A theory of memory replay and generalization performance in neural networks

Andrew M. Saxe*, Madhu Advani* *Equal contributions

Center for Brain Science, Harvard University

Overview

Why might memories be stored in hippocampus and replayed into neocortex? The Complementary Learning Systems Theory (McClelland, McNaughton, O'Reilly, 1995) holds that this two stage process allows new information to be gradually incorporated without catastrophically interfering with prior knowledge. Yet fundamental questions remain: is replay always beneficial? how much replay is optimal? and how much benefit can replay confer?

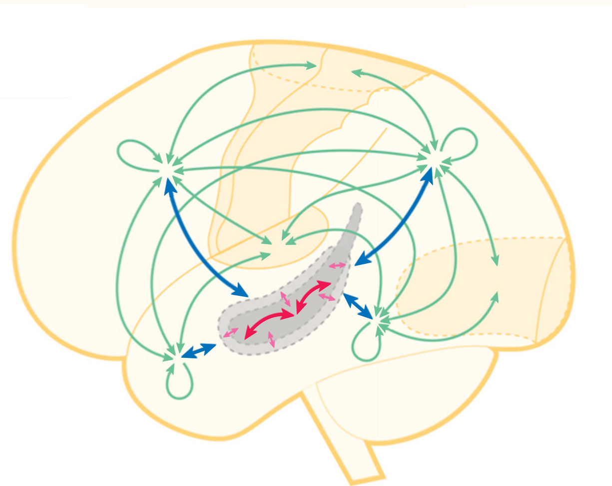
We develop a theory of the impact of experience replay on generalization performance based on an average analysis of simple neural networks. We derive exact solutions to the learning dynamics resulting from two learning strategies: online learning, in which each example is used once and discarded; and batch learning, in which all examples are stored (for instance, in hippocampus) and replayed repeatedly (for instance, during sleep).

We find that replay can be decisively better when training experience is scarce. Further, too much replay can lead to overfitting. There is therefore an optimal amount of replay that depends on the signal-to-noise ratio of the task to be learned. Our theory makes predictions about how the amount of replay should depend on task parameters if the brain is optimally managing learning; and more broadly, our results suggest a normative explanation for a two-stage memory system: replay enables better generalization from limited training experience.

Background

Striking findings of retrograde amnesia from lesions to hippocampus have long suggested that multiple subsystems interact in the formation of long term memories.

The Complementary Learning Systems Theory holds that hippocampus is specialized to rapidly store the specifics of an experience, while neocortex is specialized to slowly learn about general structure across many experiences.



Kumaran, Hassabis, & McClelland, 2017

- Rapid storage of experiences in hippocampus
- Replay from hippocampus to neocortex
- Slow learning in cortex that integrates over many experiences

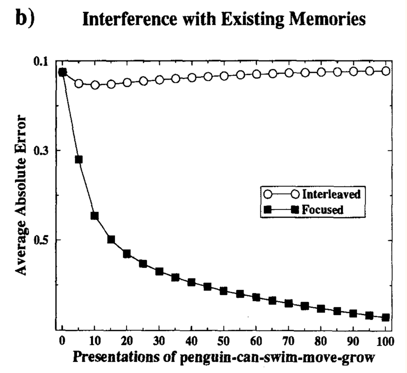
Catastrophic interference

Simulations have shown that these dual systems can prevent catastrophic interference between previously stored knowledge and new experience.

- Phase 1: Train NN on many examples
- Phase 2: Train NN on one new example

- Catastrophic interference: Phase 2 substantially disrupts knowledge acquired during phase 1

- Replay beneficial for these *supervised learning* (not *reinforcement learning*) tasks



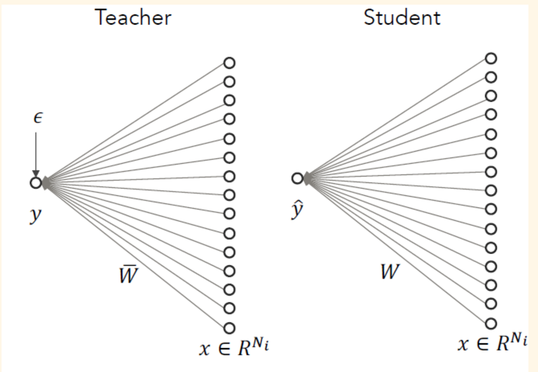
McClelland, McNaughton, O'Reilly, 1995

Prior results focus on training error, i.e., mistakes on the exact examples experienced during training. How does this impact *generalization* error to unseen examples?

