MINIMISING THE RISK OF COLLATERAL DAMAGE WITH THE KALMAN-LEVY FILTER

S. W. Sims [a], J. F. Ralph [a], K. L. Edwards [b].

[a] *Department of Electrical Engineering and Electronics, The University of Liverpool, Brownlow Hill, Liverpool, L69 3GJ, United Kingdom.*

[b] *Avionics and Vetronics Centre, Sensors and Electronics Division, QinetiQ Ltd., Farnborough, GU14 0LX, United Kingdom.*

Keywords: *Kalman filters, optimal state estimation, collateral damage.*

Abstract

This paper examines the use of Kalman filters and variants of the standard Kalman filter in guidance and navigation systems. In particular, the paper is concerned with the properties of a filter that has been developed recently for use with systems with non-Gaussian stable noise or where the noise converges slowly to the Gaussian limit. In such situations, minimisation of the estimated variance (which is the basis of the standard Kalman filter) is not compatible with the minimisation of the large errors that could be associated with collateral damage.

1 Introduction

The optimality of the Kalman filter is well known for the estimation of observed quantities in linear systems in the presence of white, Gaussian noise. For more complicated nonlinear systems the nonlinearities can often be linearised to produce a piecewise linear model that can be used as the basis of the extended Kalman filter. However, the situation is not quite so clear when the noise is non-Gaussian or correlated. In some situations, where the noise can be represented by a nonlinear process, the extended Kalman filter can be used to include the dynamics of the noise. This, and the use of the basic Kalman filter itself, is based upon the validity of the central limit theorem [1]. That is, the errors in the estimated state of the system will tend toward a Gaussian distribution, even if the errors in the individual observations are non-Gaussian.

This version of the central limit theorem is valid in the limit of infinite observations and for noise distributions with finite variance. For situations where the number of observations is finite the tail of the error distribution can be significantly non-Gaussian [2], and in situations where the noise processes have an infinite variance the error distribution converges to a stable Levy distribution [2] (a Gaussian being the most common example of a stable Levy distribution). These considerations, together with the fact that many physical phenomena seem to have noise distributions that obey a generalised non-Gaussian Levy law, have inspired recent work on generalisations of the Kalman filter to include situations where the underlying distributions have infinite variance [3,4] and where the convergence of the tail of the error distribution toward a Gaussian tail is very slow [2]. One such example is the Kalman-Levy filter recently proposed by Sornette and Ide [5].

Whilst the Kalman filter aims to minimise the variance of the estimated quantities (which usually correspond to the 'small' errors toward the centre of the distribution), the Kalman-Levy filter aims to minimise the 'large' errors in the tails of the distributions. In their paper, Sornette and Ide demonstrate that the simultaneous minimisation of the small and large errors is only possible when the underlying distributions are Gaussian.

For situations where the variance is infinite, or where the tails have incompletely converged to the Gaussian limit, the minimisation of the large errors is non-optimal in the sense of minimum variance. This fact has profound implications in the study of weapon guidance [6] and the design of advanced targeting systems [7], since it implies that it is not possible to maximise the terminal accuracy (minimised variance or 'small' errors) and to minimise the risk of collateral damage (minimised tails or 'large' errors) unless the errors are sufficiently Gaussian, and the system is sufficiently linear, for the differences to be negligible. This paper examines the implications of this result for the types of guidance-navigation systems used in advanced air-to-ground weapons.

2 Filters and α -Stable Processes

The representation of a series of measurements in the presence of noise by a stochastic process should be familiar to most people with a technical or scientific background. Such a process is normally represented by a Probability Density Function (PDF) that can be characterised in several different ways: either using an explicit analytical form, a set of moments of the distribution, or a characteristic function (the Fourier transform of the PDF). In many cases, it is assumed that the underlying value to be measured is fixed (or at least slowly varying when compared to the noise) and that the measurement noise is responsible for fluctuations about this mean value. By taking more measurements and constructing some sort of average, it is possible to improve the estimate of the measured quantity or system state. The conventional way to do this is to weight different measurements according to their expected error; penalising less accurate measurements. (Although the state to be measured is usually assumed to be slowly varying, different sensors may be employed with different noise characteristics and the noise in any sensor can vary from measurement to measurement). This is the basis of the widely used Kalman filter, which calculates a weighted average that minimises the variance (or covariance) of the measurement noise.

