Portfolio Approximation *

Thomas A. Knox

January 20, 2015

Abstract

I consider approximating the behavior of a portfolio across a set of scenarios. No asset pricing model is assumed; accuracy is measured by worst-case approximation error or average squared approximation error across the scenario set. Problems of this type have engaged researchers for centuries, resulting in tables of bond values and the concepts of immunization and of duration. Solutions to these problems are portfolio approximations; they are useful in valuing a portfolio across scenarios, compressing data on scenario valuations into risk measures, and constructing a portfolio with scenario valuations close to a target profile. The novel approximations I provide here are the first to have verifiable optimality properties for portfolios of bonds, interest rate swaps, credit default swaps, and equities. I introduce a quartet of new approximations: numerically-optimal, interpolatory, Mathieu, and oblate.

The numerically-optimal approximation earns its name by achieving, to within rounding error, a lower bound that I derive and compute for the worst-case error of any approximation. It can be interpreted as assembling a portfolio using a small set of simple approximating cashflow streams to approximate any given cashflow stream's behavior. This could obviously be used to immunize a portfolio against scenario risk; such an immunization would inherit the worst-case error optimality of the approximation.

The interpolatory approximation is optimal for mean-square approximation error. It can also be interpreted as assembling an approximating portfolio, and its immunization of a target cashflow stream across the scenario set would inherit its mean-square optimality. The Mathieu approximation is nearly optimal for worst-case approximation error and is nested: the rank-$(n+1)$ approximation is just the rank-n approximation plus an additional term. The oblate approximation is optimal for mean-square approximation error and nested.

The approximations introduced here can calculate scenario valuations more than 1,000,000 times faster than a direct method. Scenario data can be compressed to three numbers with worst-case error of less than 4 parts per 100,000. My approximations can be 20 times (rank-three), 70 times (rank-four), 300 times (rank-five), or 350,000 times (rank-twelve) more accurate in worst-case error than standard Taylor-series approximations.

1 Introduction

Portfolio approximation is useful in valuing a portfolio across scenarios, compressing data on scenario valuations into risk measures, and immunizing a portfolio against scenario risk. No asset pricing model is assumed; accuracy is measured by worst-case approximation error or average squared approximation error across a scenario set. I introduce four new approximations: numerically-optimal, Mathieu, interpolatory, and oblate. They are the first to have verifiable optimality properties in the sense described above, and they can be classified using the following simple table:

	Worst-case Error	**Avg. Squared Error**
Portfolio	Numerically-optimal	Interpolatory
Nested	Mathieu	Oblate

The novel approximations in the "worst-case error" column are numerically-optimal or (for the Mathieu approximation) near-optimal for worst-case approximation error. The approximations in the "average squared error" column, which are also new, are optimal for mean-square approximation error. The "portfolio" row indicates the two approximations (numerically-optimal and interpolatory) that can be interpreted as providing a simple approximating portfolio, which might be used to immunize a cashflow stream against scenario risk. The "nested" row includes the two approximations (Mathieu and oblate) for which the rank-$(n+1)$ approximation is just the rank-n approximation plus one new term.

Before delving deeper into these approximations, it may be useful to explore a simple example. Consider approximating the behavior of an idealized Treasury bond in parallel shifts of the interest rate curve. This is a very special case of the general problem that I analyze in this paper (I analyze shifts of arbitrary shape and any form of cashflow stream), but it shows many of the key results. Figure 1 illustrates this approximation problem across shifts from -500 basis points to 500 basis points using approximations of rank three (so that every approximation treated in the figure has three degrees of freedom; for a Taylor series, this means using base value, duration, and convexity). The top left panel simply shows that the shift is parallel; the top right panel shows the present values of the cashflows of the bond being approximated, which is an idealized version of a 30-year Treasury bond.

The middle left panel provides the approximating portfolio constructed by one of my new approximations, the numerically-optimal approximation. The rank-three numerically optimal approximation uses three approximating cashflow streams (it cannot use more), each having cashflows paid at the same four times. The approximating cashflow times chosen for numerically-optimal worst-case error approximation are the two most extreme available (the times corresponding to 6 months and to 30 years) and two intermediate times (one in the belly of the curve, the other in the long end). Although the approximation does not order the streams in any particular way, there are clearly a "level annuity" stream (Stream 2 in the panel, in which the present values of the cashflows across the approximating cashflow times all have the same sign and are similar

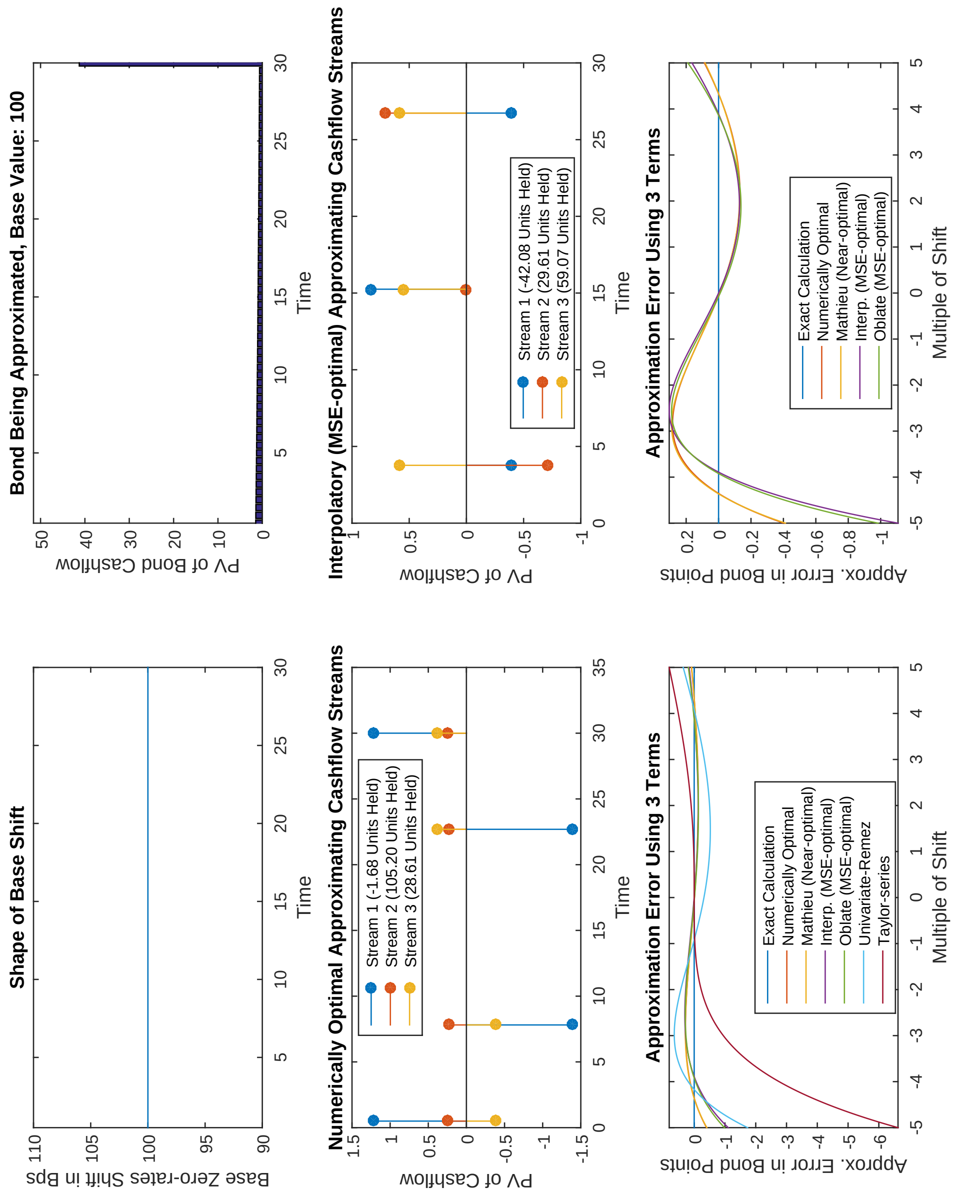

Figure 1: Approximating a Parallel-shift Profile

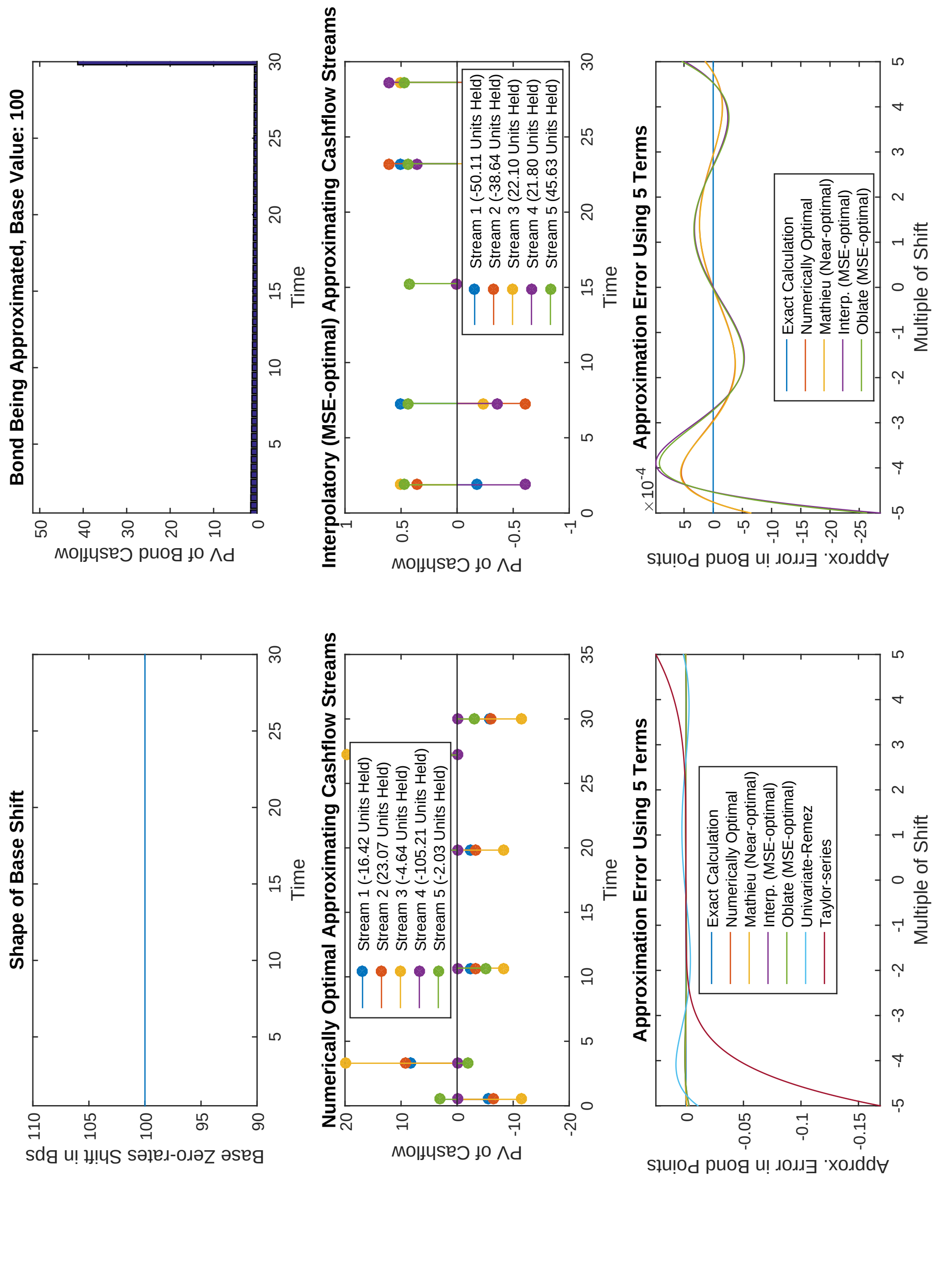

Figure 2: Rank Five Approximation of a Parallel-shift Profile

in magnitude), a “flattener” stream (Stream 3 in the panel, in which the present values of the cashflows at the first two approximating cashflow times are negative and the present values of the cashflows at the second two approximating cashflow times are positive), and a “butterfly” stream (Stream 1 in the panel, in which the present values of the cashflows at the middle two approximating cashflow times are negative and the present values of the cashflows at the first and last approximating cashflow times are positive). Note that the approximation is indifferent to changing the signs of all of the present values in a given approximating cashflow stream, since the approximating portfolio is permitted to have short positions in an approximating cashflow stream, but the shape of the present values over approximating cashflow times is very important.

Unsurprisingly, the numerically-optimal approximating portfolio is focused on Stream 2, the “level annuity” discussed above (matching the coupon stream), and on Stream 3, the “flattener” discussed above (which relates to the principal strip). It is worth noting that the present value of the approximating cashflow stream is always, by construction, equal to the approximation’s value for the target portfolio with no scenario movement.

The middle right panel of Figure 1 shows the analogous approximating portfolio for another of my new approximations, the interpolatory approximation, which also has three approximating cashflow streams but which uses only three common cashflow times. The “level annuity” (here, Stream 3), “flattener” (here, Stream 2), and “butterfly” (here, Stream 1) approximating streams are, as in the numerically-optimal case of the middle left panel, clearly identifiable. Interestingly, the interpolatory approximation (with its focus on average squared error) weights these streams rather differently, choosing a much larger butterfly position than the numerically-optimal approximation (with its focus on worst-case error) does.

In the bottom left panel of Figure 1, I show approximation error across parallel shifts for different approximations. The shift multiples of -5 to 5 mean that I consider parallel shifts of -500 to 500 basis points. That 1,000 basis-point range in rates is quite wide relative to standard scenario analysis ranges; I wish to show how well my approximations perform even under challenging conditions. The first feature that emerges from the bottom left panel of Figure 1 is the terrible performance of the Taylor series. It makes errors of more that 6.62 bond points across the scenario set; since the hypothetical bond is valued at par, this corresponds to 6.62 percent return errors across the scenario set. The univariate-Remez approximation is also deeply problematic; it incurs a nontrivial approximation error even when no scenario shift has happened!

Discarding the obviously inadequate Taylor series and univariate-Remez approximations and rescaling the vertical axis to get a better view of the four approximations I introduce in this work, I come to the bottom right panel of Figure 1. The error incurred by the numerically-optimal and Mathieu approximations is excellent; the magnitude of error is less than 0.42 points across a 1,000 basis-point range of rates. The numerically-optimal and Mathieu approximations are more than 16 times more accurate than the Taylor series. The interpolatory and oblate approximations perform well, but not as well, with maximum absolute errors of 1.107 (to three decimals) and 0.984, respectively; the presence of an important cashflow at an extreme time (the principal strip) in the target cashflow stream creates challenges for them.

The sizes of errors discussed above should be viewed in the context of the wide range of rates considered; for smaller ranges or for a higher approximating rank, the error of the numerically-optimal and Mathieu approximations is reduced to negligible levels. Even with only three degrees of freedom (rank-three approximation), the numerically-optimal and Mathieu approximations can perform well enough across 1,000 basis-point rate ranges to be useful for high-level risk management of 30-year cashflow streams.

To illustrate the reduction of approximation error to negligible levels, see Figure 2. This figure revisits the analysis of Figures 1 but increases the rank of all approximations to five. All approximations become much more accurate, but the relative advantage of the numerically-optimal and Mathieu approximations grows even greater. The maximal absolute error of the Taylor series approximation here is 0.1688086 points; this seems reasonable until one views the corresponding errors for the numerically-optimal approximation, at 0.0006286 points, and the Mathieu approximation, at 0.0006441 points. The error for the Taylor series is not large, but for the same number of terms one can achieve negligible error with the numerically-optimal or Mathieu approximations. The numerically-optimal approximation is more than 268 times more accurate than the Taylor series approximation here; this relative advantage only grows as approximation rank increases further.

I return to this example and several variants of it in Subsection 11.1 below, where slope shifts and curvature shifts are used in addition to parallel shifts and the whole bond, the coupon stream, and the principal strip are considered separately.

In order to obtain the four novel approximations introduced here (numerically-optimal, Mathieu, interpolatory, and oblate), I study optimal portfolio approximations using n-widths. For portfolios of bonds, interest rate swaps, credit default swaps, and equities, the map from cashflows to scenario values is represented using a family of integral operators, the exponential product operators with kernels $\exp(cxy)$, where c indexes the family and $x, y \in [-1, 1]$. Given a set of cashflow streams and a norm on scenario valuations, the linear n-width of an exponential product operator is the smallest scenario approximation error, across rank-n linear approximating operators, for the hardest cashflow stream to approximate. An optimal rank-n linear approximation is one that achieves the linear n-width.

The set of cashflow streams that are signed measures with total variation bounded by a constant and the supremum norm on scenario valuations lead to linear n-widths that call for *optimal worst-case error rank-reduction*. I introduce methods to compute lower bounds on these linear n-widths, to build novel rank-n linear approximations, and to verify that the new rank-n linear approximations achieve the lower bounds to within rounding error.

Because of the specific form of the novel numerically-optimal approximation, the rank-n approximation it produces for any given cashflow stream and scenario set is a portfolio of n different cashflow streams, each having cashflows at the same $n + 1$ times (as is clear in the example above). In approximating the values of different cashflow streams across the same scenario set, the rank-n numerically-optimal approximation uses the same n approximating cashflow streams but weights them differently to produce distinct approximating portfolios. The approximating portfolio holdings are linear in the cashflow stream being approximated, so the approximating holdings for a portfolio of cashflow streams are the sums of the approximating holdings for the individual cashflow streams in the portfolio.

