

Confidence-Aware Decision-Making and Control for Tool Selection*

Ajith Anil Meera¹ and Pablo Lanillos^{1,2}

Abstract—Self-reflecting about our performance (e.g., how confident we are) before doing a task is essential for decision making, such as selecting the most suitable tool or choosing the best route to drive. While this form of awareness—thinking about our performance or metacognitive performance—is well-known in humans, robots still lack this cognitive ability. This reflective monitoring can enhance their embodied decision power, robustness and safety. Here, we take a step in this direction by introducing a mathematical framework that allows robots to use their control self-confidence to make better-informed decisions. We derive a mathematical closed-form expression for control confidence for dynamic systems (i.e., the posterior inverse covariance of the control action). This control confidence seamlessly integrates within an objective function for decision making, that balances the: i) performance for task completion, ii) control effort, and iii) self-confidence. To evaluate our theoretical account, we framed the decision-making within the tool selection problem, where the agent has to select the best robot arm for a particular control task. The statistical analysis of the numerical simulations with randomized 2DOF arms shows that using control confidence during tool selection improves both real task performance, and the reliability of the tool for performance under unmodelled perturbations (e.g., external forces). Furthermore, our results indicate that control confidence is an early indicator of performance and thus, it can be used as a heuristic for making decisions when computation power is restricted or decision-making is intractable. Overall, we show the advantages of using confidence-aware decision-making and control scheme for dynamic systems.

I. INTRODUCTION

Humans make decisions not only based on performance but also on their confidence in decisions [1]. For example, humans might prefer to use a known (confident) hammer to hit a nail, instead of a high-performance nail gun. Similarly, while driving, she might prefer a known route instead of an unknown path that might be optimal in terms of distance [2]. Thanks to this capability to self-assess our performance (metacognitive performance) [3], we make decisions that can handle (or adapt to) uncertainties [1] (e.g., a bent nail, road closings). Whilst confidence is an integral part of decision-making in humans [4], and can provide more decision power and robustness, it is not widely used in robotics. Decision-making in robotics usually optimizes a performance-based objective [5], such as the quadratic function of state and control for optimal control [6] (e.g., LQR).

The concept of confidence has recently gained relevance in robotic applications, such as autonomous driving [2], shared control [7], [8] and collision avoidance [9], where

they employed empirical and learning methods [10] to measure the confidence in decisions. These methods usually evaluate human predictability confidence and not the robot's confidence in executing its own actions. In essence, in any decision-making problem (e.g., selecting the best action or classifying images) the idea is to be able to compute the posterior uncertainty of making a decision [11]—providing the true underlying distribution. In control applications, we frame this confidence as how sure the robot is about making an optimal/best action. Importantly, while in discrete decision-making problems, there are ways to calculate these confidence estimates [12], computing them for continuous-time dynamical systems is less explored, as they mainly focus on the posterior distribution of predictions [13], or do not use the control confidence [14].

Here, we draw inspiration from the metacognitive approach to cognition [4], [15], to propose and derive an analytical expression to compute control confidence in dynamical systems. Thus, allowing the agent to use performance and confidence estimates to make better-informed decisions. To this end, we use a probabilistic approach to control, based on the Free Energy Principle (FEP) [16], [17], which permits the calculation of the control confidence as the posterior inverse covariance of the control action. Hence, we provide a solution in continuous time, state and action space that incorporates the robot's self-confidence within the decision-making process, taking the dynamics of the system into account. The core theoretical contributions are:

- 1) a closed-form expression for the control confidence,
- 2) a confidence-aware objective that balances between performance, control effort and control confidence,
- 3) a mathematical framework for decision-making and control using this objective.

These theoretical contributions are supported by its evaluation of the working example of the tool selection problem [18], where the robot has to select the tool that best performs the task. Without loss of generality, we define the set of selectable tools as a subset of possible robot arms—See Fig. 1. Statistical evaluation of this decision-making problem shows the usefulness of including control confidence. In particular, results show that the selected tool/arm using our proposed criteria is less sensitive to perturbations (while doing the task) than tools selected using the performance-based objective. Our results also point towards the use of control confidence as an early and easy-to-compute indicator of performance, without any prediction horizon unrolling. This increases its practical relevance, especially for fast decision-making under limited computation time/power.

*This research is supported by the Metatool project (Grant agreement 101070940) under the EIC pathfinder program. Corresponding author: ajith.anilmeera@donders.ru.nl

^{1,2}Donders Institute, Radboud University Nijmegen, The Netherlands

²Cajal International Center for Neuroscience, Spanish Research Council.

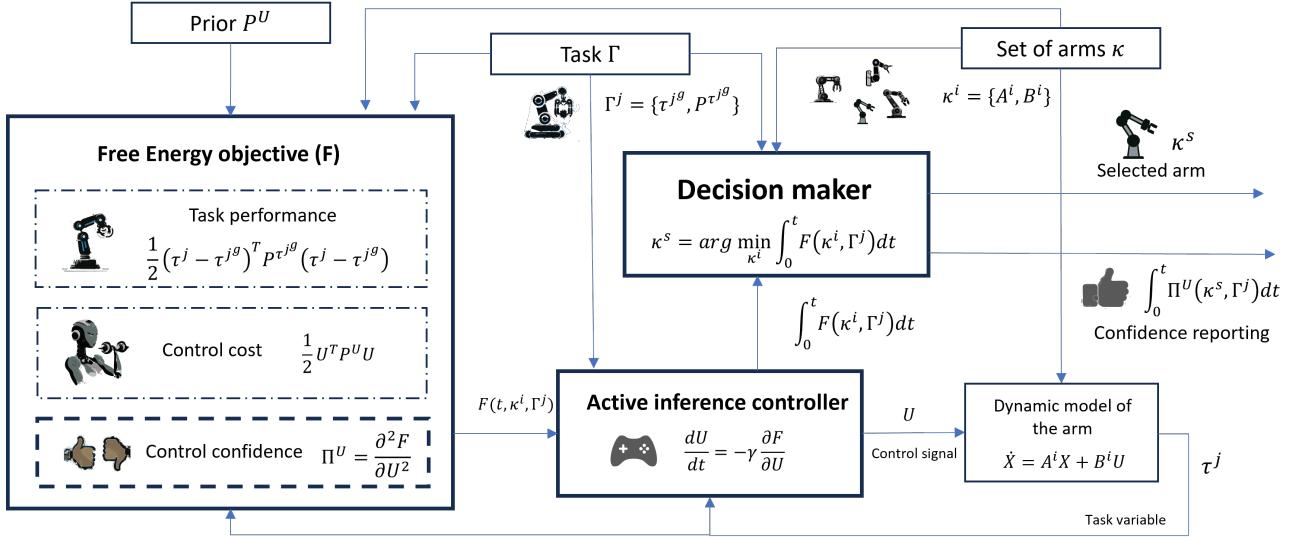


Fig. 1: The schematic of our proposed confidence aware decision making and control scheme. The robot makes a decision and reports its confidence. The decision making criteria simulates the controller to evaluate the free energy objective that balances task performance, control effort and control confidence. The self evaluation (performance) and self reflection on its decision (confidence) makes the robot metacognitive. The first two gradients of F fully describes our controller design (controller and control confidence), while the integral of F fully describes our decision maker. See Section III for the mathematical details.

