# Maximum Likelihood Estimation of Scale

John A. Stuller and Abhijit Patait

Department of Electrical Engineering
University of Missouri-Rolla, Rolla, MO 65401-0040
stuller@ee.umr.edu, patait@ee.umr.edu

## Abstract

This paper introduces the solution to the problem of maximum likelihood (ML) scale estimation. The result is obtained using *coupled* Karhunen-Loeve expansions, which were recently introduced in [1-3] to solve the problem of ML displacement (or time delay) estimation. The coupled Karhunen-Loeve expansions lead directly to intuitively reasonable signal processors associated with ML displacement and ML scale estimation. Simulation results that demonstrate the performance of ML scale estimation are included in this paper[1].

## 1 Introduction

The problems of scale and displacement estimation arise in diverse applications such as zoom and translational motion estimation in image processing and bearing/bearing-rate estimation in passive sonar. Scale and displacement are also fundamental concepts in wavelet signal representation, and therefore, the problems of estimating scale and displacement seem complementary. A simple and intuitively satisfying solution to the problem of maximum-likelihood (ML) displacement estimation has been recently obtained using the new concept of coupled Karhunen-Loeve expansions (*CKLE*) [1-3]. In the present paper, we use the CKLE to find the maximum-likelihood (ML) estimate of scale.

We define the problem of *scale estimation* as that of estimating an unknown scale factor, $a > 0$, from observations:

$$\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} = \begin{bmatrix} s(t) \\ s(at) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix}; \; t \in \mathcal{T} \qquad (1)$$

where $s$ is a random signal process, $n$ is noise, and the observation domain is $\mathcal{T} = [-T/2, T/2]$ where $T$ is the observation time. Note that $s(at)$ is a time-expanded (or low frequency) version of $s(t)$ for $a < 1$, and a time-compressed (or high frequency) version of $s(t)$ for $a > 1$. Different regions of the signal are included in $r_1(t)$ and $r_2(t)$ for $a \neq 1$ because the observation interval, $\mathcal{T}$, is the same for both observations.

[1]This work was supported in part by NSF Grant MIP-9223020.

If we represent the observations (1) with a vector $\boldsymbol{r}$, then the ML estimate of scale, $\widehat{a}_{\mathrm{ML}}$, is given by the $A$ for which the conditional probability density function (pdf) or likelihood function, $p_{\boldsymbol{r}|a}(\boldsymbol{R}|A)$, is maximum. We use upper case to denote assumed quantities and arguments of pdfs. To obtain $p_{\boldsymbol{r}|a}(\boldsymbol{R}|A)$, we must have a statistical description for the signal and noise. We assume that $s(t)$ is a zero-mean gaussian process having covariance function $E\{s(t_1)\, s(t_2)\} = K_s(t_1, t_2)$, and $n_1(t)$ and $n_2(t)$ are independent normal processes having power spectral density functions $\mathcal{N}_1/2$ and $\mathcal{N}_2/2$ respectively.

The problem of scale estimation is complementary to that of displacement estimation. Displacement estimation is well known in passive sonar, where the displacement, $d$, is the propagation time-delay of a signal received at two points in space. The observations have the form:

$$\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} = \begin{bmatrix} s(t) \\ s(t-d) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix}; \; t \in \mathcal{T} \qquad (2)$$

Historically, the elements in $\boldsymbol{r}$ for the passive sonar application were equated to the coefficients in complex Fourier series expansions of $r_1(t)$ and $r_2(t)$ [4-6]. Assuming stationary processes and the limit $T \to \infty$, this Fourier representation led to a *generalized cross correlator* as the ML estimator of delay [6]. Extensions of this classic result were obtained by equating the elements $\boldsymbol{r}$ to the coefficients in a generalized Karhunen-Loeve expansion of $[r_1(t), r_2(t)]$ [7,8]. Coupled Karhunen-Loeve expansions were recently shown to provide simpler, more intuitive, processor structures for time delay estimation (*TDE*) for arbitrary $T$, [1-3].

By applying a method similar to that in [6], Knapp and Carter found the ML estimator of scale for stationary processes in the limit $T \to \infty$ [9]. In the present paper, we apply the CKLE to find the ML estimator of scale for nonstationary processes and any $T$. The theoretical development is given in Sections 2 and 3. The experimental performance of a simulated ML scale estimator is described in Section 4.

