SEMI-PARAMETRIC COMPRESSION OF PIECEWISE SMOOTH FUNCTIONS

Varit Chaisinthop and Pier Luigi Dragotti

Electrical and Electronic Engineering
Imperial College London, Exhibition Road, London SW7-2AZ, UK
{varit.chaisinthop02,p.dragotti}@imperial.ac.uk

ABSTRACT

This paper introduces a new wavelet-based compression scheme that combines the use of linear approximation and parametric estimation. Our proposed scheme differs from the conventional wavelet-based schemes in two ways: first, the encoder uses linear approximation and second, the decoding process is non-linear as it is combined with parametric estimation. We consider a simple model of one-dimensional (1-D) piecewise smooth function and show that, with our scheme, it is possible to achieve the same decay in the distortion-rate bound as conventional wavelet-based schemes that employ non-linear approximation. A practical compression algorithm that achieves the distortion bound and uses the new concept of sampling of signal with finite rate of innovation is also presented together with the simulation results.

1. INTRODUCTION

Wavelet theory has had a profound impact on modern signal processing theory, particularly in the area of signal approximation and compression (see [1] for reviews). The performance of wavelets in compression, in relationship to their approximation properties, has been extensively studied by means of mathematical analysis [2]. Usually, the analysis involves an important class of signals which includes functions with finite degree of smoothness. The efficiency of a compression scheme is normally measured by the decay rate of its distortion-rate curve $D(R)$.

It is well known that non-linear approximation of 1-D piecewise smooth functions using a wavelet basis provides superior estimates to linear approximation with the same basis. Let us first assume that we have a smooth continuous function that satisfies a uniform Lipschitz condition of order α. Then, the distortion-rate bound of a wavelet-based compression scheme that allocates the bits using a linear approximation strategy can be shown to be [2]:

$$D(R) \leq c_1 R^{-2\alpha}.$$

If we now assume that a function is piecewise smooth with a finite number of discontinuities and each piece is α-Lipschitz. The distortion of the same wavelet-based compression scheme now follows [2]:

$$D(R) \leq c_2 R^{-2\alpha} + c_3 R^{-1}, \quad (1)$$

where the decay of R^{-1} is due to the discontinuities.

In contrast, for the same piecewise smooth function, the distortion of a wavelet-based compression scheme that employs a non-linear approximation strategy is given by [2]:

$$D(R) \leq c_4 R^{-2\alpha} + c_5 \sqrt{R} 2^{-c_6 \sqrt{R}}. \quad (2)$$

It is clear that at high rates, the decay of R^{-1} dominates in (1) whereas $R^{-2\alpha}$ dominates in (2). Therefore, a compression scheme that allocates the bits in accordance with a non-linear approximation strategy performs better as the distortion decays faster.

It has also been shown, however, that a better approximation result does not necessarily lead to better compression algorithms. In this paper, we show that, for piecewise smooth functions, a compression scheme that uses linear approximation can indeed achieve the same decay in the distortion-rate function as the schemes that use non-linear approximation. This is achieved by combining the parametric estimation procedure into the decoding process, which allows the decoder to estimate the locations of the discontinuities. As a result, the effect of discontinuities is removed from the distortion rate bound. This "semi-parametric" decoding is in fact a non-linear process.

In the next section, we present our model of the signal as well as a compression strategy. We then compute a theoretical estimate of the distortion-rate function of the proposed compression scheme in Section 3. A constructive compression algorithm is then presented in Section 4 followed by simulation results in Section 5. Section 6 presents simulation results of an extended compression algorithm that covers a wider class of signals. Finally, conclusions are drawn in Section 7.

2. MODELING AND COMPRESSION STRATEGY

2.1 Signal Model

We consider a simplified model of a piecewise smooth function $f(t)$, $t \in (0,1)$, consisting of two components; a step function and a smooth function with α-Lipschitz type behavior. It is therefore possible to write $f(t)$ as:

$$f(t) = s(t) + f_\alpha(t), \quad (3)$$

where $s(t)$ is a step function:

$$s(t) = \begin{cases} 0 & t < t_0, \\ A & t \geq t_0 \end{cases} \quad (4)$$

and $f_\alpha(t)$ is a α-Lipschitz smooth function, which can be written as:

$$f_\alpha(t) = p_\alpha(t) + \varepsilon_\alpha(t).$$

The function $p_\alpha(t)$ is a polynomial of degree $m = \lfloor \alpha \rfloor$ and $\varepsilon_\alpha(t)$ is such that:

$$\forall t \in (0,1) \text{ and } \forall \nu \in [0,1], \quad |\varepsilon_\alpha(t)| \leq K|t-\nu|^\alpha$$

with a constant $K > 0$. Thus, in this simplified model, the discontinuity is represented by $s(t)$ and the smooth pieces are represented by $f_\alpha(t)$. Note, however, that the following analysis can be generalized for functions whose discontinuity is represented by a piecewise polynomial function.

2.2 Linear Approximation

Let $\psi(t)$ be a wavelet function and let $\varphi(t)$ be the corresponding scaling function. The expansion of a function $f(t)$, $t \in \mathbb{R}$, is then given by:

$$f(t) = \sum_n c_{M,n} \varphi_{M,n}(t) + \sum_{m=-\infty}^{M} \sum_n d_{m,n} \psi_{m,n}(t),$$

where the coefficients $\{d_{m,n}\}$ are given by the inner product:

$$d_{m,n} = \langle f(t), \psi_{m,n}(t) \rangle$$

with the basis element:

$$\psi_{m,n}(t) = 2^{-m/2}\psi\left(2^{-m}t - n\right), \quad m,n \in \mathbb{Z}$$

and similarly for $\{c_{M,n}\}$ and $\varphi_{M,n}(t)$. Here, the low and high-pass coefficients are represented by $\{c_{M,n}\}$ and $\{d_{m,n}\}$ respectively. The N-term linear approximation of $f(t)$, assuming N is large, can be obtained by keeping all the coefficients from level J onwards i.e.:

$$\hat{f}(t) = \sum_n c_{M,n}\varphi_{M,n}(t) + \sum_{m=J}^{M}\sum_n d_{m,n}\psi_{m,n}(t). \tag{5}$$

2.3 Compression Strategy

Intuitively, from (3), one can approximate $f(t)$ by estimating the step function $s(t)$ and the smooth function $f_\alpha(t)$ separately. While a linear wavelet approximation procedure can be used to approximate $f_\alpha(t)$, the step function $s(t)$ can instead be retrieved by estimating the location t_0 and the amplitude A. The reconstruction of $s(t)$ is, therefore, a parametric estimation problem.

Let us divide the high-pass coefficients $\{d_{m,n}\}$ into two sets; $n \in N_s$ and $n \in N_\alpha$. The coefficients in the set N_s are influenced by the discontinuity of $s(t)$ and are said to be in the 'cone of influence'. The coefficients in N_α are, instead, outside of the cone of influence and the wavelet coefficients decay as $d_{m,n} \approx 2^{m(\alpha+1/2)}$ [1]. We now propose the following 'semi-parametric' compression scheme.

Algorithm 1 *The semi-parametric compression scheme for a function $f(t)$ described by equation (3) is as follows:*

1. ***N-term linear approximation****: the encoder approximates $f(t)$ as shown in (5).*
2. ***Quantization****: the coefficients $\{c_{M,n}\}$ and $\{d_{m,n}\}_{m\in(J,M)}$ are uniformly quantized and transmitted.*
3. ***Parametric estimation****: the decoder estimates the location t_0 and the amplitude A of $s(t)$ from the received quantized coefficients $\{\bar{c}_{M,n}\}$ and $\{\bar{d}_{m,n}\}_{m\in(J,M)}$.*
4. ***Cone of influence prediction****: the decoder predicts the coefficients $\{d_{m,n}\}_{m\in(-\infty,J-1)}$ as follows:*

$$\hat{d}_{m,n} = \begin{cases} d^s_{m,n} & n \in N_s, \\ 0 & n \in N_\alpha, \end{cases} \qquad m = -\infty,\ldots,J-1,$$

   where $\{d^s_{m,n}\}$ are the wavelet coefficients of the reconstructed step function $\hat{s}(t)$.

