Filtering of the ARMAX process with generalized t -distribution noise : the influence function approach

Ho, Weng Khuen; Ling, Keck Voon; Vu, Hoang Dung; Wang, Xiaoqiong

2014

Ho, W. K., Ling, K. V., Vu, H. D., & Wang, X. (2014). Filtering of the ARMAX Process with Generalized t -Distribution Noise: The Influence Function Approach . Industrial & Engineering Chemistry Research, 53(17), 7019-7028.

<https://hdl.handle.net/10356/103383>

<https://doi.org/10.1021/ie401990x>

© 2014 American Chemical Society. This is the author created version of a work that has been peer reviewed and accepted for publication by Industrial & Engineering Chemistry Research, American Chemical Society. It incorporates referee' s comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [DOI: <http://dx.doi.org/10.1021/ie401990x>].

Downloaded on 26 Apr 2023 20:19:13 SGT

Filtering of the ARMAX Process with Generalized t -Distribution Noise: The Influence Function Approach

Weng Khuen Ho,[†] Keck Voon Ling,[‡] Hoang Dung Vu,^{*,†} and Xiaoqiong Wang[†]

[†]National University of Singapore, Singapore 117576

[‡]Nanyang Technological University, Singapore 639798

ABSTRACT: The commonly made assumption of Gaussian noise is an approximation to reality. In this paper, influence function, an analysis tool in robust statistics, is used to formulate a recursive solution for the filtering of the ARMAX process with generalized t -distribution noise. By being a superset encompassing Gaussian, uniform, t , and double exponential distributions, generalized t -distribution has the flexibility of characterizing noise with Gaussian or non-Gaussian statistical properties. The filter is formulated as a maximum likelihood problem, but instead of solving the optimization problem numerically, influence function approximation is used to obtain a recursive solution to reduce the computational load and facilitate real-time implementation. The influence function equations derived are also useful in determining the variance of the filter and the impact of outliers.

1. INTRODUCTION

In the statistical analysis of time series, the autoregressive-moving-average with exogenous inputs model (ARMAX) with Gaussian noise is commonly used. However, the Gaussian noise assumption is an approximation to reality. The occurrence of outliers, transient data in steady-state measurements, instrument failure, human error, model nonlinearity, etc. can all induce non-Gaussian data.¹ Indeed whenever the central limit theorem is invoked, the central limit theorem being a limit theorem can at most suggest approximate normality for real data.² However, even high-quality model data may not fit the Gaussian distribution and the presence of a single outlier can spoil the statistical analysis completely for the case of least-squares estimation² including the Kalman filter.³

The generalized t -distribution (GT) was employed in the data reconciliation problem to model random noise.^{1,4,5} GT distribution was also used in econometrics^{6–9} to model random noise in the parameter estimation problem. By being a superset encompassing Gaussian, uniform, t and double exponential distributions, GT distribution has the flexibility to characterize noise with Gaussian or non-Gaussian statistical properties. The problem of estimation with GT noise was solved numerically using the Newton–Raphson or the expectation maximization algorithm.^{1,4–9} Unlike recursive algorithms such as the recursive least-squares estimator, it is not suitable for real-time applications.

In this paper, influence function (IF), an analysis tool in robust statistics,^{2,10} is used to formulate a recursive algorithm that gives an approximate solution making it suitable for real-time and online implementation. Specifically, the problem is formulated as the filtering of the ARMAX process with GT noise. Other well-known approaches^{11–13} for handling non-Gaussian noise include the approach of particle filters which is based on point mass or particle representation of probability densities.

The IF was used in ref 14 to analyze parameter estimation with GT noise. Instead of using the IF as an analysis tool to analyze a given estimator, this paper makes use of the IF to

synthesize or construct an estimator. The other difference is that, while ref 14 studied the estimation of the parameters in the transfer function, this paper estimates the states or output of the transfer function.

The main contribution of this paper is in sections 3 and 4, where we use IF approximation to derive a recursive solution for the maximum likelihood estimation of the ARMAX process with GT noise. We also show how the IF can be used to analyze the filter, specifically how it can predict the filter output due to outliers and the variance of the output. To put things in perspective, it will be shown through an example that if the noise is Gaussian then the proposed ARMAX filter is equivalent to the Kalman filter.¹⁵ Otherwise, the ARMAX filter has the extra degrees of freedom to model the noise.

2. MAXIMUM LIKELIHOOD ESTIMATION OF THE ARMAX PROCESS WITH GT NOISE

The ARMAX process and maximum likelihood estimation with GT distribution^{1,4–9} were already given in the literature. In this section we only give the equations necessary for the derivation of the recursive algorithm using IF approximation in section 3.

2.1. ARMAX Process. Consider the single-input single-output ARMAX process:

$$A(q^{-1})y(k) = B(q^{-1})u(k) + C(q^{-1})\varepsilon(k) \quad (1)$$

where

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n}$$

$$B(q^{-1}) = b_1q^{-1} + b_2q^{-2} + \dots + b_{n_b}q^{-n_b}$$

$$C(q^{-1}) = 1 + c_1q^{-1} + \dots + c_nq^{-n}$$

Received: June 24, 2013

Revised: January 10, 2014

Accepted: March 29, 2014

Figure 1. Different choices of the GT distribution shape parameters p and q can give different well-known distributions.

$k = 1, \dots, N$ is the sampling instance, $n_B \leq n$, and q^{-1} is the backward shift operator, i.e., $q^{-1}y(k) = y(k-1)$. The polynomial C may be multiplied by an arbitrary power of q as this does not change the correlation structure of $C(q^{-1})$. This is used to normalize C so that $\deg C = \deg A = n$. The input and output are given by $u(k)$ and $y(k)$, respectively.

Let the noise $\varepsilon(k)$ be modeled by the zero-mean GT probability density function⁶

$$f(\varepsilon) = \frac{p}{2\sigma q^{1/p} \beta(1/p, q) \left(1 + \frac{|\varepsilon|^p}{q\sigma^p}\right)^{q+1/p}} \quad (2)$$

where σ is the scale parameter and p and q are the shape parameters. The beta function is given by $\beta(a, b) = \int_0^1 z^{a-1} (1-z)^{b-1} dz$. By different choices of p and q , GT can represent a wide range of distributions.^{6,16} The relationships between GT, Gaussian, uniform, t , and double exponential distributions are shown in Figure 1.

