

# Message passing-based inference in an autoregressive active inference agent

Wouter M. Kouw<sup>1</sup>, Tim N. Nisslbeck<sup>1</sup>, and Wouter L.N. Nijuten<sup>1,2</sup>

<sup>1</sup>Eindhoven University of Technology, Eindhoven, Netherlands

<sup>2</sup>Lazy Dynamics B.V., Eindhoven, Netherlands

w.m.kouw@tue.nl

**Abstract.** We present the design of an autoregressive active inference agent in the form of message passing on a factor graph. Expected free energy is derived and distributed across a planning graph. The proposed agent is validated on a robot navigation task, demonstrating exploration and exploitation in a continuous-valued observation space with bounded continuous-valued actions. Compared to a classical optimal controller, the agent modulates action based on predictive uncertainty, arriving later but with a better model of the robot’s dynamics.

**Keywords:** Intelligent agents · Free energy minimization · Active inference · Autoregressive models · Factor graphs · Message passing

## 1 Introduction

Active inference is a comprehensive framework that unifies perception, planning, and learning under the free energy principle, offering a promising approach to designing autonomous agents [2,18]. We present the design of an active inference agent implemented as a message passing procedure on a Forney-style factor graph [10,8]. The agent is built on an autoregressive model, making continuous-valued observations and inferring bounded continuous-valued actions [7,15]. We show that leveraging the factor graph approach produces a distributed, efficient and modular implementation [1,3,17,22].

Probabilistic graphical models have long been a unifying framework for the design and analysis of information processing systems, including signal processing, optimal controllers, and artificially intelligent agents [4,12,5,16,8]. Many famous algorithms can be written as message passing algorithms, including Kalman filtering, model-predictive control, and dynamic programming [12,16]. However, it can be a challenge to formulate new algorithms due to the requirement of local access to variables and the difficulty of deriving backwards messages. We highlight some of these challenges, and contribute with

- the derivation of expected free energy minimization in a multivariate autoregressive model with continuous-valued observations and bounded continuous-valued actions (Sec. 4.2), and

- the formulation of the planning model as a factor graph with marginal distribution updates based on messages passed along the graph (Figure 3).

We validate the proposed design on a robot navigation task, comparing the agent to an adaptive model-predictive controller.

## 2 Problem statement

We focus on the class of discrete-time stochastic nonlinear dynamical systems with state  $z_k \in \mathbb{R}^{D_z}$ , control  $u_k \in \mathbb{R}^{D_u}$ , and observation  $y_k \in \mathbb{R}^{D_y}$  at time  $k$ . Their evolution is governed by a state transition function  $f$  and an observation function  $g$ :

$$z_k = f(z_{k-1}, u_k) + w_k, \quad y_k = g(z_k) + v_k, \quad (1)$$

where  $w_k, v_k$  are stochastic contributions. The agent only receives noisy outputs  $y_k \in \mathbb{R}^{D_y}$  from a system and sends control inputs  $u_k \in \mathbb{U} \subset \mathbb{R}^{D_u}$  back. It must drive the system to output  $y_*$  without knowledge of the system's dynamics. Performance is measured with free energy (which in the proposed model is equal to the negative log evidence), Euclidean distance to goal, and the 2-norm magnitude of controls, over the course of a trial of length  $T$ .

## 3 Model specification

The model is autoregressive in nature, meaning that the system output at time  $k$  is predicted from the system input  $u_k$ ,  $M_u$  previous system inputs  $\bar{u}_k$  and  $M_y$  previous system outputs  $\bar{y}_k$ :

$$\bar{u}_k = \begin{bmatrix} u_{k-1} \\ \vdots \\ u_{k-M_u} \end{bmatrix}, \quad \bar{y}_k = \begin{bmatrix} y_{k-1} \\ \vdots \\ y_{k-M_y} \end{bmatrix}, \quad x_k = \begin{bmatrix} u_k \\ \bar{u}_k \\ \bar{y}_k \end{bmatrix}. \quad (2)$$

The vector  $x_k$  is the concatenation of these elements and has dimension  $D_x = D_u(M_u+1) + D_y M_y$ . Our likelihood function is based on a Gaussian distribution

$$p(y_k | \Theta, u_k, \bar{u}_k, \bar{y}_k) = \mathcal{N}(y_k | A^\top x_k, W^{-1}), \quad (3)$$

where  $A \in \mathbb{R}^{D_x \times D_y}$  is a regression coefficient matrix and  $W \in \mathbb{R}_+^{D_y \times D_y}$  is a precision matrix. Let  $\Theta = (A, W)$  refer to the parameters jointly. Their prior distribution is a matrix normal Wishart distribution [21, D175]:

$$p(\Theta) = \mathcal{MNW}(A, W | M_0, A_0^{-1}, \Omega_0^{-1}, \nu_0) \quad (4)$$

$$= \mathcal{MN}(A | M_0, A_0^{-1}, W^{-1}) \mathcal{W}(W | \Omega_0^{-1}, \nu_0). \quad (5)$$

The prior distributions over control inputs are independent Gaussian distributions, as are the goal prior distributions for future observations:

$$p(u_k) = \mathcal{N}(u_k | 0, \Upsilon^{-1}), \quad p(y_t | y_*) = \mathcal{N}(y_t | m_*, S_*), \quad (6)$$

where  $\Upsilon$  is a precision matrix and  $y_* = (m_*, S_*)$  are the goal mean vector and covariance matrix.