5. ***Final reconstruction****: $f(t)$ is reconstructed from the inverse wavelet transform of the coefficients $\{\bar{c}_{M,n}\}$, $\{\bar{d}_{m,n}\}_{m\in(J,M)}$ and $\{\hat{d}_{m,n}\}_{m\in(-\infty,J-1)}$.*

By using parametric estimation, the discontinuity effect of the step function is removed by the decoder. This is because the decoder is able to 'predict' the coefficients in the cone of influence, which decays slowly otherwise, from the reconstructed step function $\hat{s}(t)$. The estimate of the distortion-rate bound of the proposed algorithm will be computed in the next section.

3. RATE-DISTORTION ANALYSIS

3.1 Bounds on Parametric Estimation of Step Function

We start by considering the following parametric estimation problem. Given a step function $s(t)$, which can be described by (4). Assuming the amplitude A is known [1], the estimator (or the decoder) has to estimate the location t_0 from a set of noisy coefficients $\{\hat{y}_n\}$ given by:

$$\hat{y}_n = y_n + \varepsilon_n, \quad n = 0,1\ldots,N-1,$$

where $\{y_n\}$ is a set of low-pass and high-pass coefficients of $s(t)$ and $\{\varepsilon_n\}$ represents independent and identically distributed (i.i.d.) additive Gaussian noise with zero mean and variance σ_ε^2.

Our aim here is to determine the Cramér-Rao bound (CRB) for the above estimation problem, which will be used in the derivation of the distortion-rate bound in the sequel. The CRB gives us the lower bound on the variance of an unbiased estimator, which defines the best possible accuracy among all unbiased methods. If we have a vector of K deterministic parameters $\Theta = (\theta_1, \theta_2, \ldots, \theta_K)$, then we have that $CRB(t_0) \leq \mathrm{E}\left[\left(\hat{\Theta}-\Theta\right)\left(\hat{\Theta}-\Theta\right)^T\right]$, where $\hat{\Theta}$ is obtained from any unbiased estimation procedure and $E[\cdot]$ denotes an expectation operator. The CRB can be calculated from the inverse of the Fisher Information Matrix $I(\Theta)$ as:

$$CRB(\Theta) = I^{-1}(\Theta) = \left(\mathrm{E}\left[\nabla l(\Theta)\nabla l(\Theta)^T\right]\right)^{-1},$$

where $l(\Theta)$ is the log-likelihood function. With our current estimation problem setup, it is possible to show that:

$$CRB(\Theta) = \sigma_\varepsilon^2\left(\sum_{n=0}^{N-1}\nabla y_n \nabla y_n^T\right)^{-1} \overset{(a)}{=} \sigma_\varepsilon^2\left(\sum_{n=0}^{N-1}\frac{\partial y_n}{\partial t_0}^2\right)^{-1}, \tag{6}$$

where (a) follows from the fact that $\Theta = t_0$ and $\nabla y_n = \left(\frac{\partial y_n}{\partial\theta_1}, \frac{\partial y_n}{\partial\theta_2}, \ldots, \frac{\partial y_n}{\partial\theta_K}\right)$.

If, for example, $\varphi(t)$ is a B-spline of order $P \geq 0$ given by:

$$\varphi(t) = \frac{1}{P!}\sum_{l=0}^{P+1}\binom{P+1}{l}(-1)^l(t-l)_+^P$$

with:

$$(t)_+^P = \begin{cases} 0 & t < 0, \\ t^P & t \geq 0, \end{cases}$$

then, using (6), it can be shown that:

$$CRB(t_0) = \sigma_\varepsilon^2\frac{C_P 2^M}{A^2} = C_P 10^{-0.1PSNR}2^M, \tag{7}$$

where $PSNR = 10\log_{10}\frac{A^2}{\sigma_\varepsilon^2}$ is the peak signal-to-noise ratio and C_P is a constant.

3.2 D(R) Estimate for Semi-Parametric Compression

We now derive the distortion-rate bound of our proposed semi-parametric compression scheme. Here, the function $f(t)$ is given by our signal model shown in (3). The decoder described in Algorithm 1 essentially reconstructs $s(t)$ and $f_\alpha(t)$ separately since the two functions are independent. Therefore, the total distortion D is given by:

$$D = D_\alpha + D_s, \tag{8}$$

where D_α and D_s are the distortion from the reconstruction of the smooth function $f_\alpha(t)$ and the step function $s(t)$ respectively. Since $f_\alpha(t)$ does not contain any discontinuity, assuming that the wavelet basis has at least $\lfloor\alpha\rfloor$ vanishing moments, a compression scheme that uses N-term wavelet linear approximation gives:

$$D_\alpha(R_1) \leq c_7 R_1^{-2\alpha}, \tag{9}$$

where R_1 is the total rate (in bits) allocated to represent $f_\alpha(t)$. Our next task is then to estimate D_s.

