JEFFREYS DIVERGENCE BETWEEN STATE MODELS: APPLICATION TO TARGET TRACKING USING MULTIPLE MODELS

Clement Magnant [1, 2]

[1] THALES SYSTEMES AEROPORTES S.A.
clement.magnant@fr.thalesgroup.com
clement.magnant@ims-bordeaux.fr

Audrey Giremus [2] *and Eric Grivel* [2]

[2] University of Bordeaux - IPB
ENSEIRB-MATMECA - IMS - UMR CNRS 5218
{audrey.giremus, eric.grivel}@ims-bordeaux.fr

ABSTRACT

In the field of recursive estimation, the choice of the state model has a significant impact on the algorithm performance. Multiple Model (MM) approaches partly address this issue. However, the improvement over a single-model based estimator directly depends on the considered model set. It was theoretically shown that using either too many or too few models can degrade the estimation accuracy. In addition, the diversity of the selected models plays a crucial role. The contribution of this paper is twofold. 1/ We propose to use the Jeffreys divergence (JD) to measure the degree of mismatch between two models. We derive its recursive expression when the state vector is *a priori* modeled as a Markov chain. Then, we focus our attention on tracking applications and provide a detailed analysis of the JD between classical motion models. 2/ We investigate the impact of the similarity between the set of possible models on the estimation error of a MM algorithm. This survey can hence serve as a guideline for model set design.

Index Terms— Jeffreys Divergence, IMM, Kalman filtering, Model selection.

1. INTRODUCTION

In the field of signal processing, recursive estimation can be based on adaptive filters (LMS, etc.), or optimal filters such as H_∞ filter and Bayesian approaches using Kalman filter or particle filters. However, for these latter methods, *a priori* modeling the system plays a key role. Hence, the practitioner has to make assumptions regarding the way the state variables have to evolve [1]. It strongly impacts the estimator performance. To relax these assumptions on the *a priori* model, multiple model (MM) based methods have been considered and three generations have been proposed. In the first one, several estimators, each one based on a specific *a priori* model, are run in parallel, but no interaction between them exists. Then, interactive multiple model (IMM) was introduced [2]. It is known to be efficient for a small set of models. To address estimation problems requiring large model sets, variable structure IMM (VS-IMM) and its variants are considered [3]. In this third MM generation, the estimator number varies in time. The unlikely models are removed while those which could be considered are activated. Hence, subsets of models and "adjacency" between them must be defined by the practitioner.

For the last years, these approaches have been used in a wide range of applications: from VOIP speech enhancement to human motion tracking, from car motion tracking in intelligent transport systems and lane changes in highways to multiple maneuvering targets in radar processing, from GPS localization to channel estimation in mobile communication systems, etc.

However, few studies have been carried out on the relevance of MM algorithms regarding a single-estimator based algorithm. In [4], Kirubarajan and Bar-Shalom have proposed to compare the IMM and the Kalman filter in some situations. More generally, there are still some open topics: how to select the models? How many models should be considered? How to define the adjacent models in the 3rd generation of MM? Answering the above questions is all the more difficult as the decision may depend on the number of available samples.

In this paper, we focus our attention on linear Gaussian estimation problems. Furthermore, we investigate the application to target tracking. Our contribution is twofold:

1/ We propose a method to evaluate the similarity between two state models. For this purpose, we suggest computing the JD between the joint distributions of successive values of the state vector based on two different models. It should be noted that we derive its analytical recursive expression. Since the JD is the symmetric version of the Kullback-Leibler (KL) divergence, we first detail the computation of the latter.

2/ Then, the special case of motion models for target tracking is investigated. We study the influence of the motion model parameters on the JD. Then, we analyze the relevance of the IMM based estimator depending on the JD between the set of competing state models.

The remainder of this paper is organized as follows: in section 2, after recalling the linear state space representation of the system, we recursively compute the JD. In section 3, we focus our attention on object tracking. We give the state space representation for the uniformly accelerated motion (UAM) and the Singer motion model and present the JD we obtain in various cases. In section 4, we analyze the performance of the IMM estimator with regards to the JD between the competing models. We also compare the IMM with a single-model based Kalman filter, by looking at the root mean square error (RMSE) between the true state vector and its estimation. Conclusions and perspectives are finally drawn in section 5.