I also develop a new near-optimal worst-case error approximation using the eigenfunctions of weighted exponential product operators, which turn out to be the Mathieu functions of classical mathematical physics. Due to this relationship, I dub the novel near-optimal approximation the "Mathieu approximation."

The set of cashflow streams that are square-integrable and the root-mean-square norm on scenario valuations lead to linear n-widths that call for rank-reduction in average squared error. Results in the n-width literature show that the eigenfunctions of the exponential product operator furnish optimal approximations in this case; these eigenfunctions are the oblate spheroidal wave functions of order zero, which are special functions of classical mathematical physics, so I name the novel approximation that uses them the "oblate approximation." Of the four approximations I introduce, the oblate approximation is closest to a well-studied approximation, the use of *prolate* spheroidal wave functions of order zero to approximate band-limited functions (see Osipov *et al.* (2013)). The prolate problem is qualitatively different, but deserves mention.

Finally, the n-width literature also contains theory around interpolatory approximations for root-mean-square error; this theory leads me to my new interpolatory approximation of the exponential product operator, which is optimal for root-mean-square (and thus mean-square) error, just as the oblate approximation is (optimal approximations in the n-width sense are generally not unique). Like the numerically-optimal approximation, the interpolatory approximation can be interpreted as providing an approximating portfolio.

Although my approach to portfolio approximation is very different from any used in the literature, a review of prior work may still be interesting from a historical perspective. Portfolio approximations have been computed and used since 1558, when Jean Trenchant first published a table of present-value factors (Goetzmann (2005), page 143). A problem of compound interest stimulated Bernoulli (1690) to place what appear to have been the first bounds on the number e. Moving forward in time, Macaulay (1938) introduced the concept of duration, which has been rediscovered and reinterpreted many times since (see Hicks (1939) and Samuelson (1945); Weil (1973) provides a nice review of the early history of duration). Redington (1952) was the first to discuss the immunization of a portfolio against interest-rate shocks. Although important contributions regarding duration were made by Fisher & Weil (1971) (a key paper on duration, as it allowed for a non-flat interest rate curve and conducted significant empirical work) and extensions were developed by Fong & Vasicek (1984), Chambers *et al.* (1988), Ho (1992), and Nawalkha & Chambers (1996), among others, the finance literature has gravitated in recent years toward problems in which an asset pricing model is assumed. This shift may have begun with the work of Vasicek (1977); it was certainly advocated in the important work of Cox *et al.* (1979) (the first work to relate duration to modern term-structure models) and Ingersoll *et al.* (1978), and it gathered force through the seminal contributions of Cox *et al.* (1985) (which was a working paper at least as early as 1978) and Heath *et al.* (1992). Jeffrey (2000) gives a helpful discussion of some more recent literature on duration and its links to term-structure models.

The focus in all of the more recent work is on derivative-based approximations (such as duration and its higher-order analogs, or sensitivities in term-structure models) that consider only infinitesimally-small shifts in variables of interest. This is clearly distinct from my approach, which is based on considering approximation across an entire range of scenarios that may involve very large shifts in the variables of interest, as the example above shows.

The prior work does include some consideration of an approximation that I use as a comparator (and which has already appeared in the example above), the Taylor-series approximation. This seems to appear first in Redington (1952). It is also explored by Chambers *et al.* (1988) from a more modern perspective.

As the example above suggested, the four novel approximations I introduce (numerically-optimal, Mathieu, interpolatory, and oblate) are all greatly superior to the standard Taylor-series approximation.

Conceptually, the numerically-optimal and Mathieu approximations (focused on worst-case error) are motivated by the same interest in *robustness* that has generated a very fruitful research program exploring uncertainty aversion in asset pricing and macroeconomics (Hansen & Sargent (1995), Hansen *et al.* (1999), Anderson *et al.* (2003), Hansen *et al.* (2006), Hansen & Sargent (2008), and Hansen & Sargent (2012) are just a few of the many important works that make up this program). A similar motivation has also led to the econometric investigations of Chamberlain (2000) and Chamberlain (2001).

Technically, the interpolatory and oblate approximations exploit eigenstructure properties of totally positive kernels. These eigenstructure properties date back to Kellogg (1918) and can be derived using Perron-Frobenius theory (Gantmacher & Krein (2002), Pinkus (1996)). Perron-Frobenius theory has also been found useful in the research of Hansen & Scheinkman (2009), Hansen (2012), Borovička *et al.* (2014), and Ross (2014).

The plan of my paper is as follows. In Section 2, I show how portfolios of bonds, interest rate swaps, credit default swaps, and equities can be put into exponential-affine form for scenario valuation purposes. In Section 3, I demonstrate that scenario valuation of any portfolio in exponential-affine form is equivalent to the application of an exponential product operator to the transformed cashflows of the portfolio. In Section 4, I review the theory of n-widths. In Section 5, I discuss the theory of Chebyshev systems and total positivity and prove a useful result. In Section 6, I derive a new lower bound on the worst-case error of any approximation to an exponential product operator and I provide novel methods for its computation. In Section 7, I introduce the first of my new approximations to the exponential product operator, the numerically-optimal approximation; I provide methods to construct and evaluate the numerically-optimal approximation, and I show using an interval-analysis method that it achieves (within rounding error) the lower bound developed in Section 6 for each c and n (approximation rank) considered. Subsection 7.1 discusses the approximating-portfolio interpretation of the numerically-optimal approximation. In Section 8, I introduce and characterize the novel oblate approximation of the exponential product operator and provide methods to build and evaluate it. In Subsection 8.5, I provide the new interpolatory approximation of the exponential product operator with methods for its construction and evaluation (and I discuss its interpretation in approximating-portfolio terms). In Section 9, I lay out the novel Mathieu approximation to the exponential product operator and give methods for its assembly and use. All four of my new approximations to the exponential product operator are compared to each other and to Taylor-series and univariate-Remez approximations in Section 10. I provide examples to illustrate the usefulness of my approximations in Section 11, and Section 12 concludes the paper.

2 Portfolios in Exponential-Affine Form

I consider portfolios whose present values across scenarios may be represented in an exponential-affine form

$$V(S) \quad = \quad \sum_{i=1}^{I} a_i \sum_{j=1}^{T_i} \exp\left(-\sum_{k=1}^{M_i} (X_{ijk} + S_{ijk})\, t_{ijk}\right) f_{ij}, \tag{1}$$

where i indexes distinct portfolio components, j indexes cashflows within a portfolio component, and k indexes state variables. The upper bounds of summation T_i and M_i are subscripted to indicate that different cashflow dates and different state variables may be relevant in different components of the portfolio (for example, different issuers in a portfolio of corporate bonds will generally have different payment dates and different default rates). X_{ijk} is the base-case value of state variable k that is relevant for cashflow j in portfolio component i (for example, X_{ijk} might be the base-case interest rate from the present to the payment date of cashflow j, the base-case default rate for issuer i from the present to the payment date of cashflow j in a portfolio of corporate bonds or credit default swaps, or the base-case dividend growth rate for stock i from the present to the payment date of cashflow j in an equity portfolio). S_{ijk} is the change in the value of state variable k that is relevant for cashflow j in portfolio component i; the X_{ijk} are fixed across scenarios, while the S_{ijk} may change from scenario to scenario. Thus, a value for all of the S_{ijk} defines a single scenario. t_{ijk} is a multiplier which typically represents the time from the present to a relevant date (such as the payment date of the cashflow or the fixing date of a relevant index).

I make no assumptions about any asset pricing model here or later. Any portfolio made up of government bonds, corporate bonds, interest rate swaps, credit default swaps, or equities (when regarded as claims to streams of future dividends) satisfies (1). Portfolios including options generally do not satsify (1).

For the base case, in which $S_{ijk} = 0$, the representation (1) imposes no real restriction for portfolios of the instruments I study; it merely expresses discount factors, survival probabilities, *etc.* in terms of continuously-compounded zero rates, default rates, and so on. However, (1) does impose a restriction on the way in which scenarios change discount factors, survival probabilities, and other relevant pricing variables: under (1), scenarios involve shifting continuously-compounded zero rates, continuously-compounded default rates, and other continuously-compounded rates relevant for valuation.

Any given scenario which generates (from a base-case set of discount factors, survival probabilities, and so on) a new set of discount factors, survival probabilities, *etc.* can always be written in terms of shifts in continuously-compounded zero rates, continuously-compounded default rates, and other continuously-compounded rates impacting valuation. With that said, if a specific scenario shift S_{ijk} is multiplied by c to form a one-parameter family of scenario shifts cS_{ijk} indexed by c (for example, if the one-parameter family represents parallel shifts to the curve of continuously-compounded Treasury zero rates), then the representation (1) does generate a different mapping from values of c to portfolio values than if the family of scenario shifts were moves in, say, semiannually-compounded rates.

My focus on changes in continuously-compounded rates does not create problems in interpreting or communicating scenario results, since it is always possible to express any one scenario in terms of some other quotation convention (e. g., semiannually- or annually-compounded rates) or to show the impact of a scenario on a benchmark such as a 10-year Treasury note yield or the five-year point of a credit default swap curve for a given issuer.

As noted above, I do not assume any asset pricing model, but affine term structure models certainly produce values in the form (1) for portfolios of the instrument types considered here (bonds, interest rate swaps, credit default swaps, or equities). The assumption of an affine term structure model is thus sufficient, but definitely not necessary, for (1) to hold; see Duffie *et al.* (2000) and Piazzesi (2010) for discussions of affine models. Even some variants of affine models, such as Campbell *et al.* (2013) and Greenwood & Vayanos (2014), provide enough structure that my approximations may be useful. Further, the exponential-affine form shows up in other contexts; my results might be employed for comparative statics in the model of Campbell *et al.* (2014) or in a more econometric setting such as that of Barndorff-Nielsen & Shephard (2001).

2.1 Treasuries

A portfolio of Treasury bonds satisfies (1), since for such a portfolio

$$V_{\mathrm{UST}}(S) \quad = \quad \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_j\right) t_j\right) f_j, \tag{2}$$

where f_j is the total payment of Treasury cashflows the portfolio will receive on date j, t_j is the time from the present to date j, r_j is the base-case zero-coupon discount rate from the present to date j, and $\tilde{S}_j$ is the scenario shift in the zero-coupon interest rate from the present to date j. This expression is of the form (1) where $I = 1$ (there is only one portfolio component), $a_1 = 1$, $T_1 = T$, $M_1 = 1$ (there is only one state variable), $X_{1j1} = r_i$, $S_{1j1} = \tilde{S}_j$, and $t_{1j1} = t_j$. Note that yield differentials between Treasury bonds do not pose a problem, as (2) requires only a list of rates $(r_1, \cdots, r_T)$ which recover the base-case value of the portfolio.

2.2 Corporate Bonds

A portfolio of one issuer's corporate (or sovereign) bonds which recover zero in the event of default satisfies (1), since the value of such a portfolio is

$$V^{\mathrm{NoR}}_{\mathrm{CORP}}(S) = \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j - \left(\lambda_j + \tilde{S}_{j2}\right) t_j\right) f_j, \tag{3}$$

where f_j is the total payment of (defaultable, zero-recovery) corporate cashflows the portfolio will receive on date j provided that the issuer is not in default as of date j, t_j is the time from the present to date j, r_j is the base-case riskless zero-coupon interest rate from the present to date j, $\tilde{S}_{j1}$ is the shift in the riskless zero-coupon interest rate from the present to date j, λ_j is the base-case average default intensity from the present to date j, and $\tilde{S}_{j2}$ is the shift in the average default intensity from the present to date j. To see that this expression is of the form (1), set $I = 1$ (there is one portfolio component), $a_1 = 1$, $T_1 = T$, $M_1 = 2$ (two state variables are relevant to the portfolio), $X_{1j1} = r_i$, $X_{1j2} = \lambda_i$, $S_{1j1} = \tilde{S}_{j1}$, $S_{1j2} = \tilde{S}_{j2}$, $t_{1j1} = t_j$, $t_{1j2} = t_j$.

As noted in the first paragraph of this section, a portfolio of zero-recovery corporate bonds which includes the bonds of multiple issuers can be put into the form (1) by treating each issuer's bonds as a separate component of the portfolio (so that I is the number of distinct issuers in the portfolio) and then handling each component using (3).

A portfolio of one issuer's corporate (or sovereign) bonds which have nonzero recovery in the event of default also satisfies (1), but showing this requires a bit more manipulation than the zero-recovery case of (3).

$$\begin{aligned} & V^{\mathrm{R}}_{\mathrm{CORP}}(S) \\ = & \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j - \left(\lambda_j + \tilde{S}_{j2}\right) t_j\right) f_j \\ + & \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j\right) \begin{pmatrix} \exp\left(-\left(\lambda_{j-1} + \tilde{S}_{j-1,2}\right) t_{j-1}\right) \\ -\exp\left(-\left(\lambda_j + \tilde{S}_{j2}\right) t_j\right) \end{pmatrix} \rho_j \end{aligned} \tag{4}$$

$$
\begin{aligned}
= & \sum_{j=1}^{T} \exp\left(-\left(r_j+\tilde{S}_{j1}\right)t_j-\left(\lambda_j+\tilde{S}_{j2}\right)t_j\right)\left(f_j-\rho_j\right) \\
& +\sum_{j=1}^{T} \exp\left(-\left(r_j+\tilde{S}_{j1}\right)t_j-\left(\lambda_{j-1}+\tilde{S}_{j-1,2}\right)t_{j-1}\right)\rho_j, \qquad (5)
\end{aligned}
$$

where in (13) all variables are defined as in (3) except that there may be some recovery in the event of default; the amount recovered by the portfolio if default occurs after date $j-1$ and on or before date j is ρ_j. There is an implicit assumption here that default occurs exactly on some date j, $j=1,\ldots,T$, but I will relax this momentarily. (5) follows from (13) by elementary rearrangement. Thus, $I=2$ (there are two portfolio components, a no-default component and a recovery component), $a_1=1$, $a_2=1$, $T_1=T$, $T_2=T$, $M_1=2$ (two state variables are relevant in the first portfolio component), $M_2=2$ (two state variables are relevant in the second portfolio component), $X_{1j1}=r_j$, $X_{1j2}=\lambda_j$, $X_{2j1}=r_j$, $X_{2j2}=\lambda_{j-1}$, $S_{1j1}=\tilde{S}_{j1}$, $S_{1j2}=\tilde{S}_{j2}$, $S_{2j1}=\tilde{S}_{j1}$, $S_{2j2}=\tilde{S}_{j-1,2}$, $t_{1j1}=t_j$, $t_{1j2}=t_j$, $t_{2j1}=t_j$, and $t_{2j2}=t_{j-1}$.

A portfolio of corporate bonds with nontrivial recoveries which includes the bonds of multiple issuers can be put into the form (1) by treating each issuer's bonds as two separate components of the portfolio (so that I is twice the number of distinct issuers in the portfolio) and then handling each issuer's bonds using (5).

2.3 Interest Rate Swaps

The value of the fixed leg of an interest rate swap can be expressed as

$$
\begin{aligned}
V_{\text{fixed}}(S) &= NC_{\text{fixed}}\sum_{j\in J}\exp\left(-\left(r_j+\tilde{S}_{j1}\right)t_j\right)\tau_j \qquad (6)\\
&= \sum_{j\in J}\exp\left(-\left(r_j+\tilde{S}_{j1}\right)t_j\right)f_j, \qquad (7)
\end{aligned}
$$

where in (6) N is the notional of the swap, C_{fixed} is the fixed-leg rate of the swap, τ_j is the accrual year fraction for the fixed-leg payment on date j (computed with the appropriate daycount convention, *e. g.*, 30 / 360 in the vanilla USD rate swap market), t_j is the time from the present to date j, r_j is the base-case zero-coupon discount rate to date j, and $\tilde{S}_{j1}$ is the scenario shift in the zero-coupon discount rate to date j. The index set $J\subset\{1,\ldots,T\}$ is used to allow the indices j to skip some of the numbers $1,\ldots,T$, since fixed-leg payment dates are typically a subset of all of the payment dates of a swap (in the USD vanilla interest rate swap market, fixed-leg payments are semiannual and floating-leg payments are quarterly). The equality (7) follows by setting $f_j\equiv NC_{\text{fixed}}\tau_j$.

The value of the floating leg of an interest rate swap admits expression as

$$\begin{aligned}
& V_{\text{float}}(S) \\
= \quad & N \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_i\right) \tilde{\tau}_j F_j && (8) \\
= \quad & N \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j\right) \left(\frac{\exp\left(-\left(p_{j-1} + \tilde{S}_{j-1,2}\right) t_{j-1}\right)}{\exp\left(-\left(p_j + \tilde{S}_{j2}\right) t_j\right)} - 1\right) && (9) \\
= \quad & \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j - \left(p_{j-1} + \tilde{S}_{j-1,2}\right) t_{j-1} + \left(p_j + \tilde{S}_{j2}\right) t_j\right) g_j \\
& - \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j\right) g_j, && (10)
\end{aligned}$$

where in (8) N is the notional of the swap, $\tilde{\tau}_j$ is the accrual year fraction for the floating-leg payment on date j (computed with the proper daycount fraction, *e. g.* , Actual / 360 in the vanilla USD rate swap market), F_j is the projected forward for the index whose fixings determine floating-leg payments (for instance, USD 3-month LIBOR), and r_j, S_{j1}, and t_j have the same meanings as in (6), but with respect to floating-leg payment dates. The equality (9) follows from representing F_j as $\frac{1}{\tilde{\tau}_j}\left(\frac{\exp(-(p_{j-1}+S_{j-1,2})t_{j-1})}{\exp(-(p_j+S_{j2})t_j)} - 1\right)$, where p_j is a zero-coupon projection rate (not a discount rate) and S_{j2} is the scenario shift in this rate. Equality (10) is obtained by rearranging terms and using the identity $\exp(x+y) = \exp(x)\exp(y)$ and setting $g_j \equiv N$.