Assuming that the amplitude A is known and that the decoder only uses the low-pass coefficients to estimate t_0, let $\{y_n\}_{n\in(0,1,\ldots,N-1)}$ denotes a set of low-pass coefficients of $f(t)$, which can be expressed as:

$$\begin{aligned} y_n &= \left\langle f(t), \varphi_{M,n}(t)\right\rangle \\ &\overset{(a)}{=} \left\langle s(t), \varphi_{M,n}(t)\right\rangle + \left\langle f_\alpha(t), \varphi_{M,n}(t)\right\rangle \\ &= y_n^s + y_n^\alpha, \end{aligned}$$

[1] The assumption that the amplitude A is known at the decoder is for the sake of clarity in our analysis, however, similar results also apply if both A and t_0 have to be estimated by the decoder.

where (a) follows from (3). We can then write the quantized coefficients as:

$$\bar{y}_n = y_n + \varepsilon_n^q = y_n^s + y_n^\alpha + \varepsilon_n^q = y_n^s + \varepsilon_n^s, \quad (10)$$

where ε_n^q represents the quantization noise. Therefore, the quantized coefficients $\bar{y}_n$ can be written as the coefficients of the step function y_n^s plus the *noise* term $\varepsilon_n^s = \varepsilon_n^q + y_n^\alpha$.

Let L be the number of coefficients that the decoder uses to estimate t_0. Since a uniform scalar quantizer is used, the variance σ_q^2 of the quantization noise $\{\varepsilon_n^q\}_{n\in(0,1,\dots,L-1)}$ is given by:

$$\sigma_q^2 = C2^{-2R_2/L}, \quad (11)$$

where C is a constant and R_2 is the total rate allocated to represent $\{y_n\}_{n\in(0,1,\dots,L-1)}$.

In order to compute D_s, we also make the following assumptions: the probability density function (PDF) of y_n^α is Gaussian with zero mean [2] and variance σ_α^2; both ε_n^q and y_n^α are independent, which implies ε_n^s is Gaussian distributed with zero mean and variance σ_ε^2, where:

$$\sigma_\varepsilon^2 = \left(\sigma_q^2 + \sigma_\alpha^2\right), \quad (12)$$

and, finally, we assume that the estimator of t_0 is a minimum variance estimator that achieves the CRB.

Let us denote the reconstructed step function with $\hat{s}(t)$ where:

$$\hat{s}(t) = \begin{cases} 0 & t < t_0 + \varepsilon_t, \\ A & t \geq t_0 + \varepsilon_t. \end{cases}$$

Here the error in the estimation of t_0 is represented by ε_t whose variance σ_t^2 is given by the CRB of the estimator shown in (6), where, from (10), σ_ε^2 is given by (12). It then follows that the mean square error, $MSE\left(\hat{s}(t)\right)$, is given by:

$$MSE\left(\hat{s}(t)\right) = \int \left(s(t) - \hat{s}(t)\right)^2 dt = \int_t^{t+\varepsilon_t} A^2 dt = A^2|\varepsilon_t|.$$

Thus, the expected distortion D_s can be computed by:

$$D_s = \mathrm{E}\left[A^2|\varepsilon_t|\right] = A^2\mathrm{E}\left[|\varepsilon_t|\right],$$

where $\mathrm{E}\left[|\varepsilon_t|\right]$ is the mean absolute deviation of ε_t. Using Jensen's inequality for concave functions, we can show that:

$$\mathrm{E}[|\varepsilon_t|] = \mathrm{E}\left[\sqrt{\left(\varepsilon_t - \mathrm{E}[\varepsilon_t]\right)^2}\right] \leq \sqrt{\mathrm{E}\left[\left(\varepsilon_t - \mathrm{E}[\varepsilon_t]\right)^2\right]} = \sigma_t$$

as $\mathrm{E}[\varepsilon_t] = 0$. Therefore, the expected distortion can be expressed as:

$$D_s \leq A^2\sigma_t = A^2\sqrt{CRB(t_0)}.$$

By using the expression for the CRB in (6) together with the relationship given in (10), we arrive at the following distortion-rate bound for the estimation of the step function:

$$\begin{aligned} D_s(R_2) &\leq A^2\left(\sigma_\varepsilon^2\right)^{\frac{1}{2}} \underbrace{\left(\sum_{n=0}^{L-1}\left(\frac{\partial y_n^s}{\partial t_0}\right)^2\right)^{-\frac{1}{2}}}_{=C} \\ &\overset{(a)}{=} c_8\left(\sigma_q^2 + \sigma_\alpha^2\right)^{\frac{1}{2}} \\ &\overset{(b)}{=} c_8\left(c_9 2^{\frac{-2R_2}{L}} + \sigma_\alpha^2\right)^{\frac{1}{2}}, \end{aligned} \quad (13)$$

[2]The PDF of y_n^α is arbitrarily assumed to be zero-mean Gaussian as this allows us to use the analytical expression of the CRB, given by (6), in the derivation of the D(R) curve (as shown in (13)). The derived distortion-rate bound is then verified with simulations.

where (a) and (b) follow from substituting in (12) and (11) respectively. The expression for the total distortion-rate bound can now be obtained by substituting (9) and (13) into (8), which gives:

$$D(R) \leq c_7 R_1^{-2\alpha} + c_8\left(c_9 2^{\frac{-2R_2}{L}} + \sigma_\alpha^2\right)^{\frac{1}{2}}, \quad (14)$$

where the total rate R, in bits, is equal to:

$$R = R_1 + R_2. \quad (15)$$

We now need to consider the following bit allocation problem; given a total rate R, how should we allocate the bits among R_1 and R_2 so that the distortion in (14) is minimized subject to the constraint in (15). This is a well known constrained optimization problem, which can be solved using a Lagrange multiplier method. One necessary condition for the optimal bit allocation is that the derivatives of the distortion D with respect to R_1 and R_2 must be equal i.e.:

$$\frac{\partial D}{\partial R_1} = \frac{\partial D}{\partial R_2}. \quad (16)$$

Assuming a high rate regime, where $c_9 2^{\frac{-2R_2}{L}} < \sigma_\alpha^2$, the distortion given in (14) can be approximated with a Taylor series expansion of the square root function as follows:

$$D(R) \approx c_7 R_1^{-2\alpha} + \frac{c_8 c_9}{2\sigma_\alpha} 2^{\frac{-2R_2}{L}} + c_8\sigma_\alpha. \quad (17)$$

By solving the equal gradient condition in (16), where D is approximately given by (17), we obtain the following rate allocation:

$$R_2 = \frac{L}{2}(2\alpha+1)\log_2 R_1 + C' \quad (18)$$

with a constant C'. The total rate R is thus given by:

$$R = R_1 + \frac{L}{2}(2\alpha+1)\log_2 R_1 + C' \approx R_1 \quad (19)$$

at high R. Therefore, by substituting (18) into (17) and using the approximation shown in (19), the overall distortion-rate function of our semi-parametric compression scheme described in Algorithm 1 is:

$$D(R) \leq c_7 R^{-2\alpha} + c_{10} R^{-(2\alpha+1)} + c_8\sigma_\alpha. \quad (20)$$

Note that the term $c_{10}R^{-(2\alpha+1)}$ represents the distortion caused by the discontinuity, which decays faster than the distortion from the encoding of the smooth function.

In comparison to a compression scheme that only uses a linear approximation strategy, where the distortion decays as R^{-1} (see (1)), our proposed scheme achieves a faster decay rate of $R^{-2\alpha}$ in the high rate regime. Furthermore, the dominating decay rate of $R^{-2\alpha}$ is comparable to a compression scheme that employs nonlinear wavelet approximation as shown in (2). Note, however, that the term $c_8\sigma_\alpha$ represents the systematic error in the modeling of our parametric estimation problem.

