}}(t - \Delta t) y_i(t) + \alpha q - d(t) \quad (10)$$

ここで,  $P_{ij}(t)$  ( $N \times N$  の行列  $\mathbf{P}(t)$ ) は次式で更新される.

$$P_{ij}(t) = P_{ij}(t - \Delta t) - \frac{\sum_k \sum_l P_{ik}(t - \Delta t) y_k(t) y_l(t) P_{lj}(t - \Delta t)}{1 + \sum_k \sum_l y_k P_{kl}(t - \Delta t) y_l(t)} \quad (11)$$

$\mathbf{P}(t)$  の初期値は  $(1/\beta)\mathbf{I}$  で与えられ, ここで  $\mathbf{I}$  は単位行列,  $\beta$  は定数である.

## 2.3 ロボットモデル

今回, 次式で与えられる 2 リンクの剛体ロボットアームを用いる (時間を表す  $t$  は省略). ただし, リンク 1 とリンク 2 は同じパラメータを有し, 重心位置はリンクの中心にあり, 重力は考慮しない.

$$\mathbf{u} = \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{F}(\mathbf{q}, \dot{\mathbf{q}}) \quad (12)$$

$$\mathbf{M}(\mathbf{q}) = \begin{bmatrix} 3ml^2 + 2ml^2 \cos q_2 & ml^2(1 + \cos q_2) \\ ml^2(1 + \cos q_2) & ml^2 \end{bmatrix} \quad (13)$$

$$\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} -ml^2(2\dot{q}_1 \dot{q}_2 + \dot{q}_2^2) \sin q_2 \\ ml^2 \dot{q}_1^2 \sin q_1 \end{bmatrix} \quad (14)$$

$$\mathbf{F}(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} b\dot{q}_1 \\ b\dot{q}_2 \end{bmatrix} \quad (15)$$

ここで,  $\mathbf{M}(\mathbf{q})$  は慣性行列,  $\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})$  は遠心力・コリオリ力項,  $\mathbf{F}(\mathbf{q}, \dot{\mathbf{q}})$  は摩擦項,  $m$  はリンクの質量,  $l$  はリンクの長さ,  $b$  は摩擦係数である.

パラメータ	説明	値
$K_P$	比例ゲイン	0.04
$K_D$	微分ゲイン	2.5
$N$	レザバー層の素子数	1000
$g_{Rec}$	リカレント重みの係数	1.5
$g_{In}$	入力重みの係数	3.0
$p$	結合確率	0.1
$\tau$	時定数	10 ms
$\beta$	逐次最小二乗法の初期値	10
$m$	リンクの質量	0.5 kg
$l$	リンクの長さ	1.8 m
$b$	摩擦係数	0.5

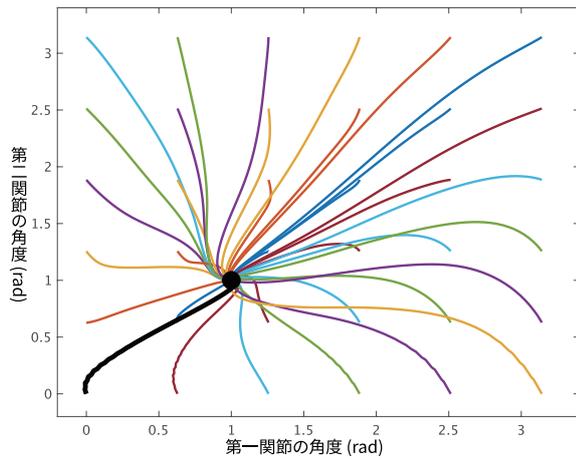


図2 初期関節角度からの軌道. 黒い円はターゲット. 黒い軌道は学習軌道であり, それ以外の軌道は未学習の初期関節角度からのテスト軌道.

### 3. 実験

#### 3.1 実験設定

実験において, 表1のパラメータを用い, シミュレーションのステップサイズを  $\Delta t = 1$  ms としてモデルの数値解を得た. シミュレーションの開始時刻を  $-250$  ms とし, 終了時刻を  $1000$  ms とした. 学習時, 初期関節角度は  $(0, 0)$  rad, ターゲット関節角度系列は  $t \leq 0$  で  $(0, 0)$  rad,  $t > 0$  で  $(1, 1)$  rad とするステップ状時系列とした. その初期関節角度から10回試行し, その間, 式(8)を用いてリードアウト重みを調節した.

#### 3.2 実験結果

外乱と観測ノイズのない場合のテスト軌道を図2に示す. 未学習の初期関節角度として, 各関節について0から  $\pi$  rad まで  $0.2\pi$  きざみの6点をとった. いずれの初期関節角度からの試行においても, ロボットが正しくターゲットに到達できていることがわかる. これは, 外乱によってロボットが未学習状態に変位した場合でもターゲットへの参照軌道を生成できることを示す.