2. RECURSIVE COMPUTATION OF THE JD

Bayesian recursive estimation techniques are based on the so-called state space representation of a system. The latter is composed of a state model that describes the evolution of the state vector over time, and an observation model that relates the state vector to the observations. When considering I models in a MM approach, we suggest introducing the subscript $i = 1, ..., I$ to make reference to the i^{th} model. Thus one has:

$$\boldsymbol{x}_{k+1} = \Phi_i \boldsymbol{x}_k + u_k \quad (1)$$
$$\boldsymbol{y}_k = H_i \boldsymbol{x}_k + b_k \quad (2)$$

where $\boldsymbol{x}_k$ is the state vector at time k, Φ_i is the transition matrix, u_k is the model noise assumed to be zero-mean Gaussian with covariance matrix Q_i. $\boldsymbol{y}_k$ denotes the observation and H_i is the observation matrix. In addition, b_k is a zero-mean Gaussian noise with covariance matrix R_i and is uncorrelated with u_k.
For the sake of simplicity, the time-invariant system is addressed, but the approach could be easily generalized to the time-varying case.
To measure the similarity between two state models, we propose to compute the JD between the joint distribution of the $N+1$ successive values of the state vector for the 1^{st} model and the 2^{nd} model respectively. We show that the latter admits a recursive expression.
In the following, let us denote $\boldsymbol{x}_{0:N} = (x_0, ..., x_N)$ the set of values of the state vector from time 0 to time N. In addition $p_1(\boldsymbol{x}_{0:N})$ and $p_2(\boldsymbol{x}_{0:N})$ are respectively the joint distributions of the $N+1$ successive values of the state vector for the 1^{st} and the 2^{nd} models. By definition, the expression of the KL divergence between both state models is:

$$KL(p_1(\boldsymbol{x}_{0:N}), p_2(\boldsymbol{x}_{0:N})) = \int_{\boldsymbol{x}_{0:N}} p_1(\boldsymbol{x}_{0:N}) \log \frac{p_1(\boldsymbol{x}_{0:N})}{p_2(\boldsymbol{x}_{0:N})} d\boldsymbol{x}_{0:N} \quad (3)$$

Then, as $\boldsymbol{x}_{0:N}$ is a markov process, $p_1(\boldsymbol{x}_{0:N})$ and $p_2(\boldsymbol{x}_{0:N})$ factorize as follows, for $i = 1$ or 2:

$$p_i(\boldsymbol{x}_{0:N}) = p_i(\boldsymbol{x}_{0:N-1}) p_i(\boldsymbol{x}_N | \boldsymbol{x}_{N-1}) \quad (4)$$

By inserting (4) in (3), one has:

$$KL(p_1(\boldsymbol{x}_{0:N}), p_2(\boldsymbol{x}_{0:N})) = \int_{\boldsymbol{x}_{0:N-1}} p_1(\boldsymbol{x}_{0:N-1}) \log \frac{p_1(\boldsymbol{x}_{0:N-1})}{p_2(\boldsymbol{x}_{0:N-1})} \int_{\boldsymbol{x}_N} p_1(\boldsymbol{x}_N | \boldsymbol{x}_{N-1}) d\boldsymbol{x}_N d\boldsymbol{x}_{0:N-1} + \int_{\boldsymbol{x}_{0:N}} p_1(\boldsymbol{x}_{0:N-1}) p_1(\boldsymbol{x}_N | \boldsymbol{x}_{N-1}) \log \frac{p_1(\boldsymbol{x}_N | \boldsymbol{x}_{N-1})}{p_2(\boldsymbol{x}_N | \boldsymbol{x}_{N-1})} d\boldsymbol{x}_{0:N} \quad (5)$$

As one has $\int_{\boldsymbol{x}_N} p_1(\boldsymbol{x}_N | \boldsymbol{x}_{N-1}) d\boldsymbol{x}_N = 1$, it follows:

$$KL(p_1(\boldsymbol{x}_{0:N}), p_2(\boldsymbol{x}_{0:N})) = KL(p_1(\boldsymbol{x}_{0:N-1}), p_2(\boldsymbol{x}_{0:N-1})) + G$$

where G is the 2^{nd} term of the addition in (5). At that stage, let us focus on that latter. After integrating out $x_{0:N-2}$, G becomes:

$$G = \int_{\boldsymbol{x}_{N-1:N}} p_1(\boldsymbol{x}_N, \boldsymbol{x}_{N-1}) \log \frac{p_1(\boldsymbol{x}_N | \boldsymbol{x}_{N-1})}{p_2(\boldsymbol{x}_N | \boldsymbol{x}_{N-1})} d\boldsymbol{x}_{N-1:N} \quad (6)$$

