

# Inferring Perception from Continuous Control

Joshua Jun Hwan Ryu  
jhryu25@stanford.edu  
Department of Psychology, ICME  
Stanford University

June. 4. 2021

## 1 Introduction

Imagine catching an annoying fly in your room. This is sometimes an amazingly difficult task that requires careful perceptual and control mechanisms. In order to successfully accomplish this task, the human agent needs to integrate uncertain time-to-time information about the fly's position. At a given split millisecond, one might have almost no idea of where the fly is, however, as more visual information makes its way through the eyes, the temporal information are properly integrated in order for one to give rise to the perception of a fly. One then needs to use this perceptual information in order to make a series of control movements: track the fly with our eyes and then eventually swat the fly with our hands. Visual information available to us is highly uncertain for a short time frame, and rich prior information from the previous information need to be integrated with the constantly changing incoming sensory information. The goal of the visual system is to infer the most likely state of the external world by integrating information from the past with the present.

In this project, I show the plausibility of the parameter estimation using simulated data. I outline the mathematical framework with given model, and the distribution of the control sequence under the given model. The inference and estimation are implemented through the EM-algorithm, maximizing the evidence lower bound (elbo). Since the model is completely Gaussian, it allows for a closed-form solution to the marginal distributions on the latents and the elbo computation. The outline of the derivations are in the appendix. I derive a closed form for the distribution of the latent states using Kalman smoothing, and then use these distribution to compute the elbo in closed form. Once the elbo is computed, I use tensorflow's autodifferentiation tools to optimize the parameters.

## 2 Task

The dataset consists of the performance various human subjects tracking a gaussian blob embedded in noise with matched spatial frequencies with a cursor controlled by either a mouse (n=5) or a trackpad (n=3). The stimulus undergoes a Brownian motion of a fixed standard deviation (which I call velocity). Various combinations of stimulus uncertainty (3 levels) and stimulus velocity (3 levels) are tested. At each condition, each subjects undergo 15 trials at a given condition. Last 10 trials are analyzed. **Input:** target blob positions (15 seconds at 60hz). **Output:** human control movements (mouse positions)

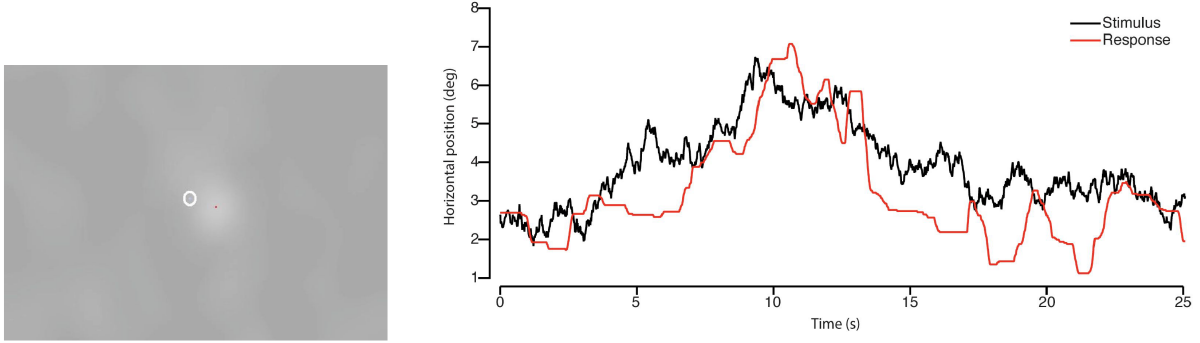


Figure 1: (Left) Sample frame from the task. The stimulus is a Gaussian blob embedded in noise with matched spatial frequencies. The subjects are instructed to fixate and track the blob with a cursor. (Right) Target (black) and cursor position (red) from a sample trial of the task.

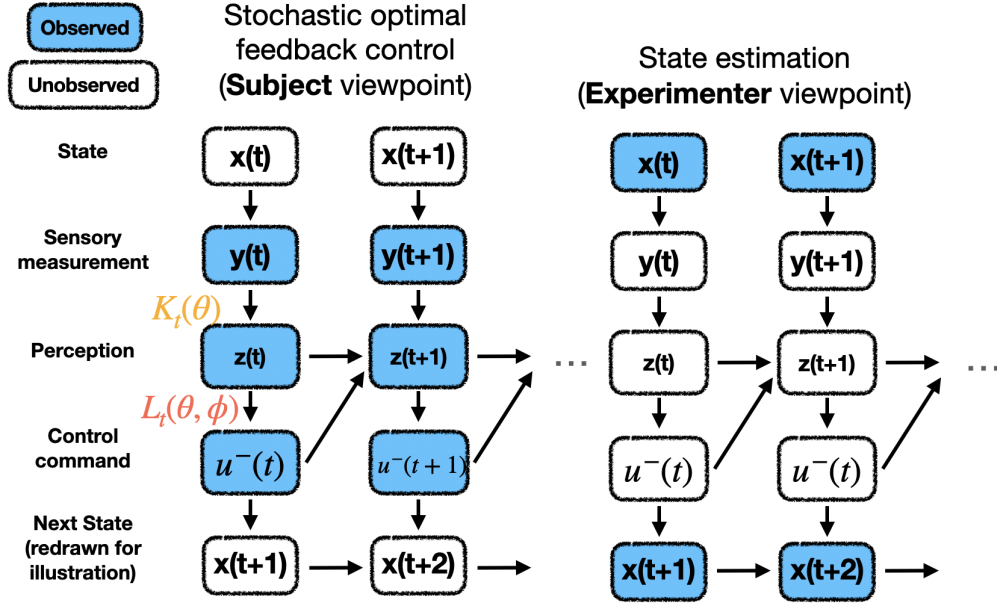


Figure 2: The generative model. The model assumes that the observer makes optimal state estimation from noisy measurements of the world for perception via using Kalman filters. The estimated states are then used to make an optimal Linear- Quadratic-Gaussian control. The experimenter, sees the output of the control and the exact state of the world. The goal of the experimenter is to find the perception and control parameters that have generated the data.

