Inference in Linear Dynamical Systems

Joaquin Rapela[*]

June 30, 2022

1 Linear dynamical systems (LDS) model

$$\begin{array}{llll}
\mathbf{x}_n = A\mathbf{x}_{n-1} + \mathbf{w}_n & \mathbf{w}_n \sim N(\mathbf{w}_n|\mathbf{0}, Q) & \mathbf{x}_n \in \mathbb{R}^m & \\
\mathbf{y}_n = C\mathbf{x}_n + \mathbf{v}_n & \mathbf{v}_n \sim N(\mathbf{v}_n|\mathbf{0}, R) & \mathbf{y}_n \in \mathbb{R}^n & n = 1 \dots N \\
\mathbf{x}_0 \sim N(\mathbf{w}_n|\mathbf{m}_0, V_0) & & &
\end{array}$$

[Figure 1 about here.]

2 Inference Problems

Prediction

$$P(\mathbf{x}_n|\mathbf{y}_1, \dots, \mathbf{y}_{n-1}) = N(\mathbf{x}_n|\mathbf{x}_{n|n-1}, P_{n|n-1}) \tag{1}$$

Filtering

$$P(\mathbf{x}_n|\mathbf{y}_1, \dots, \mathbf{y}_n) = N(\mathbf{x}_n|\mathbf{x}_{n|n}, P_{n|n}) \tag{2}$$

Smoothing

$$P(\mathbf{x}_n|\mathbf{y}_1, \dots, \mathbf{y}_N) = N(\mathbf{x}_n|\mathbf{x}_{n|N}, P_{n|N}) \tag{3}$$

3 Kalman Filter

The Kalman filter algorithm addresses the prediction (Eq. 1) and filtering (Eq. 2) inference problems. It is an iterative algorithm, which alternates between computing the mean and covariance of the prediction distribution and computing the mean and covariance of the filtering distribution.

[*] j.rapela@ucl.ac.uk

$$\begin{aligned}
\mathbf{x}_{0|0} &= \mathbf{m}_0 && \text{init filtered mean}\\
P_{0|0} &= V_0 && \text{init filtered covariance}\\
\mathbf{x}_{n+1|n} &= A\mathbf{x}_{n|n} && \text{prediction mean}\\
P_{n+1|n} &= AP_{n|n}A^{\intercal} + Q && \text{prediction covariance}\\
\mathbf{y}_{n|n-1} &= C\mathbf{x}_{n|n-1} && \text{predicted observation}\\
\tilde{\mathbf{y}}_n &= \mathbf{y}_n - \mathbf{y}_{n|n-1} && \text{residual}\\
S_n &= CP_{n|n-1}C^{\intercal} + R && \text{residual covariance}\\
\mathbf{x}_{n|n} &= \mathbf{x}_{n|n-1} + K_n\tilde{\mathbf{y}}_n && \text{filtering mean}\\
K_n &= P_{n|n-1}C^{\intercal}S_n^{-1} && \text{Kalman gain}\\
P_{n|n} &= (I_M - K_nC)P_{n|n-1} && \text{filtering covariance}
\end{aligned}$$

Inference of predicted and filtered means and covariances proceeds in a forward fashion, inferring the prediction and filtering distributions from the first to the last state, as shown in the next figure.

[Figure 2 about here.]

4 Kalman Smoother

The Kalman smoother algorithm addresses the smoothing (Eq. 3) inference problem.

$$\begin{aligned}
x_{n|N} &= x_{n|n} + C_n(x_{n+1|N} - x_{n+1|n}) && \text{smoothed mean}\\
P_{n|N} &= P_{n|n} + C_n(P_{n+1|N} - P_{n+1|n})C_n^{\intercal} && \text{smoothed covariance}\\
C_n &= P_{n|n}A^{\intercal}P_{n+1|n}^{-1}
\end{aligned}$$

Inference of smoothed mean and covariances proceeds in a backward fashion: $x_{N|N}, P_{N|N} \rightarrow x_{N-1|N}, P_{N-1|N} \rightarrow x_{N-2|N}, P_{N-2|N} \rightarrow \ldots \rightarrow x_{1|N}, P_{1|N}$. The initial mean and covariances (i.e., $x_{N|N}, P_{N|N}$) are initialized from the last step of the Kalman filter.

5 Derivation of Kalman filter equations

Claim 1. $\mathbf{x}_{n|n-1} = A\mathbf{x}_{n-1|n-1}$

Proof.

$$\begin{aligned}
\mathbf{x}_{n|n-1} &= E\{\mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\}\\
&= E\{A\mathbf{x}_{n-1} + \mathbf{w}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} \tag{4}\\
&= AE\{\mathbf{x}_{n-1}|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} + E\{\mathbf{w}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} \tag{5}\\
&= A\mathbf{x}_{n-1|n-1} + E\{\mathbf{w}_n\} \tag{6}\\
&= A\mathbf{x}_{n-1|n-1} \tag{7}
\end{aligned}$$

□

Notes:

- Eq. 4 arises from the state equation of the LDS model in Section 1,
- Eq. 5 holds because the expectation distributes over sums
- Eq. 6 uses the definition of $\mathbf{x}_{n-1|n-1}$ and the fact that the state noise, $\mathbf{w}_n$, is independent of previous observations.
- Eq. 7 follows due to the zero mean of $\mathbf{w}_n$.

Claim 2. $P_{n|n-1} = AP_{n-1|n-1}A^\intercal + Q$

Proof.

$$\begin{aligned}
P_{n|n-1} =&\text{Cov}\{\mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} \\
=&\text{Cov}\{A\mathbf{x}_{n-1}+\mathbf{w}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} && (8)\\
=&\text{Cov}\{A\mathbf{x}_{n-1}|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\}+\text{Cov}\{\mathbf{w}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} && (9)\\
=&A\ \text{Cov}\{\mathbf{x}_{n-1}|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\}\ A^\intercal+\text{Cov}\{\mathbf{w}_n\} && (10)\\
=&A\ P_{n-1|n-1}\ A^\intercal+Q && (11)
\end{aligned}$$

□

Notes:

1. Eq. 8 used the state equation of the LDS model in Section 1,
2. Eq. 9 is true because $\mathbf{w}_n$ is independent from $\mathbf{x}_{n-1}$,
3. Eq. 10 holds by the property $\text{Cov}\{A\mathbf{x}\} = A\ \text{Cov}\{\mathbf{x}\}'A^\intercal$ and because $\mathbf{w}_n$ is independent of previous observations.
4. Eq. 10 applied the definitions of $P_{n-1|n-1}$ and Q.

