Expected free energy as an information constraint

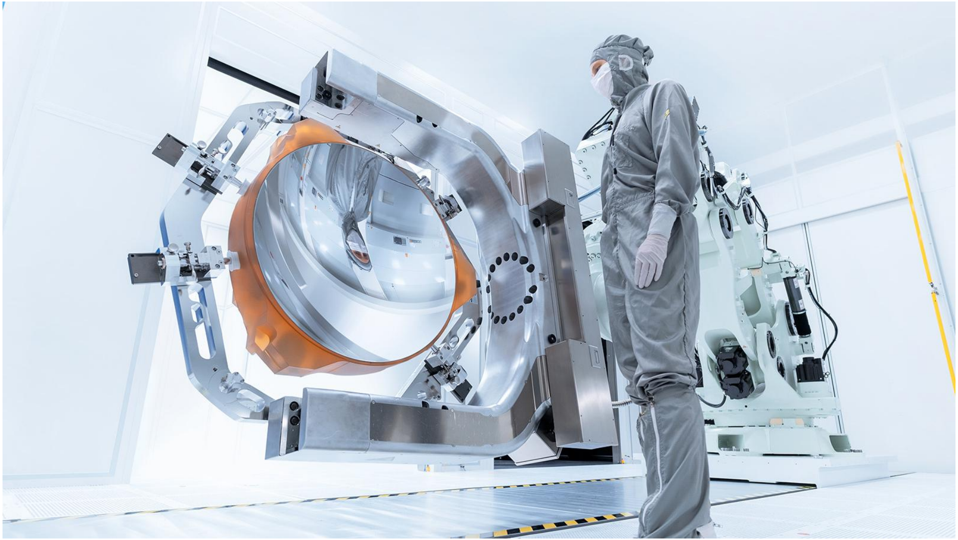
Wouter Kouw — Bayesian Intelligent Autonomous Systems lab
BIASlab Colloquium 16 september 2026

TU/e

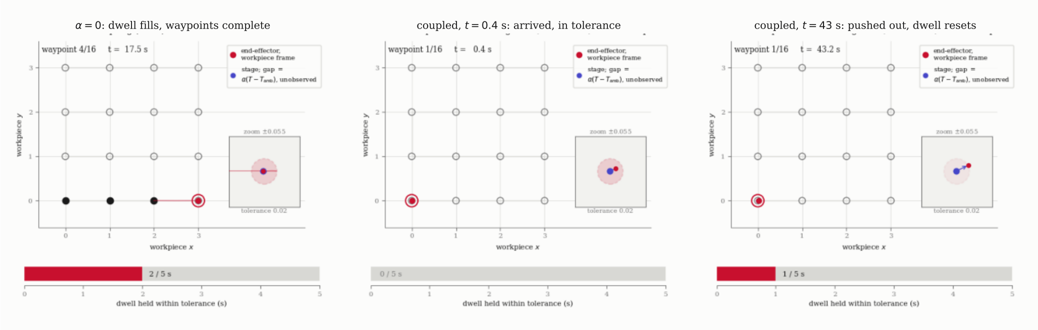
Outline

- **Problem statement** — thermal stage positioning
- **Idea** — epistemics from an information constraint
- **Hierarchy** — an upper level that acts on the lower level's epistemic drive
- **Experiment** — HARxI on thermal stage
- **Discussion** — limitations, todo's

Sharpening lenses



Laser stage positioning



Equations of motion

Position, velocity and temperature evolve over discrete-time ($\Delta t = 10$ ms) as

$$\begin{aligned}
p_{k+1} &= p_k + \Delta t\, v_k \\
v_{k+1} &= (1 - \Delta t\, b_v/m)\, v_k + (\Delta t/m)\, u_k + w_k^v \\
T_{k+1} &= T_k + \Delta t\big[-\kappa(T_k - T_{\mathrm{amb}}) + \eta p_k \big] + w_k^T
\end{aligned}$$

where w are zero-mean Gaussian distributed noises, T_{amb} is ambient temperature, $\kappa, \eta, \alpha, m, b_v$ are parameters and u_k is a control signal.