The variance is the standard measure for the expected error. Its usefulness is based on the Central Limit Theorem, which (in its most popular form) says that the distribution of a sum or average of random variables tends toward a Gaussian or Normal distribution as the number of random variables in the sum is increased. This is important because a Gaussian distribution has a well defined mean value (which will hopefully represent the desired measurement), a finite variance (representing the square of the expected error), and zero higher-order moments. This means that the distribution of the weighted average should end up being the same, whatever the original noise distribution (or distributions). All that is required is an estimate of the variance of the noise and sufficient measurements for the distribution of the average to converge sufficiently close to the desired Gaussian form.

Unfortunately, the popular form of the Central Limit Theorem is a simplification of the true situation. There is a larger class of distributions that are non-Gaussian and stable under addition. These are the α-stable distributions (often referred to as Levy distributions) that do not converge to a Gaussian limit when summed and have an infinite variance, which makes them harder to characterise by an expected error value in the conventional sense. As noise, they are distinguished by occasional very large fluctuations, giving rise to a very long tailed distribution.

Many systems contain noise that can be approximated by Gaussian distributions when the fluctuations are small enough for the system to be almost linear. However, most systems are nonlinear to some extent and will contain some form of non-Gaussian noise, although not all will be α-stable. In cases where the nonlinearities are large and/or where multiplicative noise is present, the fluctuations can have a Levy distribution (or an approximate or 'truncated' Levy distribution).

2.1 Symmetric a -Stable Processes

Most probability distributions can be characterised by an infinite set of moments about the mean value (variance, skewness, kurtosis and other higher order moments). In some cases, the higher order moments are either zero (as with the Gaussian distribution) or are related to the lower order moments (as is the case with the Poisson distribution). For some distributions, such as the log normal distribution, this representation in terms of moments can lead to ambiguities because they are not defined uniquely by their moments, but these distributions are not considered here. For other distributions, including the α-stable distributions, some or all of the moments are infinite. The Gaussian distribution is the only stable distribution that has a finite variance, and some α-stable distributions do not even have a well-defined mean. To make matters worse, there are only a few cases where the PDF for a Levy distribution can be written in closed form.

Instead, the most common way to specify an α-stable distribution is through its characteristic function ($0 < a \leq 2$) [8]:

$$\hat{L}_a(k) = \begin{cases} \exp\left\{imk - g|k|^a\left[1 + ib\frac{k}{|k|}\tan\left(\frac{ap}{2}\right)\right]\right\} & for\ a \neq 1 \\ \exp\left\{imk - g|k|^a\left[1 + ib\frac{k}{|k|}\left(\frac{2}{p}\right)\ln|k|\right]\right\} & for\ a = 1 \end{cases} \quad (1)$$

which contains four important quantities:

- a, the index of stability.
- b, the skewness parameter.
- g, the scale parameter.
- m, the location parameter.

The index of stability determines the behaviour of the tail of the distribution, the smaller the value of a, the longer the tail. The skewness parameter determines the asymmetry of the distribution, but in the cases considered in this paper attention will be restricted to situations where $b = 0$ and the distributions are symmetric.

The scale parameter determines the width of the distribution compared to some length scale $l \approx C^{1/a}$, where C is referred to as the scale parameter or 'tail covariance' and is defined by [2,5]:

$$C = \frac{pg}{a^2\Gamma(a-1)\sin\left(\frac{ap}{2}\right)} \quad (2)$$

The location parameter is the equivalent of the mean value. For $1 < a \leq 2$, the mean is defined and is equal to m. However, for $0 < a \leq 1$ the mean is not well-defined whilst the location parameter is defined and it plays an equivalent role. For simplicity's sake, all of the examples discussed in this paper will be centred on $m = 0$, but the generalisation to other situations is straightforward.

It is worth noting that for the special case $a = 2$, the distribution reduces to the Gaussian case, where the tail covariance reduces to the variance and the characteristic length scale is equal to the standard deviation associated with the variance.

Although there are only a few examples where an α-stable distribution can be written in closed form, it is possible to say some general things about their general form. In particular, the tail of the distribution $P(x)$ is always a power law, that is

$$P(x) \approx \frac{C}{|x|^{1+a}} \quad (3)$$

for $x \to \pm\infty$ (for $b = 0$ and $m = 0$ at least).

