MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

AFOSR·TR· 87-1530

Technical Report
Contract No. AFOSR-87-0224
July 1, 1987 - June 30, 1988

OUTLIER RESISTANT PREDICTIVE SOURCE ENCODING
FOR A GAUSSIAN STATIONARY NOMINAL SOURCE

Submitted to:

Air Force Office of Scientific Research/NM
Building 410
Bolling Air Force Base
Washington, D.C. 20332-6448

Attention: Major Brian W. Woodruff, NM

Submitted by:

P. Kazakos
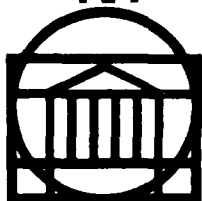Professor

DTIC
SELECTED
OCT 2 7 1987
S D

Report No. UVA/525682/EE88/102

September 1987

SCHOOL OF ENGINEERING AND
APPLIED SCIENCE

DEPARTMENT OF ELECTRICAL ENGINEERING

UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22901

87 10 15 050

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	None

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	Approved for Public Release; Distribution Unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
UVA/525682/EE88/102	**AFOSR·TR· 8 7 - 1 5 3 0**

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
University of Virginia Dept. of Electrical Engr.		Air Force Office of Scientific Rsearch/PKZ

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
Thornton Hall Charlottesville, Virginia 22901	Building 410 Bolling Air Force Base Washington, D.C. 20332-6448

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
Air Force Office of Scientific Research/NM		AFOSR-87-0224

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
Building 410 Bolling Air Force Base Washington, D.C. 20332-6448	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
	61102F	2304	A·5	

11. TITLE (Include Security Classification)
Outlier Resistant Predictive Source Encoding For A Gaussian Stationary Nominal Source

12. PERSONAL AUTHOR(S)
P. Kazakos

13a. TYPE OF REPORT Journal Annual Technical	13b. TIME COVERED FROM 07/01/87 TO 06/30/88	14. DATE OF REPORT (Year, Month, Day) 87 September 18	15. PAGE COUNT 39

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A sequence of qualitatively robust predictive source encoders, for a Gaussian stationary source with outlier contaminated observation data, is proposed and analyzed. Performance measures include mean difference-sequence distortion and output entropy at the nominal Gaussian source, as well as breakdown point and influence function. The proposed sequence of predictive encoders attains strictly positive breakdown point and uniformly bounded influence function, at the expense of increased mean difference-squared distortion and differential entropy, at the Gaussian nominal source.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☑ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Major Brian W. Woodruff	(202) 767-5027	NM

DD FORM 1473, 84 MAR
83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

~~Annual~~ Technical Report
Contract No. AFOSR-87-0224
July 1, 1987 - June 30, 1988

OUTLIER RESISTANT PREDICTIVE SOURCE ENCODING
FOR A GAUSSIAN STATIONARY NOMINAL SOURCE

Submitted to:

Air Force Office of Scientific Rsearch/NM
Building 410
Bolling Air Force Base
Washington, D.C. 20332-6448

Attention: Major Brian W. Woodruff, NM

Submitted by:

P. Kazakos
Professor

Department of Electrical Engineering

SCHOOL OF ENGINEERING AND APPLIED SCIENCE

UNIVERSITY OF VIRGINIA

CHARLOTTESVILLE, VIRGINIA

Accesion For		
NTIS CRA&I	☑	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

DTIC COPY INSPECTED 7

Report No. UVA/525682/EE88/102

September 1987

Copy No. 5

TABLE OF CONTENTS

20
C
22a

DD f

iii

I. INTRODUCTION

Predictive source encoding with distortion is considered, for an analog source, in the presence of an outlier model. In particular, a stationary Gaussian source is assumed, and observation data that are a mixture of source data and outlier data. The objective then is to design a sequence of predictive source encoders which attain satisfactory mean difference-squared distortion in both the presence and the absence of outlier data, subject to an output entropy constraint. As compared to the optimal at the Gaussian source sequence of predictive encoders, the tradeoff is increased mean difference-squared distortion and differential output entropy at the nominal Gaussian source, at the gain of good mean distortion performance in the presence of outliers, (for parametric source encoding studies, see [1]).

II. PRELIMINARIES

Let $[\mu_o, X, R]$ be a discrete-time, stationary and zero mean real source, where R denotes the real line, where X is the name of the source, and where μ_o is its measure. Let X_i, i=1,2,..., denote random variables generated by the source, let x_i, i=1,2..., denote realizations of those variables, and let $X_i^j \overset{\Delta}{=} [X_i,...,X_j]^T$ and $x_i^j \overset{\Delta}{=} [x_i,...,x_j]^T$, for $j \geq i$. Let R^n denote n one-sided multiples of the real line. Let the measure μ_o be known, and let us then call $[\mu_o, X, R]$ the nominal source.

We now consider the outlier model. Then, if $[\mu, Y, R]$ denotes the observation process, if Y_i denotes the i-th random variable generated by this process with y_i denoting its realization, and if Y_i^j and y_i^j denote vectors as in the above paragraph, we have:

$$Y_i = (1-V_i)X_i + V_i Z_i \ , \ i=1,2,... \tag{1}$$

; where X_i is the i-th random variable generated by the nominal source, where $\{Z_i\}$ is a sequence of random variables whose measure is unknown, and where the variables $\{V_i\}$ are i.i.d. and binary, with:

$$P(V_i = 0) = 1-\varepsilon \ , \ P(V_i = 1) = \varepsilon \tag{2}$$

for some ε such that $0 \leq \varepsilon < 1$. The sequence $\{V_i\}$ determines the contamination law, and the sequence $\{Z_i\}$ corresponds to the contaminating process, which is not necessarily stationary. If $\varepsilon = 0$, then the observation process is identical to the nominal source $[\mu_o, X, R]$.

We will assume that the nominal source and the sequence $\{Z_i\}$ are both absolutely continuous. We then denote by $f_o^m(y_1^m)$, the m-dimensional density function induced by the nominal source at the vector point y_1^m. We denote by $f_\varepsilon^m(y_1^m)$ the m-dimensional density function of the random vector Y_1^m at the vector point y_1^m, where Y_i is as in (1) and V_i is as in (2). Let us define the following class of m-dimensional density funcitons:

$$F_\varepsilon^m = \{f^m : f^m = (1-\varepsilon)^m f_o^m + [1-(1-\varepsilon)^m]h^m \ , \tag{3}$$

2

where h^m is any m–dimensional density function}

It can be easily seen then that $f_\varepsilon^m \varepsilon F_\varepsilon^m$. That is, F_ε^m is an enlargement of the class of m-dimensional densities that are generated by the outlier model in (1) and (2). An alternative form of the class F_ε^m is as follows:

$$F_\delta^m = \{ f^m : f^m(y_1^m) - (1-\delta)f_0^m(y_1^m) \geq 0 \ ; \ \forall y_1^m \varepsilon R^m$$

$$\int_{R^m} f^m(y_1^m) dy_1^m = 1 \}$$

(4)

; where

$$\delta \overset{\Delta}{=} 1 - (1-\varepsilon)^m \ : \ 0 \leq \delta \leq 1$$

(5)

Let C_ε denote the class of observation processes generated by the outlier model in (1) and (2), and let us signify a process $[\mu, Y, R]$ by its measure μ. Then, $\mu \varepsilon C_\varepsilon$, means that the process $[\mu, Y, R]$ is contained in class C_ε, and clearly $\mu_0 \varepsilon C_\varepsilon$, where μ_0 is the nominal source $[\mu_0, X, R]$.

We consider predictive source coding with distortion for the nominal source μ_0, when the observation process μ belongs to the class C_ε. In particular, for every given infinite observation sequence y_1^∞, we wish to design a sequence $\{v_{m,y_1^m}\}_{m \geq 1}$ of generally stochastic operations, such that v_{m,y_1^m} maps the datum x_{m+1} of the nominal source μ_0. Let us denote by $\{v_m\}_{m \geq 1}$, the sequence of the above operations when the infinite observation sequence y_1^∞ varies in R^∞. Let us denote by $\mu_{\{v_m\}}$ the process induced by $\{v_m\}_{m \geq 1}$ when the observation sequences are generated by the process μ, where $\mu \varepsilon C_\varepsilon$. Then, we are looking for sequences $\{v_m\}_{m \geq 1}$, which satisfy the following properties:

(a) For every μ in C_ε, the entropy $H(\mu_{\{v_m\}})$ of the process $\mu_{\{v_m\}}$ is bounded from above by a given finite number.

(b) There exists some constant $D < E_{\mu_0}\{X^2\}$, such that for every μ in C_ε, the difference-squared mean distortion induced by the sequence $\{v_m\}_{m \geq 1}$ is bounded from above by D. That is, if for given

3

$\mu \varepsilon C_\varepsilon$, Z_{k+1} denotes the (k+1)-th random element from the process $\mu_{\{v_m\}}$, then,

$$E_{\mu_{\{v_m\}}} \{(X_{k+1} - Z_{k+1})^2\} \le D \ ; \ \forall k, \ \forall \mu \varepsilon C_\varepsilon \tag{6}$$

; where X_{k+1} is generated by the nominal source μ_o.

(c) The sequence $\{v_m\}_{m \ge 1}$ induces entropy and difference-squared mean distortion continuities at the nominal source μ_o. That is, given $\eta > 0$, where exists $\gamma > 0$, such that if μ is a process γ-close to μ_o in an appropriate measure, then

$$|H(\mu_{o,\{v_m\}}) - H(\mu_{\{v_m\}})| < \eta \tag{7}$$

$$|E_{\mu_{o\{v_m\}}} \{(X_{k+1} - Z_{k+1})^2\} - E_{\mu_{\{v_m\}}} \{(X_{k+1} - W_{k+1})^2\}| < \eta \ ; \ \forall k \tag{8}$$

; where in (8), X_{k+1} is generated by μ_o, Z_{k+1} is generated by $\mu_{\{v_m\}}$, and W_{k+1} is generated by $\mu_{\{v_m\}}$.

Property (c) corresponds to qualitative robustness, see ([2],[3],[4],[5]), where the appropriate measure of closeness between the processes μ_o and μ is the Prohorov distance with an empirical Prohorov metric, (see [4],[5]). If property (c) is satisfied, then the sequence $\{v_m\}_{m \ge 1}$ is called <u>qualitatively robust</u> at <u>μ_o.</u> From the results in [4] and [6], we conclude that $\{v_m\}_{m \ge 1}$ is qualitatively robust at μ_o within the class of stationary processes μ, if it satisfies the following sufficient continuity conditions, where Π_{γ_l} denotes Prohorov distance with metric $\gamma_1(x,y) \stackrel{\Delta}{=} |x-y|$, and where $\gamma_l(x_1^l, y_1^l) \stackrel{\Delta}{=} l^{-1} \sum_{i-1}^{1} |x_i - y_i|$.

(A) <u>Pointwise continuity.</u> That is, given finite m, given $\eta > o$, given x_1^m, there exists $\delta > o$, such that $y_1^m : \gamma_m(x_1^m, y_1^m) < \delta$ implies $\Pi_{\gamma_1}(v_{m,x_1^m}, v_{m,y_1^m}) < \eta$.

