Neural optimal feedback control with local learning rules
Friedrich

Optimal Feedback Control (OFC)

- Optimal Feedback Control (OFC) is a framework from control theory that describes how to compute actions (or controls) to achieve a goal in the most efficient way, especially in systems with uncertainty, like noise or delayed feedback.
  - OFC continuously adjusts actions based on real-time sensory feedback, using an internal model to predict future states and minimize a cost (like error or effort).

Kalman Filtering and Linear-Quadratic Regulator (LQR)

- Classical approaches like Kalman filtering and linear-quadratic regulators (LQR) work well in theory, but they're not biologically plausible.
- The **Kalman Filter (for Estimation)** is an algorithm used to **estimate** the hidden state of a system (like position or velocity) from noisy observations.
  - It combines predictions from a model with actual noisy observations to get the best estimate of the current state.
  - Example: You're tracking a moving object (like your hand). You don't see it clearly due to noisy visual input, but you can predict where it's going based on how it was moving. The Kalman filter fuses that prediction and the noisy observation to guess where your hand really is.
- The **Linear-Quadratic Regulator (LQR) (for Control)** is a method to compute the optimal action (**control**) to apply to a system to minimize a cost, assuming the system's dynamics and noise are known and linear.
  - Given a current estimate of the state, it outputs a control action that minimizes a long-term cost function (usually quadratic—e.g., penalties on error and effort).
  - Example: It helps you smoothly move your hand to a target while using the least amount of energy and keeping the motion accurate.
- Why these aren't biologically plausible:
  - They require exact knowledge of system dynamics and noise statistics.
  - They often involve matrix operations, like inversion and solving Riccati equations—hard to do with real neurons.
  - They assume instant access to clean data (**no delays**) and global updates across the system.
  - Learning in these frameworks isn't local (i.e., neurons can't just use their own input/output to update weights).

- So, while Kalman filters and LQRs are mathematically elegant, the brain probably doesn't implement them directly, which is why papers like this one try to build **neural approximations that use biologically plausible rules**.

Core Idea of Bio-OFC

- The authors aim to model how the brain might solve control problems like moving your hand to a target despite sensory noise and delay, without needing to know the exact equations of motion or noise levels. Classical approaches like Kalman filtering and linear-quadratic regulators (LQR) work well in theory, but they're not biologically plausible.
- So they build a network that:
    - Uses adaptive Kalman filtering for estimating the system state from delayed/noisy observations.
    - Uses policy gradient (a reinforcement learning method) for learning a control policy, avoiding backpropagation and non-local updates.
    - Learns both the system dynamics (A, B, C) and Kalman gain (L) using online local learning rules, meaning updates are only based on information locally available at the synapse.

1. **Bio-OFC Architecture**:

   - It contains neurons representing:
     - **State estimates** ($\hat{x}_t$),
     - **Prediction errors** (e_t),
     - **Control actions** (u_t).
   - **Kalman filtering is modified** to work with delayed observations ($y_{t-\tau}$).

2. **Learning Rules**:

   - **Local plasticity rules** update system parameters using Hebbian-like updates.
   - Control weights are updated using **policy gradients with eligibility traces**, simulating dopamine-like global signals.

3. **No Prior Knowledge Assumed**:

   - Doesn't assume knowledge of noise covariances or dynamics.
   - Works **online** with **single-phase learning** (no alternating "learning" and "execution" phases).

4. **Biological Plausibility**:

   - **No weight transport** or centralized computation.
   - Supports **delayed feedback**, something prior models ignored.

5. **Experiments**:

   - **Double integrator system** (a basic control problem),
   - **Human hand-reaching task** (comparison with motor adaptation experiments),
   - **2D winged flight** (tests performance in nonlinear and delayed scenarios).

-

Table 1: **Limitations of previously proposed neural implementations of OFC.** Presence or absence of different properties in previously proposed neural models, and their comparison to Bio-OFC. Guide to symbols: ✓: true, ✗: false, ✓✗: partially true, N/A: not applicable.

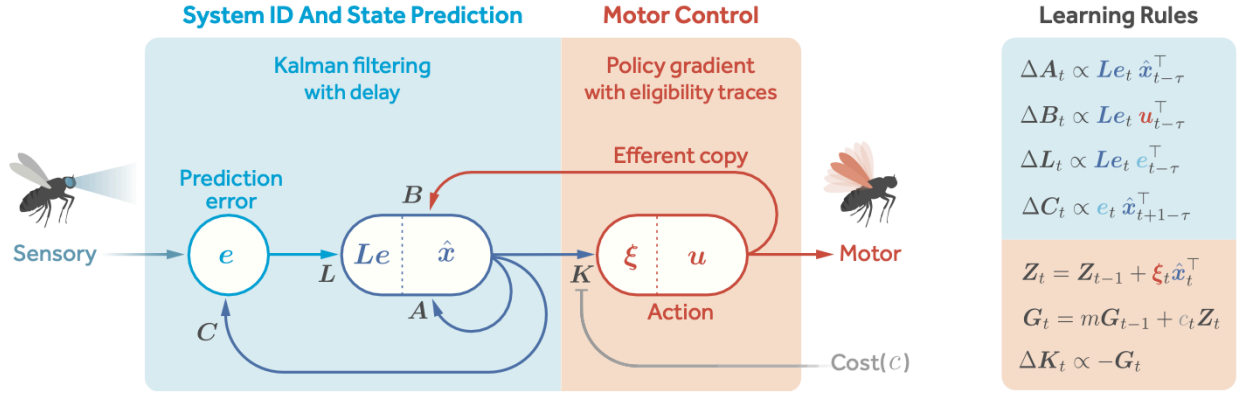|	[4]	[6]	[5]	[7]	[8]	Bio-OFC
delayed sensory feedback	✗	✗	✗	✗	✗	✓
control included	✗	✓	✗	✗	✗	✓
noise covariance agnostic	✗	✓	✗	✗	✗	✓
online system identification	✗	✓	✗	✓	✓✗	✓
local learning rules	N/A	✓	N/A	✗	✓	✓
tractable latent size	✗	✓	✗	✓	✓	✓
absence of inner loop	✓	✗	✓	✓	✗	✓
single phase learning/execution	N/A	✗	N/A	✓	✓	✓

-

Figure 1: **The circuit and learning rules of the Bio-OFC algorithm.** Our circuit is comprised of two main parts. First (in blue), the circuit performs Kalman filtering. Then (in red), the circuit performs control using policy gradients with eligibility traces. Triangular arrowheads denote synaptic connections and the flat arrowhead denotes the modulatory effect of the cost signal.