Near Optimal Rational Approximations of Large Data Sets

by

Anil Damle

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Masters of Science

Department of Applied Mathematics

2011

This thesis entitled:
Near Optimal Rational Approximations of Large Data Sets
written by Anil Damle
has been approved for the Department of Applied Mathematics

_____

Gregory Beylkin

_____

Lucas Monzón

_____

James Curry

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Damle, Anil (M.S., Applied Mathematics)

Near Optimal Rational Approximations of Large Data Sets

Thesis directed by Prof. Gregory Beylkin

We introduce a new computationally efficient algorithm for constructing near optimal rational approximations of large data sets. In contrast to wavelet-type approximations often used for the same purpose, these new approximations are effectively shift invariant. On the other hand, when dealing with large data sets the complexity of our current non-linear algorithms for computing near optimal rational approximations prevents their direct use.

By using an intermediate representation of the data via B-splines, followed by a rational approximation of the B-splines themselves, we obtain a suboptimal rational approximation of data segments. Then, using reduction and merging algorithms for data segments, we arrive at an efficient procedure for computing near optimal rational approximations for large data sets. A motivating example is the compression of audio signals and we provide several examples of compressed representations produced by the algorithm.

Acknowledgements

I would like to thank my advisor Gregory Beylkin for his supervision of this thesis. I would also like to thank my committee members Lucas Monzón and James Curry. Finally, I would also like to thank Terry Haut for providing me with code. This research was partially supported by NSF grant DMS-100995, DOE/ORNL grant 4000038129 and NSF MCTP grant DMS-0602284.

My educational experience has been greatly enriched by the relationships I have formed with students and faculty at the University of Colorado, and I would also like to thank all those who have positively impacted my educational experience. Specifically, I would like to thank Anne Dougherty for all her support and guidance as I progressed through my education, and Scot Douglass for the invaluable educational moments my interactions with him have enabled. Finally, I would like to thank my family for their support in all my educational endeavors.

Contents

Chapter

Figures

Figure

Chapter 1

Introduction

The recent development of new algorithms for the construction of near optimal rational approximations of functions, see e.g., [5, 6, 7] for an overview of algorithms and applications, provide a basis for further investigation of near optimal rational approximations. A natural application of these rational approximations is the development of a scheme that compresses and represents signals of long duration. We use sampled music as a motivating example but our approach is not limited to such signals. Signals obtained by continuous monitoring, for example seismic global monitoring, may also be compressed via our approach. We view such compression as the first step in signal analysis where the near optimal rational approximation provides a convenient representation suited for further processing.

The hurdle that must be addressed for the development of such compression scheme is that current algorithms do not scale well to large problems. Current algorithms require the computation of a Singular Value Decomposition (SVD) on a matrix whose size is proportional to the number of samples in a signal. Since a few minutes of audio sampled at 44.1 kHz contains millions of samples, the problem is computationally infeasible. In this thesis we develop a new algorithm for the construction of near optimal rational approximations that scales well to large data sets.

The algorithm first constructs a suboptimal rational representation for the sampled data using a B-spline representation as an intermediate step; then, computes a near optimal rational approximation by taking advantage of the new reduction algorithm in [15]. The resulting combination of algorithms scales well to large problems and, thus, facilitates the development of a compression

scheme for signals of long duration.

While we do not fully develop a compression scheme here, the results show the potential for a scheme which would be based on near optimal rational approximations. A complete compression scheme requires, in addition, quantization and arithmetic coding which we do not discuss here.

The applications of optimal rational approximations extend well beyond their potential for compression. Specifically, we note that the traditional wavelet-type decompositions are not shift invariant, complicating any further processing of the decomposed data. In contrast, optimal rational approximations are shift invariant. Also, as shown in [6], poles of near optimal rational approximations concentrate near the singularities of the signal and the location of the poles carry precise information about local frequency content of the signal. These properties of near optimal rational approximations make them potentially advantageous for the further processing and analysis of signals.

We first describe the overall algorithm and give some historical context for it in Chapter 2. Specifically, we discuss at some length the relation between the optimal and near optimal rational approximations used here and wavelets. The technical details and the full development is in Chapter 3 which contains a version of the paper written by the author with Gregory Beylkin, Lucas Monzón and Terry Haut. Finally, Chapter 4 contains some concluding remarks.

Chapter 2

Preliminaries

2.1 Linear and Nonlinear approximations

Wavelet-type decompositions have been used extensively for the development of compression schemes. The computational cost of wavelet decompositions is linear in the number of signal samples so that wavelet schemes scale well to large problems. For this reason we use wavelets, specifically a scheme based on B-splines, as an intermediate step in our construction. Importantly, the ability of wavelet type decompositions to compress data with a finite number of algebraic type singularities is justified via rational approximations (see e.g. [16]). Additional information may be found in [11, 17], and the references therein.

Many compression schemes are based on nonlinear approximations. In contrast to the so-called linear approximations, where a function is projected onto a pre-determined linear subspace, nonlinear approximations are adaptive in the choice of approximating functions. Wavelets may be used for both, linear and nonlinear approximations. In the case of linear approximations, wavelets are used in exactly the same manner as any other basis. Nonlinear approximations using wavelets, besides computing a wavelet decomposition, arrange the resulting coefficients with respect to their magnitudes and select only the N largest coefficients for the approximation. Such approach results in a much more efficient, adaptive representation of functions.

An early example of a multiresolution style basis that has a structure of a wavelet basis is the Haar basis [14] (circa 1910). However, the basis functions do not have a well concentrated Fourier transform since they are discontinuous. More recent interest in wavelets (see e.g., [11]) was

motivated by the desire to obtain bases with functions well concentrated in both, the space and the Fourier domains. The first construction of a basis of this type may be found in [22]. An important paper connecting wavelet constructions with filter banks [21] (used in electrical engineering) and the so-called Laplacian pyramid [8] (used in image processing) appeared in 1988 [10]. This paper by Daubechies laid the mathematical foundation for using wavelets as a tool for compression. Important results on using wavelets for compression of signals may be found in [12].

Optimal rational approximation is another form of nonlinear approximation. The theoretical considerations for such approximations began to form in the 1960's. Newman in [19] showed that rational approximations resulted in an efficient representation for $f(x) = |x|$, a function with a simple algebraic singularity. In the late 1960's Adamjan, Arov and Krein developed theory for the construction of optimal rational approximations [1, 2, 3] in the L^∞ norm. They associate a function with a Hankel matrix constructed from their Fourier coefficients and show that the L^∞ function norm is equivalent to the L^2 matrix norm for such Hankel matrix. Rational approximations are constructed by finding a singular vector associated with a singular value that determines the accuracy of approximation in the L^∞ norm. The poles of such optimal rational approximation are found as the roots (inside the unit disk) of a polynomial constructed using the entries of the singular vector. When the size of Hankel matrices in AAK theory is large and the desired accuracy requires one to compute its small singular values, this method may not be practical.

In [5] Beylkin and Monzón developed a method for the construction of near optimal exponential approximations for functions defined on an interval by using equally spaced samples of the function. These exponential approximations yield optimal rational approximations for the Fourier transform of the function as shown in [6]. Further work, [7], addresses additional application areas of these approximations. As demonstrated in [6], poles of near optimal rational approximations concentrate near the singularities of functions. The algorithms developed in these papers are sufficient for dealing with relatively small data sets since the complexity of algorithms is $\mathcal{O}\left(N_s^3\right)$, where N_s is the number of samples. This computational complexity prevents the direct use of these algorithms for large data sets.