(B) <u>Asymptotic continuity</u> at μ_o. That is, given $\zeta > 0$, $\eta > 0$, there exist integers n_o and l, some $\delta > 0$, and for each $n > n_o$ some $\Delta^n \varepsilon R^n$ with $\mu_o(\Delta^n) > 1 - \eta$, such that for each $x^n \varepsilon \Delta^n$ and y^n such that $\inf \{\alpha : \# [i : \gamma_l(x_i^{i+l-1}, y_i^{i+l-1}) > \alpha] \le n\alpha\} < \delta$, it is implied that $\Pi_{\gamma_l}(v_{n,x^n}, v_{n,y^n}) < \zeta$.

We point out that if for each given x^n and each n, the operation v_{n,x^n} is deterministic, then the Prohorov distance $\Pi_{\gamma_l}(v_{n,x^n}, v_{n,y^n})$ reduces to $|v_{n,x^n} - v_{n,y^n}|$.

4

From now on, we will assume that the nominal source is Gaussian, zero mean, and stationary, with given spectral density. In section III, we will outline the parametric version of our approach, when the observation process is known and predictive source encoding is sought. In section IV, we will design predictive encoding operations for finite dimensionalities of the observation sequences. In the same section, we will also study the performance of those operations, both at the nominal source and in the presence of contaminating processes. In section V, we will consider extensions of the operations found in section IV, for asymptotically long observation sequences. In the same section, we will also study performance issues of those extensions. In section VI, we draw some conclusions.

III. THE PARAMETRIC APPROACH

In this section, we consider the case where the nominal and the observation processes are both known and mutually dependent, and predictive encoding is sought, for entropy reduction. We will denote the nominal and the observation processes, μ_o and μ, respectively, and we will assume that they are absolutely continuous. We will then denote by $f_\mu^m(y_1^m)$ the m-dimensional density function of the observation process, at the vector point y_1^m. We will denote by $f_{\mu_o\mu}(x \mid y_1^m)$ the conditional density at the point x of the datum X_{m+1} from the nominal process μ_o, given the observation vector y_1^m from the observation process μ. We will also adopt the difference-squared distortion criterion.

Given the above, let us initially assume that no entropy reduction is sought. Then, as well known, the sequence $\{v_m\}_{m\geq 1}$ of mappings that minimize mean distortion are deterministic and given by conditional expectations. That is, given m and y_1^m, we have

$$v_{m,y_1^m} = E_{\mu_o\mu}\{X_{m+1} \mid y_1^m\} = \int_R x f_{\mu_o\mu}(x \mid y_1^m)dx \underset{=}{\Delta} m_{\mu_o\mu}(y_1^m) \tag{9}$$

and for Z_{k+1} denoting the (k+1)-th element from the process $\mu_{\{v_m\}}$, the induced by the operations in (9) mean distortion is:

$$e_m(\mu_o,\mu) \underset{=}{\Delta} E_{\mu_{\{v_m\}}}\{(X_{k+1} - Z_{k+1})^2\} = E_{\mu_o}\{X_{k+1}^2\} - \int_{R^m} f_\mu^m(y_1^m)m_{\mu_o\mu}^2(y_1^m)dy_1^m \tag{10}$$

Let us now assume that in upper bound, log M, on the entropy of the process $\mu_{\{v_m\}}$ is given. Then, we design a sequence $\{v_m\}_{m\geq 1}$ of stochastic mappings, as follows:

Step 1

We select a set $\{A_i, 1\leq i\leq M\}$ of intervals on the real line with $A_i \cap A_j = 0$; $\forall i\neq j$, $\bigcup_{1\leq i\leq M} A_i = R$, and $\int_{A_i} f_{\mu_o}(x)dx = M^{-1}$, where $f_{\mu_o}(x)$ is the one-dimensional density of the process μ_o, at the point x.

Step 2

Using the set $\{A_i, i \le i \le M\}$ of Step 1, we design the sequence $\{v_m\}_{m \ge 1}$ of stochastic mappings so that, given m and y_1^m, the mapping v_{m,y_1^m} is a stochastic channel, mapping the sequence y_1^m onto a set $\{v_i, 1 \le i \le M\}$ of scalar real values; it maps y_1^m onto v_i, with probability:

$$p_{i,\mu_o,\mu}(y_1^m) \overset{\Delta}{=} \int_{A_i} f_{\mu_o\mu}(x|y_1^m)dx \tag{11}$$

The set $\{v_i ; 1 \le i \le M\}$ is selected to minimize the mean difference-squared distortion. That is,

$$D_{m,\mu_o,\mu}(\{v_i\}) \overset{\Delta}{=} \int_{R^m} dy_1^m f_\mu(y_1^m) \sum_{i=1}^M p_{i,\mu_o,\mu}(y_1^m) \int_R (x-v_i)^2 f_{\mu_o\mu}(x|y_1^m)dx =$$

$$= \inf_{\{a_i, 1 \le i \le M\}} D_{m,\mu_o,\mu}(\{a_i\}) \tag{12}$$

Then, it is easily found that,

$$v_i = \left[\int_{R^m} dy_1^m f_\mu(y_1^m) p_{i,\mu_o,\mu}(y_1^m) \right]^{-1} \int_{R^m} dy_1^m f_\mu(y_1^m) p_{i,\mu_o,\mu}(y_1^m) m_{\mu_o\mu}(y_1^m) \tag{13}$$

$$D_{m,\mu_o,\mu}(\{v_i\}) = E_{\mu_o}\{X_{m+1}^2\} - \sum_{i=1}^M \left[\int_{R^m} dy_1^m f_\mu(y_1^m) p_{i,\mu_o,\mu}(y_1^m) \right]^{-1} \cdot$$

$$\cdot \left[\int_{R^m} dy_1^m f_\mu(y_1^m) p_{i,\mu_o,\mu}(y_1^m) m_{\mu_o\mu}(y_1^m) \right]^2 \ge$$

$$\ge c_m(\mu_o,\mu); \quad \forall m \tag{14}$$

; where $c_m(\mu_o,\mu)$ is as in (10) and where $m_{\mu_o\mu}(y_1^m)$ is the conditional expectation in (9). Due to (14), we conclude that the stochastic mappings in Step 2 induce higher mean difference-squared distortion than that induced by the conditional expectations in (9), for the gain of reduced output entropy. As the number M increases to asymptotically large values, the mean distortion $D_{m,\mu_o,\mu}(\{v_i\})$ approaches $c_m(\mu_o,\mu)$, and the output entropy increases to the entropy of the nominal process.

7

Let the nominal process μ_o be zero mean and stationary Gaussian with variance per datum r_0^2, and let the observation process μ be μ_o. Let then ρ_m^2 denote the mean-squared error induced by the optimal at μ_o mean-squared one-step predictor, when the size of the observation vector is m. Let the interval A_i in (11) be (a_i, b_i), where $b_i > a_i$. Then, we easily find that the expressions in (13) and (14) take the following form, where $\phi(x)$ and $\Phi(x)$ denote respectively the density function and the distribution of the zero mean and unit variance Gaussian random variable, at the point x:

$$v_i = \left[r_o - r_o^{-1} \rho_m^2 \right] \left[\Phi\left(\frac{b_i}{r_o} \right) - \Phi\left(\frac{a_i}{r_o} \right) \right]^{-1} \left[\phi\left(\frac{a_i}{r_o} \right) - \phi\left(\frac{b_i}{r_o} \right) \right] \tag{15}$$

$$D_{m,\mu_o,\mu}(\{v_i\}) = r_0^2 - r_0^{-2}(r_0^2 - \rho_m^2)^2 \sum_{i=1}^{M} \left[\Phi\left(\frac{b_i}{r_0} \right) - \Phi\left(\frac{a_i}{r_0} \right) \right]^{-1} \left[\phi\left(\frac{b_i}{r_0} \right) - \phi\left(\frac{a_i}{r_0} \right) \right]^2 \tag{16}$$

IV. FINITE DIMENSIONALITY OBSERVATION SEQUENCES

In this section, we consider the outlier model, as exhibited by the observation process in (1) and (2), and we assume that the nominal process is stationary zero mean Guassian. We then wish to design predictive encoding operations v_m, for $1 \leq m \leq l$, where l is some given finite integer. We want the designed operations to satisfy properties (a), (b), and (c) in section II. For given finite l, we adopt a saddle-point game theoretic approach, based on the parametric scheme in section III. We first assume that the processes in the class C_ε in (1) and (2) are all absolutely continuous, and we denote by $f_o^m(y_1^m)$ the m-dimensional density function of the nominal Gaussian process μ_0, at the vector point y_1^m. Then, given l, we consider an enlargement, F_δ^l, of the class of l-dimensional densities generated by the model in (1) and (2), as that in (4). In particular, we consider l-dimensional densities, f, of the observation process, such that $f \in F_\delta^l$, where:

$$F_\delta^l = \{f : f(y_1^m) - (1-\delta)f_o^l(y_1^m) \geq 0 \; ; \; \forall y_1^m \varepsilon R^m,$$

$$\int_{R^l} f(y_1^m)dy_1^m = 1\} \tag{17}$$

$$\delta \overset{\Delta}{=} 1 - (1-\varepsilon)^l \; : \; 0 \leq \delta < 1 \tag{18}$$

Let an upper bound, log M, on the output entropy be given. Then, we wish to design predictive encoding operations which satisfy this bound for every process in class F_δ^l, and which induce mean difference-squared distortion that is upper bounded by a given bound, for every $f \in F_\delta^l$. Our approach evolves from the parametric scheme in section III, and goes as follows:

Step 1

Select a set $\{A_i, 1 \leq i \leq M\}$ of intervals on the real line with $A_i \cap A_j = 0 ; \forall i \neq j$, $\bigcup_{1 \leq i \leq M} A_i = R$, and $\int_{A_i} f_o(x)dx = M^{-1}$, where f_{μ_o} is the one-dimensional density of the Gaussian nominal process μ_0, at the point x.

9

Step 2

Using the set $\{A_i, 1 \le i \le M\}$ in Step 1, and given a process μ whose density function belongs to the class F_δ^l, we form the set $\{p_{i,\mu}, 1 \le i \le M\}$ of probabilities as follows,

$$\text{Given } y_1^m \text{ in } R^m : p_{i,\mu}(y_1^m) \overset{\Delta}{=} \int_{A_i} f_{\mu,\mu}(x \mid y_1^m)dx, \quad 1 \le i \le M \tag{19}$$

Let N_M denote the set of sets $\{a_i ; 1 \le i \le M\}$ of M real numbers. We then consider the following class, D, of mappings $v_l = v_l(\mu, \{a_i\})$, that is generated by varying μ in F_δ^l and $\{a_i\}$ in N_M:

Given μ in F_δ^l and $\{a_i\}$ in N_M, given observation sequence y_1^m, v_{l,y_1^m} maps the sequence y_1^m onto the value a_i, with probability $p_{i,\mu}(y_1^m)$, as in (19). Given $\{a_i\}$ in N_M, given μ_1 and μ_2 in F_δ^l, let $D_l(\mu_1, \mu_2, \{a_i\})$ denote the mean difference-squared distortion induced by the operation $v_l(\mu_2, \{a_i\})$ in D, at the observation process μ_1. Then,

$$D_l(\mu_1, \mu_2, \{a_i\}) = \int_{R^m} dy_1^m f_{\mu_1}(y_1^m) \sum_{i=1}^M p_{i,\mu_2}(y_1^m) \int_R (x - a_i)^2 f_{\mu_2,\mu_1}(x \mid y_1^m)dx \tag{20}$$

We are then searching for a triple $(\mu_1^*, \mu_2^*, \{v_i\})$, such that $\mu_1^* \varepsilon F_\delta^l$, $\mu_2^* \varepsilon F_\delta^l$, $\{v_i\} \varepsilon N_M$, and:

$$\forall \mu_1, \mu_2 \varepsilon F_\delta^l ; \ D_l(\mu_1, \mu_2^*, \{v_i\}) \le D_l(\mu_1^*, \mu_2^*, \{v_i\}) \le D_l(\mu_1^*, \mu_2^*, \{a_i\}) ; \ \forall \{a_i\} \varepsilon N_M \tag{21}$$

Then, we select the $v_l^* = v_l(\mu_2^*, \{v_i\})$ encoding scheme for the class F_δ^l.