2.2 Almost Stable Processes

As mentioned in a previous section, systems which contain multiplicative noise or which exhibit very nonlinear behaviour are most likely to contain noise that has a Levy distribution. However, in many cases the noise will only approximate to an α-stable distribution. The very large tails that are present in α-stable distributions can contain very large fluctuations, and whilst the theoretical noise distribution goes to infinity, there will generally be some upper cut-off that limits the size of the fluctuations present in any real system. Reference 2 gives the distribution of energy in earthquakes as an example, where the energy released in an earthquake is approximately Levy distributed, but an infinitely long tail would require the presence of earthquakes with an infinite energy.

The result of these physical limitations or cut-offs would be that the Levy distribution predicted by whatever theoretical model was used to simulate the behaviour of the system would be truncated at some point. The resultant truncated Levy distribution would always have a finite variance and would not be α-stable. However, for situations where the approximation to a true α-stable distribution is fairly good (i.e. the cut-off is very large compared to the characteristic length scale) the convergence to a Gaussian stable distribution can be very slow. In fact, for many applications with a finite number of measurements, the convergence to a Gaussian stable distribution could be hardly noticeable.

For some types of cut-off the convergence rates can be calculated [2]. As a general rule, the maximum expected value of a Levy process increases with the number of measurements N as $(NC)^{1/a}$. As long as the cut-off is significantly larger than this length scale, so that the cut-off has a negligible effect, the truncated distribution should be a good approximation to a true Levy distribution [2].

2.3 Unstable Distributions with Power Law Tails

In practice, the distribution of the sum or average of noisy measurements tends to converge to a Gaussian distribution from the middle outwards. This means that even a distribution with a power law tail that vanishes quicker than a Gaussian distribution ($a > 2$ in equation 3) may not be sufficiently converged after a finite number of measurements for the result of a Kalman filter to have a Gaussian distribution.

Reference 2 calculates the cross over between the Gaussian centre of the distribution and the power law tails using a power series expansion of the characteristic function. The crossover for a series of N independent and identically distributed measurements is found to be approximately proportional to $\sqrt{N \ln N}$.

This property is important in the context of this paper because it is the tail of the distribution that is mainly associated with the risk of collateral damage. For advanced targeting systems and autonomous guided weapon systems, very large deviations from the desired trajectory or impact point are more likely to cause collateral damage. As a result of the finite number measurements that can be taken, any guidance system that is based around a Kalman filter is likely to have state estimates that have distributions with non-Gaussian tails (either due to incomplete convergence to a stable Gaussian distribution or due to the presence of α-stable noise). This is a general property of linear filters, including the Kalman-Levy proposed by Sornette and Ide [5]. The difference between the Kalman-Levy filter and conventional Kalman filters (including extended Kalman filters that work by linearising a nonlinear system of equations about a set of state estimates) is that the Kalman-Levy filter aims to minimise the weight associated with the tails of the distribution by minimising the tail covariance C.

3 The Kalman-Levy Filter

The general form of the Kalman-Levy filter is similar to the standard Kalman filter in that is consists of a series of steps that can be associated with:

1. State and error prediction.
2. Measurement prediction.
3. Calculation of the Kalman gain.
4. State and error update.

In the Kalman filter, the Kalman gain matrix is calculated by minimising the estimated covariance (error) matrix. In the Kalman-Levy filter, the Kalman gain is calculated by minimising the tail covariance matrix.

To simplify things, this paper will be restricted to one-dimensional problems, although the general form of the filter is given in reference 5. The multi-dimensional filter is slightly more complex to implement, but the one-dimensional case embodies all of the principal properties of the algorithm.

3.1 Tail Covariance

At the heart of the filter is the parameter C that characterises the weight associated with the tails of the distribution (see equation (3)). Rather than estimating the variance of the errors in the state estimates, as in the Kalman filter, the Kalman-Levy filter predicts and updates an estimate of the tail covariance associated with the state estimates. The important properties of the tail covariance are that it represents the characteristic length scale $l \approx C^{1/a}$ associated with the errors in the states, and the tail covariance of the sum of two random variables,

$$x = Ax_1 + Bx_2 \tag{4}$$

with the same index of stability (a) is given by [2],

$$C = |A|^a C_1 + |B|^a C_2 \tag{5}$$

This is an important result when calculating the tail covariance for a system with noise that has a power law tail for the Kalman-Levy filter (see equations (7) and (11)), and for calculating the expected tail covariance for the outputs from a Kalman filter in the presence of Levy noise.