Student-Teacher Model in the High-Dimensional Regime

Linear student-teacher formalism

A 'teacher' network generates labeled data examples which are fed to a 'student' network for learning (Seung, Sompolinsky, & Tishby, 1992; Amari et al., 1995; Advani & Saxe, 2017)



# of examples: P
Target outputs: $y \in R^{1xP}$
Inputs: $X \in R^{NxP}$
Label noise: $\epsilon \in R^{1xP}$
Teacher weights: $\overline{w} \in R^{1xN_i}$
Student weights: $w \in R^{1xN_i}$

Dataset of P examples:

$$X = [x^1, x^2, \dots, x^P]$$

$$y = \overline{w}X + \varepsilon$$

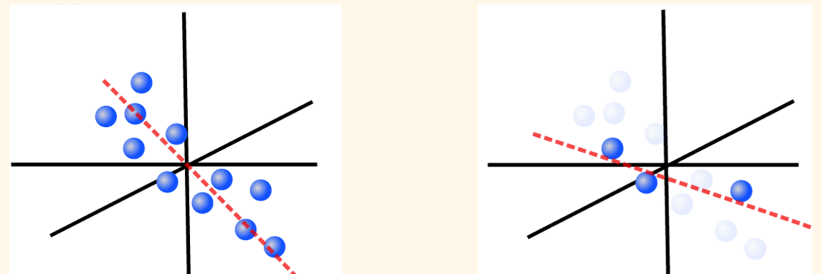
$\overline{w}_j \sim N(0, \sigma_{\overline{w}}^2)$ $\quad \varepsilon^\mu \sim N(0, \sigma_\varepsilon^2)$

$x_j^\mu \sim N\left(0, \frac{1}{N_i}\right)$ $\quad w_0 \sim N(0, \sigma_w^2)$

All drawn IID

High-dimensional regime

$\alpha = \frac{P}{N} =$ # samples / # parameters



Goal: understand how we learn from limited data.
Contrast performance of online learning with replay learning.

Seung, Sompolinsky, & Tishby, 1992; Amari et al., 1995; LeCun, Kanter, Solla, 1991; Baldi & Chauvin, 1991; Advani & Ganguli, 2016; Advani & Saxe, 2017

Online learning

In online learning, weights are updated after each example and then the example is *discarded* and cannot be visited again:

$$w_{p+1} = w_p + \eta e_p x_p^T \qquad p = 1, \cdots, P$$

Learning rate: η
Error on example p: $e_p = y_p - w_p x_p$

Can derive average generalization error vs # of examples:

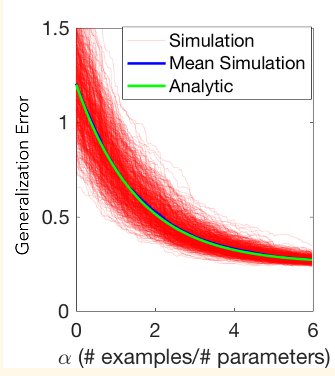
$$\frac{\langle E_g(\alpha) \rangle}{\sigma_{\bar{w}}^2} = (1 + \mathrm{INR} + 1/\mathrm{SNR}) e^{-\alpha\eta(2-\eta)} + \frac{1}{\mathrm{SNR}} \frac{1}{1-\eta/2} \left[1 - e^{-\alpha\eta(2-\eta)}\right]$$

Signal-to-noise ratio: $SNR = \frac{\sigma_{\overline{w}}^2}{\sigma_e^2}$

Initialization-noise-ratio: $INR = \frac{\sigma_w^2}{\sigma_{\overline{w}}^2}$

Results

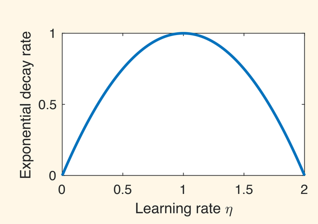
Equations accurately describe mean generalization dynamics in the high-dimensional regime.