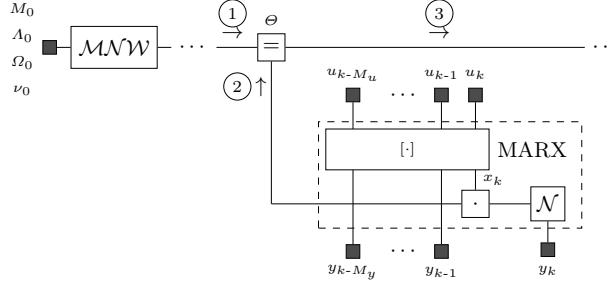
## 4 Inference

### 4.1 Learning

We use Bayesian filtering to update parameter beliefs given  $y_k, u_k$  [19,15]:

$$\underbrace{p(\Theta | \mathcal{D}_k)}_{\text{posterior}} = \frac{\overbrace{p(y_k | \Theta, u_k, \bar{u}_k, \bar{y}_k)}^{\text{likelihood}}}{\underbrace{p(y_k | u_k, \mathcal{D}_{k-1})}_{\text{evidence}}} \underbrace{p(\Theta | \mathcal{D}_{k-1})}_{\text{prior}}. \quad (7)$$

where  $\mathcal{D}_k = \{y_i, u_i\}_{i=1}^k$  is short-hand for data up to time  $k$ . Note that the memories  $\bar{u}_k, \bar{y}_k$  are subsets of  $\mathcal{D}_{k-1}$ . The evidence term is the evaluation of the observation  $y_k$  under the predictive distribution, obtained by marginalizing the likelihood over the parameters [14].



**Fig. 1.** Forney-style factor graph of one time step (separated by dots) of Bayesian filtering. Edges represent random variables and nodes operations on those variables. Black squares represent observed variables or set parameters, and the dotted box represents a custom node, composed of the nodes within. Message 1 is the prior belief over parameters and message 2 the likelihood-based update. These are multiplied at the equality node, yielding the marginal posterior distribution (message 3).

We express the Bayesian filtering procedure as message passing on the factor graph<sup>1</sup> shown in Figure 1. Message ① is the prior distribution on  $\Theta$ ,

$$\textcircled{1} = \mathcal{MNW}(A, W | M_{k-1}, A_{k-1}, \Omega_{k-1}, \nu_{k-1}). \quad (8)$$

Message ② originates from the MARX likelihood function and is an improper matrix normal Wishart distribution [15, Lemma 2],

$$\textcircled{2} = \mathcal{MNW}(A, W | \bar{M}_k, \bar{A}_k^{-1}, \bar{\Omega}_k^{-1}, \bar{\nu}_k), \quad (9)$$

with parameters based on data and buffers at time  $k$ ,

$$\bar{\nu}_k = 2 - D_x + D_y, \quad \bar{A}_k = x_k x_k^\top, \quad \bar{M}_k = (x_k x_k^\top)^{-1} x_k y_k^\top, \quad \bar{\Omega}_k = 0_{D_y \times D_y}. \quad (10)$$

<sup>1</sup> For excellent introductions to the factor graph approach, see [11,20].

It is improper because  $\bar{\omega}_k$  is singular. But when multiplied with the prior distribution, it produces the conjugate posterior distribution exactly [15, Thm. 1]. This multiplication occurs in the equality node and produces message (3):

$$(3) = p(\Theta | \mathcal{D}_k) = \mathcal{MNW}(A, W | M_k, \Lambda_k^{-1}, \Omega_k^{-1}, \nu_k). \quad (11)$$

The parameters of this distribution are

$$\nu_k = \nu_{k-1} + 1, \quad (12)$$

$$\Lambda_k = \Lambda_{k-1} + x_k x_k^\top, \quad (13)$$

$$M_k = (\Lambda_{k-1} + x_k x_k^\top)^{-1} (\Lambda_{k-1} M_{k-1} + x_k y_k^\top), \quad (14)$$

$$\Omega_k = \Omega_{k-1} + y_k y_k^\top + M_{k-1}^\top \Lambda_{k-1} M_{k-1} - \quad (15)$$

$$(\Lambda_{k-1} M_{k-1} + x_k y_k^\top)^\top (\Lambda_{k-1} + x_k x_k^\top)^{-1} (\Lambda_{k-1} M_{k-1} + x_k y_k^\top).$$

Marginalizing the Gaussian likelihood in Eq. 3 over the parameter posterior distribution (Eq. 11) yields a multivariate location-scale T-distribution [14]:

$$p(y_k | u_k, \mathcal{D}_k) = \int p(y_k | \Theta, u_k, \bar{u}_k, \bar{y}_k) p(\Theta | \mathcal{D}_k) d\Theta = \mathcal{T}_{\eta_k}(y_k | \mu_k(u_k), \Sigma_k(u_k)), \quad (16)$$

with  $\eta_k = \nu_k - D_y + 1$  degrees of freedom and a mean and covariance of

$$\mu_k(u_k) = M_k^\top \begin{bmatrix} u_k \\ \bar{u}_k \\ \bar{y}_k \end{bmatrix}, \quad \Sigma_k(u_k) = \frac{1}{\nu_k - D_y + 1} \Omega_k \left( 1 + \begin{bmatrix} u_k \\ \bar{u}_k \\ \bar{y}_k \end{bmatrix}^\top \Lambda_k^{-1} \begin{bmatrix} u_k \\ \bar{u}_k \\ \bar{y}_k \end{bmatrix} \right). \quad (17)$$

The subscripts under  $\mu$  and  $\Sigma$  indicate which parameters were used, i.e., here they refer to  $M_k$ ,  $\Lambda_k$ ,  $\Omega_k$  and  $\nu_k$ .