Remark   If an encoding scheme $v_l^* = v_l(\mu_2^*, \{v_i\})$ in D exists, such that it satisfies (21), then it is guaranteed that the maximum mean difference-squared distortion that it induces in F_δ^l is $\sup_{\mu \varepsilon F_\delta^l} D_l(\mu, \mu_2^*, \{v_i\})$, subject to the existence of the latter supremum. By construction, the mapping v_l^* also attains maximum entropy in F_δ^l that is bounded from above by log M.

Let $f_o(x \mid y_1^m)$ denote the conditional density of the Gaussian nominal process for the datum X_{m+1} at the point x, given the past sequence y_1^m from the same process. Let $f_o(y_1^m)$ denote the m-dimensional density of the Gaussian nominal process at the vector point y_1^m, and let Q_m be the m-dimensional autocovariance matrix of the process. Let us also then define:

10

$$m_o(y_1^m) \overset{\Delta}{=} \int_R f_o(x \mid y_1^m) dx$$

$$p_{oi}(y_1^m) \overset{\Delta}{=} \int_{A_i} f_o(x \mid y_1^m) dx \tag{22}$$

We then express a theorem whose proof is in the Appendix.

Theorem 1

Given the class F_δ^l in (17), and for every $\delta: 0 \le \delta < 1$, the game in (21) has a solution $(\mu_1*, \mu_2*, \{v_i\})$. If $f_j*(y_1^l)$, j=1,2 denotes the l-dimensional density function of the pocess μ_j*, j=1,2 at the vector point y_1^l, then this solution is as follows:

$$f*(y_1^l) \overset{\Delta}{=} f_1*(y_1^l) = f_2*(y_1^l) = (1-\delta) f_o(y_1^l) . \max(1, \lambda_l^{-1} \{(y_1^l)^T Q_l^{-1} y_1^l\}^{1/2})$$

$$; \text{where, } \lambda_l : \int_{R^l} f*(y_1^l) dy_1^l = 1 \tag{23}$$

and for,

$$q_i(y_1^l) \overset{\Delta}{=} M^{-1}[1 - \min(1, \lambda_l\{(y_1^l)^T Q_l^{-1} y_1^l\}^{-1/2})] +$$

$$+ p_{oi}(y_1^l) \min(1, \lambda_l\{(y_1^l)^T Q_l^{-1} y_1^l\}^{-1/2}) \tag{24}$$

$$v_i = M(1-\delta) \int_{R^l} dy_1^l \, f_o(y_1^l) m_o(y_1^l) q_i(y_1^l) \tag{25}$$

Then,

$$\forall \mu \varepsilon F_\delta^l \; ; \; D_l(\mu, \mu_2*, \{v_i\}) \le D_l(\mu_2*, \mu_2*, \{v_i\}) \overset{\Delta}{=} D_{l,max} =$$

$$= E_{\mu_o}\{X^2\} - (1-\delta)^2 M \sum_{i=1}^M \left[\int_{R^l} dy_1^l \, f_o(y_1^l) m_o(y_1^l) q_i(y_1^l) \right]^2 \tag{26}$$

The encoding scheme v_l* is as follows:

Given an observation sequence y_1^l, v_l* maps it onto v_i with probability $q_i(y_1^l)$.

11

Given l and M, an encoding scheme v_l consists of a set $\{a_i, 1 \le i \le M\}$ of values, and for every observation sequence y_1^l a set $\{p_i(y_1^l), 1 \le i \le M\}$ of probabilities, such that y_1^l is mapped onto a_i with probability $p_i(y_1^l)$. Given l, given some encoding scheme v_l, given an absolutely continuous observation process with arbitrary dimensionality densities, f, let $D_l(f, v_l)$ denote the mean difference-squared distortion induced when v_l is deployed, f is the density of the observation process, and a datum from the nominal Gaussian source is predictively encoded. Let v_l^o denote the optimal at the Gaussian observation process encoding scheme. That is, given an observation sequence y_1^l, v_o^l maps y_1^l onto u_i, with probability $p_{oi}(y_1^l)$, where, given set $\{A_i, i \le i \le M\}$, $p_{oi}(y_1^l)$ is as in (22), and where for $m_o(y_1^l)$ as in (22):

$$u_i = \left[\int_{A_i} f_o(x)dx \right]^{-1} \int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)p_{oi}(y_1^l) \qquad (27)$$

Let the common set $\{A_i, 1 \le i \le M\}$ be used by both the scheme v_l^o and the scheme v_l^* in Theorem 1, and let this set be such that $\int_{A_i} f_o(x)dx = M^{-1}$; $\forall i$. Let f_o denote the arbitrary dimensionality density of the nominal Gaussian source, and let $m_o(y_1^l)$ and $p_{oi}(y_1^l)$ be as in (22) and $q_i(y_1^l)$ be as in (24). Then, by substitution, we easily obtain:

$$D_l(f_o, v_l^o) = E_{\mu_o}\{X^2\} - M \sum_{i=1}^{M} \left[\int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)p_{oi}(y_1^l) \right]^2 \qquad (28)$$

$$D_l(f_o, v_l^*) = E_{\mu_o}\{X^2\} - (1-\delta)M \sum_{i=1}^{M} \left[\int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)q_i(y_1^l) \right]^2 \cdot$$

$$\cdot \left[2 - (1-\delta)M \int_{R^l} dy_1^l f_o(y_1^l)q_i(y_1^l) \right] \qquad (29)$$

Let I^l denote the l-dimensional vector whose elements are all equal to one. Let z denote some scalar real number, and let us then consider a density f, such that, $f(y_1^l) = (1-\zeta)f_o(y_1^l) + \zeta\delta(zI^l)$, where ζ given and such that $0 \le \zeta < 1$, where f_o is the density of the Gaussian nominal source, and where $\delta(\cdot)$ denotes delta function. Given l, given an encoding scheme v_l, let $D_l(f_o, \zeta, z, v_l)$ denote the mean difference-squared distortion induced by v_l, when the observation density is such that $f(y_1^l) = (1-\zeta)f_o(y_1^l) + \zeta\delta(zI^l)$ and a

datum from the Gaussian nominal source is predictively encoded. Then, for $D_l(f_o, v_l^o)$ as in (28) and for $D_l(f_o, v_l^*)$ as in (29), we obtain by substitution:

$$D_l(f_o,\zeta,z,v_l^o) - D_l(f_o,v_l^o) \stackrel{\Delta}{=} V_l(f_o,\zeta,z,v_l^o) =$$

$$= \zeta M \sum_{i=1}^{M} \left[\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) p_{oi}(y_1^l) \right]^2 \left[1 + M p_{oi}(zI^l) \right] \tag{30}$$

$$D_l(f_o,\zeta,z,v_l^*) - D_l(f_o,v_l^*) \stackrel{\Delta}{=} V_l(f_o,\zeta,z,v_l^*) =$$

$$= \zeta(1-\delta)M \sum_{i=1}^{M} [\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q_i(y_1^l)]^2 [2 - (1-\delta)M \int_{R^l} dy_1^l f_o(y_1^l) q_i(y_1^l) + M(1-\delta) q_i(zI^l)] \tag{31}$$

The functions in (30) and (31) represent changes in mean difference-squared distortion, when the observation process shifts from the one corresponding to the nominal source to a mixed process, which with probability $(1-\zeta)$ is the nominal source and which generates deterministic z-amplitude data with probability ζ. The rates of those changes at $\zeta = 0$ are the Influence Functions, $I_l(f_o,z,v_l^o)$ and $I_l(f_o,z,v_l^*)$, of respectively the encoding schemes v_l^o and v_l^*, at the nominal source μ_o and the amplitude value z. That is,

$$I_l(f_o,z,v_l^o) \stackrel{\Delta}{=} \frac{dV_l(f_o,\zeta,z,v_l^o)}{d\zeta} \Big|_{\zeta=0} =$$

$$= M \sum_{i=1}^{M} \left[\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) p_{oi}(y_1^l) \right]^2 [1 + M p_{oi}(zI^l)] \tag{32}$$

$$I_l(f_o,z,v_l^*) \stackrel{\Delta}{=} \frac{dV_l(f_o,\zeta,z,v_l^*)}{d\zeta} \Big|_{\zeta=0} =$$

$$= (1-\delta)M \sum_{i=1}^{M} \left[\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q_i(y_1^l) \right]^2 \left[2 - (1-\delta)M \int_{R^l} dy_1^l f_o(y_1^l) q_i(y_1^l) + M(1-\delta) q_i(zI^l) \right] \tag{33}$$

Given l, given an encoding scheme v_l, given the nominal density f_o, given z and ζ, let us consider the mean difference-squared distortion $D_l(f_o,\zeta,z,v_l)$. Let us allow the value $|z|$ to go to infinity, and let us then find the maximum value ζ for which $D_l(f_o,\zeta,\pm\infty, v_l) \leq E_{\mu_o}\{X^2\}$. This latter value is the

Breakdown Point of the encoding scheme v_l, at μ_o. It represents the highest frequency of extreme amplitude, $(\pm \infty)$, deterministic outlier values that the encoding scheme can tolerate, before it becomes useless; that is, before the observation sequences provide no information about the source data. We now express a lemma, whose proof is in the Appendix.

Lemma 1

Given M, consider a set $\{A_i, 1 \leq i \leq M\}$ of intervals on the real line with $A_i \cap A_j = 0$; $\forall i \neq j$, $\bigcup_{1 \leq i \leq M} A_i = R$, and $\int_{A_i} f_o(x) dx = M^{-1}$, where $f_o(x)$ is the one-dimensional density of the Gaussian nominal source. Let in addition $A_1 = (-\infty, -a)$ and $A_M = (a, \infty)$ for $a > 0$. Let v_l^o be the optimal at the Gaussian process encoding scheme, and let v_l^* be as in Theorem 1. Then, given l, the breakdown points ζ_l^o and ζ_l^* of the schemes v_l^o and v_l^*, respectively, are given by the following expressions.

$$J_l^o = \{1 + MP_{0l}^2 \left[\sum_{i=1}^{M} P_{oi}^2 \right]^{-1} \}^{-1} \tag{34}$$

$$\zeta_l^* = \{1 + (1-\delta) \left[\sum_{i=1}^{M} Q_{oi}^2 \right] \left[\sum_{i=1}^{M} Q_{oi}^2 [2-(1-\delta)M \int_{R^l} dy_1^l f_o(y_1^l) q_i(y_1^l)] \right]^{-1} \}^{-1} \tag{35}$$

; where, for $m_o(y_1^l)$ and $p_{oi}(y_1^l)$ as in (22) and $q_i(y_1^l)$ as in (24),

$$P_{oi} \overset{\Delta}{=} \int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) p_{oi}(y_1^l) \tag{36}$$

$$Q_{oi} \overset{\Delta}{=} \int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q_i(y_1^l) \tag{37}$$

Remarks For finite dimensionalities of the observation sequence, the encoding operation v_l^* clearly satisfies the pointwise continuity property (A) in section II; thus, it is qualitatively robust. As exhibited by expression (24) in Theorem 1, for $\{(y_1^l)^T Q_l^{-1} y_1^l\}^{1/2}$ relatively small, the operation v_l^* maps sequences y_1^l onto the set of values in (25), using the optimal at the Gaussian nominal source conditional probabilities. As $\{(y_1^l)^T Q_l^{-1} y_1^l\}^{1/2}$ increases, however, the operation v_l^* uses a mixture of such

mapping probabilities, and asymptotically, (for $(y_1^l)^T Q_l^{-1} y_1^l \to \infty$), it maps the sequences y_1^l, using the unconditional nominal density function $f_{\mu_o}(x)$. Thus, it disregards extreme observation values, offering protection to data outliers, at the expense of reduced mean difference-squared performance at the nominal source.