4. CONSTRUCTIVE COMPRESSION ALGORITHM

We start this section by introducing a practical parametric estimation technique based on the recently developed concept of sampling of signals with finite rate of innovation (FRI). FRI signals are, loosely speaking, a class of signals or functions $f(t)$ that can be described by a finite number of free parameters over a given interval $t \in [t_a, t_b]$. The definition and sampling schemes of FRI signals are given in details in [3, 4]. Clearly, a step function $s(t)$ also belongs to this class of function as it can be completely described by at most two parameters, the location t_0 and the amplitude A.

Let us present one of the key results from the sampling schemes of FRI signals described in [3, 4]. Given a function $f(t)$ and a scaling function $\varphi(t)$. In the context of sampling, the samples or the coefficients are given by:

$$y_n = \langle f(t), \varphi(t/T - n) \rangle,$$

where T is the sampling period. Assume that $\varphi(t)$ together with its shifted versions can reproduce polynomials of maximum degree P i.e. $\varphi(t)$ satisfies:

$$\sum_{n \in \mathbb{Z}} c_n^p \varphi(t/T - n) = t^p \quad p = 0, 1, ..., P \tag{21}$$

for a proper set of coefficients $\{c_n^p\}$. The polynomial reproduction coefficients can be calculated as:

$$c_n^p = \langle t^p, \tilde{\varphi}(t/T - n) \rangle,$$

where $\tilde{\varphi}(t)$ is the dual of $\varphi(t)$. It then follows that the continuous moment m_p of order p of the signal $f(t)$ is given by:

$$\begin{aligned} m_p &= \int f(t) t^p dt \\ &= \left\langle f(t), \sum_{n \in \mathbb{Z}} c_n^p \varphi(t/T - n) \right\rangle \\ &= \sum_{n \in \mathbb{Z}} c_n^p \langle f(t), \varphi(t/T - n) \rangle \\ &= \sum_{n \in \mathbb{Z}} c_n^p y_n. \end{aligned} \tag{22}$$

Therefore, given $\{c_n^p\}$, one can retrieve the continuous moments of $f(t)$ from the coefficients $\{y_n\}$ provided that $f(t)$ lies in the region where the condition given by (21) is satisfied.

In addition, a sampling scheme for a piecewise constant signal was presented in [4]. Let $z_n^{(1)}$ denotes a finite difference $y_{n+1} - y_n$. It was shown that:

$$z_n^{(1)} = y_{n+1} - y_n = \left\langle \frac{df(t)}{dt}, \varphi(t/T - n) * \beta_0(t/T - n) \right\rangle,$$

where $\beta_0(t)$ denotes a zeroth order B-spline function. Therefore, from (22), we have that:

$$m'_p = \sum_{n \in \mathbb{Z}} c'^p_n z_n, \tag{23}$$

where m'_p is the p-th order continuous moment of $\frac{df(t)}{dt}$ and $\{c'^p_n\}$ are the polynomial reproduction coefficients of the new scaling function $\varphi(t/T - n) * \beta_0(t/T - n)$.

Assuming a step function $s(t)$ as shown in (4), it is easy to show that the zeroth and first order moments of $\frac{ds(t)}{dt}$ are given by $m'_0 = A$ and $m'_1 = At_0$. Hence, the location t_0 can be calculated as:

$$t_0 = \frac{m'_1}{m'_0}. \tag{24}$$

We now present a FRI-based parametric estimation algorithm that uses the results in (23) and (24).

Algorithm 2 *Given a set of noisy low-pass coefficients $\{\bar{c}_{M,n}\}$ of a step function $s(t)$, where $\bar{c}_{M,n} = \langle s(t), \varphi_{M,n}(t) \rangle + \varepsilon_n^s$ and $\varphi(t)$ satisfies the polynomial reproduction condition in (21), the location of the step t_0 can be estimated as follows:*

1. ***Finite difference**: the finite difference is obtained as:*

$$\bar{z}_n = \bar{c}_{M,n+1} - \bar{c}_{M,n}.$$

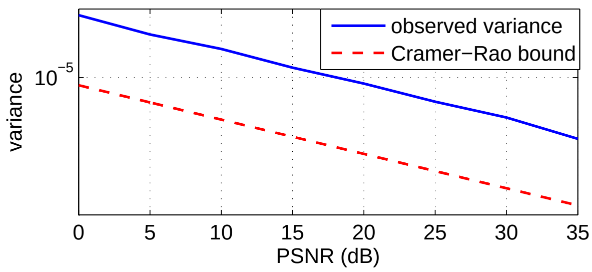

Figure 1: Plots of the observed variance (log scale) in the estimation of t_0 and the corresponding CRB. The estimator uses the FRI-principle described in Algorithm 2.