2.2. Diophantine Equation. The Diophantine equation^{17–19} or identity can be used to isolate the noise term in the ARMAX process. The Diophantine equation^{17–19} is given as

$$C(q^{-1}) = E(q^{-1})A(q^{-1}) + q^{-j}F(q^{-1}) \quad (3)$$

where $C(q^{-1})$ is asymptotically stable and

$$E(q^{-1}) = 1 + e_1 q^{-1} + \dots + e_{j-1} q^{-(j-1)}$$

$$F(q^{-1}) = f_0 + f_1 q^{-1} + \dots + f_{n_F} q^{-n_F}$$

$$n_F = n - 1$$

Using eq 3 for $j = 1$, eq 1 becomes

$$y(k+1) = \frac{F(q^{-1})}{C(q^{-1})}y(k) + \frac{B(q^{-1})}{C(q^{-1})}u(k+1) + \varepsilon(k+1) \quad (4)$$

Multiplying by q^{-1} , the current measurement can be obtained from eq 4 as

$$y(k) = \frac{F(q^{-1})}{C(q^{-1})}y(k-1) + \frac{B(q^{-1})}{C(q^{-1})}u(k) + \varepsilon(k) \quad (5)$$

As it was found to be more convenient to work in the state space, expressing eq 5 in the state-space form gives

$$x(k+1) = \Phi x(k) + \Gamma u(k) + \Omega y(k) \quad (6)$$

$$y(k) = Hx(k) + \varepsilon(k) \quad (7)$$

where

$$\Phi = \begin{bmatrix} -c_1 & 1 & 0 & \dots & \dots & 0 \\ -c_2 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -c_{n-1} & 0 & \dots & \dots & 0 & 1 \\ -c_n & 0 & \dots & \dots & \dots & 0 \end{bmatrix}$$

$$\Gamma = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{N_B} \\ 0 \end{bmatrix}$$

$$\Omega = \begin{bmatrix} c_1 - a_1 \\ c_2 - a_2 \\ \vdots \\ c_n - a_n \end{bmatrix}$$

$$H = [1 \ 0 \ \dots \ 0]$$

Iterating from the initial value $x(1)$, eqs 6 and 7 give

$$x(2) = \Phi x(1) + \Gamma u(1) + \Omega y(1)$$

$$x(3) = \Phi^2 x(1) + \Phi \Gamma u(1) + \Gamma u(2) + \Phi \Omega y(1) + \Omega y(2)$$

$\vdots$

$$x(N) = \Phi^{N-1} x(1) + \bar{x}(N) \quad (8)$$

$$y(N) = H \Phi^{N-1} x(1) + H \bar{x}(N) + \varepsilon(N) \quad (9)$$

where

$$\bar{x}(N) = \sum_{k=1}^{N-1} \Phi^{k-1} \Gamma u(N-k) + \sum_{k=1}^{N-1} \Phi^{k-1} \Omega y(N-k) \quad (10)$$

2.3. Maximum Likelihood Estimation. Given N measurements $y(k)$, $k = 1, \dots, N$, the initial condition, $x(1)$, can be estimated using eq 9 in the minimization of the following maximum likelihood cost function.

$$J = -\sum_{k=1}^N \ln f(\varepsilon(k))$$

$$= -\sum_{k=1}^N \ln f(y(k) - H\Phi^{k-1}x(1) - H\bar{x}(k))$$

This can be done by differentiating with respect to $x(1)$.

$$\frac{\partial J}{\partial x(1)} = \psi(\varepsilon) = -(pq + 1) \sum_{k=1}^N (H\Phi^{k-1})^T \frac{\varepsilon(k)|\varepsilon(k)|^{p-2}}{q\sigma^p - |\varepsilon(k)|^p} \quad (11)$$

where $p > 1$ and setting

$$\psi(\varepsilon) = 0 \quad (12)$$

Equation 12 can be solved for $x(1)$ numerically using the Newton–Raphson or the expectation maximization algorithm.²⁰ Unlike recursive algorithms such as the recursive least-squares estimator, eq 12 is not suitable for real-time applications. For example, in real-time control, the information is used by the controller to calculate the control signal for the next sampling instance. The number of iterations required by eq 12 to converge to a solution can be different for different samples; hence there is no guarantee that the information is available before the next sampling instance.

3. INFLUENCE FUNCTION APPROXIMATION

In this section, we introduce the influence function to approximate and solve eq 12 recursively.

Consider the function $x = f(h)$. The first-order Taylor series expansion

$$x = \left. \frac{dx}{dh} \right|_{h=0} h$$

makes use of the gradient $(dx/dh)|_{h=0}$ to give the approximate value of x at h . Consider $\hat{x}(1)$, the asymptotic value of the estimate of $x(1)$. Let $\hat{x}(1)$ be associated with the probability density function of $(1-h)f(\varepsilon) + h\delta(\varepsilon)$. Likewise the Taylor series expansion

$$\hat{x}(1) = \left. \frac{d\hat{x}(1)}{dh} \right|_{h=0} h \quad (13)$$

makes use of the gradient $(d\hat{x}(1)/dh)|_{h=0}$ to give the approximate value of $\hat{x}(1)$ at h . The gradient term in eq 13 known as the influence function (IF) is defined in refs 2 and 10 as

$$\begin{aligned} \text{IF}(\varepsilon) &= \left. \frac{\partial \hat{x}(1)}{\partial h} \right|_{h=0} \\ &= - \left(\int_{-\infty}^{\infty} \frac{\partial \psi(\varepsilon)}{\partial \hat{x}(1)} f(\varepsilon) d\varepsilon \right)^{-1} \psi(\varepsilon) \Big|_{x(1)=0} \end{aligned} \quad (14)$$

where

$$\begin{aligned} \frac{\partial \psi(\varepsilon)}{\partial \hat{x}(1)} &= \sum_{k=1}^N (H\Phi^{k-1})^T H\Phi^{k-1} \\ &\quad \frac{(pq + 1)[(p-1)q\sigma^p - |\varepsilon(k)|^p]|\varepsilon(k)|^{p-2}}{(q\sigma^p + |\varepsilon(k)|^p)^2} \end{aligned}$$

Derivation of eq 14 is given in the Appendix. When $h = 0$, the associated probability density function of $\hat{x}(1)$ is $f(\varepsilon)$ and the usual assumption of zero initial condition for the ARMAX transfer function is made i.e. $x(1) = 0$.

3.1. Recursive Algorithm. The solution for $\hat{x}(1)$ can be written in the form of a recursive algorithm. Substituting eqs 11, 14 and $h = 1$ into eq 13 gives

$$\begin{aligned} \hat{x}(1|N) &= \text{IF}(\varepsilon) \\ &= \left(\sum_{k=1}^N (H\Phi^{k-1})^T H\Phi^{k-1} \right)^{-1} \left(\sum_{k=1}^N (H\Phi^{k-1})^T z(k) \right) \end{aligned} \quad (15)$$

where

$$\begin{aligned} z(k) &= \left(\int_{-\infty}^{+\infty} \frac{[(p-1)q\sigma^p - |\varepsilon(k)|^p]|\varepsilon(k)|^{p-2}}{(q\sigma^p + |\varepsilon(k)|^p)^2} f(\varepsilon) d\varepsilon \right)^{-1} \\ &\quad \frac{\varepsilon(k)|\varepsilon(k)|^{p-2}}{q\sigma^p + |\varepsilon(k)|^p} \Big|_{x(1)=0} \end{aligned} \quad (16)$$

and $\hat{x}(1|N)$ denotes the estimate of $x(1)$ at sample N .