Combining the results of (7) and (10) gives the following representation for the value of a single interest rate swap:

$$\begin{aligned}
& V_{\text{rateswap}} \\
= & \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j\right) h_j \\
- & \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j - \left(p_{j-1} + \tilde{S}_{j-1,2}\right) t_{j-1} + \left(p_j + \tilde{S}_{j2}\right) t_j\right) g_j, \quad (11)
\end{aligned}$$

where in (11) I have defined $h_j \equiv g_j + f_j 1\{j \in J\}$ (so that h_j is g_j plus, if j is a fixed-leg payment date, f_j). To show that (11) is of the form (1), set $I = 2$ (there are two portfolio components, the first depending only on discount rates and the second depending on the projections of the floating-leg index as well as discount rates), $a_1 = 1$, $a_2 = -1$, $T_1 = T$, $T_2 = T$, $M_1 = 1$ (one state variable is relevant in the first portfolio component), $M_2 = 3$ (three state variables are relevant in the second portfolio component), $X_{1j1} = r_j$, $X_{2j1} = r_j$, $X_{2j2} = p_{j-1}$, $S_{1j1} = \tilde{S}_{j1}$, $S_{2j1} = \tilde{S}_{j1}$, $S_{2j2} = \tilde{S}_{j-1,2}$, $S_{2j3} = \tilde{S}_{j2}$, $t_{1j1} = t_j$, $t_{2j1} = t_j$, $t_{2j2} = t_{j-1}$, and $t_{2j3} = t_j$.

An expression for the value of a portfolio of interest rate swaps follows from summing across the swaps in the portfolio:

$$
\begin{aligned}
& V_{\mathrm{IRS}}(S) \\
= & \sum_{n=1}^{U} \exp\left(-\left(r_n + S_{n1}\right) t_n\right) \bar{h}_n \\
- & \sum_{n=1}^{U} \exp\left(-\left(r_n + S_{n1}\right) t_n - \left(p_{n-1} + S_{n-1,2}\right) t_{n-1} + \left(p_n + S_{n2}\right) t_n\right) \bar{g}_n,
\end{aligned} \tag{12}
$$

where the equality (12) follows from summation of (11) across swaps in the portfolio, so that U is the total number of floating-leg payment dates across the portfolio, n indexes the floating-leg payment dates of the portfolio, $\bar{h}_n$ is the sum across swaps in the portfolio of all of the h_j terms in (11) that occur on date n, and $\bar{g}_n$ is the sum across swaps in the portfolio of all of the g_j terms in (11) that occur on date n. The same mapping I used to show that (11) is of the form (1) demonstrates that (12) is also of the form (1).

When comparing (12) to the Treasury portfolio representation (2), it is important to note that the r_i terms in the two expressions may come from different zero-coupon discounting curves; in the Treasury case, the appropriate zero-coupon curve is the one which prices the Treasury portfolio correctly, while an interest rate swap is a bilateral contract, so the appropriate discount curve involves varying amounts of credit risk based on the credit support annex (CSA) of the ISDA master agreement between the two parties to the swap. The simplest case involves a CSA which specifies that only very high-quality collateral (cash, Treasuries) may be posted; in this case any mark-to-market on the swap is fully and reliably collateralized, limiting counterparty credit risk to one-day changes in mark-to-market in the event of default. In this situation the standard market practice is to discount based on rates in the overnight indexed swap (OIS) market, which references overnight interbank rates typically linked to central bank policy (the Federal funds rate in USD, Eonia in EUR, and Sonia in GBP).

2.4 Credit Default Swaps

The arguments I made concerning corporate bonds (with recovery in the event of a default) also work for credit default swaps, so the value of a portfolio of credit default swaps referencing a given set of bonds can be expressed as

$$
\begin{aligned}
& V_{\mathrm{CDS}}^{\mathrm{R}}(S) \\
= \quad & \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j - \left(\lambda_j + \tilde{S}_{j2}\right) t_j\right) f_j \\
+ \quad & \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j\right) \begin{pmatrix} \exp\left(-\left(\lambda_{j-1} + \tilde{S}_{j-1,2}\right) t_{j-1}\right) \\ -\exp\left(-\left(\lambda_j + \tilde{S}_{j2}\right) t_j\right) \end{pmatrix} g_j
\end{aligned} \tag{13}
$$

$$
\begin{aligned}
= & \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j - \left(\lambda_j + \tilde{S}_{j2}\right) t_j\right) (f_j - g_j) \\
& + \sum_{j=1}^{T} \exp\left(-\left(r_j + \tilde{S}_{j1}\right) t_j - \left(\lambda_{j-1} + \tilde{S}_{j-1,2}\right) t_{j-1}\right) g_j, \qquad (14)
\end{aligned}
$$

is of the form (1); f_j is the flow to be received (if positive) or paid (if negative) if there is no default on or before date j while g_j is the flow to be received (if positive) or paid (if negative) if default occurs after date $j-1$ but on or before date j. For a single credit default swap in which protection is purchased (a pay fixed, receive floating position), the f_j equal the negative of the year fraction from $j-1$ to j times the contractual spread times the swap notional, while the g_j are positive and equal swap notional times the quantity (one minus recovery fraction). The interpretation of the other variables is as in (5).

As in the case of corporate bonds, a portfolio of credit default swaps referencing different sets of bonds (credit default swaps referencing different issuers) can be put into the form (1) by dividing the portfolio by issuer and regarding each issuer's credit default swaps as two separate components of the portfolio which are then dealt with using (14).

2.5 Equities

An equity portfolio can be regarded as a claim on a future stream of dividends. When viewed in this way, a portfolio of equities can be expressed in the form

$$
V_{\text{EQUITY}}(S) = \sum_{i=1}^{I} a_i \sum_{j=1}^{T} \exp\left(-\left(r_j + S_{j1}\right) t_j + \left(\gamma_{ij} + S_{ij2}\right) t_j\right) d_{ij}, \qquad (15)
$$

where i indexes the number of different equities in the portfolio, a_i is the number of shares of stock i that the portfolio owns (it is positive for long positions and negative for short positions), j indexes future dividend payment dates (across all issuers) for the portfolio, r_j is the riskless zero-coupon discounting rate from the present to date j, S_{j1} is the scenario shift in the riskless zero-coupon discounting rate to date j, t_j is the time from the present to date j, γ_{ij} is the dividend growth rate of stock i from the present to date j, and S_{ij2} is the scenario shift in the dividend growth rate of stock i to date j, and $d_{ij} = d_i$, the current per-share dividend of stock i, if date j is a dividend payment date for stock i (otherwise $d_{ij} = 0$, and this is why the j subscript is necessary). The expression (15) is clearly of the form (1).

3 Exponential Product Operators and Portfolios in Exponential-Affine Form

Definition 3.1. The exponential product operators are the family of integral operators

$$P_c [f] (y) \equiv \int_{-1}^{1} \exp (cxy) df (x) , \tag{16}$$

where the family is indexed by the real parameter $c > 0$, f may be any signed measure of bounded total variation on $[-1, 1]$, and $y \in [-1, 1]$.

I first consider the value of a simple portfolio of the form (1), with one portfolio component ($I = 1$) and one state variable ($M_1 = 1$), across a one-parameter family of scenarios. The parameter $y \in [-1, 1]$ indexes the family; y multiplies the base scenario $(S_1, S_2, \ldots, S_T)$. $V (y)$ denotes the portfolio value as a function of y.

Proposition 3.1. *If the value of a portfolio as a function of a scenario multiplier $y \in [-1, 1]$ satisfies*

$$V (y) = \sum_{j=1}^{T} \exp (- (X_j + yS_j) t_j) f_j, \tag{17}$$

then

$$V (y) = P_c \left[\tilde{f}\right] (y) , \tag{18}$$

where $c = \max_{j=1,\ldots,T} |S_j t_j|$ and $\tilde{f}$ is the signed measure on $[-1, 1]$ that places mass $\tilde{f}_j = \exp (-X_j t_j) f_j$ on point $-S_j t_j /c$ for $j = 1, \ldots, T$.

Proof.

$$\begin{aligned}
V (y) &= \sum_{j=1}^{T} \exp (- (X_j + yS_j) t_j) f_j & (19)\\
&= \sum_{j=1}^{T} \exp (-yS_j t_j) \tilde{f}_j & (20)\\
&= \sum_{j=1}^{T} \exp (-cy\omega_j) \tilde{f}_j & (21)\\
&= \int_{-1}^{1} \exp (-cy\omega) dg (\omega) & (22)\\
&= \int_{-1}^{1} \exp (cxy) d\tilde{f} (x) & (23)\\
&= P_c \left[\tilde{f}\right] (y) , & (24)
\end{aligned}$$

where the first equality is by the assumption of the proposition, the second equality follows from the definition $\tilde{f}_j = \exp(-X_j t_j) f_j$ in the statement of the proposition, the third equality is a result of setting $c = \max_{j=1,\ldots,T} |S_j t_j|$ (as in the statement of the proposition) and letting $\omega_j = S_j t_j / c$, the fourth equality recognizes that for each $j = 1, \ldots, T$ the quantity ω_j has absolute value no greater than one by its definition (so the sum can be expressed as an integral with respect to the signed measure g placing mass $\tilde{f}_j$ on the point $\omega_j = S_j t_j / c$), the fifth equality changes variables from ω to $x = -\omega$ and thus changes measures from g to $\tilde{f}$ as defined in the statement of the proposition, and the sixth equality is by Definition 3.1. $\square$

Proposition 3.1 shows that the problem of approximating a portfolio of the exponential-affine form (1), with one component and one state variable, is equivalent to the problem of approximating an exponential product operator. The two measures of approximation error that I will consider are worst-case error (the supremum over $x, y \in [-1, 1]$ of the absolute value of approximation error) and root-mean-square error.

Worst-case error is appropriate for risk measurement, where attention is focused on extreme outcomes. A crucial equality in my analysis of worst-case error is, for any approximation f_{approx}:

$$\begin{aligned} &\sup_{x,y \in [-1,1]} \left| \exp(cxy) - f_{\text{approx}}(x, y) \right| \\ = &\sup_{h: \|h\|_{TV} \leq 1} \left\{ \sup_{y \in [-1,1]} \left| \int_{-1}^{1} \left(\exp(cxy) - f_{\text{approx}}(x, y) \right) dh(x) \right| \right\}, \end{aligned} \tag{25}$$

where M_{TV} is the set of all signed measures on $[-1, 1]$ with bounded total variation (this is the normed space of signed measures equipped with the total variation norm $\|\cdot\|_{TV}$). To demonstrate this equality, simply note that the lefthand side is no greater than the righthand side, since any (x, y) achieving the supremum on the lefthand side can be matched on the righthand side by using the same y and an h that places unit measure on the x. But if the righthand side were greater than the lefthand side, there would have to exist some y and some x_h in the support of h, both in $[-1, 1]$, such that $|\exp(cx_h y) - f_{\text{approx}}(x_h, y)|$ was greater than the lefthand side, a contradiction (since the lefthand side is the supremum over such x and y).

Using (25), I recast the search for the best approximation to the exponential product kernel under worst-case error measurement as a hunt for an integral operator that approximates the exponential product operator (16), where both the approximation and the exponential product operator are viewed as mapping the set of signed measures on $[-1, 1]$ with bounded total variation to the set of continuous functions on $[-1, 1]$.

In approaching practical approximation problems, it is usually helpful to recenter the variables of interest (the $S_j t_j$ and the y of Proposition 3.1) as well as normalizing their scales (as is done in Proposition 3.1). Consider a variant of the setting of Proposition 3.1 in which the value of a portfolio as a function of a scenario multiplier $y \in [m_y - r_y, m_y + r_y]$ satisfies (17) (only the range of y has changed). Then

$$\begin{aligned}
V(y) &= \sum_{j=1}^{T} \exp\left(-\left(X_j + yS_j\right)t_j\right) f_j && (26)\\
&= \sum_{j=1}^{T} \exp\left(-yS_j t_j\right) \tilde{f}_j && (27)\\
&= \exp\left(m_{St} y\right) \sum_{j=1}^{T} \exp\left(-\underline{c}\tilde{y}\omega_j\right) \exp\left(m_y\left(S_j t_j - m_{St}\right)\right) \tilde{f}_j && (28)\\
&= \exp\left(m_{St} y\right) \sum_{j=1}^{T} \exp\left(-\underline{c}\tilde{y}\omega_j\right) \tilde{g}_j && (29)\\
&= \exp\left(m_{St} y\right) \int_{-1}^{1} \exp\left(-\underline{c}\tilde{y}\omega\right) d\tilde{g}\left(\omega\right) && (30)\\
&= \exp\left(m_{St} y\right) \int_{-1}^{1} \exp\left(\underline{c}x\tilde{y}\right) d\tilde{h}\left(x\right) && (31)\\
&= \exp\left(m_{St} y\right) P_{\underline{c}}\left[\tilde{h}\right]\left(\tilde{y}\right), && (32)
\end{aligned}$$

where the first two lines are as in the proof of Proposition 3.1, and the third equality follows from defining $\tilde{y} \equiv (y - m_y)/r_y$ (so that $\tilde{y} \in [-1, 1]$), $\omega_j \equiv (S_j t_j - m_{St})/r_{St}$ (where m_{St} is the average of the maximum and minimum of the $S_j t_j$ and r_{St} is half of the difference between that maximum and that minimum), and $\underline{c} = r_m r_y$. The fourth equality lets $\tilde{g}_j \equiv \exp\left(m_y\left(S_j t_j - m_{St}\right)\right)\tilde{f}_j$ (noting that m_y is a constant with respect to y, it is just the midpoint of the range of y). The fifth, sixth, and seventh lines follow exactly as in the proof of Proposition 3.1.

The utility of (32) comes from the size of $\underline{c}$, which may be significantly smaller (with $P_{\underline{c}}$ thus significantly easier to approximate) than if no recentering had been done (of course, one could always rescale without recentering, in the spirit of Proposition 3.1, but the resulting c would never be smaller than $\underline{c}$ and would typically be larger). The added computational complexity of this approach is minimal relative to the accuracy gained.

As the analysis in Section 2 shows, many instruments are of the form (1) but with more than one component or more than one state variable (or both). I will show in Section 11 that the link demonstrated by Proposition 3.1 has important implications even in these more complex cases.

4 n-Widths

The study of n-widths is the branch of approximation theory that focuses on the best (in some specific sense) approximation of a given set by any finite-dimensional subspace. This contrasts with the literature on metric entropy, which studies the best approximation of a given set by any *finite*, rather than finite-dimensional, set. Some workers in approximation theory use the term "diameter" to mean "n-width," but I shall follow the bulk of the literature in using "n-width." The notation of this section primarily follows Pinkus (1985a); other monographs concerned with n-widths include Korneichuk (1991) (Chapter 8), Lorentz (1986) (Chapter 9), Lorentz *et al.* (1996) (Chapters 13 and 14), and Tikhomirov (1990) (Chapter 3). Pinkus (1985a), Chapter II, Section 7 provides basic definitions of n-widths of mappings of unit balls, and I follow him in the definitions below.

The study of n-widths began with Kolmogorov (1936), and the following definition essentially belongs to him though I use a version provided by Pinkus (1985a) (Definition 7.1 on page 29 in Section 7 of Chapter II).

Definition 4.1. Let $T \in L(X, Y)$, the set of continuous linear operators from X to Y (where both X and Y are normed linear spaces). The Kolmogorov n-width is defined as

$$d_n(T(X); Y) = \inf_{Y_n} \sup_{\|x\|_X \leq 1} \inf_{y \in Y_n} \|x - y\|_Y, \tag{33}$$

where the leftmost infimum is taken over all n-dimensional subspaces Y_n of Y. If

$$d_n(T(X); Y) = \sup_{\|x\|_X \leq 1} \inf_{y \in Y_n} \|x - y\|_Y \tag{34}$$

for a subspace Y_n of dimension no greater than n, then Y_n is an optimal subspace for $d_n(T(X); Y)$.

After the work of Kolmogorov (1936), renewed interest in n-widths was sparked by Tikhomirov (1960) (and later Tikhomirov (1969)). Tikhomirov (1960) introduced the idea of a linear n-width, which will be particularly important in my investigation of portfolio approximation. Pinkus (1985a) gives the following definition (Definition 7.3 on page 30 in Section 7 of Chapter II).

Definition 4.2. Let $T \in L(X, Y)$, the set of continuous linear operators from X to Y (where both X and Y are normed linear spaces). The linear n-width is defined as

$$\delta_n(T(X); Y) = \inf_{P_n} \sup_{\|x\|_X \leq 1} \|Tx - P_n x\|_Y, \tag{35}$$

where the infimum is taken over all continuous linear operators taking X into Y of rank n or less. If

$$\delta_n(T(X); Y) = \sup_{\|x\|_X \leq 1} \|Tx - P_n x\|_Y \tag{36}$$

for a continuous linear operator P_n of rank no greater than n, then P_n is an optimal linear operator for $\delta_n(T(X); Y)$.

Gel'fand, in conversations with Tikhomirov around the work in Tikhomirov (1960), prompted Tikhomirov to introduce the notion of a Gel'fand n-width (see Tikhomirov (1990), page 188 in Section 1.2 of Chapter 3, for a brief discussion). When considering n-widths of mappings of the unit ball, the Gel'fand n-width is defined by Pinkus (1985a) (Definition 7.2 on page 30 in Section 7 of Chapter II) as follows.

