ECTO Design Notes (SATSA & Dental Cohorts)

Anderson M. Rodriguez

October 7, 2025

1 Purpose & Scope

These notes document the design rationale behind the ECTO pipeline: converting longitudinal psychometric item distributions into an information signal (Shannon entropy, in bits), using that only to *seed initial conditions*, and then evolving a minimal autonomous ODE for two trait states N, P coupled to a cumulative stress/constraint state E_{stress}. The aim is a small, interpretable, falsifiable *phenomenological* law at the cohort level.

2 Data → Information → State

1. **Inputs:** Wave-wise Likert distributions for selected items (per cohort).
2. **Entropy extraction:** For each wave, compute Shannon entropy H (bits) of the item distribution. Within each dataset, min–max normalize to obtain $H^* \in [0, 1]$.
3. **Initialization (seeding only):** Use $H^*(t_0)$ to set $N(t_0) = N_0(H^*)$, $P(t_0) = P_0(H^*)$; set $E_{\text{stress}}(t_0) = E_0$. No time-varying inputs are used in the main ALife runs; the ODE then evolves *autonomously*.

3 Time & Units

Integration is in calendar time (years) at the observed timestamps. States N, P, E_{stress} and H^* are dimensionless. Rate parameters μ, γ have units year^{-1}. Parameters $\alpha, \beta, c_1, c_2, c_3, G, K$ are dimensionless. In the main runs we set $G = 1$ for comparability; K is tuned/fitted per run.

4 Dynamical Law (ECTO ODE)

$$\frac{dN}{dt} = \mu N - (\alpha N + \beta^2 P)N, \qquad \frac{dP}{dt} = \mu P - \beta P\left(\frac{c_1 P + c_2 N + c_3 E_{\text{stress}}}{G}\right),$$
$$\frac{dE_{\text{stress}}}{dt} = \gamma\, E_{\text{stress}}\left(\frac{N}{N+K}\right).$$

Notes: Expanding makes the interaction structure explicit:

$$\dot{N} = \mu N - \alpha N^2 - \beta^2 PN, \quad \dot{P} = \mu P - \tfrac{\beta c_1}{G} P^2 - \tfrac{\beta c_2}{G} NP - \tfrac{\beta c_3}{G} E_{\text{stress}} P.$$

5 Parameter Roles (phenomenological)

- μ: innovation/influx baseline sustaining trait expression.
- α: self-limiting constraint on N (logistic-like).
- β: cross-trait coupling scale (squared in $\dot{N}$ to keep damping nonnegative).
- c_1, c_2: within-/cross-trait cost weights affecting P.
- c_3: stress sensitivity of P.
- G: capacity/scale (set G=1 in main runs).
- γ: stress amplification (growth) rate.
- K: saturation for the $N/(N+K)$ kernel (higher K = slower stress sensitivity onset).

6 Interpretation (eco-evo analogues)

Labels such as "pleiotropy," "selection/constraint," and "environmental feedback" are cohort-level *phenomenological analogues* rather than biochemical mechanisms. The terms encode: self-limits (N^2, P^2), antagonistic trade-offs (NP), and a cumulative constraint/stress integrator E_{stress} with saturating sensitivity to N.

7 Well-posedness & Invariants (brief)

On the nonnegative orthant $N, P, E_{\text{stress}} \geq 0$, the vector field is locally Lipschitz; solutions exist and are unique on finite intervals. With nonnegative initial conditions, N, P remain nonnegative; E_{stress} is monotone nondecreasing with instantaneous rate bounded by γE_{stress}. (A relaxing variant with $-\lambda E_{\text{stress}}$ is a planned ablation, not used in the ALife runs.)

8 Validation Protocol

1. **Fit/Calibrate:** Choose/fix parameters to reproduce trajectory *shape* on a subset of waves.
2. **Held-out (LOO):** Hold out one wave, fit on the rest, report held-out RMSE/R^2.
3. **Artifacts:** Save metrics (`metrics_*.csv`, `loo_*.csv`) and figures (`fig_*.png`); print Python/NumPy/SciP versions.

9 Ablations & Sensitivity (planned)

- **Information measure swap:** Shannon → KL/Fisher/Algorithmic; compare held-out error.
- **Stress relaxation:** add $-\lambda E_{\text{stress}}$; test recovery dynamics.
- **Kernel variants:** replace $N/(N+K)$ by alternative saturating kernels.
- **Parameter sensitivity/identifiability:** CIs via profile likelihood or bootstrap; Sobol/PRCC.
- **Baselines:** ARIMA/VAR, LV-like systems fit to same normalized trajectories.

10 Reproducibility

Runs are deterministic (no stochastic forcing). Use tight solver tolerances and record environment:

`Python X.Y.Z | NumPy a.b.c | SciPy d.e.f`

Entropies are min–max normalized *within each dataset*; Dental starts near 0 and SATSA near 1 by design to compare *dynamics*, not absolute scale.

11 At-a-Glance Pitch for Collaboration:

One scalar (bits) → tiny autonomous ODE → held-out generalization. Bring any longitudinal distributions (wave-wise histograms); we will seed initial state with H^*, evolve the ODE, and report held-out error.