Grow-Prune-Freeze Networks: Adaptive & Continual Learning Technique for Olfactory Navigation

Kordel K. France, Ovidiu Daescu

Keywords: olfaction, policy gradient, temporal difference learning, robotics, world models

Summary

Training data for olfaction is scattered through disparate, non-standardized datasets that limit the ability to build representative world models. Olfactory navigation is a highly dynamic and non-stationary task that benefits from real-time continual learning. We introduce an adaptive framework called *Grow-Prune-Freeze (GPF) networks* that enable an agent to continually learn through growing, pruning, and freezing early layers of its policy in response to world complexity. Grounding GPFs in non-linear random matrix theory, we show that the work of Pennington & Worah (2017) can be extended from single hidden layers to n-layer continual-learning models, and that eigenvalue composition of network weights is preserved as successive layers are added. We show that GPFs based on Expected SARSA achieve a 94% success rate on turbulent plume navigation—a partially observable, non-stationary task representative of the "big world" challenges that motivate adaptive learning in robotics—and provide supporting methodology for applying GPFs in other world models. Further experiments amount evidence that GPFs may generalize well to other machine learning tasks such as reinforcement learning in Atari, image classification, and autoregressive language models. We open source all code and data to encourage improvements on and more research in olfactory robotics.

Contribution(s)

1. A framework for adaptive and continual reinforcement learning agents operating in partially observable, non-stationary environments and grounded in random matrix theory (RMT).
   Context: We show that foundational pruning mechanisms from LeCun et al. (1989) can be used to improve policy gradient methods whilst simultaneously growing a multi-layer perceptron and freezing earlier layers to hedge against catastrophic forgetting.
2. RMT analysis showing spectral eigenvalue invariance for indefinitely adapting models, providing a mechanism for resource-efficient learning in robotics that enables an agent to allocate representational capacity where environmental signals are complex or changing.
   Context: Work by Pennington & Worah (2017) and Martin & Mahoney (2021) show that spectral invariance among eigenvalues of weight matrices can be preserved among neural networks with a single hidden layer. We extend their methods to show that their assumptions support neural networks with multiple hidden layers.
3. Tuning criterion for introduced parameters to stabilize continual learning with GPFs and an Expected SARSA policy.
   Context: Prior work by France et al. (2026) highlights the potential of Expected SARSA in olfactory navigation and we show that their findings extrapolate well to our GPF framework.
4. Empirical demonstration that GPFs achieve 98% source-finding success on a partially observable olfactory navigation task, generalizing to 94% on held-out trials across diverse random initializations and notable performance on non-olfactory tasks.
   Context: GPFs were specifically designed to solve a problem in partially observable olfactory tasks. However, we demonstrate our methods over multiple tasks to build support for generalization of GPFs in other environments, but do not claim state-of-the-art performance across all benchmarks.

Grow-Prune-Freeze Networks: Adaptive & Continual Learning Technique for Olfactory Navigation

Kordel K. France[1], **Ovidiu Daescu**[1]
kordel.france@utdallas.edu, ovidiu.daescu@utdallas.edu

[1]**Department of Computer Science, University of Texas at Dallas, USA**

Abstract

Training data for olfaction is scattered through disparate, non-standardized datasets that limit the ability to build representative world models. Olfactory navigation is a highly dynamic and non-stationary task that benefits from real-time continual learning. We introduce an adaptive framework called *Grow-Prune-Freeze (GPF) networks* that enable an agent to continually learn through growing, pruning, and freezing early layers of its policy in response to world complexity. Grounding GPFs in non-linear random matrix theory, we show that the work of Pennington & Worah (2017) can be extended from single hidden layers to n-layer continual-learning models, and that eigenvalue composition of network weights is preserved as successive layers are added. We show that GPFs based on Expected SARSA achieve a 94% success rate on turbulent plume navigation—a partially observable, non-stationary task representative of the "big world" challenges that motivate adaptive learning in robotics—and provide supporting methodology for applying GPFs in other world models. Further experiments amount evidence that GPFs may generalize well to other machine learning tasks such as reinforcement learning in Atari, image classification, and autoregressive language models. We open source all code and data to encourage improvements on and more research in olfactory robotics.

1 Introduction

Modern deep learning architectures are typically static, limiting their applicability in dynamic and non-stationary tasks. In addition, many of these modern models are based on transformers trained on

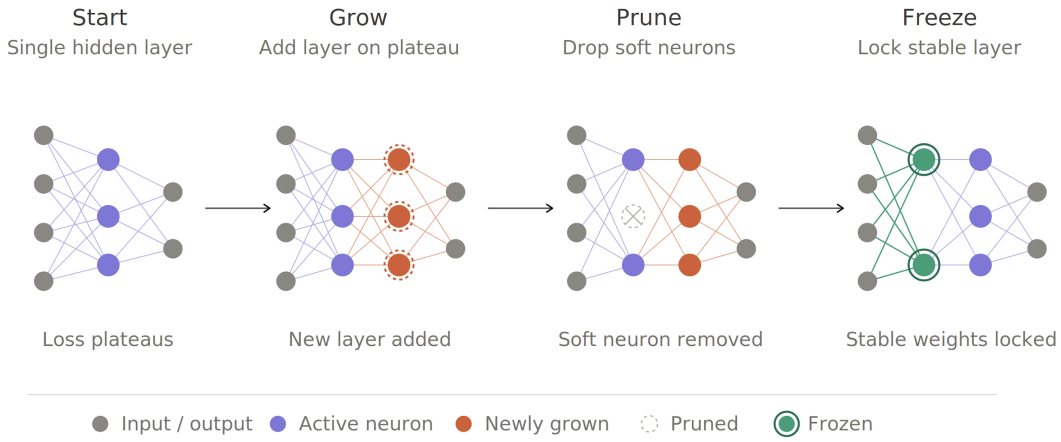

Figure 1: Conceptualization of a Grow-Prune-Freeze (GPF) neural network.

petabytes of internet data. The luxury of big data does not exist among the unstandardized and data-starved world of olfaction which discourages work in representative world models. Furthermore, molecular dynamics cause an environment to constantly evolve, and some environments are more dynamic than others which correspondingly warrant more compute in order to correlate molecular patterns within the air. Tasks like olfactory navigation significantly benefit from runtime learning. As a result, we became drawn toward machine learning methods that could continuously update their own priors based on batches of observed data, and adjust their own architectures according to the observed data complexity. These desires led us to develop *Grow-Prune-Freeze (GPF) networks*, grounded in known work from random matrix theory and scaled from first-principles mathematics.

