Phased Pruning in Neural Networks Recapitulates Selectivity–Fragility Trade-Offs in Brain Development

Authors:

Ngo Cheung, FHKAM(Psychiatry)

Affiliations:

[1] Independent Researcher

Corresponding Author:

Ngo Cheung, MBBS, FHKAM(Psychiatry)

Hong Kong SAR, China

Tel: 98768323

Email: info@cheungngomedical.com

Conflict of Interest: None declared.

Funding Declaration: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics Declaration: Not applicable.

Abstract

Synaptic pruning refines neural circuits during development, but altered trajectories are implicated in neurodevelopmental disorders. Here, we use a task-gated neural network to show that the timing of pruning critically determines functional outcomes. Networks subjected to aggressive pruning only after an initial period of high connectivity develop enhanced resistance to interfering inputs—achieving near-perfect performance under conflicting task demands—yet exhibit marked fragility to internal noise. In contrast, moderate pruning preserves robustness but allows greater interference. These trade-offs emerge prominently when late pruning follows early overgrowth, providing a buffer for refinement into highly selective circuits. The findings reveal developmental phasing as a key driver of circuit specialization versus resilience, offering a mechanistic explanation for phenotypic duality in disorders such as autism spectrum disorder (where intense focus co-occurs with sensory fragility) and highlighting principles relevant to efficient sparse neural networks.

Introduction

During early life, the brain generates a surplus of synaptic connections; later, many of these links are removed so that efficient, finely tuned circuits emerge. Classic quantitative work in human cortex showed that roughly half of all excitatory synapses formed in infancy disappear over childhood and adolescence [1,2]. This selective loss, known as synaptic pruning, depends on neural activity and is essential for normal cognition [3].

Autism spectrum disorder (ASD) has long been associated with alterations in this pruning program. Post-mortem and imaging studies report that young children on the spectrum often retain an excess of synapses, whereas many adults show the opposite pattern, with reduced synaptic markers and

weakened long-range connectivity [4,5,6,7]. One interpretation is a two-phase disturbance: pruning lags during early development, then overshoots during adolescence [3,8].

Computational experiments offer a way to test how timing and magnitude of pruning shape circuit function. Recent machine-learning work demonstrates that dense neural networks contain small "winning tickets" whose lucky initial weights let them learn as well as the full model after pruning [9]. Although conceived for efficiency, the method resembles developmental refinement—initial overconnectivity followed by selective elimination—yet few studies have asked how different pruning schedules influence network behavior in tasks that mimic real-world sensory demands.

Here we simulate developmental trajectories with task-gated neural networks. First, models grow with varied initial synaptic densities; later, they undergo moderate (typical) or aggressive (ASD-like) pruning. We test whether heavy late pruning after early overgrowth can yield the blend of strengths—such as reduced interference—and weaknesses—such as fragility to perturbation—reported in autism. By linking molecular insights on faulty pruning to observable network performance, the work aims to bridge biology and computation.

Methods

Model architecture and task design

We built a feed-forward network that mirrors early brain over-connectivity and later pruning (Figure 1). The model is a multilayer perceptron with three hidden layers (256, 256 and 128 ReLU units) followed by a four-way classifier. Inputs combine two sensory features with a two-unit cue that tells the network which rule to apply. A cue of [1 0] selects the "visual" rule; a cue of [0 1] selects the

"voice" rule. This arrangement resembles cognitive-flexibility models that switch behaviour on the basis of context [10].

