

Supplementary information S5 (box): Policies and cost

This box describes a scheme that ensures agents are attracted to locations in state-space, using prior expectations about the motion of hidden states; $\tilde{x}(t) = [x, x']^T \in X$ comprising position and velocity. This formulation of how an ensemble density can be restricted to an attractive subset of state-space $A \subset X$ rests on the Fokker-Planck description (see Frank 2004) of how the density changes with time

$$\dot{p}(\tilde{x} | m) = \gamma \nabla^2 p - p \nabla \cdot f - f \cdot \nabla p$$

At equilibrium, $\dot{p}(\tilde{x} | m) = 0$ and

$$p(\tilde{x} | m) = \frac{\gamma \nabla^2 p - f \cdot \nabla p}{\nabla \cdot f} \quad \text{S5.1}$$

Notice that as the divergence $\nabla \cdot f$ increases, the sojourn time (i.e., the proportion of time a state is occupied) falls. Crucially, at the peaks of the ensemble density, the gradient is zero and its curvature is negative, which means the divergence must be negative (from Equation S5.1)

$$\left. \begin{array}{l} p > 0 \\ \nabla p = 0 \\ \nabla^2 p < 0 \end{array} \right\} \Rightarrow \nabla \cdot f < 0 \quad \text{S5.2}$$

This provides a simple and general mechanism to ensure peaks of the ensemble density lie in, and only in $A \subset X$. This is assured if $\nabla \cdot f(\tilde{x}) < 0$ when $\tilde{x} \in A$ and $\nabla \cdot f(\tilde{x}) \geq 0$ otherwise. We can exploit this using the generic equations of motion

$$f = \begin{bmatrix} x' \\ cx' - \partial_x \varphi \end{bmatrix} \Rightarrow \nabla \cdot f = c \quad \text{S5.3}$$

This flow describes the Newtonian motion of a unit mass in a potential energy well $\varphi(x, \theta)$, where cost plays the role of negative dissipation or friction. Crucially, under this policy or flow, divergence is simply cost; meaning the associated ensemble density can only have maxima in regions of negative cost. This provides a means to specify attractive regions $A \subset X$ by assigning them negative cost

$$\begin{aligned} c(x) &\leq 0 : x \in A \\ c(x) &> 0 : x \notin A \end{aligned} \quad \text{S5.4}$$

Put simply, this scheme ensures that agents are expelled from high-cost regions of state-space and get 'stuck' in attractive regions.

In summary, the previous supplementary information box (S4) showed that any flow can be described in terms of a scalar value-function (and vector potential W), from which an implicit cost-function can be derived. In this box (S5), we have addressed the inverse problem of how cost can be used to constrain flow, ensuring that it leads to attractive, low-cost states. The ensuing policy or flow can be used in a generative model of flow or state-transitions to provide predictions that action fulfils, under the free-energy principle. A full discussion of these and related ideas will be presented in Friston et al (in preparation).

Reference

Frank TD (2004). *Nonlinear Fokker-Planck Equations: Fundamentals and Applications*. Springer Series in Synergetics (Berlin: Springer)