### 3 The Generative model

We are interested in the following variables:

- $x_t$  := external state of the world (x,y position, velocity and force for stimulus and tracker)
- $y_t$  := (latent) observer's noisy measurements of the state (position of the stimulus and tracker)
- $z_t$  := (latent) observer's perceptual estimation of the state (position estimates)
- $u_t$  := observer's exerted force

We can describe the generative process using the following dynamical system:

World	$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \boldsymbol{\xi}_t$
Measurement	$\mathbf{y}_t = H\mathbf{x}_t + \boldsymbol{\omega}_t$
Perception	$\mathbf{z}_{t+1} = A\mathbf{z}_t - BL(t, \theta, \phi)\mathbf{z}_t + K(t+1, \theta)(\mathbf{y}_t - H\mathbf{z}_t) + \boldsymbol{\epsilon}_t$
Control	$\mathbf{u}_t = -L(t, \theta, \phi)\mathbf{z}_t + \boldsymbol{\epsilon}_t$

The noise terms are all stationary, independent and normally distributed.

### 4 EM Algorithm

To estimate the parameters, I maximize the elbo  $L(q, \theta, \phi$ :

$$\begin{aligned}
\log P(\{x_t\}_{t=1}^T, \theta, \phi) &= \log P(\{x_t\}|\theta, \phi) + \log P(\theta, \phi) \\
&\geq \mathbb{E}_{q(\{z_{t-1}\})} [\log P(\{x_t, z_{t-1}\}|\theta, \phi) + \log q(\{z_{t-1}\})] + \log P(\theta, \phi) \\
&\triangleq L(q, \theta, \phi)
\end{aligned}$$

In the expectation step, we find the probability of the latent variables given all the observations  $p(z_t|x_{1:T})$ . This is done through Kalman smoothing (check appendix for details).

In the maximization step, we use the expectations to calculate the elbo (check appendix for details). Then, assuming that the perceptual and control filters are differentiable with respect to the parameters, we use autodifferentiation tools in tensorflow to run gradient descent on the negative elbo.

### 5 Inference and Estimation

In this section, we check whether the algorithm is able to recover the true parameters. For each simulation, we simulate the model behavior from a set of true parameters. In figure 5, we show the results of the E-step of the algorithm with the true parameters. The latent state estimates are highly dependent on the control parameters and noise. Observing the control that arises from the latent state significantly decreases the uncertainty of the latent state (e.g. the filtered and the smoothed marginals as opposed to prior marginals).

Having checked that the algorithm recovers reasonable latent state estimates, we then run the E-step and M-step with random initialization of the parameters on the data simulated with true parameters. While there seems to be some numerical instabilities (yet to figure out.), for many simulated data, we see that the elbo for the true parameter is the highest. Figure 5 shows the results for one such simulation. Finally figure 5 shows the results of the gradient ascent on the elbos for different model initializations.