Definition 4.3. Let $T \in L(X, Y)$, the set of continuous linear operators from X to Y (where both X and Y are normed linear spaces). The Gel'fand n-width is defined as

$$d^n (T(X); Y) = \inf_{L^n} \sup_{\substack{\|x\|_X \leq 1 \\ x \in L^n}} \|Tx\|_Y , \tag{37}$$

where the infimum is taken over all subspaces L^n of X of codimension n, that is, all subspaces L^n such that there are n continuous, linearly independent linear functionals $\{F_i\}_{i=1}^n$ on X for which

$$L^n = \{x : x \in X, F_i x = 0, i = 1, \ldots, n\} . \tag{38}$$

If

$$d^n (T(X); Y) = \sup_{\substack{\|x\|_X \leq 1 \\ x \in L^n}} \|Tx\|_Y \tag{39}$$

for a subspace L^n of codimension no greater than n, then L^n is an optimal subspace for $d^n (T(X); Y)$.

Gel'fand n-widths are important in interpolation and recovery problems.

The notion of a Bernstein n-width is useful in providing lower bounds for other n-widths; Pinkus (1985a) gives the definition below (his Definition 2.2 on page 149 in Section 2 of Chapter V).

Definition 4.4. Let $T \in L(X, Y)$, the set of continuous linear operators from X to Y (where both X and Y are normed linear spaces). The Bernstein n-width is defined as

$$b_n (T(X); Y) = \sup_{Y_{n+1}} \inf_{\substack{Tx \in Y_{n+1} \\ Tx \neq 0}} \frac{\|Tx\|_Y}{\|x\|_X}, \tag{40}$$

where Y_{n+1} is any subspace of span $\{Tx : x \in X\}$ of dimension at least $n+1$. If

$$b_n (T(X); Y) = \inf_{\substack{Tx \in Y_{n+1} \\ Tx \neq 0}} \frac{\|Tx\|_Y}{\|x\|_X} \tag{41}$$

for a subspace Y_{n+1} of dimension no less than $n+1$, then Y_{n+1} is an optimal subspace for $b_n (T(X); Y)$.

There are other n-widths defined in the literature (such as the Aleksandrov n-width and the projective n-width), but the four defined above will be most useful here. They are related by the following inequalities:

Theorem 4.1. *Let $T \in L(X, Y)$, the set of continuous linear operators from X to Y (where both X and Y are normed linear spaces). Then*

$$\delta_n (T(X); Y) \geq d_n (T(X); Y) \geq b_n (T(X); Y) \tag{42}$$

and

$$\delta_n (T(X); Y) \geq d^n (T(X); Y) \geq b_n (T(X); Y). \tag{43}$$

Proof. As shown in Pinkus (1985a), Proposition 1.6 on page 13 in Section 1 of Chapter II, $d_n (T(X); Y) \geq b_n (T(X); Y)$ (this proposition in Pinkus (1985a) is actually in the more general context of a subset A of Y that need not be a mapping of a unit ball, and as Pinkus (1985a) observes on page 149 in Section 2 of Chapter V, it thus holds in the mapping-of-a-unit-ball context). This is a nontrivial inequality that ultimately requires the profound topological result of Borsuk (1933) known as the Borsuk antipodality theorem; Tikhomirov (1960) was the first to use the Borsuk antipodality theorem to directly develop a lower bound for n-widths, though an equivalent result appeared earlier in Krein *et al.* (1948) (this earlier paper is difficult to obtain, and the book of Gohberg & Krein (1969), which reproduces its argument, is frequently cited). The inequality $\delta_n (T(X); Y) \geq d_n (T(X); Y)$ appears on page 21 in Section 4 of Chapter II of Pinkus (1985a) in the more general context of a subset A of Y that need not be a mapping of a unit ball, but it certainly holds in the mapping-of-a-unit-ball context as noted in the Remark on page 30 in Section 7 of Chapter II of Pinkus (1985a). Indeed, it is clear that the best linear approximation can never be superior to the best approximation without any linearity restriction.

Pinkus (1985a), Proposition 3.5 on page 19 in Section 3 of Chapter II, shows that $d^n (T(X); Y) \geq b_n (T(X); Y)$ (this proposition in Pinkus (1985a) is actually in the more general context of a subset A of Y that need not be a mapping of a unit ball, and thus holds in the mapping-of-a-unit-ball context). The inequality $\delta_n (T(X); Y) \geq d^n (T(X); Y)$ is Proposition 7.7 on page 32 in Section 7 of Chapter II of Pinkus (1985a). □

4.1 Optimal Computational Methods and n-Widths

The importance of n-widths in studying the optimality of computational methods has long been appreciated. In Traub & Woźniakowski (1980), the Gel'fand n-width (Section 6 of Chapter 2, pages 41 - 47), the Kolmogorov n-width (Section 4 of Chapter 7, pages 157 - 159), and the linear n-width (Section 5 of Chapter 3, pages 64 - 67, where it is called the "linear Kolmogorov n-width" and is defined slightly differently) are each related to the optimality of a distinct class of methods.

The compressed sensing literature has appreciated the usefulness of prior work on Gel'fand n-widths, in particular; see the seminal work of Candès &

Tao (2006) and Donoho (2006) (where the significance of Gel'fand n-widths is discussed) and the interesting proof offered by Baraniuk *et al.* (2008).

The linear n-width can be used to investigate the optimality of fast multipole methods, as introduced by Greengard & Rokhlin (1987), though I have not found any discussion of this in the literature. Fast multipole methods use analytical techniques to achieve approximate rank reduction of a matrix whose (i, j) entry is $K(x_i, y_j)$ for some kernel K and some arguments $x_1 \leq x_2 \leq \cdots \leq x_N$ and $y_1 \leq y_2 \leq \cdots \leq y_M$. Beatson & Greengard (1997) provide a concise review of the fast multipole method and some of its applications, while Rokhlin (1995) gives an even more focused discussion. The fast multipole method has been deemed one of the ten most significant computational methods of the twentieth century (Cipra (2000)).

The goal of fast multipole methods is to permit rapid approximate matrix-vector multiplication through approximate matrix rank reduction; the motivation for the pioneering work of Greengard & Rokhlin (1987) was the computational challenge of particle simulation with a large number of particles. The computational complexity of matrix-vector multiplication is $O(MN)$ in general, but fast multipole methods allow the reduction of this computational complexity to $O(M + N)$ at the cost of introducing some approximation error. The great success of fast multipole methods has led to many generalizations and extensions, of which the fast Laplace transforms of Rokhlin (1988) (using Chebyshev polynomials) and Strain (1992) (using Laguerre functions) are closest to my investigations here; the kernel in these works is effectively the same as the exponential product kernel, but the domain and range of the operator they study differ from the domain and range of the exponential product operator. As is usual in work on fast multipole methods, neither Rokhlin (1988) nor Strain (1992) investigates questions of optimality for their respective methods, though both do show how their methods improve upon direct, unstructured matrix multiplication.

Studies in the literature on fast multipole methods typically introduce an approximate rank-reduction technique for matrices arising from a given kernel, prove that it achieves a certain computational complexity for matrix-vector multiplication as a function of allowed approximation error, and demonstrate the usefulness of the technique in practice. Using linear n-widths, one could also investigate optimality of the approximate rank-reduction technique used (given norms on the domain and range of the linear transformation to be approximated). From an n-width perspective, research on fast multipole methods is focused on constructing approximations that produce upper bounds on the linear n-widths of integral operators having specific kernels; however, these approximations may or may not have optimality properties.

My work here could be regarded as introducing fast multipole methods for exponential product operators that have verifiable optimality properties. From this point of view, it is useful to note that typical fast multipole approximations allow handling of kernel singularities (which certainly arise in the motivating problem of Greengard & Rokhlin (1987)). The additional computational effort involved in handling possible kernel singularities is not necessary in my problem, where the kernel $\exp(cxy)$ for $x, y \in [-1, 1]$ of an exponential product operator is analytic. The smoothness of the exponential product kernel leads to an absence, in my numerically-optimal and near-optimal approximations, of some of the computations often present in fast multipole methods.

5 Chebyshev Systems and Total Positivity

5.1 Chebyshev Systems

Chebyshev systems of functions are vital to the study of approximation in the supremum norm (that is, approximation with the goal of minimizing the worst possible error). This has led to their extensive exploration over the course of many decades. As Karlin & Studden (1966) note (on page 2 in Section 1 of Chapter I), significant progress in the theory of Chebyshev systems was made by Chebyshev (of course), Bernstein, Descartes, Haar, Laguerre, and de la Vallée Poussin. Gantmacher & Krein (2002), Karlin & Studden (1966), Karlin (1968), Kreĭn & Nudel'man (1977), and Zielke (1979) offer thorough discussions of Chebyshev systems, though most texts on approximation theory make some mention of them; see Achieser (1956), Cheney (1982), Lorentz (1986), Meinardus (1967), Powell (1981), Rivlin (1969), and Singer (1970). For a historical perspective on approximation theory, which includes a history of work on Chebyshev systems, see Steffens (2006) (particularly page 146).

Because of their importance in studying n-widths, Pinkus (1985a) also discusses Chebyshev systems in Section 1 of his Chapter III. I largely follow the notation and treatment of Pinkus (1985a), though I use the more modern romanization "Chebyshev" rather than the formerly traditional "Tchebycheff" (note that Pinkus (2010) also employs the "Chebyshev" romanization on page 88 in Section 1 of Chapter 4). I also avoid the use of abbreviations such as "T-system" for "Chebyshev system," since T will have other meanings at different points in this paper, and I use "Markov system" in place of "complete Chebyshev system."

It is important to note that the concept of a Chebyshev system is distinct from the Chebyshev polynomials (though the latter are very useful in approximation theory and will be employed in Section 9). For any $n > 0$, the first n Chebyshev polynomials form a Chebyshev system, but there are many Chebyshev systems which are not sets of Chebyshev polynomials; for example, the first n monomials $1, x, x^2, \ldots, x^{n-1}$ form the most familiar Chebyshev system. More importantly for my work here, the set $\{\exp c\rho_j x\}_{j=1}^{n}$ is a Chebyshev system on any interval (for any $c \neq 0$ and for any $\rho_1 < \cdots < \rho_n$).

It is useful to have a concise notation for certain determinants when working with Chebyshev systems.

Definition 5.1. For any set of functions $u_1(x), u_2(x), \ldots, u_n(x)$ all of which are defined at each of $x_1, x_2, \ldots, x_n$, I use the notation

$$U\begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} = \begin{vmatrix} u_1(x_1) & u_1(x_2) & \cdots & u_1(x_n) \\ u_2(x_1) & u_2(x_2) & \cdots & u_2(x_n) \\ \cdots & \cdots & \cdots & \cdots \\ u_n(x_1) & u_n(x_2) & \cdots & u_n(x_n) \end{vmatrix}.$$

This notation is used by Pinkus (1985a) (page 39 in Section 1 of Chapter III, as part of his Definition 1.1). It is distinct from the notation of Karlin & Studden (1966) (page 1 in Section 1 of Chapter I) only in that indices begin with one rather than beginning with zero.

Definition 5.2. Let I be an interval (open, closed, or half-open). $u_1, \ldots, u_n$ is a weak Chebyshev system on I if $U\begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ is either nonnegative for all $x_1 < \cdots < x_n \in I$ or nonpositive for all $x_1 < \cdots < x_n \in I$.

Definition 5.2 is essentially Definition 1.1 on page 39 and 40 in Section 1 of Chapter III in Pinkus (1985a), though it appears elsewhere in the literature with slight variations, *e. g.*, Definition 2.1 on pages 3 and 4 in Section 2 of Chapter I in Karlin & Studden (1966) (where nonnegativity is specified and the interval I is assumed to be closed).

Definition 5.3. Let I be an interval (open, closed, or half-open). $u_1, \ldots, u_n$ is a Chebyshev system on I if $U\begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ is either strictly positive for all $x_1 < \cdots < x_n \in I$ or strictly negative for all $x_1 < \cdots < x_n \in I$.

Definition 5.3 matches the initial portion of Definition 1.3 on page 40 in Section 1 of Chapter III in Pinkus (1985a), but the content of this definition is standard in the literature. Lorentz (1986) notes that his property 2 on page 24 in Section 4 of Chapter 2, which matches the definition above, is equivalent to his definition of a Chebyshev system. Gantmacher & Krein (2002) show their definition is equivalent to the one above; see their Lemma 2 on page 137 in Section 5 of Chapter III. It is common to find the romanization "Tchebycheff" used in place of "Chebyshev," particularly in earlier works; see the definition given on page 74 in Section 4 of Chapter 3 in Cheney (1982), the initial part of Definition 1.1 on page 1 in Section 1 of Chapter I in Karlin & Studden (1966) (where strict positivity is specified and the interval I is assumed to be closed), and the definition given on page 24 in Section 4 of Chapter I in Karlin (1968) (where I is assumed to be open). In Singer (1970), on page 182 in Section 1 of Chapter II, Chebyshev systems are described but the romanization "Čebyšev" is used (Kreĭn & Nudel'man (1977) also use this romanization; they introduce Chebyshev systems on pages 31 and 32 in Section 1 of Chapter II). Sometimes the notion of a nested sequence of subspaces is included in the definition of a Chebyshev system; see Section 1 of Chapter 4 in Part I of Meinardus (1967). The condition satisified by a Chebyshev system is often labeled the "Haar condition" or "Haar's condition" (see pages 67 through 73 in Chapter II of Achieser (1956), page 74 in Section 4 of Chapter 3 in Cheney (1982), and pages 76 through 77 in Section 3 of Chapter 7, along with all of Appendix A, in Powell (1981)).

Definition 5.4. Let I be an interval (open, closed, or half-open). $u_1, \ldots, u_n$ is a Markov system on I if $u_1, \ldots, u_k$ is a Chebyshev system on I for each $k = 1, \ldots, n$.

Definition 5.4 uses the term "Markov" rather than "complete Tchebycheff," but it describes the same property as the latter portion of Definition 1.3 on page 40 in Section 1 of Chapter III in Pinkus (1985a) (and is similar to the definition of "complete Tchebycheff" that forms part of Definition 1.1 on page 1 in Section 1 of Chapter I in Karlin & Studden (1966), where strict positivity is specified and the interval I is taken to be closed). My terminology matches that of Gantmacher & Krein (2002) (Definition 2 on page 181 in Section 3 of Chapter IV, where the interval is assumed to be open) and Kreĭn & Nudel'man (1977) (page 43 in Section 4 of Chapter II). As Karlin (1968) notes (on page 274 in Section 1 of Chapter 6), the term "Haar system" is sometimes used for what I call a Markov system.

Definition 5.5. Let I be an interval (open, closed, or half-open). $u_1, \ldots, u_n$ is a Descartes system on I if $U\begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} > 0$ for each $1 \leq i_1 < \cdots < i_k \leq n$, $k = 1, \ldots, n$ and $x_1 < \cdots < x_n \in I$.

Definition 5.5 is essentially Definition 1.9 on page 44 in Section 1 of Chapter III of Pinkus (1985a) (who notes that his specification of the sign of all the relevant determinants is atypical). A Descartes system under my definition is one under the (slightly more general) definitions of Karlin & Studden (1966) (Definition 4.5 on page 25 in Section 4 of Chapter I) and Kreĭn & Nudel'man (1977) (page 39 in Section 2 of Chapter II), with the caveat that they assume that the interval I is closed.

An important example of a Descartes system, and one very relevant to this work, is provided by the following theorem.

Theorem 5.1. *For any $\rho_1 < \cdot < \rho_n \in \mathbb{R}$, the functions $\exp(c\rho_1 x), \ldots, \exp(c\rho_n x)$ form a Descartes system on any interval $I \subset \mathbb{R}$.*

Proof. This result follows directly from display (3.1), and the argument below it, in Karlin & Studden (1966) (on page 9 in Section 3 of Chapter I). As Karlin & Studden (1966) note, this is a special case of a classical result. □

In considering Chebyshev systems of smooth functions, it is helpful to have a notation for determinants which extend by continuity a suitable scaling of the determinants in Definition 5.1. These determinants involve derivatives of the functions in the system whenever one or more pairs of x values coincide. Given an interval I, the set $C^{n-1}(I)$ means, as usual, the set of functions that are $n-1$ times continuously differentiable on I.

Definition 5.6. Let I be an interval (open, closed, or half-open). For any set of functions $u_1(x), u_2(x), \ldots, u_n(x) \in C^{n-1}(I)$ and for any $x_1 \leq x_2 \leq \cdots \leq x_n \in I$,

$$U^* \begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

means $U \begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$, but with the columns of the determinant corresponding to any equal x values replaced by successive derivatives of the u_i. Thus, if $x_1 = x_2 = x_3 = x$, column 1 would be $u_i(x)$, column 2 would be $\frac{d}{dx} u_i(x)$, and column 3 would be $\frac{d^2}{dx^2} u_i(x)$. Thus, if $x_1 < x_2 < \cdots < x_n$, $U^* \begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ is identical to $U \begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$. At the other extreme, if $x_1 = x_2 = \cdots = x_n = x$, then $U^* \begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ is the determinant of the matrix whose (i,j) element is $\frac{d^{j-1}}{dx^{j-1}} u_i(x)$.

This definition is given by Pinkus (1985a) (page 43 in Section 1 of Chapter III) and is also used by Karlin & Studden (1966) (see page 5 in Section 2 of Chapter I) and Karlin (1968) (see page 274 in Section 1 of Chapter 6, which references page 48 in Section 1 of Chapter 2).