Power flows *only while dwelling* on the waypoint, and it must be held within a range of $\tau = 0.02$ for a $5\,$s dwell.

$$p_k^{\mathrm{wp}} = p_k + \alpha\,(T_k - T_{\mathrm{amb}})$$

Total trial time is 200 seconds.

Model specification: autoregression

- Let y_k be a noisy observation of position p_k, and $\bar{y}_k$ the buffer of previous observations. Then:

$$p(y_k \,|\, \Theta, u_k, \bar{y}_k) = \mathcal{N}\big(y_k \,|\, A^\top x_k,\; W^{-1}\big), \qquad x_k = \begin{bmatrix} u_k \\ \bar{y}_k \end{bmatrix} \in \mathbb{R}^{D_x}$$

- Matrix Normal-Wishart prior for autoregression parameters A, W.
- Symmetric-Beta control prior on $[-1, 1]^{D_u}$.
- Bayesian filtering is in closed form*, with an exact posterior predictive:

$$p(y_{k+1} \,|\, u_{k+1}, \mathcal{D}_k) = \mathcal{T}_{\eta_k}\Big(y_{k+1} \,|\, M_k^\top x_{k+1},\; \eta_k^{-1}\Omega_k(1 + \underbrace{x_{k+1}^\top \Lambda_k^{-1} x_{k+1}}_{\text{leverage } \ell_{k+1}})\Big)$$

*Nisslbeck & Kouw (2025), *Factor graph-based online Bayesian identification and component evaluation for multivariate autoregressive exogenous input models, Entropy.*

Expected free energy

Let τ denote a sequence of times into the future, i.e., $\tau = k+1, \dots$. Then, the EFE functional may be written as*

$$G[q] = \mathbb{E}_{q(x_\tau, u_\tau, \Theta)} \left[\mathbb{E}_{q(y_\tau \,|\, x_\tau, \Theta)} \left[\ln \frac{q(x_\tau, u_\tau, \Theta)}{p(y_\tau, u_\tau, x_\tau, x_t, \Theta \,|\, \mathcal{D}_t)} \right] \right] ,$$

with an EFE function

$$G(u_k) = \underbrace{\mathbb{D}_{\mathrm{KL}} \left[q(y_k \,|\, u_k) \,\|\, p^*(y_k) \right]}_{\text{risk}} + \underbrace{\mathbb{E}_{q(x_k, \Theta \,|\, u_k)} \left[\mathbb{H} \left[p(y_k \,|\, x_k, \Theta, u_k) \right] \right]}_{\text{ambiguity}} .$$

This can also be decomposed into goal-seeking and information-seeking terms:

$$G(u_k) \;=\; \underbrace{\mathbb{E}_{q(y_k \,|\, u_k)} \left[-\ln p^*(y_k) \right]}_{\text{goal cross-entropy}} - \underbrace{\mathbb{I} \left[x_k, \Theta \,;\, y_k \,|\, u_k \right]}_{\text{information gain}}$$

*Kouw (2023), *Information-seeking polynomial NARX model-predictive control through expected free energy minimization.*

Why this does not sit on a factor graph

The obstruction

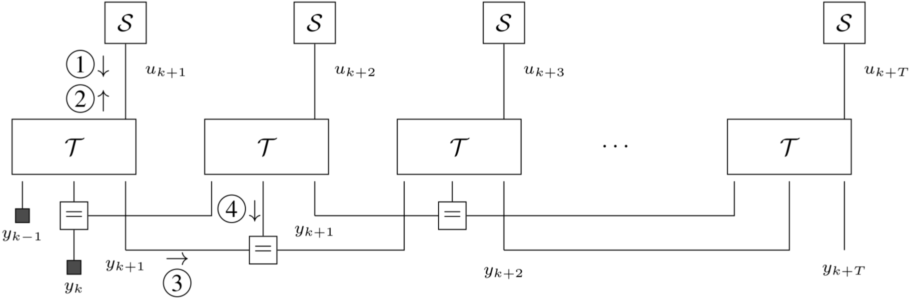
$\mathbb{E}_{q(y_\tau \mid x_\tau, \Theta)}[\cdot]$ costs $G[q]$ its KL structure — no local decomposition. Fold it back in, and the emission entropy cancels the emission energy: ambiguity vanishes.