Training speed cannot be arbitrarily increased by increasing learning rate: rate of exponential decay is $\eta(2-\eta)$ which is bounded above by 1

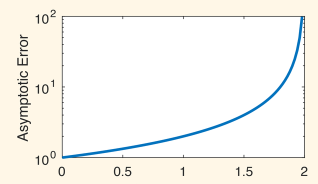
Large α limit performance depends on learning rate:

- Large rates yield fast learning but large error
- Small rates yield slow learning but can attain more accurate error



Given a fixed amount of training data (α finite), the optimal learning rate depends on SNR and INR

In the limit $SNR \to \infty$, the optimal learning rate is $\eta = 1$, showing that learning cannot be faster than exponential decay in the # of examples



Offline replay-based learning

In offline learning, all examples are stored and repeatedly used to update the weights, either using batch gradient descent or SGD with small learning rate

$$\tau \dot{w}(t) = yX^T - wXX^T$$

Inverse learning rate: τ
Time in *epochs* (passes through entire dataset): t

Average generalization error depends on correlations within finite dataset and can be derived with random matrix theory in the high-dimensional limit (Advani & Saxe, 2017):

$$\frac{\langle E_g(t) \rangle}{\sigma_{\bar{w}}^2} = \int \rho^{\mathrm{MP}}(\lambda) \left[(1 + \mathrm{INR})\, e^{-\frac{2\lambda t}{\tau}} + \frac{1}{\lambda \cdot \mathrm{SNR}} (1 - e^{-\frac{\lambda t}{\tau}})^2\right] d\lambda + \frac{1}{\mathrm{SNR}}$$

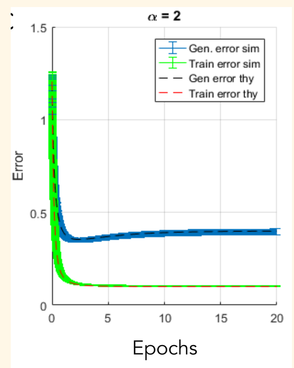
Marchenko-Pasteur distribution:

$$\rho^{\mathrm{MP}}(\lambda) = \frac{1}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} + 1_{\alpha<1}(1-\alpha)\delta(\lambda)$$

$$\lambda_\pm = (\sqrt{\alpha} \pm 1)^2$$

Results

Equations accurately describe mean generalization dynamics in the high-dimensional regime.



Replay can cause **overfitting** to correlations in specific batch, such that there is an optimal amount of replay.

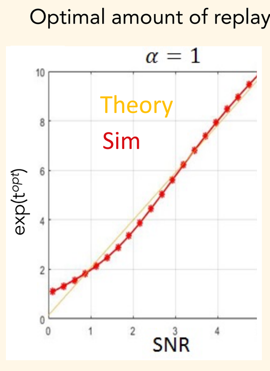
Asymptotic performance depends on α:

- Worst at $\alpha = 1$ (catastrophic overfitting)

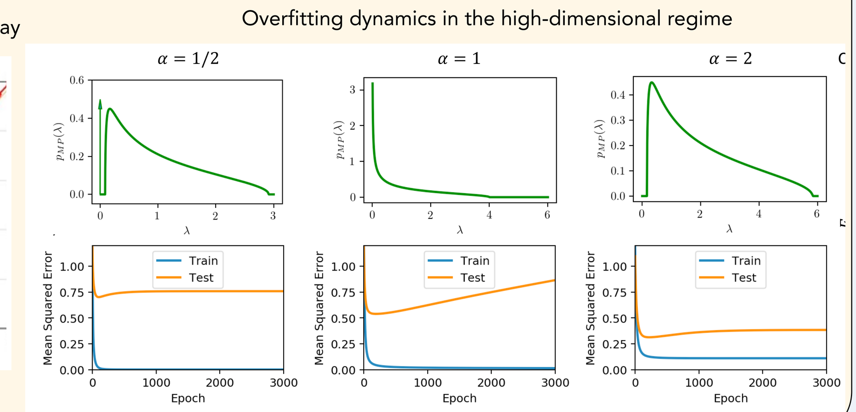
Optimal stopping time depends on SNR and INR