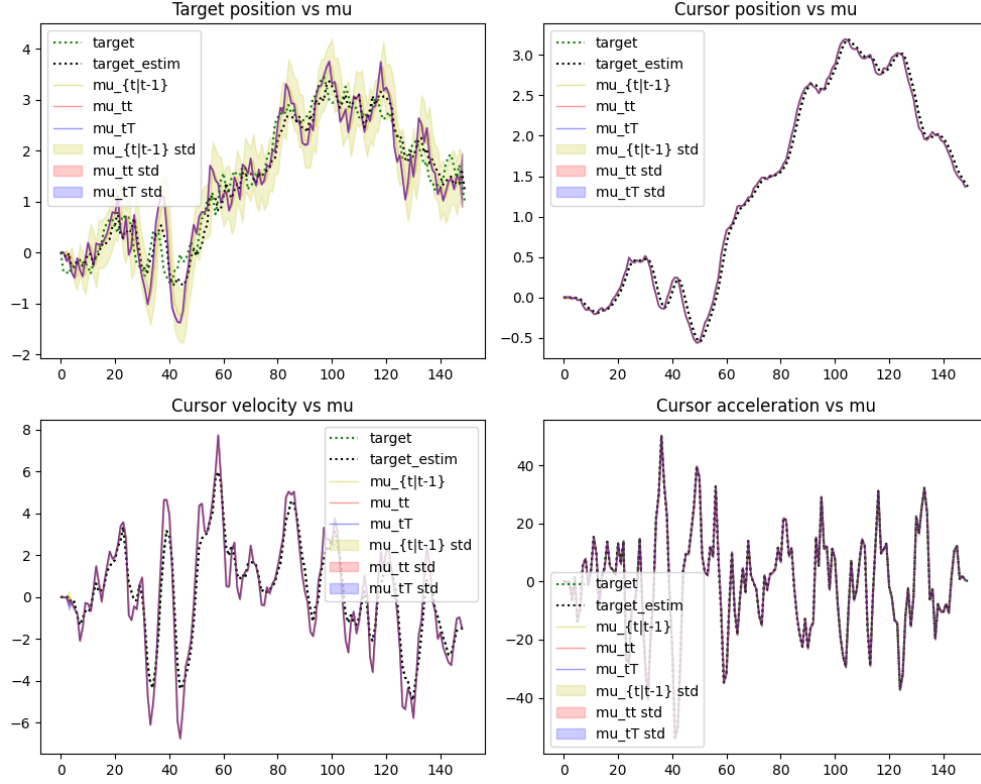


Figure 3: Probability and 3 std of state perception  $p(z_t)$  under the model with true parameters. (Green dotted): Simulated target trajectory. (Black dotted): Simulated target perception. (Yellow) prior for filtered marginals:  $p(z_t|x_{1:t})$ , given true parameters. (Red) filtered marginals:  $p(z_t|x_{1:t+1})$ , given true parameters. (Blue) smoothed marginals:  $p(z_t|x_{1:T})$ . Each panel indicate the different dimensions of the perceptual state (target position, cursor position, cursor velocity, cursor acceleration).

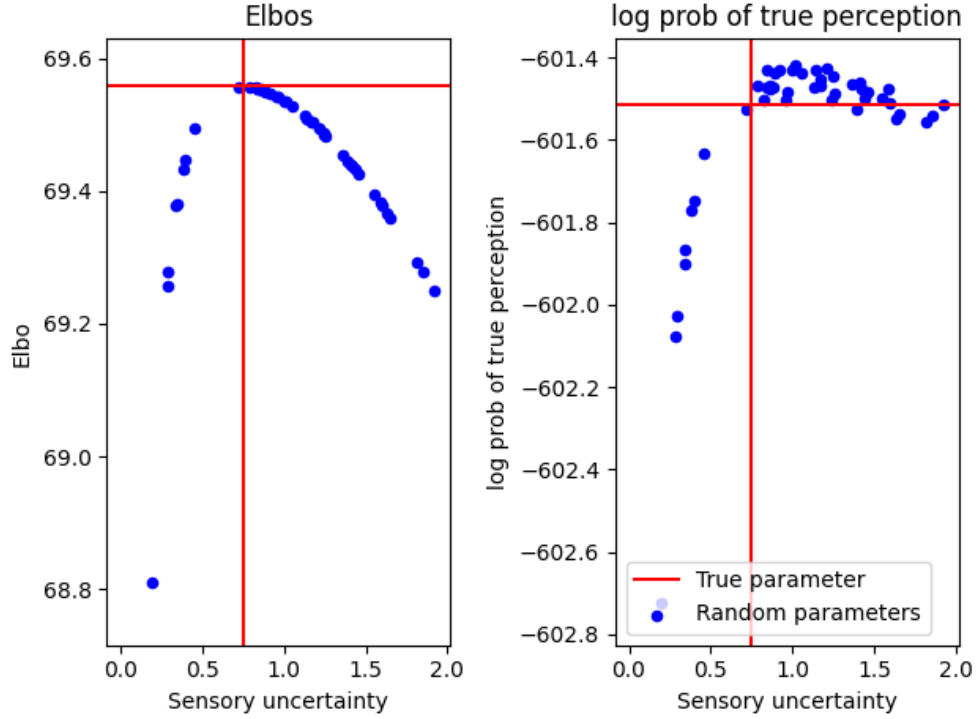


Figure 4: Testing the elbo calculation on simulated data. We simulated ideal observer behavior from a set of true parameters, and used suboptimal models with random parameter initializations to calculate elbo and to check the probability of getting the simulated behavior with the random parameters. (Left) Elbo per timepoint for models with different parameter values. The model with true parameters has the highest elbo, for this specific true parameter. (Right) log probability of simulated perception  $z_t$  under the calculated smoothed marginals  $p(z_t|x_{1:T})$  for different model parameters.

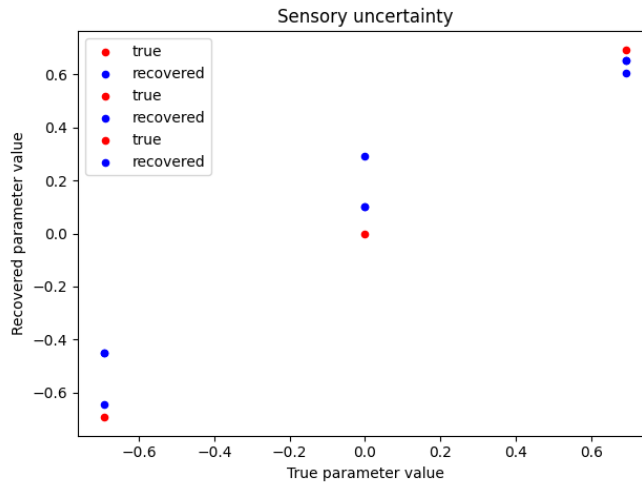


Figure 5: We initialized the models with random sensory parameters and initialized the control parameter at the true value. We plot the recovered parameters of model with 10 iterations of the EM algorithm.