II. PROBLEM FORMULATION

We use the tool selection problem as the driving example of confidence-aware decision-making. The tool selection problem consists of selecting a tool (defined as a robotic arm) from a set of p tools $\kappa = \{\kappa^1, \kappa^2, \dots, \kappa^p\}$, such that it best completes a task using its controller, by fulfilling the desired goal conditions. We restrict the set of possible tools to those whose dynamics can be modelled using a linear state space system of the form:

$$\dot{X} = AX + BU, \quad (1)$$

where $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times r}$ are the matrices defining the system dynamics, $U \in \mathbb{R}^{r \times 1}$ is the control input to the system and $X \in \mathbb{R}^{n \times 1}$ is the hidden state. Hence, every tool is fully described by its matrices $\kappa^i = \{A^i, B^i\}$.

We consider three types of tasks $\Gamma = \{\Gamma^1, \Gamma^2, \Gamma^3\}$ in terms of goal condition τ^g : reach a desired i) constant goal state X^g (e.g., reaching task), ii) constant goal state velocity \dot{X}^g (e.g., screw tightening tasks), and iii) constant goal acceleration \ddot{X}^g (e.g., constant force tasks like lifting, pushing). Thus, the problem reduces to select the tool κ^i , for a given task Γ^j , such that the task variable τ best reaches the desired goal τ^g within a tolerance (desired goal covariance) of $\Sigma^{\tau^g} = (P^{\tau^g})^{-1}$. The equations of motion of the tool (robot arm) used in the results are derived in Appendix A. In Appendix B, they are linearized to a form as given in (1).

III. CONFIDENCE-AWARE DECISION MAKING AND CONTROL

This section lays down the theoretical contributions that allow the robot to use confidence within the decision-making process. Figure 1 shows the schematic of our proposed

decision-making framework driven by the tool selection problem and the three main theoretical contributions: i) the design of a task-based probabilistic controller that minimizes the free energy objective, ii) the usage of this controller to derive a closed-form expression for control confidence, and iii) introduce a decision-making process that balances performance, control cost and control confidence. We fully describe our controller design and the control confidence using the first two gradients of free energy and our decision maker using the free energy integral.

We mathematically define control confidence—the robot’s confidence in the control action being the optimal one for task completion—as the posterior inverse covariance of the control action. To derive the closed-form expression of control confidence, we employ the mathematical framework of active inference, as it provides a way to evaluate it: the posterior inverse covariance is the curvature of the free energy objective.

A. Active inference controller design

We derive an active inference controller so that the closed-form expression for the control confidence (posterior inverse covariance of control action) can be computed. To this end, we introduce a novel form of the free energy objective from first principles aimed at task completion with high performance, minimal control effort and high confidence. This objective is minimized by the controller to select the best actions [19], [17].

Free energy optimization is a way to approximate Bayesian inference. The goal is to estimate the Bayesian optimal value for the parameter α such that it best explains the observation y . According to Bayes rule, the posterior

distribution $p(\alpha/y)$ of parameter α , given the measurement y is given by $p(\alpha/y) = p(\alpha, y)/p(y)$. Since the computation of $p(y) = \int p(y, \alpha)d\alpha$ is intractable for large search spaces of α , variational methods use a recognition density $q(\alpha)$ to closely approximate $p(\alpha/y)$ by minimizing the Kullback–Leibler (KL) divergence between both the distributions. This procedure results in the minimization of an objective function called free energy, given by [16]:

$$F = - \int q(\alpha) \ln p(y, \alpha)d\alpha + \int q(\alpha) \ln q(\alpha)d\alpha. \quad (2)$$

FEP is based on the idea that the Bayesian optimal estimate of α is the one that minimises the free energy objective F . The active inference controller (based on FEP) follows the gradient of F for the controller design.

1) *Task-based active inference*: We consider the problem of evaluating the control action U , to perform the task Γ^j , by controlling the task variable τ^j to reach the goal τ^{j^g} , within a desired level of uncertainty or prior covariance $\Sigma^\tau = (P^{\tau^g})^{-1}$. We proceed towards modelling the distributions inside (2), to compute F for our problem definition in Section II. We model the posterior distribution of control actions using a Gaussian distribution. So, the recognition density becomes $q(U) = \mathcal{N}(U : \mu^U, (\Pi^U)^{-1})$. The notation P is used for the prior precision (or inverse covariance) and Π is used for the posterior precision. We enable the robot to fuse its prior information into the scheme by modelling the prior distribution over control actions as a Gaussian distribution on U , written as $p(U) = \mathcal{N}(U : \eta^U, (P^U)^{-1})$. We assume that we can directly measure the task variable ($y = \tau^j$), and choose the control action as the unknown parameter to be computed ($\alpha = U$). We model the posterior distribution of the task variable under the robot’s action using a Gaussian distribution, given by $p(\tau^j/U) = \mathcal{N}(\tau^j : \tau^{j^g}, (P^{\tau^j})^{-1})$. The free energy in (2) is simplified using the distributions, and the constants are dropped to get (See Appendix C for the detailed derivation):

$$F = \underbrace{\frac{1}{2}(\tau^j - \tau^{j^g})^T P^{\tau^{j^g}} (\tau^j - \tau^{j^g})}_{\text{performance}} + \underbrace{\frac{1}{2}U^T P^U U}_{\text{control cost}} - \underbrace{\frac{1}{2} \ln |\Pi^U|}_{\text{confidence}}. \quad (3)$$

The resulting free energy objective can be seen as a sum of three terms: i) performance measure (the prior precision weighted deviation of the task variable from the goal), ii) control cost (the prior precision weighted control effort), and iii) confidence measure (the level of confidence in the chosen control action). When $\eta^U = 0$, minimizing F implies, maximizing performance, minimizing control effort and maximizing confidence. This objective can be used to design the controller. Our agent optimizes F to choose the best control actions via gradient descent on free energy:

$$\frac{dU}{dt} = -\gamma \frac{\partial F}{\partial U}, \quad (4)$$

where γ is the learning rate.