## 4.2 Actions

**Planning** We start by building a generative model for the input and output at time  $t = k + 1$ :

$$p(y_t, \Theta, u_t | \mathcal{D}_k) = p(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t) p(\Theta | \mathcal{D}_k) p(u_t). \quad (18)$$

Note that  $\bar{u}_t$  and  $\bar{y}_t$  are absent on the left-hand side because, at time  $t = k + 1$ , these buffers are subsets of  $\mathcal{D}_k$ . We want the agent to pursue a target, a specific future observation. To do so, we first isolate the marginal distribution  $p(y_t)$ ,

$$p(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t) p(\Theta | \mathcal{D}_k) = p(\Theta | y_t, u_t, \mathcal{D}_k) p(y_t), \quad (19)$$

and then constrain it to be the goal prior,  $p(y_t) \rightarrow p(y_t | y_*)$ . We use Bayes' rule in the reverse direction to relate the distribution over parameters given the future output and input, to known distributions:

$$p(\Theta | y_t, u_t, \mathcal{D}_k) = \frac{p(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t) p(\Theta | \mathcal{D}_k)}{p(y_t | u_t, \mathcal{D}_k)}. \quad (20)$$

To obtain an approximate marginal posterior distribution for the action  $u_t$ , we form an expected free energy functional,

$$\mathcal{F}_k[q] = \mathbb{E}_{q(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t)} \left[ \mathbb{E}_{q(\Theta, u_t)} \left[ \ln \frac{q(\Theta, u_t)}{p(y_t, \Theta, u_t, | y_*, \mathcal{D}_k)} \right] \right]. \quad (21)$$

The variational model is  $q(y_t, \Theta, u_t, \bar{u}_t, \bar{y}_t) = q(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t)q(\Theta)q(u_t)$ . The likelihood and parameter factors are not free variational distributions but fixed to the same form as the likelihood and parameter factors of the generative model:

$$q(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t) = p(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t) = \mathcal{N}(y_t | A^\top x_t, W^{-1}) \quad (22)$$

$$q(\Theta) = p(\Theta | \mathcal{D}_k) = \mathcal{MNW}(A, W | M_k, A_k^{-1}, \Omega_k^{-1}, \nu_k). \quad (23)$$

We then minimize this expected free energy functional with respect to the variational distribution  $q(u_t)$ :

$$q^*(u_t) = \arg \min_{q \in Q} \mathcal{F}_k[q]. \quad (24)$$

where  $Q$  represents the set of candidate variational distributions.

**Theorem 1.** *The optimal variational posterior  $q^*(u_t)$  under the free energy functional defined in (21) is proportional to a prior times a likelihood,*

$$q^*(u_t) \propto p(u_t) \exp(-G(u_t)), \quad (25)$$

where  $G$  is the sum of a mutual information and a cross-entropy term

$$G(u_t) = -\mathbb{E}_{p(y_t, \Theta | u_t, \mathcal{D}_k)} \left[ \ln \frac{p(y_t, \Theta | u_t, \mathcal{D}_k)}{p(y_t | u_t, \mathcal{D}_k)p(\Theta | \mathcal{D}_k)} \right] - \mathbb{E}_{p(y_t | u_t, \mathcal{D}_k)} \left[ \ln p(y_t | y_*) \right]. \quad (26)$$

The proof can be found in Appendix A.

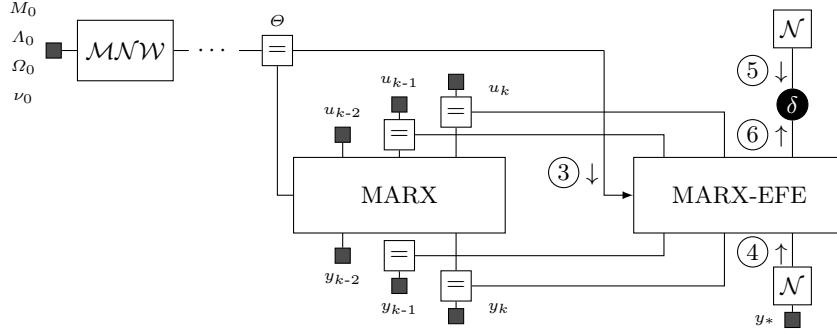
**Corollary 1.** *The expected free energy function  $G(u_t)$  evaluates to:*

$$G(u_t) = \text{constants} - \frac{1}{2} \ln |\Sigma_t(u_t)| + \frac{1}{2} \text{Tr} \left[ S_*^{-1} \left( \Sigma_t(u_t) \frac{\eta_t}{\eta_t - 2} + \Xi(u_t) \right) \right]. \quad (27)$$

where  $\Xi(u_t) = (\mu_t(u_t) - m_*)(\mu_t(u_t) - m_*)^\top$ .

The proof is also in Appendix A.

Figure 2 provides an example of how this inference process can be mapped to a factor graph, using  $M_y = M_u = 2$ . The node marked "MARX" is the composite node depicted in Figure 1. It is now connected to another composite node marked "MARX-EFE", which is connected to previous observations, parameters, goal prior parameters  $y_*$  and the to-be-taken action  $u_t$ . Message ③ is the same as in Figure 1, namely the parameter posterior distribution (Sec. 4.1). Note that during planning, the parameters are not updated. This is indicated by a *directed* edge from the equality node to the MARX-EFE node. Message ④ is the goal



**Fig. 2.** Factor graph of the 1-step ahead planning model. The left half of the graph is the same as in Figure 1. The parameter posterior (message 3) is passed forwards to the MARX-EFE node, which takes in message 4 from the goal prior node and produces message 6 containing the exponentiated EFE function. Combined with message 5 from the control prior node, this produces the variational control posterior. The  $\delta$  circle denotes a collapse of the posterior to a Dirac delta distribution [9].

prior distribution (Eq. 6), (5) is the control prior distribution (Eq. 6) and message (6) is the unnormalized exponentiated expected energy;

$$(4) = \mathcal{N}(y_t | m_*, S_*) , \quad (5) = \mathcal{N}(u_t | 0, \Upsilon^{-1}) , \quad (6) = \exp(-G(u_t)) . \quad (28)$$