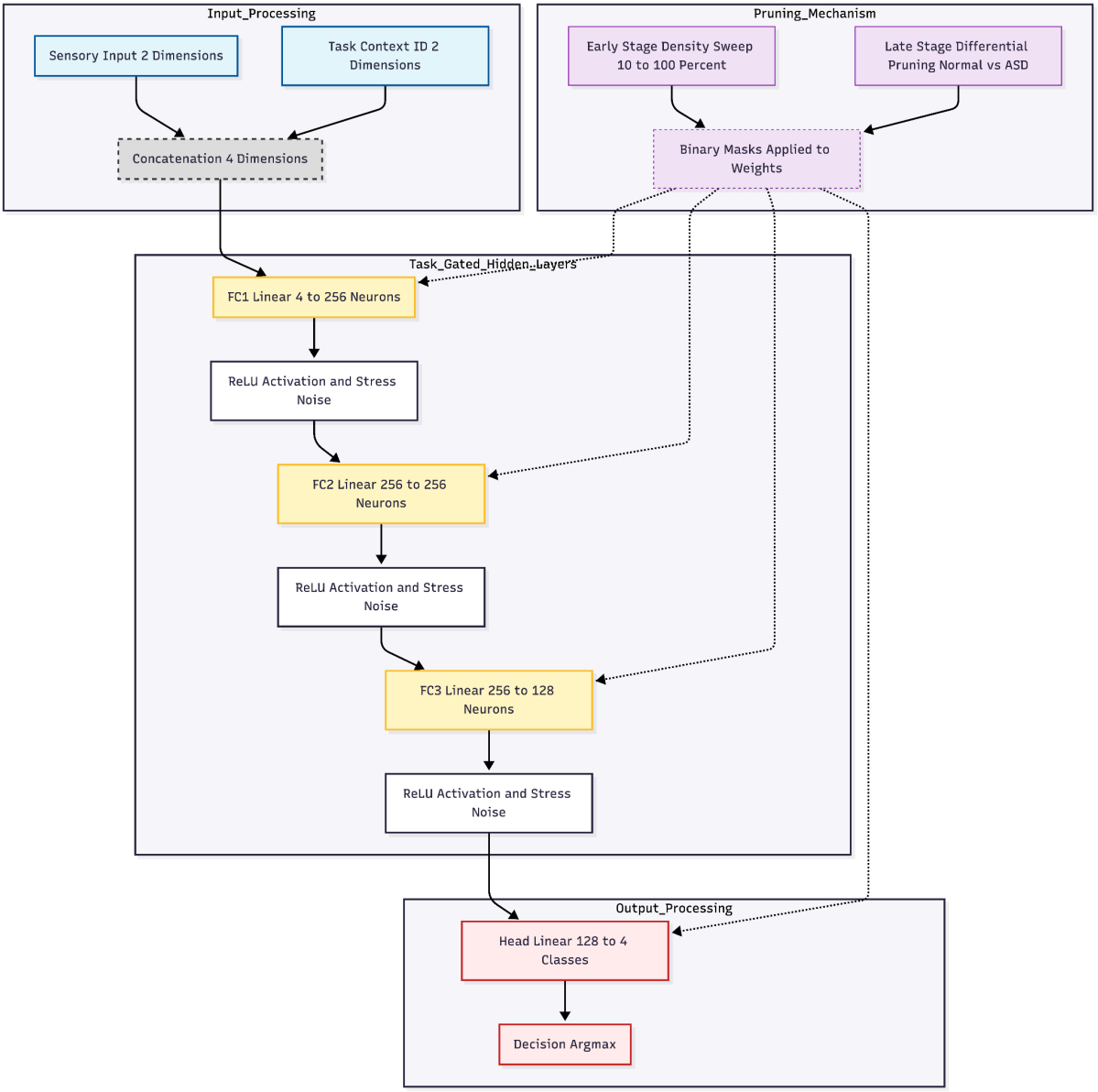

Figure 1: *Task-Gated Network Architecture and Developmental Pruning Framework. The model processes a 4-dimensional input vector combining sensory data and task context. Information propagates through three fully connected hidden layers with 256, 256, and 128 neurons respectively. ReLU activation functions are followed by a noise injection step to simulate neural stress. The TaskGatedPruningManager controls network connectivity via binary masks. The experimental design involves an Early Stage density sweep followed by a Late Stage differential pruning phase, where the Normal condition removes 20 percent of remaining weights and the ASD condition removes 50 percent.*