Notice that eq 15 gives the well-known least-squares estimates $\hat{x}(1|N)$ from the minimization of the least-squares loss function

$$V = \frac{1}{2} \sum_{k=1}^N (z(k) - H\Phi^{k-1}\hat{x}(1|N))^2$$

and the recursive version in eqs 18 and 19 with the covariance matrix

$$P(1|N) = \left(\sum_{k=1}^N (H\Phi^{k-1})^T H\Phi^{k-1} \right)^{-1} \quad (17)$$

are given in many textbooks that discuss least squares.¹⁷ Equations 6 and 9 are then used to obtain $\bar{x}(N)$ and $\hat{y}(N|N)$ in eqs 20 and 21, respectively.

The derivation is complete and the recursive ARMAX filter algorithm for $N = 1, 2, 3, \dots$ is summarized in eqs 18–21.

ARMAX filter:

$$\begin{aligned} P(1|N) &= P(1|N-1) \\ &\quad - \frac{P(1|N-1)(H\Phi^{N-1})^T H\Phi^{N-1}P(1|N-1)}{1 + H\Phi^{N-1}P(1|N-1)(H\Phi^{N-1})^T} \end{aligned} \quad (18)$$

$$\begin{aligned} \hat{x}(1|N) &= \hat{x}(1|N-1) + P(1|N)(H\Phi^{N-1})^T \\ &\quad [z(N) - H\Phi^{N-1}\hat{x}(1|N-1)] \end{aligned} \quad (19)$$

$$\bar{x}(N+1) = \Phi\bar{x}(N) + \Gamma u(N) + \Omega y(N) \quad (20)$$

$$\hat{y}(N|N) = H\Phi^{N-1}\hat{x}(1|N) + H\bar{x}(N) \quad (21)$$

The covariance of $\bar{x}(1)$ and estimate $\hat{y}(N)$ at sample N are denoted by $P(1|N)$ and $\hat{y}(N|N)$, respectively. For initialization, $P(1|0)$ can be set as an identity matrix multiplied by some large number and $x(1) = \bar{x}(1) = 0$.

3.2. Mean, Variance, and Outlier. Let the actual noise be associated with probability density function $g(\varepsilon)$ which is not necessarily equal to $f(\varepsilon)$, the noise model used in the design of

Table 1. Parameters of the ARMAX Process and ARMAX Filter in the Examples

	example	1	2	3	4
	Figure		2a	3	8
ARMAX model	A	$1 + aq^{-1}$	$1 + aq^{-1}, a = -0.9$	$1 + aq^{-1}, a = -0.6$	$1 + aq^{-1}, a = -0.987$
	B	bq^{-1}	$bq^{-1}, b = 0.1$	$bq^{-1}, b = 0.4$	$bq^{-1}, b = 0.037$
	C	$1 + cq^{-1}$	$1 + cq^{-1}, c = a$	$1 + cq^{-1}, c = -0.8$	$1 + cq^{-1}, c = a$
ARMAX Filter	$f(\varepsilon)$	$N(0, \alpha)$	$t_3(0, 0.1)$	$t_3(0, 0.1)$	eq 2
	Φ	$-c$	$-c$	$-c$	$-c$
	H	1	1	1	1
	Γ	b	b	b	b
	Ω	$c - a$	$c - a$	$c - a$	$c - a$
	p	2	2	2	2
	q	∞	1.5	1.5	3.433
	σ	$\alpha\sqrt{2}$	$0.1\sqrt{2}$	$0.1\sqrt{2}$	0.1636
	actual noise	$g(\varepsilon)$	$f(\varepsilon)$	$\begin{cases} \delta(\varepsilon_1) & k = k_1 = 2; \quad \varepsilon_1 = -1 \\ f(\varepsilon) & k \neq k_1 \end{cases}$	$f(\varepsilon)$

the filter. The mean is then given by the following first-order von Mises expansion^{21,22}

$$\hat{x}(1|N) = \int_{-\infty}^{\infty} \text{IF}(\varepsilon) g(\varepsilon) d\varepsilon \quad (22)$$

and the variance

$$\text{var } \hat{x}(1|N) = \int_{-\infty}^{\infty} \text{IF}^T(\varepsilon) \text{IF}(\varepsilon) g(\varepsilon) d\varepsilon \quad (23)$$

The assumption that $g(\varepsilon)$ is the same for all k is commonly made. Here we extend to the case where $g(\varepsilon)$ could be different for different samples k denoted by $g_k(\varepsilon)$. The case where $g(\varepsilon)$ could be different for different samples k is useful for the analysis of outliers (see section 4.3). Hence, instead of integrating $\text{IF}(\varepsilon)$ with $g(\varepsilon)$ in eq 22, we first substitute eq 11 into eq 14 and then integrate $\text{IF}(\varepsilon)$ with different $g_k(\varepsilon)$ for different k giving

$$\hat{x}(1|N) = \left(\int_{-\infty}^{\infty} \frac{\partial \psi(\varepsilon)}{\partial \hat{x}(1)} f(\varepsilon) d\varepsilon \right)^{-1} (pq + 1) \sum_{k=1}^N (H\Phi^{k-1})^T \int_{-\infty}^{\infty} \frac{\varepsilon(k)|\varepsilon(k)|^{p-2}}{q\sigma^p + |\varepsilon(k)|^p} g_k(\varepsilon) d\varepsilon \Big|_{x(1)=0} \quad (24)$$

In section 4 examples will be given to illustrate the properties of the ARMAX filter such as the equivalence to the Kalman filter if we design with Gaussian noise in mind by choosing $p = 2$, $q = \infty$ (see section 4.1), and the variance of the filter output (see sections 4.2 and 4.4).

4. EXAMPLES

Four examples are given to illustrate the properties of the ARMAX filter and the IF analysis. For easy reference, the parameters of the ARMAX process and ARMAX filter are summarized in Table 1. The parameters Φ , H , Γ , and Ω depend on A , B , and C of the ARMAX process. In examples 1, 2, and 3 the parameters p , q , and σ are chosen according to Figure 1 to give $f(\varepsilon)$. In example 4, p , q , and σ are obtained by fitting the GT distribution of eq 2 to the experimental data. The ARMAX filter is designed with $p = 2$, $q = \infty$ for Gaussian noise in example 1, $p = 2$, $q = 1.5$ for t_3 noise in examples 2 and 3, and $p = 2$, $q = 3.433$ for the noise in example 4. Note that the distribution $f(\varepsilon)$ used for the filter design need not be the same

as $g(\varepsilon)$, the distribution of the actual noise in the last row of Table 1. One thousand simulation runs were conducted in examples 2 and 3, and 100 experimental runs were conducted in example 4 to give the variance of the estimate. The simulation is started with $P(1|0) = 1000I$ and $\hat{x}(1|0) = 0$.