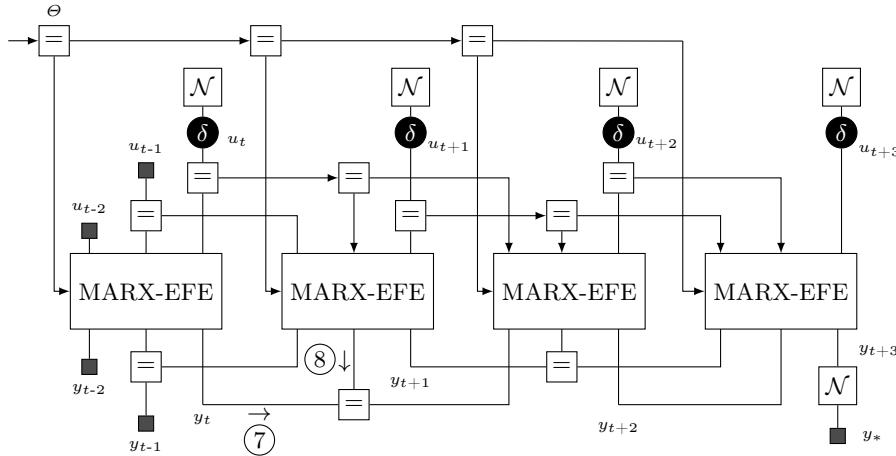
Note that message 6 is the result of the EFE derivation (Thm. 1) and not the result of minimizing the Bethe free energy, a point discussed in more detail in Section 6.

Normalizing  $q^*(u_t)$ , i.e., the product of (5) and (6), requires integrating over  $u_t$ . This is challenging and avoided by collapsing the approximate posterior to its maximum a posteriori point-mass distribution,  $q^*(u_t) \approx \delta(u_t - \hat{u}_t)$  where

$$\hat{u}_t = \arg \max_{u_t \in \mathbb{U}} q^*(u_t) = \arg \min_{u_t \in \mathbb{U}} \frac{1}{2} u_t^\top \Upsilon u_t + G(u_t) . \quad (29)$$

We believe this is justified because collapsing the posterior to a point estimate is anyway required to pass controls to actuators.

**Horizon** Extending the time horizon is challenging, requiring additional marginalizations that complicate the above results (see Sec. 6 for an extended discussion). In this paper, we adopt a simpler approach and generalize the planning factor graph (Figure 2) by including additional MARX-EFE nodes. Figure 3 shows an extension with  $M_y = M_u = 2$  and a time horizon of  $H = 4$ . The main difference is that, for  $t > k + 2$ , the buffer  $\bar{y}$  will no longer contain delta distributed variables. Note that the buffer  $\bar{u}$  will still contain delta variables, because we constrain the marginal action posteriors to be delta's (Eq. 29). Inspecting the second MARX-EFE node in Figure 3 reveals that the only change from the



**Fig. 3.** Factor graph of a 4-step ahead planning model, showing repeated MARX-EFE node from Figure 2. Some buffer variables are now latent as well. Message 7 is the posterior predictive over  $y_t$  carrying forward system output predictions given a selection control input. Message 8 is a predictive likelihood over  $y_t$  sent backwards from the node at time  $t + 1$ . Together the forward and backward pass of predictive messages generates a sequence of goal priors.

MARX-EFE node in Figure 2 is the incoming message for  $y_t$ . This message is the posterior predictive distribution (given a selected action  $\hat{u}_t$ ) sent out by the first MARX-EFE node in the planning graph;

$$\textcircled{7} = p(y_t | \hat{u}_t, \mathcal{D}_t) = \mathcal{T}_{\eta_k}(y_t | \mu_k(\hat{u}_t), \Sigma_k(\hat{u}_t)). \quad (30)$$

This message is incorporated into the MARX-EFE node function through a variational approximation:

$$\begin{aligned} & p(y_{t+1} | u_{t+1}, \bar{u}_{t+1}, \tilde{y}_{t+1}, \Theta) \\ & \approx \exp \left( \mathbb{E}_{p(y_t | \hat{u}_t, \mathcal{D}_t)} [\ln p(y_{t+1} | u_{t+1}, \bar{u}_{t+1}, \bar{y}_{t+1}, \Theta)] \right) \end{aligned} \quad (31)$$

$$\propto \exp \left( -\frac{1}{2} (y_{t+1}^\top W y_{t+1} - 2 y_{t+1}^\top W A^\top \mathbb{E}_{p(y_t | \hat{u}_t, \mathcal{D}_t)} [u_{t+1} \bar{u}_{t+1} y_t \hat{y}_{t-1}]^\top) \right) \quad (32)$$

$$\propto \mathcal{N}(y_{t+1} | A^\top [u_{t+1} \bar{u}_{t+1} \mu_k(\hat{u}_t) \hat{y}_{t-1}]^\top, W^{-1}). \quad (33)$$

Note that this is, in essence, still the MARX likelihood function except with the mean of the posterior predictive  $\mu_k(\hat{u}_t)$  instead of an observed value for  $y_t$  in  $\bar{y}_{t+1}$ . We mark this change with  $\tilde{y}$  instead of  $\bar{y}$ .

For the backwards message from the MARX-EFE node at  $t + 1$  towards the variable  $y_t$ , we first utilize the same variational approximation as above but now

with respect to the variational factor  $q(y_{t+1}) = \mathcal{N}(y_{t+1} | m_{t+1}, S_{t+1})$  (Eq. 39);

$$\begin{aligned} & \exp(\mathbb{E}_{q(y_{t+1})}[\ln p(y_{t+1} | u_{t+1}, \bar{u}_{t+1}, \bar{y}_{t+1}, \Theta)]) \\ & \propto \exp\left(-\frac{1}{2}\mathbb{E}_{q(y_{t+1})}[y_{t+1}^\top W y_{t+1} - 2y_{t+1} W A^\top [u_{t+1} \bar{u}_{t+1} y_t \hat{y}_{t-1}]^\top]\right) \quad (34) \end{aligned}$$

$$\propto \mathcal{N}(m_{t+1} | A^\top [u_{t+1} \bar{u}_{t+1} y_t \hat{y}_{t-1}]^\top, W^{-1}). \quad (35)$$