Then the above equation (6) can be reformulated as follows[1]:

$$G = \int_{\boldsymbol{x}_{N-1}} p_1(\boldsymbol{x}_{N-1}) \times \int_{\boldsymbol{x}_N} p_1(\boldsymbol{x}_N | \boldsymbol{x}_{N-1}) \log \frac{p_1(\boldsymbol{x}_N | \boldsymbol{x}_{N-1})}{p_2(\boldsymbol{x}_N | \boldsymbol{x}_{N-1})} d\boldsymbol{x}_N d\boldsymbol{x}_{N-1} = \mathrm{E}_{p_1(x_{N-1})}[KL(p_1(\boldsymbol{x}_N | \boldsymbol{x}_{N-1}),\ p_2(\boldsymbol{x}_N | \boldsymbol{x}_{N-1}))] \quad (7)$$

In (7), $p_1(\boldsymbol{x}_N | \boldsymbol{x}_{N-1})$ and $p_2(\boldsymbol{x}_N | \boldsymbol{x}_{N-1})$ are hence useful. In our case, given (1) for $i = 1$ or 2, one has:

$$p_i(\boldsymbol{x}_N | \boldsymbol{x}_{N-1}) \sim \mathcal{N}(\Phi_i \boldsymbol{x}_{N-1},\ Q_i) \quad (8)$$

Then, using the formula of the KL divergence between two multivariate normal densities[2] $\mathcal{N}_1$ and $\mathcal{N}_2$, it ensues:

$$G = \int \frac{1}{2}[Tr(Q_2^{-1} Q_1) + [(\Phi_2 - \Phi_1)\boldsymbol{x}_{N-1}]^T Q_2^{-1} [(\Phi_2 - \Phi_1)\boldsymbol{x}_{N-1}] - l - \log \frac{\det Q_1}{\det Q_2}]\ p_1(\boldsymbol{x}_{N-1}) d\boldsymbol{x}_{N-1} \quad (9)$$

If one introduces $\Delta\Phi$ as $\Phi_2 - \Phi_1$, this leads to:

$$G = \frac{1}{2}[\mathrm{Tr}(Q_2^{-1} Q_1) - l - \log \frac{\det Q_1}{\det Q_2} + \int [\Delta\Phi \boldsymbol{x}_{N-1}]^T Q_2^{-1} [\Delta\Phi \boldsymbol{x}_{N-1}] p(\boldsymbol{x}_{N-1}) d\boldsymbol{x}_{N-1}] \quad (10)$$

After rearranging the terms and integrating, we obtain:

$$G = \frac{1}{2}[\mathrm{Tr}(Q_2^{-1} Q_1) - l - \log \frac{\det Q_1}{\det Q_2} + \mathrm{E}_{p_1(\boldsymbol{x}_{N-1})}\ [\mathrm{Tr}(\boldsymbol{x}_{N-1} \boldsymbol{x}_{N-1}^T (\Delta\Phi^T Q_2^{-1} \Delta\Phi))]] \quad (11)$$

Then, by linearity of the trace operator, we have:

$$G = \frac{1}{2}[\mathrm{Tr}(Q_2^{-1} Q_1) - l - \log \frac{\det Q_1}{\det Q_2} + \mathrm{Tr}\ [P_{N-1}^1 (\Delta\Phi^T Q_2^{-1} \Delta\Phi))]] \quad (12)$$

[1] $\mathrm{E}_{p_1(x_{N-1})}$ denotes the expectation over $p_1(x_{N-1})$.
[2] The KL divergence between two multivariate normal densities is:

$$KL(\mathcal{N}_1, \mathcal{N}_2) = \frac{1}{2}\ [\mathrm{Tr}(Q_2^{-1} Q_1) + (\mu_2 - \mu_1)^T Q_2^{-1} (\mu_2 - \mu_1) - l - \log \frac{\det Q_1}{\det Q_2}\ ]$$

where μ_1, μ_2 and Q_1, Q_2 are the means and covariances of both densities, l is the dimension of the state vector and Tr is the trace of the matrix.