2) *Robotic arm control*: Here we derive the task-based active inference controller for a robot arm to reach the task goal in joint space. The free energy of an agent that wants to achieve a particular goal angular position θ^g , goal angular velocity $\dot{\theta}^g$ and goal angular acceleration $\ddot{\theta}^g$, with goal precisions (inverse covariance P^{θ^g} , $P^{\dot{\theta}^g}$ and $P^{\ddot{\theta}^g}$), and control priors (η^U , P^U) is:

$$F = \frac{1}{2} \left[(\theta - \theta^g)^T P^{\theta^g} (\theta - \theta^g) + (\dot{\theta} - \dot{\theta}^g)^T P^{\dot{\theta}^g} (\dot{\theta} - \dot{\theta}^g) + (\ddot{\theta} - \ddot{\theta}^g)^T P^{\ddot{\theta}^g} (\ddot{\theta} - \ddot{\theta}^g) - \ln |\Pi^U| + (U - \eta^U)^T P^U (U - \eta^U) \right]. \quad (5)$$

The core tenet of this controller is that taking control action U that minimises F , the states of the system (θ , $\dot{\theta}$ and $\ddot{\theta}$) move close to their respective goals. When the control prior is set to $\eta^U = 0$, the controller tries to minimize the control effort (or magnitude) of U , where P^U modulates the weightage of the control effort within F . F is similar to the objective J^1 from optimal control [6], with an additional term of $\ln |\Pi^U|$, which describes the control confidence, i.e., the confidence of the agent on the chosen control action to take the system to the goal (task completion) optimally.

To compute the control actions we differentiate (5) and substitute it in (4). Thus, giving the update rule for the controller:

$$\dot{U} = -\frac{\partial \theta}{\partial U}^T P^{\theta^g} (\theta - \theta^g) - \frac{\partial \dot{\theta}}{\partial U}^T P^{\dot{\theta}^g} (\dot{\theta} - \dot{\theta}^g) - \frac{\partial \ddot{\theta}}{\partial U}^T P^{\ddot{\theta}^g} (\ddot{\theta} - \ddot{\theta}^g) - P^U (U - \eta^U). \quad (6)$$

The update equations of this controller to perform the three tasks are derived in Appendix E.

B. Control confidence

Inspired from the dynamics expectation maximization algorithm [20], [21], we derive a closed-form solution for the optimal posterior precision (inverse covariance) of control action as the curvature of F (derivation in Appendix D):

$$\Pi^U = \frac{\partial^2 F}{\partial U^2}. \quad (7)$$

The precision of control action indicates the degree of certainty that the controller has over its control action. Technically, it is the inverse covariance of the control action. Higher Π^U indicates tighter bounds (or uncertainty) on the control U , which intuitively means that the controller is confident about the control action, while a lower Π^U indicates a larger uncertainty over the control action, and lower confidence in control. In Section IV-C, we will show that a higher Π^U indicates a higher stability in tool performance under perturbations on the control signal U .

The closed form expression for the control confidence in (7), for position task can be derived as (see Appendix E for

¹ $J = \int_0^\infty (x^T Q x + u^T R u) dt$ where Q and R are weights.

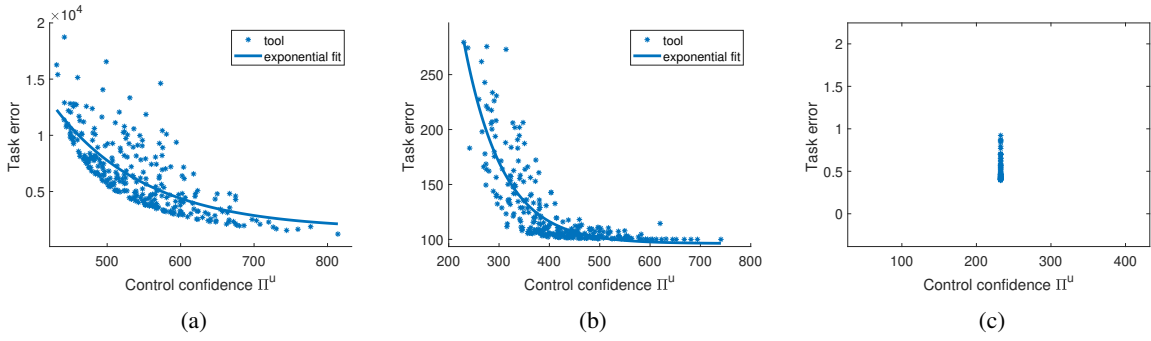


Fig. 2: Task error vs control confidence for 300 randomly selected tools (with exponential curve fit) for three tasks a) position task. b) velocity task and c) acceleration task. On average, the tools with higher control confidence performs better at the task than the tools with lower control confidence, for task 1 and 2. However, the controller for task 3 has the same confidence on all tools because Π^U in (9) is independent of the tool properties for a 2DOF arm (refer Appendix E.3).

full derivation):

$$\Pi^U = (A^{-1}(\frac{\partial \dot{X}}{\partial U} - B))^T P^{X^g} (A^{-1}(\frac{\partial \dot{X}}{\partial U} - B)) + P^{\dot{\theta}^g} + P^U, \quad (8)$$

The control confidence of the tool Π^U depends on i) the dynamics of the tool (A , B), ii) on the prior precisions of the goal (P^{X^g} , $P^{\dot{\theta}^g}$), and iii) the agent's prior control confidence (P^U). Since Π^U is independent of X and U , and is dependent on constants representing the robot arm and the priors of the agent, Π^U can be precomputed, and hence remains a constant while in operation. Therefore, the confidence in tool use can be evaluated before the tool use, increasing the practical relevance of this expression.