Then we marginalize over the parameter posterior distribution,

$$(8) \propto \mathbb{E}_{p(\Theta | \mathcal{D}_k)}[\mathcal{N}(m_{t+1} | A^\top [u_{t+1} \bar{u}_{t+1} y_t \hat{y}_{t-1}]^\top, W^{-1})] \quad (36)$$

$$\propto \mathcal{T}_{\bar{\eta}}(m_{t+1} | \bar{\mu}_{t+1}(y_t), \bar{\Sigma}_{t+1}(y_t)), \quad (37)$$

with  $\bar{\eta} = \nu_k - D_y + 1$  degrees of freedom and mean and covariance

$$\bar{\mu}_{t+1}(y_t) = M_k^\top \begin{bmatrix} u_{t+1} \\ \bar{u}_{t+1} \\ y_t \\ \hat{y}_{t-1} \end{bmatrix}, \quad \bar{\Sigma}_{t+1}(y_t) = \frac{1}{\bar{\eta}} \Omega_k \left( 1 + \begin{bmatrix} u_{t+1} \\ \bar{u}_{t+1} \\ y_t \\ \hat{y}_{t-1} \end{bmatrix}^\top A_k^{-1} \begin{bmatrix} u_{t+1} \\ \bar{u}_{t+1} \\ y_t \\ \hat{y}_{t-1} \end{bmatrix} \right). \quad (38)$$

In essence, this distribution scores which values of  $y_t$  best predict  $y_{t+1}$ , with  $m_{t+1}$  as a pseudo-observation. At the  $y_t$  edge, we perform a variational factor update based on the product of messages (7) and (8):

$$q(y_t) \propto \mathcal{T}_{\eta_k}(y_t | \mu_k(\hat{u}_t), \Sigma_k(\hat{u}_t)) \mathcal{T}_{\bar{\eta}}(m_{t+1} | \bar{\mu}_{t+1}(y_t), \bar{\Sigma}_{t+1}(y_t)). \quad (39)$$

This product is not part of a known parametric family of distributions. We perform a Laplace approximation to produce  $q(y_t) \approx \mathcal{N}(y_t | m_t, S_t)$  where [13]:

$$m_t = \arg \max_{y_t} \ln \mathcal{T}_{\eta_k}(y_t | \mu_k(\hat{u}_t), \Sigma_k(\hat{u}_t)) \mathcal{T}_{\bar{\eta}}(m_{t+1} | \bar{\mu}_{t+1}(y_t), \bar{\Sigma}_{t+1}(y_t)) \quad (40)$$

$$S_t^{-1} = -\nabla_{y_t}^2 \ln \mathcal{T}_{\eta_k}(y_t | \mu_k(\hat{u}_t), \Sigma_k(\hat{u}_t)) \mathcal{T}_{\bar{\eta}}(m_{t+1} | \bar{\mu}_{t+1}(y_t), \bar{\Sigma}_{t+1}(y_t)) \Big|_{y_t=m_t}. \quad (41)$$

This Gaussian variational factor effectively becomes a goal prior for time  $t$ . So, in the extended time horizon, we see that the forward and backward passes over the future observations generate a sequence of intermediate goal priors. As such, at each future time point, the agent needs only to solve a 1-step ahead expected free energy minimization problem.

## 5 Experiments

We perform simulation experiments in which agents have to reach a target state in a single trial. We refer to the proposed free energy minimizing agent as *MARXEFE*<sup>2</sup>. The benchmark is the same agent but with controls found by minimizing a standard model-predictive control cost function;

$$\hat{u}_{k+1:k+H}^{\text{MPC}} = \arg \min_{u_{k+1:k+H} \in \mathbb{U}^H} \sum_{t=k+1}^{k+H} u_t^\top \Upsilon u_t + (\mu_t(u_t) - m_*)^\top (\mu_t(u_t) - m_*), \quad (42)$$

---

<sup>2</sup> Agent built with RxInfer; <https://github.com/biaslab/IWAI2025-MARXEFE-MP>

This agent will be called *MARX-MPC*. We evaluate the probabilities of the system output  $y_k$  under the goal prior distribution,  $p(y_k | y_*)$ . Additionally, we evaluate model evidence, i.e., the probability of the system output observation under the agent's predictive distribution,  $p(y_k | u_k, \mathcal{D}_k)$ .

**System** Consider a linear Gaussian dynamical system where the state vector  $z_k$  contains the two-dimensional position and velocity of a robot. The robot's state transition and measurement functions are

$$f(z_{k-1}, u_k) = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} z_{k-1} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix} u_k, \quad g(z_k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} z_k, \quad (43)$$

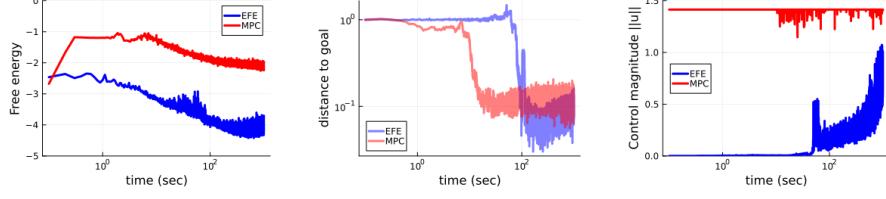
where  $\Delta t$  is the time step size. Its covariance matrices are

$$Q = \begin{bmatrix} \frac{\Delta t^3}{3} \varsigma_1 & 0 & \frac{\Delta t^2}{2} \varsigma_1 & 0 \\ 0 & \frac{\Delta t^3}{3} \varsigma_2 & 0 & \frac{\Delta t^2}{2} \varsigma_2 \\ \frac{\Delta t^2}{2} \varsigma_1 & 0 & \Delta t \varsigma_1 & 0 \\ 0 & \frac{\Delta t^2}{2} \varsigma_2 & 0 & \Delta t \varsigma_2 \end{bmatrix}, \quad R = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}, \quad (44)$$

with  $\varsigma = [10^{-6} \ 10^{-6}]^\top$  and  $\rho = [10^{-3} \ 10^{-3}]^\top$ .