Asymptotic Performance

Let us assume that the number of values onto which observation sequences y_1^l are mapped is asymptotically large. That is, $M \to \infty$. We are then interested in the performance of the encoding schemes v_l^o and v_l^*, for given finite l. From expressions (28), (29), (30), (31), (32), and (33), and taking limits, we find:

Define, the scalars A_l and ρ_l as follows:

$$A_l : \int dy_1^l f_o(y_1^l) f_o(x \mid y_1^l) m_o(y_1^l) = A_l x f_o(x)$$

$$\rho_l^2 = E_{\mu_o}\{[X_{l+1} - m_o(X_1^l)]^2\} \tag{38}$$

Then,

$$\lim_{M \to \infty} D_l(f_o, v_l^o) = (1 - A_l^2) E_{\mu_o}\{X^2\} \tag{39}$$

$$\lim_{M \to \infty} \{D_l(f_o, \zeta, z, v_l^o) - D_l(f_o, v_l^o)\} = \zeta A_l^2\{[1 + \rho_l^3]\zeta E_{\mu_o}\{X^2\} + \rho_l m_o^2(zl')\} \tag{40}$$

Define,

$$q(x, y_1^l) \stackrel{\Delta}{=} [1 - \min(1, \lambda_l\{(y_1^l)^T Q_l^{-1} y_1^l\}^{-1/2})] f_o(x) +$$

$$+ \min(1, \lambda_l\{(y_1^l)^T Q_l^{-1} y_1^l\}^{-1/2}) f_o(x \mid y_1^l) \tag{41}$$

Then,

$$\lim_{M \to \infty} D_l(f_o, v_l^*) = E_{\mu_o}\{X^2\} - 2(1-\delta)\int_R dx f_o^{-1}(x)\left[\int_R dy_1^l f_o(y_1^l) m_o(y_1^l) q(x, y_1^l)\right]^2$$

15

$$+ (1-\delta)^2 \int_R dx f_o^{-2}(x) [\int_{R^l} dy_1^l f_o(y_1^l) q(x,y_1^l)][\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2 \geq$$

$$\geq (1-A_l^2) E_{\mu_o} \{X^2\} \tag{42}$$

$$\lim_{M\to\infty} \{D_l(f_o,\zeta,z,v_l^*) - D_l(f_o, v_l^*)\} =$$

$$= \zeta 2(1-\delta) \int_R dx f_o^{-1}(x) \, [\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2$$

$$- \zeta(1-\delta)^2 \int_R dx f_o^{-2}(x) [\int_{R^l} dy_1^l f_o(y_1^l) q(x,y_1^l)][\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2$$

$$+ \zeta(1-\delta)^2 \int_R dx f_o^{-2}(x) q(x,zI^l)[\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2 \tag{43}$$

From the above expressions, and noting that $\lim_{|z|\to\infty} q(x,zI^l) = f_o(x)$, we also find, denoting by Q_l the

$l \times l$ autocovariance matrix of the nominal Gaussian source:

Define,

$$c_l \overset{\Delta}{=} \{(I^l)^T Q_l^{-1} I^l \}^{-1/2} \tag{44}$$

Then,

$$\lim_{M\to\infty} I_l(f_o,z,v_l^o) = A_l^2 \{[1+\rho_l^3] E_{\mu_o} \{X^2\} + z^2 \rho_l m_o^2(I^l)\} \tag{45}$$

$$\lim_{M\to\infty} I_l(f_o,z,v_l^*) =$$

$$= (1-\delta)\{2+(1-\delta)[1-\min(1,\lambda_l c_l |z|^{-1})]\} \int_R dx f_o^{-1}(x) [\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2$$

$$+ (1-\delta)^2 \min(1,\lambda_l c_l |z|^{-1}) \int_R dx f_o^{-2}(x) f_o(x|zI^l) [\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2$$

$$-(1-\delta)^2 \int_R dx f_o^{-2}(x)[\int_{R^l} dy_1^l f_o(y_1^l) q(x,y_1^l)][\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2 \tag{46}$$

$$\lim_{M\to\infty} \zeta_l^o = 0 \tag{47}$$

$$\lim_{M\to\infty} \zeta_l^* = \{1 + (1-\delta)\left[\int_R dx f_o^{-1}(x) [\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2 \right] \cdot$$

16

$$\cdot \left[\int_R dx [\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2 [2f_o^{-1}(x) - (1-\delta)f_o^{-2}(x) \int_{R^l} dy_1^l f_o(y_1^l) q(x,y_1^l)] \right]^{-1} \}^{-1} \quad (48)$$

Let us define,

$$m_l \overset{\Delta}{=} \max_{y_1^l : (y_1^l)^T Q_l^{-1} y_1^l = 1} |m_o(y_1^l)| \quad (49)$$

Then, we can express the following lemma, whose proof is in the Appendix.

Lemma 2

The limit influence function in (46), and the limit breakdown point in (48), that the encoding operation v_l^* induces, are bounded as below:

$$\lim_{M \to \infty} I_l(f_o, z, v_l^*) \le (1-\delta)(3-\delta)4\lambda_l^2 m_l^2 -$$

$$- (1-\delta)^2 \int dx f_o^{-2}(x) [\int_{R^l} dy_1^l f_o(y_1^l)] q(x,y_1^l) [\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2$$

$$; \forall z \quad (50)$$

$$\frac{2\delta}{1+\delta} \le \lim_{M \to \infty} \zeta_l^* \le (1+\delta)4\lambda_l^2 m_l^2 \{ (1+\delta)4\lambda_l^2 m_l^2 + (1-\delta) \int_R dx f_o^{-1}(x) [\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2 \}^{-1} \quad (51)$$

Thus, asymptotically, $(M \to \infty)$, the optimal at the nominal source encoding operation has breakdown point zero, and quadratic influence function. On the other hand, the operation v_l^* has then uniformly bounded influence function and strictly positive breakdown point. Remarks   As compared to the optimal at the nominal source operation v_l^o, the operation v_l^* is asymptotically, $(M \to \infty)$, superior in terms of breakdown point and influence function performances. This is at the expense of mean difference-squared distortion and differential entropy performances, at the nominal Gaussian source. Indeed, as it can be easily seen, asymptotically, $(M \to \infty)$, the process induced by v_l^* and the Gaussian measure μ_o has higher differential entropy than the process induced by v_l^o and μ_o. In addition,

$\lim_{M\to\infty} D_l(f_o,v_l{}^*) > \lim_{M\to\infty} D_l(f_o,v_l^o)$, and from (26) we conclude:

$$\lim_{M\to\infty} D_l(f_o,v_l{}^*) \le E_{\mu_o}\{X^2\} - (1-\delta)^2 \int_R dx f_o^{-1}(x) [\int_{R^l} dy_1^l f_o(y_1^l) m_o(y_1^l) q(x,y_1^l)]^2;$$

$$;\forall f \epsilon F_\delta^l \qquad\qquad (52)$$

Given l, let $H_\mu(v_l)$ denote the differential entropy induced asymptotically, $(M\to\infty)$, by the encoding scheme v_l at the observation process μ. Let $H_{\mu_o,\zeta,z}(v_l)$ denote the differential entropy induced asymptotically, $(M\to\infty)$ by v_l, when the observation sequence is generated by the nominal source μ_o, with probability $(1-\zeta)$, and it consists of deterministic, amplitude-z data, with probability ζ. Let ρ_l be as in (38), and let us define,

$$r_0^2 \stackrel{\Delta}{=} E_{\mu_o}\{X^2\} \quad , \quad \sigma_l^2 \stackrel{\Delta}{=} \rho_l^2 r_0^{-2}$$

$$g(y_1^l) \stackrel{\Delta}{=} \min(1,\lambda_l\{(y_1^l)^T Q_l^{-1} y_1^l\}^{-1/2}) \qquad\qquad (53)$$

Then, we can express the following lemma, whose proof is in the Appendix.

Lemma 3

Let μ be some absolutely continuous observation process. Given l, let $f(y_1^l)$ denote the density function of this process, at the vector point y_1^l. For,

$$B_\mu(v_l^*) \stackrel{\Delta}{=} 2^{-1} \int_{R^l} dy_1^l f(y_1^l)[1-g(y_1^l)]g(y_1^l)[-2+\sigma_l^2+\sigma_l^{-2}+\rho_l^{-2}(1+\sigma_l^2)m_o^2(y_1^l)]$$

$$-[ln\sigma_l]\int_{R^l} dy_1^l f(y_1^l)[1-g(y_1^l)] \qquad\qquad (54)$$

the differential entropies $H_\mu(v_l{}^*)$ and $H_{\mu_o,\zeta,z}(v_l{}^*)$ are bounded from above as follows:

$$H_\mu(v_l{}^*) \le 2^{-1}[1 + ln2\pi\rho_l^2] + B_\mu(v_l{}^*) \qquad\qquad (55)$$

$$H_{\mu_o,\zeta,z}(v_l{}^*) \le 2^{-1}[1+ln2\pi\rho_l^2] + (1-\zeta)B_{\mu_o}(v_l{}^*) +$$

$$+ \zeta 2^{-1} g(zI^{l})[1-g(zI^{l})][-2+\sigma_{l}^{2} + \sigma_{l}^{-2} + \rho_{l}^{-2}(1+\sigma_{l}^{2})z^{2}m_{o}^{2}(I^{l})]$$

$$-\zeta[1-g(zI^{l})]\ln\sigma_{l} \qquad (56)$$

For $|z| \to \infty$, we find a tighter bound on $H_{\mu_{o},\zeta,z}(\nu_{l}^{*})$, as follows:

$$\lim_{|z| \to \infty} H_{\mu_{o},\zeta,z}(\nu_{l}^{*}) = (1-\zeta)H_{\mu_{o}}(\nu_{l}^{*}) - \zeta \int_{R} dx f_{o}(x)\ln f_{o}(x) \leq$$

$$\leq (1-\zeta)2^{-1}[1+\ln 2\pi\rho_{l}^{2}] + \zeta 2^{-1}[1+\ln 2\pi r_{0}^{2}] + (1-\zeta)B_{\mu_{o}}(\nu_{l}^{*})$$

$$= 2^{-1}[1 + \ln 2\pi\rho_{l}^{2}] + (1-\zeta)B_{\mu_{o}}(\nu_{l}^{*}) - \zeta \ln \sigma_{l} \qquad (57)$$

We note that the differential entropy $H_{\mu}(\nu_{l}^{o})$ induced asymptotically, $(M \to \infty)$, at the nominal source by the optimal at the nominal predictive operation ν_{l}^{o} is bounded as follows:

$$H_{\mu_{o}}(\nu_{l}^{o}) = 2^{-1}[1 + \ln 2\pi\rho_{l}^{2}] \qquad (58)$$

Also,

$$H_{\mu_{o},\zeta,z}(\nu_{l}^{o}) = 2^{-1}[1 + \ln 2\pi\rho_{l}^{2}] \quad ; \quad \forall \zeta, z \qquad (59)$$

We point out that when the nominal Gaussian source is k-order Markov, then we select $l=k$, and we deploy the predictive operation ν_{k}^{*} in Theorem 1, for $l=k$.