Similarly, we can derive the control confidence for velocity and acceleration control as:

$$\begin{aligned} \Pi^U &= ([I_2 \ O_2] A^{-1} B)^T P^{\dot{\theta}^g} ([I_2 \ O_2] A^{-1} B) + P^{\dot{\theta}^g} + P^U, \\ \Pi^U &= ([I_2 \ O_2] A^2 (A^{-1}(\frac{\partial \dot{X}}{\partial U} - B) + AB))^T P^{\ddot{\theta}^g} \\ &\quad ([I_2 \ O_2] A^2 (A^{-1}(\frac{\partial \dot{X}}{\partial U} - B) + AB)) + P^U. \end{aligned} \quad (9)$$

C. Confidence-aware decision making

This section introduces a confidence aware objective function that can be used for high level robot decision making. Again, we use tool selection as the driving example – selecting the best tool from a set of tools. Conventional methods follow a performance based objective where the error to goal and the control cost is penalized within the objective [6]. The main contribution of this work is the introduction of a control confidence term into this decision making criteria, enabling the agent to make high level decisions by considering their confidence in each decisions made. We propose the use of a high level decision making criteria that uses the performance measure, control cost and control confidence of a low level controller, to make high level decisions about the tool selection. The tool selection problem involves the selection of a tool (robot arm) κ^i from a set of tools $\kappa = \{\kappa^1, \kappa^2, \dots, \kappa^p\}$, that best performs the task

Γ^j . From Section B, we can fully describe a robot arm as a tool using its matrices $\kappa^i = \{A^i, B^i\}$. We propose a free energy based tool selection criteria to balance between the performance, control cost and control confidence as:

$$\kappa^s = \arg \min_{\kappa^i} \int_0^t F(\kappa^i, \Gamma^j) dt. \quad (10)$$

The chosen tool maximises the task performance with minimal control effort and maximum confidence in action, leading to the task completion. In addition to task performance, the agent is now aware of its self confidence in actions while using the tool, making the decision making process metacognitive. The next section will deal with demonstrating the usefulness of this criteria in simulation.

IV. RESULTS

We evaluate how control confidence affects decision-making under the tool selection problem. Results show the following:

- 1) Sec IV-B shows that control confidence is correlated with task performance. This means that it may be used as an early indicator of performance, for instance, to use it as a heuristic for tool comparison.
- 2) Sec IV-C shows that control confidence indicates the reliability of a tool for its task performance under a constant control perturbation.
- 3) Sec IV-D shows how different objective functions impact the decision making (i.e, selects different tools).
- 4) Sec IV-E statistically shows the advantage of using F over J for decision making – F chooses better tools.

A. Implementation details

All tools are considered to be 2DOF arms. The dynamic system used for the robot arm follows the derivations from Appendix A and B. The arms are defined by mass and length parameters $\{m_1, m_2, l_1, l_2\}$, which are randomly selected from the parameter domain $[0.1, 0.6]$. *ode45* in MATLAB was used to solve the differential equations of the controller and the dynamics for 5 seconds with a sampling time of $dt = 0.05$ with initial conditions of states as $X_0 = [-\frac{\pi}{2} \ -\frac{\pi}{2} \ 0]^T$

and of torque as $T_0 = [0.1 \ 0.1]^T$. Three tasks are considered: i) task 1 (position task) with goal position $\theta^g = [\frac{\pi}{3} \ \frac{\pi}{6}]^T$ and, zero goal velocity and acceleration, ii) task 2 (velocity task) with goal velocity $\dot{\theta}^g = [0 \ 2]^T$ and zero acceleration, and iii) task 3 (acceleration task) with goal acceleration $\ddot{\theta}^g = [0 \ 0.1]^T$. The goal precisions are set (coarse-tuned for stable control) as $P^{\theta^g} = \text{diag}(100, 100)$ for task 1, $P^{\dot{\theta}^g} = \text{diag}(50, 50)$, $P^{\ddot{\theta}^g} = \text{diag}(1, 1)$ and $P^u = \text{diag}(1, 1)$ for task 1 and 2 and, $P^{\ddot{\theta}^g} = \text{diag}(10, 10)$ and $P^u = \text{diag}(0, 0)$ for task 3. The performance of each tool on a given task is evaluated using (weighted) task error as the metric. For example, the task error for task 1 is $\sum_t (\theta - \theta^g)^T P^{\theta^g} (\theta - \theta^g)$. Higher the task error, lower the tool performance on the given task.

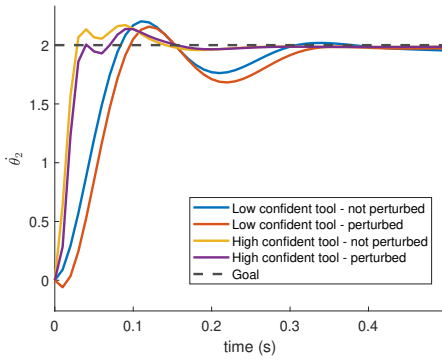


Fig. 3: High confident tool (yellow and purple) settles near the goal $\dot{\theta}_2^g$ (in dotted black) faster than the low confident tool (red and blue). The low-confident tool shows more variability in performance than the high-confident tool under a constant control perturbation of $\Delta T = -0.8Nm$ on the joints.

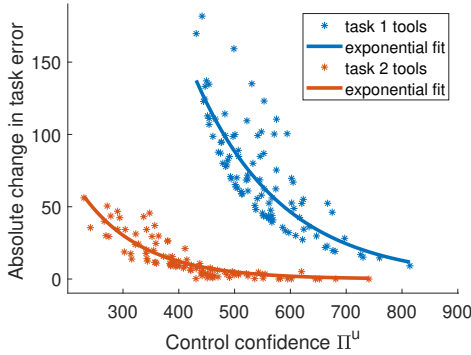


Fig. 4: Higher the control confidence on the tool, lower the absolute change in task error, under a control perturbation. This implies that the tool performance is less influenced by control perturbations for tools with higher Π^u . Therefore, the control confidence represents the reliability of the tool for task performance, under control perturbations.

B. Control confidence and tool performance

This section aims to establish the correlation between control confidence and tool performance in simulation. 300

random tools were sampled and the controller was evaluated for three tasks. Fig. 2 reports the clear correlation between the control confidence and the task error for tasks 1 and 2. On average, the tools with high control confidence performs better for tasks 1 and 2. However, for task 3, the controller is equally confident about all tools, which is in line with the similarity in their performance for all tools. In conclusion, control confidence can be used as an early indicator for tool performance.