## 6 Appendix: Model assumptions and details

I summarize some of the assumptions of the model here:

- Humans receive very noisy visual information about the world in a small time frame.
- Humans continuously integrate noisy visual information in an optimal fashion, given the world dynamics, weighing its uncertainty via the Kalman filter.
- Perception and control are modular (e.g. Firestone and Scholl, 2016 [7]). Planning for movement does not directly change our perceptual processes.
- Human control is limited by our biology and control noise. Our arms cannot move at infinite speeds and are parsimonious with respect to making movements. Motor movement have been shown to minimize jerk or endpoint variance., and are subject to multiplicative motor noise (e.g. Wolpert, 2000 [4]). For the sake of the project, I add a quadratic loss on the control force and add a Gaussian control noise.
- Human control movements are approximately linear (e.g. Todorov, 2002 [5]).

For the scope of the paper, I omit the details of the perceptual and control filters, since the method employed is general and can be applied to any perception and control models that are differentiable with respect to the parameters. A version can be found in Todorov, 2005 [6], with the only difference being that multiplicative noise was omitted here for simplicity.

The noise are distributed as follows:

- Dynamics noise (determined by experimenter)  $\xi_t \sim \mathcal{N}(0, \Omega_x)$
- Measurement noise  $\omega_t \sim \mathcal{N}(0, \Omega_y)$
- Perceptual estimation noise  $z_t \sim \mathcal{N}(0, \Omega_z)$
- Control noise  $\varepsilon_t \sim \mathcal{N}(0, \Omega_u)$

The dynamics are determined by:

- Dynamics:  $A, B$  determine the stimulus dynamics and control gain
- Measurement:  $H$  determine the sensory gain
- Perception: Perception is determined by the dynamics  $A, B$  and the Kalman filter  $K(t, \theta)$ . This is the optimal linear filter for perception, given that that observer has full access to their intended control and the external dynamics.
- Control: Control is determined by the optimal linear filter  $L(t, \phi)$  from the linear-quadratic-gaussian (LQG) controller. The controller minimizes an expected tracking error subject to control costs.

We can convert the dynamical systems to conditional distributions on the variables of interest:

$$p(u_t \mid z_t; \theta, \phi) = \mathcal{N}(u_t \mid -L(t, \theta, \phi)z_t, \Omega_u)$$

$$p(\mathbf{x}_{t+1} \mid \mathbf{z}_t, \mathbf{x}_t; \theta, \phi) = \int_{u_t} p(\mathbf{x}_t \mid u_t, \mathbf{z}_t, \mathbf{x}_t; \theta, \phi) p(u_t \mid \mathbf{z}_t; \theta, \phi)$$

$$\begin{aligned}
& \sim \int_{u_t} \mathcal{N}(x_{t+1} | Ax_t + Bu_t, \Omega_x) \mathcal{N}(u_t | -L(t, \theta, \phi)z_t, \Omega_u) \\
& \sim \mathcal{N}(x_{t+1} | Ax_t - BL(t, \theta, \phi)z_t, \Omega_x + B\Omega_u B^T) \\
p(\mathbf{z}_t | \mathbf{z}_{t-1}; \mathbf{x}_t, \theta, \phi) &= \int_{y_t} p(\mathbf{z}_t | \mathbf{z}_{t-1}, \mathbf{y}_t; \mathbf{x}_t, \theta, \phi) p(\mathbf{y}_t; \mathbf{x}_t, \theta, \phi) \\
&\sim \int_{y_t} \mathcal{N}(z_t | Az_{t-1} - BL(t-1, \theta, \phi)z_{t-1} + K(t, \theta)(y_t - H\mathbf{z}_{t-1}), \Omega_z) \mathcal{N}(y_t | H\mathbf{x}_t, \Omega_y) \\
&\sim \mathcal{N}(Az_{t-1} - BL(t-1, \theta, \phi)z_{t-1} + K(t, \theta)H(\mathbf{x}_t - \mathbf{z}_{t-1}), \Omega_z + K(t, \theta)\Omega_y K(t, \theta)^T) \\
&\triangleq \mathcal{N}(a_{z_t|t-1}z_{t-1} + b_{z_t|t-1}, \Omega_{z_t|t-1})
\end{aligned}$$

where

$$a_{z_t|t-1} := A - BL(t-1, \theta, \phi) - K(t, \theta)H$$

$$b_{z_t|t-1} := K(t, \theta)H\mathbf{x}_t$$

$$\Omega_{z_t|t-1} := \Omega_z + K(t, \theta)\Omega_y K(t, \theta)^T$$

Assume other prior distributions:

$$\begin{aligned}
p(z_0; \theta) &\triangleq \mathcal{N}(0, \Omega_{z_0}) \\
p(\theta) &\triangleq \mathcal{N}(\text{someval}, \text{largeval}\Omega_\theta) \\
p(\phi) &\triangleq \mathcal{N}(\text{someval}, \text{largeval}\Omega_\phi)
\end{aligned}$$

Since the elbo takes into consideration many timepoints and batches, the prior distribution on the parameters become almost negligible.

## 7 Appendix: Inference details

We are interested in finding the parameters that maximize the joint probability of seeing the observer control with the parameters  $\arg \max_{\{\theta, \phi\}} \log P(\{\mathbf{u}_t\}, \theta, \phi | \{x_t\})$ . In order to do the inference, we need to marginalize the latent variables and then optimize the objective. I adopt the EM algorithm for this, and alternate between approximating the marginal distribution on the latent variables (E-step) and optimizing the evidence lower bound (elbo):

$$\begin{aligned}
\log P(\{x_t\}_{t=1}^T, \theta, \phi) &= \log P(\{x_t\} | \theta, \phi) + \log P(\theta, \phi) \\
&\geq \mathbb{E}_{q(\{z_{t-1}\})} [\log P(\{x_t, z_{t-1}\} | \theta, \phi) + \log q(\{z_{t-1}\})] + \log P(\theta, \phi) \\
&\triangleq L(q, \theta, \phi)
\end{aligned}$$

(I omit the dependence on the stimulus  $x_t$ , and write them only as needed.)