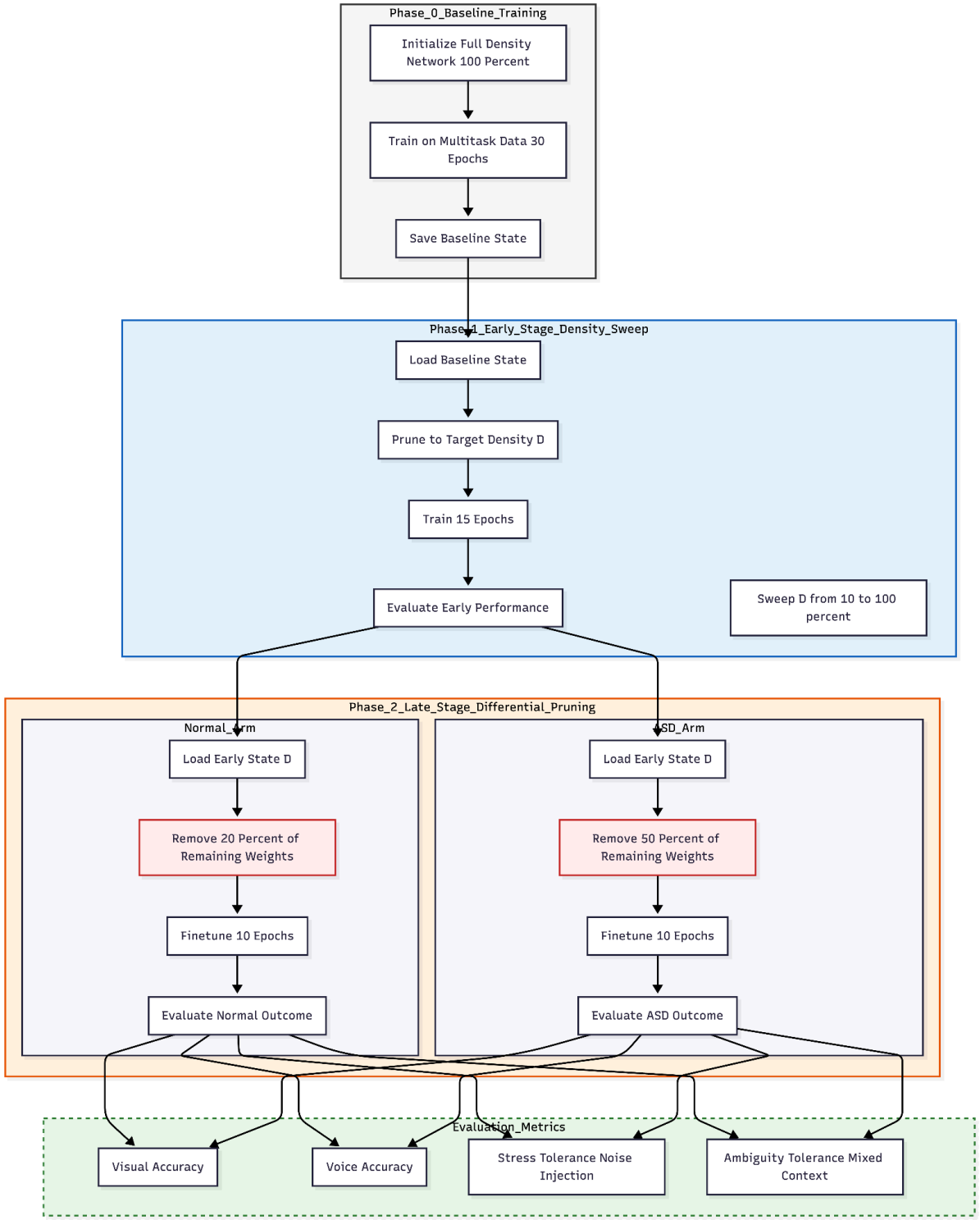

Figure 2: *Experimental Design for Developmental Pruning Simulation. The study follows a three-phase progression. Phase 0 establishes a fully connected baseline trained on multitask data. Phase 1 performs a density sweep, creating distinct models with densities ranging from 10% to 100% to simulate varying degrees of early development. Phase 2 branches into two conditions for each density level: a 'Normal' trajectory removing an additional 20% of active weights, and an 'ASD' trajectory removing an additional 50% of active weights. Both arms undergo identical fine-tuning before final evaluation across four key metrics: visual accuracy, voice accuracy, stress tolerance, and ambiguity tolerance.*

Synthetic data were drawn from four Gaussian clusters centred at (−3, −3), (3, 3), (−3, 3) and (3, −3) with noise $\sigma = 0.8$. In the visual task each cluster maps to its own label, whereas in the voice task the labels are permuted (2, 3, 0, 1), forcing the same stimulus to demand different responses depending on the cue. Training sets contained 12 000 samples per task and test sets 4 000. All code ran in PyTorch on a fixed-seed CPU environment to ensure repeatability.

Pruning procedure

To emulate synaptic elimination we adopted global magnitude pruning, masking the smallest-weight connections while leaving biases intact. Each weight matrix carries a binary mask that persists throughout training. Two modes were used. First, a global density target removed weights until only a chosen proportion remained. Second, fractional pruning discarded a fixed share of whatever weights were still present—20 % for the "normal" trajectory and 50 % for the "ASD-like" trajectory. Masks were reapplied after every optimiser step so that sparsity never lapsed.

Training protocol

A dense baseline network (≈100 000 parameters) was trained jointly on both tasks for 30 epochs with Adam (learning rate 0.001) and cross-entropy loss. Batches alternated between visual and voice data, each carrying the proper cue. After baseline training reached near-perfect accuracy, the model's weights served as the starting point for all pruning experiments. Fine-tuning at fixed sparsity used Adam at 0.0005 while honouring the masks.

Experimental design

The study unfolded in three phases (Figure 2).