19

V. ASYMPTOTICALLY LONG OBSERVATION SEQUENCES

In this section, we consider the same outlier model and the same Guassian source, as in section IV, but we include asymptotically long observation sequences. In the presence of such sequences, the precise modelling of the observation processes that evolve from the outlier model in (1) and (2) is an impossible task. On the other hand, enlargements of the class of observation processes, as those in (17), misrepresent the actual class when long observation sequences are considered. In fact, when the length l of the observation sequences tends to infinity, the class F_δ^l in (17) represents the case where the observation process is the nominal source, with probability $(1-\delta)$, and it is some other process, with probability δ; that is, no data mixing is then included, and the outlier model is not then a member of the class. For non Markovian Gaussian nominal source, and asymptotically long observation sequences, we thus extend the predictive operations of section IV adhocly, but in an intuitively satisfactory fashion.

Given l finite, given k, given the observation sequence y_1^{kl}, and for Q_l denoting the l-dimensional autocovariance matrix of the nominal Gaussian source, let us define,

$$a_{i,l}(y_1^{kl}) \triangleq \{ \left[y_{il+1}^{(i+1)l} \right]^T Q_l^{-1} y_{il+1}^{(i+1)l} \}^{1/2} \quad ; \quad 0 \le i \le k-1 \tag{60}$$

For λ_l as in (23) in Theorem 1, let us also define,

$$z_{il+1}^{(i+1)l}(y_1^{kl}) \triangleq \min\left[1, \frac{\lambda_l}{a_{i,l}(y_1^{kl})} \right] y_{il+1}^{(i+1)l}$$

$$\left[z_1^{kl}(y_1^{kl}) \right]^T = \left[z_1^l(y_1^{kl}),...,z_{(k-1)l+1}^{kl}(y_1^{kl}) \right] \tag{61}$$

Let us now consider the following two mapping densities, that map the observation sequence y_1^{kl} onto the real line, for predictive encoding of the datum X_{kl+1} from the nominal source:

$$q'(x,y_1^{kl}) \triangleq \sum_{m=0}^{k-1} \sum_{\{i_j,1\le j\le m\}} \prod_{j=0}^{m} \min\left[1, \frac{\lambda_l}{a_{i_j,l}(y_1^{kl})} \right] \prod_{j=m+1}^{k-1} \left[1-\min\left[1, \frac{\lambda_l}{a_{i_j,l}(y_1^{kl})} \right] \right] \cdot$$

$$\cdot [f_0(x \mid y_{i_j,l}^{(i_j+1)l}, 1\le j\le m) - f_0(x)] + f_0(x) \tag{62}$$

20

$$q^*(x, y_1^{k/}) \overset{\Delta}{=} \left[1 - k^{-1} \sum_{j=0}^{k-1} \min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{k/})} \right) \right] f_o(x)$$

$$+ k^{-1} \sum_{j=0}^{k-1} \min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{k/})} \right) f_o(x \mid z_1^{k/}(y_1^{k/})) \qquad (63)$$

The mapping density in (62) is an intuitively pleasing extension of the operation v_l^* in Theorem 1, but very complex, both in terms of implementation and in terms of analysis. In addition, it does not provide a clear indication as to the mapping values, when their number M is finite. The mapping density in (63) is much simpler. It also has intuitively pleasing characteristics as well: For $\lambda_l \to \infty$, it converges to the optimal at the nominal source mapping. It also disregards extreme data values, using the unconditional density $f_o(x)$ in its mapping, when $k^{-1} \sum_{i=0}^{k-1} \min\left(1, \frac{\lambda_l}{a_{i,l}(y_1^{k/})} \right) \to 0$. In addition, $q^*(x, y_1^{k/})$ provides easy extensions of the mapping values in (25), when M is finite. In conclusion, we propose the following predictive encoding scheme for non Markovian Gaussian nominal sources, and arbitrarily long observation sequences:

Encoding Scheme

Gven M, select a set $\{A_i, 1 \leq i \leq M\}$ of intervals on the real line with $A_i \cap A_j = 0$; $\forall i \neq j$, $\bigcup_{1 \leq i \leq M} A_i = R$, and $\int_{A_i} f_o(x)dx = M^{-1}$, $\forall i$.

Select some finite natural number l, and given $\delta : 0 \leq \delta < 1$, find the positive constant λ_l, as in (23). Then, given k, given an observation sequence $y_1^{k/}$, map $y_1^{k/}$ onto v_i^* with probability $q^*_i(y_1^{k/})$, where for $p_{oi}(y_1^m)$ as in (22), for $z_1^{k/}(y_1^{k/})$ as in (61), and for $a_{i,l}(y_1^{k/})$ as in (60) the values $\{v_i^*, 1 \leq i \leq M\}$ and the probabilities $q_i^*(y_1^{k/})$ are as follows:

$$q_i^*(y_1^{k/}) = M^{-1} \left[1 - k^{-1} \sum_{j=0}^{k-1} \min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{k/})} \right) \right] +$$

$$+ k^{-1} \sum_{j=0}^{k-1} \min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{k/})} \right) p_{oi}\left(z_1^{k/}(y_1^{k/}) \right) \qquad (64)$$

21

$$v_i^* = M(1-\delta) \int_{R^u} dy_1^{kl} f_o(y_1^{kl}) m_o(y_1^{kl}) q_i^*(y_1^{kl}) \qquad (65)$$

<u>Remarks</u>   Given l, given length kl of observation sequences, we will denote the above encoding scheme $v_{l,k}^*$. We will denote by $\{v_{l,k}^*, k \geq 1\}$ the sequence of encoders evolving from $v_{l,k}^*$, for varying k values. We note that the scheme utilizes l-size disjoint blocks of observed data, where l may be considered as a design parameter. In addition, it bounds disjoint l-size blocks of data in $p_{oi}(z_1^{kl}(y_1^{kl}))$, for all i. This is in contrast to the scheme in section IV, and is needed for asymptotic, $(k \rightarrow \infty)$, qualitative robustness.

We now express a lemma, whose proof is in the Appendix.

<u>Lemma 4</u>

Let $\{b_{im}\}$ be the one step prediction coefficients of the nominal Gaussian source, when m-size observation sequences are given. Let $\{b_{im}\}$ be such that, $\sum_{i=1}^{M} |b_{im}| < c^* < \infty$; $\forall m$. Then, the sequence $\{v_{l,k}^*, k>1\}$ of predictive encoders is qualitatively robust at the nominal Gaussian source. That is, it satisfies both continuity conditions (A) and (B) in section II.

Let $D_{l,k}(f_o, v_{l,k}^*)$ denote the mean difference-squared distortion induced by the encoding scheme $v_{l,k}^*$, at the nominal Gaussian source. Let $D_{l,k}(f_o, \zeta, z, v_{l,k}^*)$ denote the mean difference-squared distortion induced by $v_{l,k}^*$, when the l-dimensional observation density is such that, $f\left(y_{jl+1}^{(j+1)l}\right) = (1-\zeta)f_o\left(y_{jl+1}^{(j+1)l}\right) + \zeta\delta(zl')$, and let then $\zeta_{l,k}^*$ be the breakdown point of $v_{l,k}^*$. Given M, we then easily find by substitution, and as in section IV:

$$D_{l,k}(f_o, v_{l,k}^*) = E_{\mu_o}\{X^2\} - (1-\delta)M \sum_{i=1}^{M} \left[\int_{R^u} y_1^{kl} f_o(y_1^{kl}) m_o(y_1^{kl}) q_i^*(y_1^{kl}) \right]^2 .$$

22

$$\cdot\left[2-(1-\delta)M\int_{R^u} dy_1^{kl}f_o(y_1^{kl})q_i{}^*(y_1^{kl})\right] \qquad (66)$$

$$D_{l,1}(f_o,\zeta,z,\,\nu_{l,1}{}^*) - D_{l,1}(f_o,\nu_{l,1}{}^*) =$$

$$= \zeta(1-\delta)M\sum_{i=1}^{M}\left[\int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)q_i{}^*(y_1^l)\right]^2 [2-(1-\delta)M\int_{R^l} dy_1^l f_o(y_1^l)q_i{}^*(y_1^l)$$

$$+ M(1-\delta)q_i{}^*(zI^l)] \qquad (67)$$

$$\zeta_{l,1}{}^* = \{1+(1-\delta)\left[\sum_{i=1}^{M}(Q_{oi}{}^*)^2\right]\left[\sum_{i=1}^{M}(Q_{oi}{}^*)^2[2-(1-\delta)M\int_{R^l}dy_1^l f_o(y_1^l)q_i{}^*(y_1^l)]\right]^{-1}\}^{-1} \qquad (68)$$

; where,

$$Q_{oi}{}^* \overset{\Delta}{=} \int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)q_i{}^*(y_1^l) \qquad (69)$$

Let us define, for $\{a_{j,l}(y_1^{kl})\}$ as in (60) and $z_1^{kl}(y_1^{kl})$ as in (61),

$$q^*(x,y_1^{kl}) \overset{\Delta}{=} \left[1-k^{-1}\sum_{j=0}^{k-1}\min\left[1,\frac{\lambda_l}{a_{j,l}(y_1^{kl})}\right]\right]f_o(x)$$

$$+ k^{-1}\sum_{j=0}^{k-1}\min\left[1,\frac{\lambda_l}{a_{j,l}(y_1^{kl})}\right]f_o(x\mid z_1^{kl}(y_1^{kl})) \qquad (70)$$

Then, if $I_{l,k}\left[f_o,z,\nu_{l,k}{}^*\right]$ denotes the influence function of the operation $\nu_{l,k}{}^*$, and in parallel to the expressions (42), (43), and (48) in section IV, we find the following asymptotic, $(M\to\infty)$, expressions:

$$\lim_{M\to\infty} D_{l,k}(f_o,\nu_{l,k}{}^*) = E_{\mu_o}\{X^2\} - 2(1-\delta)\int_R dx f_o^{-1}(x)\left[\int_{R^u} dy_1^{kl}f_o(y_1^{kl})m_o(y_1^{kl})q^*(x,y_1^{kl})\right]^2$$

$$+ (1-\delta)^2\int_R dx f_o^{-2}(x)\left[\int_{R^u}dy_1^{kl}f_o(y_1^{kl})q^*(x,y_1^{kl})\right]\left[\int_{R^u}dy_1^{kl}f_o(y_1^{kl})m_o(y_1^{kl})q^*(x,y_1^{kl})\right]^2 \qquad (71)$$

$$\lim_{M\to\infty} I_{l,1}(f_o,z,v_{l,1}^*) = 2(1-\delta)\int_R dx f_o^{-1}(x)\left[\int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)q^*(x,y_1^l)\right]^2$$

$$-(1-\delta)^2\int_R dx f_o^{-2}(x)\left[\int_{R^l} dy_1^l f_o(y_1^l)q^*(x,y_1^l)\right]\left[\int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)q^*(x,y_1^l)\right]^2$$

$$+(1-\delta)^2\int_R dx f_o^{-2}(x)q^*(x,zI^l)\left[\int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)q^*(x,y_1^l)\right]^2 \tag{72}$$

$$\lim_{M\to\infty} \zeta_{l,1}^* = \left\{1+(1-\delta)\left[\int_R dx f_o^{-1}(x)\left[\int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)q^*(x,y_1^l)\right]^2\right]\right\}\cdot$$

$$\left[\int_R dx\left[\int_{R^l} dy_1^l f_o(y_1^l)m_o(y_1^l)q^*(x,y_1^l)\right]^2\left[2f_o^{-1}(x)-(1-\delta)f_o^{-2}(x)\int_{R^l} dy_1^l f_o(y_1^l)q^*(x,y_1^l)\right]\right]^{-1}\}^{-1} \tag{73}$$

<u>Remarks</u>   The asymptotic expressions in (72) and (73) correspond to l-size observation blocks and asymptotically many mapping values $\{v_i^*\}$. For l-order Markov nominal Gaussian sources, those expressions represent the asymptotic, ($M\to\infty$), influence function and breakdown point induced by the encoding scheme $\{v_{l,k}^*\}$ at the nominal source, for any k. Comparing expressions (71), (72), and (73), with expressions (42), (43), and (48), in section IV, we can draw the following conclusions:

The encoding scheme in Theorem 1 induces smaller mean difference-squared distortion at the nominal source, than the scheme $v_{l,1}^*$ does. However, the breakdown point of the former is generally smaller than the breakdown point of the latter. The influence function of $v_{l,1}^*$ is bounded, and it converges to its bound slower than the scheme in Theorem 1 does.