4.1. Example 1: Kalman Filter Connection. This example shows that if the ARMAX filter is designed with Gaussian noise in mind then it is equivalent to the Kalman filter although it was formulated through maximum likelihood estimation with GT noise and IF approximation.

4.1.1. Kalman Filter. Consider the ARMAX process with Gaussian noise in example 1 of Table 1. A state-space representation is given by

$$\begin{aligned} x(k+1) &= -ax(k) + bu(k) + (c-a)\varepsilon(k) \\ y(k) &= x(k) + \varepsilon(k) \end{aligned} \quad (25)$$

The Kalman filter¹⁷ for the above state-space model of eq 25 is given in eqs 26–29.

Kalman filter for first-order ARMAX process:

$$\begin{aligned} \hat{x}(N|N) &= \hat{x}(N|N-1) + \frac{p(N|N-1)}{1 + p(N|N-1)} \\ &\quad [y(N) - \hat{x}(N|N-1)] \end{aligned} \quad (26)$$

$$\begin{aligned} \hat{x}(N+1|N) &= bu(N) - ay(N) + \frac{c}{1 + p(N|N-1)} \\ &\quad [y(N) - \hat{x}(N|N-1)] \end{aligned} \quad (27)$$

$$p(N+1|N) = \frac{p(N|N-1)c^2}{1 + p(N|N-1)} \quad (28)$$

$$\hat{y}(N|N) = \hat{x}(N|N) \quad (29)$$

4.1.2. ARMAX Filter. The ARMAX filter is designed for the ARMAX process with Gaussian noise of standard deviation α . According to Figure 1, the GT parameters to model the Gaussian noise are $p = 2$, $q = \infty$, and $\sigma = \alpha\sqrt{2}$ as shown in Table 1. Equations 16 and 9 give $z(N) = \varepsilon(N) = y(N) - \bar{x}(N)$, and eqs 18–21 give the ARMAX filter eqs 30–33.

Figure 2. Simulation results of example 2.

ARMAX filter for first-order ARMAX process with Gaussian noise:

$$p(1|N) = \frac{p(1|N-1)}{1 + p(1|N-1)c^{2(N-1)}} \quad (30)$$

$$\hat{x}(1|N) = \hat{x}(1|N-1) + p(1|N)(-c)^{N-1}[y(N) - \bar{x}(N) - (-c)^{N-1}\hat{x}(1|N-1)] \quad (31)$$

$$\bar{x}(N+1) = -c\bar{x}(N) + bu(N) + (c-a)y(N) \quad (32)$$

$$\hat{y}(N|N) = (-c)^{N-1}\bar{x}(1|N) + \bar{x}(N) \quad (33)$$

where from eq 17

$$p(1|N) = \left(\sum_{k=1}^N c^{2(k-1)} \right)^{-1} \quad (34)$$

4.1.3. The Connection. To connect the ARMAX filter with the Kalman filter, we will now show that the Kalman filter equations, eqs 26–29, can be obtained from the ARMAX filter equations, eqs 30–33.

Multiplying eq 31 by $(-c)^{N-1}$ and then adding $\bar{x}(N)$ to both sides of the equation gives

$$\begin{aligned} & (-c)^{N-1}\hat{x}(1|N) + \bar{x}(N) \\ &= (-c)^{N-1}\hat{x}(1|N-1) + \bar{x}(N) + p(1|N)(-c)^{2(N-1)} \\ & \quad [y(N) - \bar{x}(N) - (-c)^{N-1}\hat{x}(1|N-1)] \end{aligned} \quad (35)$$

Multiplying eq 31 by $(-c)^N$ and then adding eq 32 gives

$$\begin{aligned} & (-c)^N\hat{x}(1|N) + \bar{x}(N+1) \\ &= (-c)^N\hat{x}(1|N-1) + p(1|N)(-c)^{2N-1} \\ & \quad [y(N) - \bar{x}(N) - (-c)^{N-1}\hat{x}(1|N-1)] - c\bar{x}(N) \\ & \quad + bu(N) + (c-a)y(N) \end{aligned} \quad (36)$$

Note that, from eq 8, $x(N) = (-c)^{N-1}x(1) + \bar{x}(N)$, so eqs 33, 35, and 36 can be written as

$$\hat{y}(N|N) = \hat{x}(N|N) \quad (37)$$

$$\begin{aligned} \hat{x}(N|N) &= \hat{x}(N|N-1) + p(1|N) \\ & \quad c^{2(N-1)}[y(N) - \hat{x}(N|N-1)] \end{aligned} \quad (38)$$

$$\begin{aligned} \hat{x}(N+1|N) &= bu(N) - ay(N) + [c + p(1|N)(-c)^{2N-1}] \\ & \quad [y(N) - \hat{x}(N|N-1)] \end{aligned} \quad (39)$$

Substitute $\hat{x}(1|N) = (\hat{x}(N+1|N) - \bar{x}(N+1))/(-c)^N$ from eq 8 into eq 15 to give

$$\begin{aligned} & \hat{x}(N+1|N) - \bar{x}(N+1) \\ &= \left(\frac{1}{c^{2N}} \sum_{k=1}^N c^{2(k-1)} \right)^{-1} \\ & \quad \left(\frac{1}{(-c)^N} \sum_{k=1}^N (-c)^{k-1} z(k) \right) \Bigg|_{x(1)=0} \end{aligned} \quad (40)$$

and corresponding to eq 17 the covariance matrix

$$P(N+1|N) = \left(\frac{1}{c^{2N}} \sum_{k=1}^N c^{2(k-1)} \right)^{-1} \quad (41)$$

Using eq 34, eq 41 becomes

$$p(1|N) = \frac{p(N+1|N)}{c^{2N}} \quad (42)$$

Substituting eq 42 into eqs 38, 39, 30, and 37 gives the Kalman filter equations, eqs 26, 27, 28, and 29, respectively. For simplicity, we have used the first-order ARMAX process as an example. It can be shown that, in general, the ARMAX filter is equivalent to the Kalman filter if the GT parameters p and q in the ARMAX filter design are chosen as 2 and ∞ respectively to model Gaussian noise.

4.2. Example 2: Variance. In this example, the ARMAX filter is designed for the ARMAX process with t_3 noise. According to Figure 1, to model the t_3 noise, the GT parameters are $p = 2$, $q = 1.5$, and $\sigma = 0.1\sqrt{2}$ as shown in Table 1.