We note that for signals, the concentration of poles near the singularities corresponds to the locations of rapid change such as occurring when a piano key is struck or at wave arrivals in seismic recordings. The location of the poles also carries information about local frequency content of the signal in a manner similar to wavelets, i.e., the logarithmic distance of these locations from the real axis corresponds to wavelet scales. In contrast to the wavelet decompositions, the optimal rational approximations are shift invariant and it is easy to translate the resulting approximations. In fact, shifting an optimal rational approximation yields the optimal approximation of the shifted function or data.

2.2 Informal Description of Algorithm

An alternative to using the sampled data is to first obtain (by some method) a suboptimal rational approximation and then construct the optimal rational approximation by using an algorithm specifically designed for this purpose. Such an algorithm was introduced in [5] as a reduction algorithm since it reduces the number of necessary poles. An important aspect of the reduction algorithm is that by starting with a suboptimal rational approximation the problem reduces to a con-eigenvalue problem for a Cauchy matrix of size $N \times N$, where N is the number of poles in the suboptimal approximation. A new algorithm by Beylkin and Haut for the reduction problem may be used to compute a near optimal rational approximation in $\mathcal{O}\left(M^2 N\right)$ operations, where M is the number of poles in the near optimal approximation [15]. Furthermore, this algorithm maintains high relative accuracy in computing the con-eigenvalues and high absolute accuracy in computing con-eigenvectors. The computational complexity of this algorithm means that, given a situation where there is a substantial reduction in the number of poles in the suboptimal rational approximation, the algorithm is essentially linear in the number of original poles N.

We set up our problem for computing near optimal rational approximations of large data sets in a way that takes advantage of the computational complexity of this new reduction algorithm. Given as input a large number of equally spaced samples, we first compute a partial wavelet decomposition of the data. Next, we develop a suboptimal rational approximation of data seg-

ments. We then use the reduction algorithm on this representation to obtain a near optimal rational approximation on (slightly overlapping) data segments. Finally, we merge approximations on adjacent segments to obtain the near-optimal approximation of the entire signal.

As the first step, we construct a B-spline decomposition of the signal. We note that B-splines were introduced in the work of [20] and used extensively ever since. With the advent of wavelet-type schemes, there are at least two associated wavelet constructions. There are the so-called Battle-Lemarié wavelets (see e.g. [11, Chapter 5]), and Strmberg wavelets [22]. Further discussions of B-spline representations and wavelets may be found in [9].

Although theoretically we may use any wavelet-type basis for this intermediate step, we found that the choice of B-splines yields a more efficient suboptimal rational approximation. We use a fast algorithm to compute the B-spline coefficients from the signal samples (see [4, 18]). Specifically, for sampled periodic functions, the B-spline coefficients may be evaluated by first computing the FFT of the data, followed by multiplication by a specific function, and then computing the inverse FFT. In order to apply the algorithm, we subdivide our signal using a partition of unity into overlapping windows so that we may treat the function within each window as periodic. The size of windows in the partition of unity should be appropriate for an efficient use of the FFT but otherwise is arbitrary. We compute the B-spline coefficients within each window and then combine the B-spline coefficients from all windows to get B-spline coefficients for the entire signal.

We construct a special suboptimal rational approximation of a B-spline by arranging the poles of the approximation on a rectangular grid aligned with the knots of the B-spline. This ensures that the number of poles in our suboptimal rational approximation of the entire signal is a constant multiple of the number of given samples. In this work we use seventh degree B-splines which require three poles under each knot. The choice of the seventh degree B-splines is dictated by the target accuracy for our final signal approximation. For higher accuracy, higher degree B-splines may be used and similar suboptimal representations may be constructed.

We then split the obtained B-spline coefficients into consecutive segments (which are un-

related to the partition of unity). The size of these segments should be selected for the optimal use of the reduction algorithm [15].Then within each segment we replace the B-splines by their suboptimal rational approximations. This yields a suboptimal rational approximation of a data segment with three times as many poles as the number of original samples. We then apply the algorithm in [15] to reduce the suboptimal rational approximation on each segment and obtain the corresponding near optimal rational approximation.

At this point we already have a representation of the entire signal that may be sufficient for most purposes, such as the first step of a compression scheme. In fact, the final step of our scheme, the merging of representations on adjacent segments, does not reduce the number of poles in a significant manner. On the other hand, if the goal is further processing and analysis of the signal, it is also important to maintain the near optimality of the approximation in the (small) overlap region between segments. In order to merge the near optimal approximations from adjacent segments, we may use the reduction algorithm once again. Only poles near the boundary between adjacent segments carry information about the overlapping region between the segments. For this reason, we need to merge only poles near the boundary between adjacent segments keeping unchanged the poles and the residues away from this boundary region. To accomplish this, we consider the function generated by the poles and associated residues in a small vicinity of the boundary between segments. In the reduction algorithm, we request a slightly higher accuracy across the support of both segments so that we preserve the overall accuracy of the approximation. The slightly higher accuracy assures that the unchanged poles and corresponding residues away from the boundary region are not impacted by this merge and, hence, we do not need to recompute their residues. Although this procedure does not yield the optimal approximation over the support of the two segments, we have observed that the approximation is near optimal, both in terms of the number of poles and their locations.

To test our algorithm, we use samples taken from an audio recording, compute the B-spline coefficients, and then examine the performance of the algorithm on a segment of 768 coefficients. For this segment we were able to achieve roughly three and a half digits of accuracy using only 44

poles. Counting four real numbers per pole, the cost of storing this signal is reduced by a factor of approximately 4.4. We note that even though the same data in the MP3 format yields the same compression factor of 4.4, in our experiment we did not quantize nor use coding to achieve our level of compression. Thus, we achieve (roughly) the same compression as the well developed MP3 format just using the near optimal rational approximation. Further work is required to make more precise comparisons to other compression schemes, but we note that our preliminary results look promising.

Chapter 3

Paper to be Submitted

Near optimal rational approximations of large data sets

Anil Damle, Gregory Beylkin, Lucas Monzón and Terry Haut

Department of Applied Mathematics, University of Colorado, Boulder, CO 80309-0526

This research was partially supported by NSF grant DMS-100995, DOE/ORNL grant 4000038129 and NSF MCTP grant DMS-0602284.

3.1 Abstract

We introduce a new computationally efficient algorithm for constructing near optimal rational approximations of large data sets. In contrast to wavelet-type approximations often used for the same purpose, these new approximations are effectively shift invariant. On the other hand, when dealing with large data sets the complexity of our current non-linear algorithms for computing near optimal rational approximations prevents their direct use.

By using an intermediate representation of the data via B-splines, followed by a rational approximation of the B-splines themselves, we obtain a suboptimal rational approximation of data segments. Then, using reduction and merging algorithms for data segments, we arrive at an efficient procedure for computing near optimal rational approximations for large data sets. A motivating example is the compression of audio signals and we provide several examples of compressed representations produced by the algorithm.