If $H_{\mu_o}(v_{l,1}^*)$ denotes the differential entropy induced by the scheme $v_{l,1}^*$ at the nominal source, and for $B_\mu(v_l^*)$ as in (54), $g(y_1^l)$ as in (53), and ρ_l as in (38), we find via methods as those in the proof of Lemma 3:

$$H_{\mu_o}(v_{l,1}^*) \le 2^{-1}\left[1+ln2\pi\rho_l^2\right] + B_{\mu_o}(v_l^*) +$$

24

$$+ 2^{-1} \int_{R^l} dy_1^l f_o(y_1^l)[1-g(y_1^l)]g(y_1^l)[m_o^2(z_1^l(y_1^l)) - m_o^2(y_1^l)] \qquad (74)$$

From the results in Lemma 3, in conjunction with (74), we conclude:

The scheme $v_{l,1}*$ induces lower differential entropy at the nominal

source, than the scheme in Theorem 1 does.

Limiting Behavior

The sequence $\{v_{l,k}*, k \geq 1\}$ of encoders in this section was designed especially for non-Markovian nominal Gaussian sources, and asympotically long observation sequences. Thus, the study of its performance characteristics for $k \to \infty$ is important. We will perform such studies, for the case where the mapping values $\{v_i*\}$ are asymptotically many; that is, for $M \to \infty$. We first express a theorem, whose proof is in the Appendix.

Theorem 2

The influence function $\lim_{M \to \infty} I_{l,k}(f_o, z, v_{l,k}*)$ is uniformly bounded from above, for every z and every k. The breakdown point $\lim_{M \to \infty} \zeta_{l,k}*$ is uniformly bounded from below by a strictly positive constant, for every k.

In view of Theorem 2, we remind the reader that the optimal at the nominal source predictive encoding operation induces asymptotic, $(M \to \infty)$, breakdown point equal to zero, and unbounded quadratic asymptotic, $(M \to \infty)$, influence function, for every dimensionality of the observation sequence. As k increases, the asymptotic, $(M \to \infty)$, mean difference-squared distortion induced by the sequence $\{v_{l,k}*\}$ of encoders at the nominal source, decreases monotonically, but remains uniformly higher than that induced by the optimal at the nominal sequence of predictive encoders. Given k, the former is given by expression (71), where the latter is given by expression (39) in section IV. Let $H_{\mu_o}(v_{l,k}*)$ denote the differential entropy induced by the encoding scheme $v_{l,k}*$ at the nominal source. Then, we express a

25

lemma, whose proof is in the Appendix. For ρ_l as in (38) and r_o and σ_l as in (53), we first define,

$$G_{k,l}(y_1^{kl}) \overset{\Delta}{=} k^{-1} \sum_{j=0}^{k-1} \min\left[1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})}\right] \tag{75}$$

Lemma 5

For $g(y_1^l)$ as in (53), and for,

$$D(\nu_{l,k}^*) \overset{\Delta}{=} \frac{1}{2} \int_{R^{kl}} dy_1^{kl} f_o(y_1^{kl}) G_{k,l}(y_1^{kl})\left[1 - G_{k,l}(y_1^{kl})\right]\left[\sigma_{kl}^{-2} + \sigma_{kl}^2 + \right.$$

$$\left. + r_0^{-2}(1+\sigma_{kl}^{-2})m_0^2(z_1^{kl}(y_1^{kl})) - 2\right]$$

$$- \left[ln\sigma_{kl}\right]\left[1 - \int_{R^l} dy_1^l f_o(y_1^l) g(y_1^l)\right] \tag{76}$$

The differential entropy $H_{\mu_o}(\nu_{l,k}^*)$ is bounded as follows:

$$H_{\mu_o}(\nu_{l,k}^*) \le \frac{1}{2}\left[1 + ln2\pi\rho_{kl}^2\right] + D(\nu_{l,k}^*) \tag{77}$$

If $\{b_{im}\}$ are the one step prediction coefficients of the nominal Gaussian source when m-size observation sequences are given, and if $\sum_{i=1}^{M} |b_{im}| < \infty$, $\forall m$, then there exists $c_l^* < \infty$, such that, $|m_o(z_1^{kl}(y_1^{kl}))| \le \lambda_l c_l^*$. Then, we find a looser upper bound on $H_{\mu_o}(\nu_{l,k}^*)$, as follows:

$$H_{\mu_o}(\nu_{l,k}^*) \le \frac{1}{2}\left[1 + ln2\pi\rho_{kl}^2\right] + C(\nu_{l,k}^*) \tag{78}$$

; where

$$C(\nu_{l,k}^*) \overset{\Delta}{=} \frac{1}{2}\left[\sigma_{kl}^{-2} + \sigma_{kl}^2 + r_0^{-2}(1+\sigma_{kl}^{-2})\lambda_l^2(c_l^*)^2 - 2\right]\left[\int_{R^l} dy_1^l f_o(y_1^l) g(y_1^l) - \right.$$

$$\left. - \int_{R^{kl}} dy_1^{kl} f_o(y_1^{kl}) G_{k,l}^2(y_1^{kl})\right] - \left[ln\sigma_{kl}\right]\left[1 - \int_{R^l} dy_1^l f_o(y_1^l) g(y_1^l)\right] \tag{79}$$

<u>Remarks</u>   The differential entropy $H_{\mu_o}(v_{l,k}^*)$ decreases monotonically with increasing k, and remains strictly higher than the differential entropy induced by the optimal at the nominal predictive encoder, at the nominal source, (given k, the latter equals $\frac{1}{2}\left[1 + ln2\pi\rho_{kl}^2\right]$). In the { italic l im }it, (k→∞), the bound in (78) can be as small as the asymptotic mean-squared error, $\lim_{n\to\infty}\lambda_n^2$, of the optimal at the nominal source one-step predictor allows. This depends on the spectral characteristics of the nominal Gaussian source.

VI. CONCLUSIONS

In this paper, we considered predictive encoders with distortion, for entropy reduction. We considered a stationary and Gaussian nominal source and we designed and analyzed qualitatively robust predictive encoders, for resistance to data outliers. Our encoders offer protection against outlier values, at the expense of increased distortion and differential entropy, at the nominal source.

APPENDIX

Proof of Theorem 1

Let μ_1 and μ_2 be given, and let f_1 and f_2 denote their corresponding densities. Let $f_{\mu_0,\mu}(x,y_1^l)$ denote joint density of the datum X_{m+1} from the nominal process, at the point x, and the random vector Y_1^l from the observation process at the vector point y_1^l. Then, from class F_0^l we conclude:

$$f_{\mu_0,\mu_2}(x,y_1^l) = (1-\delta)f_o(x,y_1^l) + \delta f_o(x)\left[f_2(y_1^l) - (1-\delta)f_o(y_1^l)\right] \qquad (A.1)$$

$$f_{\mu_0,\mu_2}(x\mid y_1^l) = \frac{f_{\mu_0,\mu_2}(x,y_1^l)}{f_2(y_1^l)} = \left[1 - \frac{(1-\delta)f_o(y_1^l)}{f_2(y_1^l)}\right]f_o(x) +$$

$$+ \frac{(1-\delta)f_o(y_1^l)}{f_2(y_1^l)}f_o(x\mid y_1^l) \qquad (A.2)$$

$$p_{i,2}(y_1^l) \overset{\Delta}{=} \int_{A_i} f_{\mu_0,\mu_2}(x\mid y_1^l)dx = M^{-1}\left[1 - \frac{(1-\delta)f_o(y_1^l)}{f_2(y_1^l)}\right] + \frac{(1-\delta)f_o(y_1^l)}{f_2(y_1^l)}p_{oi}(y_1^l) \qquad (A.3)$$

Let us define,

$$b_i^{(1,2)} = (1-\delta)\left[\int dy_1^l f_1(y_1^l)p_{i,2}(y_1^l)\right]^{-1}\int dy_1^l f_o(y_1^l)m_o(y_1^l)p_{i,2}(y_1^l) \qquad (A.4)$$

Then, we easily find,

$$D_l\left[\mu_1,\mu_2,\{b_i^{(1,2)}\}\right] = E_{\mu_*}\{X^2\} - (1-\delta)^2 \sum_{i=1}^{M}\left[\int dy_1^l f_1(y_1^l)p_{i,2}(y_1^l)\right]^{-1} \cdot$$

$$\cdot \left[\int dy_1^l f_o(y_1^l)\, m_o(y_1^l)p_{i,2}(y_1^l)\right]^2 \le$$

$$\le D_l(\mu_1,\mu_2,\{a_i\}) \; ; \; \forall\{a_i\}\epsilon N_M \qquad (A.5)$$

29

$$\sum_i \left[\int dy_1^l f_1(y_1^l) p_{i,2}(y_1^l) \right]^{-1} \left[dy_1^l\, f_0(y_1^l) m_0(y_1^l) p_{i,2}(y_1^l) \right]^2 \geq$$

$$\geq \left\{ \sum_i \int dy_1^l f_1(y_1^l) p_{i,2}(y_1^l) \right\}^{-1} \sum_i \left[\int dy_1^l f_0(y_1^l) m_0(y_1^l) p_{i,2}(y_1^l) \right]^2 =$$

$$= \sum_i \left[\int dy_1^l f_0(y_1^l) m_0(y_1^l) p_{i,2}(y_1^l) \right]^2 \tag{A.6}$$

with equality in (A.6) if $f_1(y_1^l) = f_2(y_1^l)$; $\forall$ $y_1^l \varepsilon R^l$. From (A.6) and (A.5) we conclude,

$$D_l \left[\mu_1, \mu_2, \{b_i^{(1,2)}\} \right] \leq D_l(\mu_2, \mu_2, \{b_i^{(2,2)}\}) =$$

$$= E_{\mu_*}\{X^2\} - (1-\delta)^2 M \sum_{i=1}^{M} \left[\int dy_1^l f_0(y_1^l) m_0(y_1^l) p_{i,2}(y_1^l) \right]^2 \tag{A.7}$$

; where,

$$b_i^{(2,2)} = (1-\delta) M \int dy_1^l f_0(y_1^l) m_0(y_1^l) p_{i,2}(y_1^l) \tag{A.8}$$

Now, $\sup\limits_{\mu_2 \varepsilon F_*^l} D_l(\mu_2, \mu_2, \{b_i^{(2,2)}\})$ corresponds to $\inf\limits_{\mu_2 \varepsilon F_*^l} M \sum_{i=1}^{M} \left[\int dy_1^l f_0(y_1^l) m_0(y_1^l) p_{i,2}(y_1^l) \right]^2$.