4.2.1. Simulation. We conducted 1000 simulation runs using the ARMAX filter (eqs 18–21) and Kalman filter (eqs 26–29). The results are shown in Figure 2, where the mean value for the 1000 runs is given by the white curve. The mean and variance

Table 2. Mean and Variance of $\hat{y}(N)$ in Figure 2

N	ARMAX filter					Kalman filter		
	eqs 12 and 21		eqs 18–21		eq 45	eqs 26–29		eq 49
	numerical solution		recursive solution		var (10^{-2})	recursive solution		var (10^{-2})
	mean	var (10^{-2})	mean	var (10^{-2})		mean	var (10^{-2})	
1	0.00	3.062	0.00	1.460	1.500	0.00	3.062	3.000
5	0.34	0.268	0.34	0.190	0.188	0.34	0.399	0.377
10	0.61	0.051	0.61	0.049	0.049	0.61	0.102	0.097
15	0.77	0.016	0.77	0.015	0.016	0.77	0.032	0.031
20	0.86	0.006	0.86	0.005	0.005	0.86	0.011	0.011

at $N = 1, 5, \dots, 20$ are tabulated in Table 2 under the column “eqs 18–21” and “eqs 26–29”. The result of solving eq 12 numerically for $\hat{x}(1|N)$ and then $\hat{y}(N|N)$ from eq 21 is also given under the column “eqs 12 and 21”. Figure 2 and Table 2 show clearly that the variance from the Kalman filter is larger than that from the ARMAX filter. The Kalman filter assumes Gaussian and not t_3 noise. This example shows that the GT parameters in the ARMAX filter can be chosen gainfully to give smaller variance.

4.2.2. IF Analysis: ARMAX Filter. The IF can be used to derive an equation to calculate the variance in Table 1. Using eq 15

$$\text{IF}(\varepsilon) = \left(\sum_{k=1}^N c^{2(k-1)-1} \left(\int_{-\infty}^{+\infty} \frac{0.03 - \varepsilon^2}{(0.03 + \varepsilon^2)^2} f(\varepsilon) d\varepsilon \right) \right)^{-1} \left(\sum_{k=1}^N \frac{(-c)^{k-1} \varepsilon(k)}{0.03 + \varepsilon(k)^2} \right)$$

Using eq 23

$$\begin{aligned} \text{var } \hat{x}(1|N) &= \left(\sum_{k=1}^N c^{2(k-1)-1} \left(\int_{-\infty}^{+\infty} \frac{0.03 - \varepsilon^2}{(0.03 + \varepsilon^2)^2} f(\varepsilon) d\varepsilon \right) \right)^{-2} \\ &\quad \left(\int_{-\infty}^{+\infty} \frac{\varepsilon^2}{(0.03 + \varepsilon^2)^2} g(\varepsilon) d\varepsilon \right) \\ &= \frac{3}{200} \frac{1 - c^{2N}}{1 - c^{2N}} \end{aligned} \quad (43)$$

where $\int_{-\infty}^{+\infty} ((0.03 - \varepsilon^2)/(0.03 + \varepsilon^2)^2) f(\varepsilon) d\varepsilon = 50/3$, $\int_{-\infty}^{+\infty} (\varepsilon^2/(0.03 + \varepsilon^2)^2) g(\varepsilon) d\varepsilon = (25/6)$, and, since ε is assumed to be a zero-mean independent random variable, $\int_{-\infty}^{+\infty} \varepsilon(j) \varepsilon(k)/((0.03 + \varepsilon(j)^2)(0.03 + \varepsilon(k)^2)) g(\varepsilon) d\varepsilon = 0$ for $j \neq k$. From eq 15 $\hat{x}(1|N)$ is zero mean as $\varepsilon(k)$ is zero mean and since $\bar{x}(N)$ is not a function of the random variable ε , eq 21 gives

$$\text{var } \hat{y}(N|N) = c^{2(N-1)} \text{var } \hat{x}(1|N) \quad (44)$$

Substituting eq 43 into eq 44 gives

$$\text{var } \hat{y}(N|N) = \frac{3}{200} \frac{c^{2(N-1)} - c^{2N}}{1 - c^{2N}} \quad (45)$$

Equation 45 is used to calculate the variance in the column “eq 45” and Table 2 shows that it is close to the values obtained from simulation in column “eqs 18–21”. Table 2 also shows that, for $N \geq 10$, the variances in the columns “eq 45” and “eqs 18–21” from IF approximation are close to the variance in

column “eqs 12 and 21” obtained by solving eq 12 numerically without approximation.

4.2.3. IF Analysis: Kalman Filter. Although the Kalman filter assumes Gaussian noise, the IF can also be used to derive an equation to calculate the variance of the Kalman filter when the actual noise is t_3 . Example 1 shows that the ARMAX filter designed with $p = 2$, $q = \infty$ is equivalent to the Kalman filter, so from eq 15

$$\text{IF}(\varepsilon) = \left(\sum_{k=1}^N c^{2(k-1)-1} \left(\sum_{k=1}^N (-c)^{k-1} \varepsilon(k) \right) \right) \quad (46)$$

Using eq 23

$$\text{var } \hat{x}(1|N) = \left(\sum_{k=1}^N c^{2(k-1)-1} \left(\int_{-\infty}^{+\infty} \varepsilon^2 g(\varepsilon) d\varepsilon \right) \right) \quad (47)$$

where $\int_{-\infty}^{+\infty} \varepsilon(j) \varepsilon(k) g(\varepsilon) d\varepsilon = 0$ for $j \neq k$ and $g(\varepsilon)$ is the GT probability density function of eq 2 with $p = 2$, $q = 1.5$, and $\sigma = 0.1\sqrt{2}$ to model the actual t_3 noise. This gives

$$\text{var } \hat{x}(1|N) = \frac{3}{100} \frac{1 - c^2}{1 - c^{2N}} \quad (48)$$

since $\int_{-\infty}^{+\infty} \varepsilon^2 g(\varepsilon) d\varepsilon = 3/100$. Using eq 48

$$\text{var } \hat{y}(N|N) = c^{2(N-1)} \text{var } \hat{x}(1|N) = \frac{3}{100} \frac{c^{2(N-1)} - c^{2N}}{1 - c^{2N}} \quad (49)$$

Equation 49 is used to calculate the variance in the column “eq 49” of Table 2. It is clear that it matched the variance from the simulation in the column “eqs 26–29”.

4.3. Example 3: Outlier. This example shows how the IF can be used to calculate the ARMAX and Kalman filter output in the presence of an outlier.

4.3.1. Simulation. Consider the ARMAX process in example 3 of Table 1. It has an outlier of $\varepsilon_1 = -1$ at $k = k_1 = 2$. Here $g(\varepsilon)$ will be different at each sample k and is given as

$$g_k(\varepsilon) = \begin{cases} \delta(\varepsilon_1) & k = k_1 \\ f(\varepsilon) & k \neq k_1 \end{cases} \quad (50)$$

where $\delta(\varepsilon_1)$ is an impulse at ε_1 to model the outlier of ε_1 at the sample $k = k_1$. The ARMAX filter is designed with $p = 2$, $q = 1.5$, and $\sigma = 0.1\sqrt{2}$ to model the t_3 noise. Unlike example 2, here $f(\varepsilon) \neq g(\varepsilon)$. The outputs $\hat{y}(N)$ of the ARMAX filter (eqs 18–21) and Kalman filter (eqs 26–29) are shown in Figures 3 and 4, respectively. The mean value of the 1000 runs is given by the yellow curve. It is clear that the Kalman filter output is greatly affected by the outlier at $k = 2$, unlike the ARMAX filter. It is known that a single outlier can spoil the statistical analysis