Phase 0: baseline. The dense model was trained as described above.

Phase 1: early pruning sweep. The dense baseline was pruned to 10 densities (from 10 % to 100 % in 10-point steps). Each sparsified network was then fine-tuned for 15 epochs, representing different degrees of early synapse loss or retention.

Phase 2: late differential pruning. Every early-stage network branched into two versions. One underwent a further 20 % weight reduction (normal trajectory); the other lost 50 % (ASD-like trajectory). Both branches were fine-tuned for 10 additional epochs. This contrast modelled moderate adolescent refinement versus excessive elimination, complementing biological reports of pruning imbalance in autism [5].

Evaluation metrics

Performance was judged on four fronts.

- Core accuracy. Classification success on held-out visual and voice tests using the correct cues.
- Stress tolerance. Gaussian noise (σ = 0.5, 1.0, 2.0, 3.0) was injected into hidden activations during the visual task to gauge robustness [11].
- Ambiguity tolerance. Visual-task accuracy was measured while gradually blending the cue toward [0.5 0.5], exposing interference between rules.
- Sensory integration. Mixed visual-voice inputs were classified under clear or ambiguous cues to probe multisensory handling.
- Statistical comparisons focused on differences between the normal and ASD-like branches at each early density. Together, these procedures link developmental timing of pruning to functional outcomes in a controlled computational setting.

Results

Early-stage pruning effects

The dense baseline achieved ceiling-level accuracy on both visual and voice rules and withstood hidden-layer noise (about 84 percent correct at the highest noise level) (Table 1). When we removed weights before adolescence, visual performance remained essentially perfect at every sparsity tested, whereas the voice rule proved more sensitive. At the most aggressive early cut (10 percent of weights kept) voice accuracy fell to chance (roughly 50 percent) but recovered to 100 percent once at least one-fifth of the original connections were present.

Table 1. *Early-Stage Performance Across Initial Densities*

Density (%)	Visual Acc (%)	Voice Acc (%)	Stress Tolerance (High Noise, %)	Ambiguity Tolerance (w=0.50, %)	Blended + Ambiguous (%)
10	100.0	50.0	45.0	100.0	56.5
20	100.0	100.0	53.5	69.3	46.9
30	100.0	100.0	66.1	75.2	47.0
40	100.0	100.0	72.8	75.5	48.5
50	100.0	100.0	76.7	82.2	50.6
60	100.0	100.0	78.5	50.8	40.3
70	100.0	100.0	80.0	49.1	37.7
80	100.0	100.0	81.9	50.2	38.5
90	99.9	100.0	84.7	75.8	48.6
100	100.0	100.0	83.7	53.2	43.1

Noise robustness grew steadily with added connections, rising from 45 percent at 10 percent density to nearly 85 percent at 90 percent density and then dipping slightly in the fully dense model. The ability to cope with an ambiguous cue (half visual, half voice) followed a U-shape. Extreme sparsity or full density produced high tolerance to cue conflict, whereas intermediate densities (about 60–80 percent of weights) yielded the poorest ambiguous performance, near 50 percent correct. The same

mid-range models also struggled with blended sensory inputs, dipping to about 38–40 percent accuracy, while the sparsest and densest networks fared better (around mid-40s to mid-50s).

Late-stage differential pruning outcomes

Applying a second round of pruning after fine-tuning exposed distinct trade-offs. Visual accuracy stayed at or near 100 percent regardless of condition, so secondary measures were more informative. Across almost all starting densities the normally pruned branch—losing 20 percent of its remaining weights—handled hidden noise better, often by 5–19 percentage points. In contrast, the aggressively pruned branch—losing 50 percent—excelled at resisting cue ambiguity when its early density had been high. For example, beginning at 40 percent density then pruning hard in adolescence raised ambiguous-cue accuracy by more than 50 points compared with the normal branch; starting at 90 percent density produced a swing of roughly 73 points in favour of the aggressive cut.

Table 2. *Late-Stage Normal vs. ASD-like Trajectories by Early Density*

Early Density (%)	Normal Final Density (%)	ASD Final Density (%)	Visual Acc Normal (%)	Visual Acc ASD (%)	Stress Tolerance Diff (Normal – ASD, %)	Ambiguity Tolerance Diff (Normal – ASD, %)
10	8.0	5.0	100.0	100.0	-4.2	+0.1
20	16.0	10.0	100.0	100.0	+9.5	+13.1
30	24.0	15.0	100.0	100.0	+19.1	+8.4
40	32.0	20.0	100.0	100.0	+15.9	+55.2
50	40.0	25.0	100.0	100.0	+12.8	+36.1
60	48.0	30.0	100.0	100.0	+8.5	-21.4
70	56.0	35.0	100.0	100.0	+12.1	+11.9
80	64.0	40.0	100.0	100.0	+11.8	+1.7
90	72.0	45.0	100.0	100.0	+4.7	-73.3
100	80.0	50.0	99.9	100.0	+4.8	-25.7