The lower bound is tight when the variational distribution on the latents  $q(z_{1:T})$  is equal to the posterior. The posterior on the latents is intractable. However, for the E-step, it suffices to calculate the latent marginals needed for the M-step. Since the joint distribution given the parameters is:

$$\begin{aligned}
P(\{z_t\}_{t=0}^{T-1}, \{\mathbf{x}_t\}_{t=1}^T) &:= P(\{z_t\}, \{\mathbf{x}_t\} | \theta, \phi, \{x_t\}) \\
&= P(z_0) \prod_{t=0}^{T-1} P(z_{t+1} | z_t) \prod_{t=0}^T P(x_{t+1} | z_t)
\end{aligned}$$

The maximization of the elbo over the parameters amounts to:

$$\arg \max_{\theta, \phi} L(q, \theta, \phi) = \arg \max_{\theta, \phi} [\mathbb{E}_{q(z_{1:T})} [\log P(\{\mathbf{u}_t, z_t\} | \theta, \phi) + \log q(\{z_t\})] + \log P(\theta, \phi)] \quad (7.1)$$

$$= \arg \max_{\theta, \phi} [\mathbb{E}_{q(z_{1:T})} [\log P(\{\mathbf{u}_t, z_t\} | \theta, \phi)] + \log P(\theta, \phi)] \quad (7.2)$$

$$= \arg \max_{\theta, \phi} \left[ \mathbb{E}_{q(z_{1:T})} \left[ \log P(z_0) + \sum_{t=0}^{T-1} \log P(z_{t+1} | z_t) + \sum_{t=0}^T \log P(u_t | z_t) \right] \right] \quad (7.3)$$

$$+ \log P(\theta, \phi) \quad (7.4)$$

$$= \arg \max_{\theta, \phi} \left[ \mathbb{E}_{q(z_0)} \log P(z_0) + \sum_{t=0}^{T-1} \mathbb{E}_{q(z_{t+1}, z_t)} \log P(z_{t+1} | z_t) \right] \quad (7.5)$$

$$+ \sum_{t=0}^T \mathbb{E}_{q(z_t)} \log P(u_t | z_t) + \log P(\theta, \phi) \quad (7.6)$$

where in the last step, we have used linearity of expectations, and take expectation over the uncaptured variables.

Thus, in the E-step we need to update the marginals and the two-slice marginals:

$$\begin{aligned} q(z_t) &:= p(z_t | u_{1:T}, \theta, \phi) \\ q(z_t, z_{t+1}) &:= p(z_t, z_{t+1} | u_{1:T}, \theta, \phi) \end{aligned}$$

## 7.1 E-step

The marginals  $p(z_t | \{u_t\})$  are calculated using an analog of forward-backward algorithm for state-space models [2, 3].

Then the distributions of interest are given by:

$$p(z_t | u_{1:T}) = \int_{z_{t+1}} p(z_t | z_{t+1}, u_{1:T}) p(z_{t+1} | u_{1:T}) \quad (7.7)$$

$$= \int_{z_{t+1}} p(z_t | z_{t+1}, u_{1:t}, \underline{u_{t+1:T}}) p(z_{t+1} | u_{1:T}) \quad (7.8)$$

$$p(z_t, z_{t+1} | u_{1:T}) = p(z_t | z_{t+1}, u_{1:T}) p(z_{t+1} | u_{1:T}) \quad (7.9)$$

$$= p(z_t | z_{t+1}, u_{1:t}, \underline{u_{t+1:T}}) p(z_{t+1} | u_{1:T}) \quad (7.10)$$

The distributions of interest are Gaussian and  $\forall s, t$  we defined them as:

- $p(z_s | u_{1:t}) \sim \mathcal{N}(\mu_{s|t}, \Sigma_{s|t})$ .
- $p(z_s, z_{s+1} | u_{1:t}) \sim \mathcal{N}(\mu_{s,s+1|t}, \Sigma_{s,s+1|t})$ .

$p(z_t | z_{t+1}, u_{1:t})$  is found by computing  $p(z_t | u_{1:t})$  and  $p(z_{t+1} | u_{1:t})$  in the forward step, then taking the Gaussian conditional of their joint distribution:

$$p(z_t, z_{t+1} | u_{1:t}) = \mathcal{N} \left( \begin{bmatrix} z_t \\ z_{t+1} \end{bmatrix} \mid \begin{bmatrix} \mu_{t|t} \\ \mu_{t+1|t} \end{bmatrix}, \begin{bmatrix} \Sigma_{t|t} & \Sigma_{t|t} a_{z_{t+1}|t}^T \\ a_{z_{t+1}|t} \Sigma_{t|t} & \Sigma_{t+1|t} \end{bmatrix} \right) \quad (7.11)$$



$$p(z_t|z_{t+1}, u_{1:t}) = \mathcal{N}(\mu_{t|t} + J_t(z_{t+1} - \mu_{t+1|t}), \Sigma_{t|t} - J_t \Sigma_{t+1|t} J_t^T) \quad (7.12)$$

$$J_t \triangleq \Sigma_{t|t} a_{z_t|t-1}^T \Sigma_{t+1|t}^{-1} \quad (7.13)$$

$p(z_{t+1}|u_{1:T})$  can be computed recursively in the backward step. And similarly,

$$p(z_s, z_{s+1}|u_{1:T}) = \mathcal{N}\left(\begin{bmatrix} z_t \\ z_{t+1} \end{bmatrix} \mid \begin{bmatrix} \mu_{t|T} \\ \mu_{t+1|T} \end{bmatrix}, \begin{bmatrix} \Sigma_{t|T} & \Sigma_{t|T} a_{z_{t+1}|t}^T \\ a_{z_{t+1}|t} \Sigma_{t|T} & \Sigma_{t+1|T} \end{bmatrix}\right)$$