## 2 Likelihood Function

We can write (1) as

$$\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} = \begin{bmatrix} s(t) \\ s(t;a) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix};\ t \in \mathcal{T} \quad (3)$$

where

$$s_1(t) = s(t), \quad (4)$$

$$s_2(t;a) = s(at) \quad (5)$$

Coupled Karhunen-Loeve expansions can be used to represent $r_1(t)$ and $r_2(t)$ by $N$-element vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. The coupled Karhunen-Loeve expansions are defined as the limit as $N \to \infty$ of:

$$r_1^N(t) = \sum_{k=1}^{N} r_{1,k}\phi_{1k}(t),\ r_2^N(t) = \sum_{k=1}^{N} r_{2,k}\phi_{2k}(t;A) \quad (6)$$

$t \in \mathcal{T}$, where

$$r_{1,k} = \int_{\mathcal{T}} r_1(t)\phi_{1k}(t)dt,\ r_{2,k} = \int_{\mathcal{T}} r_2(t)\phi_{2k}(t;A)dt \quad (7)$$

The representation vectors are the coefficient vectors $\boldsymbol{r}_i = r_{i,1}, r_{i,2}, \cdots r'_{i,N}$, $i = 1,2$. The functions $\phi_{1k}(t)$, $k = 1, 2, \ldots N$, are the normalized solutions (eigenfunctions) of

$$\lambda_{1k}^s \phi_{1k}(t) = \int_{\mathcal{T}} K_{s_1}(t,\tau)\phi_{1k}(\tau)d\tau; \quad (8)$$

where $\lambda_{1k}^s = \mathrm{VAR}\{s_{1,k}\}$, $K_{s_1}(t,\tau) = E\{s_1(t)s_1(\tau)\}$, and $t \in \mathcal{T}$. The functions $\phi_{2k}(t;A)$ are the normalized solutions (eigenfunctions) of

$$\lambda_{2k}^e(A)\phi_{2k}(t;A) = \int_{\mathcal{T}} K_{s_2|\boldsymbol{r}_1,a}(t,\tau;\boldsymbol{R}_1,A)\phi_{2k}(\tau;A)d\tau; \quad (9)$$

$t \in \mathcal{T}$, where

$$K_{s_2|\boldsymbol{r}_1,a}(t,\tau;\boldsymbol{R}_1,A) = E\Big\{[s_2(t;a) - \widehat{s}_2(t;A,\boldsymbol{r}_1)] \times$$

$$[s_2(\tau;a) - \widehat{s}_2(\tau;A,\boldsymbol{r}_1)]\Big| a = A, \boldsymbol{r}_1 = \boldsymbol{R}_1\Big\}, \quad (10)$$

with

$$\widehat{s}_2(t;A,\boldsymbol{R}_1) = \mathrm{E}\{s_2(t;a) \mid a = A,\ \boldsymbol{r}_1 = \boldsymbol{R}_1\}; \quad (11)$$

and $\lambda_{2k}^e(A) = \mathrm{VAR}\{s_{2,k} \mid a = A,\ \boldsymbol{r}_1 = \boldsymbol{R}_1\}$. The CKLE leads to a tractable factorization of the likelihood function

$$p_{\boldsymbol{r}_1,\boldsymbol{r}_2|a}(\boldsymbol{R}_1,\boldsymbol{R}_2|A) = p_{\boldsymbol{r}_1}(\boldsymbol{R}_1)p_{\boldsymbol{r}_2|\boldsymbol{r}_1,a}(\boldsymbol{R}_2 \mid \boldsymbol{R}_1,\ A) \quad (12)$$

and this yields the log-likelihood function [3],

$$\ell = \ell_R^{(1)} + \ell_B^{(1)} + \ell_R^{(2|1)} + \ell_B^{(2|1)} \quad (13)$$

where

$$\ell_R^{(1)} = \frac{1}{\mathcal{N}_1}\int_{\mathcal{T}} r_1(t)\widehat{s}_{1|1}(t)dt, \quad (14)$$

$$\ell_B^{(1)} = -\frac{1}{\mathcal{N}_1}\int_{\mathcal{T}} \xi_c^{s_1}(t)dt, \quad (15)$$

$$\ell_R^{(2|1)} = \frac{2}{\mathcal{N}_2}\int_{\mathcal{T}} r_2(t)\widehat{s}_{2|1}(t;A)dt - \frac{1}{\mathcal{N}_2}\int_{\mathcal{T}} \widehat{s}_{2|1}^2(t;A)dt$$

$$+\frac{1}{\mathcal{N}_2}\int_{\mathcal{T}}[r_2(t) - \widehat{s}_{2|1}(t;A)]\widehat{e}(t;A)dt, \quad (16)$$

$$\ell_B^{(2|1)} = -\frac{1}{\mathcal{N}_2}\int_{\mathcal{T}} \xi_c^e(t|A)dt. \quad (17)$$