3.2 Introduction

We develop algorithms for constructing near optimal rational representations of functions using as input a large number of equally spaced samples. Examples of such data sets include, among others, digitized versions of a musical recordings and continuous seismic records. Optimal or near optimal rational approximations provide both, a method for data compression as well as a useful representation for further data analysis. Rational approximations are more efficient than wavelet decompositions of data sets and, in fact, the ability of wavelets to compress signals may be justified via optimal rational approximations, see e.g., [16, Chapter 11]. Furthermore, in contrast to the wavelet decompositions, rational functions are closed under translations and the optimal rational approximations are shift invariant. In fact, shifting an optimal rational approximation yields the optimal approximation of the shifted function or data.

For functions given analytically or for functions described by a relatively small number of samples, there are several methods for obtaining their near optimal rational approximations [5, 6, 7]. For large data sets these methods are impractical due to their computational complexity. On the other hand, a wavelet decomposition of a large data set does not present a difficulty since its computational cost is linear in the number of samples. We use these facts in our approach, where we first compute a wavelet decomposition of the data. We then convert the result into near optimal rational approximations on slightly overlapping data segments. Next, we merge approximations on adjacent segments together. We show that the poles and the residues of the rational approximations away from the overlap region are affected only within the approximation accuracy.

For the wavelet decomposition, we use a B-spline basis as it offers a simple and efficient method for a transition to a suboptimal rational representation. For such transition we construct a particular rational approximation of B-splines, where the poles are arranged on a rectangular grid aligned with the location of spline knots. Given a suboptimal rational approximation on a segment, we compute the near optimal approximation using a new algorithm in [15].

We use sampled music as a motivating example but our approach is not limited to such

signals. Signals obtained by continuous monitoring, for example seismic global monitoring, may also be compressed via our approach. We view such compression as the first step in signal analysis where the near optimal rational approximation offers a good starting point. As shown in [6], poles of near optimal rational approximations concentrate near the singularities of functions. For signals, this corresponds to locations of rapid change such as occurring when a piano key is struck or at wave arrivals in seismic recordings. The location of the poles also carries information about local frequency content of the signal in a manner similar to wavelets, i.e., the logarithmic distance of these locations from the real axis corresponds to wavelet scales.

We start in Section 3.3 by providing the background information pertaining to the new algorithm that is presented. Specifically we briefly present key existing algorithms that facilitate the development of our new algorithm. Section 3.5 describes in detail the algorithm for constructing near optimal rational approximations of large data sets. We present an overview of the algorithm and then examine the specifics of each step. Section 3.6 contains numerical examples that validate the performance of the algorithm. Finally, Section 3.7 contains concluding remarks.

3.3 Preliminary Considerations

3.3.1 An algorithm for finding a near optimal rational approximation

We start describing the method in [5] to obtain a near optimal rational approximation samples of the Fourier transform of the function. For functions with a fast decaying Fourier transform this method is closely connected to theory developed by Adamjan, Arov and Krein (AAK) [1, 2, 3].

Given samples $\hat{f}\left(a\frac{k}{2N}\right), \quad k = 0, 1, \ldots, 2N$, (that sufficiently oversample $\hat{f}(\xi)$ on the interval $\xi \in [0, a]$) we seek a representation of $\hat{f}(\xi)$ of the form

$$\hat{f}(\xi) \approx \sum_{m=1}^{M} w_m e^{-\eta_m \xi}. \tag{3.1}$$

The algorithm proceeds as follows:

- Construct a $N + 1 \times N + 1$ Hankel matrix of the form $H_{kl} = \widehat{f}\left(a\frac{k+l}{2N}\right)$.

- Find a singular vector $\mathbf{u} = (u_0, \ldots, u_N)$ that solves the con-eigenvalue problem $\mathbf{Hu} = \sigma\bar{\mathbf{u}}$, with σ selected according to the target accuracy ϵ. We may find $\mathbf{u}$ by solving the eigenvalue problem for

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{0} & \mathbf{H} \\ \mathbf{H}^* & \mathbf{0} \end{bmatrix}, \tag{3.2}$$

  which yields singular values **H,** $\sigma_0 \geq \sigma_1 \geq \cdots \geq \sigma_M \geq \cdots \geq \sigma_N$**.** We choose the singular value σ_M, so that $\frac{\sigma_M}{\sigma_0} \approx \epsilon$. Typically the singular values decay exponentially fast so that $M \ll N$.

- Compute M appropriate roots of the polynomial $u(z) = \sum_{n=0}^{N} u_n z^n$ and denote them γ_m. Given values γ_m, η_m in (3.1) may be recovered as

$$\eta_m = 2N \log \gamma_m, \tag{3.3}$$

  where we use the principle value of the logarithm.

- Compute the weights w_m in (3.1) by solving the least-squares Vandermonde system

$$\sum_{m=1}^{M} w_m \gamma_m^k = \widehat{f}\left(a\frac{k}{2N}\right), \quad 0 \leq k \leq 2N. \tag{3.4}$$

If no a priori information is available, we may use all N roots of $u(z)$, and then determine the M necessary roots by selecting those with corresponding weights of magnitude greater than the target accuracy.

Applying the inverse Fourier transform to (3.1) yields a rational representation

$$f(x) = -2\mathcal{R}e\left(\sum_{m=1}^{M} \frac{w_m}{2\pi i x - \eta_m}\right) = -2\sum_{m=1}^{M} \frac{v_m(2\pi x - \theta_m) - u_m\tau_m}{(2\pi x - \theta_m)^2 + \tau_m^2}, \tag{3.5}$$

where $w_m = u_m + iv_m$ and $\eta_m = \tau_m + i\theta_m$. Note that there is a misprint in equation (5) of [6]. As illustrated in [6] the positions of the poles η_m carry information about the location of singularities of the function f. Furthermore, the representation for translates of (3.5) are readily obtained by simply shifting the poles $\eta_m/2\pi i$.

3.3.2 Algorithm for reduction of a suboptimal rational approximation

An effective and accurate algorithm for reducing the number of poles of a rational function while maintain some target accuracy is given in [15]. The formulation of the problem may be found in [5] and is based on results in [3].

We note that even though we present this algorithm for rational trigonometric functions, a similar algorithm may be used for approximating rational functions defined on a real interval. Indeed, instead of considering as input residues and poles or a rational trigonometric function, we may consider as input weights and exponents of the Fourier transform of the rational function. After performing the reduction we may analytically obtain the reduced residues and poles.

We start with a real valued rational function $f(z)$,

$$f(z) = \sum_{j=1}^{n} \frac{d_j}{z - \mu_j} + \sum_{j=1}^{n} \frac{\overline{d_j} z}{1 - \overline{\mu_j} z} + d_0, \tag{3.6}$$

with $d_0 \in \mathbb{R}$, $d_j, \mu_j \in \mathbb{C}$ and $|\mu_j| < 1$. Our goal is to find a rational function $r(z)$ of the form

$$r(z) = \sum_{j=1}^{m} \frac{r_j}{z - \eta_j} + \sum_{j=1}^{m} \frac{\overline{r_j} z}{1 - \overline{\eta_j} z} + d_0,$$

with fewer poles than $f(z)$ such that

$$\left| r(e^{2\pi i x}) - f(e^{2\pi i x}) \right| < \epsilon \quad \forall x \in [0, 1).$$

The steps to compute such an approximation are:

- For the Cauchy matrix $C_{ij}(\mu_i, d_j)$,

  $$C_{ij} = \frac{a_i b_j}{x_i + y_j}, \;\; i, j = 1, \ldots, n,$$

  with $a_i = \sqrt{d_i}/\mu_i$, $b_j = \sqrt{\overline{d_j}}$, $x_i = \mu_i^{-1}$, and $y_j = -\bar{\mu_j}$. Solve the con-eigenproblem $Cu = \sigma_m \bar{u}$ for a con-eigenvalue σ_m and con-eigenvector $u = (u_1 \;\; u_2 \cdots u_n)^t$. Here $0 < \sigma_m \leq \epsilon$ and the con-eigenvalues are denoted by $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$.