3.2 The Kalman-Levy Filter

As with the Kalman filter, the Kalman-Levy filter contains five internal vectors and matrices, and requires five inputs. The five internal quantities are:

(i). The state vector $\{x_i\}$ for each time step or measurement, $i = 0\ldots n$.
(ii). The state tail covariance matrices $\{C_i\}$.
(iii). The Kalman gain matrices $\{K_i\}$, which are used to weight of the state updates.
(iv). The predicted state vectors $\{x_{i|i-1}\}$.
(v). The predicted state tail covariance matrices $\{C_{i|i-1}\}$.

Whilst the five inputs are:

(i). A series of measurement vectors $\{y_i\}$.
(ii). The measurement tail covariances $\{Y_i\}$.
(iii). The measurement matrices $\{H_i\}$ to relate the measurements y_i to the estimated state vectors x_i.
(iv). The dynamical matrices $\{M_i\}$ to predict the time evolution of the system states.
(v). The dynamical noise (tail covariance) matrices $\{N_i\}$ to allow for errors in the (linear) dynamical prediction process.

Of course, for a one-dimensional filter, the vectors and matrices are simply numbers.

For $a > 1$, the four steps of the Kalman-Levy filter are given by:

1. State and error prediction

$$x_{i|i-1} = M_{i-1}x_{i-1} \tag{6}$$

$$C_{i|i-1} = |M_{i-1}|^a C_{i-1} + N_{i-1} \tag{7}$$

2. Measurement prediction

$$y_{i|i-1} = H_i x_{i|i-1} \tag{8}$$

3. Calculation of the Kalman gain

$$K_i = \frac{H_i^{-1}}{\left(1 + \left(\frac{Y_i^{\frac{1}{a}}}{H_i C_{i|i-1}^{\frac{1}{a}}}\right)^{\frac{a}{a-1}}\right)} \tag{9}$$

4. State and error update

$$x_i = (1 - K_i H_i)x_{i|i-1} + K_i y_i \tag{10}$$

$$C_i = |1 - K_i H_i|^a C_{i|i-1} + |K_i|^a Y_i \tag{11}$$

The main differences between the Kalman-Levy filter and the standard Kalman filter, in the $a > 1$ regime, are the calculation of the Kalman gain and the way in which the error matrices (tail covariances) are calculated, using equation (5). For situations where $a \leq 1$, the absolute deviation from the mean and the mean are not well defined, and the Kalman-Levy update reduces to accepting the single measurement with the lowest value of the tail covariance, at the expense of all other measurements [5].

3.3 Optimality

The Kalman-Levy filter reduces to the standard Kalman filter when $a = 2$, and will provide an optimal estimate of the state when the input noise is Gaussian or when there are sufficient measurements for the tails of the distribution for the state estimate to have converged to a Gaussian distribution. In more general cases, the Kalman filter will minimise the variance (if it is well defined) and the Kalman-Levy filter will minimise the weight of the tail through the tail covariance. It can be proved that the Kalman-Levy filter minimises the tail covariance and that this minimum is unique [5]. It is also shown that simultaneously minimising the variance and the tail covariance is only possible when the measurement noise is Gaussian. If a filter is required to minimise the weight of the power law tails of a distribution, then it must be a Kalman-Levy filter, whether or not the power law is α-stable or Gaussian stable.

4 Comparison of Kalman-Levy and Kalman filters

One of the problems in proposing modifications to the basic Kalman filter is that the performance of the Kalman filter itself is generally very good. Even in situations where the performance is not optimal, it normally offers a good estimate of the system states that is fairly robust to errors in the system parameters (dynamical prediction matrix or measurement noise). By comparison, the estimates generated by the extended Kalman filter for nonlinear system can be very sensitive to errors in some of these parameters.

As suggested by section 3.3, the Kalman-Levy filter does offer an improvement over the standard Kalman filter for situations where the underlying distributions are non-Gaussian and have power law tails, but the average performance benefits are relatively small, typically less than about 10% [5]. The main advantages are that the estimated/predicted errors are far more accurate in the Kalman-Levy filter and that the performance is less sensitive to errors in the estimates of the error matrices.