In the (14) – (17),

$$\widehat{s}_{1|1}(t) = \int_{\mathcal{T}} h_{1|1}(t,\tau)r_1(\tau)d\tau, \quad (18)$$

$$\widehat{s}_{2|1}(t;A) = \int_{\mathcal{T}} h_{2|1}(t,\tau;A)r_1(\tau)d\tau, \quad (19)$$

and

$$\widehat{e}(t;A) = \int_{\mathcal{T}} h_e(t,\tau;A)[r_2(\tau) - \widehat{s}_{2|1}(\tau;A)]d\tau \quad (20)$$

are, respectively, the noncausal MMSE estimates of $s_1(t)$ from $r_1(\tau)$; of $s_2(t;a)$ from $r_1(\tau)$ given $a = A$; and of $e(t;A) = s_2(t;a) - \widehat{s}_{2|1}(t;A)$ from $r_2(\tau) - \widehat{s}_{2|1}(\tau;A)$ given $a = A$, where $t, \tau \in \mathcal{T}$. $\xi_c^{s_1}(t)$ and $\xi_c^e(t|\,A)$ are, respectively, the mean square errors resulting from the causal MMSE estimates of $s_1(t)$ from $r_1(\tau)$ and of $e(t;A)$ from $r_2(\tau) - \widehat{s}_{2|1}(\tau;A)$ given $a = A$. Since $\ell_R^{(1)}$ and $\ell_B^{(1)}$ do not depend on $A$, the MLE of $a$ is the $A$ maximizing

$$\ell' = \ell_R^{(2|1)} + \ell_B^{(2|1)} \quad (21)$$

If we assume that an upper bandwidth limit on $r_2(t)$ exists with negligible distortion of $s_2(t;a)$, then we can subtract $\frac{1}{\mathcal{N}_2}\int_{\mathcal{T}} r_2^2(t)dt$, from $\ell'$ without affecting $\widehat{a}_{\mathrm{ML}}$. After multiplying this difference by $-\mathcal{N}_2/T$, we obtain

$$\delta = -\frac{\mathcal{N}_2}{T}\left[\ell' - \frac{1}{\mathcal{N}_2}\int_{\mathcal{T}} r_2^2(t)dt\right] = \delta_1 - \frac{\mathcal{N}_2}{T}\ell_B^{(2|1)} \quad (22)$$

where

$$\delta_1 = \frac{1}{T}\int_{\mathcal{T}}[r_2(t) - \widehat{s}_{2|1}(t;A)][r_2(t) - \widehat{s}_{2|1,2}(t;A)]dt \quad (23)$$

with

$$\widehat{s}_{2|1,2}(t;A) = \widehat{s}_{2|1}(t;A) + \widehat{e}(t;A). \quad (24)$$

$\ell'$ of (21) is maximum when $\delta$ of (22) is minimum. Sequential estimates of $s_2(t;a)$ are involved in obtaining $\delta_1$. First, a MMSE estimate, $\widehat{s}_{2|1}(t;A)$ of $s_2(t;a)$ is made from $r_1(\tau)$, $-T/2 \le \tau \le T/2$ based on the assumption that $a = A$. Then $r_2(t)$ is used in "innovations", $r_2(\tau) - \widehat{s}_{2|1}(\tau; A)$, $-T/2 \le \tau \le T/2$, to form the MMSE estimate, $\widehat{e}(t;A)$, of the estimation error $e(t;A) = s_2(t;a) - \widehat{s}_{2|1}(t;A)$ given $a = A$. The quantity $\widehat{s}_{2|1,2}(t;A)$ is the MMSE estimate of $s_2(t;a)$ from *both* $r_1(\tau)$ and $r_2(\tau)$, where $-T/2 \le \tau \le T/2$, based on the assumption that $a = A$. Since the processes involved are gaussian, the signal estimates $\widehat{s}_{2|1}(t;A)$ and $\widehat{s}_{2|1,2}(t;A)$ can be obtained using optimum, noncausal linear filters.