次に, 観測ノイズのない場合, 初期関節角度を  $(0, 0)$  rad とし, 時刻  $0.5$  s で逆方向のインパルス状の外乱ト

ルク (各関節に大きさ  $0.5$  Nm) を加えたテスト試行の第二関節の  $r(t)$ ,  $u(t)$ ,  $q(t)$  を図3に示す.  $\alpha = 0$  では, 外乱にかかわらず  $r(t)$  が一定であり, 外乱に対して大きな  $u(t)$  (ピークが  $0.13$  Nm) が生成された. 一方,  $\alpha = 1$  では, 外乱に対して  $r(t)$  が変化し, 小さな  $u(t)$  (ピークが  $0.011$  Nm) となった.  $t = 0$  ms からのターゲットに対して  $\alpha$  の違いにより  $q(t)$  は変化しなかったが,  $t = 500$  ms からの外乱に対しては,  $\alpha = 1$  のとき  $\alpha = 0$  のときよりターゲットへの復帰が遅れた.

最後に, 外乱のない場合, 初期関節角度を  $(0, 0)$  rad とし, 学習時とテスト時に時刻  $1$  ms 以降に平均  $0$ , 分散  $0.1$  のガウス観測ノイズを与えたテスト試行の第二関節の  $r(t)$ ,  $u(t)$ , 実際の関節角度を図4に示す.  $\alpha = 1$  の場合に  $r(t)$  がノイジーになり, これは  $r(t)$  が, 観測ノイズを含む  $q(t)$  を陽に含むからである. しかし, PD制御に用いる誤差を計算する際に  $q(t)$  が減算され, ノイズの直接的な影響は除去される. ノイズを含む  $r(t)$  はレザバー層に入力されるが, レザバーコンピューティングがノイズに頑健であることで, 望ましいリードアウト出力が得られた. したがって,  $\alpha = 1$  の場合の  $u(t)$  と実際の関節角度において, ノイズの影響が低減された.

### 4. 結論

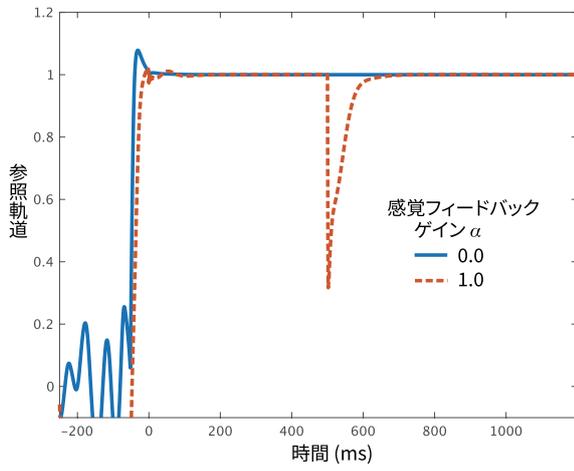
本研究では, レザバーコンピューティングを用いて, 現在のロボットの状態に応じた参照軌道を生成することで, 外乱や観測ノイズの直接的な影響を除去し, それらに小さな力で応答するフィードバック制御システムを提案した. 2リンクロボットアームで提案手法を評価し, 1例の軌道の学習から大域的に安定なベクトル場を獲得でき, 外乱に対する生成ピークトルクがおよそ  $1/10$  になり, また, 観測ノイズに対してロバストになることを示した.

人間が周囲にいるような開かれた環境ではたらくロボットのニーズはますます増えていくと考えられる. そのようなロボット開発においては, 未知の外乱になじみ (柔らかく応答し), 観測ノイズに頑健である制御系の設計は不可欠であり, 提案手法はその基盤を提供する. また, 提案手法は極めてシンプルであり, モデル同定や制御器設計をほとんど必要としないため, ロボット以外の制御系を含めて幅広い応用が期待される. 今後, より複雑な軌道への適用とロボット実装を通して, 様々な状況における適用と評価を行う.

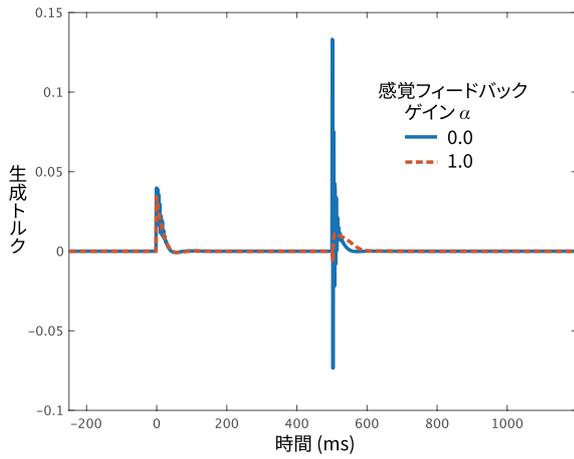
謝辞 本研究は, JST, CREST (JPMJCR17A4) の支援を受けたものであり, 国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務 (JPNP16007) の結果得られたものである.