- Better SNR allows more replay before overfitting

Optimal amount of replay



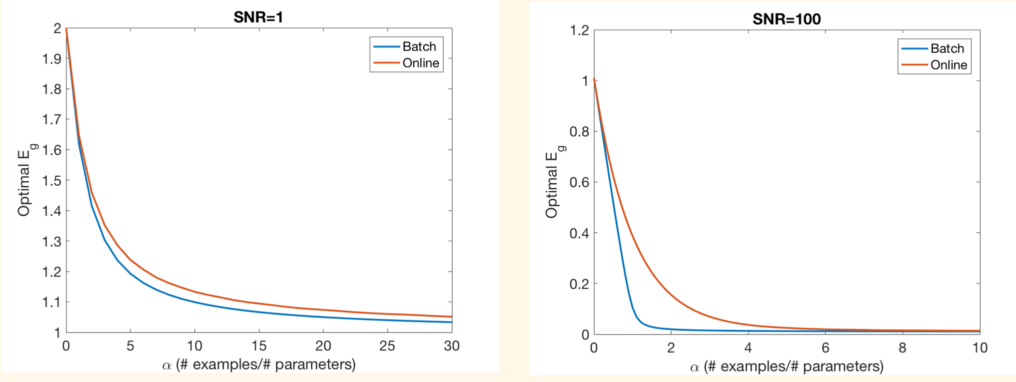
Overfitting dynamics in the high-dimensional regime



Optimal replay vs optimal online learning

When is replay a better strategy than online learning?

- Optimize out learning rate, stopping time for each setting



These results show that replay confers a decisive advantage when

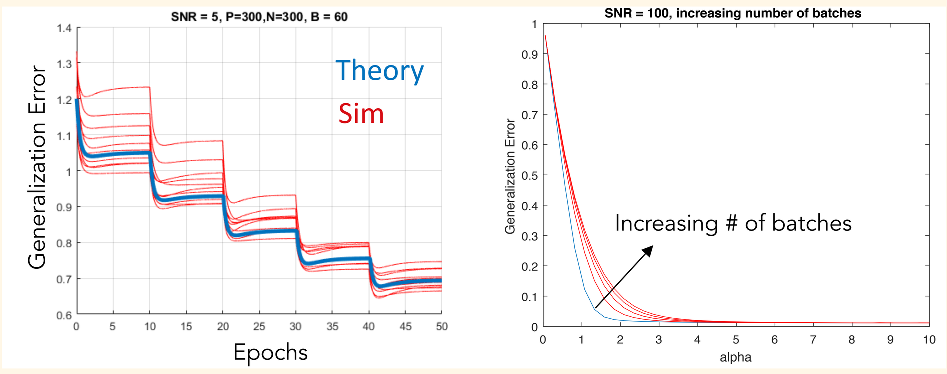
- Training data is scarce: $\alpha \leq 1$
- The rule to be learned is reliable: $SNR \to \infty$

Hence replay can yield better generalization from limited data.

Minibatch replay

Examples arrive in blocks with replay after each.

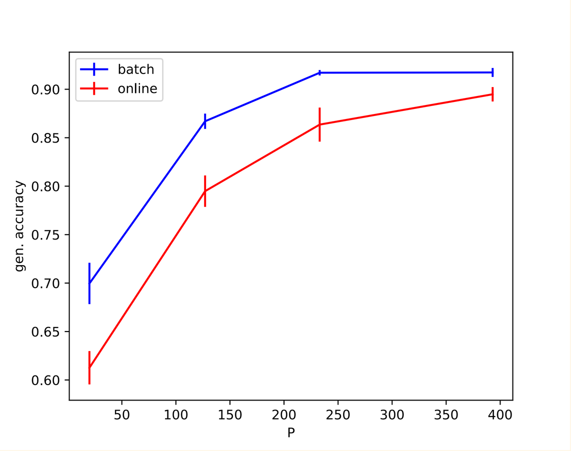
- A model for replay of a day's training experiences during sleep
- Performance interpolates between online and batch



Replay in nonlinear networks on real-world data

To see if our results hold in more realistic settings, we trained nonlinear ReLU networks on the MNIST handwritten digit dataset.

- Architecture: 1 hidden layer, 100 hidden units, softmax output
- Binary classification: 7s vs 9s
- Online learning rates: optimized over 15 log-spaced values



Optimal replay yields ~10% accuracy improvement for few training samples.

Conclusions

- A replay-based memory system can attain superior generalization performance compared to an online learner with no ability to store training examples.
- The advantage of replay is largest for small datasets ($\alpha \leq 1$) with accurate labels ($SNR \to \infty$), a plausible regime for many ecological tasks.
- Replay must be discontinued at some point to guard against overfitting, and the optimal amount of replay is predicted to scale with the SNR of the task.
- Our results suggest a normative preference for a dual system architecture even without multiple training phases or tasks that catastrophically interfere, and for a basic classification/regression setting rather than a RL setting

Funding: We thank the Swartz program in theoretical neuroscience at Harvard.