# 3 Signal Estimates

The determination of $\widehat{s}_{2|1}(t;A)$ is straightforward. We first apply an optimum linear smoother to obtain $\widehat{s}(t)$ from $r_1(\tau)$, $-T/2 \le t, \tau \le T/2$. For any $A < 1$, we make the substitution $t \to At$ to obtain the MMSE

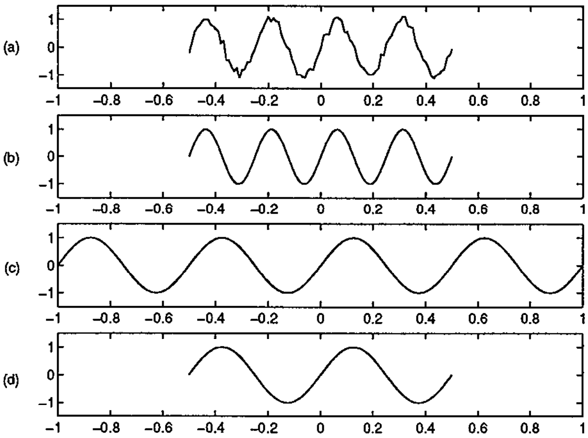


Figure 1: Derivation of $\hat{s}(At); A = 0.5, T = 1$. (a)$r_1(t)$, (b)$\hat{s}(t)$, (c)$\hat{s}(At)$ (Expanded from (b)), and (d)$\hat{s}(At), t \in [-T/2, T/2]$.

estimate of the time-expanded signal $\widehat{s}(At)$, $-T/2 \le At \le T/2 \Rightarrow -T/(2A) \le t \le -T/(2A)$. Because $n_2(t)$ is independent of both $s(t)$ and $n_1(t)$, we set $\widehat{s}_{2|1}(t;A) = \widehat{s}(At)$, $-T/2 \le t \le T/2$. This process is illustrated in Figure 1 for $A = a = 0.5$. (In Figures 1-8, the observation interval, $T$, is taken to be 1. We depict the signals as sinusoidal and the noise as small ripples for simplicity.) Figures 1a and 1b illustrate $r_1(t)$ and $\widehat{s}(t)$. The time-expanded signal estimate $\widehat{s}(At)$ is shown in Figures 1c and 1d, for $-T/(2A) \le t \le -T/(2A)$ and $-T/2 \le t \le T/2$, respectively.

For $A > 1$, we need to optimally extrapolate $\widehat{s}(t)$, $-T/2 \le t \le T/2$ forward in time to obtain $\widehat{s}(t)$ for $T/2 < t \le AT/2$ and backward in time to obtain $\widehat{s}(t)$ for $-AT/2 \le t < -T/2$. We then make the substitution $t \to At$ to obtain the MMSE estimate of the time-compressed signal $\widehat{s}(At)$, $-AT/2 \le At \le AT/2 \Rightarrow -T/2 \le t \le T/2$. Finally, we set $\widehat{s}_{2|1}(t;A) = \widehat{s}(At)$, $-T/2 \le t \le T/2$. These steps are illustrated in Figure 2 for $A = a = 2$. Figure 2a and 2b depict $r_1(t)$ and $\widehat{s}(t)$, $-T/2 \le t \le T/2$. Figure 2c depicts the optimally extrapolated $\widehat{s}(t)$ for $-AT/2 \le t \le AT/2$. The compressed version of this optimal signal estimate, $\widehat{s}_{2|1}(t;A) = \widehat{s}(At)$, $-T/2 \le t \le T/2$, is shown in Figure 2d.

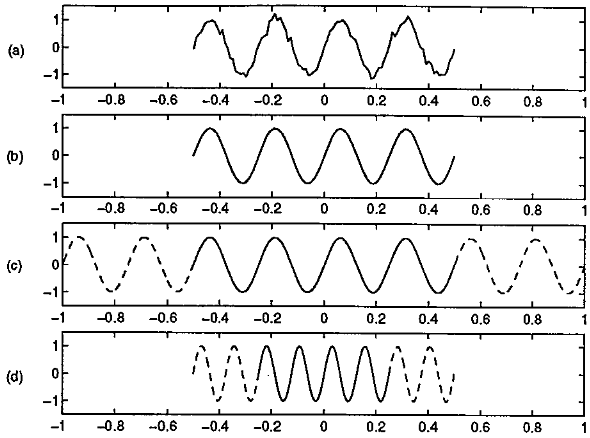


Figure 2: Derivation of $\hat{s}(At); A = 2.0, T = 1$. (a)$r_1(t)$, (b)$\hat{s}(t)$, (c)$\hat{s}(t)$ with extrapolation (dashed), and (d)$\hat{s}(At), t \in [-T/2, T/2]$, (compressed from (c)).