The covariance is very difficult to estimate in the presence of α-stable noise or truncated Levy noise. In the presence of true α-stable noise, the second order moment is infinite, but the ratio of the variances needed for the standard Kalman filter can be found. However, the ratio is frequently unreliable. In the case of truncated Levy noise, the variance is finite and, for a sufficiently large sample, it has an expected value that is proportional to Cd^{2-a} [2], where d is the cut-off. The ratio of the variances should therefore converge to the ratio of the two tail convergences. However, the fact that these distributions include very large fluctuations means that the ratio of the variances calculated from some finite sample drawn from a truncated Levy distribution is very unreliable. In fact, the fluctuations in the estimated ratio increase as the cut-off is increased.

In situations where a is larger than 2, in the Gaussian stable regime, the variance of the noise is well defined and can be estimated directly from the noise characteristics or from a finite sample of the noise. However, the ratio of covariances used in the Kalman filter is not necessarily equal to (or even similar to) the ratio of tail covariances used in the Kalman-Levy filter. The variance of the noise is dependent on the properties of the whole distribution, the centre of the distribution and the tails, whilst the tail covariances are only dependent on the tails of the distribution. In fact, due to the rapid decay of the tails in this regime, the variances tend to be dominated by the centres of the distributions.

In this paper, three separate cases are examined:

1. Constant measurement errors.
2. Alternating measurement errors.
3. Linearly reducing measurement errors.

Each of these examples is chosen because it represents a case that might occur in an autonomous guidance system. For simplicity, all noise is scaled so that the characteristic length scale of the initial measurement noise is one, and so that the ratio of the tail covariances is equal to the ratio of the variances. The second of these assumptions means that the differences between the Kalman and Kalman-Levy filters will be relatively small. In practice, because of the unreliability of the estimates of the variances for $\alpha < 2$ and the dominance of the centre of the distribution for $\alpha > 2$, the actual differences would be expected to be much greater than shown here.

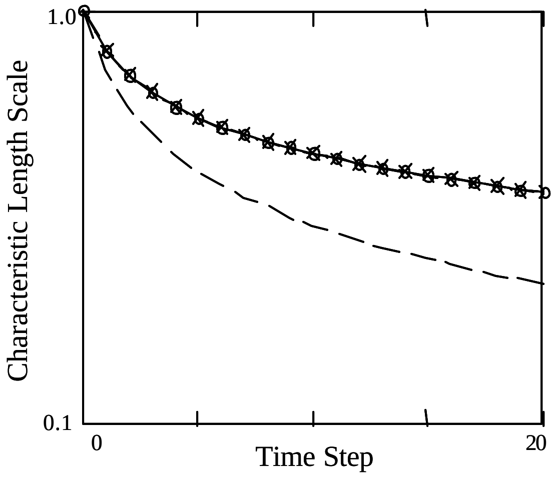

Figure 1 – Expected characteristic length scale ($l \approx C^{1/\alpha}$, log scale) as a function of time step for constant measurement error (α = 1.5): output error for Kalman filter (circles), output error for the Kalman-Levy filter (crosses), the estimated error for the Kalman-Levy filter (solid line), and the estimated error for the Kalman filter (dashed line).

4.1 Constant measurement errors

The simplest case is where the measurement errors have a constant variance or tail covariance. In this case, the true output error (as distinct from the output error estimated by the filters) is equal for the Kalman and Kalman-Levy filter. This is because each of the filters gives equal weight to each of the measurements, as should be expected, leading to identical Kalman gains at each time step. The difference is in the estimated error produced by each filter. The estimated error produced by the Kalman-Levy filter is equal to the true error in the output, but the estimated error is significantly smaller than the true error for Kalman filter for $\alpha < 2$ (see figure 1 for an example with $\alpha = 1.5$) and significantly larger than the true error for $\alpha > 2$. This means that the Kalman filter does produce an optimum output for this simple case, but that its own estimate of the expected error is wildly inaccurate. This can be very important since it can give the filter an undue confidence in the state estimate, which may be particularly important if this error estimate is to be used by another system for combining or fusing information from different sources.