Claim 3. $\mathbf{x}_{n|n} = \mathbf{x}_{n|n-1} + K_n\tilde{\mathbf{y}}_n$ *and* $P_{n|n} = (I - K_nC)P_{n|n-1}$.

Proof. Define the random variables $\mathbf{x} = \mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}$ and $\mathbf{y} = \mathbf{y}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}$. Then $\mathbf{x}|\mathbf{y} = \mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_n$ and the mean and covariance that we want to find, $\mathbf{x}_{n|n}$ and $P_{n|n}$, are those of $\mathbf{x}|\mathbf{y}$. Thus, we want to compute the mean, $\mu_{\mathbf{x}|\mathbf{y}} = \mathbf{x}_{n|n}$, and covariance, $\Sigma_{\mathbf{x}|\mathbf{y}} = P_{n|n}$, of $\mathbf{x}|\mathbf{y}$.

Because $\mathbf{x}_n$ and $\mathbf{y}_n$ are jointly Gaussian, then $\mathbf{x}$ and $\mathbf{y}$ are also jointly Gaussian. Then, $\mu_{\mathbf{x}|\mathbf{y}}$ and $\Sigma_{\mathbf{x}|\mathbf{y}}$ are (Bishop, 2016, Chapter 2)

$$\mu_{\mathbf{x}|\mathbf{y}} = \mu_{\mathbf{x}} + \Sigma_{\mathbf{xy}}\Sigma_{\mathbf{yy}}^{-1}(\mathbf{y}_n - \mu_{\mathbf{y}}) \tag{12}$$

$$\Sigma_{\mathbf{x}|\mathbf{y}} = \Sigma_{\mathbf{xx}} - \Sigma_{\mathbf{xy}}\Sigma_{\mathbf{yy}}^{-1}\Sigma_{\mathbf{yx}} \tag{13}$$

Thus, to compute $\mu_{\mathbf{x}|\mathbf{y}}$ and $\Sigma_{\mathbf{x}|\mathbf{y}}$ we need to calculate $\mu_{\mathbf{x}}$, $\mu_{\mathbf{y}}$, $\Sigma_{\mathbf{xx}}$, $\Sigma_{\mathbf{xy}}$ and $\Sigma_{\mathbf{yy}}$.

$$\mu_{\mathbf{x}} = E\{\mathbf{x}\} = E\{\mathbf{x}_n|\mathbf{y}_1,\ldots\mathbf{y}_{n-1}\} = \mathbf{x}_{n|n-1} \tag{14}$$

$$\begin{aligned}
\mu_{\mathbf{y}} = E\{\mathbf{y}\} = E\{\mathbf{y}_n|\mathbf{y}_1,\ldots\mathbf{y}_{n-1}\} = E\{C\mathbf{x}_n + \mathbf{v}_n|\mathbf{y}_1,\ldots\mathbf{y}_{n-1}\} = \\
= CE\{\mathbf{x}_n|\mathbf{y}_1,\ldots\mathbf{y}_{n-1}\} + E\{\mathbf{v}_n|\mathbf{y}_1,\ldots\mathbf{y}_{n-1}\} = C\mathbf{x}_{n|n-1} + E\{\mathbf{v}_n\} = C\mathbf{x}_{n|n-1} = \mathbf{y}_{n|n-1} && (15)
\end{aligned}$$

Notes:

- The penultimate equality in Eq. 15 uses the definition of $\mathbf{x}_{n|n-1}$ and the fact that $\mathbf{v}_n$ is independent of previous observations.
- The last equality in Eq. 15 holds because the mean of $\mathbf{v}_n$ is zero. Claim 1.

$$\begin{aligned}
\Sigma_{\mathbf{yy}} &= \mathrm{Cov}(\mathbf{y}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}) \\
&= \mathrm{Cov}(C\mathbf{x}_n+\mathbf{v}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}) \\
&= \mathrm{Cov}(C\mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1})+\mathrm{Cov}(\mathbf{v}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}) \\
&= C\ \mathrm{Cov}(\mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1})\ C^\intercal+\mathrm{Cov}(\mathbf{v}_n) \qquad (16)\\
&= CP_{n|n-1}C^\intercal+R \qquad (17)
\end{aligned}$$

Notes:

- As in Eq. 9, Eq. 16 holds by the property $\mathrm{Cov}\{A\mathbf{x}\}=A\ \mathrm{Cov}\{\mathbf{x}\}\ A^\intercal$ and because $\mathbf{v}_n$ is independent of previous observations.

$$\begin{aligned}
\Sigma_{\mathbf{xy}} &= \mathrm{cCov}(\mathbf{x}_n,\mathbf{y}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}) \\
&= \mathrm{cCov}(\mathbf{x}_n,C\mathbf{x}_n+\mathbf{v}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}) \\
&= \mathrm{cCov}(\mathbf{x}_n,C\mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1})+\mathrm{cCov}(\mathbf{x}_n,\mathbf{v}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}) \qquad (18)\\
&= \mathrm{Cov}(\mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1})C^\intercal+0 \qquad (19)\\
&= P_{n|n-1}C^\intercal \qquad (20)
\end{aligned}$$

Notes:

- the first term in Eq. 19 arises from the first term of Eq. 18 since

$$\begin{aligned}
\mathrm{cCov}(\mathbf{x}_n,C\mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}) &= E\{(\mathbf{x}_n-\mu_{\mathbf{x}})(C\mathbf{x}_n-C\mu_{\mathbf{x}})^\intercal|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} \\
&= E\{(\mathbf{x}_n-\mu_{\mathbf{x}})(\mathbf{x}_n-\mu_{\mathbf{x}})^\intercal C^\intercal|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} \\
&= E\{(\mathbf{x}_n-\mu_{\mathbf{x}})(\mathbf{x}_n-\mu_{\mathbf{x}})^\intercal|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\}C^\intercal \\
&= \mathrm{Cov}(\mathbf{x}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1})C^\intercal
\end{aligned}$$

- the second term in Eq. 19 arises from the second term of Eq. 18 since

$$\begin{aligned}
\mathrm{cCov}(\mathbf{x}_n,\mathbf{v}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}) &= E\{(\mathbf{x}_n-\mathbf{x}_{n|n-1})\mathbf{v}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} \\
&= E\{(\mathbf{x}_n-\mathbf{x}_{n|n-1})|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\}E\{\mathbf{v}_n|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\} \\
&= E\{(\mathbf{x}_n-\mathbf{x}_{n|n-1})|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\}E\{\mathbf{v}_n\} \\
&= E\{(\mathbf{x}_n-\mathbf{x}_{n|n-1})|\mathbf{y}_1,\ldots,\mathbf{y}_{n-1}\}0 \\
&= 0
\end{aligned}$$

  the second line follows from the first one because $\mathbf{v}_n$ is independent of $\mathbf{x}_n$, and the third line follows from the second one because $\mathbf{v}_n$ is independent of previous observations.

$$\Sigma_{\mathbf{xx}} = \mathrm{Cov}(\mathbf{x}_n|\mathbf{y}_1, \dots, \mathbf{y}_{n-1}) = P_{n|n-1} \tag{21}$$

Having calculated $\mu_{\mathbf{x}}$, $\mu_{\mathbf{y}}$, $\Sigma_{\mathbf{xx}}$, $\Sigma_{\mathbf{xy}}$ and $\Sigma_{\mathbf{yy}}$ we now use Eqs. 12, 13, 14, 15, 17, 20, and 21 to obtain $\mathbf{x}_{n|n}$ and $P_{n|n}$.

$$\begin{aligned}
\mathbf{x}_{n|n} = \mu_{\mathbf{x}|\mathbf{y}} &= \mu_{\mathbf{x}} + \Sigma_{\mathbf{xy}}\Sigma_{\mathbf{yy}}^{-1}(\mathbf{y}_n - \mu_{\mathbf{y}}) \\
&= \mathbf{x}_{n|n-1} + P_{n|n-1}C^\intercal S_n^{-1}(\mathbf{y}_n - \mathbf{y}_{n|n-1}) \\
&= \mathbf{x}_{n|n-1} + K_n\tilde{\mathbf{y}}_n \\
P_{n|n} = \Sigma_{\mathbf{x}|\mathbf{y}} &= \Sigma_{\mathbf{xx}} - \Sigma_{\mathbf{xy}}\Sigma_{\mathbf{yy}}^{-1}\Sigma_{\mathbf{yx}} \\
&= P_{n|n-1} - P_{n|n-1}C^\intercal S_n^{-1}CP_{n|n-1} \\
&= (I - P_{n|n-1}C^\intercal S_n^{-1}C)P_{n|n-1} \\
&= (I - K_nC)P_{n|n-1}
\end{aligned}$$

□

6 Joint normality of the states and observations in the LDS

Claim 4. *The state and observation random variables of an LDS* $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{y}_1, \dots, \mathbf{y}_N\}$ *are jointly normal.*

Proof. Note that a set of real random variables $\mathcal{Z} = \{z_1, \dots, z_N\}$ is jointly normal if and only if their joint probability distribution is a multivariate Normal distribution, if and only if the logarithm of this joint probability distribution is a quadratic function of the random variables in $\mathcal{Z}$ (i.e., $\ln P(z_1, \dots, z_N) = k + \sum_{i=1}^N k_1(i)z_i + \sum_{i=1}^N \sum_{j=1}^N k_2(i,j)z_iz_j$, with $k_1(i)$ and $k_2(i,j)$ real numbers).

Thus, to prove this claim it suffice to show that property P_n: "$\log P(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n, \mathbf{y}_1, \dots, \mathbf{y}_n)$ is a quadratic function of the components of $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n, \mathbf{y}_1, \dots, \mathbf{y}_n\}$" holds for any positive integer n. We show this by induction.

P_1 :

$$\begin{aligned}
\ln P(\mathbf{x}_0, \mathbf{x}_1, \mathbf{y}_1) =& \ln P(\mathbf{y}_1|\mathbf{x}_1) \ln P(\mathbf{x}_1|\mathbf{x}_0) \ln P(\mathbf{x}_0) \\
=& K - \frac{1}{2}(\mathbf{y}_1 - C\mathbf{x}_1)^\intercal R^{-1}(\mathbf{y}_1 - C\mathbf{x}_1) \\
& - \frac{1}{2}(\mathbf{x}_1 - A\mathbf{x}_0)^\intercal Q^{-1}(\mathbf{x}_1 - A\mathbf{x}_0) \\
& - \frac{1}{2}(\mathbf{x}_0 - \mathbf{m}_0)^\intercal Q^{-1}(\mathbf{x}_0 - \mathbf{m}_0)
\end{aligned} \tag{22}$$

P_1 follows from the observation that the components of $\mathbf{x}_0$, $\mathbf{x}_1$ and $\mathbf{y}_1$ are combined quadratically in Eq. 22.