We next consider $\widehat{s}_{2|1,2}(t;A)$. The determination of $\widehat{s}_{2|1,2}(t;A)$ is facilitated by applying a reversible transformation $[r_1(t), r_2(t)] \leftrightarrow [r_1'(t), r_2'(t)]$. We derive a sufficient statistic, $\ell(t)$, from $[r_1'(t), r_2'(t)]$ and obtain $\widehat{s}_{2|1,2}(t;A)$ by optimally smoothing $\ell(t)$. The forms of the reversible transformation and the sufficient statistic depend on $A$. For $A \le 1$,

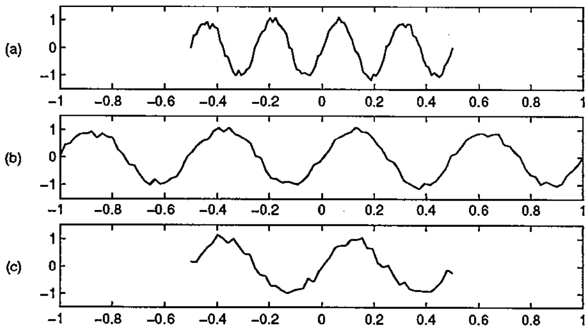


Figure 3: Steps leading to $r_1'(t)$ and $r_2'(t)$ for $A < 1$. The figure assumes $A = a = 0.5, T = 1$. (a)$r_1(t)$, (b)$r_1(At)$ (expanded from (a)), and (c)$r_2(t)$.

$$\begin{bmatrix} r_1'(t) \\ r_2'(t) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} r_1(At) + r_2(t) \\ r_1(At) - r_2(t) \end{bmatrix}; \tag{25}$$

where $|t| \leq T/2$; and

$$\begin{bmatrix} r_1'(t) \\ r_2'(t) \end{bmatrix} = \begin{bmatrix} r_1(At) \\ 0 \end{bmatrix}; \tag{26}$$

where $T/2 \leq |t| \leq T/(2A)$. The steps leading to $r_1'(t)$ and $r_2'(t)$ are illustrated in Figure 3 for $A = a = 0.5$. Figures 3a and 3b illustrate $r_1(t)$ and $r_1(At)$ respectively. Figure 3c illustrates $r_2(t)$.

Figures 4 and 5 illustrate $r_1'(t)$ and $r_2'(t)$, respectively, as defined in (25)-(26).

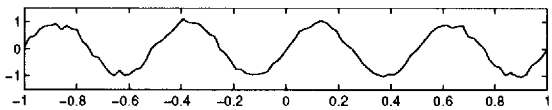


Figure 4: $r_1'(t)$ of (25)-(26).

If $a = A$, then (25)-(26) simplify to

$$\begin{bmatrix} r_1'(t) \\ r_2'(t) \end{bmatrix} = \begin{bmatrix} s(At) + n_1'(t) \\ n_2'(t) \end{bmatrix}; \tag{27}$$

where $|t| \leq T/2$; and

$$\begin{bmatrix} r_1'(t) \\ r_2'(t) \end{bmatrix} = \begin{bmatrix} s(At) + n_1(At) \\ 0 \end{bmatrix}; \tag{28}$$

where $T/2 < |t| \leq T/(2A)$. $n_1'(t)$ and $n_2'(t)$ are defined by

$$\begin{bmatrix} n_1'(t) \\ n_2'(t) \end{bmatrix} \triangleq \frac{1}{2} \begin{bmatrix} n_1(At) + n_2(t) \\ n_1(At) - n_2(t) \end{bmatrix}; \tag{29}$$

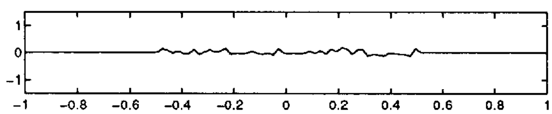


Figure 5: $r_2'(t)$ of (25)-(26).