- Find the roots ν_j inside the unit disk of

$$v(z) = \frac{1}{\sigma_m} \sum_{j=1}^{n} \frac{\sqrt{\overline{d_j}}\,\overline{u_j}}{1 - \overline{\mu_j} z}.$$

  Note that there should be exactly m zeros of $v(z)$ inside the unit disk based on results from [3].

- Finally solve for the residuals r_k of $r(z)$ by solving the $m \times m$ linear system

$$\sum_{j=1}^{m} \frac{1}{1 - \nu_j \overline{\nu_k}}\, r_j = \sum_{j=1}^{n} \frac{d_j}{1 - \mu_j \overline{\nu_k}}.$$

Using this algorithm we obtain $\|f - r\| \approx \sigma_m$, which provides a near optimal representation of $f(z)$ using only $2m$ poles. The computational complexity of this algorithm is $\mathcal{O}\left(m^2 n\right)$, where m is the number of resulting poles and n is the original number of poles. Since typically $m \ll n$, this algorithm is linear in its practical use.

3.3.3 Spline representations

Finding rational approximation for large data sets directly via algorithms described above is not practical due to their computational cost. For this reason we use an intermediate representation via B-splines as a step towards computing the rational approximation. Although theoretically we may use any wavelet-type basis, the choice of B-splines reduces the computational cost of this intermediate step.

We recall the definition of the m^{th} degree B-spline as

$$\beta_m(x) = \beta_{m-1}(x) * \beta_0(x),$$

with

$$\beta_0(x) = \begin{cases} 1, & |x| \leq \frac{1}{2} \\ 0, & \text{otherwise}, \end{cases}$$

(see e.g., [9]). It is easy to show that β_m is a piecewise polynomial of degree m with knots on the integers. We use only B-splines of odd degree. To represent periodic functions, we use periodized versions of B-splines. Let us introduce the 1-periodic function

$$a_m(\omega) = \sum_{j\in\mathbb{Z}} |\widehat{\beta}_m(\omega + j)|^2 = \sum_{l=-\frac{m-1}{2}}^{\frac{m-1}{2}} \beta_m(l) e^{-2\pi i\omega}.$$

Given a uniformly sampled 1-periodic function f, we seek the coefficients α_j such that

$$f\left(\frac{k}{2N}\right) = \sum_{j=0}^{2N} \alpha_j \beta_m(k-j), \quad k = 0, \ldots, 2N. \tag{3.7}$$

The algorithm in [4, 18] rapidly computes the coefficients α_j in (3.7) using the Fast Fourier Transform (FFT). It performs the following steps:

- Set $f_k = f(\frac{k}{2N})$ and compute, for $k = 0, \ldots, 2N$,

  $$\widehat{f_k} = \sum_{n=0}^{2N} f_n e^{\frac{-2\pi i}{2N+1} kn}$$

  using the FFT.

- Compute, for $k = 0, \ldots, 2N$,

  $$\widehat{\alpha}_k = \frac{\widehat{f_k}}{a_m(\frac{k}{2N+1})}.$$

- The B-spline coefficients are now obtained via the FFT as

  $$\alpha_j = \frac{1}{2N+1} \sum_{n=0}^{2N} \widehat{\alpha}_n e^{\frac{2\pi i}{2N+1} jn}, \quad j = 0, \ldots, 2N$$

This algorithm requires $\mathcal{O}(N \log N)$ operations. The details may be found in the appendix in [18].

3.4 Rational representation of B-splines

In this section we construct rational approximations of B-splines. In our construction we force the imaginary parts of the poles to be of the form $2\pi l$, $l \in \mathbb{Z}$ so that these imaginary parts

coincide with that of the knots of the B-spline. As we show below, this reduces the cost of intermediate computations.

Specifically, we are looking for a suboptimal rational approximation of the form (3.5) such that

$$\left| \beta_m(x) + 2 \sum_{l=-\frac{m+1}{2}}^{l=\frac{m+1}{2}} \sum_{k=1}^{M_0} \frac{v_{k,l}(2\pi x - 2\pi l) - u_{k,l}\tau_k}{(2\pi x - 2\pi l)^2 + \tau_k^2} \right| \leq \epsilon. \tag{3.8}$$

Since

$$\sum_{l=-\frac{m+1}{2}}^{l=\frac{m+1}{2}} \sum_{k=1}^{M_0} \frac{v_{k,l}(2\pi x - 2\pi l) - u_{k,l}\tau_k}{(2\pi x - 2\pi l)^2 + \tau_k^2} = \mathcal{R}e \left(\sum_{l=-\frac{m+1}{2}}^{l=\frac{m+1}{2}} \sum_{k=1}^{M_0} \frac{w_{k,l}}{2\pi i x - \eta_{k,l}} \right),$$

we have as poles $\eta_{k,l} = 2\pi i\, l + \tau_k$ and as weights $w_{k,l} = u_{k,l} + i v_{k,l}$. We note that with this constraint on the imaginary parts of $\eta_{k,l}$, the poles of the approximation are arranged on a rectangular grid.

We start by computing a near optimal rational approximation of B-splines following the approach in [6]. For $m \leq 7$, we evaluate $\widehat{\beta}$ at a sufficient number of samples, specifically

$$h_n = \widehat{\beta}_m \left(\frac{n}{16} \right), \quad n = 0, 1, \ldots, 400, \tag{3.9}$$

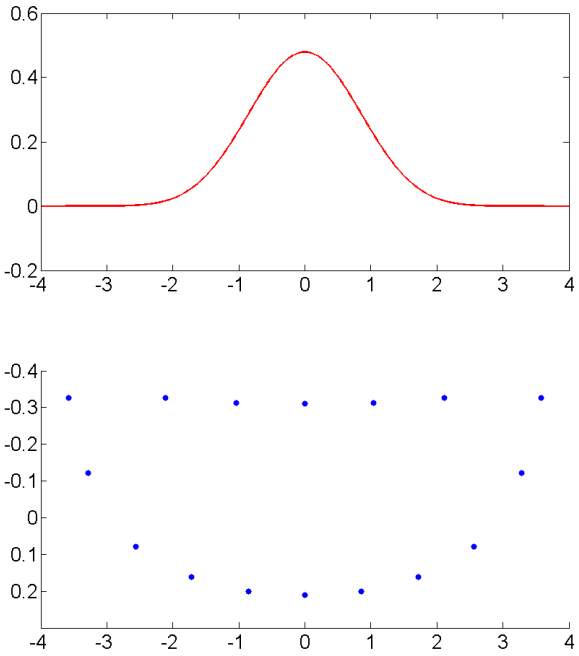
where

$$\widehat{\beta}_m(\xi) = \left(\frac{\sin \pi\xi}{\pi\xi} \right)^{m+1}, \tag{3.10}$$

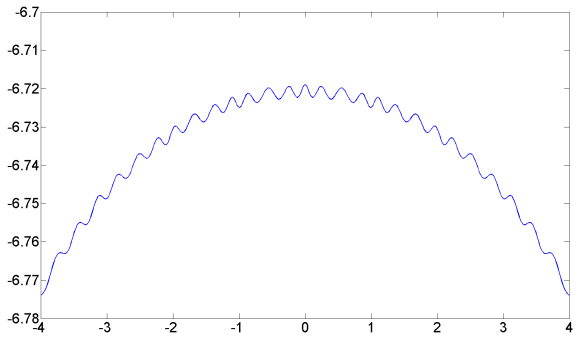
and use the algorithm in Section 3.3.1.