$P_n \rightarrow P_{n+1}$:

$$\begin{aligned}
\ln P(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{x}_{n+1}, \mathbf{y}_1, \ldots, \mathbf{y}_n, \mathbf{y}_{n+1}) =& \ln P(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}) + \\
& \ln P(\mathbf{x}_{n+1}|\mathbf{x}_n) + \\
& \ln P(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}_1, \ldots, \mathbf{y}_n) \\
=& K - \frac{1}{2}(\mathbf{y}_{n+1} - C\mathbf{x}_{n+1})^\intercal R^{-1}(\mathbf{y}_{n+1} - C\mathbf{x}_{n+1}) \\
& - \frac{1}{2}(\mathbf{x}_{n+1} - A\mathbf{x}_n)^\intercal R^{-1}(\mathbf{x}_{n+1} - A\mathbf{x}_n) \\
& + \ln P(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}_1, \ldots, \mathbf{y}_n) \qquad (23)
\end{aligned}$$

P_{n+1} follows from the observation that the components of $\mathbf{x}_n$, $\mathbf{x}_{n+1}$ and $\mathbf{y}_{n+1}$ are combined quadratically in the first two lines of Eq. 23 and, by the inductive hypothesis P_n, the elements of $\mathbf{x}_0, \ldots, \mathbf{x}_n \mathbf{y}_1, \ldots \mathbf{y}_n$ are combined quadratically in the last line of Eq. 23.

□

7 Evaluation

7.1 Simulations

We compare the accuracy of the Kalman filter and smoother with that of the method of finite differences, to infer velocities and accelerations of a simulated object following the dynamics of the Discrete Wiener Process Acceleration (DWPA) model[1]. We used the following parameters in the simulations:

Name	*Value*
$\mathbf{x}_0$	$[0, 0]$
V_0	diag([0.001, 0.001])
$\gamma = \gamma_1 = \gamma_2$	1.0
$\sigma = \sigma_1 = \sigma_2$	varied

We simulated $N = 10,000$ samples with a step size $dt = 0.001$.

7.1.1 Lower noise ($\sigma = 1\text{e}{-10}$)

We simulated measurements from the two-dimensional DWPA model with a standard deviation for the noise of the observations set to $\sigma = 1\text{e}{-10}$. Figure 3 shows the state positions in blue and the noise-corrupted measurements in black.

[Figure 3 about here.]

We next inferred velocities and accelerations from the simulated measurements (Figure 4). For inference we used the Kalman filter, Kalman smoother and the finite difference method. For these low-noise simulations, all velocity and acceleration estimates were very accurate.

[Figure 4 about here.]

[1] https://github.com/joacorapela/lds_python/blob/master/docs/tracking/tracking.pdf

7.1.2 Medium noise ($\sigma = 1e{-3}$)

We simulated measurements from the two-dimensional DWPA model with a standard deviation for the noise of the observations set to $\sigma = 1e{-3}$. Figure 5 shows the state positions in blue and the noise-corrupted measurements in black.

[Figure 5 about here.]

We next inferred velocities and accelerations from the simulated measurements (Figure 6). For these medium-noise simulations, all velocity estimates were accurate. The Kalman filter and smoother estimates of acceleration were also accurate, but the finite difference estimates of acceleration were not.

[Figure 6 about here.]

7.1.3 Higher noise ($\sigma = 1e{-1}$)

We simulated measurements from the two-dimensional DWPA model with a standard deviation for the noise of the observations set to $\sigma = 1e{-1}$. Figure 7 shows the state positions in blue and the noise-corrupted measurements in black.

[Figure 7 about here.]

We next inferred velocities and accelerations from the simulated measurements (Figure 8). For these high-noise simulations, velocity and acceleration estimates by the Kalman filter and smoother were accurate, but those from the finite difference method were not.

[Figure 8 about here.]

7.1.4 Conclusions

It is remarkable that the finite difference method breaks down when adding very little noise to the true measurements. As we increased the amount of noise, this break down happened earlier for accelerations than for velocities. The Kalman filter and smoother were robust to the amount of noise considered here, both for the estimation of velocities and accelerations.

7.2 Foraging mouse

Figure 9 shows the measured and Kalman filtered and smoothed positions of a mouse foraging in a circular arena.

[Figure 9 about here.]

7.3 Mouse running on a maze

Figure 10 shows the measured and Kalman filtered and smoothed positions of a mouse running in a maze.

[Figure 10 about here.]

References

Bishop, C. M. (2016). *Pattern recognition and machine learning*. Springer-Verlag New York.

List of Figures

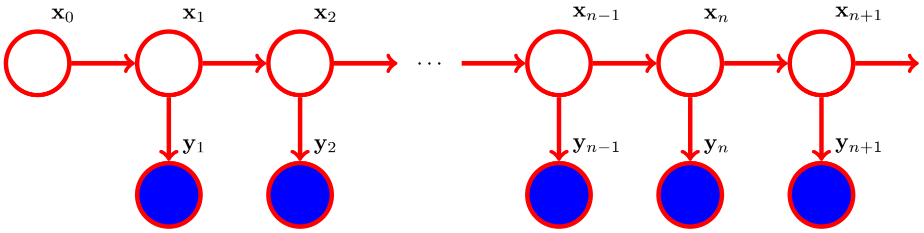


Figure 1: Graphical model for linear dynamical systems

$$\begin{tikzcd}
 & \text{Filtering Distribution} & \text{Prediction Distribution} \\
\mathbf{m}_0, V_0 \arrow[r] & \mathbf{m}_{0|0}, P_{0|0} \arrow[r] & \mathbf{m}_{1|0}, P_{1|0} \arrow[dl, "\mathbf{y}_1"'] \\
 & \mathbf{m}_{1|1}, P_{1|1} \arrow[r] & \mathbf{m}_{2|1}, P_{2|1} \\
 & \vdots & \vdots \\
 & \mathbf{m}_{n-1|n-1}, P_{n-1|n-1} \arrow[r] & \mathbf{m}_{n|n-1}, P_{n|n-1} \arrow[dl, "\mathbf{y}_n"'] \\
 & \mathbf{m}_{n|n}, P_{n|n} &
\end{tikzcd}$$

Figure 2: Order of calculation of the prediction and filtering distributions in the Kalman filtering algorithm.