Figure 3. ARMAX filter output $\hat{y}(N)$. Figure 4. Kalman filter output $\hat{y}(N)$.

completely for the case of least-squares estimation² including the Kalman filter.³

4.3.2. IF Analysis. Equation 24 can be used to draw the yellow or mean curves in Figures 3 and 4. With $g_k(\varepsilon)$ of eq 50 and $p = 2$, eq 24 gives

$$\begin{aligned} \hat{x}(1|N) = & \left(\sum_{k=1}^N c^{2(k-1)-1} \int_{-\infty}^{+\infty} \frac{q\sigma^2 - \varepsilon^2}{(q\sigma^2 + \varepsilon^2)^2} f(\varepsilon) d\varepsilon \right)^{-1} \\ & \left[\int_{-\infty}^{+\infty} \frac{(-c)^{k_1-1} \varepsilon(k_1)}{q\sigma^2 + \varepsilon(k_1)^2} \delta(\varepsilon_1) d\varepsilon \right. \\ & \left. + \sum_{k=1, k \neq k_1}^N \int_{-\infty}^{+\infty} \frac{(-c)^{k-1} \varepsilon(k)}{q\sigma^2 + \varepsilon(k)^2} f(\varepsilon) d\varepsilon \right], \quad N \geq k_1 \end{aligned} \quad (51)$$

Note that the second term in the square brackets of eq 51 is zero because the expectation of $\varepsilon(k)$ for $k \neq k_1$ is zero. Simplifying gives

$$\begin{aligned} \hat{x}(1|N) = & \frac{(-c)^{k_1-1} (1 - c^2)}{1 - c^{2N}} \left(\frac{\varepsilon_1}{q\sigma^2 + \varepsilon_1^2} \right) \\ & \left(\int_{-\infty}^{+\infty} \frac{q\sigma^2 - \varepsilon^2}{(q\sigma^2 + \varepsilon^2)^2} f(\varepsilon) d\varepsilon \right)^{-1}, \quad N \geq k_1 \end{aligned} \quad (52)$$

From eq 21

$$\begin{aligned} \hat{y}(N|N) = & \frac{(-c)^{N+k_1-2} (1 - c^2)}{1 - c^{2N}} \left(\frac{\varepsilon_1}{q\sigma^2 + \varepsilon_1^2} \right) \\ & \left(\int_{-\infty}^{+\infty} \frac{q\sigma^2 - \varepsilon^2}{(q\sigma^2 + \varepsilon^2)^2} f(\varepsilon) d\varepsilon \right)^{-1} + \bar{x}(N), \quad N \geq k_1 \end{aligned} \quad (53)$$

Note that, according to Figure 1, the t_3 and Gaussian distribution are modeled by setting $q = 1.5$ and $q = \infty$, respectively. Therefore, the yellow curve in Figure 3 is obtained by substituting $q = 1.5$ and $\sigma = 0.1\sqrt{2}$ into eq 53 to give

$$\begin{aligned} \hat{y}(N|N) = & \frac{(-c)^{N+k_1-2} (1 - c^2)}{300(1 - c^{2N})} \left(\frac{\varepsilon_1}{0.03 + \varepsilon_1^2} \right) + \bar{x}(N), \\ N \geq & k_1 \end{aligned}$$

where $\int_{-\infty}^{+\infty} (q\sigma^2 - \varepsilon^2)/(q\sigma^2 + \varepsilon^2)^2 f(\varepsilon) d\varepsilon = 300$. The yellow curve in Figure 4 is obtained by substituting $q = \infty$ into eq 53 to give

$$\hat{y}(N|N) = \frac{(-c)^{N+k_1-2} (1 - c^2)}{1 - c^{2N}} \varepsilon_1 + \bar{x}(N), \quad N \geq k_1$$

4.4. Example 4: Liquid Level Estimation Experiment.

Consider the liquid level estimation problem commonly encountered in chemical processes in the coupled tank of Figure 5. The transfer function between the liquid level in tank 1, $y(k)$, and the control voltage, $u(k)$, at a sampling interval of 1 s is given as

$$y(k) = \frac{0.037q^{-1}}{1 - 0.987q^{-1}} u(k) + \varepsilon(k) \quad (54)$$

Figure 5. Photo of the coupled tank.

The polynomials A , B , and C in the ARMAX model can be obtained by comparing eqs 54 and 1 and are given in Table 1 in the column for example 4.

4.4.1. Experiment. One thousand measurements of the liquid level $y(k)$ were collected as shown in Figure 6 when the control voltage $u(k)$ was held constant at 2 V. The histogram of the measurements $y(k)$ after subtracting the mean are plotted in Figure 7 and is considered as the noise, $\varepsilon(k)$, distribution.

The maximum likelihood criterion can be used to find the parameters of the GT probability density function. In this

Figure 6. Measurement $y(N)$ for the liquid level estimation experiment.

Figure 7. Maximum likelihood criterion was used to fit a GT distribution (solid line) and Gaussian distribution (dashed line) to the noise distribution.

example we fixed $p = 2$ and then used the maximum likelihood criterion to find the other parameters q and σ of the GT probability density function $f(\varepsilon)$ of eq 2 by maximizing the objective function

$$J_f = \sum_{k=1}^{1000} \ln \frac{p}{2\sigma q^{1/p} \beta(1/p, q) \left(1 + \frac{|\varepsilon(k)|^p}{q\sigma^p}\right)^{q+1/p}}$$

This gives $q = 3.433$ and $\sigma = 0.1636$, and the resultant GT distribution is superimposed on the distribution in Figure 7. Using maximum likelihood, a Gaussian distribution was also fitted to the histogram. It is evident in Figure 7 that the GT distribution gives a better fit. In this example, p was not obtained by maximizing the objective function J_f but was simply chosen as 2 because we compared several other values of this parameter and found it made little difference. References 1 and 4 give further detailed discussion on the choice and determination of p , q , and σ .

With the control voltage $u(k) = 2$ V, 1000 measurements were collected and divided into 100 runs of 10 measurements

each. For each run, the ARMAX filter (eqs 18–21) and Kalman filter (eqs 26–29) were used to estimate the liquid level $\hat{y}(k)$. The results are shown in Figures 8 and 9, and the variances are

Figure 8. ARMAX filter estimate $\hat{y}(N)$.

Figure 9. Kalman filter estimate $\hat{y}(N)$.

tabulated in Table 3 in the rows labeled “exptl value”. Figures 8 and 9 and Table 3 show that the variance from the ARMAX filter is about 10% smaller than that from the Kalman filter. This example shows that the GT parameters in the ARMAX filter can be chosen gainfully to give smaller variance.