An example of a near optimal rational approximation of a B-spline of degree $m = 3$ may be found in [6]. As observed in that paper, the poles concentrate towards the locations of the knots of the B-spline since its third derivative is discontinuous at these points. In our application we would like to use a higher degree B-spline to lessen the impact of discontinuities and obtain fewer poles. We seek a suboptimal rational representation with poles in the locations indicated in (3.8) and use the near optimal approximation to select the τ_k in (3.8).

In Figure 3.1 we present the results for a near optimal approximation of a 7th degree B-spline using the same approach as in [6]. Since the poles, $\eta_{k,l}/2\pi i$, appear in complex conjugate pairs, in Figure 3.1 we display only those with negative imaginary part (on a $\log_{10}$ scale).



(a)



(b)

Figure 3.1: (a) Near optimal rational representation of a B-spline of degree 7 using 16 nodes (displayed on a $\log_{10}$ scale for their imaginary part). (b) Associated error on a $\log_{10}$ scale.

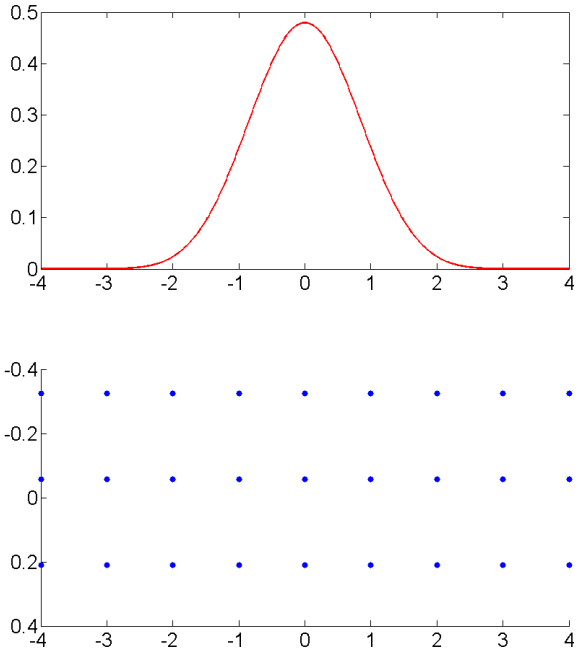
Taking into account that the poles closer to the real line are responsible for the high frequency information of the representation, whereas those furthest away capture the lower frequency content, we limit the range for the imaginary parts of our suboptimal poles by using the corresponding maximum, τ^+, and minimum, τ^-, of the near optimal poles. We also add a third set of poles with

$$\tau = e^{\frac{1}{2}\left(\log \tau^+ + \log \tau^-\right)}.$$

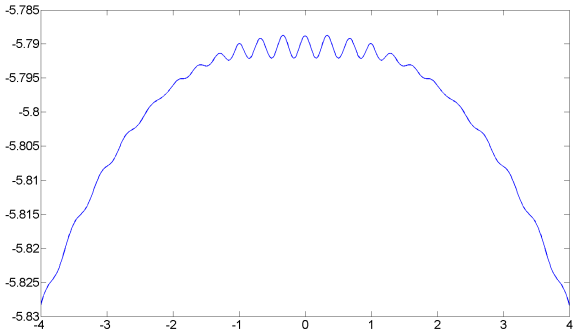
We select the real part for all of these poles as in (3.8), i.e. at location $2\pi l$, where $l = -\frac{m+1}{2}, \ldots, \frac{m+1}{2}$ (recall that m is odd). The computation of the residuals leads to an ill-conditioned Vandermonde system (3.4). We compute a solution with minimal l_1-norm using the optimization package [13]. The resulting absolute error is shown in Figure 3.2(b), where the target accuracy for approximating the B-spline is $\epsilon \leq 10^{-5}$.

We make an additional observation about our suboptimal B-spline approximation. While we have designed this approximation to be accurate within the support of the B-spline, it is actually accurate in the L^∞-norm over the whole real line. In Figure 3.3 we illustrate this behavior. This property is particularly important for the merging algorithm.

We use this suboptimal representation of the B-spline to transform a B-spline decomposition of the original signal to a suboptimal rational representation. The particular choice of pole locations implies that the number of poles of the resulting suboptimal representation exceeds the number of B-spline coefficients in (3.7) only by a factor of 3. In fact, our choice of B-splines was motivated by this moderate increase in the number of terms in comparison to other wavelet-type decompositions. The choice of the 7th degree spline is dictated by the target accuracy for our final signal approximation. For greater accuracy higher order B-splines may be used and their approximation may be obtained by the same procedure.



(a)



(b)

Figure 3.2: Rational representation of a B-spline of degree 7 using 27 nodes arranged on rectangular grid (displayed using $\log_{10}$ scale for their real part) (a) and associated error (b) on $\log_{10}$ scale.

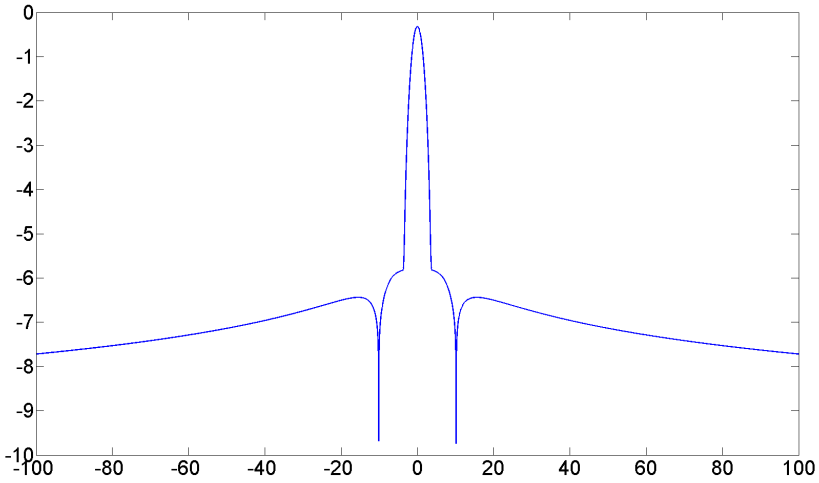


Figure 3.3: Magnitude of suboptimal rational approximation of 7th degree B-spline over a large interval ($\log_{10}$ scale). Note that outside of the support of the B-spline its rational approximation is accurate within the target accuracy.