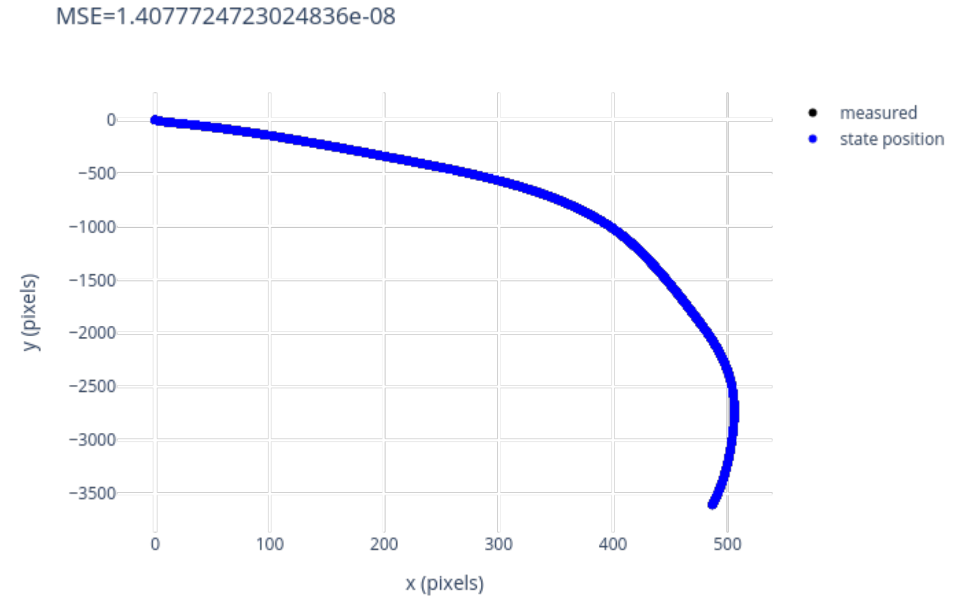
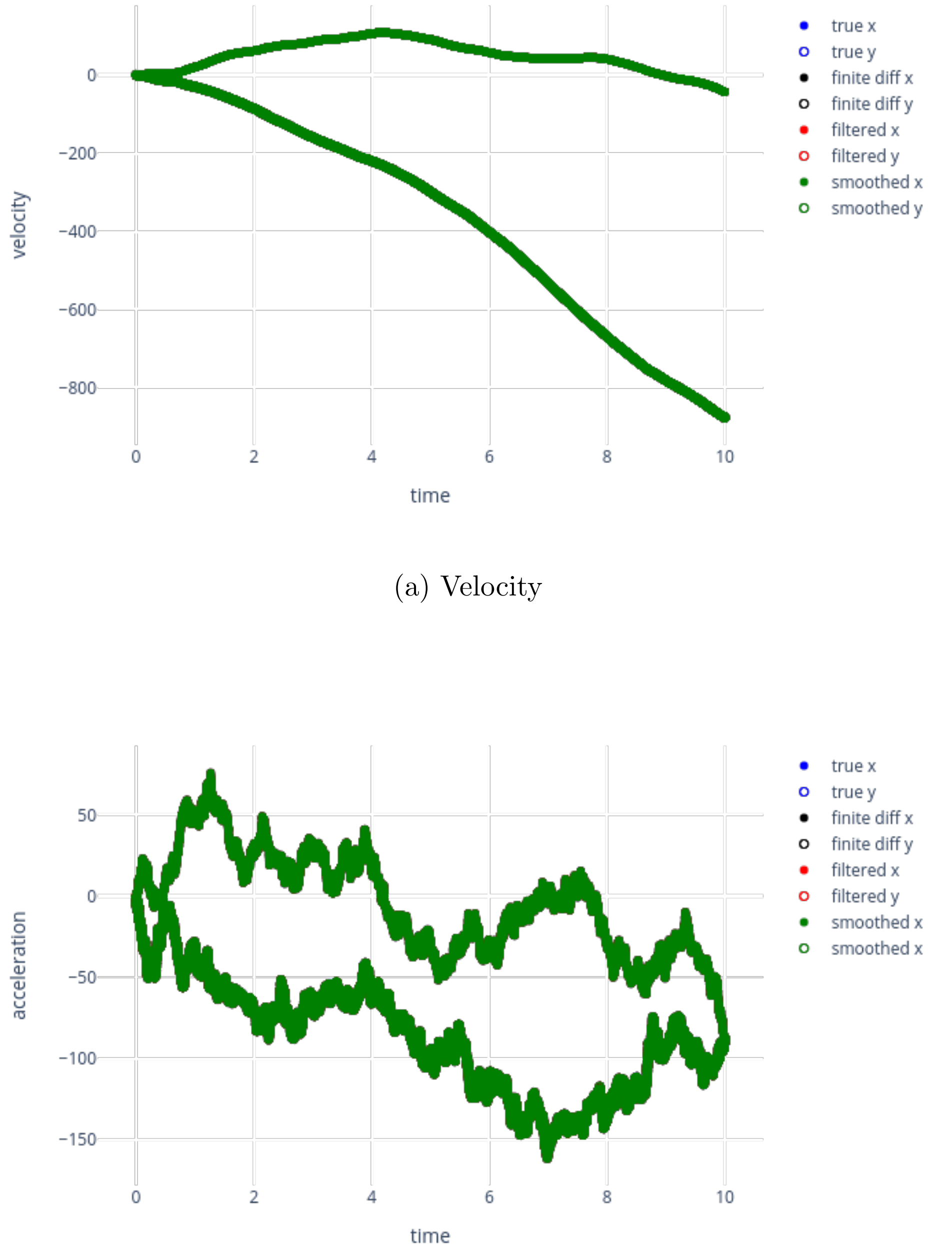


Figure 3: Noise corrupted measurements (black) and state positions (blue) simulated with low noise (standard deviation $\sigma = 1\mathrm{e}{-10}$) using a two-dimensional DWPA model. The noise was so low that the differences between measurements and state positions cannot be appreciated visually. The mean-squared error (MSE) between measurements and state positions is indcated in the title. Click on the image to view its interactive version.