### 7.1.1 Forward step: Kalman filtering

Forward step involves calculating the filtered marginal probability of the current latent variable given the current and past observations  $p(z_t|u_{1:t})$ . While I do not derive Kalman filter here, I outline them in the appendix. Details can also be found in [2] and [1].

As an inductive assumption, assume that  $p(z_{t-1}|x_{1:t}) \sim \mathcal{N}(\mu_{t-1|t-1}, \Sigma_{t-1|t-1})$ . Then, the distributions can be computed recursively as:

$$\begin{aligned} p(z_t|x_{1:t}) &\triangleq \mathcal{N}(\mu_{t|t-1}, \Sigma_{t|t-1}) \\ \mu_{t|t-1} &\triangleq a_{z_t|t-1} \mu_{t-1|t-1} + b_{z_t|t-1} \\ \Sigma_{t|t-1} &\triangleq a_{z_t|t-1} \Sigma_{t-1|t-1} a_{z_t|t-1}^T + \Omega_{z_t|t-1} \\ p(z_t|x_{1:t+1}) &\triangleq \mathcal{N}(\mu_{t|t}, \Sigma_{t|t}) \\ \mu_{t|t} &\triangleq \mu_{t|t-1} + R_t r_t \\ \Sigma_{t|t} &\triangleq (I + R_t(-BL_t)) \Sigma_{t|t-1} \\ R_t &\triangleq \Sigma_{t|t-1}(-BL_t)^T (BL_t \Sigma_{t|t-1} (BL_t)^T + \Omega_x + B\Omega_u B^T)^{-1} \quad (\text{Kalman filter}) \\ r_t &\triangleq x_{t+1} - (Ax_t - BL_t \mu_{t|t-1}) \quad (\text{innovation}) \end{aligned}$$

### 7.1.2 Backward step: Kalman smoothing

The backward step involves calculating the smoothed marginals  $p(z_t|u_{1:T})$ . Again, I omit the derivations and just outline the results here. The results follow from equations 7.8 and 7.13. Details can also be found in references [2] and [1]. Unlike the forward backward algorithm in hidden markov models that computes  $p(x_{t+1:T}|z_t)$  in the backwards step, in the backwards step, it is more common and more computationally tractable to compute  $p(z_t|x_{t+1:T})$  directly [2].

As an inductive hypothesis, assume that  $p(z_{t+1}|x_{1:T}) \sim \mathcal{N}(\mu_{t+1|T}, \Sigma_{t+1|T})$ . Then, the smoothed marginals can be computed recursively as:

$$\begin{aligned} p(z_t|x_{1:T}) &\triangleq \mathcal{N}(\mu_{t|T}, \Sigma_{t|T}) \\ \mu_{t|T} &\triangleq \mu_{t|t} + J_t (\mu_{t+1|T} - \mu_{t+1|t}) \\ \Sigma_{t|T} &\triangleq \Sigma_{t|t} + J_t (\Sigma_{t+1|T} - \Sigma_{t+1|t}) J_t^T \\ J_t &\triangleq \Sigma_{t|t} a_{z_t|t-1}^T \Sigma_{t+1|t}^{-1} \end{aligned}$$

## 7.2 M-step

In the M-step, we want to calculate the the expected ELBO over the calculated marginals, as in equation 7.6 and then maximize it with respect to the parameters. Similar calculations can be found in [3].

$$\arg \max_{\theta, \phi} L(q, \theta, \phi) = \arg \max_{\theta, \phi} \left[ \mathbb{E}_{q(z_0)} \log P(z_0) + \sum_{t=0}^{T-2} \mathbb{E}_{q(z_{t+1}, z_t)} \log P(z_{t+1} | z_t) \right. \\ \left. + \sum_{t=1}^{T-1} \mathbb{E}_{q(z_t)} \log P(x_t | z_{t-1}) + \log P(\theta, \phi) \right]$$

Calculating each summand and ignoring constants

$$\begin{aligned} \mathbb{E}_{q(z_0)} \log P(z_0) &\propto -\frac{1}{2} \log |\Omega_{z_0}| - \mathbb{E}_{q(z_0)} \frac{1}{2} z_0^T \Omega_{z_0}^{-1} z_0 \\ &= -\frac{1}{2} \log |\Omega_{z_0}| - \mathbb{E}_{q(z_0)} \frac{1}{2} \text{Tr}(\Omega_{z_0}^{-1} z_0 z_0^T) \\ &= -\frac{1}{2} \log |\Omega_{z_0}| - \frac{1}{2} \text{Tr}(\Omega_{z_0}^{-1} \mathbb{E}_{q(z_0)}(z_0 z_0^T)) \\ &\quad \text{by linearity of trace} \\ &= -\frac{1}{2} \log |\Omega_{z_0}| - \frac{1}{2} \text{Tr} \left( \Omega_{z_0}^{-1} \left( \Sigma_{0|T} + \mu_{0|T} \mu_{0|T}^T \right) \right) \end{aligned}$$