2. ***Moments estimation**: the zeroth and first order moments of $\frac{ds(t)}{dt}$ are estimated as:*

$$\bar{m}'_p = \sum_{n \in \mathbb{Z}} c'^p_n \bar{z}_n \qquad p = 0, 1.$$

3. ***Location estimation**: the location of the step is approximated as:*

$$\bar{t}_0 = \frac{\bar{m}'_1}{\bar{m}'_0}.$$

Therefore, a practical semi-parametric compression scheme can be constructed using Algorithm 1, where the parametric estimation step is implemented with Algorithm 2. This allows the decoder to approximate the step function from the quantized low-pass coefficients $\{\bar{c}_{M,n}\}$. The simulation results of this compression scheme are shown in the next section.

5. SIMULATION RESULTS

Let us first compare the variance of an estimator that uses the FRI principle presented in the previous section to the CRB described in section 3. In our simulation, we applied a ten-level bi-orthogonal 2.2 wavelet decomposition to the step function $s(t)$, where the scaling function $\varphi(t)$ is given by the first order B-spline. Gaussian noise was then added to the low-pass coefficients $\{c_{M,n}\}$. The value of the amplitude A was assumed to be known and the location t_0 was estimated using Algorithm 2. Figure 1 shows the plots of the observed variance and the CRB, which is given by (7). It is clear that the decays on both plots are the same even though the estimator does not achieve the bound. Thus, the squared error of the estimate of t_0 decays as $10^{-0.1PSNR}$.

We now present the simulation results of the proposed semi-parametric compression scheme described in Algorithm 1, where the parametric estimation step is implemented with Algorithm 2. The piecewise smooth function was generated with the model described by (3), where the degree of smoothness was set to $\alpha = 1.75$. Once again, we used the bi-orthogonal 2.2 wavelet decomposition with a first order B-spline scaling function. The rates were then allocated in accordance with (18). Figure 2 shows the distortion-rate plot of our proposed semi-parametric compression scheme in comparison with the plot for a scheme based on linear approximation. At high rates, our scheme achieves a decay rate of $R^{-2\alpha} = R^{-3.5}$, which is in line with our analysis in section 3. In contrast, the linear approximation based scheme decays as R^{-1}. Therefore, we were able to achieve the same decay in the distortion-rate curve as a compression scheme that employs a non-linear approximation strategy. Note also that the systematic error in the parametric estimation was insignificant in this simulation. Finally, the reconstructed functions are shown in Figure 3.

6. EXTENSION TO WIDER CLASS OF SIGNALS

The distortion-rate analysis presented in this paper can be generalized to cover a wider, less restrictive, class of signals whose dis-

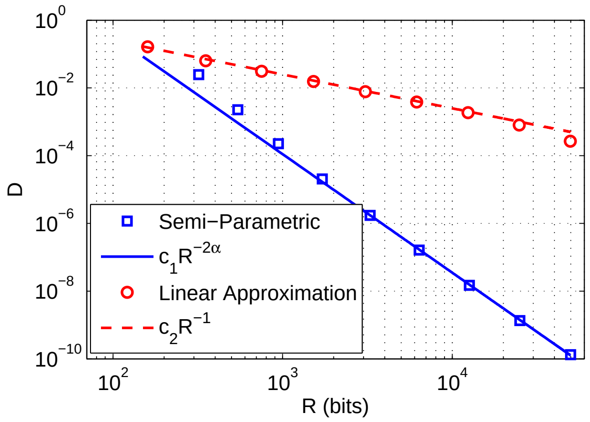

Figure 2: Distortion-rate plots (log scale); at high rates, the proposed semi-parametric compression scheme has the decay rate of $R^{-2\alpha}$ whereas the distortion of a linear approximation based scheme decays as R^{-1}.