(a) Velocity

(b) Acceleration

Figure 4: Estimated velocities (a) and accelerations (b) from low-noise simulations. Estimates were obtained using the finite differences method, the Kalman filter and the Kalman smoother. Velocity and Acceleration estimates using all methods were very accurate. Click on the image to view its interactive version.

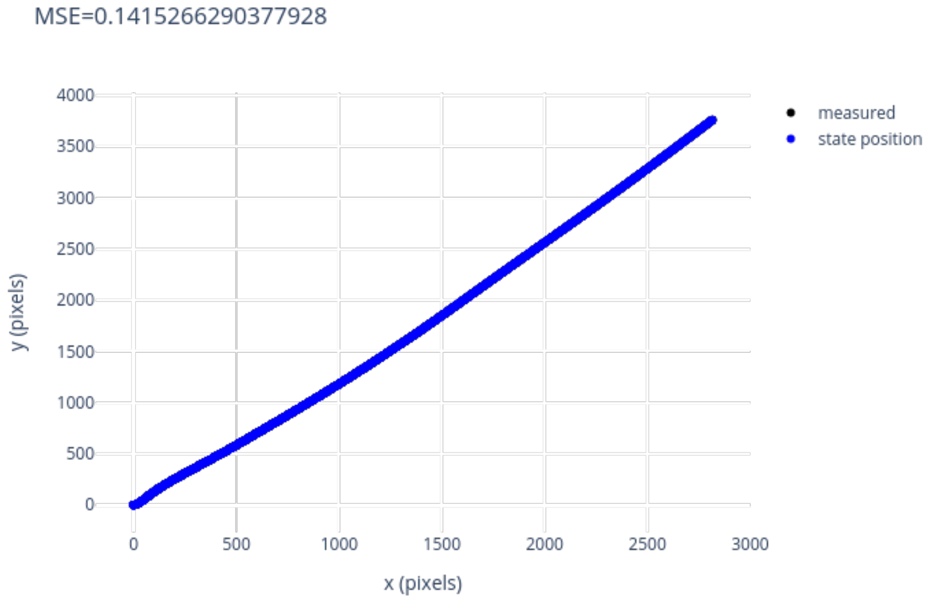
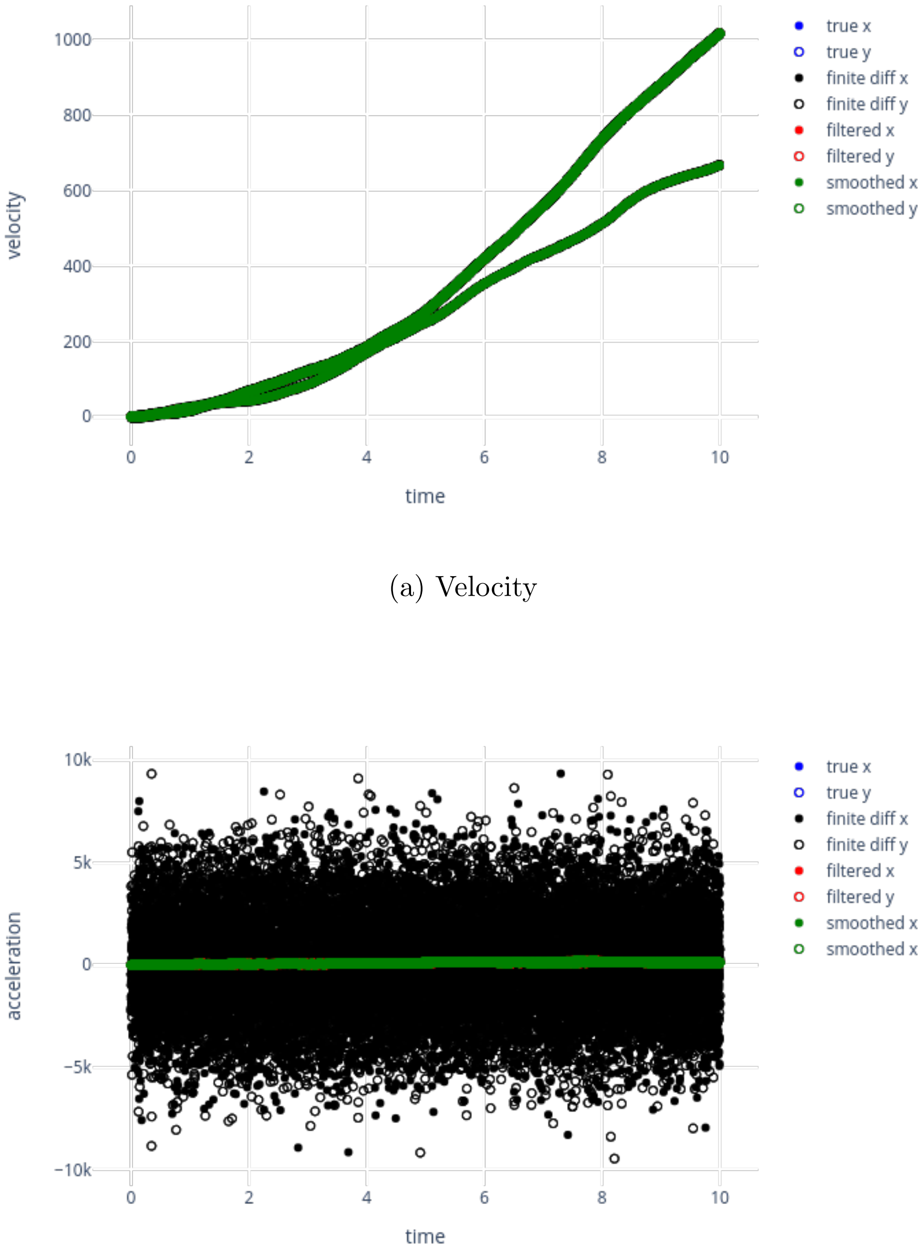


Figure 5: Noise corrupted measurements (black) and state positions (blue) simulated with medium noise (standard deviation $\sigma = 1e{-}3$) using a two-dimensional DWPA model. The noise was still so low that the differences between measurements and state positions cannot be appreciated visually. The MSE between measurements and state positions is indicated in the title. Click on the image to view its interactive version.