**Prior parameters** The prior parameters are weakly informative;  $\nu_0 = 100$ ,  $M_0 = 1/(D_x D_y) \cdot I_{D_x \times D_y}$ ,  $\Lambda_0 = 10^{-2} \cdot I_{D_x}$ ,  $\Omega_0 = I_{D_y}$ , and  $\Upsilon = 10^{-6} \cdot I_{D_u}$ . The system starts at  $z_0 = [0 \ 0 \ 0 \ 0]^\top$  and the goal prior has mean  $m_* = [0 \ 1]^\top$  and covariance matrix  $S_* = 10^{-6} \cdot I_{D_y}$ . Buffers are fixed at  $M_u = M_y = 2$  and the time horizon at  $H = 3$ . Controls are limited to  $\mathbb{U} = [-1, 1]$  for  $T = 10000$  steps at  $\Delta t = 0.1$ .

**Results** Figure 4 shows the experimental results comparing MARX-EFE to MARX-MPC. The left figure shows that MARX-EFE consistently scores a smaller free energy than MARX-MPC, demonstrating that it cares more strongly about accurately predicting its next observation. The middle figure shows the distance to goal over the duration of the trial where, on average, MARX-MPC reaches the goal sooner than MARX-EFE. MARX-MPC does not care about making accurate predictions, only to close the gap to the target as quickly as possible. It is successful in that regard but struggles to park itself on the target exactly because it ignored opportunities to learn the finer parts of robot's dynamics earlier in the trial. The MARX-EFE agent ultimately gets closer than MARX-MPC because - by the time it gets to the goal - it has a much better model of the robot's dynamics. The right figure shows the 2-norm of the controls, highlighting that MARX-MPC consistently utilizes maximum power ( $\max \|u_t\|_2 = \sqrt{2}$ ) to get closer. The MARX-EFE agent takes very small actions in the beginning, when it is uncertain of their outcomes, and slowly takes larger actions when its uncertainty shrinks. We interpret this behaviour as some form of "caution".



**Fig. 4.** MARX-EFE (blue) vs. MARX-MPC (red) over a trial of 1000 seconds, compared in terms of free energy (left), Euclidean distance to goal ( $m_*$ ; middle) and 2-norm of controls (right). Results are averaged over 10 experiments. MARX-EFE initially takes smaller actions, aiming to improve parameters and predictions first. It arrives at the goal later than MARX-MPC but is better able to park on the goal itself. MARX-EFE's actions are small initially but increase in magnitude as uncertainty shrinks.

## 6 Discussion

For planning, we observe that the expectation over the future observation is actually not necessary in the above model. If the likelihood is incorporated into the numerator of Eq. 21 as well, i.e.,

$$\mathcal{F}_k[q] = \mathbb{E}_{q(y_t, \Theta, u_t)} \left[ \ln \frac{q(y_t, \Theta, u_t)}{p(y_t, \Theta, u_t | y_*, \mathcal{D}_k)} \right], \quad (45)$$

then - following the same steps as in Appendix A - the EFE function becomes:

$$G(u_t) = \mathbb{E}_{p(y_t | u_t, \mathcal{D}_k)} [\ln p(y_t | u_t, \mathcal{D}_k)] + \mathbb{E}_{p(y_t | u_t, \mathcal{D}_k)} [-\ln p(y_t | y_*)]. \quad (46)$$

In both the mutual information in Eq. 26 and in the entropy of the posterior predictive above, the only term that depends on  $u_t$  is the variance of the posterior predictive  $\Sigma_k(u_t)$ . All other terms drop out due to the translation invariance of differential entropies. Thus, we find the same solution when using a standard free energy functional instead of an expected free energy functional [22].

*Limitations* In Sec. 4.2 we avoided forming the joint posterior predictive distribution over all future outputs in the horizon;

$$p(y_{k+1:k+H} | u_{k+1:k+H}, \mathcal{D}_k) = \int p(\Theta | \mathcal{D}_k) \prod_{t=k+1}^{k+H} p(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t) d\Theta. \quad (47)$$

This marginalization is a challenge because blocks of the autoregressive coefficient start to nest in both the mean and covariance matrix of the joint Gaussian likelihood. To illustrate this, consider an example with  $M_y = 1$  such that  $\bar{y}_t = y_{t-1}$ . The joint distribution of the likelihoods for  $k+1$  and  $k+2$  is Gaussian distributed with mean vector and covariance matrix

$$\begin{bmatrix} A_1^\top u_{k+1} + A_2^\top \bar{u}_k + A_3^\top y_k \\ A_1^\top u_{k+2} + A_2^\top \bar{u}_{k+1} + A_3^\top (A_1^\top u_{k+1} + A_2^\top \bar{u}_k + A_3^\top y_k) \end{bmatrix} \begin{bmatrix} W^{-1} & W^{-1} A_3 \\ A_3^\top W^{-1} & A_3^\top W^{-1} A_3 + W^{-1} \end{bmatrix}, \quad (48)$$

where  $A_i$  represent row-indexed blocks of the coefficient matrix  $A$ . Marginalizing this joint future likelihood over  $p(\Theta | \mathcal{D}_k)$  is difficult but a solution would avoid the variational and Laplace approximation errors in Eqs. 31, 34, and 40.

## 7 Conclusion

We designed an active inference agent with continuous-valued actions as a message passing procedure on a factor graph. The forward and backward pass of predictions over future system outputs generates a sequence of intermediate goal priors. Each node in the planning graph only has to solve a 1-step ahead EFE minimization problem to find appropriate controls. The agent successfully navigates a robot to a goal position under unknown dynamics.