4.4.2. IF Analysis for ARMAX Filter. The IF can be used to derive an equation to calculate the variance in Table 3. Using eq 15

$$IF(\varepsilon) = \left(\sum_{k=1}^N c^{2(k-1)-1} \left(\int_{-\infty}^{+\infty} \frac{0.0919 - \varepsilon^2}{(0.0919 + \varepsilon^2)^2} f(\varepsilon) d\varepsilon \right)^{-1} \right. \\ \left. \left(\sum_{k=1}^N \frac{(-c)^{k-1} \varepsilon(k)}{0.0919 + \varepsilon(k)^2} \right) \right)$$

Using eq 23

Table 3. Variance ($\times 10^{-3}$) of $\hat{y}(N)$ in Figures 8 and 9

	N	1	2	3	4	5	6	7	8	9	10
ARMAX filter	exptl value	17.7	8.2	4.8	4.1	3.1	2.5	2.2	1.8	1.7	1.5
(Figure 8)	eq 57	16.8	8.3	5.5	4.0	3.2	2.6	2.2	1.9	1.7	1.5
Kalman filter	exptl value	20.0	9.6	5.8	4.9	3.6	3.1	2.6	2.2	1.9	1.7
(Figure 9)	eq 61	18.7	9.2	6.1	4.5	3.6	2.9	2.5	2.1	1.9	1.7

$$\begin{aligned} \text{var } \hat{x}(1|N) &= \left(\sum_{k=1}^N c^{2(k-1)} \right)^{-1} \\ &\quad \left(\int_{-\infty}^{+\infty} \frac{0.0919 - \varepsilon^2}{(0.0919 + \varepsilon^2)^2} f(\varepsilon) d\varepsilon \right)^{-2} \\ &\quad \left(\int_{-\infty}^{+\infty} \frac{\varepsilon^2}{(0.0919 + \varepsilon^2)^2} g(\varepsilon) d\varepsilon \right) \\ &= 0.0168 \frac{1 - c^2}{1 - c^{2N}} \end{aligned} \quad (55)$$

where $\int_{-\infty}^{+\infty} ((0.0919 - \varepsilon^2)/(0.0919 + \varepsilon^2)^2) f(\varepsilon) d\varepsilon = 7.57$, $\int_{-\infty}^{+\infty} (\varepsilon^2/(0.0919 + \varepsilon^2)^2) g(\varepsilon) d\varepsilon = 0.963$, and, since ε is assumed to be a zero mean-independent random variable, $\int_{-\infty}^{+\infty} \varepsilon(j) \varepsilon(k) / ((0.0919 + \varepsilon(j)^2)(0.0919 + \varepsilon(k)^2)) g(\varepsilon) d\varepsilon = 0$ for $j \neq k$. From eq 15 $\hat{x}(1|N)$ is zero mean as $\varepsilon(k)$ is zero mean and since $\bar{x}(N)$ is not a function of the random variable ε , eq 21 gives

$$\text{var } \hat{y}(N|N) = c^{2(N-1)} \text{var } \hat{x}(1|N) \quad (56)$$

Substituting eq 55 into eq 56 gives

$$\text{var } \hat{y}(N|N) = 0.0168 \frac{c^{2(N-1)} - c^{2N}}{1 - c^{2N}} \quad (57)$$

Equation 57 is used to calculate the variance in row “eq 57” of Table 3. It is close to the experimental values given in the row just above.

4.4.3. IF Analysis for Kalman Filter. Although the Kalman filter assumes Gaussian noise, the IF can also be used to derive an equation to calculate the variance of the Kalman filter when the actual noise is not Gaussian. Example 1 shows that the ARMAX filter designed with $p = 2$, $q = \infty$ is equivalent to the Kalman filter, so from eq 15

$$\text{IF}(\varepsilon) = \left(\sum_{k=1}^N c^{2(k-1)} \right)^{-1} \left(\sum_{k=1}^N (-c)^{k-1} \varepsilon(k) \right) \quad (58)$$

Using eq 23

$$\text{var } \hat{x}(1|N) = \left(\sum_{k=1}^N c^{2(k-1)} \right)^{-1} \left(\int_{-\infty}^{\infty} \varepsilon^2 g(\varepsilon) d\varepsilon \right) \quad (59)$$

where $\int_{-\infty}^{\infty} \varepsilon(j) \varepsilon(k) g(\varepsilon) d\varepsilon = 0$ for $j \neq k$ and $g(\varepsilon)$ is the GT probability density function of eq 2 with $p = 2$, $q = 3.433$, and $\sigma = 0.1636$ to model the noise. This gives

$$\text{var } \hat{x}(1|N) = 0.0189 \frac{1 - c^2}{1 - c^{2N}} \quad (60)$$

since $\int_{-\infty}^{\infty} \varepsilon^2 g(\varepsilon) d\varepsilon = 0.0189$. Using eq 56

$$\text{var } \hat{y}(k|N) = c^{2(N-1)} \text{var } \hat{x}(1|N) = 0.0189 \frac{c^{2(N-1)} - c^{2N}}{1 - c^{2N}} \quad (61)$$

Equation 61 is used to calculate the variance in the row “eq 61” of Table 3. It is close to the experimental values given in the row just above.

The values in Table 3 show that the variance of the estimate from the ARMAX filter is about 10% smaller than the one from the Kalman filter. If the noise is non-Gaussian and can be modeled by the GT distribution, then the ARMAX filter with GT noise model can produce more accurate estimate of the process output $y(k)$ because of the more accurate noise model. The ARMAX filter can be used gainfully in control systems. For example, in adaptive control, the process output $y(k)$ is fed back to the adaptive controller which makes use of the information to adapt itself to meet performance criteria. The more accurate estimate of $y(k)$ from the ARMAX filter can then be used gainfully by the same adaptive controller to enhance performance.

4.4.4. Outlier Analysis. While collecting the measurements for example 4, an outlier was observed and this set of data shown in Figure 10 was taken out for further analysis. The

Figure 10. Outlier analysis. (x) measurement; (—) ARMAX filter estimate; (---) Kalman filter estimate.

occurrence of an outlier is clearly not restricted to liquid level measurements but can also be expected in measurements of all sorts, for instance, temperature and pressure measurements common in chemical engineering. Although the mean estimate for the ARMAX filter and the Kalman filter are both 6.1 cm as shown in Figures 8 and 9, it took only one outlier at $N = 2$ in Figure 10 to cause the estimate from the Kalman filter to deviate from 6.1 to 6.6 cm, and not returning to 6.1 cm even at $N = 10$. On the other hand, the ARMAX filter estimate was hardly affected by the outlier.

5. CONCLUSION

The IF is employed to give an approximate solution to the maximum likelihood estimation problem in the ARMAX filter. The solution is recursive, making it suitable for online and real-time implementation. We also used the IF to analyze the output of the filter designed with the GT noise model instead of the usual Gaussian noise model. Equations derived are useful in determining the variance of the estimates and the impact of outliers. If the noise is modeled by the Gaussian distribution, then the proposed filter reduces to the Kalman filter. Otherwise

the GT distribution has the extra degree of freedom to model non-Gaussian noise. If we do not know the distribution of the noise then we can use the Kalman filter, but if there is information then it can be used gainfully in the GT distribution framework to take into account the non-Gaussian noise.