where the matrix P^1_{N-1} satisfies $P^1_{N-1} = \mathrm{E}_{p_1(\boldsymbol{x}_{N-1})}\left[\boldsymbol{x}_{N-1}\boldsymbol{x}^T_{N-1}\right]$. Given (1), it can be computed recursively as follows:

$$P^1_{N-1} = \Phi_1 P^1_{N-2}\Phi^T_1 + Q_1 \quad (13)$$

By combining (2), (12) and (13), $KL(p_1(\boldsymbol{x}_{0:N}), p_2(\boldsymbol{x}_{0:N}))$ can be recursively computed. As one can operate similarly for $KL(p_2(\boldsymbol{x}_{0:N}), p_1(\boldsymbol{x}_{0:N}))$, we can derive the JD from the KL:

$$JD(p_1(\boldsymbol{x}_{0:N}), p_2(\boldsymbol{x}_{0:N})) = KL(p_1(\boldsymbol{x}_{0:N}), p_2(\boldsymbol{x}_{0:N})) + KL(p_2(\boldsymbol{x}_{0:N}), p_1(\boldsymbol{x}_{0:N})) \quad (14)$$

Thus, one deduces:

$$JD(p_1(\boldsymbol{x}_{0:N}), p_2(\boldsymbol{x}_{0:N})) = JD(p_1(\boldsymbol{x}_{0:N-1}), p_2(\boldsymbol{x}_{0:N-1})) + A + B \quad (15)$$

where:

$$A = -l + \frac{1}{2}[\mathrm{Tr}(Q_1^{-1}Q_2) + \mathrm{Tr}(Q_2^{-1}Q_1)]$$

$$B = \frac{1}{2}[\mathrm{Tr}\,[P^1_{N-1}(\Delta\Phi^T Q_2^{-1}\Delta\Phi)] + \mathrm{Tr}\,[P^2_{N-1}(\Delta\Phi^T Q_1^{-1}\Delta\Phi)]]$$

In (15), the term A is constant, whereas B evolves in time. In the sequel, we take advantage of the recursive equation (15) to illustrate the influence of the terms A and B. For this purpose, we study motion models in the field of object tracking.

3. JD COMPUTATION FOR OBJECT TRACKING

3.1. System presentation

First, let us recall the UAM and the Singer motion model. Given (1) and (2), the state vector satisfies in both cases:

$$\boldsymbol{x}_k = [x_k, \dot{x}_k, \ddot{x}_k]^T \quad (16)$$

where x_k denotes the position, $\dot{x}_k$ the velocity and $\ddot{x}_k$ the acceleration.
In addition, the transition matrices and the observation matrices are defined as follows:

$$\Phi_{UAM} = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \quad (17)$$

where T is the sampling period.

$$\Phi_{Sin} = \begin{bmatrix} 1 & T & \frac{1}{\alpha^2}(\alpha T - 1 - \rho) \\ 0 & 1 & \frac{1}{\alpha}(1-\rho) \\ 0 & 0 & \rho \end{bmatrix} \quad (18)$$

where $\rho = e^{(-\alpha T)}$ and $\alpha = \frac{1}{\tau_m}$ with τ_m is the Singer time constant. The covariance matrices of both motions can be respectively defined as:

$$Q_{UAM} = \sigma^2_{UAM}\begin{bmatrix} T^5/20 & T^4/8 & T^3/6 \\ T^4/8 & T^3/3 & T^2/2 \\ T^3/6 & T^2/2 & T \end{bmatrix} \quad (19)$$

$$Q_{Sin} = 2\alpha\sigma^2_{Sin}\begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \quad (20)$$

where σ^2_{UAM} and $2\alpha\sigma^2_{Sin}$ are the jerk variances, and:

$$\begin{aligned}
q_{11} &= \frac{1}{2\alpha^5}(2\alpha T - 2\alpha^2T^2 + 2\alpha^3T^3/3 - 4\alpha T\rho - \rho^2 + 1) \\
q_{12} &= q_{21} = \frac{1}{2\alpha^4}(\alpha^2T^2 + 1 + \rho^2 + \rho(2\alpha T - 2) - 2\alpha T) \\
q_{22} &= \frac{1}{2\alpha^3}(2\alpha T - 3 + 4\rho - \rho^2) \\
q_{13} &= q_{31} = \frac{1}{2\alpha^3}(1 - 2\alpha T\rho - \rho^2) \\
q_{23} &= q_{32} = \frac{1}{2\alpha^2}(1-\rho)^2 \text{ and } q_{33} = -\frac{1}{2\alpha}(\rho^2 - 1)
\end{aligned}$$

In the next sub-section, let us compare various models and analyze the influence of T, τ_m, σ^2_{Sin} and σ^2_{UAM} on the JD.