(a) Velocity

(b) Acceleration

Figure 6: Estimated velocities (a) and accelerations (b) from medium-noise simulations. Estimates were obtained using the finite differences method, the Kalman filter and Kalman smoother. Estimates of velocity by all methods were accurate. Estimates of accelerations by the Kalman filter and smoother were also accurate, but not those by the finite difference method. Click on the image to view its interactive version.

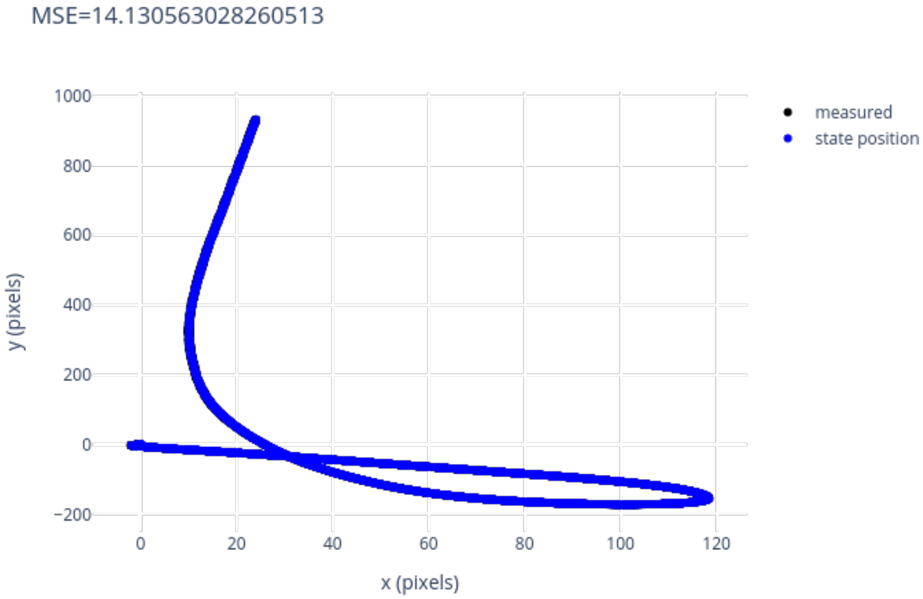
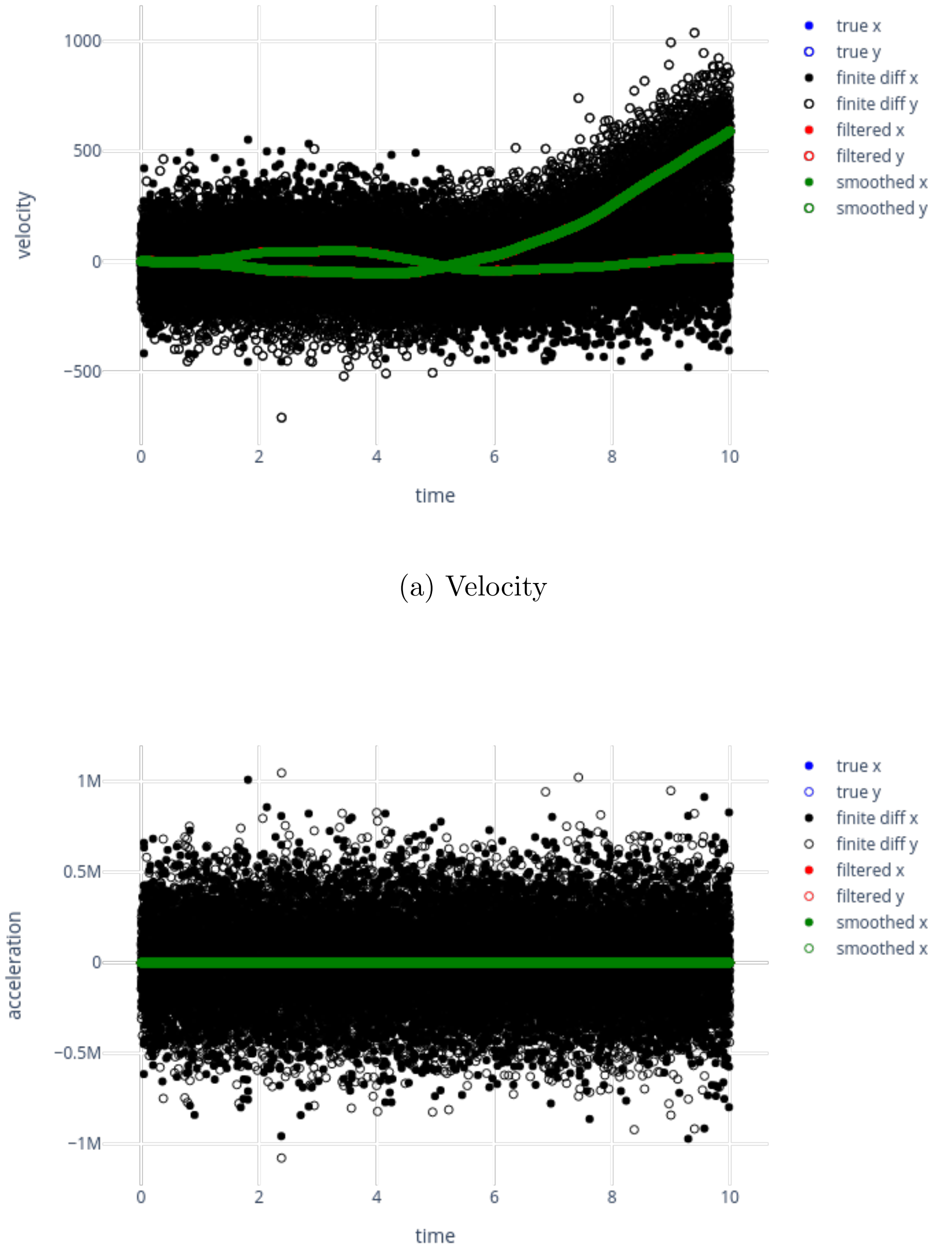