- Leave EFE out of the passing — Schwöbel et al. (2018)
- Point-mass constraint on predicted outcomes — van de Laar et al. (2022)
- Non-standard message — Kouw et al. (2025)
- Epistemic priors — Nuijten et al. (2025)

This talk

An ordinary Bethe free energy, with the epistemic drive as an *inequality constraint*.

Unrolled planning graph



Bethe free energy, and one extra constraint

Bethe free energy of planning graph:

$$B[q] = \sum_{t\in\tau} \mathbb{E}_{q_t^y}\left[\ln \frac{q_t^y}{f_t^y}\right] + \sum_{t\in\tau} \mathbb{E}_{q_t^u}\left[\ln \frac{q_t^u}{f_t^u}\right] - \sum_{v} (d_v - 1)\,\mathbb{E}_{q(v)}\left[\ln q(v)\right]$$

$d_{y_t} = 1 + \min(M_y, T - t), \quad d_{u_t} = 2.$

Constraints:

- **Marginalization** and **normalization**, as usual.
- **Form constraints**: $q(y_t \mid u_t, \bar{y}_t) = \mathcal{T}_{\eta_k}(y_t \mid \mu_k, \Sigma_k)$ and $q(y_{k+T}) = \mathcal{N}(y_{k+T} \mid m_*, S_*)$.

Information constraint

$$\mathbb{E}_{q(u_t)q(\bar{y}_t)}\left[\mathbb{I}\left[\Theta, y_t \mid u_t, \bar{y}_t\right]\right] \;\geq\; \beta_t \equiv -\mathbb{H}\left[q(y_{k+T})\right]$$

Stationary action posterior

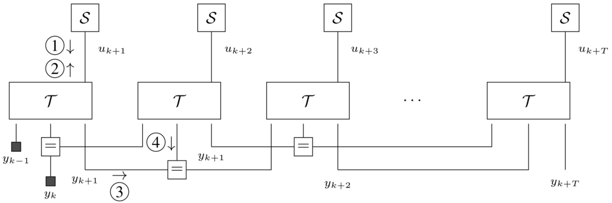
The Lagrangian (Bethe plus constraints) is purely over $q(u_t)$ and its stationary point is:

Theorem

$$q^*(u_t) \propto \underbrace{p(u_t)}_{\text{control prior}} \underbrace{\left(1 + \bar{\ell}_t(u_t)\right)^{\beta_t D_y/2}}_{\text{information-seeking}} \underbrace{\exp\left(-\int \lambda_t(y_t)\, \mathbb{E}_{q(\bar{y}_t)}\left[\tilde{p}(y_t \mid u_t, \bar{y}_t)\right] \mathrm{d}y_t\right)}_{\text{goal-seeking}}$$

- β_t is a KKT multiplier, inactive when it is 0
- Defining one β_t for each time-step in the planning graph means one 1-D bisection, $\sim 3\,$ms.
- At $\beta_t = 1$, we recover expected-free-energy exactly.
- Action selection by $u_t^* = \arg\max_{u_t \in [-1,1]^{D_u}} \ln q(u_t)$.

Message passing



- ① symmetric-Beta control prior, down onto the action edge
- ② back up it: leverage × goal matching, so $q^*(u_t) \propto$ ① · ②
- ③ forward: predictive marginalized over the buffer belief — moment-matched Gaussian
- ④ backward: predictive kernel against $q(u_t)$ and the goal-side message

Intermediate goals

Mode and Hessian of ③·④ give $\mathcal{N}(m_t^*, S_t^*)$, a Gaussian intermediate goal prior.