Application of calculus of variation gives that f^* in (23) attains the latter infimum. The proof of the theorem is now complete.

Proof of Lemma 1

From (24), we have $\lim\limits_{z \to \pm\infty} q_i(zI^l) = M^{-1}$; $\forall i$. Also, $\lim\limits_{z \to \pm\infty} p_i(zI^l) = \begin{cases} 1 & ; \quad i=1,M \\ 0 & ; \quad \text{otherwise} \end{cases}$

Substituting the above in (30) and (31), in conjunction with (28) and (29), we find that $D_l(f_0, \zeta, \pm\infty, v_l^0) \leq E_{\mu_*}\{X^2\}$; $\forall \zeta \leq \zeta_l^0$ and $D_l(f_0, \zeta, \pm\infty, v_l^*) \leq E_{\mu_*}\{X^2\}$; $\forall \zeta \leq \zeta_l^*$.

Proof of Lemma 2

We easily conclude, for $q(x, y_1^l)$ as in (41):

$$q(x, y_1^l) < f_0(x) + f_0(x \mid y_1^l)$$

and thus,

$$2f_o^{-1}(x) - (1-\delta)f_o^{-2}(x)\int dy_1^l f_o(y_1^l)q(x,y_1^l) > 2\delta f_o^{-1}(x) \tag{A.9}$$

Also,

$$\int dy_1^l f_o(y_1^l)m_o(y_1^l)q(x,y_1^l) \leq$$

$$\leq f_o(x)\int dy_1^l f_o(y_1^l)|m_o(y_1^l)|\min(1,\lambda_l\{(y_1^l)^T Q_l^{-1} y_1^l\}^{-1/2})$$

$$+ \int dy_1^l f_o(y_1^l)|m_o(y_1^l)|\min(1,\lambda_l\{(y_1^l)^T Q_l^{-1} y_1^l\}^{1/2})f_o(x|y_1^l)$$

$$\leq 2\lambda_l m_l f_o(x) \tag{A.10}$$

Applying (A.9) to (48), we find,

$$\lim_{M\to\infty} \zeta_l^* > \frac{2\delta}{1+\delta} \tag{A.11}$$

Applying (A.10) to (48) and (46), we find,

$$\lim_{M\to\infty} \zeta_l^* < (1+\delta)4\lambda_l^2 m_l^2 \ \{(1+\delta)4\lambda_l^2 m_l^2 + (1-\delta)\int_R dx f_o^{-1}(x)\left[\int_{R'} dy_1^l f_o(y_1^l)m_o(y_1^l)q(x,y_1^l)\right]^2 \}^{-1}$$

$$\tag{A.12}$$

$$\lim_{M\to\infty} I_l(f_o,z,v_l^*) \leq (1-\delta)4\lambda_l^2 m_l^2\left\{ 3-\delta-(1-\delta)\min(1,\lambda_l c_l |z|^{-1}\right\}$$

$$+ (1-\delta)^2\, 4\lambda_l^2 m_l^2 \min(1,\lambda_l c_l|z|^{-1})$$

$$- (1-\delta)^2\int_R dx f_o^{-2}(x)\left[\int_{R'} dy_1^l f_o(y_1^l)q(x,y_1^l)\right]\left[\int_{R'} dy_1^l f_o(y_1^l)m_o(y_1^l)q(x,y_1^l)\right]^2$$

$$= (1-\delta)(3-\delta)4\lambda_l^2 m_l^2$$

$$- (1-\delta)^2\int_R dx f_o^{-2}(x)\left[\int_{R'} dy_1^l f_o(y_1^l)q(x,y_1^l)\right]\left[\int_{R'} dy_1^l f_o(y_1^l)m_o(y_1^l)q(x,y_1^l)\right]^2 \tag{A.13}$$

Proof of Lemma 3

Clearly, for $q(x,y_1^l)$ as in (41), we have,

$$-H_\mu(v_l^*) \geq \int_{R'} dy_1^l f(y_1^l)\int_R dx q(x,y_1^l) \log q(x,y_1^l) \tag{A.14}$$

31

; where,

$$\int_R dxq(x,y_1^I)logq(x,y_1^I) = \left[1-\min(1,\lambda_I\{(y_1^I)^TQ_I^{-1}y_1^I\}^{-1/2}) \right] \int_R dxf_o(x)logq(x,y_1^I)$$

$$+ \min(1,\lambda_I,\{(y_1^I)^TQ_I^{-1}y_1^I\}^{-1/2})\int_R dxf_o(x|y_1^I)logq(x,y_1^I) \tag{A.15}$$

Let us define,

$$g(y_1^I) \overset{\Delta}{=} \min(1,\lambda_I\{(y_1^I)^TQ_I^{-1}y_1^I\}^{-1/2}) \tag{A.16}$$

$$r_0^2 \overset{\Delta}{=} E_{\mu_*}\{X^2\} \tag{A.17}$$

Then, from (A.15) and the convexity of the logarithmic function, we obtain:

$$C(y_1^I) \overset{\Delta}{=} \int_R dxq(x,y_1^I)logq(x,y_1^I) = [1-g(y_1^I)]\int_R dxf_o(x)logq(x,y_1^I) +$$

$$+ g(y_1^I)\int_R dxf_o(x|y_1^I)logq(x,y_1^I) \geq$$

$$\geq [1-g(y_1^I)]^2\int_R dxf_o(x)logf_o(x) + [1-g(y_1^I)]g(y_1^I)\int_R dxf_o(x)logf_o(x|y_1^I)$$

$$+ [1-g(y_1^I)]g(y_1^I)\int_R dxf_o(x|y_1^I)logf_o(x) + g^2(y_1^I)\int_R dxf_o(x|y_1^I)logf_o(x|y_1^I)$$

$$= -2^{-1}[1-g(y_1^I)]^2[1 + log2\pi r_0^2] - 2^{-1}g^2(y_1^I)[log2\pi\rho_I^2 + 1]$$

$$-2^{-1}g(y_1^I)[1-g(y_1^I)]\{log2\pi r_0^2 + log2\pi\rho_I^2 + \rho_I^{-2}[r_0^2 + m_o^2(y_1^I)] + r_0^{-2}[\rho_I^2 + m_o^2(y_1^I)]\}$$

$$= -2^{-1}[1 + log2\pi\rho_I^2] + 2^{-1}[1-g(y_1^I)]log\sigma_I^2 -$$

$$-2^{-1}g(y_1^I)[1-g(y_1^I)][-1 + \sigma_I^2 + \sigma_I^{-2} + \rho_I^{-2}(1 + \sigma_I^2)m_o^2(y_1^I)] \tag{A.18}$$

32

; where, *log* is the natural logarithm *ln*, and where,

$$\sigma_l^2 \overset{\Delta}{=} r_0^{-2}\rho_l^2 \qquad (A.19)$$

Substituting (A.18) in (A.14), we find (54). Similarly, we find,

$$-H_{\mu_{\ast}\zeta_{\ast}}(\nu_l{}^\ast) \geq (1-\zeta)\int_{R^l}dy_1^l f_o(y_1^l)\int_R dx\, q(x,y_1^l)\log q(x,y_1^l) +$$

$$+\zeta\int_R dx\, q(x,zI^l)\log q(x,zI^l) \geq -(1-\zeta)2^{-1}[1 + \log 2\pi\rho_l^2] -$$

$$-(1-\zeta)B_{\mu_{\ast}}(\nu_l{}^\ast) - \zeta 2^{-1}[1 + \log 2\pi\rho_l^2] + \zeta 2^{-1}[1-g(2I^l)]\log\sigma_l^2 -$$

$$-\zeta 2^{-1}g(zI^l)[1-g(2I^l)][-1 + \sigma_l^2 + \sigma_l^{-2} + \rho_l^{-2}(1 + \sigma_l^2)z^2 m_o^2(I^l)] \qquad (A.20)$$

Proof of Lemma 4

The mapping $q_i{}^\ast(y_1^{kl})$ is clearly pointwise continuous for every finite k and every i, since

$$\min\left[\frac{\lambda_l}{a_{j,l}(y_1^{kl})}\right]$$ and $p_{oi}(y_1^{kl})$ are both pointwise continuous, for every i and j, and every finite k.

Let now k be given, and let then x_1^{kl} and y_1^{kl} be two sequences such that $\lambda_l\left[y_{jl+1}^{(j+1)l}, x_{il+1}^{(j+1)l}\right] \leq \alpha, 0 \leq \alpha \leq 1$, for $k(1-\alpha)$ of the k i's. Then,

$$|m_o(z_1^{kl}(y_1^{kl})) - m_o(z_1^{kl}(x_1^{kl}))| \leq \alpha c^\ast + \lambda_l\alpha c^\ast = \alpha c^\ast(1 + \lambda_l)$$

and given, x_1^{kl}, given $\varepsilon_1 > 0$, there exists $\alpha_1 > 0$, such that,

$$\{\#j : \gamma_l(y_{jl+1}^{(j+1)l}, x_{jl+1}^{(j+1)l}) > \alpha_1\} < k\alpha_1 \quad \text{implies,}$$

$$|p_{oi}(z_1^{kl}(y_1^{kl})) - p_{oi}(z_1^{kl}(x_1^{kl}))| < \varepsilon_1 \quad ; \quad \forall i \qquad (A.21)$$

Similarly, given x_1^{kl}, given $\varepsilon_2 > 0$, there exists $\delta_2 > 0$ and $\delta_3 > 0$, such that,

$$\gamma_l\left(x_{jl+1}^{(j+1)l}, y_{jl+1}^{(j+1)l}\right) < \delta_2 \text{ implies } |\min\left(1, \frac{\lambda_l}{a_{j,l}(x_1^{kl})}\right) - \min\left(1, \lambda_l \text{overea}_{j,l}(y_1^{kl})\right)| < \varepsilon_2 \quad (A.22)$$

$$\{\#j : \gamma_l(x_{jl+1}^{(j+1)l}, y_{jl+1}^{(j+1)l}) < k\delta_3 \text{ and } \gamma_l\left(x_{jl+1}^{(j+1)l}, y_{jl+1}^{(j+1)l}\right) < \delta_3$$

$$\text{imply, } |\min\left(1, \frac{\lambda_l}{a_{j,l}(x_1^{kl})}\right) p_{oi}(z_1^{kl}(x_1^{kl})) - \min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})}\right) p_{oi}(z_1^{kl}(y_1^{kl}))| < \varepsilon_2 \quad (A.23)$$

Given x_1^{kl}, let now y_1^{kl} be such that:

$$\{\#j : \gamma_l(y_{jl+1}^{(j+1)l}, x_{jl+1}^{(j+1)l}) > \alpha\} \le k\alpha, \text{ for some } \alpha \text{ such that } \alpha \le \min(\delta_2, \delta_3)$$