3.2. JD in various cases

Two cases are investigated:

1/ Let us compare two UAMs defined by two different jerk variances $\sigma^2_{UAM,i}$, $i = 1$ or 2. In this case, $\Delta\Phi = 0$. According to (15), the JD is hence a linear expression of the number of samples. It only depends on the quantity A which is the slope. Given (19), it can be expressed by means of the jerk variance ratio. Indeed, the higher the ratio is, the higher the slope of the curve is, as confirmed by Fig. 1. In addition, if the variance ratio is equal to 1, then $A = 0$; as B is also equal to 0, the JD is necessary null. It should be noted that when comparing two UAM motions, as the JD only depends on the jerk variance ratio, there is an infinity of set of jerk variances that leads to the same JD.

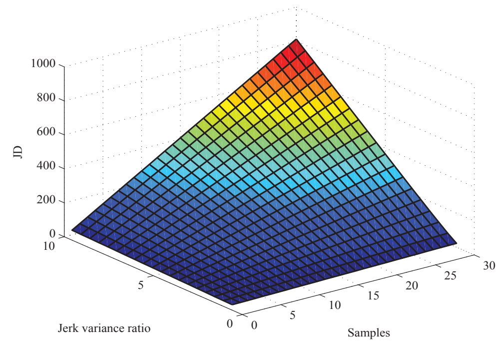

Fig. 1. Influence of the jerk variance ratio for UAMs on the JD

2/ Let us now compare a UAM and a Singer motion. In that case $\Delta\Phi \neq 0$. As a consequence, the JD is a non-linear expression of the motion sample number. Nevertheless, depending on the parameters such as T, τ_m, σ^2_{Sin} and σ^2_{UAM}, the quantity B may be negligible compared to A. Let us look more carefully

at the influence of these parameters on the JD:
Influence of the jerk variance ratio: we have studied various cases. In this paper and without loss of generality, let us set the model parameters as follows: $T = 1s$, $\tau_m = 30s$, $2\alpha\sigma^2_{Sin} = 0.1g$, where g denotes the gravity constant equal to $9.81m.s^{-2}$. The jerk variance ratio evolves from 1 to 10. From Fig. 2, the JD increases with both the variance ratio and the number of samples. With this choice of parameters, $\Delta\Phi$ is close to zero, hence B is rather negligible compared to A. This leads to a quasi-linear behavior of the JD.

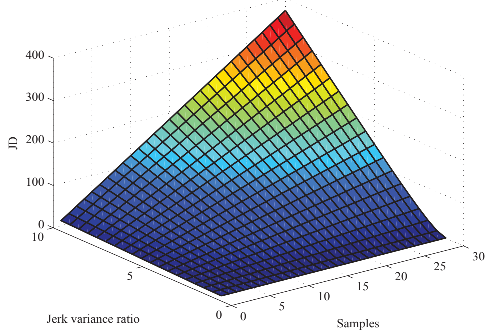

Fig. 2. Influence of the jerk variance ratio for a UAM and a Singer model on the JD

Influence of τ_m: here $T = 1s$, $\sigma^2_{UAM} = 2\alpha\sigma^2_{Sin} = 0.1g$. The value of τ_m evolves from 1 to 5s. The JD increases with both the number of samples and the value of τ_m due to the contribution of the terms A and B. When τ_m becomes high, the JD is close to zero. This result confirms what is commonly known about UAM and Singer models. Indeed, they are embedded models: for a high value of τ_m, both Singer transition and covariance matrices converge to the ones of the UAM [3].