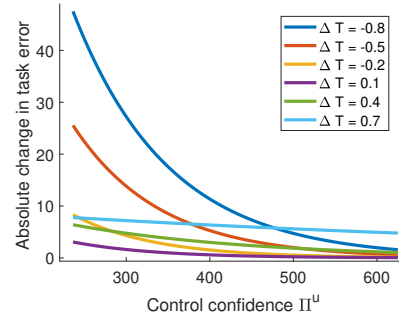


Fig. 5: Tools with low control confidence show high variability in task performance when subjected to different levels of constant control perturbations during the velocity task. Tools with high control confidence show relatively low variability.

C. Control confidence and control perturbations

This section aims to show that the performance of tools with high control confidence are less effected by control perturbations on the tool. Fig. 3 shows that the low confident tool shows high variability in performance than the high confident tool (under a constant control perturbation of $\Delta T = -0.8Nm$ on the joints). Moreover, the high confident tool reaches the goal faster than the low confident tool, highlighting the significance of the confidence. 100 random tools were sampled and the change in task error when the tool is subjected to a constant control perturbation of $\Delta T = -0.8$ were analysed for tasks 1 and 2. Fig. 4 shows that the tools with higher control confidence show less changes in task performance under a control perturbation, whereas the tools with lower control confidence show a higher variability in task performance. Fig. 5 shows that the trend remains the same for a range of perturbations, with higher variability when the perturbation resists the controller (perturbing torque is negative and opposes the controller torque). In conclusion, control confidence can be used as a measure of variability in task performance under control perturbations.

D. Tool selection criteria using control confidence

We show that our tool selection criterion produces different decision-making outcomes than just optimizing for performance, i.e., it selects different tools. While the objective function (usually) used in control engineering (J) balances between performance and control effort, our tool selection objective (\bar{F}) balances between performance, control effort and control confidence. Selecting a tool that the agent is

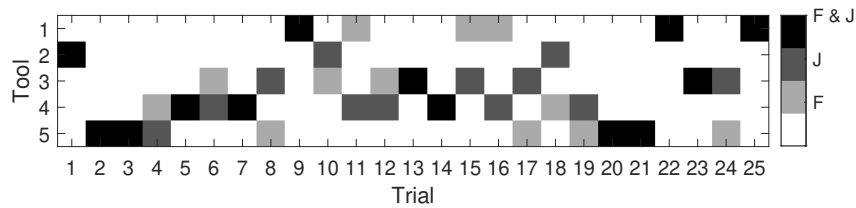


Fig. 6: Selected tool/arm depending on the objective function used for decision-making. The experiment uses 25 trials, with tools selected from a randomized set of 5 tools. Tool not selected (white), tool selected by F (light-grey), tool selected by J (dark-grey), tool selected by both F and J (black). The presence of grey blocks shows that both objectives can produce different decision outcomes, i.e., different tool selections. This is because, J might sometimes select a tool with low confidence, but F would prefer not to select it. Therefore, the tools selected by F has a better control confidence, and consequently, a better robustness to control perturbations (Figure 7).

confident about is an important aspect of tool selection. Fig. 6 shows the tools selected among the randomly generated toolset of 5 tools each for 25 trials. The presence of grey blocks indicates that both F and J do not select the same tool. Therefore, control confidence plays a significant role in selecting a different tool than the one selected through pure performance.

E. Benchmarking tool selection

We statistically show the advantage of our tool selection objective over the one used in control engineering. This is done by showing that the tool selected using F has a superior tolerance to control perturbations than the tool selected using J . Fig. 7 shows the benchmarking for tool selection for task 2 using three objectives: F , J and a pure confidence-based tool selection. 10 tools each were randomly selected for 50 trials and the performance of the tools selected by all three objectives were analysed. The tools selected by J showed lower control confidence and had a higher sensitivity to a constant control perturbation. The confidence-based tool selection showed similar confidence, task error and sensitivity to that of F . Therefore, control confidence can be used as an early predictor for tool selection, without simulating the controller performance on the tool.

V. CONCLUSION

Metacognitive performance evaluation (evaluation of one’s self-confidence in their decisions) is a fundamental cognitive feature that differentiates decision-making in humans and robots. We drew inspiration from this to introduce a mathematical framework for confidence-aware decision-making and control for tool selection. We derived a closed-form expression for the control confidence, which can be precomputed by the robot using the known system dynamics. Using this confidence, the proposed objective function that the decision-making employs seamlessly combines task performance, control cost and control confidence. Through both qualitative and statistical numerical experiments, we showed the advantages and properties of confidence-aware decisions in control. Our decision-making approach chose tools that presented less variability in task performance under control perturbations (external force) than the tools

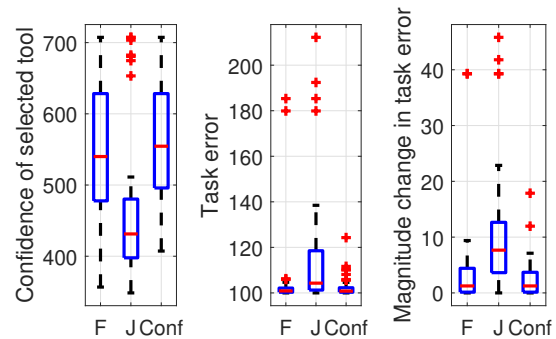


Fig. 7: The statistical analysis of our confidence-aware decision-making framework. We evaluate confidence, task error and the variability in task performance under a constant control perturbation (external joint torque). When compared to J , the F objective (on average) selects tools with a high control confidence (leftmost figure), and low variability in task performance under control perturbations (rightmost figure). The F and Conf objectives have a lower average task error (middle figure) than J , highlighting the importance of using confidence in decision making. The tools selected by F maintains the reliability of a pure confidence-based objective, without giving up on the performance.

selected by the performance-based objective from optimal control. Relevantly, we showed that the control confidence can be precomputed and that it is correlated with task performance, suggesting its use as an early indicator/heuristic of tool performance for a given task. Future research will focus on applying the proposed theoretical framework to realistic tasks, such as industrial applications. Besides, we are investigating how to online adapt the robot’s confidence during task execution.