**Acknowledgments.** The authors gratefully acknowledge support from the Eindhoven Artificial Intelligence Systems Institute.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## A Appendix

*Proof.* Using the factorisation of the variational model, the expected free energy functional can be re-arranged to isolate the expectation over  $q(u_t)$ :

$$\mathcal{F}_t[q] = \mathbb{E}_{q(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t)} \left[ \mathbb{E}_{q(\Theta, u_t)} \left[ \ln \frac{q(\Theta, u_t)}{p(y_t, \Theta, u_t | y_*, \mathcal{D}_k)} \right] \right] \quad (49)$$

$$= \mathbb{E}_{q(u_t)} \left[ \ln \frac{q(u_t)}{p(u_t)} + \mathbb{E}_{q(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t)q(\Theta)} \left[ \ln \frac{q(\Theta)}{p(y_t, \Theta, u_t | y_*, \mathcal{D}_k)} \right] \right] \quad (50)$$

$$= \mathbb{E}_{q(u_t)} \left( \ln \frac{q(u_t)}{p(u_t) \exp(-G(u_t))} \right), \quad (51)$$

for  $G(u_t) = \mathbb{E}_{q(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t)q(\Theta)} \left[ \ln q(\Theta)/p(y_t, \Theta | u_t, y_*, \mathcal{D}_k) \right]$  and the identity  $G(u_t) = \ln(1/\exp(-G(u_t)))$ . Constraining  $q(u_t)$  to be a probability distribution over the space of affordable controls  $\mathbb{U}$  is done with a Lagrange multiplier:

$$\mathcal{L}[q, \gamma] = \mathcal{F}_t[q] + \gamma \left( \int_{\mathbb{U}} q(u_t) du_t - 1 \right). \quad (52)$$

The stationary solution  $q^*(u_t)$  of the Lagrangian is found at  $\delta \mathcal{L}[q, \gamma]/\delta q = 0$  [23]). Let  $\delta q(u_t) = \epsilon \phi(u_t)$  be a variation with  $\phi$  a continuous and differentiable test function. Then the variational derivative can be found with:

$$\int_{\mathbb{U}} \frac{\delta \mathcal{L}[q, \gamma]}{\delta q} \phi(u_t) du_t = \frac{d \mathcal{L}[q(u_t) + \epsilon \phi(u_t), \gamma]}{d \epsilon} \Big|_{\epsilon=0} \quad (53)$$

$$= \int_{\mathbb{U}} \left( \ln \frac{q(u_t)}{p(u_t) \exp(-G(u_t))} + 1 + \gamma \right) \phi(u_t) du_t. \quad (54)$$

Setting the variational derivative to 0, yields

$$\ln \frac{q(u_t)}{p(u_t) \exp(-G(u_t))} + 1 + \gamma = 0 \rightarrow q(u_t) = \exp(-\gamma - 1)p(u_t) \exp(-G(u_t)). \quad (55)$$

Plugging this into the constraint gives  $\exp(-\gamma - 1) = 1 / \int_{\mathbb{U}} p(u_t) \exp(-G(u_t)) du_t$ . As such, we have:

$$q(u_t) = \frac{p(u_t) \exp(-G(u_t))}{\int_{\mathbb{U}} p(u_t) \exp(-G(u_t)) du_t}. \quad (56)$$

Using (18), (19) and (22),  $G(u_t)$  can be simplified to a negative mutual information plus a cross-entropy term:

$$G(u_t) = \mathbb{E}_{p(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t) p(\Theta | \mathcal{D}_k)} \left[ \ln \frac{p(\Theta | \mathcal{D}_k) p(y_t | u_t, \mathcal{D}_k)}{p(y_t | \Theta, u_t, \bar{u}_t, \bar{y}_t) p(\Theta | \mathcal{D}_k) p(y_t | y_*)} \right] \quad (57)$$

$$= -\mathbb{E}_{p(y_t, \Theta | u_t, \mathcal{D}_k)} \left[ \ln \frac{p(y_t, \Theta | u_t, \mathcal{D}_k)}{p(y_t | u_t, \mathcal{D}_k) p(\Theta | \mathcal{D}_k)} \right] + \mathbb{E}_{p(y_t | u_t, \mathcal{D}_k)} \left[ -\ln p(y_t | y_*) \right]. \quad (58)$$

*Proof.* We split the mutual information into a joint entropy minus the entropy of the posterior predictive and that of the parameter posterior [13]:

$$\begin{aligned} -\mathbb{E}_{p(y_t, \Theta | u_t, \mathcal{D}_k)} \left[ \ln \frac{p(y_t, \Theta | u_t, \mathcal{D}_k)}{p(y_t | u_t, \mathcal{D}_k) p(\Theta | \mathcal{D}_k)} \right] &= \mathbb{E}_{p(y_t | u_t, \mathcal{D}_k)} \left[ \ln p(y_t | u_t, \mathcal{D}_k) \right] \\ &\quad + \mathbb{E}_{p(\Theta | \mathcal{D}_k)} \left[ \ln p(\Theta | \mathcal{D}_k) \right] - \mathbb{E}_{p(y_t, \Theta | u_t, \mathcal{D}_k)} \left[ \ln p(y_t, \Theta | u_t, \mathcal{D}_k) \right]. \end{aligned} \quad (59)$$