Hierarchies

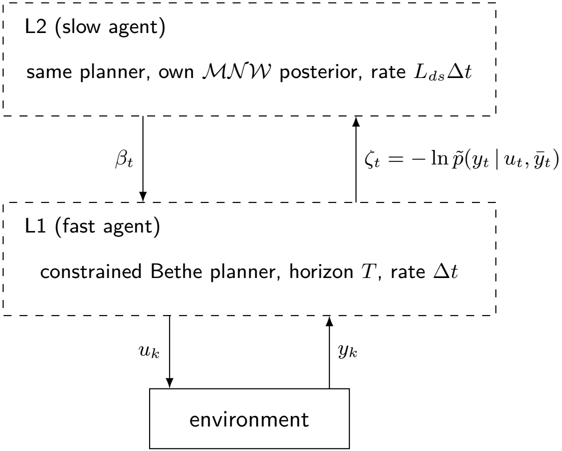
The literature contains a number of ways to design and develop hierarchies for agents:

- Multi-timescale MPC, singular perturbation: exchange *set-points*.
- Options, hierarchical RL: learn an *opaque coupling*.
- Hierarchical active inference: hand *goal priors* down.

Claim

A higher level should act on the lower level's *desire to explore*, not on its goals. And β is an ordinary scalar — so it can be an action.

Nesting: the upper level acts on β



L2's *observation* is L1's prediction error; L2's *action* is L1's information multiplier.

Agents and baselines

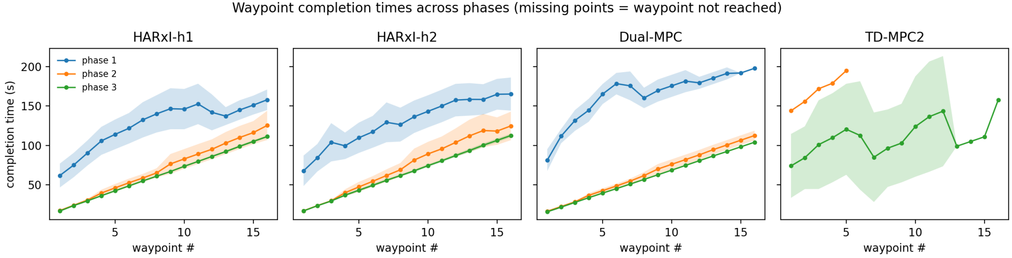
Ours

- HARXI-H1 — β solved per plan call by KKT bisection.
- HARXI-H2 — adds L2; every $L_{ds} = 10$ steps it emits β.

Baselines

- DUAL-MPC — same model, same posterior, *fixed* probing weight.
- TD-MPC2 — Hansen et al. (ICLR 2024).
- LQR / ORACLE — the references from before.

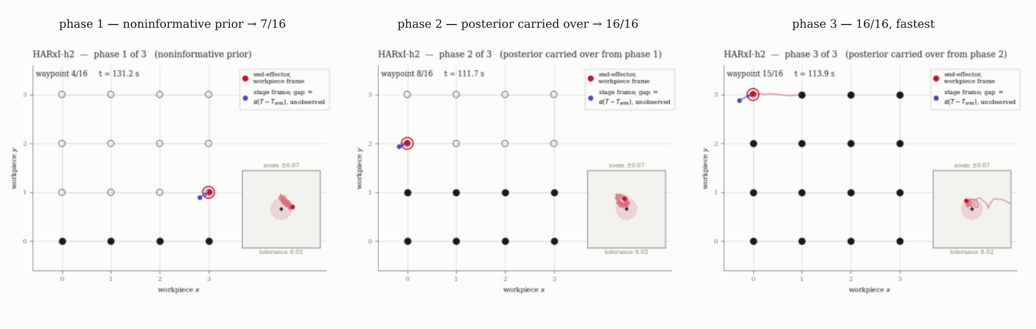
Results



- H2 is already near its phase-3 curve by phase 2.
- Dual-MPC needs all of phase 1 to clear the grid.
- TD-MPC2 never completes a phase-1 trial.