When early density was low (10–20 percent) the two late-pruning strategies converged: stress tolerance differences shrank and ambiguity scores were nearly identical. Thus, large initial networks

that later lost many connections became highly selective, ignoring conflicting cues but proving fragile to internal noise. More modest late pruning preserved redundancy, supporting noise resilience at the price of greater interference when cues were mixed.

Discussion

Interpreting the trade-offs in pruning dynamics

The simulations reveal a striking balance between selectivity and resilience once adolescent pruning begins. Networks that started with abundant early connections and then lost half of their remaining weights behaved much like a finely tuned filter: they withstood cue conflicts almost effortlessly, maintaining close to 80 percent accuracy when the rule signal was fully ambiguous. Yet that refinement came at a price. As soon as hidden-layer noise was introduced, performance deteriorated far more than in moderately pruned counterparts. A similar duality is often noted clinically. Individuals on the autism spectrum can show impressive focus and immunity to distraction, but they may also become overwhelmed when environments turn chaotic or unpredictable [3,12].

The networks that shed only a fifth of their remaining weights offered the opposite profile. They retained redundancy that buffered them against internal noise but allowed more interference when cues conflicted. This pattern resembles typical development, in which flexible integration is valued even if it occasionally lets noise leak through. Notably, the most dramatic divergence between the two pruning schedules emerged only when the system began with high early density. If early connectivity was already sparse, both late-stage strategies converged on fragility, implying that an initial surplus may be a prerequisite for later over-pruning to yield highly specialized—rather than simply impaired—circuits. This echoes experimental evidence that early synaptic overgrowth followed by excessive adolescent elimination can underlie autistic phenotypes [5,8].

The results as a whole support a two-phase view of autism. Hyper-connectivity during childhood, possibly influenced by mTOR signaling or other growth-promoting pathways, results in an unstable circuit landscape. During adolescence, microglia and associated mechanisms seem to overcompensate, removing long-range connections and creating a sparse architecture that is good at isolating task-relevant inputs but not so good at dealing with noise and change [7]. Timing therefore matters: pruning that occurs after sufficient overgrowth can sharpen function, whereas equivalent pruning applied earlier risks outright failure. Recent longitudinal imaging supports this view, documenting early cortical thickening that later gives way to accelerated thinning in many autistic brains [8].

Finally, the model highlights why a single label such as "under-pruned" or "over-pruned" is too coarse. Depending on region and developmental window, the same brain may pass through both states. Grasping these temporal subtleties may guide interventions, potentially redirecting emphasis towards mitigating early overgrowth or moderating later elimination [13].

Reconciling simulation findings with polygenic evidence

The contrasting behaviours that appeared in our network—high selectivity but poor noise tolerance after aggressive late pruning versus broader resilience but greater cross-talk after moderate pruning—track closely with the pathway-level genetics now emerging for autism spectrum disorder (ASD) and attention-deficit/hyperactivity disorder (ADHD) (Figure 3). Cheung's cross-disorder GWAS dissection (Cheung, 2026) shows that common ASD risk is concentrated where glutamatergic signalling and pruning genes intersect, whereas ADHD risk is linked to pruning genes that sit largely outside excitatory pathways [14]. In the simulations, an early period of dense connectivity created the substrate for an "over-pruning" phase that cut interference to near zero yet left the model fragile whenever random noise was introduced. That same trade-off—clarity versus adaptability—fits the

idea that hyper-excitation in ASD invites microglia to remove too many synapses, producing circuits that excel at filtering but struggle with change.