The state transition terms are computed as follows:

$$\begin{aligned} \sum_{t=0}^{T-1} \mathbb{E}_{q(z_{t+1}, z_t)} \log P(z_{t+1} | z_t) &\propto -\frac{1}{2} \sum_{t=0}^{T-1} \left[ \log |\Omega_{z_{t+1}|t}| + \text{Tr} \left( \Omega_{z_{t+1}|t}^{-1} \mathbb{E}_{q(z_{t+1}, z_t)} \text{Quad} \right) \right] \\ &= -\frac{1}{2} \sum_{t=0}^{T-1} \left[ \log |\Omega_{z_{t+1}|t}| + \text{Tr} \left( \Omega_{z_{t+1}|t}^{-1} \left( \Sigma_{t+1|T} - a_{z_{t+1}|t} \Sigma_{t|T} a_{z_{t+1}|t}^T \right. \right. \right. \\ &\quad \left. \left. \left. + (\mu_{t+1} - a_{z_{t+1}|t} \mu_t - b_{z_{t+1}|t})(\mu_{t+1} - a_{z_{t+1}|t} \mu_t - b_{z_{t+1}|t})^T \right) \right) \right] \end{aligned} \quad (7.14)$$

$$\begin{aligned} &= -\frac{1}{2} \sum_{t=0}^{T-1} \left[ \log |\Omega_{z_{t+1}|t}| + \text{Tr} \left( \Omega_{z_{t+1}|t}^{-1} \left( P_{t+1} \right. \right. \right. \\ &\quad - a_{z_{t+1}|t} P_{t+1,t}^T - P_{t+1,t} a_{z_{t+1}|t}^T + a_{z_{t+1}|t} P_t a_{z_{t+1}|t}^T \\ &\quad - \mu_{t+1|T} b_{z_{t+1}|t}^T - b_{z_{t+1}|t} \mu_{t+1|T}^T + a_{z_{t+1}|t} \mu_{t|T} b_{z_{t+1}|t}^T + b_{z_{t+1}|t} \mu_{t|T}^T a_{z_{t+1}|t}^T \\ &\quad \left. \left. \left. + b_{z_{t+1}|t} b_{z_{t+1}|t}^T \right) \right) \right] \end{aligned} \quad (7.15)$$

$$P_{t+1} \triangleq \Sigma_{t+1|T} + \mu_{t+1|T} \mu_{t+1|T}^T$$

$$P_t \triangleq \Sigma_{t|T} + \mu_{t|T} \mu_{t|T}^T$$

$$P_{t+1,t} \triangleq (a_{z_{t+1}|t} \Sigma_{t|T} + \mu_{t+1|T} \mu_{t|T}^T)$$

$$P_{t,t+1} \triangleq \Sigma_{t|T} a_{z_{t+1}|t}^T + \mu_{t|T} \mu_{t+1|T}^T = P_{t+1,t}^T$$

The last equation (7.15) follows the equations from [3]. Empirically, the second equation (7.14), in terms of the quadratic form of the expected smoothed marginals seems to give more numerically stable results.

I derive the transitions. here Note that the joint expectations marginalizes out except the covariances, which we know from the joint distribution. Explicitly, the quadratic form is in the transition probability is:

$$\begin{aligned}
Quad &= (z_{t+1} - a_{z_{t+1}|t} z_t - b_{z_{t+1}|t})(z_{t+1} - a_{z_{t+1}|t} z_t - b_{z_{t+1}|t})^T \\
&= z_{t+1} z_{t+1}^T + a_{z_{t+1}|t} z_t z_t^T a_{z_{t+1}|t}^T + b_{z_{t+1}|t} b_{z_{t+1}|t}^T \\
&\quad - z_{t+1} z_t^T a_{z_{t+1}|t}^T - z_{t+1} b_{z_{t+1}|t}^T - a_{z_{t+1}|t} z_t z_{t+1}^T - b_{z_{t+1}|t} z_{t+1}^T \\
&\quad + a_{z_{t+1}|t} z_t b_{z_{t+1}|t}^T + b_{z_{t+1}|t} z_t^T a_{z_{t+1}|t}^T \\
\mathbb{E}_{q(z_{t+1}, z_t)} Quad &= \mathbb{E}_{q(z_{t+1}, z_t)} (z_{t+1} - a_{z_{t+1}|t} z_t - b_{z_{t+1}|t})(z_{t+1} - a_{z_{t+1}|t} z_t - b_{z_{t+1}|t})^T \\
&= \mathbb{E}_{q(z_{t+1})} z_{t+1} z_{t+1}^T + \mathbb{E}_{q(z_t)} a_{z_{t+1}|t} z_t z_t^T a_{z_{t+1}|t}^T + b_{z_{t+1}|t} b_{z_{t+1}|t}^T \\
&\quad - \mathbb{E}_{q(z_{t+1}, z_t)} z_{t+1} z_t^T a_{z_{t+1}|t}^T - \mathbb{E}_{q(z_{t+1})} z_{t+1} b_{z_{t+1}|t}^T - \mathbb{E}_{q(z_{t+1}, z_t)} a_{z_{t+1}|t} z_t z_{t+1}^T - b_{z_{t+1}|t} z_{t+1}^T \\
&\quad + \mathbb{E}_{q(z_t)} a_{z_{t+1}|t} z_t b_{z_{t+1}|t}^T + \mathbb{E}_{q(z_t)} b_{z_{t+1}|t} z_t^T a_{z_{t+1}|t}^T \\
&= (\Sigma_{t+1|T} + \mu_{t+1|T} \mu_{t+1|T}^T) + a_{z_{t+1}|t} (\Sigma_{t|T} + \mu_{t|T} \mu_{t|T}^T) a_{z_{t+1}|t}^T \\
&\quad - (a_{z_{t+1}|t} \Sigma_{t|T} + \mu_{t+1|T} \mu_{t|T}^T) a_{z_{t+1}|t}^T - a_{z_{t+1}|t} (\Sigma_{t|T} a_{z_{t+1}|t}^T + \mu_{t|T} \mu_{t+1|T}^T) - \mu_{t+1|T} b_{z_{t+1}|t}^T \\
&\quad - b_{z_{t+1}|t} \mu_{t+1|T}^T + a_{z_{t+1}|t} \mu_{t|T} b_{z_{t+1}|t}^T + b_{z_{t+1}|t} \mu_{t|T}^T a_{z_{t+1}|t}^T + b_{z_{t+1}|t} b_{z_{t+1}|t}^T \\
&= (\mu_{t+1} - a_{z_{t+1}|t} \mu_t - b_{z_{t+1}|t})(\mu_{t+1} - a_{z_{t+1}|t} \mu_t - b_{z_{t+1}|t})^T + \Sigma_{t+1|T} - a_{z_{t+1}|t} \Sigma_{t|T} a_{z_{t+1}|t}^T \\
&= P_{t+1} + a_{z_{t+1}|t} P_t a_{z_{t+1}|t}^T - a_{z_{t+1}|t} P_{t,t+1} - P_{t,t+1}^T a_{z_{t+1}|t}^T \\
&\quad - \mu_{t+1|T} b_{z_{t+1}|t}^T - b_{z_{t+1}|t} \mu_{t+1|T}^T + a_{z_{t+1}|t} \mu_{t|T} b_{z_{t+1}|t}^T + b_{z_{t+1}|t} \mu_{t|T}^T a_{z_{t+1}|t}^T \\
&\quad + b_{z_{t+1}|t} b_{z_{t+1}|t}^T
\end{aligned}$$

Lastly, the likelihoods terms are summed as follows:

$$\begin{aligned}
\sum_{t=0}^T \mathbb{E}_{q(z_t)} \log P(x_t | z_t) &\propto -\frac{1}{2} \sum_{t=0}^T \left[ \log |\Omega_\varepsilon| + \text{Tr} \left( \Omega_\varepsilon^{-1} \mathbb{E}_{q(z_t)} (x_t - Ax_{t-1} + BL(t, \theta, \phi) z_t)(x_t - Ax_{t-1} + BL(t, \theta, \phi) z_t)^T \right) \right] \\
&= -\frac{1}{2} \sum_{t=0}^T \left[ \log |\Omega_\varepsilon| + \text{Tr} \left( \Omega_\varepsilon^{-1} (x_t - Ax_{t-1} + BL(t, \theta, \phi) \mu_{t|T})(x_t - Ax_{t-1} + BL(t, \theta, \phi) \mu_{t|T})^T \right. \right. \\
&\quad \left. \left. + BL(t, \theta, \phi) \Sigma_{t|T} (BL(t, \theta, \phi))^T \right) \right]
\end{aligned}$$

Assuming that  $K(t, \theta)$  and  $L(t, \theta, \phi)$  are differentiable functions of the parameters  $\{\theta, \phi\}$ , we can then use gradient ascent on this objective function to calculate the optimal parameters.

## 8 Appendix: Problems with the algorithm

If any of the transitions are deterministic (for instance acceleration, to velocity to position) or have very small noise, it can cause problems in the loglikelihood estimation. Small variances make the condition

number of matrix inversions very large, causing numerical instabilities. Deterministic dynamics also have undefined probability density. These probably need to be modeled separately. I add small noise to all dynamics to bypass the deterministic nature., but there may still be numerical instabilities.

Numerical stability.

note to myself: Numerical stability: Thrun et al. 2006; Simon 2006. take care of 0 probabilities and other small probabilities properly (some of these matrices are sparse)

## References

- [1] M.I. Jordan *Graphical models*. 2007
- [2] K.P. Murphy *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012
- [3] R.H. Shumway, D.S. Stoeffer *AN APPROACH TO TIME SERIES SMOOTHING AND FORECASTING USING THE EM ALGORITHM*. Journal of Time Series Analysis, 3 (4): 253-264, 1982.
- [4] Daniel M Wolpert, Zoubin Ghahramani. *Computational principles of movement neuroscience*. Nature Neuroscience 3 (11), 1212-1217,. 2000.
- [5] E. Todorov. *Optimal Feedback Control as a theory of motor coordination*. Nature Neuroscience, 5 (11). 2002
- [6] E. Todorov. *Stochastic Optimal Control and Estimation Methods Adapted to the Noise Characteristics of the Sensorimotor System*. Neural Computation 17. 2005
- [7] C. Firestone and B.J. Scholl. *Cognition does not affect perception: Evaluating the evidence for top-down effects*. Brain and Behavioral Sciences. 2016.