Definition 5.7. Let I be an interval (open, closed, or half-open). $u_1, \ldots, u_n \in C^{n-1}(I)$ is an extended Chebyshev system on I if $U^* \begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ is either strictly positive for all $x_1 \leq x_2 \leq \cdots \leq x_n \in I$ or strictly negative for all $x_1 \leq x_2 \leq \cdots \leq x_n \in I$. If $u_1, \ldots, u_k$ is an extended Chebyshev system for $k = 1, \ldots, n$, then $u_1, \ldots, u_n$ is an extended Markov system.

Definition 5.7 matches Definition 1.8 on page 43 in Section 1 of Chapter III in Pinkus (1985a), though Pinkus (1985a) uses "extended complete Tchebycheff system" rather than extended Markov system. With the same exception, the definition above also agrees with the terminology of Karlin (1968) (pages 274 and 275 in Section 1 of Chapter 6). A version of this definition that assumes I is closed and demands strict positivity appears in Karlin & Studden (1966) (Definition 2.4 on page 6 in Section 2 of Chapter I). A concept of "degree" (Karlin (1968)) or "order" (Karlin & Studden (1966)) of an extended Chebyshev system appears in the latter two sources, but I do not require this additional nuance.

The function system of Theorem 5.1 is not only Descartes but also extended Markov, as the following theorem makes clear.

Theorem 5.2. *For any $\rho_1 < \cdot < \rho_n \in \mathbb{R}$, the functions $\exp(c\rho_1 x), \ldots, \exp(c\rho_n x)$ form an extended Markov system on any interval $I \subset \mathbb{R}$.*

Proof. This result is a direct consequence of Corollary 3.2 on page 53 in Section 3 of Chapter III in Pinkus (1985a). $\square$

One of the crucial properites of Chebyshev systems (and related systems of functions) concerns zero-counting. Different varieties of Chebyshev systems (weak Chebyshev systems, Chebyshev systems, and extended Chebyshev systems) have different zero-counting properties. To capture these differences, I need the following definitions of particular types of zeros.

Definition 5.8. Let $f(x_0) = 0$, where f is a real-valued function defined on an interval I (open, closed, or half-open). Then x_0 is a *nodal zero* of the function f if either x_0 is an endpoint of the interval I (if I has endpoints) or, if x_0 is in the interior of I, then for every $\epsilon > 0$ small enough, $f(x_0 - \epsilon) f(x_0 + \epsilon) < 0$. If x_0 is not a nodal zero, it is termed a *nonnodal zero*.

This echoes the Definition 1.5 of Pinkus (1985a) (on page 42 in Section 1 of Chapter III), and also matches the first portion of Definition 4.3 on page 22 in Section 4 of Chapter I in Karlin & Studden (1966). Gantmacher & Krein (2002) (page 137 in Section 5 of Chapter III) give a similar definition, though they use "node" rather than "nodal zero" and "antinode" in place of "nonnodal zero." Informally, a nodal zero of a function is an isolated zero of that function and, moreover, an isolated zero *crossing* of that function.

Definition 5.9. Let I be an interval (open, closed, or half-open), and let f be a continuous function defined on I. Then $S(f)$ is the number of times that f changes sign on I.

The above definition echoes Definition 1.2 of Pinkus (1985a) (on page 40 in Section 1 of Chapter II).

It is sometimes helpful to consider the sign changes of finite lists of real numbers.

Definition 5.10. For $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ with $\sum_{i=1}^{n} x_i^2 > 0$, $S^-(\boldsymbol{x})$ is the number of sign changes in the sequence $x_1, x_2, \ldots, x_n$ after any zeros in the sequence have been deleted. $S^+(\boldsymbol{x})$ is the largest number of sign changes in the sequence $x_1, x_2, \ldots, x_n$ that can be obtained by replacing each zero in the sequence with +1 or -1.

This is Definition 2.2 of Pinkus (1985a) (on page 45 in Section 2 of Chapter III).

Definition 5.11. Let I be an interval (open, closed, or half-open), and let f be a continuous function defined on I. Then $Z(f)$ is the number of distinct zeros of f on I.

Definition 5.11 is Definition 1.4 of Pinkus (1985a) (page 41 in Section 1 of Chapter II) and also appears in Karlin & Studden (1966) with the added assumption that I is a closed interval (see their Definition 4.2 on page 21 in Section 4 of Chapter I).

Definition 5.12. Let I be an interval (open, closed, or half-open), and let f be a continuous function defined on I. Then $\widetilde{Z}(f)$ is the number of zeros of f on I where nodal zeros are counted once and nonnodal zeros are counted twice.

Definition 5.12 matches Definition 1.6 of Pinkus (1985a) (page 42 in Section 1 of Chapter II) and also echoes the latter portion of Definition 4.3 on page 23 in Section 4 of Chapter I in Karlin & Studden (1966) (though Karlin & Studden (1966) assume the interval I to be closed).

Definition 5.13. Let I be an interval (open, closed, or half-open), and let $f \in C^{n-1}(I)$. Then $Z^*(f)$ is the number of zeros of f on I counting multiplicities (in the usual way, where the multiplicity of any zero $x_0 \in I$ can be at most n).

Definition 5.13 is that of Pinkus (1985a) (Definition 1.7 on page 42 in Section 1 of Chapter III) and is similar to that used by Karlin & Studden (1966) (Definition 4.4 on page 24 in Section 4 of Chapter I), though Karlin & Studden (1966) assume a closed interval and simply specify that f must be "sufficiently differentiable."

It is now possible to express the relationships between the definitions of function systems above and zero-counting properties.

Proposition 5.1. *$u_1, \ldots, u_n$ is a weak Chebyshev system on the interval I (open, closed, or half-open) if and only if $S(u) \leq n-1$ for every function u of the form $u = \sum_{i=1}^n a_i u_i$ with constants a_i such that $\sum_{i=1}^n a_i^2 > 0$.*

Proof. This is Proposition 1.1 on page 40 in Section 1 of Chapter III in Pinkus (1985a), though no formal proof is given there. □

Proposition 5.2. *$u_1, \ldots, u_n$ is a Chebyshev system on the interval I (open, closed, or half-open) if and only if $\widetilde{Z}(u) \leq n-1$ for every function u of the form $u = \sum_{i=1}^n a_i u_i$ with constants a_i such that $\sum_{i=1}^n a_i^2 > 0$.*

Proof. The proposition is Proposition 1.5 on page 42 in Section 1 of Chapter III in Pinkus (1985a), and a proof is provided there. This result also appears as Theorem 4.2 on page 23 in Section 4 of Chapter I in Karlin & Studden (1966). Note that Karlin & Studden (1966) state the result for a system of $n+1$ functions, obtaining a bound on $\widetilde{Z}(u)$ of n, and they must qualify the equivalence because their definition of a Chebyshev system involves strict positivity of $U\begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$, rather than strict positivity or strict negativity as in Definition 5.3. □

Proposition 5.3. *If $u_1, \ldots, u_n$ is a Descartes system on the* closed *interval I, then $\widetilde{Z}(u) \leq S^-(\boldsymbol{a})$ for every function u of the form $u = \sum_{i=1}^n a_i u_i$ with a vector $\boldsymbol{a}$ of constants such that $\sum_{i=1}^n a_i^2 > 0$.*

Proof. This is the first portion of Corollary 2.5 in Pinkus (1985a) (on page 49 in Section 2 of Chapter III). Karlin & Studden (1966) prove a similar result as their Theorem 4.4 (on page 25 in Section 4 of Chapter I), attributing it to the 1950 edition of Gantmacher & Krein (2002), though they prove equivalence with Z in place of $\widetilde{Z}$. □

Proposition 5.3 is a generalization of the classical Descartes rule of signs for ordinary polynomials.

Proposition 5.4. *$u_1, \ldots, u_n$ is an extended Chebyshev system on the interval I (open, closed, or half-open) if and only if $Z^* (u) \leq n-1$ for every function u of the form $u = \sum_{i=1}^n a_i u_i$ with constants a_i such that $\sum_{i=1}^n a_i^2 > 0$.*

Proof. This is Proposition 1.6 on page 43 in Section 1 of Chapter III in Pinkus (1985a), who notes that the "proof is totally analogous to the proof of Proposition 1.3." Proposition 1.3 of Pinkus (1985a) is proven on page 41. Karlin & Studden (1966) provide a proof of this result, which is their Theorem 4.3, on pages 24 and 25 in Section 4 of Chapter I. Note that Karlin & Studden (1966) state the result for a system of $n+1$ functions, obtaining a bound on $Z^* (u)$ of n, and they must qualify the equivalence because their definition of an extended Chebyshev system involves strict positivity of $U^* \begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$, rather than strict positivity or strict negativity as in Definition 5.7. □

5.2 Approximation in the Supremum Norm Using Chebyshev Systems: Equioscillation

Chebyshev systems are of great importance in classical approximation theory: they allow clear conditions for the unicity of best approximation and they provide a vital characterization of the best approximant.

Theorem 5.3 (Haar's Theorem)**.** *The linear combination $u^* = \sum_{i=1}^n a_i^* u_i$ that solves the problem*

$$\min_{a_1,\ldots,a_n} \max_{x \in [b,d]} \left| f(x) - \sum_{i=1}^n a_i u_i (x) \right| \tag{44}$$

is uniquely determined for every function $f \in C([b,d])$ if and only if $u_1, \ldots, u_n$ is a Chebyshev system on $[b,d]$.

Proof. This is result, initially obtained by Haar (1917), is part (i) of Theorem 1.1 on page 280 in Section 1 of Chapter IX in Karlin & Studden (1966) (the proof is given in Section 3 of Chapter IX). Note that Karlin & Studden (1966) introduce a possible sign change for one of the functions in the system (and I do not) because their definition of a Chebyshev system demands strict positivity, rather than either strict positivity or strict negativity as in Definition 5.3, of the determinants $U \begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ for all $x_1 < \cdots < x_n$ in the interval $[b,d]$. □

Theorem 5.4 (The Equioscillation Theorem)**.** *If $u_1, \ldots, u_n$ is a Chebyshev system on $[b,d]$, then for any $f \in C([b,d])$ the unique linear combination $u^* = \sum_{i=1}^n a_i^* u_i$ that solves the problem (44) has the property that there exist $n+1$ points $\{x_i\}_{i=1}^{n+1}$ with $b \leq x_1 < x_2 < \cdots < x_{n+1} \leq d$ such that*

$$(-1)^i \epsilon \left(f(x_i) - u^*(x_i) \right) = \max_{x \in [b,d]} \left| f(x) - u^*(x) \right|, \tag{45}$$

where ϵ is either +1 or -1. Furthermore, u^* is the only linear combination of the $u_1, \ldots, u_n$ that has this property.

Proof. This result has a rich history, reviewed in Steffens (2006) and briefly treated by Trefethen (2013) (page 74 in Chapter 10), that stretches back to Chebyshev himself. The version I give here is part (ii) of Theorem 1.1 on page 280 in Section 1 of Chapter IX in Karlin & Studden (1966) (the proof is given in Section 3 of Chapter IX) except for the last sentence of my statement, which is demonstrated on page 280 in Section 1 of Chapter IX in Karlin & Studden (1966) immediately below their Theorem 1.1. Note that Karlin & Studden (1966) begin their indexing with zero while I begin my indexing with one, so they have $n+1$ functions in their Chebyshev system and $n+2$ points of alternation, while I have n functions in my Chebyshev system and $n+1$ points of alternation. □

As Trefethen (2013) observes (page 119 in Chapter 16), "Everybody remembers . . ., the equioscillation theorem, from the moment they first see it." This theorem is also sometimes termed the "alternation" or "equiripple" theorem.

Theorem 5.5 (The Snake Theorem)**.** *Suppose that $u_1, \ldots, u_n$ is a Chebyshev system on $[b, d]$ and that f and g are two continuous functions on $[b, d]$ such that there exists a linear combination $v(x) = \sum_{i=1}^{n} a_i^v u_i(x)$ with $f(x) > v(x) > g(x)$ for all $x \in [b, d]$. Then there exists a unique linear combination $\underline{u} = \sum_{i=1}^{n} \underline{a}_i u_i$ with the properties*

1. *$f(x) \geq \underline{u}(x) \geq g(x)$ for all $x \in [b, d]$ and*
2. *there exist n points $b \leq \underline{s}_1 < \underline{s}_2 < \cdots < \underline{s}_n \leq d$ such that*

$$\underline{u}\left(\underline{s}_{n-i}\right) = \begin{cases} f\left(\underline{s}_{n-i}\right) & i = 0, 2, 4, \ldots \\ g\left(\underline{s}_{n-i}\right) & i = 1, 3, 5 \ldots \end{cases} . \tag{46}$$

There also exists a unique linear combination $\overline{u} = \sum_{i=1}^{n} \overline{a}_i u_i$ with the properties

1. *$f(x) \geq \overline{u}(x) \geq g(x)$ for all $x \in [b, d]$ and*
2. *there exist n points $b \leq \overline{s}_1 < \overline{s}_2 < \cdots < \overline{s}_n \leq d$ such that*

$$\overline{u}\left(\overline{s}_{n-i}\right) = \begin{cases} f\left(\overline{s}_{n-i}\right) & i = 0, 2, 4, \ldots \\ g\left(\overline{s}_{n-i}\right) & i = 1, 3, 5 \ldots \end{cases} . \tag{47}$$

Proof. This is Theorem 10.2 on page 72 (proven on page 73) in Section 10 of Chapter II in Karlin & Studden (1966). The result was originally obtained by Karlin (1963), and the name "the snake theorem" is due to Kreĭn & Nudel'man (1977) (see below for more on the origins of the name). Note that I have been somewhat more verbose than Karlin & Studden (1966) in the statement of the theorem, and that Karlin & Studden (1966) state their theorem for a T-system, which they define similarly to my definition of a Chebyshev system in Definition 5.3, but with the additional restriction that the determinant

$U\begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ be strictly positive on $[b,d]$ (as noted under Definition 5.3). In Definition 5.3, I require only that the determinant $U\begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ be of one strict sign (either strictly positive or strictly negative) on $[b,d]$. However, the theorem continues to hold for Chebyshev systems as defined in Definition 5.3: if the determinant $U\begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ is strictly positive on $[b,d]$, then the result of Karlin & Studden (1966) applies directly. If the determinant $U\begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ is strictly negative on $[b,d]$, simply choose one of the u_i, $i=1,\ldots,n$, and multiply it by -1. The resulting new system has a determinant $U_{\text{new}}\begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$ that is strictly positive on $[b,d]$ (by the multilinearity of the determinant), and any linear combination of the functions from the new system is also a linear combination of the original functions, so the Karlin & Studden (1966) result applied to the sign-changed new system yields the desired result for the original system. □

Remark 5.1. The name "the snake theorem" comes from Kreĭn & Nudel'man (1977) (pages 368 and 369 in Section 6 of Chapter IX), who give a memorable motivation for this choice of name for the result. I paraphrase their interpretation: they think of any linear combination of the $u_1,\ldots,u_n$ as a snake. The functions f and g form the two walls of a tunnel, through which a snake (v) may crawl without touching the walls. Given the assumptions of the theorem, there are two "most sinuous snakes" (this phrase is used, in quotes, by Kreĭn & Nudel'man (1977); in my notation, these two most sinuous snakes are $\underline{u}$ and $\overline{u}$) each of which touches the walls, alternating from one wall to the other, at least n times.

Theorem 5.6. *If $u_1,\ldots,u_{n+1}$ is a Descartes system on $[b,d]$, the linear combination $u^{LD}=\sum_{i=1}^{n+1} a_i^{LD}u_i$ that solves the problem*

$$\min_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1} \max_{x\in[b,d]} \left|\sum_{i=1}^{n+1} a_i u_i(x)\right| \tag{48}$$

is unique up to multiplication by -1 (that is, it is unique up to sign). u^{LD} has the property that there exist $n+1$ points $\left\{x_i^{LD}\right\}_{i=1}^{n+1}$ with $b\le x_1^{LD}<x_2^{LD}<\cdots<x_{n+1}^{LD}\le d$ such that

$$(-1)^i \epsilon u^{LD}\left(x_i^{LD}\right) \quad = \quad \max_{x\in[b,d]}\left|u^{LD}(x)\right|, \tag{49}$$

where ϵ is either +1 or -1. Furthermore, u^{LD} and $-u^{LD}$ are the only linear combinations of the $u_1,\ldots,u_{n+1}$ that have this property. Finally, $\boldsymbol{a}^{LD}$ has no zero elements and its elements alternate in sign.

Proof. Apply Theorem 5.5 to the system $u_1,\ldots,u_{n+1}$ (this system is a Chebyshev system because it is a Descartes system, see Definition 5.5) with $f(x)\equiv 1$

and $g(x) \equiv -1$. This produces the functions $\underline{u}$ and $\overline{u}$ of Theorem 5.5, which are unique in their alternation properties as detailed in that theorem. Because of the uniqueness of $\underline{u}$ and $\overline{u}$, and because I have chosen f and g so that $f = -g$, $\underline{u} = -\overline{u}$; indeed, $-\overline{u}$ satisfies both properties of $\underline{u}$ provided in Theorem 5.5, and, since $\underline{u}$ is the unique linear combination of $u_1, \ldots, u_{n+1}$ that satisfies these properties, it follows that $\underline{u} = -\overline{u}$ (this is a more verbose version of an argument given by Karlin & Studden (1966), with a typographical error, on page 282 in Section 2 of Chapter IX). Of course, this means that $\underline{\boldsymbol{a}} = -\overline{\boldsymbol{a}}$. By Theorem 5.5, $\underline{u}$ alternates between -1 and 1 at $n+1$ points; thus, $\underline{u}$ has at least n zeros. Since $u_1, \ldots, u_{n+1}$ is a Descartes, and thus Chebyshev, system, $\underline{u}$ has no more than n zeros, so it has exactly n zeros. By Proposition 5.3, this implies that $S^-(\underline{\boldsymbol{a}}) \geq n$; since the number of sign changes in a vector of length $n+1$ cannot exceed n, $S^-(\underline{\boldsymbol{a}}) = n$. The definition of $S^-(\underline{\boldsymbol{a}})$ is the number of sign changes after any zeros are discarded (see Definition 5.10), so $\underline{\boldsymbol{a}}$ cannot contain any zeros (if it did, then $S^-(\underline{\boldsymbol{a}})$ would discard them, in which case the remaining vector could not have n sign changes). That is, $\underline{a}_i \neq 0$ for $i = 1, \ldots, n+1$.