Our motivation is to develop a useful method for robotic runtime learning in partially observable world models, with a specific emphasis on olfactory navigation: a robot must locate a chemical source by following a turbulent odor plume in sparse-signal conditions using an edge-deployable policy-gradient model. Growing layer-by-layer enables a network to adapt capacity to plume complexity, pruning keeps the model compact for onboard inference, and freezing preserves temporal dynamics learned from early plume encounters to hedge against catastrophic forgetting. We selected foundational temporal difference learning and pruning methods to emphasize that the performance observed in our experiments is credited to *GPFs* rather than an exotic learning architecture. In further support of *GPFs* as a generalized method, we elaborate on their intuitive extension to other learning tasks in the Supplementary Material.

2 Grow-Prune-Freeze Networks

Continual learning requires adapting indefinitely to non-stationary streams while maintaining *stability* (retention of prior knowledge) and *plasticity* (ability to learn new tasks). Standard deep learning suffers catastrophic plasticity loss in continual settings as shown by Dohare et al. (2024), and while adaptive pruning methods such as those in Lasby et al. (2024); Frankle & Carbin (2019) show promise, they lack theoretical grounding for long-term spectral stability.

Grow-Prune-Freeze neural networks offer a framework that automatically (1) grows hidden layers when validation loss plateaus; (2) prunes "soft" neurons according to temporal belief values; and (3) freezes stable layers to prevent forgetting while preserving plasticity.

We root *GPFs* in random matrix theory (RMT) and analogize the mathematics to theories of deep learning architectures. Specifically, we show that the work of Pennington & Worah (2017) can be extended from single-layer neural networks to n-layer continuous-learning deep neural networks that incrementally add layers as they learn while preserving eigenvalue spectrality. Through this, we also explore how a *GPF's* low-rank perturbations maintain Marchenko-Pastur spectral densities with the appropriate shape, ensuring dynamical isometry and gradient propagation. We then show that principles from Tao (2012); Martin & Mahoney (2021) hold as a result. The final model partially addresses the three plasticity correlates identified by Dohare et al. (2024): (1) weight magnitude growth, (2) dead units, and (3) effective rank collapse.

2.1 Relevant Background

Random matrix theory is the bedrock from which *GPFs* were inspired and built. Pennington & Worah (2017) derive spectral densities for non-linear activations in basic neural networks with a single hidden layer, and Cohen-Karlik et al. (2022) link low-rank bias to generalization. Martin & Mahoney (2021) cover RMT as it pertains to deep learning by showing how model width and *iid* weight assumptions can be checked and controlled. We extend these works to non-equilibrium ensembles induced through the the continual adding of layers in multilayer perceptrons.

Dynamic Sparse Training (DST) by Lasby et al. (2024) alternates pruning and regrowth of neural networks, while MorphNet from Gordon et al. (2018) optimizes sparsity via regularization. Growing methods like *DEN* from Yoon et al. (2018) add neurons according to loss plateaus. General prin-

ciples for constructing adaptive neural networks are recently enumerated by Mirzadeh et al. (2022) and backed by empirical support on recent model architectures.

Elastic Weight Consolidation from Kirkpatrick et al. (2017) highlights weight-specific penalization during learning; *Progressive Nets* by Rusu et al. (2016; 2022) expand layers laterally to support additional inputs. Dohare et al. (2024) demonstrate plasticity loss across optimizers and propose continual backpropagation (CBP) via selective reinitialization. Additional work from Li & Ji (2021) and Lyle et al. (2023) denote other challenges in building adaptive architectures that continually learn. *GPFs* build on top of principles established in each of these works by using random matrix theory as the grounding ideology and *Columnar Constructive Netwroks* established by Javed et al. (2023) as the inspiration for continually growing layers.

Fehring et al. (2025) show grow-and-prune methods may be empirically effective but lack theoretical characterization of how structural adaptation affects network dynamics. Magnitude-based pruning from LeCun et al. (1989), the Lottery Ticket Hypothesis by Frankle & Carbin (2019), and parameter freezing from Foster et al. (2017) all give supporting evidence that sparse, stabilized networks preserve critical knowledge. In *GPFs*, pruning and freezing are integrated with growth to form a continuous cycle of adaptive architecture. We select LeCun's magnitude-based pruning as the specific method for our evaluation.

3 Theoretical Analysis

The *GPF* framework assumes a very simple neural network to start: an input layer, an output layer, and a single hidden layer. Consider a feedforward network with $L_t = 1$ layer at time $t = 0$, input dimension n_0, hidden widths $n_1, \ldots, n_{L_t-1}$, and output n_{L_t}. Weights $W^\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ are initialized from Kaiming distribution. *GPFs* grow depth only as their environment demands it, preserving the eigenvalue spectrality of previously learned weights.

To facilitate this, we define a threshold ω_l within the validation loss. As it learns, if the network experiences no change in loss above this threshold, a new hidden layer ℓ_i is added, where $i > 0, i \in \mathbb{R}$. At time k, there is a "patience" variable ρ_k such that there must be ρ_k epochs with validation loss all below this threshold in order to add a new layer.

The above can be denoted as follows: Add layer $\ell_i = L_t + 1$ if validation loss $\mathcal{L}_{\text{val}}$ satisfies:

$$\mathcal{L}_{\text{val}}^{(t-k)} - \mathcal{L}_{\text{val}}^{(t)} < \omega_l, \quad \forall i \in [t-k+1, t], \tag{1}$$

where $\omega_l > 0$ is the stagnation threshold.