REFERENCES

- [1] N. Yeung and C. Summerfield, “Metacognition in human decision-making: confidence and error monitoring,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, no. 1594, pp. 1310–1321, 2012.
- [2] F. Bontje, “Are you sure? modelling the confidence of a driver in left-turn gap acceptance decisions,” *TU Delft*, 2023.
- [3] S. M. Fleming, R. J. Dolan, and C. D. Frith, “Metacognition: computation, biology and function,” pp. 1280–1286, 2012.

- [4] S. M. Fleming and N. D. Daw, "Self-evaluation of decision-making: A general bayesian framework for metacognitive computation." *Psychological review*, vol. 124, no. 1, p. 91, 2017.
- [5] C. Parkan and M.-L. Wu, "Decision-making and performance measurement models with applications to robot selection," *Computers & Industrial Engineering*, vol. 36, no. 3, pp. 503–523, 1999.
- [6] B. D. Anderson and J. B. Moore, *Optimal control: linear quadratic methods*. Courier Corporation, 2007.
- [7] H. Saeidi, J. D. Opfermann, M. Kam, S. Raghunathan, S. Léonard, and A. Krieger, "A confidence-based shared control strategy for the smart tissue autonomous robot (star)," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1268–1275.
- [8] H. Saeidi and Y. Wang, "Trust and self-confidence based autonomy allocation for robotic systems," in *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 6052–6057.
- [9] D. Fridovich-Keil, A. Bajcsy, J. F. Fisac, S. L. Herbert, S. Wang, A. D. Dragan, and C. J. Tomlin, "Confidence-aware motion prediction for real-time collision avoidance," *The International Journal of Robotics Research*, vol. 39, no. 2-3, pp. 250–265, 2020.
- [10] Z. Cao, S. Xu, H. Peng, D. Yang, and R. Zidek, "Confidence-aware reinforcement learning for self-driving cars," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7419–7430, 2021.
- [11] J. Moon, J. Kim, Y. Shin, and S. Hwang, "Confidence-aware learning for deep neural networks," in *international conference on machine learning*. PMLR, 2020, pp. 7034–7044.
- [12] G. Papadopoulos, P. J. Edwards, and A. F. Murray, "Confidence estimation methods for neural networks: A practical comparison," *IEEE transactions on neural networks*, vol. 12, no. 6, pp. 1278–1287, 2001.
- [13] A. Lederer, Q. Hao, and S. Hirche, "Confidence regions for simulations with learned probabilistic models," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 3947–3952.
- [14] W. M. Kouw, "Information-seeking polynomial narx model-predictive control through expected free energy minimization," *IEEE Control Systems Letters*, 2023.
- [15] C. Hesp, R. Smith, T. Parr, M. Allen, K. J. Friston, and M. J. Ramstead, "Deeply felt affect: The emergence of valence in deep active inference," *Neural computation*, vol. 33, no. 2, pp. 398–446, 2021.
- [16] K. Friston, "The free-energy principle: a unified brain theory?" *Nature reviews neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.
- [17] P. Lanillos, C. Meo, C. Pezzato, A. A. Meera, M. Baioumy, W. Ohata, A. Tschantz, B. Millidge, M. Wisse, C. L. Buckley *et al.*, "Active inference in robotics and artificial agents: Survey and challenges," *arXiv preprint arXiv:2112.01871*, 2021.
- [18] A. Anil Meera and P. Lanillos, "Towards metacognitive robot decision making for tool selection," in *International Workshop on Active Inference*. Springer, 2023, pp. 31–42.
- [19] G. Oliver, P. Lanillos, and G. Cheng, "An empirical study of active inference on a humanoid robot," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 2, pp. 462–471, 2021.
- [20] A. Anil Meera and M. Wisse, "Dynamic expectation maximization algorithm for estimation of linear systems with colored noise," *Entropy*, vol. 23, no. 10, p. 1306, 2021.
- [21] K. J. Friston, N. Trujillo-Barreto, and J. Daunizeau, "Dem: a variational treatment of dynamic systems," *Neuroimage*, vol. 41, no. 3, pp. 849–885, 2008.

APPENDIX

A. Tool Dynamics (generative process)

This section details the mathematical model that describes the dynamics of a 2DOF robot arm (in the vertical plane), which will be used as the tool. The free body diagram of the links with input torques T_1 and T_2 applied to joints A and B is shown in Fig. 8. The force balance equation on link 1 in x and y directions is written as:

$$m_1 \ddot{y}_1 = V_A - V_B - m_1 g, \quad m_1 \ddot{x}_1 = H_A - H_B. \quad (11)$$

The moment balance equation at joint A can be written as:

$$\begin{aligned} \frac{1}{3} m_1 l_1^2 \ddot{\theta}_1 &= T_1 - T_2 + H_B l_1 \sin \theta_1 - V_B l_1 \cos \theta_1 \\ &\quad - \frac{1}{2} m_1 g l_1 \cos \theta_1. \end{aligned} \quad (12)$$

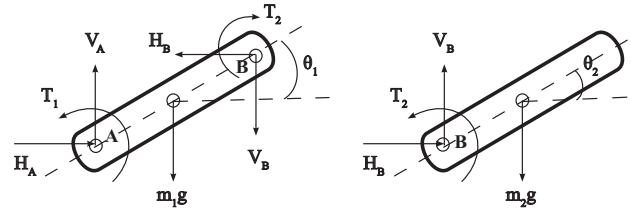


Fig. 8: Free body diagram of the links of a 2DOF robot arm.

Similarly, the force balance and moment balance equations for link 2 can be written as:

$$m_2 \ddot{y}_2 = V_B - m_2 g, \quad m_2 \ddot{x}_2 = H_B, \quad (13)$$

$$\frac{1}{3} m_2 l_2^2 \ddot{\theta}_2 = T_2 - \frac{1}{2} m_2 g l_2 \cos \theta_2. \quad (14)$$

The constraint equations for the arm to maintain its structural integrity is written as:

$$x_2 = l_1 \cos \theta_1 + \frac{1}{2} l_2 \cos \theta_2, \quad y_2 = l_1 \sin \theta_1 + \frac{1}{2} l_2 \sin \theta_2. \quad (15)$$