The sufficient statistic for the estimate $\widehat{s}_{2|1,2}(t; A)$ for $A \leq 1$ is

$$\ell(t) = r_1'(t) - \widehat{n}_1'(t) \tag{30}$$

where $|t| \leq T/(2A)$, and

$$\widehat{n}_1'(t) = \frac{\mathcal{N}_1 A^{-1} - \mathcal{N}_2}{\mathcal{N}_1 A^{-1} + \mathcal{N}_2} r_2'(t) \tag{31}$$

is the MMSE estimate of $n_1'(t)$ from $r_2'(t) \equiv n_2'(t)$.

For $A \geq 1$,

$$\begin{bmatrix} r_1'(t) \\ r_2'(t) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} r_1(At) + r_2(t) \\ r_1(At) - r_2(t) \end{bmatrix}; \tag{32}$$

where $|t| \leq T/(2A)$; and

$$\begin{bmatrix} r_1'(t) \\ r_2'(t) \end{bmatrix} = \begin{bmatrix} r_2(t) \\ 0 \end{bmatrix}; \tag{33}$$

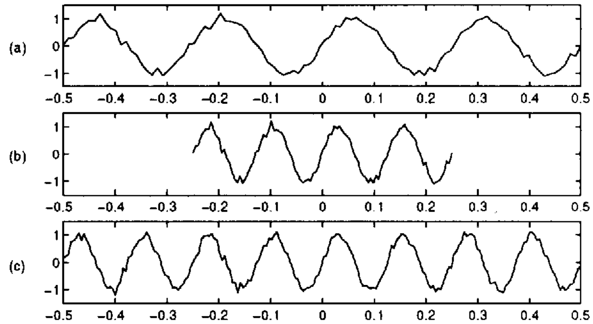


Figure 6: Steps leading to $r_1'(t)$ and $r_2'(t)$ for $A \geq 1$. The figure assumes $A = a = 2, T = 1$. (a)$r_1(t)$, (b)$r_1(At)$ (compressed from (a)), and (c)$r_2(t)$.

where $T/(2A) < |t| \leq T/2$. The steps leading to $r_1'(t)$ and $r_2'(t)$ are illustrated in Figure 6 for $A = a = 2$. Figures 6a and 6b illustrate $r_1(t)$ and $r_1(At)$ respectively. Figure 6c illustrates $r_2(t)$. Figures 7 and 8 illustrate $r_1'(t)$ and $r_2'(t)$, respectively, as defined in (32)-(33).

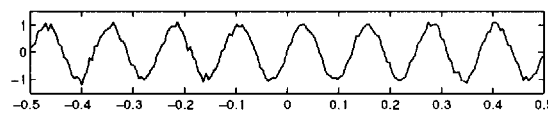


Figure 7: $r_1'(t)$ of (32)-(33).

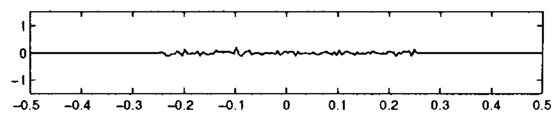


Figure 8: $r_2'(t)$ of (32)-(33).

If $a = A$, then (32)-(33) simplify to

$$\begin{bmatrix} r_1'(t) \\ r_2'(t) \end{bmatrix} = \begin{bmatrix} s(At) + n_1'(t) \\ n_2'(t) \end{bmatrix}; \tag{34}$$

where $|t| \leq T/2$; and

$$\begin{bmatrix} r_1'(t) \\ r_2'(t) \end{bmatrix} = \begin{bmatrix} s(At) + n_2(t) \\ 0 \end{bmatrix}; \tag{35}$$

where $T/(2A) < |t| \leq T/2$. The sufficient statistic for the estimate $\widehat{s}_{2|1,2}(t; A)$ for $A \geq 1$ is given by (30) where $|t| \leq T/2$.

Simulation results for ML scale estimation are presented in the following section. Corresponding simulation results for ML *displacement* estimation appear in [1-3,8].

# 4 Simulation Experiments

We investigated the performance of ML scale estimation by means of digital simulation experiments, written in MATLAB© [10]. We assumed that the continuous-time (*CT*) signal, $s(t)$, being simulated was a stationary, zero-mean, gaussian process with autocorrelation function

$$R_s(\tau) = \sigma_s^2 \exp\{-k|\tau|\} \tag{36}$$

where $\sigma_s^2$ is the variance and $k$ is the 3 dB bandwidth of the signal process in radians per second. The *noise equivalent bandwidth* [11] is $0.25k$ in Hertz, and the signal process *correlation time* is $\tau_c = k^{-1}$. Hence, there are $kT$ correlation intervals in the observation time $T$. The CT noise processes, $n_1(t)$ and $n_2(t)$, were assumed to have equal power spectra, $\mathcal{N}_1/2 = \mathcal{N}_2/2 \triangleq \mathcal{N}/2$.