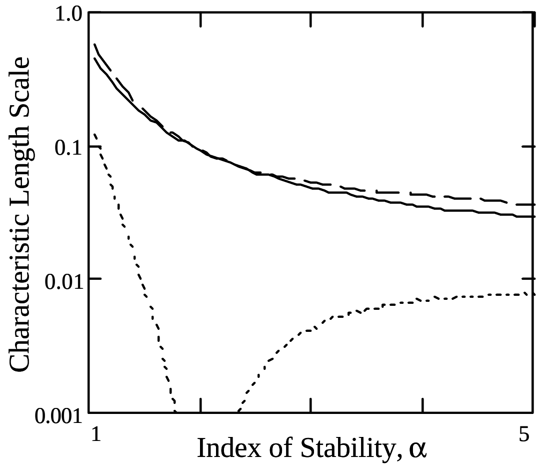

Figure 2 – Expected characteristic length scale ($l \approx C^{1/\alpha}$, log scale) as a function of α for alternating measurement errors after 50 time steps: output error for Kalman filter (dashed line), output error for the Kalman-Levy filter (solid line), and the difference between the Kalman and Kalman-Levy errors (dotted line).

4.2 Alternating measurement errors

In cases where measurements are being combined from several sensors, each sensor is likely to have a different characteristic error. The second example is intended to reflect this process, where the measurement errors alternate between two values. Figure 2 shows the characteristic length scale of true (output) error for both the Kalman and Kalman-Levy filters, together with the difference between the two, as a function of the index of stability (α). Since the tail covariance and the characteristic length scale of the noise process are related by $l \approx C^{1/\alpha}$, then one or the other must be fixed for comparison. In Figure 2, the characteristic length of the input errors is fixed and alternates between one and one-half at successive measurements.

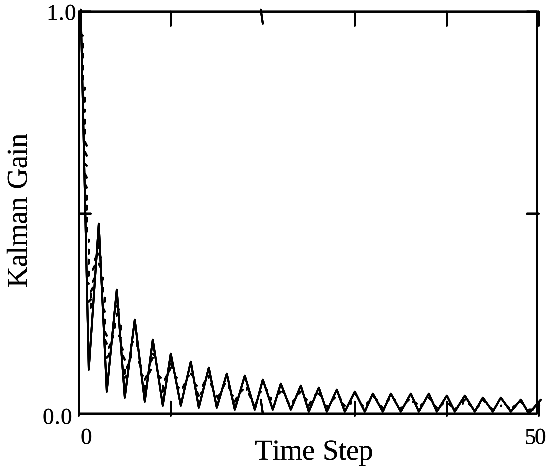

Figure 3 – Kalman gain (linear scale) as a function of time step for alternating measurement error (a = 1.5): Kalman filter (dotted line) and the Kalman-Levy filter (solid line).

As can be seen in Figure 2, the difference between the two filters is greatest when a << 2 or when a >> 2. Near a = 2, the differences disappear, as required. Even so, the differences are not extremely large, but this is due to the fact that the ratio for the variances in the Kalman filter is chosen to give an optimal Kalman filter solution. In practice, the Kalman filter would produce significantly worse output errors, for the reasons outlined above.

4.3 Linearly reducing measurement errors

The third example is most likely to reflect the type of errors that might occur in an advanced targeting system or an autonomous guidance system. Many systems use angle and/or angle rate sensors to estimate positions, either the relative positions of airframes and potential targets or the position of objects relative to some reference datum (target localisation). Examples of these types of system are passive ranging algorithms based on angle rate measurements and active (laser) targeting. In these cases the angular errors (e.g. sensor-airframe alignment errors, errors due to angular resolution limitations) are approximately constant but the distance between the object and the sensor may vary as a function of time. If the sensor is approaching the object the positional error should decrease approximately linearly with distance. This type of situation is represented in Figures 4 and 5. The characteristic length scale associated with the measurement tail covariance decreases linearly so that the tail covariance is given by,