TU/e

HARxI solving the task



Solved, controlled, and still hand-set

Solved

β by complementary slackness. One scalar root-find.

Controlled

The same β as the *action* of a slower agent.

Still hand-set

The goal prior — and hence the floor.

Open

Have we removed a hyperparameter, or moved it?

Conclusion

- Action selection as *constrained* Bethe free energy minimization.
- One information constraint, one KKT multiplier — solved, not tuned.
- In the autoregressive case the constraint *is* the leverage.
- A knob is something another agent can hold.

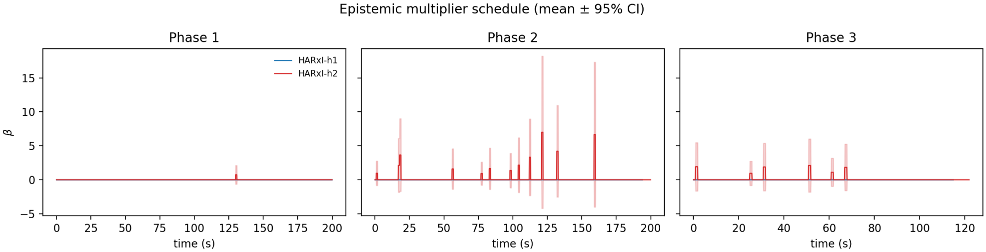
Next: deeper hierarchies, richer likelihoods, and the quadruped.

Thanks!

References

- Da Costa, Parr, Sajid, Veselic, Neacsu, Friston (2020). Active inference on discrete state-spaces: a synthesis. *J. Math. Psych.*
- Yedidia, Freeman, Weiss (2005). Constructing free-energy approximations and generalized belief propagation algorithms. *IEEE Trans. Inf. Theory.*
- Şenöz, van de Laar, Bagaev, de Vries (2021). Variational message passing and local constraint manipulation in factor graphs. *Entropy.*
- van de Laar, Koudahl, van Erp, de Vries (2022). Active inference and epistemic value in graphical models. *Frontiers in Robotics and AI.*
- van de Laar, Şenöz, Özçelikkale, Wymeersch (2021). Chance-constrained active inference. *Neural Computation.*
- Nuijten, Lukashchuk, van de Laar, de Vries (2025). A message passing realization of expected free energy minimization. *IWAI.*
- Kouw, Nisslbeck, Nuijten (2025). Message passing-based inference in an autoregressive active inference agent. *IWAI.*
- Nisslbeck, Kouw (2025). Factor graph-based online Bayesian identification for MARX models. *Entropy.*
- Hansen, Su, Wang (2024). TD-MPC2: scalable, robust world models for continuous control. *ICLR.*
- Boyd, Vandenberghe (2004). *Convex Optimization.* Cambridge University Press.

Backup: the β schedule



L2 leaves β at the KKT-inactive floor almost everywhere and raises it in short bursts. Phase 1 has nothing to react to yet; from phase 2 the carried-over posterior makes prediction error informative.

Backup: filtering updates

$$
\begin{aligned}
\Lambda_k &= \Lambda_{k-1} + x_k x_k^\top, \qquad \nu_k = \nu_{k-1} + 1,\\
M_k &= (\Lambda_{k-1} + x_k x_k^\top)^{-1}(\Lambda_{k-1} M_{k-1} + x_k y_k^\top),\\
\Omega_k &= \Omega_{k-1} + y_k y_k^\top + M_{k-1}^\top \Lambda_{k-1} M_{k-1}\\
&\quad - (\Lambda_{k-1} M_{k-1} + x_k y_k^\top)^\top (\Lambda_{k-1} + x_k x_k^\top)^{-1} (\Lambda_{k-1} M_{k-1} + x_k y_k^\top).
\end{aligned}
$$

Under persistent excitation the posterior contracts around (A^*, W^*) at the parametric Bernstein–von Mises rate.