■ APPENDIX: DERIVATION OF THE INFLUENCE FUNCTION

By taking expectation, eq 12 can be written as

$$\int_{-\infty}^{+\infty} \psi(\varepsilon) f(\varepsilon) d\varepsilon = 0 \quad (62)$$

To study the change in $\hat{x}(1)$ when the distribution changes from $f(\varepsilon)$ to a new distribution $f_1(\varepsilon)$, replace $f(\varepsilon)$ in eq 62 by $(1-h)f(\varepsilon) + hf_1(\varepsilon)$, where $0 \leq h \leq 1$, giving

$$\int_{-\infty}^{+\infty} \psi(\varepsilon)((1-h)f(\varepsilon) + hf_1(\varepsilon)) d\varepsilon = 0 \quad (63)$$

Differentiating eq 63 with respect to h gives

$$\begin{aligned} \frac{\partial}{\partial h} \left(\int_{-\infty}^{+\infty} \psi(\varepsilon)((1-h)f(\varepsilon) + hf_1(\varepsilon)) d\varepsilon \right) &= 0 \\ \int_{-\infty}^{+\infty} \psi(\varepsilon)(-f(\varepsilon) + f_1(\varepsilon)) d\varepsilon &+ \left(\int_{-\infty}^{+\infty} \frac{\partial \psi(\varepsilon)}{\partial \hat{x}(1)} ((1-h)f(\varepsilon) + hf_1(\varepsilon)) d\varepsilon \right) \frac{\partial \hat{x}(1)}{\partial h} = 0 \end{aligned} \quad (64)$$

Let $h = 0$ and using eq 62, eq 64 reduces to

$$\left. \frac{\partial \hat{x}(1)}{\partial h} \right|_{h=0} = - \left(\int_{-\infty}^{+\infty} \frac{\partial \psi(\varepsilon)}{\partial \hat{x}(1)} f(\varepsilon) d\varepsilon \right)^{-1} \int_{-\infty}^{+\infty} \psi(\varepsilon)(f_1(\varepsilon)) d\varepsilon \Big|_{x(1)=0} \quad (65)$$

When $h = 0$, the associated probability density function of $\hat{x}(1)$ is $f(\varepsilon)$ and the usual assumption of zero initial condition for the ARMAX transfer function is made, i.e., $x(1) = 0$. Let $f_1(\varepsilon) = \delta(\varepsilon)$ an impulse function at ε and eq 65 reduces to the influence function

$$\begin{aligned} \text{IF}(\varepsilon) &= \left. \frac{\partial \hat{x}(1)}{\partial h} \right|_{h=0} \\ &= - \left(\int_{-\infty}^{+\infty} \frac{\partial \psi(\varepsilon)}{\partial \hat{x}(1)} f(\varepsilon) d\varepsilon \right)^{-1} \psi(\varepsilon) \Big|_{x(1)=0} \end{aligned}$$

which is eq 14.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: hoangdung@nus.edu.sg.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research is funded by the Republic of Singapore National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in

the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore.

■ REFERENCES

- (1) Wang, D.; Romagnoli, J. A. A Framework for Robust Data Reconciliation Based on a Generalized Objective Function. *Ind. Eng. Chem. Res.* **2003**, *42*, 3075–3084.
- (2) Hampel, F. R.; Ronchetti, E. M.; Rousseeuw, P. J.; Stahel, W. A. *Robust Statistics: The Approach Based on Influence Functions*; Wiley: New York, 1986; p 536.
- (3) Sorenson, H. W. Least-squares estimation: from Gauss to Kalman. *IEEE Spectrum* **1970**, *7*, 63–68.
- (4) Wang, D.; Romagnoli, J. Generalized T distribution and its applications to process data reconciliation and process monitoring. *Trans. Inst. Meas. Control* **2005**, *27*, 367–390.
- (5) Yan, H.; Ho, W. K.; Ling, K. V.; Lim, K. W. Multi-Zone Thermal Processing in Semiconductor Manufacturing: Bias Estimation. *IEEE Trans. Ind. Inf.* **2010**, *6*, 216–228.
- (6) McDonald, J. B.; Newey, W. K. Partially adaptive estimation of regression models via the generalized t distribution. *Econometric Theory* **1988**, *4*, 428–457.
- (7) Theodossiou, P. Financial Data and the Skewed Generalized T Distribution. *Manage. Sci.* **1998**, *44*, 1650–1661.
- (8) Bali, T. G.; Theodossiou, P. A conditional-SGT-VaR approach with alternative GARCH models. *Ann. Oper. Res.* **2006**, *151*, 241–267.
- (9) Hansen, C.; McDonald, J. B.; Newey, W. K. Instrumental Variables Estimation With Flexible Distributions. *J. Bus. Econ. Stat.* **2010**, *28*, 13–25.
- (10) Hampel, F. R. The Influence Curve and its Role in Robust Estimation. *J. Am. Stat. Assoc.* **1974**, *69*, 383–393.
- (11) Kitagawa, G. A Self-Organizing State-Space Model. *J. Am. Stat. Assoc.* **1998**, *93*, 1203–1215.
- (12) Arulampalam, M.; Maskell, S.; Gordon, N.; Clapp, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **2002**, *50*, 174–188.
- (13) Crisan, D.; Rozovskii, B. *The Oxford Handbook of Nonlinear Filtering*; Oxford University Press: Oxford, U.K., 2011.
- (14) Ho, W. K.; Vu, H. D.; Ling, K. V. Influence Function Analysis of Parameter Estimation with Generalized t Distribution Noise Model. *Ind. Eng. Chem. Res.* **2013**, *52*, 4168–4177.
- (15) Kalman, R. A new approach to linear filtering and prediction problems. *J. Basic Eng.* **1960**, *82*, 35–45.
- (16) McDonald, J. B. Partially adaptive estimation of ARMA time series models. *Int. J. Forecast.* **1989**, *5*, 217–230.
- (17) Åström, K. J.; Wittenmark, B. *Computer-Controlled Systems*, 3rd ed.; Prentice Hall, Inc.: Upper Saddle River, NJ, USA, 1997.
- (18) Landau, I. D.; Lozano, R.; M'Saad, M.; Karimi, A. *Adaptive Control*; Communications and Control Engineering; Springer: London, 2011.
- (19) Camacho, E.; Bordons, C. *Model Predictive Control*; Springer Verlag: Berlin, 1998.
- (20) McLachlan, G. J.; Krishnan, T. *The EM Algorithm and Extensions*, 2nd ed.; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2008.
- (21) von Mises, R. On the Asymptotic Distribution of Differentiable Statistical Functions. *Ann. Math. Stat.* **1949**, *18*, 309–348.
- (22) Turrin Fernholz, L. On multivariate higher order von Mises expansions. *Metrika* **2001**, *53*, 123–140.