Each neuron's input weight $w_{i,j}^\ell$ has belief $h_{i,j}^\ell \in [0, 1)$. This belief value is such that if the absolute value of a neuron's weight is consistently above a positive value ω_w (or that neuron is constantly fired) for multiple successive epochs, then it is incremented, and thus becomes "hardened". New layers initialize $n_{L_t+1} = n_{L_t}$ neurons with $h_{i,j}^{(L_t+1)} = 0$. We perform network updates according to the following:

$$h_{i,j}^\ell(t) = h_{i,j}^\ell(t-1) + \eta_h \cdot \mathbb{1}\left\{|w_{i,j}^\ell(t)| > m\right\}, \tag{2}$$

where $\eta_h = 1/(k_{max})$ normalizes to approach but never reach 1 according to the current training epoch k_{max}. Hardened neurons are asymptotically closer to 1 and adverse to pruning. A belief value of 1 is associated with a neuron considered as "frozen", which we discuss shortly.

Neurons with persistently high belief are hardened and resistant to pruning. We follow the magnitude-based pruning technique established by LeCun et al. (1989) with the exception that a layer's belief values protect useful low-magnitude weights that encode sparse but relevant patterns as shown by Frankle & Carbin (2019).

Parameter ω_p is a layer-wise pruning threshold. After at least ρ_b epochs have occurred since the addition of layer ℓ_i, all layers $\ell_0 - \ell_{i-1}$ prior to the one just added will be pruned such that all neurons within said layers containing belief values below some threshold of $|\omega_p|$ will be removed from the network, or their weights set to zero.

Succinctly, we prune neuron j in layer $\ell_i < \ell_k$ if:

$$h^{\ell}_{i,j} < \omega_p \quad \text{and} \quad |w^{\ell}_{i,j}| < d_i. \tag{3}$$

Newly added layers start with neurons whose belief values each have $h = 0$ and are exempt from pruning until ρ_b patience epochs have elapsed; the grow-prune cycle continues indefinitely.

Another set of tunable parameters are ρ_f and ω_f, where $\rho_f < k_{\textbf{max}}$ and $\omega_f \in [0, 1]$. We freeze layer ℓ_i if, for a patience of ρ_f epochs, $100.0 * \omega_f\%$ weights satisfy:

$$|w^{\ell}_{i,j}(k_{\textbf{max}} - q) - w^{\ell}_{i,j}(k_{\textbf{max}})| < \omega_f, \quad \forall q \in [k_{\textbf{max}} - \rho_f + 1, k_{\textbf{max}}]. \tag{4}$$

Freezing follows methods given by Dohare et al. (2024) to prevent forgetting while maintaining plasticity in newer layers. Optionally, batch size may decay as $b_t = b_0 \cdot \exp(-\eta_{bs}\mathcal{L}^{(t)}_{\text{val}})$ to account for evolving network size and convergence rate as shown by Adachi et al. (2024). As we soon show, freezing stable layers also ensures that the spectral distribution of the Gram matrix is preserved.

3.1 Training GPFs

As the network begins training, it does not know how many layers it will eventually add. One may cap the layer count, but we argue that a properly developed *GPF* will cease growing layers at the appropriate time if designed correctly. A continually learning model may, in some instances, maintain the same neuron count at time t_0 as at time t_n, but with more layers and increased sparsity. We desire sparse models that grow in depth but necessarily and sufficiently parameterize themselves to the data they perceive. In this manner, the concepts of training and testing being two separate processes are foregone. For *GPFs*, feedback is directly incorporated according to its experience or training data and there is no separate test phase. We design *GPFs* to be employed in scenarios where there is too little data for conventional training mechanisms and/or where the data may contain temporal patterns that change over time according to new dynamics introduced into the environment, as is the case with plume tracking in olfactory robotics.

3.2 Analysis via Random Matrix Theory

Pennington & Worah (2017) showed that the distribution of eigenvalues from the data covariance matrix is idempotent after passing through a single nonlinear layer of a neural network. We aim to show that this principle extrapolates to any number of layers $L \in [1, \infty]$ as long as (1) stable depth-wise growth occurs within the network according to Javed et al. (2023), (2) pruning is done according to the Hessian approximation in LeCun et al. (1989), and (3) layers are frozen at the appropriate time to preserve pruned knowledge. We show how principles from Schoenholz et al. (2016) govern the limit for how deep a network can be given a set of hyperparameters, and how decisions on growth can occur based on their assumptions.

We follow the notation of Pennington & Worah (2017), reviewed in full in Supplementary Section C. The core object at each layer ℓ is the empirical Gram matrix $M^{(\ell)} = \frac{1}{m}[Y^{(\ell)}]^\top Y^{(\ell)}$ where $Y^{(\ell)} = f(W^{(\ell)}Y^{(\ell-1)})$, and its Stieltjes transform $s^{(\ell)}(z) = \frac{1}{m}\operatorname{Tr}((M^{(\ell)} - zI)^{-1})$.

We extend the above to the multi-layer case. For a network with L layers and weight matrices $W^{(1)}, \ldots, W^{(L)}$, define the activation recursion $Z^{(\ell)} = W^{(\ell)}Y^{(\ell-1)}$, $Y^{(\ell)} = f(Z^{(\ell)})$, $Y^{(0)} = X$, under the assumptions established by Pennington & Worah (2017) (highlighted in Supplementary Section C). The layer-ℓ Gram matrix and Stieltjes transform are:

$$M^{(\ell)} \;=\; \frac{1}{m}\,[Y^{(\ell)}]^\top Y^{(\ell)}, \tag{5}$$

$$s^{(\ell)}(z) \;=\; \frac{1}{m}\operatorname{Tr}\big((M^{(\ell)} - zI)^{-1}\big). \tag{6}$$

The nonlinearity of activation functions precludes many principles from random matrix theory applying to neural networks. While many of the tools for assessing moments and spectral densities from matrix theory still apply, we show how these tools can be further employed to analyze space folds with ReLU activations (and friends) as demonstrated by Lewandowski et al. (2025).