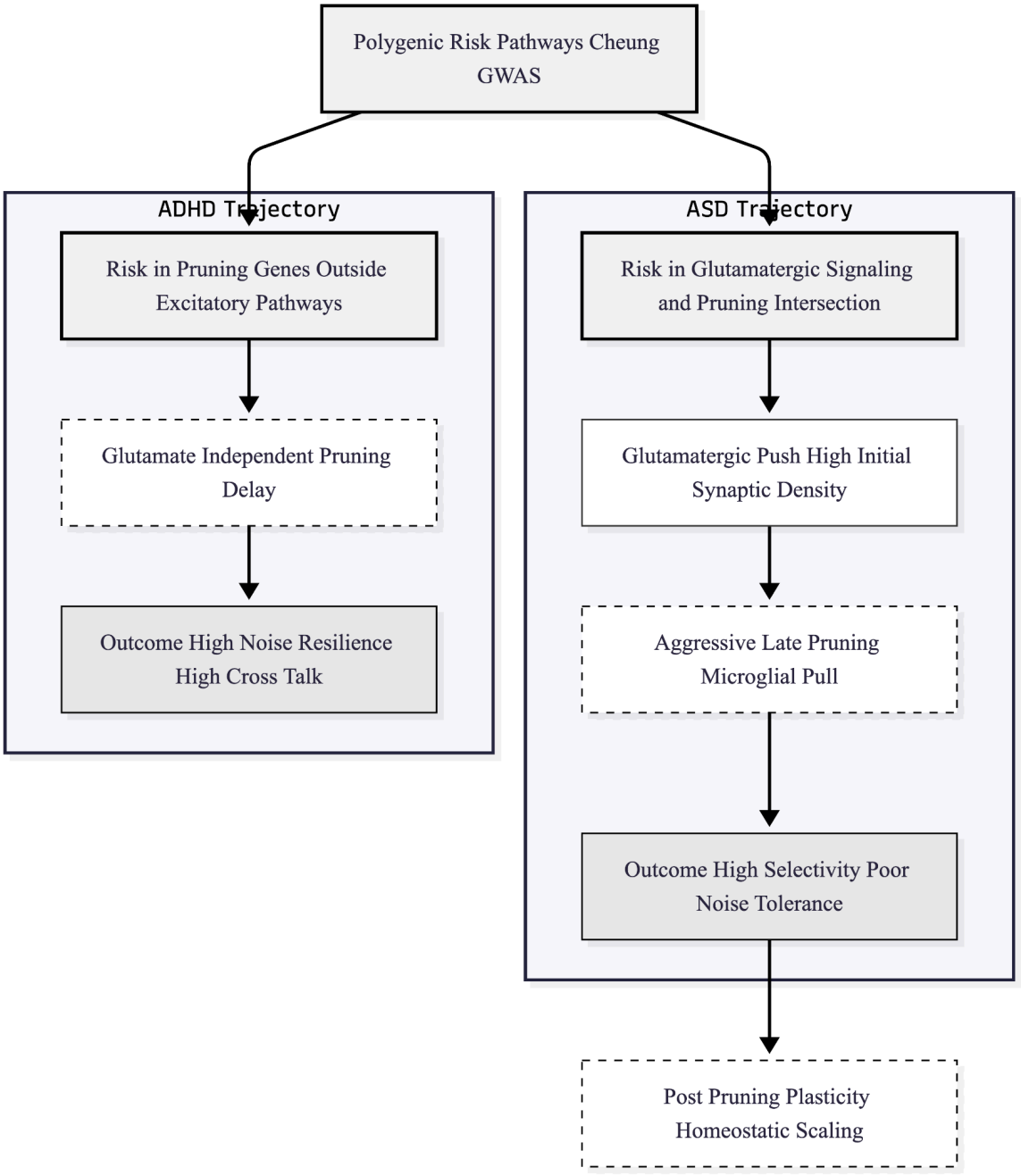

Figure 3. *Divergent pruning trajectories informed by polygenic risk pathways. The model synthesizes simulation results with cross-disorder GWAS evidence. Left: ASD risk concentrates at the intersection of glutamatergic signaling and pruning genes. This generates a "push-pull" dynamic where initial hyper-excitation and density are followed by aggressive elimination, resulting in specialized circuits that filter interference well but are fragile to hidden-layer noise. Post-pruning plasticity allows for the recovery of primary accuracy. Right: ADHD risk involves glutamate-independent pruning genes, leading to a delay or moderation in pruning. These networks retain redundancy, offering resilience against internal noise but suffering from cross-talk and distractibility when cues conflict.*

On the other hand, networks that only lost a few links kept redundant links, were able to handle noise in the hidden layer, but got more confused when task cues didn't match. This profile echoes the glutamate-independent pruning delay proposed for ADHD, where immature, over-connected networks foster distractibility rather than rigid focus. The fact that the largest functional split between the two pruning schedules appeared only when initial density was high supports a sequential model: a glutamatergic "push" first expands the synaptic pool in ASD, and only then can an aggressive pruning "pull" carve out the specialised, low-interference architecture seen later in development.

A caveat concerns post-pruning recovery. Our most severe condition still permitted fine-tuning, allowing the model to regain near-perfect primary accuracy despite heavy loss of weights. Real brains may rely on similar but biologically limited forms of plasticity—homeostatic scaling, dendritic spine turnover—to stabilise function after over-elimination. This continuing adaptation could help explain why many autistic individuals show strong, domain-specific skills alongside marked sensitivity to sensory variability.