For any $j = 1, \ldots, n+1$, consider $u^{(j)} \equiv \sum_{i=1}^{n+1} a_i^{(j)} u_i$, where $a_i^{(j)} \equiv \underline{a}_i / \underline{a}_j$. This construction is well-defined because $\underline{a}_j \neq 0$ by the preceding paragraph. Then $u^{(j)} = u_j + \sum_{\substack{i=1 \\ i \neq j}}^{n+1} a_i^{(j)} u_i$. Setting $\widehat{a}_i^{(j)} \equiv -a_i^{(j)}$ for $i \neq j$, $i = 1, \ldots, n+1$, gives the expression $u^{(j)} = u_j - \sum_{\substack{i=1 \\ i \neq j}}^{n+1} \widehat{a}_i^{(j)} u_i$. Now, since $\underline{u}$ alternates between -1 and 1 at $n+1$ points, $u^{(j)}$ alternates between $-1/\underline{a}_j$ and $1/\underline{a}_j$ at $n+1$ points. As the representation $u^{(j)} = u_j - \sum_{\substack{i=1 \\ i \neq j}}^{n+1} \widehat{a}_i^{(j)} u_i$ shows, it is also the error in an approximation of u_j by u_i, $i \neq j$, $i = 1, \ldots, n+1$. Because $u_1, \ldots, u_{n+1}$ is a Descartes system, u_i, $i \neq j$, $i = 1, \ldots, n+1$ is a Chebyshev system. Thus, by Theorem 5.4, the only linear combination of the u_i, $i \neq j$, $i = 1, \ldots, n+1$ whose difference from u_j has this property is the best approximation, so $u^{(j)}$ is the error in the best approximation of u_j by u_i, $i \neq j$, $i = 1, \ldots, n+1$:

$$\max_{x \in [b,d]} \left| u^{(j)}(x) \right| = \min_{\boldsymbol{a}} \max_{x \in [b,d]} \left| u_j(x) - \sum_{\substack{i=1 \\ i \neq j}}^{n+1} a_i u_i(x) \right| . \tag{50}$$

I now assert that

$$u^{\dagger} = \left(\frac{1}{\sum_{i=1}^{n+1} |\underline{a}_i|} \right) \underline{u} \tag{51}$$

is, in fact, the u^{LD} of the statement of this theorem. By construction, $\sum_{i=1}^{n+1} \left| a_i^{\dagger} \right| = \sum_{i=1}^{n+1} |\underline{a}_i| / \sum_{i=1}^{n+1} |\underline{a}_i| = 1$. The equioscillation property (49) follows from the alternation of $\underline{u}$ between -1 and 1 at $n+1$ points. Because $-\underline{u} = \overline{u}$, it is immediate that $-u^{\dagger}$ has the same equioscillation property, but with ϵ in (49) having the opposite sign from that in the equioscillation for $u^{\dagger}$. Further, the uniqueness results of Theorem 5.5 affirm that no other linear combinations of the $u_1, \ldots, u_{n+1}$ can equioscillate between these values. The property of $\boldsymbol{a}^{LD}$ claimed in the last sentence of the statement of the theorem follows from the fact that $S^-(\underline{\boldsymbol{a}}) = n$ as shown in the first paragraph of this proof (since $\boldsymbol{a}^\dagger$ is just a nonzero scalar multiple of $\underline{\boldsymbol{a}}$).

It remains to show that $u^\dagger$ and $-u^\dagger$ solve the problem (48). Suppose not; I will demonstrate a contradiction, but first I must show that the minimum over $\boldsymbol{a}$ in (48) is attained.

By Berge (1963) (Theorem 1 on page 115 in Section 3 of Chapter VI, taking the mapping Γ in that theorem to be the map assigning to each value of $\boldsymbol{a}$ the clearly non-empty set $[b,d]$), $\max_{x\in[b,d]} \left|\sum_{i=1}^{n+1} a_i u_i(x)\right|$ is a lower semi-continuous function of $\boldsymbol{a}$. By Berge (1963) (Theorem 2 on page 116 in Section 3 of Chapter VI, taking the mapping Γ in that theorem to be the map assigning to each value of $\boldsymbol{a}$ the clearly non-empty set $[b,d]$), it is also an upper semi-continuous function of $\boldsymbol{a}$. Thus, $\max_{x\in[b,d]} \left|\sum_{i=1}^{n+1} a_i u_i(x)\right|$ is a continuous function of $\boldsymbol{a}$. (This is often termed "the maximum theorem," and I use entirely similar logic in Theorem 6.1 below.) Because $\left\{\boldsymbol{a} : \sum_{i=1}^{n+1} |a_i| = 1\right\}$ is compact (as a closed and bounded subset of $\mathbb{R}^{n+1}$), the minimum over $\boldsymbol{a}$ in (48) is attained.

Because the minimum in (48) is attained, there is some $\tilde{u} = \sum_{i=1}^{n+1} \tilde{a}_i u_i$ that attains it. Further, $\sum_{i=1}^{n+1} |\tilde{a}_i| = 1$. Either there is some $j \in \{1, \ldots, n+1\}$ such that $|\tilde{a}_j| > \left|a_j^\dagger\right|$ or $|\tilde{a}_j| = \left|a_j^\dagger\right|$ for every $j = 1, \ldots, n+1$ (if neither of these statements holds, $1 = \sum_{i=1}^{n+1} |\tilde{a}_i| < \sum_{i=1}^{n+1} \left|a_i^\dagger\right| = 1$, a contradiction). I will consider these two cases separately.

Case 1: $|\tilde{a}_j| > \left|a_j^\dagger\right|$ for some $j = 1, \ldots, n+1$. Clearly $\tilde{a}_j \neq 0$, so $\tilde{a}_i^{(j)} = -\tilde{a}_i/\tilde{a}_j$, $i \neq j$, is well-defined.

$$\begin{aligned}
& \max_{x\in[b,d]} \left| u_j(x) - \sum_{\substack{i=1\\ i\neq j}}^{n+1} \tilde{a}_i^{(j)} u_i(x) \right| \\
= \quad & \frac{1}{|\tilde{a}_j|} \max_{x\in[b,d]} \left| \sum_{i=1}^{n+1} \tilde{a}_i u_i(x) \right| && (52)\\
\leq \quad & \frac{1}{|\tilde{a}_j|} \max_{x\in[b,d]} \left| \sum_{i=1}^{n+1} a_i^\dagger u_i(x) \right| && (53)\\
< \quad & \frac{1}{\left|a_j^\dagger\right|} \max_{x\in[b,d]} \left| \sum_{i=1}^{n+1} a_i^\dagger u_i(x) \right| && (54)
\end{aligned}$$

$$= \max_{x\in[b,d]} \left| u_j(x) - \sum_{\substack{i=1 \\ i\neq j}}^{n+1} \widehat{a}_i^{(j)} u_i(x) \right| \tag{55}$$

$$= \max_{x\in[b,d]} \left| u^{(j)}(x) \right|, \tag{56}$$

where the first equality is by the definition of $\tilde{a}_i^{(j)}$, the first inequality is by the assumption that $\max_{x\in[b,d]} \left|\sum_{i=1}^{n+1} \tilde{a}_i u_i(x)\right| \leq \max_{x\in[b,d]} \left|\sum_{i=1}^{n+1} a_i^{\dagger} u_i(x)\right|$, the second inequality is due to the assumption (for the current Case 1) that $|\tilde{a}_j| > \left|a_j^{\dagger}\right|$ (and by the fact that $a_j^{\dagger} \neq 0$ as shown above), the second equality is by the definition of $\widehat{a}_i^{(j)}$ above (noting that $a_i^{\dagger}/a_j^{\dagger} = \underline{a}_i/\underline{a}_j$ because the factor $\sum_{i=1}^{n+1} |\underline{a}_i|$ is present in both the numerator and denominator of $a_i^{\dagger}/a_j^{\dagger}$), and the final equality is by the definition of $u^{(j)}$ above. As shown above in (50), $u^{(j)}$ is the error in the best approximation of u_j by u_i, $i \neq j$, $i = 1, \ldots, n+1$, so the display above is a contradiction and Case 1 cannot hold.

Case 2: $|\tilde{a}_j| = \left|a_j^{\dagger}\right|$ for every $j = 1, \ldots, n+1$. For each $j = 1, \ldots, n+1$, note that $a_j^{\dagger} \neq 0$ (as shown above) and proceed exactly as in Case 1 but with equality, rather than strict inequality, holding on line (54) to obtain

$$\max_{x\in[b,d]} \left| u_j(x) - \sum_{\substack{i=1 \\ i\neq j}}^{n+1} \tilde{a}_i^{(j)} u_i(x) \right|$$

$$\leq \max_{x\in[b,d]} \left| u_j(x) - \sum_{\substack{i=1 \\ i\neq j}}^{n+1} \widehat{a}_i^{(j)} u_i(x) \right| \tag{57}$$

$$= \max_{x\in[b,d]} \left| u^{(j)}(x) \right|. \tag{58}$$

By Theorem 5.4, the best approximation of u_j by u_i, $i \neq j$, $i = 1, \ldots, n+1$ is unique; as in Case 1, this best approximation was shown via (50) to be $\sum_{\substack{i=1 \\ i\neq j}}^{n+1} \widehat{a}_i^{(j)} u_i(x)$. Thus, $\tilde{a}_i^{(j)} = \widehat{a}_i^{(j)}$, so $\tilde{a}_i/\tilde{a}_j = a_i^{\dagger}/a_j^{\dagger}$ for all $i \neq j$, $i = 1, \ldots, n+1$ (as in Case 1, $a_i^{\dagger}/a_j^{\dagger} = \underline{a}_i/\underline{a}_j$ because the factor $\sum_{i=1}^{n+1} |\underline{a}_i|$ is present in both the numerator and denominator of $a_i^{\dagger}/a_j^{\dagger}$). But $|\tilde{a}_j| = \left|a_j^{\dagger}\right|$ by the assumption of this case, so either $\tilde{a}_i = a_i^{\dagger}$ for all $i \neq j$, $i = 1, \ldots, n+1$ or $\tilde{a}_i = -a_i^{\dagger}$ for all $i \neq j$, $i = 1, \ldots, n+1$. The above argument can be made for *every* $j = 1, \ldots, n+1$, so either $\tilde{a}_i = a_i^{\dagger}$ for all $i = 1, \ldots, n+1$ or $\tilde{a}_i = -a_i^{\dagger}$ for all $i = 1, \ldots, n+1$. This is equivalent to the statement that either $\tilde{u} = u^{\dagger}$ or $\tilde{u} = -u^{\dagger}$, as desired. $\square$

Remark 5.2. Theorem 5.6 appears to be new, though Theorem 5.5 (which is critical to my proof of Theorem 5.6) has been used in handling different constrained approximation problems considered in the literature (see Pinkus (1976) and Keener (1993), for example). A different but related problem is handled using distinct logic as part of Proposition 5.8 on page 193 in Section 5 of Chapter V in Pinkus (1985a) (this proposition is itself a review of Proposition 3.1 of Micchelli & Pinkus (1979)). Theorem 5.6 is central to my method for computing a lower bound on certain key approximation errors, as described in Section 6 below.

Remark 5.3. It is easy to generalize Theorem 5.6 to the constraint $\|\boldsymbol{a}\| = 1$ for any norm $\|\cdot\|$ on $\mathbb{R}^{n+1}$ such that $|x_i| \leq |y_i|$, $i = 1, \ldots, n+1$, and $|x_j| < |y_j|$ for some j together imply that $\|\boldsymbol{x}\| < \|\boldsymbol{y}\|$, but I do not need the additional generality in what follows.

5.3 Total Positivity

The theory of total positivity requires some cumbersome notation, but provides "surprisingly powerful tools" (Pinkus (1985a), page 39). For readers interested in going beyond the results that I specifically reference here, the books of Gantmacher & Krein (2002), Karlin & Studden (1966), and Karlin (1968) provide thorough coverage of earlier work on total positivity. Chapter III of Pinkus (1985a) and Pinkus (1996) offer more concise treatments of portions of the field. Pinkus (2010) and Fallat & Johnson (2011) are focused on the discrete (matrix) case rather than the continuous (kernel) case. Finally, the importance of the early work of Kellogg (1918) deserves mention; it may have been the first work on total positivity, though the term (in German as *total positiv*) was introduced by Schoenberg (1930) (Pinkus (2010), in his Afterword, notes the crucial contributions of Schoenberg, Krein, Gantmacher, and Karlin to the theory of total positivity).

The reader is warned that there are two distinct sets of terminology used in the literature on total positivity (and there is confusing overlap between the terms used). In Gantmacher & Krein (2002), the term "totally positive" is reserved for matrices whose every minor is strictly positive, while "totally nonnegative" is employed for matrices whose every minor is nonnegative. In the work of Karlin & Studden (1966), Karlin (1968), Pinkus (1985a) (Chapter III), Pinkus (1996), and Pinkus (2010), as well as many journal articles written in English (such as Melkman & Micchelli (1978), Micchelli & Pinkus (1977b), Micchelli & Pinkus (1977c), Micchelli & Pinkus (1977a), and Micchelli & Pinkus (1978)) the term "strictly totally positive" is used to refer to a matrix whose every minor is strictly positive, while "totally positive" refers to a matrix whose every minor is nonnegative. The same confusing overlap of terms extends to the continuous case, where terminology is used to describe the kernels of integral operators. Pinkus (2010) gives a helpful history of these overlapping terminologies on page 33 in Section 5 of Chapter 1. I employ the terminology of Karlin & Studden (1966), Karlin (1968), Pinkus (1985a) (Chapter III), Pinkus (1996), and Pinkus (2010) in this work.

A compact notation for determinants is helpful when working with totally positive kernels; I follow Karlin (1968) (page 1 in Chapter 0, where only the square-matrix case is considered) and Pinkus (1985a) (page 45 in Section 2 of Chapter 3), who employ Fredholm's notation (originally used in the continuous case for kernels).

Definition 5.14. Given the m-row and n-column matrix

$$A \quad = \quad \left(\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array} \right),$$

the minors of this matrix are denoted

$$A \left(\begin{array}{cccc} i_1 & i_2 & \cdots & i_p \\ k_1 & k_2 & \cdots & k_p \end{array} \right) \quad = \quad \left| \begin{array}{cccc} a_{i_1 k_1} & a_{i_1 k_2} & \cdots & a_{i_1 k_p} \\ a_{i_2 k_1} & a_{i_2 k_2} & \cdots & a_{i_2 k_p} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_p k_1} & a_{i_p k_2} & \cdots & a_{i_p k_p} \end{array} \right|,$$

where

$$\begin{array}{c} 1 \le i_1 < i_2 < \cdots < i_p \le m \\ 1 \le k_1 < k_2 < \cdots < k_p \le n. \end{array}$$

Now I need definitions characterizing certain classes of matrices, which come from Karlin (1968) (pages 11 and 12 in Section 1 of Chapter 1) and Pinkus (1985a) (page 47 in Section 2 of Chapter III).

Definition 5.15. A square matrix $A = (a_{ik})_1^n$ will be called *strictly totally positive* (abbreviated STP) if and only if $\forall p \in \{1, 2, \ldots, n\}$,

$$A \left(\begin{array}{cccc} i_1 & i_2 & \cdots & i_p \\ k_1 & k_2 & \cdots & k_p \end{array} \right) > 0 \quad \text{for} \quad 1 \le \begin{array}{ccccccc} i_1 & < & i_2 & < & \cdots & < & i_p \\ k_1 & < & k_2 & < & \cdots & < & k_p \end{array} \le n.$$

Definition 5.16. A square matrix $A = (a_{ik})_1^n$ will be called *totally positive* (abbreviated TP) if and only if $\forall p \in \{1, 2, \ldots, n\}$,

$$A \left(\begin{array}{cccc} i_1 & i_2 & \cdots & i_p \\ k_1 & k_2 & \cdots & k_p \end{array} \right) \ge 0 \quad \text{for} \quad 1 \le \begin{array}{ccccccc} i_1 & < & i_2 & < & \cdots & < & i_p \\ k_1 & < & k_2 & < & \cdots & < & k_p \end{array} \le n.$$

An intermediate concept, between STP and TP, is often useful; the following definition echoes Karlin (1968) (page 35 in Section 6 of Chapter 1) and Pinkus (1985a) (page 50 in Section 2 of Chapter III) (though I use the term "oscillatory" rather than the term "oscillating" used by Karlin (1968) and Pinkus (1985a)).

Definition 5.17. A square matrix $A = (a_{ik})_1^n$ will be called *oscillatory* if and only if A is totally positive and a positive integer k exists such that A^k is strictly totally positive.