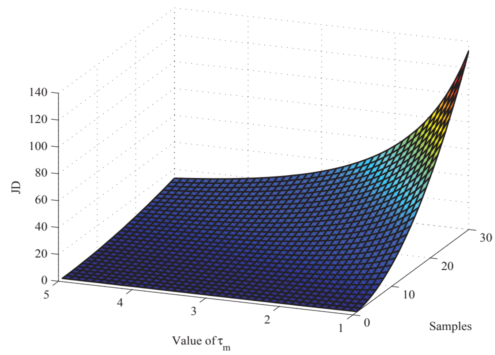

Fig. 3. Influence of the Singer model constant on the JD

Influence of T: here $\sigma^2_{UAM} = 2\alpha\sigma^2_{Sin} = 0.1g$, $\tau_m = 30s$. Due to (17) and (18), the higher T is, the higher the JD is. See Fig. 4. Moreover, if the value of T is low, $\Delta\Phi$ is nearly zero, then the JD becomes almost linear. In addition, A tends to $-l + \frac{l}{2}(\frac{\sigma^2_{UAM}}{2\alpha\sigma^2_{Sin}} + \frac{2\alpha\sigma^2_{Sin}}{\sigma^2_{UAM}})$. So, in our case A tends to zero and the JD is close to zero. If T increases, $\Delta\Phi$ is non-negligible, the JD becomes slightly non-linear.

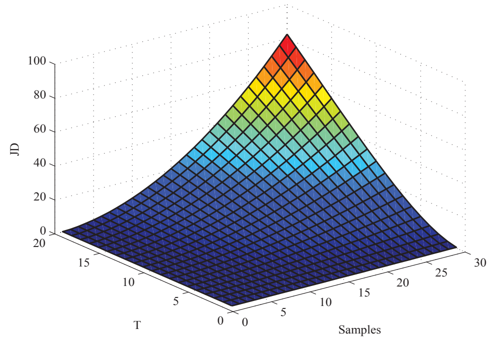

Fig. 4. Influence of the sampling period on the JD

In the next section, we take advantage of the previous study to discuss the choice of an IMM estimator over a Kalman filter, as well as for model set design when using the IMM.

4. JD VS IMM

This section addresses the issue of designing an efficient estimation algorithm to retrieve an object trajectory, based on the preliminary study on the JD in section 3. We consider two simulation scenarios.
In the 1st place, our purpose is to investigate when it is relevant to use an IMM instead of a standard Kalman filter. Various cases are studied and differ by the JD between the actual motion models of the trajectory we wish to estimate. The latter are assumed to be known, which is not realistic in practice. Note that this is a complementary study to the one proposed in [1].
In the 2nd place, we analyze the impact of the JD between the IMM models when we have few information on the unknown trajectory.

First scenario. The reference trajectory (denoted *Ref.*) is composed of $N = 500$ samples and switches between two models, one being a UAM and the other a Singer model. The sampling period is $T = 1s$ and the probability of a motion model change at a given time instant is set to 0.05. The Singer model parameters are fixed and equal to $\sigma_{Sin} = 0.1g$ and $\tau_m = 60s$ whereas the standard deviation σ_{UAM} is made to vary so that different JD between both models can be considered. Only a noisy observation of this trajectory is available where the measurement noise variance is $R = 200$. Then, an IMM based exactly on the above-mentioned models and transition probabilities is implemented to carry out the estimation. It is compared to two

Kalman filters that are independently run by the practitioner. One is based only on the UAM state model, and denoted *Kal.* 1, whereas the other is based only on the Singer model, and referred to as *Kal.* 2. The estimation accuracy is evaluated by computing the RMSE by averaging over 200 realizations of the reference trajectory and the measurement noise. The results are reported in Table 1 where the parameters of the motion models as well the corresponding JD between them are given in the first rows. The value that serves as an indicator for the JD is taken arbitrarily 20 samples after the beginning of the motion. It should be noted that, as a reference, we have also used the 3 algorithms to estimate trajectories generated entirely by a single model: either a UAM model (denoted *UAM traj.*) or a Singer model (denoted *Sin. traj.*).

By analyzing Table 1, unsurprisingly it appears that for single-model based trajectories (namely *UAM traj.* and *Sin. traj.*), the Kalman filter based on the proper model outperforms the other algorithms. However, the IMM yields close results. It should be noted that when σ^2_{Sin} increases, the RMSE also increases due to a higher uncertainty on the unknown trajectory. As for *Ref.*, we can observe that, whatever the JD, the IMM is more accurate than the single-model based Kalman filters. Furthermore, *Kal.* 1 is always more accurate than *Kal.* 2. Indeed, the Singer model is embedded in the UAM. Thus, although *Kal.* 1 is outperformed by *Kal.* 2 for Singer motion phases, its error does not grow unbounded. Last but not least, the difference between the RMSE of the IMM and *Kal.* 1 is all the more significant as the JD is high. As a conclusion, if the JD between the switching state models is small, the IMM does not improve the estimation compared to a standard Kalman filter. On the contrary, the IMM becomes more relevant as the JD is significant.

