

# Amendments

## 1 Major Comments

- Chapter 1

1. In the introduction, the motivation of this thesis and the usage of real GPS data are clearly written. However, there is no separate description of the data. It would be easier to engage simulation studies if there is a subsection about the data.

Added separate descriptions in Section 1.2 on page 2.

2. Before Section 1.6, add a section outlining your research problems and significant contributions in the thesis work.

Stated the problem in Section 1.2 on page 2 and outlined the original contributions in Section 1.7 on page 15.

- Chapter 2.

1. Is this physical plausibility guaranteed by the objective function?

In the objective function, it is guaranteed that the V-spline is smooth. A mathematical explanation of “adjusted penalty term”, added in Section 2.2.3 on page 24, guarantees the reconstruction is physical plausible.

2. Theorem 1 describes the basic feature of the optimal V-spline but the “linear outside knots” feature of the optimal solution is mathematically undefined or ambiguous.

Define the meaning of “linear outside knots” in Theorem 1 on page 20.

3. There are some zig-zags on the reconstructed trajectories in Figures 2.14,2.15. Is this a true reflection of the reality or an artefact introduced by the reconstruction? Is it possible to visualise some close-ups of the reconstructed trajectories and contrast them with a “ground-truth”?

Unfortunately, we only have measurements. The true trajectories are unknown. These zig-zags in Figures 2.14,2.15 are reconstructions by V-splines. We believe that these zig-zags reflect a true behaviour of a tractor, because along this lane there are more measurements. On the contrary, along the lane with fewer measurements these zig-zags are gone.

4. There seems no comparison made on the two approaches for generating Figures 2.14,2.15.

In Section 2.5.2, the differences are elaborated.

- Chapter 3.

1. How good will it be? Can the same real-world dataset in Chapter 2 be used here for testing? Theoretically, V-spline and its Bayes estimates give the same reconstruction with the same data set, regardless simulation data or real-world data.

2. This is by far the most technical and difficult chapter to follow. Not all terms are adequately defined, for example I did not see a definition of a reproducing kernel.

In fact, all terms have been defined in this chapter. The conjecture has been resolved and added to Section 3.4.

- Chapter 4.

1. The simulation study in Section 4.4 is interesting but it seems entirely out of place. It is relevant to Chapter 5. In Chapter 4, we review state-space models and a number of filtering methods for combined state and parameter estimation that have been proposed in the literature. However, there are some disadvantages in these filtering methods, such as the sufficient statistics are not available in complex process. That is our motivation to propose the algorithm in Chapter 5.

2. Improperly cited references, equations and notations have been resolved.

- Chapter 5.

1. In page 92, (Section 5.2.4), does the equation 5.6 mean that a joint distribution of  $x_1, \dots, x_t, y_1, \dots, y_t$  is a multivariate normal distribution with zero mean and covariance matrix of  $\Sigma$ ?

Specifically, we suppose that the forward map and the observation model are linear and homogeneous, and the noise is Gaussian. It is mentioned in Section 5.2 on page 96.

2. In Equation 5.9,  $0.5 \ln |\Sigma^{-1}| = 0.5 \ln(\text{tr}(B)) - \ln(\text{tr}(L)) + \ln(\text{tr}(R))$  is only true when  $B$ ,  $L$  and  $R$  are diagonal matrices. Are they always diagonal matrices? Doesnt it rather depend on models?

In fact, matrix  $B$  is diagonal and  $L, R$  are lower triangular. The determinants of  $B, L$  and  $R$  are the product of their diagonal elements.

3. Section 5.3.2, it is unclear how accurate the MH-based parameter space estimation is.

In the two simulation studies in Section 5.4, it shows that the proposed MH-based approach has a good estimation on both parameter and state.

4. Reduce unnecessary repetitions. For instance, Equations 5.9, 5.29, 5.45, and 5.59(?) seem identical.

Repeated equations have been removed.

5. Alg 5.2, which is a sliding-window delayed acceptance method. Why havent the results of this chapter compared with the spline solution of Chapter 2?