Evaluating the exponential product kernel provides useful examples of strictly totally positive matrices, as the following theorem shows.

Theorem 5.7. *For any* $\begin{array}{ccccccc} x_1 & < & x_2 & < & \cdots & < & x_n \\ y_1 & < & y_2 & < & \cdots & < & y_n \end{array} \in \mathbb{R}$*, the matrix*

$$\left(\begin{array}{cccc} \exp(cx_1y_1) & \exp(cx_1y_2) & \cdots & \exp(cx_1y_n) \\ \exp(cx_2y_1) & \exp(cx_2y_2) & \cdots & \exp(cx_2y_n) \\ \cdots & \cdots & \cdots & \cdots \\ \exp(cx_ny_1) & \exp(cx_ny_2) & \cdots & \exp(cx_ny_n) \end{array} \right) \tag{59}$$

is STP*.*

Proof. This is an immediate consequence of Definition 5.15 and display (3.1), and the argument below it, in Karlin & Studden (1966) (on page 9 in Section 3 of Chapter I). ☐

STP and TP matrices enjoy variation-diminishing properties that are nicely described by Pinkus (1985a) and Pinkus (2010).

Proposition 5.5. *If A is a strictly totally positive $n \times n$ matrix and $\boldsymbol{x} \neq \mathbf{0}$, $\boldsymbol{x} \in \mathbb{R}^n$, then*

$$S^+(A\boldsymbol{x}) \quad \leq \quad S^{-1}(\boldsymbol{x}). \tag{60}$$

Proposition 5.5 appears as a portion of part (ii) of Theorem 2.4 on page 47 in Section 2 of Chapter III in Pinkus (1985a). It is also a special case of Theorem 3.3 on page 77 in Section 1 of Chapter 3 in Pinkus (2010).

Proposition 5.6. *If A is a nonsingular totally positive $n \times n$ matrix, then*

$$S^-(A\boldsymbol{x}) \quad \leq \quad S^{-1}(\boldsymbol{x}) \tag{61}$$

for all vectors $\boldsymbol{x} \in \mathbb{R}^n$.

Proposition 5.6 is included in part (ii) of Theorem 2.3 on page 47 in Section 2 of Chapter III in Pinkus (1985a). It is also a special case of Theorem 3.4 on page 81 in Section 1 of Chapter 3 in Pinkus (2010).

Moving to the continuous case involves the kernels $K(x,y)$ of integral operators $P[f](y) = \int_a^b K(x,y) f(x)\,dx$. There is a clear analog, used by Karlin (1968) (page 11 in Section 1 of Chapter1) and Pinkus (1985a) (page 52 in Section 3 of Chapter III), to the notation given in Definition 5.14:

Definition 5.18. Given a kernel $K(x,y)$ defined on $x, y \in [a,b]$, I use the following notation:

$$K\left(\begin{array}{cccc} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{array} \right) \quad = \quad \left| \begin{array}{cccc} K(x_1,y_1) & K(x_1,y_2) & \cdots & K(x_1,y_n) \\ K(x_2,y_1) & K(x_2,y_2) & \cdots & K(x_2,y_n) \\ \cdots & \cdots & \cdots & \cdots \\ K(x_n,y_1) & K(x_n,y_2) & \cdots & K(x_n,y_n) \end{array} \right|,$$

where

$$a \leq \begin{array}{ccccccc} x_1 & < & x_2 & < & \cdots & < & x_n \\ y_1 & < & y_2 & < & \cdots & < & y_n \end{array} \leq b.$$

There are continuous analogs to Definitions 5.15 and 5.16, which come from Karlin (1968) (pages 11 and 12 in Section 1 of Chapter 1) and Pinkus (1985a) (page 52 in Section 3 of Chapter III).

Definition 5.19. A kernel $K(x,y) \in C([a,b] \times [a,b])$ will be called *strictly totally positive* (abbreviated STP) if and only if $\forall n \geq 1$,

$$K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} > 0 \quad \text{for} \quad a \leq \begin{array}{ccccccc} x_1 & < & x_2 & < & \cdots & < & x_n \\ y_1 & < & y_2 & < & \cdots & < & y_n \end{array} \leq b.$$

Definition 5.20. A kernel $K(x,y) \in C([a,b] \times [a,b])$ will be called *totally positive* (abbreviated TP) if and only if $\forall n \geq 1$,

$$K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} \geq 0 \quad \text{for} \quad a \leq \begin{array}{ccccccc} x_1 & < & x_2 & < & \cdots & < & x_n \\ y_1 & < & y_2 & < & \cdots & < & y_n \end{array} \leq b.$$

The exponential product kernel is a strictly totally positive kernel, as the following theorem shows.

Theorem 5.8. *For any $c > 0$, the exponential product kernel* $\exp(cxy)$ *is STP on $\mathbb{R} \times \mathbb{R}$ (and thus on $[a,b] \times [a,b]$ for any $[a,b] \subset \mathbb{R}$).*

Proof. This is an immediate consequence of Definition 5.19 and display (3.1), and the argument below it, in Karlin & Studden (1966) (on page 9 in Section 3 of Chapter I). Alternatively, it is implied by Corollary 3.2 on page 53 in Section 3 of Chapter III in Pinkus (1985a). □

It will also be useful to consider generalizations of total positivity and stricty total positivity.

Definition 5.21. A kernel $K(x,y) \in C([a,b] \times [a,b])$ will be called *strictly sign regular* (abbreviated SSR) if and only if $\forall n \geq 1$,

$$\epsilon_n K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} > 0 \quad \text{for} \quad a \leq \begin{array}{ccccccc} x_1 & < & x_2 & < & \cdots & < & x_n \\ y_1 & < & y_2 & < & \cdots & < & y_n \end{array} \leq b.$$

Here ϵ_n is either 1 or -1 and changes only with n, not with any x_i or y_i.

Definition 5.22. A kernel $K(x,y) \in C([a,b] \times [a,b])$ will be called *sign regular* (abbreviated SR) if and only if $\forall n \geq 1$,

$$\epsilon_n K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} \geq 0 \quad \text{for} \quad a \leq \begin{array}{ccccccc} x_1 & < & x_2 & < & \cdots & < & x_n \\ y_1 & < & y_2 & < & \cdots & < & y_n \end{array} \leq b.$$

Here ϵ_n is either 1 or -1 and changes only with n, not with any x_i or y_i.

These definitions come from Karlin (1968) (page 12 in Section 1 of Chapter 1).

It will be convenient to define a specific interval that depends on the properties of the kernel under consideration. I echo Gantmacher & Krein (2002)'s notation here, but I follow Hogan & Lakey (2012) in adding a subscript to clarify the dependence of the interval on the specific kernel considered.

Definition 5.23. For any kernel $K(x,y)$ defined on $x, y \in [a,b]$, the interval I_K is defined as:

$$I_K \quad \equiv \quad \begin{cases} (a,b) & \text{if} \quad K(a,a) = K(b,b) = 0 \\ (a,b] & \text{if} \quad K(a,a) = 0, K(b,b) \neq 0 \\ [a,b) & \text{if} \quad K(a,a) \neq 0, K(b,b) = 0 \\ [a,b] & \text{if} \quad K(a,a) \neq 0, K(b,b) \neq 0 \end{cases}.$$

Gantmacher & Krein (2002) (pages 178 - 179) give two equivalent analogs of Definition 5.17 for the kernel $K(x,y)$ of an integral operator $P[f](y) = \int_a^b K(x,y) f(x)\,dx$.

Definition 5.24. The kernel $K(x,y) \in C([a,b] \times [a,b])$ is called *oscillatory* if and only if the following three conditions hold.

$$K(x,y) > 0 \;\forall x, y \in I_K \qquad \text{unless}\, x = a, y = b \,\text{or}\, x = b, y = a \quad (62)$$

$$K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} \geq 0 \quad \text{for} \quad a < \begin{matrix} x_1 < x_2 < \cdots < x_n \\ y_1 < y_2 < \cdots < y_n \end{matrix} < b \quad (63)$$

$$K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} > 0 \quad \text{for} \quad a < x_1 < x_2 < \cdots < x_n < b. \quad (64)$$

The equivalent definition to Definition 5.24 is, as Gantmacher & Krein (2002) demonstrate:

Definition 5.25. The kernel $K(x,y) \in C([a,b] \times [a,b])$ is called *oscillatory* if and only if for any collection of $x_1 < x_2 < \cdots < x_n \in I_K$ among which there is at least one interior point, the matrix $A = (K(x_i, x_k))_1^n$ is oscillatory.

Note that I have made explicit the inequalities $x_1 < x_2 < \cdots < x_n$ (these are clear in Gantmacher & Krein (2002) from context including their footnote 9 on page 179, but may be helpful to specify here). It is evident from the (equivalent) definitions that any STP kernel is oscillatory.

The following is a slightly more general analog of Definition 5.17 for the kernel $K(x,y)$ generating the integral operator $P[f](y) = \int_a^b K(x,y) f(x)\,dx$. Gantmacher & Krein (2002) (page 180) call kernels satisfying the conditions below *Kellogg kernels*, which led me to this definition. It omits only the first criterion from the definition of an oscillatory kernel (see Definition 5.24). As Gantmacher & Krein (2002) point out, the notion of an oscillatory kernel is needed only for integral operators in which the measure used is not Lebesgue measure; since I use Lebesgue measure, I make use of the more general notion of a Kellogg kernel.

Definition 5.26. The kernel $K(x,y) \in C([a,b] \times [a,b])$ is called *Kellogg* if and only if the following two conditions hold.

$$K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} \geq 0 \quad \text{for} \quad a < \begin{matrix} x_1 < x_2 < \cdots < x_n \\ y_1 < y_2 < \cdots < y_n \end{matrix} < b \quad (65)$$

$$K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} > 0 \quad \text{for} \quad a < x_1 < x_2 < \cdots < x_n < b. \quad (66)$$

Clearly, any STP kernel is automatically Kellogg, so the exponential product kernel is Kellogg; I summarize this in Lemma 5.1 below.

Lemma 5.1. *For any $c > 0$, the exponential product kernel* $\exp(cxy)$*, defined on $x, y \in [-1, 1]$, is symmetric, continuous, Kellogg, and strictly positive.*

Proof. The kernel $\exp(cxy)$ is clearly symmetric, continuous and strictly positive for $x, y \in [-1, 1]$. For any $c > 0$, it is STP by Theorem 5.8; thus, it is Kellogg for any $c > 0$ by Definition 5.26. $\square$

I now summarize a key result in the literature on the eigenstructure of Kellogg kernels.

Definition 5.27. Any eigenvalue λ and an associated eigenfunction ψ (such that $\|\psi\| = \int_{-1}^{1} \psi^2(t)\, dt = 1$) of an operator P whose domain is contained in $L^2[-1, 1]$ satisfy:

$$\lambda\psi(y) \quad = \quad P[\psi](y). \tag{67}$$

If P is, additionally, an integral operator with kernel $K(x, y)$ defined on $[-1, 1] \times [-1, 1]$ then

$$\lambda\psi(y) \quad = \quad \int_{-1}^{1} K(x, y)\, \psi(x)\, dx. \tag{68}$$

Under appropriate regularity assumptions given below, the eigenvalues and eigenfunctions of the integral operator $P[f](y) = \int_{-1}^{1} K(x, y) f(x)\, dx$ associated with a Kellogg kernel will satisfy the following very useful set of properties.

Definition 5.28. A linear operator P on a Hilbert space is *positive definite with simple eigenvalues and Markov eigenfunctions* if and only if

1. All of the eigenvalues λ_i of P are positive and simple:
$$\lambda_0 > \lambda_1 > \lambda_2 > \cdots \quad \text{with} \quad \lambda_i > 0, i = 0, 1, 2, \ldots.$$
2. The eigenfunction $\psi_0(x)$ corresponding to the largest eigenvalue λ_0 does not have any zeros in $[-1, 1]$.
3. For $j = 1, 2, 3, \ldots$, the eigenfunction $\psi_j(x)$ corresponding to λ_j (the $(j+1)^{\text{st}}$-largest eigenvalue) has exactly j nodes in $(-1, 1)$ and no other zeros in $[-1, 1]$.
4. For every pair of integers k and m with $0 \le k \le m$ and for every set of arbitrary real numbers c_i $\left(i = k, k+1, \ldots, m; \sum_{i=k}^{m} c_i^2 > 0\right)$, the function $\psi(x) = \sum_{i=k}^{m} c_i \psi_i(x)$ has at most m zeros in $[-1, 1]$ and at least k nodes in $(-1, 1)$.
5. The nodes of two adjacent eigenfunctions $\psi_j(x)$ and $\psi_{j+1}(x)$ alternate for $j = 1, 2, \ldots$.

Remark 5.4. In most of this paper, indices begin with 1. For eigenvalues, eigenfunctions, and orthogonal polynomials, however, I adhere to the seemingly-uniform convention in the literature on the special functions of classical mathematical physics and index beginning with 0, as in Definition 5.28. To do otherwise would, I believe, invite great confusion in relating my results to the literature.

The following theorem is just a special case of a very powerful result provided in Gantmacher & Krein (2002) (Theorem 3 of Chapter IV, Section 3), and it is one of the key tools used in my analysis.

Theorem 5.9. *Suppose that the kernel $K(x,y) \in C([a,b]\times[a,b])$, is symmetric, Kellogg, and strictly positive for $x,y \in [-1,1]$. Then the associated integral operator $P[f](y) = \int_{-1}^{1} K(x,y) f(x)\,dx$ is positive definite with simple eigenvalues and Markov eigenfunctions.*

Proof. Since the kernel $K(x,y)$ is strictly positive for $x,y \in [-1,1]$, Definition 5.23 implies that the interval $I_K = [-1,1]$. Further, the kernel $K(x,y)$ is Kellogg, so by Definition 5.26

$$K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} \geq 0 \quad \text{for} \quad a < \begin{matrix} x_1 & < & x_2 & < & \cdots & < & x_n \\ y_1 & < & y_2 & < & \cdots & < & y_n \end{matrix} < b$$

$$K\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} > 0 \quad \text{for} \quad a < x_1 < x_2 < \cdots < x_n < b.$$

These facts, and the assumed continuity and symmetry of the kernel $K(x,y)$ for $x,y \in [-1,1]$, supply the assumptions of Gantmacher & Krein (2002)'s Theorem 3 of Chapter IV, Section 3, which then provides the desired conclusion once three facts are noted: the I of Gantmacher & Krein (2002)'s theorem is $[-1,1]$ given the assumptions made, the function $\sigma(s)$ in Gantmacher & Krein (2002)'s theorem is simply $\sigma(s) = s$ here (and is thus certainly strictly increasing as Gantmacher & Krein (2002) require in their theorem), and Gantmacher & Krein (2002) use the notation $\phi(x) = \lambda \int_a^b K(x,s)\,\phi(s)\,d\sigma(s)$ for an eigenvalue-eigenfunction pair of an integral equation, so the λ_i of their theorem is actually $\frac{1}{\lambda_i}$ here (and all rankings of eigenvalues reverse accordingly). □

Corollary 5.1. *For any $c > 0$, the exponential product operator is positive definite with simple eigenvalues and Markov eigenfunctions (as defined in Definition 5.28).*

Proof. The exponential product operator has the kernel $K(x,y) = \exp(cxy)$, defined on $[-1,1]\times[-1,1]$, and for any $c > 0$ this kernel satisfies all of the assumptions of Theorem 5.9 by Lemma 5.1. □

5.4 Total Positivity and n-Widths

Total positivity has been quite useful in analyzing the n-widths of integral operators. Melkman & Micchelli (1978) proved that each Kolmogorov n-width

($n = 1, 2, \ldots$) of an integral operator with a nondegenerate totally positive kernel (a condition weaker than strict total positivity, see Pinkus (1985a), Definition 5.1 on page 108 in Section 5 of Chapter IV) that maps $L^2[0,1]$ to $L^2[0,1]$ is achieved by two optimal interpolatory subspaces (as well as the usual eigenfunction subspace, see Section 8 below). This constructively disproved the assertion of Kolmogorov (1936) regarding the uniqueness of the optimal subspace for each Kolmogorov n-width in Hilbert-space settings (though this assertion was previously known to be false). The results of Melkman & Micchelli (1978) are nicely reviewed and summarized in Pinkus (1985a) (Section 5 of Chapter IV).

Going beyond Hilbert-space settings, Micchelli & Pinkus (1978) determined the exact Kolmogorov, Gel'fand, linear, and Bernstein n-widths of an integral operator with a nondegenerate totally positive kernel mapping $L^\infty[0,1]$ to any $L^q[0,1]$, $1 \le q \le \infty$; they also found the exact n-widths of an integral operator with a nondegenerate totally positive kernel mapping any $L^p[0,1]$, $1 \le p \le \infty$, to $L^1[0,1]$. This extended prior work on integral operators from $L^\infty[0,1]$ to $L^\infty[0,1]$ (Micchelli & Pinkus (1977b)) and from $L^1[0,1]$ to $L^1[0,1]$ (Micchelli & Pinkus (1977c)) by the same authors. The optimal subspaces and linear operators are also determined; they are interpolatory for Kolmogorov and linear n-widths and have corresponding forms for Gel'fand and Bernstein n-widths. Pinkus (1985a) reviews this work in Section 2 of Chapter V (Section 3 of Chapter V covers extensions also obtained by Micchelli & Pinkus (1978)). Pinkus (1985b) extends these results to the case of integral operators with nondegenerate totally positive kernels mapping $L^p[0,1]$ to $L^p[0,1]$, $1 \le p \le \infty$.