$$Y_i = (1 - 0.02i)^a \qquad (12)$$

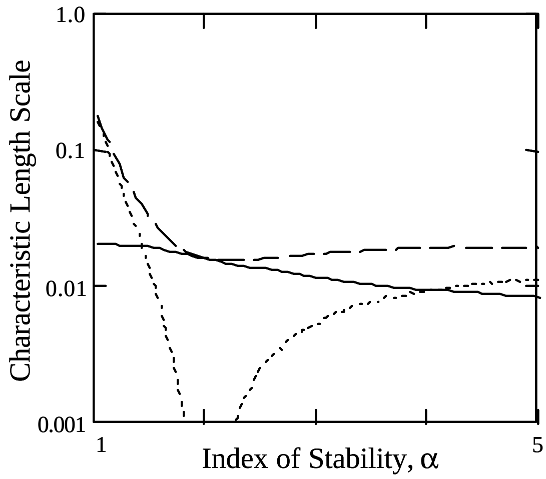

Figure 4 – Expected characteristic length scale ($l \approx C^{1/a}$, log scale) as a function of a for linearly reducing measurement errors after 49 time steps: output error for Kalman filter (dashed line), output error for the Kalman-Levy filter (solid line), and the difference between the Kalman and Kalman-Levy errors (dotted line).

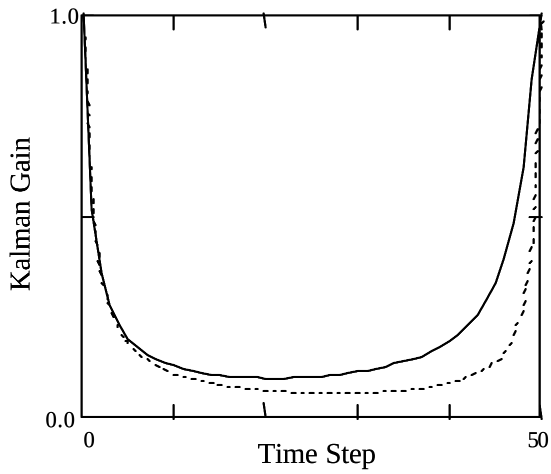

Figure 5 – Kalman gain (linear scale) as a function of time step for linearly reducing measurement error (a = 1.5): Kalman filter (dotted line) and the Kalman-Levy filter (solid line).

In this case the measurement error reduces to zero at the 50th time step, but immediately prior to this (the 49th time step) the differences between the Kalman filter and the Kalman-Levy filter are relatively large (see Figure 4). These differences are due to the undue confidence that is associated with the estimates and predicted states in the Kalman filter. When a << 2, the Kalman filter is over confident in the current/predicted state estimate and gives a larger weight to the predicted state vector than it should do (indicated by the lower Kalman gain values in Figure 5). During the 49 time steps, the effect of this over confidence accumulates, giving rise to the comparatively large expected error for the Kalman filter in Figure 4. When a >> 2, the reverse is true. Rather than being over confident, the Kalman filter is over cautious about the predicted state estimates.

5 Collateral Damage and Terminal Accuracy

The Kalman filter aims to minimise the variance of the state estimates, whilst the Kalman-Levy filter aims to minimise the tail covariance of the state estimates. It is the tail covariance that governs the weight of the power law tails, and therefore the amount of probability that is found in these tails. It will therefore determine the expected number of solutions that have large deviations from the desired solution. For applications involving guided weapon systems [6] or advanced targeting systems [7] such large deviations are far more likely to be concerned with collateral damage than the small deviations around the desired solution.

For Gaussian stable noise, the small deviations dominate the variance because they are normally far more frequent, and when one attempts to maximise the accuracy of a guidance or targeting system it is these small deviations that tend to influence the optimisation process. Such issues are likely to become more important as these systems become more sophisticated. Recent advances in information and data fusion [7,9,10] and autonomous guidance and navigation systems [11] offer significant advantages in terms of flexibility and performance. However, as these systems become more sophisticated, they tend to become more reliant on the accuracy and availability of estimates for the noise present in the underlying processes, and as they become more complex the nature of this noise becomes harder to characterise. For example, one of the new techniques that has been developed for both weapon guidance and air-to-ground targeting is scene-matching and area correlation [7,11]. This type of system correlates a database of ground features (normally generated from maps and/or reconnaissance imagery) with imagery generated by an onboard camera. This allows the position of objects found in the onboard imagery to be referenced against an external database and can be used to assist the aircraft's navigation system. In a guided weapon system it can also be used for terminal phase aim point optimisation. The problem with characterising the noise in this type of system is that the properties of the noise are very dependent on the features that appear in the scene. For example, areas of ground with a grid plan layout tend to pose particular problems for scene-matching algorithms because there are generally a large number of possible ways to maximise the correlations between the database and the imagery. As such, these systems might be expected to generate noise distributions with very large tails, of the type discussed in this paper, and therefore be good candidates for the Kalman-Levy filter.