It is discussed in Section 5.6 on page 136.

6. Have you used the real GPS data in Chapter 5? Then can you clearly indicate which result is for the real data?

The application to real data is demonstrated in Section 5.5.7 on page 131.

## 2 Minor Comments

- In Section 1, please check the references. The first example is that in Section 1.2, the first sentence should be Smoothing spline ..... reconstruction. See Eubank (2004) and Durbin and Koopman (2012) for details. There are many of these in Section 1.

The error has been resolved. All citation errors have been resolved.

- Page 5, Algorithm 1.1 : the notation  $t_{k+1}, t_{2k+1}, \dots$  is not distinguishable to  $t_2, \dots$ . Please use a different notation.

This issue has been fixed.

- Page 5, Second paragraph in Section 1.4 : Please rewrite For example a posterior estimation of  $x_t \sim p(x_t | y_{1:t})$  ...for incorrect estimates .....

The sentence has been rewritten.

- Page 7 : Describe  $R_t$  and  $Q_t$ .

Added description for  $R_t$  and  $Q_t$ .

- Page 9 : The first equation,  $E[p(y_t | y_{1:t1}, \theta)]$ ?

It has been fixed  $E[y_t | y_{1:t1}, \theta]$ .

- Page 26, Equation (2.35) : Isnt it  $1/(n1)$  instead of  $1/n$ ?

No, it is  $1/n$ . Equation (2.35) calculates the mean errors on the entire interval.

- P66:  $x_{k1}$  instead of  $x_{t1}$  in displayed equation.

The typo has been resolved.

- Sec 4.2.2 displayed equation has  $dx_x$ .

Should be  $dx_t$ .

- Section 4.2 makes reference to a target-tracking system which is undefined.  
It is defined in Section 4.2 on page 69.
- Final displayed equation before sec 4.2.3.  
It has been fixed.
- Page 69 onwards uses  $w(i)$  for the normalized importance weight whereas equation (4.4) uses tilde for the normalized weight.  
It uses  $w$  for weight and  $\tilde{w}$  for normalized weight.
- Alg 4.3 line 5 in  $\alpha$ : how is  $p(x_k, x_{k1} | y_{1:t})$  available to be used?  
According to equation 4.29, it is proportional to  $p(y_t | x_t)p(x_t | x_{t-1})p(x_{t-1} | y_{1:t-1})$
- Page 81, Algorithm 4.8 : “Draw  $\theta \sim p(\theta | y_{1:k})$ ” seems to duplicate to the previous sentence.  
It has been removed.
- Page 84 : It was observed that the LW filter has a larger distance to the true parameter, can you describe the reason?  
LW filter is trying to kill particle degeneracy. However, it is not doing well comparing with other filtering methods.
- Page 92, Equation for  $\Sigma$  : Change  $B$  to  $B_t$ .  
It has been fixed.
- Alg 4.7: in line 3 the observation  $y_{t+1}$  depends on the sufficient stat instead of  $x_{t+1}$ ? No particle indices present in line 4.  
Yes, it depends on sufficient statistics. Particle indices are added to the equation.
- Page 104,  $\Sigma^{-1}$  is  $(2t + 1) \times (2t + 1)$  and  $\Sigma$  expression is for  $2t \times 2t$ . Please change it for the correct dimension.  
It should be  $2t \times 2t$ .
- Page 120. Section 5.5.5 : What is Eff?  
Eff represents for “Efficiency”. An explanation is added to Section 5.5.5.
- Page 120, Second sentence in Section 5.5.5 : What is same dataset? Please write in details.  
Added an explanation: which is demonstrated in Figure 5.8.
- Page 129 : Figures 5.15-5.17 are not discussed in the section.  
Replaced by new figures and discussions.
- Appendix C : It is not clear how Appendix C is relevant to the main part. Please describe it in the main sections.  
It is relevant to “data cleaning” method, which is mentioned in Section 2.5 on page 42.