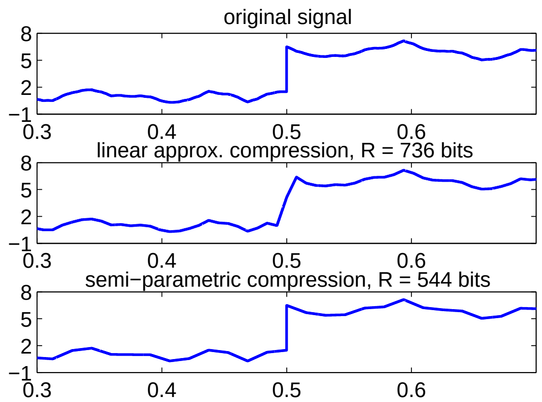

Figure 3: Plots of the original signal, the reconstructed signals with linear approximation scheme at $R = 736$ bits and semi-parametric scheme at $R = 544$ bits.

continuity is described by a piecewise polynomial function rather than a step function. The bound on the performance of the optimal compression scheme for piecewise polynomial signals was derived in [5]. By using the results from the FRI-based sampling scheme of piecewise polynomial signals presented in [4], the compression strategy shown in section 4 can be extended to include signals with piecewise polynomial discontinuity. The details are omitted here due to a limited space.

Figure 4 and Figure 5 show the simulation results of the extended semi-parametric compression scheme, where the signal was constructed from a piecewise quadratic function and a smooth α-Lipschitz function. At high rates, the modified scheme also achieves a decay rate of $R^{-2\alpha}$.

7. CONCLUSION

In this paper, we have introduced a new wavelet-based compression scheme for piecewise smooth functions, where the decoder uses a parametric estimation technique. While the encoding process is based on a linear approximation strategy, our distortion-rate analysis and simulation results have shown that the proposed semi-parametric compression scheme can achieve the same decay in the distortion-rate curve as conventional wavelet-based schemes that employ a non-linear approximation strategy.

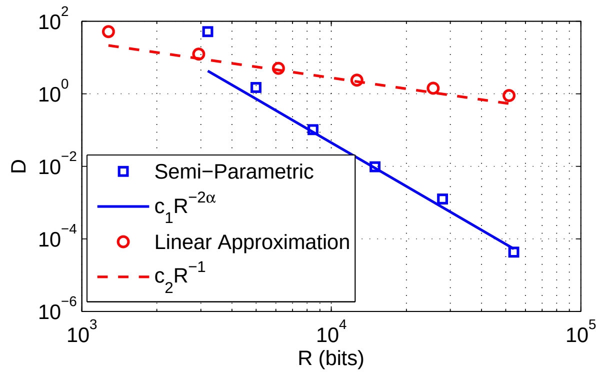

Figure 4: Distortion-rate plots (log scale); at high rates, the extended semi-parametric compression scheme for signals with piecewise polynomial discontinuity also achieves the decay rate of $R^{-2\alpha}$.

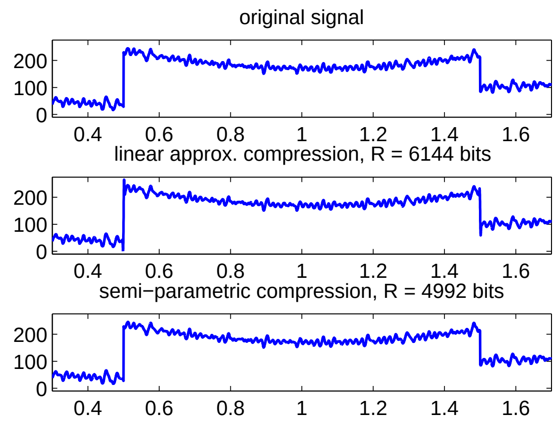

Figure 5: Plots of the original signal, the reconstructed signals with linear approximation scheme at $R = 6144$ bits and extended semi-parametric scheme at $R = 4992$ bits.

REFERENCES

[1] Martin Vetterli, "Wavelets, approximation, and compression," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 59–73, 2001.

[2] A. Cohen, I. Daubechies, O.G. Guleryuz, and M.T. Orchard, "On the importance of combining wavelet-based nonlinear approximation with coding strategies," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1895–1921, Jul 2002.

[3] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, 2002.

[4] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Transactions on Signal Processing*, vol. 55(5), pp. 1741–1757, May 2007.

[5] A. Maleki and G. Carlsson, "ε-entropy of piecewise polynomial functions and tree partitioning compression," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1181–1184, April 2008.