6 Conclusions

This paper has considered the use of a variant of the conventional Kalman filter, the Kalman-Levy filter proposed by Sornette and Ide, for the minimisation of the risk of collateral damage. This filter aims to minimise the weight of the tails of the error distributions associated with the filtered state vectors. As a result the Kalman-Levy filter minimises the errors associated with large fluctuations, which are more likely to be associated with collateral damage.

Where the measurement noise is Gaussian, the Kalman-Levy filter reduces to the conventional Kalman filter, but where the measurement noise has a power law tail the Kalman-Levy filter minimises the weight of the tail. This is true if the noise has either a long (α-stable) tail or a short (α > 2) tail. However, in the case of the short tail, this minimisation is at the expense of an increase in the variance of the state estimates. In that case of a long tail, the variance is not strictly defined, and any estimate of the variance used in the Kalman filter is likely to be unreliable.

The interesting aspect of this filter is that it shows that the desire to maximise the accuracy of a sensor or guidance system (as characterised by the variance) is incompatible with the desire to minimise the large deviations unless the measurement noise is Gaussian. Of course, in many situations the differences between the Kalman filter and the Kalman-Levy filter can be small, but there are some situations where the differences are significant. Given the widespread desire to reduce the risk of collateral damage, this fundamental incompatibility needs further consideration.

7 References

[1] F.A.Aliev, L.Ozbek, 'Evaluation of the convergence rate in the central limit theorem for the Kalman filter', *IEEE Transactions on Automatic Control*, vol.44, pp.1905-1909 (1999).

[2] D.Sornette, 'Critical Phenomena in Natural Sciences', Springer Series in Synergetics, Springer-Verlag, Heidelberg (2000).

[3] A.Le Breton, M.Musiela, 'A generalisation of the Kalman filter to models with infinite variance', *Stochastic Processes and Their Applications*, vol. 47, pp.75-94 (1993).

[4] H.Ahn, R.E.Feldman, 'Optimal filtering of a Gaussian signal in the presence of Levy noise', *SIAM Journal on Applied Mathematics*, vol.60, pp.359-369 (1999).

[5] D.Sornette, K.Ide, 'The Kalman-Levy filter' *Physica D*, vol. 151 pp. 142-174 (2002).

[6]. P.Zarchan, 'Tactical and Strategic Missile Guidance, 3rd Ed.', Progress in Astronautics and Aeronautics Vol.176 (AIAA, 1997).

[7]. J.F.Ralph, E.M.Januarius, M.I.Smith, K.L.Edwards, M.Bernhardt, 'Performance limits for multi-platform scene-referenced navigation systems', 'Sensor Fusion: Architectures, Algorithms and Applications V' Ed. B.B.Dasarthy, SPIE Press, Vol. 4385, pp.292; J.F.Ralph, M.I.Smith, M.Bernhardt, C.E.West, C.R.Angell, S.W.Sims, 'Distributed Air-to-Ground Targeting', 'Sensor Fusion: Architectures, Algorithms and Applications VI' Ed. B.B.Dasarthy, SPIE Press, Vol. 4731, pp.216.

[8] G.Samorodnitsky, M.S.Taqqu, 'Stable non-Gaussian random processes', Chapman & Hall, New York (1994).

[9] K.L.Edwards, S.J.Lloyd, J.F.Ralph, 'Air-to-ground weapon aiming – a synopsis and a look to the future' (The Aeronautical Journal 105, 53 (2001)).

[10] E.g. Papers contained in: F.A.Sadjadi (Ed.) 'Selected Papers on Sensor and Data Fusion', SPIE Milestone Series vol. MS124, (1996).

[11] P.R.C.Collins, P.Greenway, D.Edmondson, M.A.Green, 'Targeting for Future Weapon Systems', Proceedings of AGARD MSP Symposium (CP-576) on 'Technologies for Precision Air Strike Operations in Rapid-Reaction and Localised-Conflict Scenarios', Seville, Spain, October 1995.