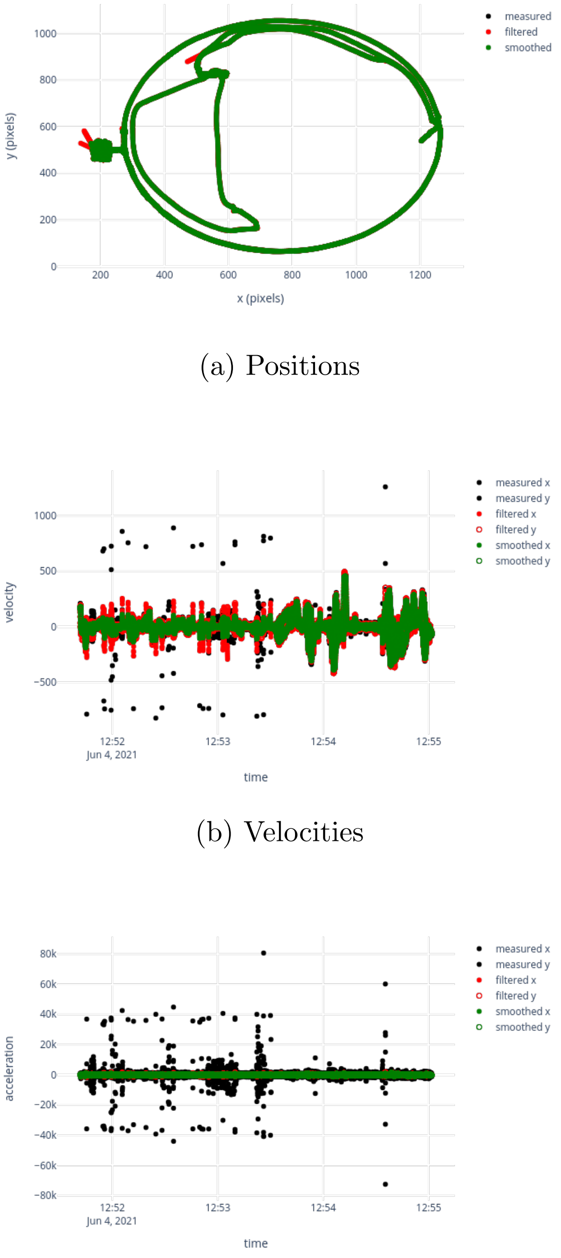
Figure 7: Noise corrupted measurements (black) and state positions (blue) simulated with high noise (standard deviation $\sigma = 1\text{e}{-1}$) using a two-dimensional DWPA model. The noise can now be appreciated visually. The MSE between measurements and state positions is indicated in the title. Click on the image to view its interactive version.



(a) Velocity

(b) Acceleration

Figure 8: Estimated velocities (a) and accelerations (b) from high-noise simulations. Estimates were obtained using the finite differences method, the Kalman filter and the Kalman smoother. Velocity and Acceleration estimates by the Kalman filter and smoother were accurate, but not those by the finite differences methods. Click on the image to view its interactive version.

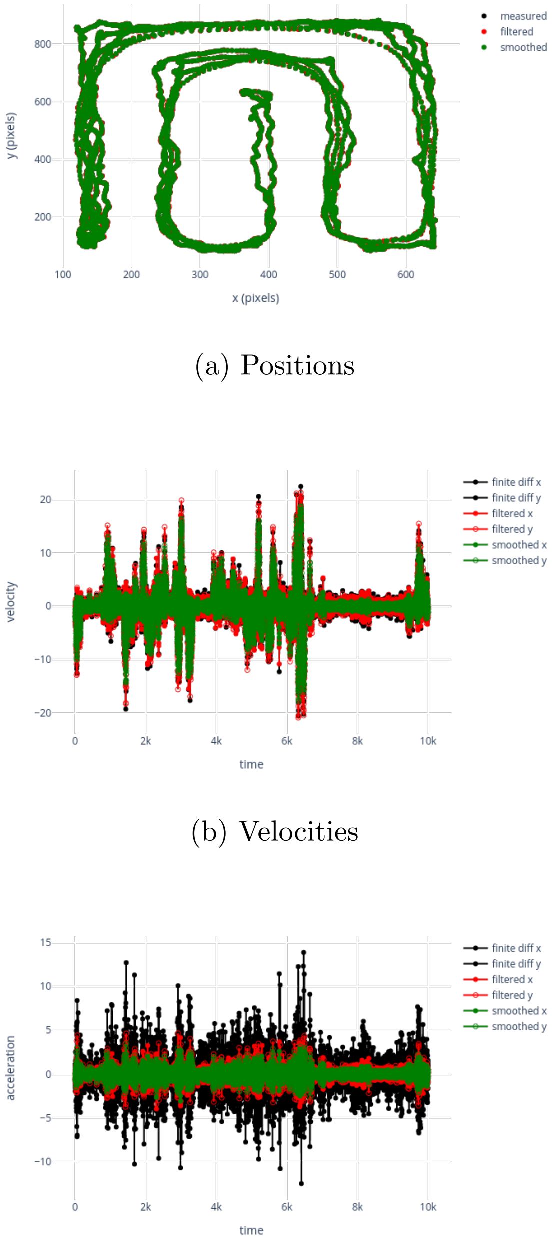


(a) Positions

(b) Velocities

(c) Accelerations

Figure 9: Kalman filtered and smoothed positions, velocities and accelerations of a mouse foraging in a circular arena. Click on the images to see their interactive versions.



(a) Positions

(b) Velocities

(c) Accelerations

Figure 10: Kalman filtered and smoothed positions, velocities and accelerations of a mouse running in a maze. Click on the images to see their interactive versions.