Then,

$$|q_i*(x_1^{kl}) - q_i*(y_1^{kl})| \le k^{-1} \sum_{j=0}^{k-1} |\min\left(1, \frac{\lambda_l}{a_{j,l}(x_1^{kl})}\right) - \min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})}\right)|$$

$$+ k^{-1} \sum_{j=0}^{k-1} |\min\left(1, \frac{\lambda_l}{a_{j,l}(x_1^{kl})}\right) p_{oi}(z_1^{kl}(x_1^{kl})) - \min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})}\right) p_{oi}(z_1^{kl}(y_1^{kl}))| \le$$

$$\le 2(1-\alpha)\varepsilon_2 + 2\alpha \quad (A.24)$$

From (A.24) we finally conclude that given x_1^{kl}, given $\varepsilon > 0$, there exists $\alpha > 0$, such that,

$$\{\#j : \gamma_l(y_{jl+1}^{(j+1)l}, x_{jl+1}^{(j+1)l}) > \alpha\} \le k\alpha \text{ implies}$$

$$|q_j*(x_1^{kl}) - q_i*(y_1^{kl})| < \varepsilon ; \quad Vi.$$

Proof of Theorem 2

Let us define,

$$q_j*(x, y_1^{kl}) \overset{\Delta}{=} \left[1 - \min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})}\right)\right] f_o(x) + \min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})}\right) f_o(x \mid z_1^{kl}(y_1^{kl})) \quad (A.25)$$

Then,

$$q_j{}^*(x,y_1^{kl}) = k^{-1} \sum_{j=0}^{k-1} q_i(x,y_1^{kl}) \qquad (A.26)$$

If $D_{l,k}(f,v_{l,k}{}^*)$ denotes the mean difference-squared distortion induced by $v_{l,k}{}^*$ at the density f, then,

$$\lim_{M \to \infty} D_{l,k}(f,v_{l,k}{}^*) = k^{-1} \sum_{j=0}^{k-1} D_f^{(j)}(f,v_{l,k}{}^*) \qquad (A.27)$$

; where,

$$D_f^{(j)}(f,v_{l,k}{}^*) \stackrel{\Delta}{=} E_{\mu_o}\{X^2\} - 2(1-\delta)\int_R dx f_o^{-1}(x)\left[\int_{R^\mu} dy_1^{kl} f_o(y_1^{kl}) m_o(y_1^{kl}) q^*(x,y_1^{kl})\right] \cdot$$

$$\cdot \left[\int_{R^\mu} dy_1^{kl} f(y_1^{kl}) q_j{}^*(x,y_1^{kl}) \int_R x f(x \mid y_1^{kl}) dx\right]$$

$$+ (1-\delta)^2 \int_R dx f_o^{-2}(x)\left[\int_{R^\mu} f_o(y_1^{kl}) m_o(y_1^{kl} q^*(x,y_1^{kl})\right]^2 \left[\int_{R^\mu} dy_1^{kl} f(y_1^{kl}) q_j{}^*(x,y_1^{kl})\right] \qquad (A.28)$$

Due to (A.27), we conclude that the influence function induced by $v_{l,k}{}^*$ is the average of the influence functions induced by the operations $\{q_j{}^*, 0 \le j \le k-1\}$. Also, if $\mu_{l,j}{}^*$ denotes the breakdown point of the operation $q_j{}^*$, then the breakdown point of $v_{l,k}{}^*$ is bounded from below by $\min_{0 \le j \le k-1} \mu_{l,j}{}^*$. From (A.28), and due to the boundness of the vector, $z_1^{kl}(y_1^{kl})$, we now conclude that there exists some positive constant, d^*, such that $\mu_{l,j}{}^* \ge d^* \lim_{M \to \infty} \zeta_{l,1}{}^*$; $\forall j$. If $I_{l,j}(f_o,z,q_j{}^*)$ denotes the influence function of the operation $q_j{}^*$ at the nominal source, we also conclude that there exists some finite constant, e^*, such that, $I_{l,j}(f_o,z,q_j{}^*) \le e^* \lim_{M \to \infty} I_{l,1}(f_o,z,v_{l,1}{}^*)$. The Theorem easily follows from the above, where $\lim_{M \to \infty} I_{l,1}(f_o,z,v_{l,1}{}^*)$ is given by (72) and $\lim_{M \to \infty} \zeta_{l,1}{}^*$ is given by (73).

Proof of Lemma 5

For $q*(x, y_1^{kl})$ as in (70), we clearly have,

$$-H_{\mu_0}(v_{l,k}*) \geq \int_{R^{kl}} dy_1^{kl} f_0(y_1^{kl}) \int_R dx q*(x, y_1^{kl}) ln q*(x, y_1^{kl}) \quad (A.29)$$

For $q_j*(x, y_1^{kl})$ as in (A.25), and due to the convexity of the logarithmic function, we have,

$$\int_R dx q*(x, y_1^{kl}) ln q*(x, y_1^{kl}) = k^{-1} \sum_{j=0}^{k-1} \int_R dx q_j*(x, y_1^{kl}) ln q*(x, y_1^{kl}) \geq$$

$$\geq k^{-2} \sum_{j=0}^{k-1} \sum_{i=0}^{k-1} \int_R dx q_j*(x, y_1^{kl}) ln q_i*(x, y_1^{kl}) \quad (A.\widehat{}0)$$

Also,

$$\int_R dx q_j*(x, y_1^{kl}) ln q_i*(x, y_1^{kl}) = \left[1 - min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})} \right) \right] \int_R dx f_0(x) ln q_i*(x, y_1^{kl})$$

$$+ min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})} \right) \int_R dx f_0(x|z_1^{kl}(y_1^{kl})) ln q_i*(x, y_1^{kl}) \geq$$

$$\geq \left[1 - min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})} \right) \right] \left[1 - min\left(1, \frac{\lambda_l}{a_{i,l}(y_1^{kl})} \right) \right] \int_R dx f_0(x) ln f_0(x)$$

$$+ \left[1 - min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})} \right) \right] min\left(1, \frac{\lambda_l}{a_{i,l}(y_1^{kl})} \right) \int_R dx f_0(x) ln f_0(x|z_1^{kl}(y_1^{kl}))$$

$$+ min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})} \right) \left[1 - min\left(1, \frac{\lambda_l}{a_{i,l}(y_1^{kl})} \right) \right] \int_R dx f_0(x|z_1^{kl}(y_1^{kl})) ln f_0(x)$$

$$+ min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})} \right) min\left(1, \frac{\lambda_l}{a_{i,l}(y_1^{kl})} \right) \int_R dx f_0(x|z_1^{kl}(y_1^{kl})) ln f_0(x|z_1^{kl}(y_1^{kl})). \quad (A.31)$$

Let us define,

$$g_{j,l}(y_1^{kl}) \stackrel{\Delta}{=} min\left(1, \frac{\lambda_l}{a_{j,l}(y_1^{kl})} \right) \quad (A.32)$$

36

Then, from (A.31), and for r_0 and ρ_m as in the proof of Lemma 3, we obtain:

$$\int_R dx q_j^*(x,y_1^{kl}) \ln q_i^*(x,y_1^{kl}) \geq -\frac{1}{2} \{ g_{i,l}(y_1^{kl}) g_{j,l}(y_1^{kl}) + [1-g_{i,l}(y_1^{kl})][1-g_{j,l}(y_1^{kl})] \}$$

$$-\frac{1}{2} \{1-g_{i,l}(y_1^{kl})\} \ln 2\pi r_0^2 - \frac{1}{2} g_{i,l}(y_1^{kl}) \ln 2\pi \rho_{kl}^2 -$$

$$-\frac{1}{2\rho_{kl}^2}[r_0^2 + m_o^2(z_1^{kl}(y_1^{kl}))]g_{i,l}(y_1^{kl}) - \frac{1}{2r_0^2}[\rho_{kl}^2 + m_o^2(z_1^{kl}(y_1^{kl}))]g_{j,l}(y_1^{kl})$$

$$+\frac{1}{2}\{ \frac{r_0^2}{\rho_{kl}^2} + \frac{\rho_{kl}^2}{r_0^2} + \left[\frac{1}{\rho_{kl}^2} + \frac{1}{r_0^2} \right] m_o^2(z_1^{kl}(y_1^{kl})) \} g_{i,l}(y_1^{kl}) g_{j,l}(y_1^{kl}) \qquad (A.33)$$

Define,

$$G_{k,l}(y_1^{kl}) \overset{\Delta}{=} k^{-1} \sum_{j=0}^{k-1} g_{j,l}(y_1^{kl}) \qquad (A.34)$$

From (A.30) and (A.33), we then obtain:

$$\int_R q^*(x,y_1^{kl}) \ln q^*(x,y_1^{kl}) \geq -\frac{1}{2} G_{k,l}^2(y_1^{kl}) - \frac{1}{2}\left[1-G_{k,l}(y_1^{kl}) \right]^2$$

$$-\frac{1}{2}\left[1-G_{k,l}(y_1^{kl}) \right] \ln 2\pi r_0^2 - \frac{1}{2} G_{k,l}(y_1^{kl}) \ln 2\pi \rho_{kl}^2$$

$$-\frac{1}{2}\left[\sigma_{kl}^{-2} + \rho_{kl}^{-2} m_o^2(z_1^{kl}(y_1^{kl})) \right] G_{k,l}(y_1^{kl})$$

$$-\frac{1}{2}\left[\sigma_{kl}^2 + r_0^{-2} m_o^2(z_1^{kl}(y_1^{kl})) \right] G_{k,l}(y_1^{kl})$$

$$+\frac{1}{2}[\sigma_{kl}^{-2} + \sigma_{kl}^2 + \rho_{kl}^{-2}(1 + \sigma_{kl}^2) m_o^2(z_1^{kl}(y_1^{kl}))]G_{k,l}^2(y_1^{kl})$$

$$= -\frac{1}{2}[1 + ln2\pi\rho_{kl}^2] + [1 - G_{k,l}(y_1^{kl})]ln\sigma_{kl}$$

$$-\frac{1}{2}G_{k,l}(y_1^{kl})[1 - G_{k,l}(y_1^{kl})][\sigma_{kl}^{-2} + \sigma_{kl}^2 + r_0^{-2}(1 + \sigma_{kl}^{-2})m_o^2(z_1^{kl}(y_1^{kl})) - 2] \qquad \text{(A.35)}$$

; where,

$$\sigma_{kl} \overset{\Delta}{=} r_0^{-1}\rho_{kl} \qquad \text{(A.36)}$$

Applying (A.35) to (A.29) we obtain the result.

38

REFERENCES

[1]   Berger, T. (1971), Rate Distortion Theory. A Mathematical Basis for Data Compression, Prentice-Hall, Englewood Cliffs, New Jersey.

[2]   Hampel, F.R. (1971), "A General Qualitative Definition of Robustness," Annals of Math. Stat., 42, 1887-1896.

[3]   Papatoni-Kazakos, P. and Gray, R.M. (1979), "Robustness of Estimators on Stationary Observations," Annals of Probability, 7, 989-1002.

[4]   Papantoni-Kazakos, P. (1987), "Qualitative Robustness in Time-Series," Information and Computation, 72, 239-269.

[5]   Papantoni-Kazakos, P. (1984), "Some Aspects of Qualitative Robustness in Time Series," Robust and Nonlinear Time Series Analysis, in Lecture Notes in Statistics, 26, 218-230, Springer-Verlag.

[6]   Papantoni-Kazakos, P. (1981), "Sliding Block Encoders that are Rho-Bar Continuous Functions of Their Input," IEEE Trans. Inform. Th., IT-27, 372-376.

DISTRIBUTION LIST

0352:ald:270R

END
DATE
FILMED
JAN
1988