3.5 Near optimal rational approximations

Let us describe the algorithm for computing near optimal rational representations for very large data sets. We only assume that the signal is negligible at the start and end but, otherwise, may have a very large number of samples. The algorithm involves the following steps:

1 In order to apply the algorithm described in Section 3.3.3, we subdivide the signal using windows into overlapping segments so that we may treat each segment as a periodic function. The size of these segments should be appropriate for an efficient use of the FFT but otherwise is arbitrary. We compute the B-spline coefficients for each segment. We combine the B-spline coefficients from each segment to get B-spline coefficients for the entire signal.

2 We group the computed B-spline coefficients from step (1) into consecutive segments (which are unrelated to the windows used in step (1)). The size of these segments should be appropriate to guarantee efficiency of the algorithm in Section 3.3.2. By using the B-spline approximation constructed in Section 3.4, we obtain a suboptimal rational representation of each segment.

3 On each segment we use the reduction algorithm in Section 3.3.2 to obtain a near optimal rational approximation for that segment.

4 We now merge the rational representations of adjacent segments. As we explain below, only adjacent segments interact with each other so that this step does not have to be done globally. Furthermore, only poles near the boundary between segments need to be merged. This step may be considered optional since the overlap of the functions associated with adjacent segments is small in comparison to the length of each segment, so that the number of extra nodes is a small percentage of the total number of nodes.

We will now describe each step in some detail.

1 We use a partition of unity as our windows, and we note that there may be significant overlap

between adjacent windows. The only requirements for the windows is a smoothly decaying transition region so that we do not introduce additionally frequency content into the signal, and sufficient decay as to obtain partitions that are appropriate for the use of the FFT. We then use the algorithm in Section 3.3.3 to compute the B-spline coefficients for each partition. Once the B-spline coefficients for each partition are found, adding components from different partitions, we obtain the B-spline coefficients for the entire sequence of sampled data. The cost of this step is $\mathcal{O}(N_{signal} \log N_{part})$, where N_{signal} is the total number of samples and N_{part} is the number of samples in each partition assuming they are of the same length. As a result we obtain a representation

$$f(x) = \sum_{j=0}^{N_{signal}-1} \alpha_j \beta_m(x-j). \tag{3.11}$$

2 Given the signal in the form (3.11), we split the sum into segments of length P for further processing,

$$f_p(x) = \sum_{j=pP}^{(p+1)P-1} \alpha_j \beta_m(x-j) = \sum_{j=0}^{P-1} \alpha_{j+pP} \beta_m(x-j-pP), \tag{3.12}$$

where

$$p = 0, \ldots, \left\lfloor \frac{N_{signal}-1}{P} \right\rfloor, \tag{3.13}$$

(we allow incomplete segments). For each segment we replace the B-splines by their suboptimal rational representation constructed in Section 3.4. For each segment we obtain the suboptimal representation

$$\widetilde{f}_p(x) = -2 \sum_{j=0}^{P-1} \left(\alpha_{j+pP} \sum_{l=-\frac{m+1}{2}}^{l=\frac{m+1}{2}} \sum_{n=1}^{M_0} \frac{2\pi v_{n,l}(x-j-pP-l) - u_{n,l}\tau_n}{4\pi^2(x-j-pP-l)^2 + \tau_n^2} \right). \tag{3.14}$$

This approximation may also be written as

$$\tilde{f}_p(x) = -2 \sum_{j=-\frac{m+1}{2}}^{P-1+\frac{m+1}{2}} \mathcal{R}e \left\{ \sum_{n=1}^{M_0} \frac{\rho_{p,n,j}}{2\pi i(x-j-pP) + \tau_n} \right\},$$

where

$$\rho_{p,n,j} = \sum_{k=\max(0,j-\frac{m+1}{2})}^{\min(P-1,j+\frac{m+1}{2})} \alpha_{k+pP} w_{n,j-k}.$$

Thus, the suboptimal approximation in each segment requires $(P + m + 1)\, M_0$ poles. For example, with our choice of seventh degree B-splines, $M_0 = 3$ (see Section 3.4).

3 For each p in (3.13), we apply to the suboptimal approximation $\widetilde{f}_p(x)$ the reduction algorithm described in Section 3.3.2. We obtain a near optimal representation with M_p^{opt}poles,

$$\tilde{f}_p^{opt}(x) = -2\mathcal{R}e \sum_{m=1}^{M_p^{opt}} \frac{w_m^p}{2\pi i x - \eta_m^p}, \qquad \left\| \widetilde{f}_p^{opt}(x) - f_p(x) \right\| < \epsilon.$$

4 In order to merge the near optimal approximations from adjacent segments, we may use the reduction algorithm once again. We note that, for our purposes, we need to merge only poles near the boundary between adjacent segments keeping unchanged the poles far away from the boundary region. To accomplish this, we consider the function generated by the poles we wish to reduce and their corresponding residues. By requesting a slightly higher accuracy across the support of both segments we preserve the overall accuracy of the approximation. In our experiments, we reduce the set of poles located at most 64 units (measured by step size of the original signal) from the midpoint of the overlapping region between adjacent segments. This selection is made to assure that the positions of the nodes possibly affected by the splitting of the data into segments are adjusted by the reduction algorithm. The slightly higher accuracy assures that the untouched poles are not impacted by this merge, and hence we do not need to recompute their weights. We obviously do not obtain the optimal approximation over the support of the two segments, but we claim that the approximation is near optimal, both in terms of the number of poles and their locations.

Given the optimal representations $\widetilde{f}_p^{opt}(x)$ for p in (3.13), we merge the adjacent represen-

tations and denote the entire merged representation as

$$\widetilde{f}(x) = \sum_{p} \widetilde{f}_p^{merged}(x).$$

We also note that the observed reduction in the number of poles within two adjacent segments is minimal and, therefore, this step may be considered optional in practice.

3.6 Numerical Examples

We have computed a series of numerical examples using the algorithm from Section 3.5. Since one of the potential applications for this method is a compression scheme, the sampled data was taken from a high quality audio recording. Generally, audio recordings are done with 16 bits per sample. This means that the maximum accuracy possible is 2^{-15}, provided that the maximum amplitude of the signal is 1. This section contains an example of finding a near optimal approximation for a single segment, $f_p(x)$, and then demonstrates the procedure of merging the approximations of adjacent segments.

3.6.1 Rational representation for a single segment

Using the algorithm in Section 3.5, we compute a near optimal rational approximations of sampled data from a high quality audio recording. First, we compute a B-spline representation for the entire signal. We then consider the performance of our algorithm to approximate, with accuracy 6×10^{-4}, the function $f_p(x)$ in (3.12). Figure 3.4 displays the rational approximation with 44 poles, the locations of those poles, and the associated error. We display the error $\left| f_p(x) - \widetilde{f}_p(x) \right|$, where $\widetilde{f}_p(x)$ is the resulting rational approximation and note that we achieve the same accuracy vis-à-vis the original signal (in a slightly smaller interval) due to the fact that the computation of a B-spline representation is accurate to machine precision.