Since entropies are invariant to translation, only the entropy of the posterior predictive affects  $G(u_t)$  [7]. The entropy of a location-scale T-distribution is [6]:

$$\begin{aligned} \mathbb{E}_{p(y_t | u_t, \mathcal{D}_k)} \left[ \ln p(y_t | u_t, \mathcal{D}_k) \right] \\ = -\mathbb{E}_{\mathcal{T}_{\eta_k}(y_t | 0, \Sigma_k(u_t))} \left[ -\ln \mathcal{T}_{\eta_k}(y_t | 0, \Sigma_k(u_t)) \right] \end{aligned} \quad (60)$$

$$= -\ln \frac{(\eta_k \pi)^{\frac{D_y}{2}}}{\Gamma(\frac{D_y}{2})} B\left(\frac{D_y}{2}, \frac{\eta_k}{2}\right) - \frac{\eta_k + D_y}{2} \left( \psi\left(\frac{\eta_k + D_y}{2}\right) - \psi\left(\frac{\eta_k}{2}\right) \right) - \frac{1}{2} \ln |\Sigma_k(u_t)|. \quad (61)$$

where  $B(\cdot)$ ,  $\Gamma(\cdot)$ ,  $\psi(\cdot)$  are the beta, gamma and digamma functions, respectively. Note that only the last term depends on  $u_t$ . The cross-entropy from posterior predictive to goal distribution is:

$$\begin{aligned} \mathbb{E}_{p(y_t | u_t, \mathcal{D}_k)} \left[ -\ln p(y_t | y_*) \right] \\ = \frac{1}{2} \left( \ln 2\pi |S_*| + \mathbb{E}_{\mathcal{T}_{\eta_t}(y_t | \mu_t(u_t), \Sigma_t(u_t))} \left[ (y_t - m_*)^\top S_*^{-1} (y_t - m_*) \right] \right) \end{aligned} \quad (62)$$

$$= \frac{1}{2} \ln 2\pi |S_*| + \frac{1}{2} \text{Tr} \left[ S_*^{-1} \left( \Sigma_t(u_t) \frac{\eta_t}{\eta_t - 2} + \Xi(u_t) \right) \right], \quad (63)$$

where  $\Xi(u_t) = (\mu_t(u_t) - m_*)(\mu_t(u_t) - m_*)^\top$ . Note that the first term is constant.

## References

1. De Vries, B., Friston, K.J.: A factor graph description of deep temporal active inference. *Frontiers in Computational Neuroscience* **11**(95) (2017)
2. Friston, K., Kilner, J., Harrison, L.: A free energy principle for the brain. *Journal of Physiology* **100**(1-3), 70–87 (2006)
3. Friston, K.J., Parr, T., de Vries, B.: The graphical brain: Belief propagation and active inference. *Network Neuroscience* **1**(4), 381–414 (2017)
4. Hewitt, C.: Viewing control structures as patterns of passing messages. *Artificial Intelligence* **8**(3), 323–364 (1977)
5. Hoffmann, C., Rostalski, P.: Linear optimal control on factor graphs—a message passing perspective—. *IFAC-PapersOnLine* **50**(1), 6314–6319 (2017)
6. Kotz, S., Nadarajah, S.: Multivariate T-distributions and their applications. Cambridge University Press (2004)
7. Kouw, W.M.: Information-seeking polynomial NARX model-predictive control through expected free energy minimization. *IEEE Control Systems Letters* **8**, 37–42 (2023)
8. van de Laar, T., Koudahl, M., van Erp, B., de Vries, B.: Active inference and epistemic value in graphical models. *Frontiers in Robotics and AI* **9** (2022)
9. van de Laar, T., Koudahl, M., de Vries, B.: Realizing synthetic active inference agents, part II: Variational message updates. *Neural Computation* **37**(1), 38–75 (2024)
10. van de Laar, T.W., de Vries, B.: Simulating active inference processes by message passing. *Frontiers in Robotics and AI* **6**(20) (2019)
11. Loeliger, H.A.: An introduction to factor graphs. *IEEE Signal Processing Magazine* **21**(1), 28–41 (2004)
12. Loeliger, H.A., Dauwels, J., Hu, J., Korl, S., Ping, L., Kschischang, F.R.: The factor graph approach to model-based signal processing. *Proceedings of the IEEE* **95**(6), 1295–1322 (2007)
13. MacKay, D.J.: Information theory, inference and learning algorithms. Cambridge University Press (2003)
14. Nisslbeck, T.N., Kouw, W.M.: Factor graph-based online Bayesian identification and component evaluation for multivariate autoregressive exogenous input models. *Entropy* **27**(7) (2025)
15. Nisslbeck, T.N., Kouw, W.M.: Online Bayesian system identification in multivariate autoregressive models via message passing. In: *IEEE European Control Conference* (2025)
16. Palmieri, F.A., Pattipati, K.R., Di Gennaro, G., Fioretti, G., Verolla, F., Buonanno, A.: A unified view of algorithms for path planning using probabilistic inference on factor graphs. *arXiv:2106.10442* (2021)
17. Parr, T., Markovic, D., Kiebel, S.J., Friston, K.J.: Neuronal message passing using mean-field, Bethe, and marginal approximations. *Scientific Reports* **9**(1889) (2019)
18. Parr, T., Pezzulo, G., Friston, K.J.: Active inference: the free energy principle in mind, brain, and behavior. MIT Press (2022)
19. Särkkä, S.: Bayesian filtering and smoothing. Cambridge University Press (2013)
20. Şenöz, İ., van de Laar, T., Bagaev, D., de Vries, B.: Variational message passing and local constraint manipulation in factor graphs. *Entropy* **23**(7) (2021)
21. Soch, J., Faulkenberry, T.J., Petrykowski, K., Allefeld, C.: The Book of Statistical Proofs (2021). <https://doi.org/https://doi.org/10.5281/zenodo.4305950>

22. de Vries, B., Nuijten, W., van de Laar, T., Kouw, W., Adamiat, S., Nisslbeck, T., Lukashchuk, M., Nguyen, H.M.H., Araya, M.H., Tresor, R., Jenneskens, T., Nikoloska, I., Ganapathy Subramanian, R., van Erp, B., Bagaev, D., Podusenko, A.: Expected free energy-based planning as variational inference. arXiv:2504.14898 (2025)
23. Wainwright, M.J., Jordan, M.I., et al.: Graphical models, exponential families, and variational inference. *Foundations and Trends® in Machine Learning* **1**(1–2), 1–305 (2008)