Let Ψ_ℓ be an activation function that acts as a *nonlinear spectral transformer*, encoding how a random layer reshapes the spectral statistics of the preceding layer. Explicit expressions for Ψ_ℓ depend on integrals involving the activation $\psi \in \Psi$ and may reduce to polynomial or rational forms for common nonlinearities. Under the moments-method derivation (counting non-linear random walks at each layer) one obtains, for each ℓ, a deterministic operator Ψ_ℓ that enables Equation 6 to be written as

$$s^{(\ell)}(z) \;=\; \Psi_\ell\big(s^{(\ell-1)}(z),\, z\big)$$

for every $z \in \mathbb{C} \setminus \mathbb{R}$, where Ψ_ℓ depends on the activation ψ, layer variance σ_ℓ^2. Thus, when depth is increased by adding layers, the transform is updated by *composition* of these operators. We can define the depth-L composition operator

$$\Psi^{(L:1)} \;=\; \Psi_L \circ \Psi_{L-1} \circ \cdots \circ \Psi_1,$$

so that the Stieltjes transform at depth L is succinctly written as

$$s^{(L)}(z) \;=\; \Psi^{(L:1)}\big(s^{(0)}(z),\, z\big). \tag{7}$$

We formalize the conditions for convergence of GPFs at any layer depth limit via Theorem 1.

Theorem 1 (Depth-Limit under Composition) *Suppose the layers are homogeneous such that $\psi_\ell = \Psi \forall \psi \in \Psi$. Assume there exists an open set $\mathcal{U} \subset \mathbb{C} \setminus \mathbb{R}$ containing the spectral support of $M^{(\ell)}$ for all ℓ, and a Banach space $(\mathcal{S}, \|\cdot\|)$ of analytic transforms on $\mathcal{U}$ such that:*

1. *Condition 1: $\Psi : \mathcal{S} \times \mathcal{U} \to \mathcal{S}$ is well-defined and analytic in its first argument for each fixed $z \in \mathcal{U}$.*
2. *Condition 2: For each fixed $z \in \mathcal{U}$, the Fréchet derivative $D_s\Psi(s, z)$ exists and there is a constant $\rho(z) < 1$ with*

$$\sup_{s \in \mathcal{S}} \|D_s\Psi(s, z)\| \le \rho(z) < 1.$$

Then for every initial Stieltjes transform $s^{(0)} \in \mathcal{S}$, the iterates $\{s^{(L)}\}_{L \ge 0}$ defined by $s^{(L)} = \Psi^{\circ L}(s^{(0)}, \cdot)$ converge in $\mathcal{S}$ to the unique fixed point $s^\star(\cdot)$ that satisfies $s^\star(z) = \Psi(s^\star(z), z)$ for all $z \in \mathcal{U}$. Consequently, the empirical spectral distribution of $M^{(L)}$ converges (as $m, n_0, \ldots, n_L \to \infty$, jointly with $L \to \infty$) to the deterministic law given by the Stieltjes transform $s^\star$.

Freezing layers locks spectral distributions into their ideal shape to further enable convergence. Our analysis assumes layer homogeneity, with new layers always initialized at identical width d. Formal convergence results along with further remarks on Theorem 1 and ReLU spectral analysis appear in Supplementary Sections D and E.

4 Experiments

We evaluate the *GPFs* on olfactory navigation - the problem of a robot localizing an odor source by following a turbulent plume back to its origin. This task is representative of the class of sparse-reward, partially observable navigation problems that motivate runtime learning algorithms in big worlds: the signal is intermittent, the environment is non-stationary, and the appropriate behavioral strategy depends on context that must be inferred from current state and recent sensory history.