Differentiating it twice yields the relationship between angular and linear accelerations as:

$$\begin{aligned} \ddot{x}_2 &= -l_1 \dot{\theta}_1^2 \cos \theta_1 - l_1 \ddot{\theta}_1 \sin \theta_1 - \frac{l_2}{2} \dot{\theta}_2^2 \cos \theta_2 - \frac{l_2}{2} \ddot{\theta}_2 \sin \theta_2, \\ \ddot{y}_2 &= l_1 \ddot{\theta}_1 \cos \theta_1 - l_1 \dot{\theta}_1^2 \sin \theta_1 + \frac{l_2}{2} \ddot{\theta}_2 \cos \theta_2 - \frac{l_2}{2} \dot{\theta}_2^2 \sin \theta_2. \end{aligned} \quad (16)$$

Substituting (16) in (13) simplifies V_B and H_B , which upon substitution in (12) gives:

$$\begin{aligned} &\left(\frac{m_1}{3} + m_2\right) l_1^2 \ddot{\theta}_1 + \frac{1}{2} m_2 l_1 l_2 \dot{\theta}_2^2 \cos(\theta_1 - \theta_2) \\ &+ m_2 l_1^2 \dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + \frac{1}{2} m_2 l_1 l_2 \dot{\theta}_2^2 \sin(\theta_1 - \theta_2) \\ &+ T_2 - T_1 + \left(\frac{m_1}{2} + m_2\right) g l_1 \cos \theta_1 = 0. \end{aligned} \quad (17)$$

Equations (17) and (14) are the nonlinear differential equations that represent the dynamics of the 2 DOF robot arm, given the torques T_1 and T_2 . They represent the generative process of the system.

B. Generative model

This section derives the simplified model of the plant (from Section A) that the robot can use for the controller design and the decision making. This is done by linearizing (17) and (14) using the Jacobians around the equilibrium point to derive its linear state space form. Equation (14) can be rearranged as:

$$\ddot{\theta}_2 = \frac{3}{m_2 l_2^2} \left(T_2 - \frac{1}{2} m_2 g l_2 \cos \theta_2 \right), \quad (18)$$

which upon substitution in (17) results in:

$$\begin{aligned} \ddot{\theta}_1 &= - \left(\left(\frac{m_1}{3} + m_2 \right) l_1^2 \right)^{-1} \left[m_2 l_1^2 \dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + T_2 - T_1 \right. \\ &\quad + \frac{3 l_1}{2 l_2} \cos(\theta_1 - \theta_2) \left(T_2 - \frac{1}{2} m_2 g l_2 \cos \theta_2 \right) \\ &\quad \left. + \frac{1}{2} m_2 l_1 l_2 \dot{\theta}_2^2 \sin(\theta_1 - \theta_2) + \left(\frac{m_1}{2} + m_2 \right) g l_1 \cos \theta_1 \right] \end{aligned} \quad (19)$$

Using (19) and (18), a combined state space equation of the form $\dot{X} = f(X, U)$ can be written, where $X = [\theta_1 \ \theta_2 \ \dot{\theta}_1 \ \dot{\theta}_2]^T$ is the combined vector of joint angles and velocities, $U = [T_1 \ T_2]^T$ is the torque vector, and $f(X, U)$ is the combined function mapping. Here $f_1(X, U) = \dot{\theta}_1$, $f_2(X, U) = \dot{\theta}_2$, and $f_3(X, U)$ and $f_4(X, U)$ are on the right hand side of (19) and (18) respectively. Since f is nonlinear, we linearize it around the equilibrium point $X^e = [-\frac{\pi}{2} \ -\frac{\pi}{2} \ 0 \ 0]^T$ as:

$$\dot{X} = f(X^e, U^e) + A(X - X^e) + B(U - U^e), \quad (20)$$

where $A = \left. \frac{\partial f}{\partial X} \right|_{X^e, U^e}$ and $B = \left. \frac{\partial f}{\partial U} \right|_{X^e, U^e}$ are the Jacobians at the equilibrium point. At equilibrium point, $f(X^e, U^e) = 0$. Upon simplification, the linear state space equation of a 2DOF robot arm can be written in the form of (1) as:

$$\dot{X} = A(X - X^e) + BU, \quad (21)$$

where A and B are:

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{3(m_1+m_2)g}{l_1(m_1+3m_2)} & \frac{9m_2g}{4l_1(m_1+3m_2)} & 0 & 0 \\ 0 & -\frac{3g}{2l_2} & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{3}{l_1^2(m_1+3m_2)} & \frac{-3(3l_1+2l_2)}{2l_1^2l_2(m_1+3m_2)} \\ 0 & \frac{3}{m_2l_2^2} \end{bmatrix}.$$

This model is used by the agent for the controller design and the decision making.

C. Free energy derivation

The F from (3) can be written as a sum of two terms (V and H):

$$F = V + H = - \int q(U) \ln p(\tau, U) dU + \int q(U) \ln q(U) dU. \quad (22)$$

Here $q(U) = \mathcal{N}(U : \mu^U, (\Pi^U)^{-1})$, $p(U) = \mathcal{N}(U : \eta^U, (P^U)^{-1})$ and $p(\tau/U) = \mathcal{N}(\tau : \tau^g, (P^{\tau^g})^{-1})$, which can be written as:

$$\begin{aligned} q(U) &= \frac{1}{\sqrt{(2\pi)^r |(\Pi^U)^{-1}|}} e^{-\frac{1}{2}(U - \mu^U)^T \Pi^U (U - \mu^U)} \\ p(U) &= \frac{1}{\sqrt{(2\pi)^r |(P^U)^{-1}|}} e^{-\frac{1}{2}(U - \eta^U)^T P^U (U - \eta^U)} \\ p(\tau/U) &= \frac{1}{\sqrt{(2\pi)^n |(P^{\tau^g})^{-1}|}} e^{-\frac{1}{2}(\tau - \tau^g)^T P^{\tau^g} (\tau - \tau^g)} \end{aligned} \quad (23)$$

Considering a normalized recognition density $\int q(U) dU = 1$, and the fact that $\int q(U) (U - \mu^U)^T \Pi^U (U - \mu^U) dU = 1$, H in (22) reduces to $H = -\frac{1}{2} \ln |\Pi^U|$. The V in (22) can be simplified using the property $E = -\ln p(\tau, U) = -\ln p(\tau/U) - \ln p(U)$, and substituting (23) in it. E simplifies to (after dropping constants):

$$E = \frac{1}{2} (\tau - \tau^g)^T P^{\tau^g} (\tau - \tau^g) + \frac{1}{2} (U - \eta^U)^T P^U (U - \eta^U) \quad (24)$$