Although this level of compression is reasonable, we note that in order to develop a complete compression scheme additional steps should be taken to encode the parameters of the near optimal

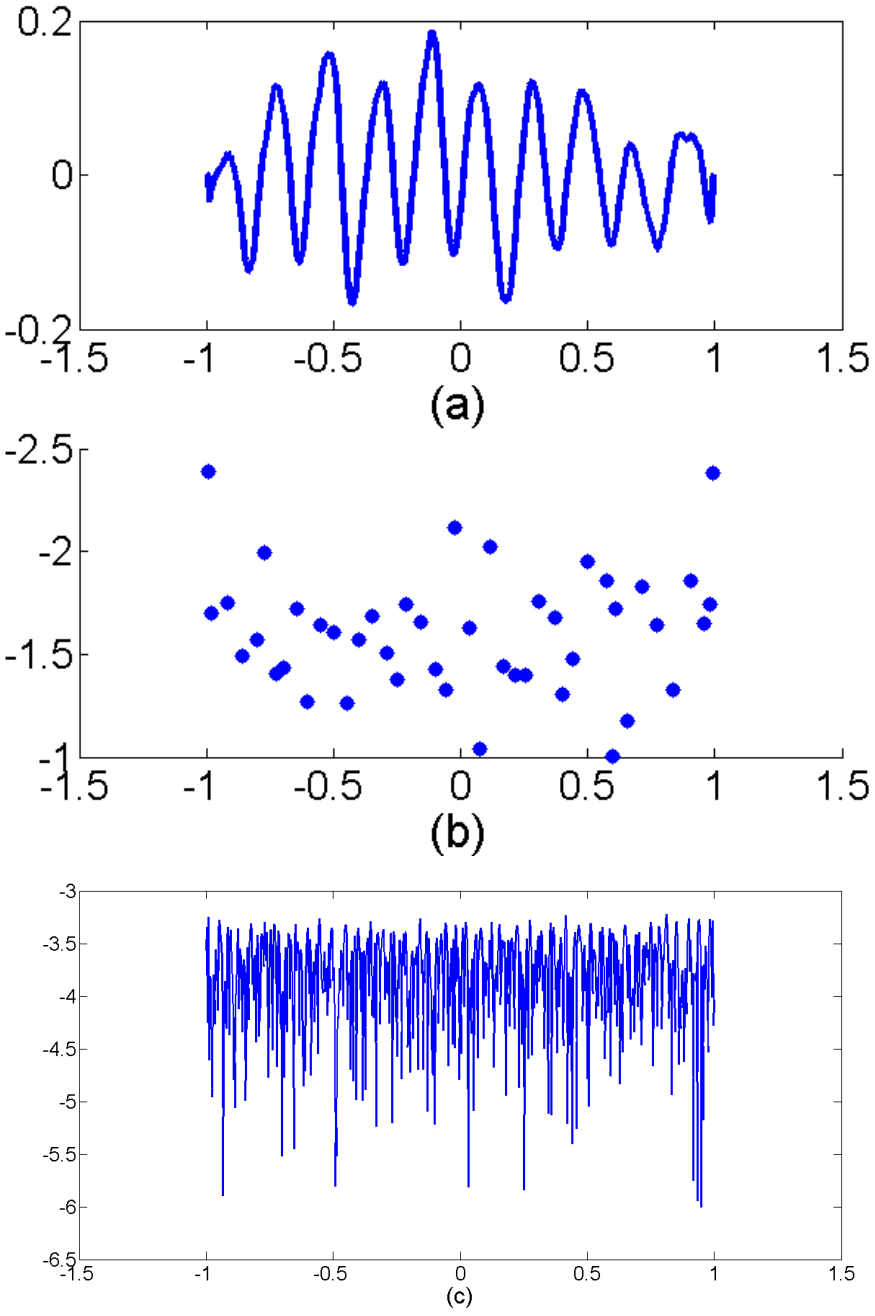


Figure 3.4: (a) Rational approximation of a 768 sampled data points using 44 nodes. (b) Pole locations (displayed using $\log_{10}$ scale for their imaginary part). (c) Associated error on $\log_{10}$ scale.

rational approximation. Furthermore, by taking into account the level of signal noise, we may reduce the number of poles in the representation. We note that the decay of the singular values of the Cauchy matrix formed for the reduction algorithm may be examined to develop an approximation for the level of noise in the signal [6].

3.6.2 Merging of adjacent segments

One of the key benefits of the approximation method used is that each pole of the near optimal rational representation only locally influences the reconstruction. For this reason merging the near optimal rational approximations of adjacent segments minimally alters the original pole locations for the two segments. To show this, we compute near optimal rational approximations for two adjacent segments, $f_p(x)$ and $f_{p+1}(x)$, each of length 512. Figure 3.5 shows the near optimal approximations of the two adjacent segments along with the error. The first segment requires 30 poles and the second requires 29 poles.

Figure 3.6 shows the representation for the merged windows, pole locations and associated error.

The poles that were within 64 sample distances of the boundary between segments were merged. The merged approximation required 58 poles, which means that a minimal reduction has taken place with respect to the total number of poles required for the two segments. This shows that due to the compact support of the B-splines, and consequently the good localization of their rational approximations, the combined representations of the individual segments provide a near optimal representation for the combined segments. Furthermore, these results demonstrate that the step of merging adjacent representations may be considered optional.

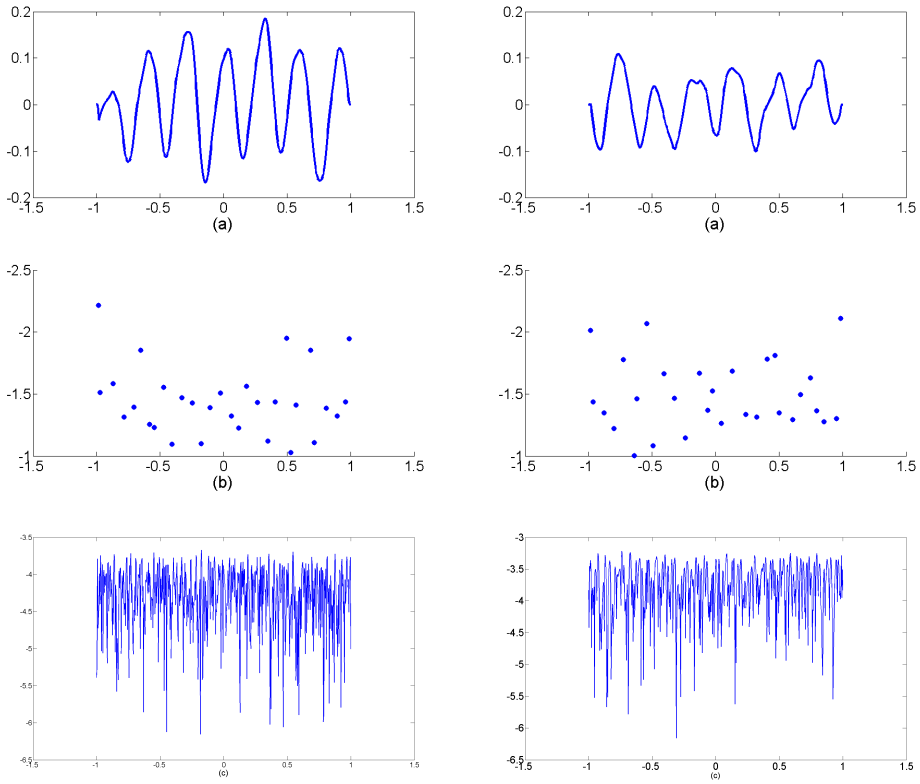
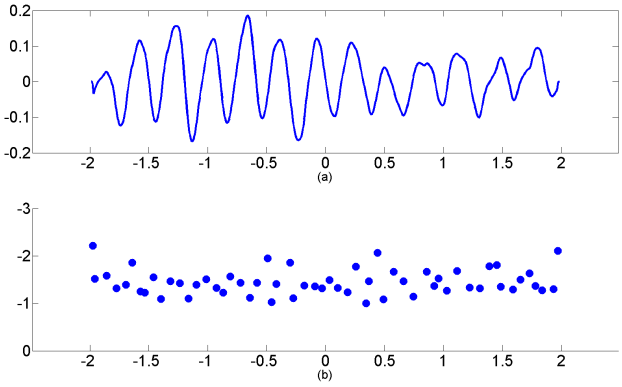
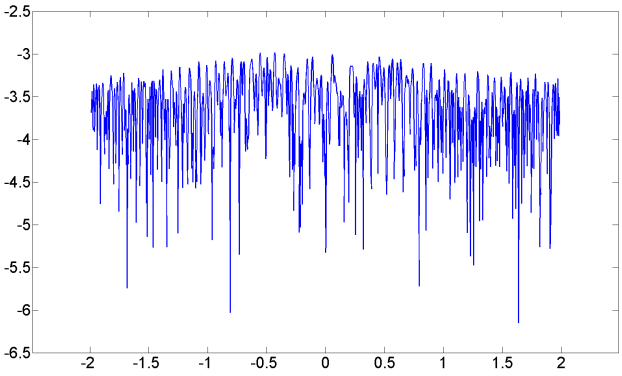