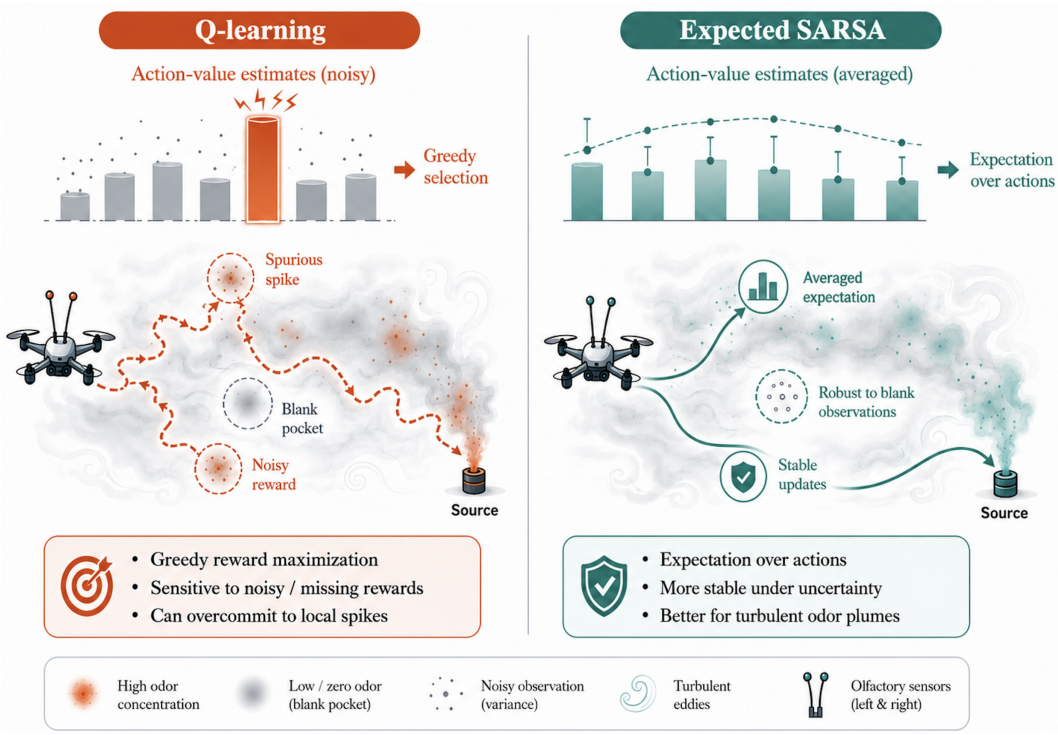

Figure 2: Q-learning vs Expected SARSA in Olfactory Navigation

4.1 Plume Simulator

We simulate the plume using a Farrell–Murlis filament model Farrell et al. (2002) in a $20\,\mathrm{m} \times 20\,\mathrm{m}$ two-dimensional domain at $10\,\mathrm{Hz}$. The source emits filaments at a Poisson-sampled $\lambda = 5$ filaments/s from a randomly chosen upwind position. Filaments advect with a coupled Ornstein–Uhlenbeck wind field with velocities of 1.0m/s, gusting durations of 2.0m/s, and spread via Gaussian diffusion rate of $0.05\,\mathrm{m}^2/\mathrm{s}$ Total concentration is clamped to 1.0 and corrupted by additive noise $\sigma_n = 10^{-3}$. Full details on the simulator construction are enumerated in Supplementary Section G.

4.2 Agent and Sensor Model

The agent carries a bilateral antenna pair (10 cm inter-antenna separation) and selects from six discrete actions: *surge forward*, *turn* $\pm 15°$, *turn around* $180°$, and *cast* $\pm 30°$, each followed by a 50 cm forward step. Olfactory inertial odometry (OIO) established by France & Daescu (2025) is used to fuse inertial and olfactory states. The antennae are modeled as fast-sampling Sensirion MICS-6814 metal oxide sensors building on those models established by recent olfactory navigation work from Dennler et al. (2024) and France et al. (2026). Episodes terminate when the agent moves within 50 cm of the source or after 1200 steps (100 s). Initial placement of the odor source is 3–10 m downwind with $\pm 30°$ heading noise. Sensor observations are discretized into a 22-dimensional state: seven concentration bins (blank + six log-spaced levels above σ_n) per antenna and eight wind-direction octants relative to heading, giving $7 \times 7 \times 8 = 392$ unique states. Further details on state- and action-space tokenization can be observed in Section G of the Supplementary Material.

4.3 Expected SARSA Policy

We build off the work established by France & Sheppard (2023) and France et al. (2026) showing the advantage Expected SARSA can provide against greedy Q-learning in partially observable environments with sparse rewards. We find that a policy that seeks the maximum reward may induce

lossy actions when rewards are sparse and not fully observable; a policy that seeks the *average* reward for a given action set supports robustness against eddy currents and blanks in a plume. Figure 2 illustrates this comparison between Q-learning and Expected SARSA agent policies.

Furthering this hypothesis, we select the latter as the navigation policy for both a canonical agent and *GPF* agent. The canonical agent contains a single hidden layer to evaluate as a control against the *GPF* agent. Both agents are identical in architecture otherwise. The Q-function is represented by a simple dense MLP with ReLU activations. The initial architecture is $22 \to 64 \to 6$ (one hidden layer of width 64) trained with the Expected SARSA update rule under an ε-greedy policy, with ε decayed from 1.0 to 0.05 multiplicatively by 0.9995 per episode. We select the Adam optimizer with learning rate $\alpha = 1e^{-3}$, and a discount factor $\gamma = 0.99$. We keep the neural network basic for two reasons: (1) our intent is to evaluate *GPF* methodology, not an exotic network architecture, and (2) a simple network ensures maximum exportability to and reproducibility in various model interpreters required by edge robotics. The belief-value pruning threshold is set as $\omega_p = 10^{-6}$.

The per-step reward r_t for both agents combines a constant time penalty of $r_{\text{time}} = -0.01$ with event signals r_{event} of $+100$, $+1.0$, and -0.05 for success, whiff, and prolonged blanks $> 2\,\text{s}$, respectively. A potential-based distance-shaping term r_{shape} (Ng et al., 1999) is delegated to provide dense gradient signal without altering the optimal policy. In culmination of the above, the final reward function is represented as $r_t = r_{\text{time}} + r_{\text{event}} + r_{\text{shape}}$. All experiments use seed 42 for the environment and seed 43 for the agent. Training runs for $4{,}500$ episodes with evaluation every 500 episodes over 200 held-out episodes. The highest-performing checkpoint is retained. Full hyperparameters for all experiments are available in Table 5 of the Supplementary Material.

5 Results

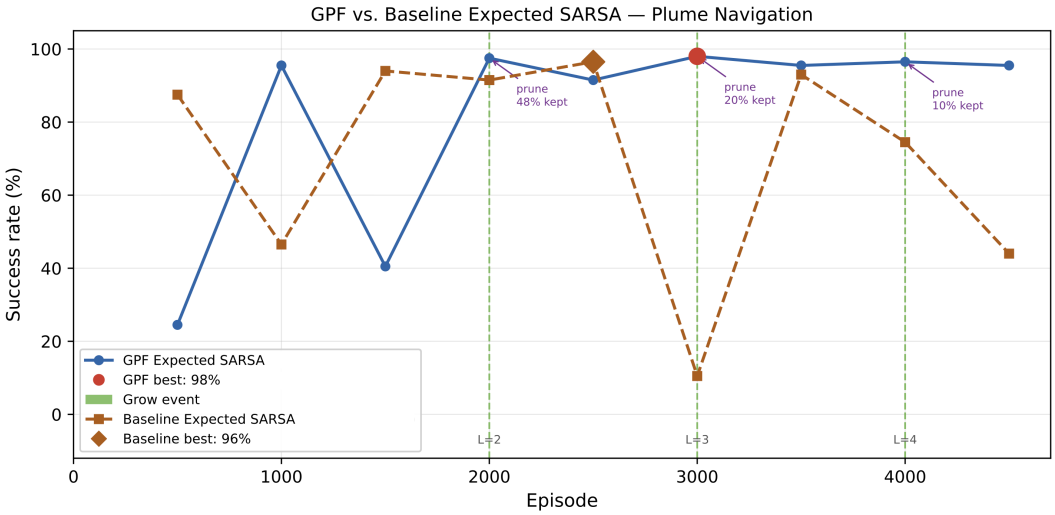

Figure 3: *GPF* Expected SARSA training curve on the plume navigation task. The orange dotted lines denote the success rate of the baseline agent; the solid blue line denotes that of the GPF agent. Dashed green lines mark grow events; purple annotations show the fraction of weights retained by each prune pass. The red dot marks the best checkpoint. The bottom access indicates layer count.

