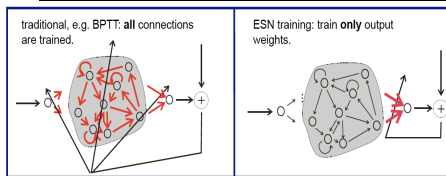


I. Reservoir Computingとは

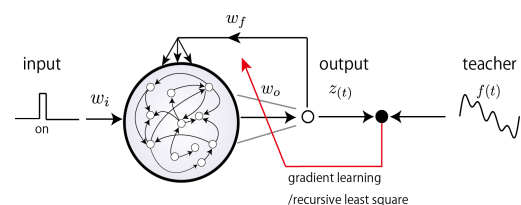
Reservoir Computing (Maas 2002, Jaeger 2003)



- 従来: 全ての結合を誤差逆伝播などで学習
Back Propagation Through Time (BPTT)法
Real Time Recurrent Learning (RTRL)法
- ESN: 入力信号をいったん高次元空間に写し、読み出し部分だけ線形モデルで推定

(Jaeger, 2009)

FORCE-learning: Sussillo and Abbott (2009)



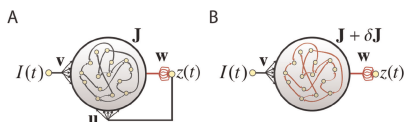
- オンラインの教師あり学習
- 勾配学習や逐次的最小二乗法など
- readout neuronからのフィードバック → 安定なダイナミクス

$$\tau \dot{x}_i = -x_i + g \sum_{j=1}^J w_{rec}^{ij} \tanh(x_j) + a_i \sum_{k=1}^K w_{in}^{ik} u_k + a_f \sum_{l=1}^L w_{fb}^{il} z_l$$

逐次的最小二乗法で推定

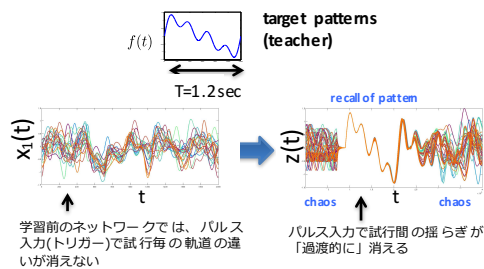
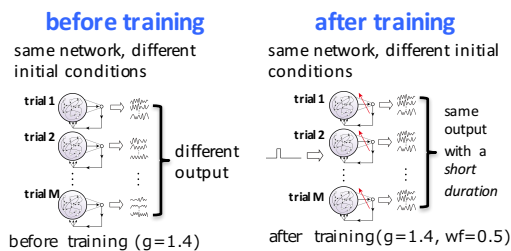
$$z_l(t) = {}^t w_{out}^l \tanh(x(t))$$

↑ RNNの内部状態 ↑ 入力 ↑ readout neuronからの出力

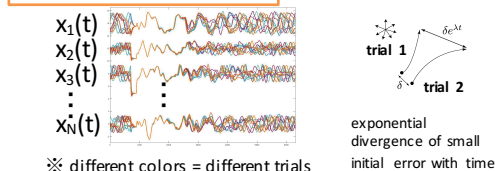


読み出し部分での学習結果を内部の学習に「転送」する方法も提案 (Sussillo and Abbott 2012)

II. 例: パターン生成



transitional consistency



III. 研究内容

パラメータによる挙動の違い

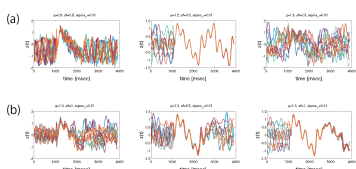


Fig. 3. (a) Outputs $z(t)$ from the trained network for several different trials and for three different values of $g = 0.9, 1.2$ and 1.5 (σ^2 is set to 0.5). (b) Outputs $z(t)$ from the trained network for several different trials and for three different values of $\sigma^2 = 0, 0.5$ and 1.0 (g is set to 1.3).

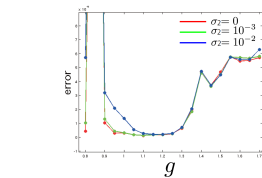


Fig. 4. The averaged error $E(z, f)$ vs. the nonlinear gain g for three different noise cases. $\sigma^2 = 0.5$. The number of 10^4 trials is used for taking the average for each parameter case.

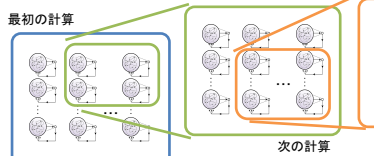
“最適”なネットワークとは?

- ・ 個々の素子の性質
- ・ ネットワーク構造
- ・ 非線形ダイナミクス



ランダムネットワーク上の非線形ダイナミクスを教師あり学習 = 同期現象の獲得プロセスとして解析

異なる初期値・パラメータの計算を同時並行に実施



- 最適パラメータの探索
- ・ 一定の閾値以下の物を破棄して繰り返し実行
 - ・ 評価
 - ・ 同期の精度・収束性
 - ・ 並列計算の効率

IV. 応用: データ同化

リザーバー集団による粒子フィルタ

- データ同化 (data assimilation): 数値シミュレーション (物理法則) + 実際の観測データの協調によるモデリング
- 粒子フィルタ (particle filter, sampling importance resampling: SIR): ベイズの定理に基づくリサンプリング → 非ガウス分布、状態と観測の非線形な関係

observation and particles

$$y_{1:t-1} : \text{observation from 1 to } t-1$$

$$\{x_{t-1}^{(n)}\}_{n=1}^N : N \text{ particles at time } t-1$$

simulation model for particles

$$x_{t|t-1}^{(n)} = F(x_{t-1|t-1}^{(n)}, v_t^{(n)})$$

new observation y_t → filtered distribution from the Bayes' rule

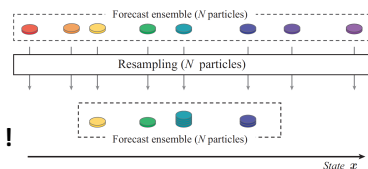
$$p(x_t | y_{1:t}) = \frac{p(x_t | y_{1:t-1}) p(y_t | x_t)}{\int p(x_t | y_{1:t-1}) p(y_t | x_t) dx_t}$$

$$\approx \frac{1}{\sum_m p(y_t | x_{t|t-1}^{(m)})} \sum_{n=1}^N p(y_t | x_{t|t-1}^{(n)}) \delta(x_t - x_{t|t-1}^{(n)})$$

$$= \sum_{n=1}^N \beta_t^{(n)} \delta(x_t - x_{t|t-1}^{(n)})$$

$$\rightarrow \{x_{t|t}^{(n)}\}_{n=1}^N$$

new particles at time t !



(Nakano et al., 2007)