$\sigma_{UAM}(\times g)$		0.012	0.041	0.108	0.237	0.471
$\sigma_{Sin}(\times g)$		0.1	0.1	0.1	0.1	0.1
JD		20	100	1000	5000	20000
UAM traj.	Kal. 1	4.92	5.97	7.18	8.55	10.53
	Kal. 2	5.13	8.15	19.08	37.78	75.35
	IMM	4.98	6.05	7.29	8.68	10.70
Sing. traj.	Kal. 1	5.26	5.50	6.33	7.37	8.53
	Kal. 2	5.12	5.12	5.12	5.12	5.12
	IMM	5.15	5.28	5.66	6.15	6.57
Ref.	Kal. 1	5.07	5.69	6.79	8.03	9.76
	Kal. 2	5.09	6.11	10.92	21.74	43.78
	IMM	5.02	5.60	6.61	7.69	9.17

Table 1. RMSE between the state vector and its estimate for 500 samples, averaged over 200 Monte Carlo simulations.

Second scenario. We generate a trajectory as a sequence of 5 different Singer models defined by the same time constant $\tau_m = 1s$ and their respective acceleration variances $\sigma^2_{Sin} = ([0.1\ 0.3\ 0.05\ 0.2\ 0.03] \times g)^2$. Each motion phase is composed of 100 samples, hence the simulation is 500 samples long. Parameters are set as follows: $T = 1s$ and $R = 200$. For the tracking, it is assumed that we have no precise information on the target motion. The purpose is to study the performance of the IMM estimator for various sets of competing models corresponding to different JD. For the sake of simplicity, we consider an IMM algorithm using only two Singer models with the true value $\tau_m = 1s$. Thus, only the model variances are undefined. For the first model, we arbitrarily set the variance to a low enough value in order to represent low maneuvering motions, *e.g.* $\sigma^2_{Sin,1} = (0.04 \times g)^2$. Conversely, the second-model variance $\sigma^2_{Sin,2}$ is made to vary and we compute the corresponding JD. To evaluate the performance of the estimator, we figure out the RMSE over 200 realizations of the measurement noise. According to Table 2, the RMSE decreases with the JD until a minimum value, then increases again. This minimum value is reached when $\sigma^2_{Sin,2}$ is close to the upperbound of the theoretical variance values, i.e. $(0.3 \times g)^2$. In the general case, it appears that it is better to overestimate the model variance rather than to underestimate it. As a consequence, we suggest the IMM user to choose models with a high enough JD, corresponding to sufficient dissimilarities. In this way, the estimator can capture a greater variety of motions.

$\sigma_{Sin,2}(\times g)$	0.10	0.20	0.30	0.40	0.50	0.70
RMSE	6.37	5.71	5.73	5.80	5.88	5.96
JD	132	691	1628	2940	4628	9128

Table 2. RMSE between the state vector and its estimate for 500 samples, averaged over 200 Monte Carlo simulations.

5. CONCLUSIONS AND PERSPECTIVES

In this paper, we proposed an approach to evaluate the similarity between two state models. For this purpose, we derived a recursive expression of the JD and we applied in the context of target tracking. Our first simulation results showed that the JD is a relevant indicator with regards to the choice of the estimation algorithm as well as the design of the model set. We are currently carrying out a more comprehensive study to derive more precise guidelines for the practitioner. Other applications will be studied. Finally, we plan to also investigate the recursive JD expression for state vectors of different sizes.

6. REFERENCES

[1] Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, YBS Publishing, Storrs, 1995.

[2] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting Multiple Model Methods in Target Tracking: A Survey," *IEEE Trans. on Aerospace and Elec. Syst.*, vol. 34, pp. 103–123, Jan. 1998.

[3] X. R. Li and V. P. Jilkov, "Survey of Maneuvering Target Tracking. Part V: Multiple Model Methods," *IEEE Trans. on Aerospace and Elec. Syst.*, vol. 41, pp. 1255–1321, Oct. 2005.

[4] T. Kirubarajan and Y. Bar-Shalom, "Kalman Filter versus IMM Estimator: When Do We Need the Latter?," *IEEE Trans. on Aerospace and Elec. Syst.*, vol. 39, pp. 1452–1457, Oct. 2003.