Figure 3 shows the evaluation success rate throughout training. The agent requires roughly 1000 episodes to acquire consistent navigating behavior. Three grow events occurred at episodes $2{,}000$, $3{,}000$, and $4{,}000$. After the first grow-prune pass (retained 47.9% of weights), the two-layer network temporarily dropped to 91.5% before recovering to $\mathbf{98.0}\%$ at episode $3{,}000$ - the best checkpoint. Subsequent grows were more aggressive (19.6% and 9.9% weight retention); the four-layer configuration peaked at 95.5% before degrading to 68.0% by episode $5{,}000$ as ε-decay compounded

ultra-sparse representation. The full episode-by-episode trajectory is in Table 4 of the Supplementary Material. Extending the episode budget from 50 s to 100 s raised the peak from 94.5% to 98.0% by eliminating timeout failures on hard initializations where the agent is far downwind or in a weak-signal region, allowing full surge–cast–surge cycles to complete as necessary.

Evaluating this checkpoint over 100 independent episodes with distinct random seeds, the agent achieved 94% success in acquiring the source of the plume. All six failures were timeouts in which the agent correctly tracked the plume but exhausted the episode budget - consistent with timing rather than navigational strategy as the remaining failure mode. We do not increase the episode budget because the molecular dynamics of the odors of interest would equalize within the plume beyond the allotted time, leaving no signal for the agent to trace. We leave varying episode budgets according to chemical diversity and plume density as an intuitive avenue for future work.

The GPF agent learned a high-performance navigation policy through structured growth of a sparse four-layer Expected SARSA policy that achieved 98% on the training distribution and 94% on held-out random initializations. These results support that the GPF grow-and-prune mechanism successfully compresses a capable policy into a sparse architecture, and that the 100-second episode budget is sufficient for the agent to develop and execute the multi-phase olfactory search strategies - surge, cast, and loiter - that characterize successful biological plume navigation.

To provide support that *GPFs* generalize beyond olfactory navigation, we perform a battery of experiments over standard image classification, Atari reinforcement learning, and autoregressive language modeling problems. These results can be found in Section J of the Supplementary Material.

6 Conclusion

GPF networks were developed out of a need for lightweight and adaptable olfactory navigation agents that could be deployed at the edge on real robots. Olfaction does not have the luxuries of other modalities in plentiful and accessible world models. As such, plume navigation tasks significantly benefit from adaptable runtime learning. *GPFs* and temporal difference learning provide an encouraging solution for this by growing depth according to observed environment complexity, prune low-saliency weights as learning occurs, and freezing stable layers - addressing catastrophic forgetting and plasticity loss while maintaining spectral invariance via Stieltjes composition.

On the olfactory navigation task, the 94% source-finding success rate supports GPFs as a practical candidate for adaptive temporal difference learning agents operating in partially observable, non-stationary environments, while also gathering evidence as a continual learning framework for the same. A larger experimental battery across diverse benchmarks and chemical plumes remains the primary open direction to validate *GPFs* on a more comprehensive level. We hope our work here inspires further research into continual learning and olfactory world models for robotics.

All code for the experiments associated with this paper is available at the following URL: `https://github.com/KordelFranceTech/Grow-Prune-Freeze-Neural-Networks`

Broader Impact Statement

Our work here progresses advanced navigation abilities for robots. Olfaction is still a largely unexplored modality. It lacks a unifying data standard, ImageNet-scale training datasets, and no formal benchmarks for tasks like scent-based navigation. Maturing olfaction to the degree of its vision and audition counterparts will enable many new capabilities for AI and robotics. However, these will be capabilities for which humanity has little experience in addressing due to the aforementioned obstacles. Additionally, enabling AI models to adapt their own architectures and continually learn as we demonstrate here with *GPFs* could enable unintended behaviors. A *GPF* model may be deployed against certain criteria and alter itself to learn action sequences adversarial to its original policy.

Consequently, we feel our work here adds to the growing call-to-action to establish meaningful multimodal benchmarks for continual learning to ensure adaptive models remain grounded as they explore the world. We strongly encourage the olfactory modality to be included in this benchmark.

References

Masaki Adachi, Satoshi Hayakawa, Martin Jørgensen, Xingchen Wan, Vu Nguyen, Harald Oberhauser, and Michael A. Osborne. Adaptive batch sizes for active learning: A probabilistic numerics approach. In Sanjoy Dasgupta, Stephan Mandt, and Yingzhen Li (eds.), *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 496–504. PMLR, 02–04 May 2024. URL https://proceedings.mlr.press/v238/adachi24b.html.

Junyeob Baek, Mingyu Jo, Minsu Kim, Mengye Ren, Yoshua Bengio, and Sungjin Ahn. Generative recursive reasoning, 2026. URL https://arxiv.org/abs/2605.19376.

Wenxuan Bao, Ruxi Deng, Ruizhong Qiu, Tianxin Wei, Hanghang Tong, and Jingrui He. Latte: Collaborative test-time adaptation of vision-language models in federated learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 133–143, October 2025.

Edo Cohen-Karlik, Avichai Ben David, Nadav Cohen, and Amir Globerson. On the implicit bias of gradient descent for temporal extrapolation. In Gustau Camps-Valls, Francisco J. R. Ruiz, and Isabel Valera (eds.), *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pp. 10966–10981. PMLR, 28–30 Mar 2022. URL https://proceedings.mlr.press/v151/cohen-karlik22a.html.