Novelty and broader implications

Combining an early-overgrowth phase with two distinct late-pruning schedules allowed us to reproduce a clinical paradox that has been hard to capture in silico: heightened resistance to distraction paired with marked vulnerability to sensory stress (Figure 4). In the model, dense networks that later lost half of their surviving weights became masters at ignoring contradictory cues—often exceeding 80 percent accuracy under full ambiguity—yet they stumbled once random noise entered the hidden layers. This mirrors the observation that many autistic individuals can focus intensely on a single stream of information while finding unpredictable environments overwhelming [3,12].

Importantly, these selective-but-fragile circuits emerged only when pruning followed an early period of exuberant connectivity. If the starting density was already low, both mild and severe pruning

produced similar fragility, suggesting that an initial surplus may act as a buffer that lets aggressive elimination carve out highly specialised pathways rather than simply degrading performance. This finding dovetails with post-mortem and imaging evidence showing early synaptic surplus followed by later hypo-connectivity in autism [5,8,7].

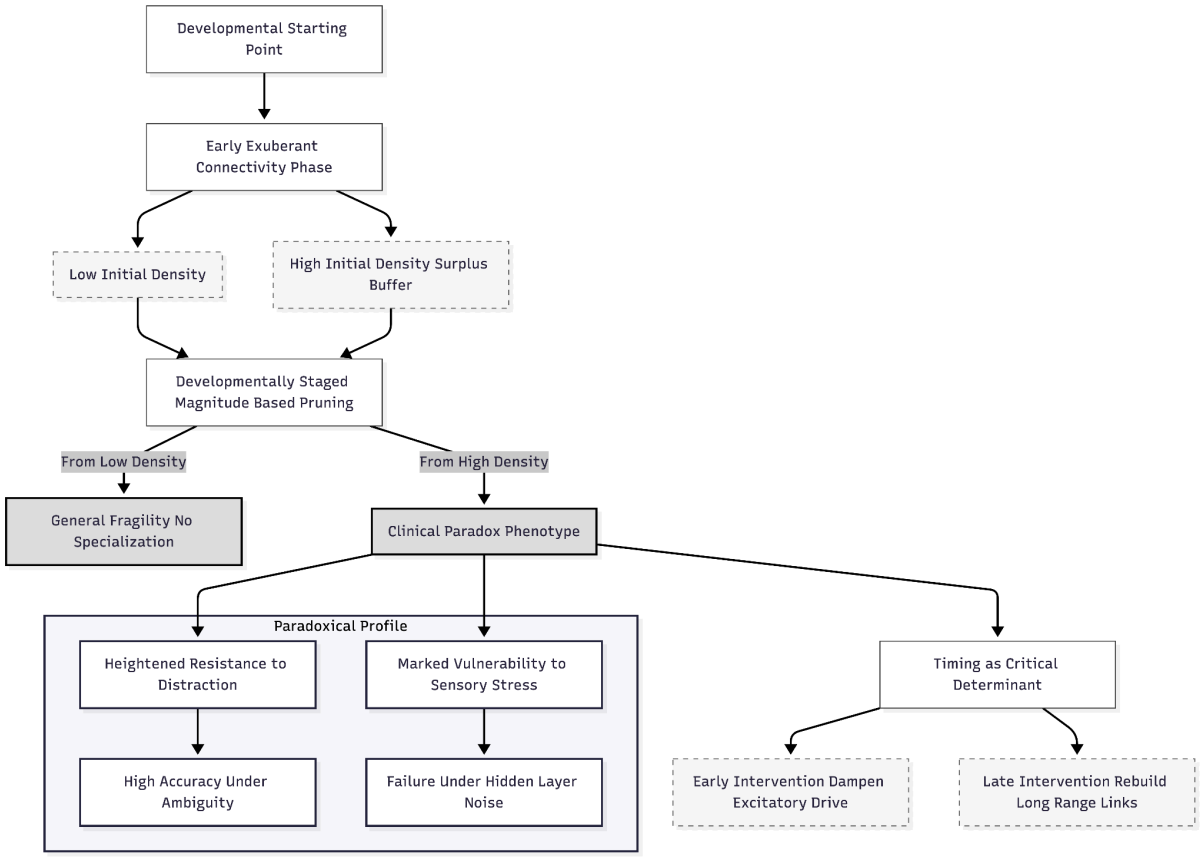

Figure 4. *Developmental trajectory of synaptic pruning and resulting phenotypes. The model demonstrates that the timing and initial magnitude of connectivity are critical determinants of network behavior. Top: An initial phase of exuberant connectivity (High Initial Density) acts as a necessary buffer. When followed by magnitude-based pruning, this surplus allows for the emergence of a "clinical paradox" phenotype observed in autism: networks become highly specialized and resistant to contradictory cues (distraction) but remain fragile to random noise (sensory stress). Bottom: In contrast, starting with low density leads to generalized fragility regardless of pruning severity. The findings suggest age-specific therapeutic targets: early interventions might focus on dampening excitatory drive to prevent over-elimination, while post-adolescent strategies might aim to rebuild connectivity without re-introducing interference.*