The discrete time signal was generated using the recursion

$$x(i) = \exp\{-k\Delta t\}\, x(i-1) + \nu(i) \tag{37}$$

where $\nu(i)$ is a zero-mean, stationary gaussian sequence, and $i_0$ was sufficiently large for $s(i\Delta t) \triangleq x(i-i_0)$ to be stationary for $i\Delta t \in \mathcal{T}$. The noise sequences, $n_1(i\Delta t)$ and $n_2(i\Delta t)$ were zero mean iid gaussian sequences, each having variance $\sigma_n^2 \equiv [\mathcal{N}/2] \times [1/\Delta t]$. Thus, the signal to noise ratio (SNR) in the noise-equivalent signal bandwidth is $\mathrm{SNR} = \sigma_s^2/[2 \times 0.25\, k\mathcal{N}/2] = 4\sigma_s^2/(k\mathcal{N}) = [2/(k\Delta t)][\sigma_s^2/\sigma_n^2] = [2N/(kT)][\sigma_s^2/\sigma_n^2]$ where $N = T/\Delta t$. We took $N = 20kT$. Thus, in decibels (dB),

$$\mathrm{SNR}\big|_{\mathrm{dB}} = \left.\frac{\sigma_s^2}{\sigma_n^2}\right|_{\mathrm{dB}} + 16 \tag{38}$$

The $N = 20kT$ samples $s_2(i\Delta t;\, a)$, (3), and $r_1(iA\Delta t)$, (32)-(34), were generated from oversampled versions, $s(j\delta t)$ and $r_1(j\delta t)$, of $s(t)$ and $r_1(t)$ respectively, where $\delta t = \Delta t/64$.

Signal estimates $\widehat{s}_{2|1}(i\Delta t; A)$ and $\widehat{s}_{2|1,2}(i\Delta t; A)$ were obtained, respectively, from the length $N$ sequences $r_1(i\Delta t)$ and $\ell(i\Delta t)$ as described in Section 3 using optimal processing [12]. The simulation experiments assumed true scale, $a$, in the range $\{0.125, 0.25, 0.5, 1, 2, 4, 8\}$, and trial scale, $A$, in the range $\{ai/8;\, i = 4, 5, \ldots, 16\}$. Any "endpoint estimates" $\hat{a} = a/2$ and $\hat{a} = 2a$ were declared anomalous and discarded, but a count of the percentage of anomalous estimates was made.

Five hundred independent simulations were conducted for each set of parameters. The scale estimates were refined using a three-point quadratic interpolation [13]. The bias, $\ell_B^{(2|1)}$, was not computed.

Tables 1-4 list the percent anomalies (PA), the experimental average (EA) and experimental standard deviation (ESD) of the scale estimates as a function of $a$, with $\sigma_s^2/\sigma_n^2|_{\mathrm{dB}} = 3$, 6, 10 and 20 and $kT = 4$.

We see from the statistical performance results given in the tables that performance improves with increasing SNR, as expected. A remarkable feature of the results is that performance degrades significantly as the true scale, $a$, is increased beyond 1. We attribute this degradation to the fact that extrapolation is re-

Table 1: Simulation Results: $\sigma_s^2/\sigma_n^2 = 3$ dB

| $a$ | PA | EA | ESD |
|---|---|---|---|
| 1/8 | 5.8 | 0.14 | .04 |
| 1/4 | 2.2 | 0.27 | .06 |
| 1/2 | 0.2 | 0.51 | .07 |
| 1 | 10. | 1.1 | .27 |
| 2 | 24. | 2.5 | .8 |
| 4 | 30. | 6.2 | 2.6 |
| 8 | 30. | 12. | 5.1 |