The relation between V and E can be expressed as $V(\tau, U) = \int dU q(U) E(\tau, U)$. Here, $E(\tau, U)$ can be approximated through the Taylor series expansion around the mean μ^U . So, $V(\tau, U) = \int dU q(U) [E(\tau, \mu^U) + \frac{\partial E(\tau, \mu^U)}{\partial U} (U - \mu^U) + \frac{1}{2} (U - \mu^U)^T \frac{\partial^2 E(\tau, \mu^U)}{\partial U^2} (U - \mu^U)]$. Since $\frac{\partial E(\tau, \mu^U)}{\partial U} = 0$ for free energy minimization, the second term drops to 0. Since $\int q(U) dU = 1$, $V(\tau, U) = E(\tau, \mu^U) + \frac{1}{2} \int dU q(U) (U - \mu^U)^T \frac{\partial^2 E(\tau, \mu^U)}{\partial U^2} (U - \mu^U)$, which reduces to $V(\tau, U) = E(\tau, \mu^U) + \frac{1}{2} \text{tr} \left((\Pi^U)^{-1} \frac{\partial^2 E(\tau, \mu^U)}{\partial U^2} \right)$. Using (24) with $\eta^U = 0$, V can be written as $V = \frac{1}{2} (\tau - \tau^g)^T P^{\tau^g} (\tau - \tau^g) + \frac{1}{2} U^T P^U U + W$, where $W = \frac{1}{2} \text{tr} \left((\Pi^U)^{-1} \frac{\partial^2 E(\tau, \mu^U)}{\partial U^2} \right)$. Substituting it in (22) after dropping constants gives:

$$F = \frac{1}{2} (\tau^j - \tau^{j^g})^T P^{\tau^{j^g}} (\tau^j - \tau^{j^g}) + \frac{1}{2} U^T P^U U - \frac{1}{2} \ln |\Pi^U| + W. \quad (25)$$

Since most of the FEP literature neglects W in the free energy term, we will drop it for calculating F . However, we will keep W in F in the next appendix. Refer [20], [21] for the details of a similar mathematical procedure (variational inference) to evaluate F .

D. Control confidence

According to FEP, the free energy gradient is zero at optimal estimates. Therefore, the free energy gradient with respect to the covariance of action ($\Sigma^U = (\Pi^U)^{-1}$) is zero for optimal action U , i.e., $\frac{\partial F}{\partial \Sigma^U} = 0$. Differentiating (25) with Σ^U and equating it to zero yields:

$$\frac{\partial F}{\partial \Sigma^U} = \frac{1}{2} \Pi^U - \frac{1}{2} \frac{\partial^2 E}{\partial U^2} = 0, \implies \Pi^U = \frac{\partial^2 E}{\partial U^2}. \quad (26)$$

Since F in (3) has the same components of U as that of E in (24), we abuse the notation to write $\Pi^U = \frac{\partial^2 F}{\partial U^2}$.

E. Task specific controllers

In this work, three variations of active inference controllers are introduced for three tasks: i) position control for task 1, ii) velocity control for task 2, and iii) acceleration control for task 3.

1) *Task 1 (position control)*: Task 1 involves controlling the robot arm to move to a goal configuration θ^g with zero goal velocity and acceleration. The controller uses all four terms of (6) with $\dot{\theta}^g = \ddot{\theta}^g = \eta^U = [0 \ 0]^T$ as:

$$\dot{U} = -\frac{\partial X}{\partial U}^T \begin{bmatrix} P^{\theta^g} & 0 \\ 0 & P^{\dot{\theta}^g} \end{bmatrix} \begin{bmatrix} \theta - \theta^g & 0 \\ 0 & \dot{\theta} \end{bmatrix} - \frac{\partial \ddot{\theta}}{\partial U}^T P^{\ddot{\theta}^g} \ddot{\theta} - P^U U \quad (27)$$

The goal is to find the unknowns $\frac{\partial X}{\partial U}$ and $\frac{\partial \ddot{\theta}}{\partial U}$. We assume a zero jerk motion for the robot using $\frac{\partial \ddot{\theta}}{\partial U} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\frac{\partial \dot{X}}{\partial U} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Differentiating (21) with respect to U gives the remaining unknown in the controller design as $\frac{\partial \dot{X}}{\partial U} = A^{-1} (\frac{\partial \dot{X}}{\partial U} - B)$.

2) *Task 2 (velocity control)*: Task 2 involves controlling the robot arm to maintain a goal velocity $\dot{\theta}^g$ with zero acceleration. The controller uses the last three terms of (6) with $\ddot{\theta}^g = \eta^U = [0 \ 0]^T$. The goal is to find the unknowns $\frac{\partial \dot{\theta}}{\partial U}$ and $\frac{\partial \ddot{\theta}}{\partial U}$. Differentiating 21 with U , and assuming $\frac{\partial \dot{X}}{\partial U} = 0$ gives $\frac{\partial \dot{\theta}}{\partial U} = -[I_2 \ O_2] A^{-1} B$. Assuming jerk free motion gives the last unknown in the update equation as $\frac{\partial \ddot{\theta}}{\partial U} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

3) *Task 3 (acceleration control)*: Task 3 involves controlling the robot arm to maintain a goal acceleration $\ddot{\theta}^g$. The controller uses the last two terms of (6) with $\eta^U = [0 \ 0]^T$. The goal is to find the unknown $\frac{\partial \ddot{\theta}}{\partial U}$. Differentiating (21) with t , and substituting (21) in it yields $\ddot{X} = A^2 (X - X^e) + ABU + B\dot{U}$. Differentiating it with U and using $\frac{\partial \dot{U}}{\partial U} = 0$ gives $\frac{\partial \ddot{X}}{\partial U} = A^2 \frac{\partial X}{\partial U} + AB$. From Section E.1 we know $\frac{\partial \dot{X}}{\partial U} = A^{-1} (\frac{\partial \dot{X}}{\partial U} - B)$ and $\frac{\partial \dot{X}}{\partial U} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Therefore, $\frac{\partial \ddot{\theta}}{\partial U} = [I_2 \ O_2] \frac{\partial \dot{X}}{\partial U} = [I_2 \ O_2] A \frac{\partial \dot{X}}{\partial U} = I_2$. Substituting $\frac{\partial \dot{X}}{\partial U}$ in (9) upon evaluation yields $\Pi^U = P^{\ddot{\theta}^g} + P^U$, which is a constant that is independent of the tool properties.