Beyond autism, the results emphasise timing as a critical determinant of pruning outcome. Age-specific interventions may therefore need to differ: dampening early excitatory drive could prevent later over-elimination, whereas post-adolescent therapies might aim to rebuild long-range

links without re-introducing excessive interference [13]. More generally, the study shows that even a simple magnitude-based pruning algorithm, applied in a developmentally staged manner, can recapitulate complex neurodevelopmental phenotypes—opening the door to similar explorations in other disorders of circuit refinement.

Limitations

The present network is intentionally compact and purely feed-forward; it lacks recurrence, layered hierarchies and biologically realistic plasticity rules found in real cortex. The Gaussian-blob stimuli create a clear conflict between two rules, but they can't copy the high-dimensional, multi-modal inputs that we come across in daily life. Pruning was exclusively influenced by weight magnitude, disregarding activity-dependent or immune-mediated mechanisms currently recognized to regulate synaptic elimination in vivo [3]. They also didn't take into account region-specific trajectories; human data show that association and sensory areas prune at different times [8]. Finally, each pruned model received a period of fine-tuning, a luxury that biological circuits may not enjoy to the same extent, potentially inflating our estimates of functional recovery.

Conclusion

Taken together, these simulations lend computational support to a compensatory over-pruning view of autism: early synaptic excess, followed by intense adolescent elimination, yields circuits that excel at filtering competing inputs yet flounder when noise rises or flexibility is required. The strong dependence on initial density underscores developmental timing as a pivotal variable, matching longitudinal evidence of phased dysregulation from infancy to adulthood [13,7]. Although modest in scale, the model demonstrates how staged pruning can generate both adaptive specialisation and hidden costs—providing a concise bridge between molecular risk factors, circuit structure and behavioural outcome. Scaling up to recurrent, sensory-rich networks with region-specific rules and incorporating glial or complement pathways represent logical next steps.

References

[1] Huttenlocher, P. R. Synaptic density in human frontal cortex—developmental changes and effects of aging. Brain Research. 1979;163:195–205.

[2] Rakic, P., et al. Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. Science. 1986;232:232-235.

[3] Faust, T. E., et al. Mechanisms governing activity-dependent synaptic pruning in the developing mammalian CNS. Nature Reviews Neuroscience. 2021;22:657–673.

[4] Pardo, C. A., & Eberhart, C. G. The neurobiology of autism. Brain Pathology. 2007;17:434–447.

[5] Tang, G., et al. Loss of mTOR-dependent macroautophagy causes autistic-like synaptic pruning deficits. Neuron. 2014;83:1131–1143.

[6] Bourgeron, T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. Nature Reviews Neuroscience. 2015;16:551–563.

[7] Matuskey, D., et al. 11C-UCB-J PET imaging is consistent with lower synaptic density in autistic adults. Molecular Psychiatry. 2025;30:1610-1616.

[8] Zielinski, B. A., et al. Longitudinal changes in cortical thickness in autism and typical development. Brain. 2014;137:1799–1812.

[9] Frankle, J., & Carbin, M. The lottery ticket hypothesis: Finding sparse, trainable neural networks. International Conference on Learning Representations. 2019.

[10] Herd, S. A., et al. A neural network model of individual differences in task-switching abilities. Neuropsychologia. 2014;62:375–389.

[11] Bishop, C. M. Training with noise is equivalent to Tikhonov regularization. Neural Computation. 1995;7:108–116.

[12] Rubenstein, J. L. R., & Merzenich, M. M. Model of autism: Increased ratio of excitation to inhibition in key neural systems. Genes, Brain and Behavior. 2003;2:255–267.

[13] Cheung, N. Timing is everything: Why the Cheung glutamatergic regimen is contraindicated in young children with autism but promising after puberty. Preprints. 2025.

[14] Cheung, N. Polygenic dissection of synaptic pruning and glutamatergic signaling: Contrasting mechanisms in ASD and ADHD. Preprints. 2026.