Figure 3.5: (a) Near optimal rational approximations on adjacent segments. (b) Pole locations (displayed using $\log_{10}$ scale for their imaginary part). (c) Associated error on $\log_{10}$ scale.



(a)



(b)

Figure 3.6: (a) Merged Representation using 59 poles. (b) Pole locations (displayed using $\log_{10}$ scale for their imaginary part). (c) Associated error on $\log_{10}$ scale.

3.7 Conclusion

We have presented a new algorithm to compute near optimal rational approximations for large data sets. Our approach combines several algorithms for this purpose. The speed of these algorithms allows us to construct such approximations even for signals generated by some continuous monitoring, as for example, seismic monitoring. The results also show promise for the development of a competitive algorithm for music compression. The building blocks of these representations contain information about local frequency content and are shift invariant. This property facilitates further processing of signals as the parameters are practically independent of the initial shift of the input data; this also opens the possibility of recognizing recurring signal features present at spatially separated points. We plan to present these further developments elsewhere.

Chapter 4

Final remarks

We have presented a new algorithm to compute near optimal rational approximations for large data sets. This development addresses the impractical computational cost of using currently available algorithms for such large data sets. Our approach combines several algorithms and this combination allows us to construct such approximations even for signals generated by a continuous monitoring, as for example, seismic monitoring. The results also show promise for the development of a competitive algorithm for music compression. Similarly to wavelet decompositions, the poles of these representations contain information about local frequency content but, unlike wavelet representations, the near optimal rational approximations are shift invariant. The shift-invariant nature of these representations may facilitate further processing of signals and opens a possibility of recognizing recurring signal features present at spatially separated points.

Bibliography

[1] V. M. Adamjan, D. Z. Arov, and M. G. Kreĭn. Infinite Hankel matrices and generalized Carathéodory-Fejér and I. Schur problems. Funkcional. Anal. i Priložen., 2(4):1–17, 1968.

[2] V. M. Adamjan, D. Z. Arov, and M. G. Kreĭn. Infinite Hankel matrices and generalized problems of Carathéodory-Fejér and F. Riesz. Funkcional. Anal. i Priložen., 2(1):1–19, 1968.

[3] V. M. Adamjan, D. Z. Arov, and M. G. Kreĭn. Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur-Takagi problem. Mat. Sb. (N.S.), 86(128):34–75, 1971.

[4] G. Beylkin and R. Cramer. Toward multiresolution estimation and efficient representation of gravitational fields. Celestial Mechanics and Dynamical Astronomy, 84(1):87–104, 2002.

[5] G. Beylkin and L. Monzón. On approximation of functions by exponential sums. Appl. Comput. Harmon. Anal., 19(1):17–48, 2005.

[6] G. Beylkin and L. Monzón. Nonlinear inversion of a band-limited Fourier transform. Appl. Comput. Harmon. Anal., 27(3):351–366, 2009. doi: 10.1016/j.acha.2009.04.003.

[7] G. Beylkin and L. Monzón. Approximation of functions by exponential sums revisited. Appl. Comput. Harmon. Anal., 28(2):131–149, 2010.

[8] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. IEEE Trans. Communications, 31(4):532–540, Apr. 1983.

[9] C. Chui. An Introduction to Wavelets. Academic Press, 1992.

[10] I. Daubechies. Orthonormal bases of compactly supported wavelets. Comm. Pure and Appl. Math., 41:909–996, 1988.

[11] I. Daubechies. Ten Lectures on Wavelets. CBMS-NSF Series in Applied Mathematics. SIAM, 1992.

[12] R.A. DeVore, B. Jawerth, and V. Popov. Compression of wavelet decompositions. American Journal of Mathematics, 114(4):737–785, 1992.

[13] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 1.21. http://cvxr.com/cvx, February 2011.

[14] A. Haar. Zur Theorie der orthogonalen Funktionensysteme. Mathematische Annalen, pages 331–371, 1910.

[15] T. S. Haut and G. Beylkin. Fast and accurate con-eigenvalue algorithm for optimal rational approximations. arXiv:1012.3196v2 [math.NA], 2011.

[16] S. Jaffard, Y. Meyer, and R.D. Ryan. Wavelets: tools for science & technology. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, revised edition, 2001.

[17] B. Jawerth and W. Sweldens. An overview of wavelet based multiresolution analyses. SIAM Rev., 36(3):377–412, 1994.

[18] B.A. Jones, G.H. Born, and G. Beylkin. A Cubed Sphere Gravity Model for Fast Orbit Propagation. AAS/AIAA Spaceflight Mechanics Meeting, Advances in the Astronautical Sciences, 134:567–584, 2009.

[19] D. J. Newman. Rational approximation of $|x|$. Michigan Math. J., 11:11–14, 1964.

[20] I. J. Schoenberg. Cardinal spline interpolation. SIAM, Philadelphia, Pa., 1973. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 12.

[21] M. J. T. Smith and T. P. Barnwell. Exact reconstruction techniques for tree-structured subband coders. IEEE Trans. ASSP, 34:434–441, 1986.

[22] J. O. Strömberg. A Modified Franklin System and Higher-Order Spline Systems on $\mathbf{R}^n$ as Unconditional Bases for Hardy Spaces. In Conference in Harmonic Analysis in Honor of Antoni Zygmund, Wadworth Math. Series, pages 475–493, 1983.