### 参考文献

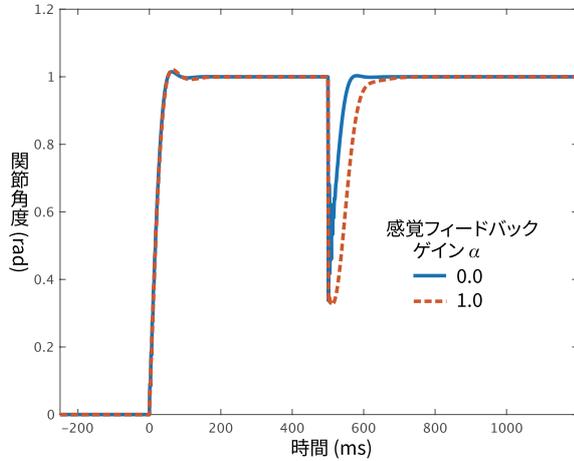
- [1] F. Seto and T. Sugihara: "Nonlinear reference shaping with endpoint position feedback for large acceleration avoidance in reaching movement," *Journal of Robotics and Mechatronics*, vol. 22, no. 2, pp. 173–178, 2010.
- [2] S.M. Khansari-Zadeh and A. Billard: "Learning stable nonlinear dynamical systems with gaussian mixture models," *IEEE Transactions on Robotics*, vol. 27, no. 5, pp. 943–957, 2011.
- [3] H. Jaeger: "The "echo state" approach to analysing and training recurrent neural networks-with an erratum note," GMD Technical Report 148, 2001.



(a) 参照軌道  $r$

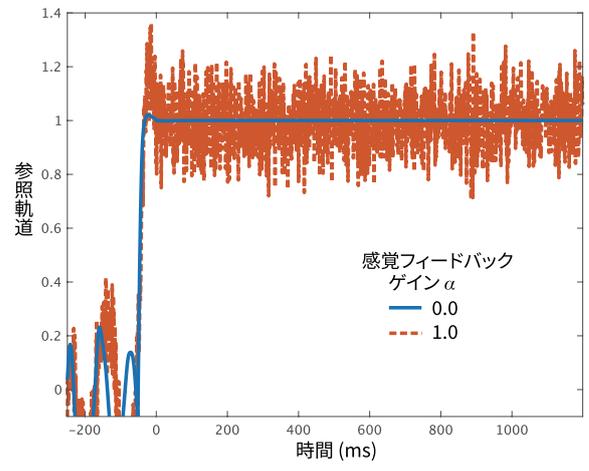


(b) 生成トルク  $u$

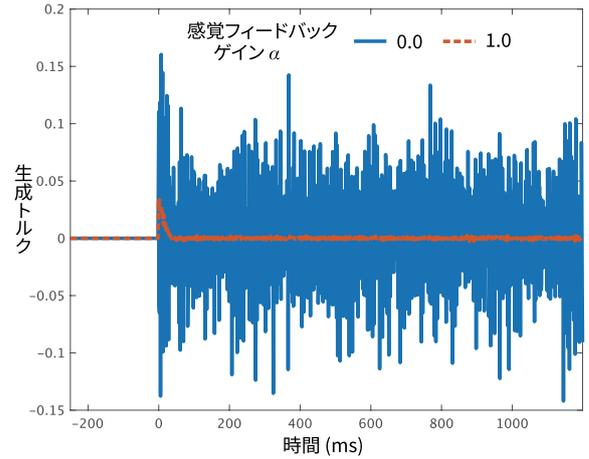


(c) 関節角度  $q$

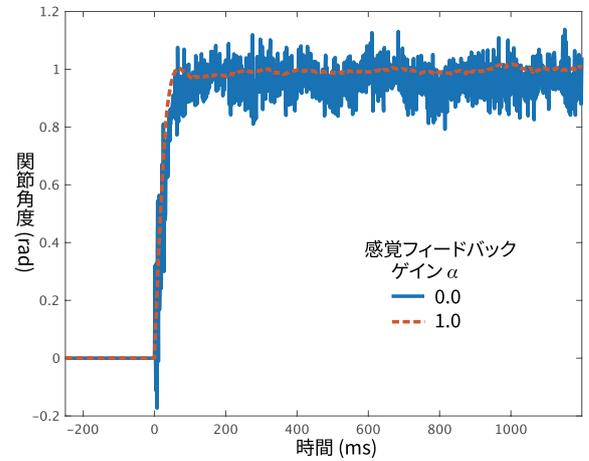
図3 インパルス状外乱応答



(a) 参照軌道  $r$



(b) 生成トルク  $u$



(c) 関節角度

図4 観測ノイズ応答

- [4] W. Maass, T. Natschläger and H. Markram: "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Computation*, vol. 14, no. 11, pp. 2351–2560, 2002.
- [5] T. Waegeman: "Adaptive Control of Compliant Robots with Reservoir Computing," Ghent University, 2013.
- [6] E.A. Antonelo and B. Schrauwen: "On learning navi-

- gation behaviors for small mobile robots with reservoir computing architectures," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 4, pp. 763–780, 2014.
- [7] D. Sussillo and L.F. Abbott: "Generating coherent patterns of activity from chaotic neural networks," *Neuron*, vol. 63, no. 4, pp. 544–557, 2009.