Nik Dennler, Damien Drix, Tom PA Warner, Shavika Rastogi, Cecilia Della Casa, Tobias Ackels, Andreas T Schaefer, André van Schaik, and Michael Schmuker. High-speed odor sensing using miniaturized electronic nose. *Science Advances*, 10(45):eadp1764, 2024. DOI: 10.1126/sciadv.adp1764.

Shibhansh Dohare, J. Fernando Hernandez-Garcia, Qingfeng Lan, Parash Rahman, A. Rupam Mahmood, and Richard S. Sutton. Loss of plasticity in deep continual learning. *Nature*, 632 (8026):768–774, August 2024. ISSN 1476-4687. DOI: 10.1038/s41586-024-07711-7. URL https://doi.org/10.1038/s41586-024-07711-7.

Jay A. Farrell, John Murlis, Xuezhu Long, Wei Li, and Ring T Cardé. Filament-based atmospheric dispersion model to achieve short time-scale structure of odor plumes. *Environmental Fluid Mechanics*, 2:143–169, 2002. URL https://api.semanticscholar.org/CorpusID:17394044.

Lukas Fehring, Marius Lindauer, and Theresa Eimer. Growing with experience: Growing neural networks in deep reinforcement learning, 2025. URL https://arxiv.org/abs/2506.11706.

D. J. Foster, S. Kale, M. Mohri, and K. Sridharan. Parameter-free online learning via model selection. In *Advances in Neural Information Processing Systems*, pp. 6022–6032, 2017.

Kordel K. France. Scentience sigma. https://apps.apple.com/us/app/scentience-sigma/id6756134989, 2025. URL https://apps.apple.com/us/app/scentience/id6741092923. Accessed: 2026-03-18.

Kordel K. France and Ovidiu Daescu. Olfactory inertial odometry: Methodology for effective robot navigation by scent. In *2025 22nd International Conference on Ubiquitous Robots (UR)*, pp. 51–58, 2025. DOI: 10.1109/UR65550.2025.11078037.

Kordel K. France and John W. Sheppard. Factored particle swarm optimization for policy co-training in reinforcement learning. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '23, pp. 30–38, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400701191. DOI: 10.1145/3583131.3590376. URL https://doi.org/10.1145/3583131.3590376.

Kordel K. France, Ovidiu Daescu, Latifur Khan, and Rohith Peddi. Chasing ghosts: A simulation-to-real olfactory navigation stack with optional vision augmentation. *Sensors*, 26(9), 2026. ISSN 1424-8220. DOI: 10.3390/s26092849. URL https://www.mdpi.com/1424-8220/26/9/2849.

Jonathan Frankle and Michael Carbin. The lottery ticket hypothesis: Finding sparse, trainable neural networks, 2019. URL https://arxiv.org/abs/1803.03635.

Ariel Gordon, Elad Eban, Ofir Nachum, Bo Chen, Hao Wu, Tien-Ju Yang, and Edward Choi. Morphnet: Fast & simple resource-constrained structure learning of deep networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

Khurram Javed, Haseeb Shah, Richard S. Sutton, and Martha White. Scalable real-time recurrent learning using columnar-constructive networks. *Journal of Machine Learning Research*, 24(256): 1–34, 2023. URL http://jmlr.org/papers/v24/23-0367.html.

Alexia Jolicoeur-Martineau. Less is more: Recursive reasoning with tiny networks, 2025. URL https://arxiv.org/abs/2510.04871.

Andrej Karpathy. char-rnn. https://github.com/karpathy/char-rnn, 2015.

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, 2017. DOI: 10.1073/pnas.1611835114. URL https://www.pnas.org/doi/abs/10.1073/pnas.1611835114.

Mike Lasby, Anna Golubeva, Utku Evci, Mihai Nica, and Yani Ioannou. Dynamic sparse training with structured sparsity. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=kOBkxFRKTA.

Yann LeCun, John Denker, and Sara Solla. Optimal brain damage. In D. Touretzky (ed.), *Advances in Neural Information Processing Systems*, volume 2. Morgan-Kaufmann, 1989. URL https://proceedings.neurips.cc/paper_files/paper/1989/file/6c9882bbac1c7093bd25041881277658-Paper.pdf.

Michal Lewandowski, Hamid Eghbalzadeh, Bernhard Heinzl, Raphael Pisoni, and Bernhard A. Moser. On space folds of reLU neural networks. *Transactions on Machine Learning Research*, 2025. ISSN 2835-8856. URL https://openreview.net/forum?id=RfFqBXLDQk.

Yang Li and Shihao Ji. Neural plasticity networks, 2021. URL https://arxiv.org/abs/1908.08118.

Clare Lyle, Zeyu Zheng, Evgenii Nikishin, Bernardo Avila Pires, Razvan Pascanu, and Will Dabney. Understanding plasticity in neural networks. In *Proceedings of the 40th International Conference on Machine Learning*, ICML'23. JMLR.org, 2023.

Charles H. Martin and Michael W. Mahoney. Implicit self-regularization in deep neural networks: Evidence from random matrix theory and implications for learning. *Journal of Machine Learning Research*, 22(165):1–73, 2021. URL http://jmlr.org/papers/v22/20-410.html.

Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-Efficient Learning of Deep Networks from Decentralized Data. In Aarti Singh and Jerry Zhu (eds.), *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pp. 1273–1282. PMLR, 20–22 Apr 2017. URL https://proceedings.mlr.press/v54/mcmahan17a.html.

Seyed Iman Mirzadeh, Arslan Chaudhry, Dong Yin, Timothy Nguyen, Razvan Pascanu, Dilan Gorur, and Mehrdad Farajtabar. Architecture matters in continual learning, 2022. URL https://arxiv.org/abs/2202.00275.

Andrew Y. Ng, Daishi Harada, and Stuart J. Russell. Policy invariance under reward transformations: Theory and application to reward shaping. In *Proceedings of the Sixteenth International Conference on Machine Learning*, ICML '99, pp. 278–287, San Francisco, CA, USA, 1999. Morgan Kaufmann Publishers Inc. ISBN 1558606122.