Table 2: Simulation Results: $\sigma_s^2/\sigma_n^2 = 6$ dB

| $a$ | PA | EA | ESD |
|---|---|---|---|
| 1/8 | 1.4 | 0.14 | 0.03 |
| 1/4 | 1.2 | 0.26 | 0.03 |
| 1/2 | 0.0 | 0.50 | 0.03 |
| 1 | 7.2 | 1.1 | 0.19 |
| 2 | 13.0 | 2.3 | 0.62 |
| 4 | 28. | 5.8 | 2.34 |
| 8 | 22. | 12. | 4.46 |

Table 3: Simulation Results: $\sigma_s^2/\sigma_n^2 = 10$ dB

| $a$ | PA | EA | ESD |
|---|---|---|---|
| 1/8 | 0.04 | 0.13 | 0.03 |
| 1/4 | 0.0 | 0.25 | 0.03 |
| 1/2 | 0.0 | 0.50 | 0.01 |
| 1 | 1.2 | 1.0 | 0.02 |
| 2 | 2.0 | 3.0 | 0.30 |
| 4 | 15. | 5.0 | 1.70 |
| 8 | 14. | 11. | 3.87 |

Table 4: Simulation Results: $\sigma_s^2/\sigma_n^2 = 20$ dB

| $a$ | PA | EA | ESD |
|---|---|---|---|
| 1/8 | 0.0 | 0.13 | 0.002 |
| 1/4 | 0.0 | 0.25 | 0.003 |
| 1/2 | 0.0 | 0.50 | 0.004 |
| 1 | 0.0 | 1.0 | 0.009 |
| 2 | 0.0 | 2.0 | 0.04 |
| 4 | 0.2 | 4.1 | 0.25 |
| 8 | 2.2 | 8.8 | 1.8 |

quired to form the signal estimate $\widehat{s}_{2|1}(t; A) = \widehat{s}\,(At)$ for $A \geq 1$ as explained in the text surrounding Figure 2.

# 5 Concluding Remarks

We have used the recently introduced CKLEs to obtain the first solution to the problem of ML scale estimation and have studied the performance of ML scale estimation by means of simulation experiments. This paper complements previous work of ML displacement estimation appearing in [1-3] and demonstrates further the usefulness of CKLEs for developing algorithms for ML parameter estimation.

# Acknowledgement

The simulation program was based on a program written by Prof. Nancy Hubing for time-delay estimation [1].

# References

[1] J. A. Stuller and N. Hubing, "Maximum-Likelihood Time Delay Estimation: A New Perspective", *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 1995.

[2] J. A. Stuller and N. Hubing, "Maximum-Likelihood Parameter and Time Delay Estimation Using Two Channels", *Proc. ICASSP 1996*, May 7-9, 1996.

[3] J. A. Stuller and N. Hubing, "New Perspectives for Maximum-Likelihood Time Delay Estimation", *IEEE Trans. on Signal Processing*, Mar. 1997.

[4] E. J. Hannan and P. J. Thompson, "The Estimation of Coherence and Group Delay", *Biometrika*, Vol. 58, pp. 469-81, 1971.

[5] W. R. Hahn and S. A. Tretter, "Optimum Processing for Delay-Vector Estimation in Passive Signal Array", *IEEE Trans. on Inform. Theory*, Vol. 19, pp. 608-614, 1973.

[6] C. H. Knapp and G. C. Carter, "The Generalized Correlation Method for Estimation of Time Delay", *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-24, Aug. 1976.

[7] J. A. Stuller, "Maximum-Likelihood Estimation Time-Varying Delay-Part I", *IEEE Trans. Acoust. Speech, Signal Processing*, ASSP-35, March 1987.

[8] B. Champagne, M. Eizenman, and S. Pasupathy, "Exact Maximum Likelihood Time Delay Estimation", *IEEE Trans. Signal Processing*, ASSP-39, pp. 1245-1257, June 1991.

[9] C. H. Knapp and G. C. Carter, "Estimation of Time Delay in the Presence of Source or Receiver Motion", *J. Acoust. Soc. Am.*, Vol. 61, No. 6, pp. 1545-1549, June 1977.

[10] Abhijit S. Patait, *Applications of Coupled Karhunen-Loeve Expansions*, M.S. Thesis, The University of Missouri-Rolla, in preparation.

[11] R. E. Ziemer and W. H. Tranter, *Principles of Communications*, Houghton Mifflin, New York, p. 361, 4th Ed., 1995.

[12] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*, Prentice Hall, 1995.

[13] J. P. Ianniello, "Time Delay Estimation via Cross Correlation in the Presence of Large Estimation Errors", *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-30, pp. 998-1003, Dec. 1982.