Team Olmo, :, Allyson Ettinger, Amanda Bertsch, Bailey Kuehl, David Graham, David Heineman, Dirk Groeneveld, Faeze Brahman, Finbarr Timbers, Hamish Ivison, Jacob Morrison, Jake Poznanski, Kyle Lo, Luca Soldaini, Matt Jordan, Mayee Chen, Michael Noukhovitch, Nathan Lambert, Pete Walsh, Pradeep Dasigi, Robert Berry, Saumya Malik, Saurabh Shah, Scott Geng, Shane Arora, Shashank Gupta, Taira Anderson, Teng Xiao, Tyler Murray, Tyler Romero, Victoria Graf, Akari Asai, Akshita Bhagia, Alexander Wettig, Alisa Liu, Aman Rangapur, Chloe Anastasiades, Costa Huang, Dustin Schwenk, Harsh Trivedi, Ian Magnusson, Jaron Lochner, Jiacheng Liu, Lester James V. Miranda, Maarten Sap, Malia Morgan, Michael Schmitz, Michal Guerquin, Michael Wilson, Regan Huff, Ronan Le Bras, Rui Xin, Rulin Shao, Sam Skjonsberg, Shannon Zejiang Shen, Shuyue Stella Li, Tucker Wilde, Valentina Pyatkin, Will Merrill, Yapei Chang, Yuling Gu, Zhiyuan Zeng, Ashish Sabharwal, Luke Zettlemoyer, Pang Wei Koh, Ali Farhadi, Noah A. Smith, and Hannaneh Hajishirzi. Olmo 3, 2025. URL https://arxiv.org/abs/2512.13961.

OpenAI, :, Sandhini Agarwal, Lama Ahmad, Jason Ai, Sam Altman, Andy Applebaum, Edwin Arbus, Rahul K. Arora, Yu Bai, Bowen Baker, Haiming Bao, Boaz Barak, Ally Bennett, Tyler Bertao, Nivedita Brett, Eugene Brevdo, Greg Brockman, Sebastien Bubeck, Che Chang, Kai Chen, Mark Chen, Enoch Cheung, Aidan Clark, Dan Cook, Marat Dukhan, Casey Dvorak, Kevin Fives, Vlad Fomenko, Timur Garipov, Kristian Georgiev, Mia Glaese, Tarun Gogineni, Adam Goucher, Lukas Gross, Katia Gil Guzman, John Hallman, Jackie Hehir, Johannes Heidecke, Alec Helyar, Haitang Hu, Romain Huet, Jacob Huh, Saachi Jain, Zach Johnson, Chris Koch, Irina Kofman, Dominik Kundel, Jason Kwon, Volodymyr Kyrylov, Elaine Ya Le, Guillaume Leclerc, James Park Lennon, Scott Lessans, Mario Lezcano-Casado, Yuanzhi Li, Zhuohan Li, Ji Lin, Jordan Liss, Lily, Liu, Jiancheng Liu, Kevin Lu, Chris Lu, Zoran Martinovic, Lindsay McCallum, Josh McGrath, Scott McKinney, Aidan McLaughlin, Song Mei, Steve Mostovoy, Tong Mu, Gideon Myles, Alexander Neitz, Alex Nichol, Jakub Pachocki, Alex Paino, Dana Palmie, Ashley Pantuliano, Giambattista Parascandolo, Jongsoo Park, Leher Pathak, Carolina Paz, Ludovic Peran, Dmitry Pimenov, Michelle Pokrass, Elizabeth Proehl, Huida Qiu, Gaby Raila, Filippo Raso, Hongyu Ren, Kimmy Richardson, David Robinson, Bob Rotsted, Hadi Salman, Suvansh Sanjeev, Max Schwarzer, D. Sculley, Harshit Sikchi, Kendal Simon, Karan Singhal, Yang Song, Dane Stuckey, Zhiqing Sun, Philippe Tillet, Sam Toizer, Foivos Tsimpourlas, Nikhil Vyas, Eric Wallace, Xin Wang, Miles Wang, Olivia Watkins, Kevin Weil, Amy Wendling, Kevin Whinnery, Cedric Whitney, Hannah Wong, Lin Yang, Yu Yang, Michihiro Yasunaga, Kristen Ying, Wojciech Zaremba, Wenting Zhan, Cyril Zhang, Brian Zhang, Eddie Zhang, and Shengjia Zhao. gpt-oss-120b & gpt-oss-20b model card, 2025. URL https://arxiv.org/abs/2508.10925.

Jeffrey Pennington and Pratik Worah. Nonlinear random matrix theory for deep learning. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/0f3d014eead934bbdbacb62a01dc4831-Paper.pdf.

Andrei A. Rusu, Neil C. Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, and Raia Hadsell. Progressive neural networks. *CoRR*, abs/1606.04671, 2016. arXiv preprint.

Andrei A. Rusu, Neil C. Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, and Raia Hadsell. Progressive neural networks, 2022. URL https://arxiv.org/abs/1606.04671.

Scentience. Scentience app: World's first olfaction-vision language model. 2025. URL https://scentience.ai/news/f/scentience-app-worlds-first-olfaction-vision-language-model. Accessed: 2025-03-02.

S. S. Schoenholz, J. Gilmer, S. Ganguli, and J Sohl-Dickstein. Deep information propagation, November 2016.

PhD Sebastian Raschka. From GPT-2 to gpt-oss: Analyzing the Architectural Advances — magazine.sebastianraschka.com. https://magazine.sebastianraschka.com/p/from-gpt-2-to-gpt-oss-analyzing-the. [Accessed 08-02-2026].

Terence Tao. Topics in random matrix theory, 2012.

Jaehong Yoon, Eunho Yang, Jeongtae Lee, and Sung Ju Hwang. Lifelong learning with dynamically expandable networks. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=Sk7KsfW0-.