My work here is different from prior studies of n-widths that exploit total positivity in two related regards: first, it concerns the "most difficult case" in which the domain of the integral operator is the set of signed measures with bounded total variation and the range of the integral operator is equipped with the supremum norm. (Pinkus (1985a) conjectures (on page 138 in Section 1 of Chapter V) that results such as those of Micchelli & Pinkus (1978) hold for integral operators from $L^p[0,1]$ to $L^q[0,1]$ where $1 \le q \le p \le \infty$; the case I consider is, in some sense, as far as possible from these "nicer" cases, since I study (the closure of) $p = 1$ and $q = \infty$.) Second, due to the difficulty of the problem and to my practical goal of optimal portfolio approximation, I have a very computational focus. I content myself with results that are verifiably numerically optimal rather than proving "with pencil and paper," as the studies cited above do, the optimality of my linear operators.

In noting my computational focus, it is worth mentioning that much of the work cited in the preceding three paragraphs is theoretical in character; outside of specific examples such as spline subspaces, it is not typical in the work cited above to give methods for actually computing optimal subspaces for use in applied problems.

The closest results to mine in the n-width literature appear to belong to Melkman (1985), who studies entire functions of exponential type which have magnitude bounded by 1 on the real line excluding some centered interval. This is, in some sense, the trigonometric version of the problem I consider; Melkman (1985) notes its importance in the context of time- and band-limited signals. In line with my comments above, Melkman (1985) is not concerned with the actual computation of optimal subspaces, but he does give a remarkable characterization of the optimal subspaces for his problem. Numerical examples show that analogs of the Melkman (1985) strategy (based on complex analysis) do not work for the problem I consider (there was no great expectation that they would, since nontrivial linear combinations of real exponentials are certainly not bounded on the real line). Pinkus (1985a) (pages 270 through 275 in Section 5 of Chapter VIII) provides a concise review of the results of Melkman (1985) (note that Melkman (1985) was not yet published when Pinkus (1985a) wrote, so Pinkus (1985a) cites a version of this work from 1982). As a final note on the work of Melkman (1985), one might be tempted to try to apply his work after making the same trigonometric substitution used in Section 9 below, resulting in the expression $\exp(c\cos(\theta)\cos(\eta))$, with $\theta, \eta \in [0,\pi]$, for the kernel of the exponential product operator. Unfortunately, that expression gives an entire function which is not of exponential type in either variable (it grows as $\exp(c_0 \exp(|\theta|))$ for a constant c_0 as $|\theta| \to \infty$ along the imaginary axis, for example, rather than growing as $\exp(c_0|\theta|)$), so the results of Melkman (1985) cannot be applied even after this substitution.

6 Computable Lower Bounds for Approximation Error

In this section, I use Definition 4.4 to show that the Bernstein n-width (for worst-case error) of the exponential product operator with parameter c is bounded below by

$$\min_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1} \max_{y\in[-1,1]} \left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right| \tag{69}$$

for any $\rho_1, \ldots, \rho_{n+1} \in [-1,1]$. I use "max" and "min" rather than "sup" and "inf" because I will show that the supremum and infimum are achieved (due to continuity and compactness). By Theorem 4.1 the linear, Kolmogorov, and Gel'fand n-widths are all bounded below by the Bernstein n-width, so (69) is a lower bound, for any $\rho_1, \ldots, \rho_{n+1} \in [-1,1]$, for these n-widths as well. I demonstrate that the expression (69), viewed as a function from $(\rho_1, \ldots, \rho_{n+1}) \in [-1,1]^{n+1}$ to $\mathbb{R}$, is a continuous function on a compact set, so the supremum of (69) over $(\rho_1, \ldots, \rho_{n+1}) \in [-1,1]^{n+1}$ is attained and provides the greatest lower bound of this type:

$$\max_{\rho_1,\ldots,\rho_{n+1}\in[-1,1]} \min_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1} \max_{y\in[-1,1]} \left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right|. \tag{70}$$

For given $\rho_1, \ldots, \rho_{n+1}$, I provide a Remez-like method to compute (69); this method makes use of my characterization of the solution of (69) in Theorem 5.6. I also give another method that uses the Remez-like method in each step of an iteration; this second method is useful in finding (70).

Definition 6.1.

$$\widetilde{\Omega}_m \equiv \{(t_1, \dots, t_m) \in [-1,1]^m : t_1 \le t_2 \le \dots \le t_m\} \tag{71}$$
$$\Omega_m \equiv \{(t_1, \dots, t_m) \in [-1,1]^m : t_1 < t_2 < \dots < t_m\} \tag{72}$$

Definition 6.2. For any $\boldsymbol{\rho} \in \widetilde{\Omega}_{n+1}$,

$$B_c(\boldsymbol{\rho}) \equiv \min_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1} \max_{y\in[-1,1]} \left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right|. \tag{73}$$

Theorem 6.1. *The function $B_c(\boldsymbol{\rho})$ is a continuous function from $\widetilde{\Omega}_{n+1}$ to $\mathbb{R}$.*

Proof. First observe that $\left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right|$ is a continuous function of $(\boldsymbol{a}, \boldsymbol{\rho}, y)$. By Berge (1963) (Theorem 1 on page 115 in Section 3 of Chapter VI, taking the mapping Γ in that theorem to be the map assigning to each value of $(\boldsymbol{a}, \boldsymbol{\rho})$ the clearly non-empty set $[-1,1]$), $\max_{y\in[-1,1]}\left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right|$ is a lower semi-continuous function of $(\boldsymbol{a}, \boldsymbol{\rho})$. By Berge (1963) (Theorem 2 on page 116 in Section 3 of Chapter VI, taking the mapping Γ in that theorem to be the map assigning to each value of $(\boldsymbol{a}, \boldsymbol{\rho})$ the clearly non-empty set $[-1,1]$), it is also an upper semi-continuous function of $(\boldsymbol{a}, \boldsymbol{\rho})$. Thus, $\max_{y\in[-1,1]}\left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right|$ is a continuous function of $(\boldsymbol{a}, \boldsymbol{\rho})$. (The results just applied are often grouped together as "the maximum theorem"; I used the same logic in the proof of Theorem 5.6) Now consider $\max_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1}\left\{-\max_{y\in[-1,1]}\left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right|\right\}$. Applying the same results of Berge (1963) noted above, but now with the map Γ defined to take every value of ρ to the clearly non-empty set $\left\{\boldsymbol{a} : \sum_{i=1}^{n+1}|a_i| = 1\right\}$, I conclude that $\max_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1}\left\{-\max_{y\in[-1,1]}\left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right|\right\}$ is a continuous function of $\boldsymbol{\rho}$. But then

$$\begin{aligned} & -\max_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1}\left\{-\max_{y\in[-1,1]}\left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right|\right\} \\ = & \min_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1} \max_{y\in[-1,1]} \left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right| \end{aligned} \tag{74}$$

is also continuous, which completes the demonstration. □

Theorem 6.2. *Let M_{TV} be the set of all signed measures on $[-1,1]$ with bounded total variation (this is the normed space of signed measures where the norm is total variation), and let $C([-1,1])$ be the space of continuous functions on $[-1,1]$ equipped with the max norm. Then, for any $\boldsymbol{\rho} \in \widetilde{\Omega}_{n+1}$ the Bernstein n-width satisfies*

$$b_n(P_c(M_{TV}); C([-1,1])) \ge B_c(\boldsymbol{\rho}). \tag{75}$$

The supremum over $\boldsymbol{\rho} \in \widetilde{\Omega}_{n+1}$ of $B_c(\rho)$ is attained, so the greatest lower bound of this type is

$$b_n\left(P_c\left(M_{TV}\right); C\left([-1,1]\right)\right) \quad \geq \quad \max_{\boldsymbol{\rho}\in\widetilde{\Omega}_{n+1}} B_c\left(\boldsymbol{\rho}\right). \tag{76}$$

Proof. Let R_{n+1} be the space of all signed measures of bounded total variation which have nonzero measure only on some subset of the points $\rho_1, \ldots, \rho_{n+1} \in [-1,1]$. Then R_{n+1} is clearly a subspace of the M_{TV}. Further,

$$\begin{aligned}
P_c R_{n+1} \quad &\equiv \quad \left\{P_c\left[\mu\right] : \mu \in R_{n+1}\right\} &(77)\\
&= \quad \left\{\int_{-1}^{1} \exp\left(cxy\right) d\mu\left(x\right) : \mu \in R_{n+1}\right\} &(78)\\
&= \quad \left\{\sum_{i=1}^{n+1} a_i \exp\left(c\rho_i y\right) : a_1, \ldots a_{n+1} \in \mathbb{R}\right\}, &(79)
\end{aligned}$$

where the first line is a definition of $P_c R_{n+1}$, the second line follows from the definition of P_c in (3.1), and the third line follows from the definition of R_{n+1} earlier in this paragraph ($\mu \in R_{n+1}$ places measure $a_i \in \mathbb{R}$ on the point ρ_i, $i = 1, \ldots, n+1$; since any $\mu \in R_{n+1}$ has nonzero measure only on some subset of $\rho_1, \ldots, \rho_{n+1} \in [-1,1]$, the integral in the second line becomes a sum involving only these ρ_i in the third line).

R_{n+1} is a subspace of M_{TV} as noted above, so $P_c R_{n+1}$ is a subspace of the image under P_c of the space of all measures of bounded total variation on $[-1,1]$. Recalling the definition of the Bernstein n-width (Definition 4.4), I now argue that

$$\begin{aligned}
&b_n\left(P_c\left(M_{TV}\right); C\left([-1,1]\right)\right)\\
&= \sup_{Y_{n+1}} \left\{ \inf_{\substack{P_c[f]\in Y_{n+1}\\ P_c[f]\neq 0}} \left\{ \frac{\sup_{y\in[-1,1]} \left|P_c\left[f\right](y)\right|}{\|f\|_{TV}} \right\} \right\} &(80)\\
&\geq \inf_{\substack{P_c[f]\in P_c R_{n+1}\\ P_c[f]\neq 0}} \left\{ \frac{\sup_{y\in[-1,1]} \left|P_c\left[f\right](y)\right|}{\|f\|_{TV}} \right\} &(81)\\
&= \inf_{\substack{P_c[f]\in P_c R_{n+1}\\ P_c[f]\neq 0}} \left\{ \sup_{y\in[-1,1]} \left| P_c\left[\frac{f}{\|f\|_{TV}}\right](y) \right| \right\} &(82)\\
&= \inf_{\substack{P_c[g]\in P_c R_{n+1}\\ \|g\|_{TV}=1}} \left\{ \sup_{y\in[-1,1]} \left|P_c\left[g\right](y)\right| \right\} &(83)\\
&= \inf_{\substack{P_c[g]=\sum_{i=1}^{n+1} a_i \exp\left(c\rho_i y\right)\\ \sum_{i=1}^{n+1}|a_i|=1}} \left\{ \sup_{y\in[-1,1]} \left|P_c\left[g\right](y)\right| \right\} &(84)\\
&= \min_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1} \left\{ \max_{y\in[-1,1]} \left| \sum_{i=1}^{n+1} a_i \exp\left(c\rho_i y\right) \right| \right\} &(85)
\end{aligned}$$

$$= B_c(\boldsymbol{\rho}), \tag{86}$$

where the first equality is by Definition 4.4 (simply substituting the relevant operator and norms into the definition), the first inequality follows from the fact that $P_c R_{n+1}$ is a subspace of span $\{P_c[f] : f \in M_{TV}\}$ of dimension $n+1$ (and is thus one of the Y_{n+1} that the supremum in the preceding line is taken over), the second equality is by the scaling property of the norm ($\alpha \sup_{y\in[-1,1]} |g(y)| = \sup_{y\in[-1,1]} |\alpha g(y)|$ for any constant α) and the linearity of P_c, the third equality recognizes that $P_c[f] \neq 0 \Rightarrow \|f\|_{TV} \neq 0 \Rightarrow \left\|\frac{f}{\|f\|_{TV}}\right\|_{TV} = 1$ and (because P_c is injective) $\left\|\frac{f}{\|f\|_{TV}}\right\|_{TV} = 1 \Rightarrow \|f\|_{TV} \neq 0 \Rightarrow P_c[f] \neq 0$, the fourth equality is by (79) and the fact that the total variation of $f \in R_{n+1}$ is just the sum of the absolute values of the measures a_i that f places on the ρ_i, the fifth equality simplifies the preceding expression and acknowledges that the supremum and infimum are attained by replacing them with a maximum and a minimum, respectively (this is by continuity, which was demonstrated in the proof of Theorem 6.1), and the sixth equality is by Definition 6.2.

I have now demonstrated the bound (75), but I must still show that (76) holds by proving that the supremum over $\boldsymbol{\rho} \in \widetilde{\Omega}_{n+1}$ of $B_c(\boldsymbol{\rho})$ is attained. The supremum is achieved because $B_c(\boldsymbol{\rho})$ is continuous as a function of $\boldsymbol{\rho}$, as proven in Theorem 6.1, and the supremum is taken over $\widetilde{\Omega}_{n+1}$ (which is a closed and bounded, and thus compact, subset of $\mathbb{R}^{n+1}$).

□

Corollary 6.1. *With M_{TV} and $C([-1,1])$ as in Theorem 6.2,*

$$d_n(P_c(M_{TV}); C([-1,1])) \geq \max_{\boldsymbol{\rho}\in\widetilde{\Omega}_{n+1}} B_c(\boldsymbol{\rho}) \tag{87}$$

$$d^n(P_c(M_{TV}); C([-1,1])) \geq \max_{\boldsymbol{\rho}\in\widetilde{\Omega}_{n+1}} B_c(\boldsymbol{\rho}) \tag{88}$$

$$\delta_n(P_c(M_{TV}); C([-1,1])) \geq \max_{\boldsymbol{\rho}\in\widetilde{\Omega}_{n+1}} B_c(\boldsymbol{\rho}), \tag{89}$$

and thus clearly

$$d_n(P_c(M_{TV}); C([-1,1])) \geq B_c(\boldsymbol{\rho}) \tag{90}$$

$$d^n(P_c(M_{TV}); C([-1,1])) \geq B_c(\boldsymbol{\rho}) \tag{91}$$

$$\delta_n(P_c(M_{TV}); C([-1,1])) \geq B_c(\boldsymbol{\rho}) \tag{92}$$

for any $\boldsymbol{\rho} \in \widetilde{\Omega}_{n+1}$.

Proof. Combine the bounds of Theorem 6.2 on the Bernstein n-width with the inequalities of Theorem 4.1. □

The lower bound demonstrated in Theorem 6.2 can be computed very efficiently by a Remez-like method.

Method 6.1. The following method computes the lower bound (75) given $\boldsymbol{\rho} \in \widetilde{\Omega}_{n+1}$.

- **Given**:
  1. $c > 0$
  2. $\boldsymbol{\rho} \in \widetilde{\Omega}_{n+1}$
  3. a tolerance
  4. a maximum allowable number of iterations
  5. optionally, a grid used in zero-finding over the interval $[-1, 1]$
- **Output**:
  1. a vector $\boldsymbol{a}$ of coefficients such that $\sum_{i=1}^{n+1} |a_i| = 1$ that achieves the minimum over such $\boldsymbol{a}$ in (75) to within numerical tolerance
  2. a vector $\boldsymbol{y}$ of the nodes at which the maximum over $y \in [-1, 1]$ in (75) is attained to within numerical tolerance; if $\boldsymbol{\rho} \in \Omega_{n+1}$, then $\boldsymbol{y} \in \Omega_{n+1}$
  3. a $1 \times (n+1)$ vector $\boldsymbol{\epsilon}$ of errors that equioscillate, so that $\epsilon_{i+1} = -\epsilon_i$ for $i = 1, \ldots, n$, to within numerical tolerance
  4. a scalar discrepancy representing the change in $\max_{i \in \{1,\ldots,n+1\}} |\epsilon_i|$ at the last iteration prior to return
  5. the number of iterations performed

Proceed as in the usual Remez method, but seek equioscillation of the linear combination $\sum_{i=1}^{n+1} v_i \exp(c\rho_i y)$ between -1 and 1 at each step of the iteration. In each iterative step, after finding the v_i rescale them to get $\boldsymbol{a}$ as described.

Method 6.1 is motivated by Theorem 5.6, which shows that the solution of the optimization problem (69) is unique (up to sign) and satisfies the alternation properties that Method 6.1 depends on.

Method 6.1 implements a nontrivial variant of a method due to Remez (sometimes labeled with his name, and sometimes called the "exchange" method), and specifically of the "multipoint" flavor of the Remez method. Cheney (1982) gives a helpful account of this method under the romanization "Remes" beginning on page 97 in Section 8 of Chapter 3. As noted by Cheney (1982), Veidinger (1960) provides a proof that the multipoint Remez method is quadratically convergent, and Meinardus (1967) (Theorem 84 on pages 111 and 112 of Section 1 in Chapter 7) proves an extension of the quadratic convergence result using a different argument. Powell (1981) gives another useful discussion of the method, focusing on a variant in which only one point is changed in each iteration and using the name "exchange" rather than "Remez" (see his Chapters 8 and 9). In the practical problems I examine later in this paper, Method 6.1 typically converges to within rounding error in five iterations or less.

Although Remez methods have been known, and extensively investigated, for decades, I am not aware of any prior work which implements any (nontrivial) variant of them to compute numerical lower bounds on n-widths of integral operators with symmetric, strictly totally positive kernels (such as the exponential product operators). It is worth noting that the results in this section do not use any properties of the exponential product operators beyond the symmetry and strict total positivity of their kernels. Any other integral operator with a symmetric and strictly totally positive kernel would also be subject to the analysis of this section, though the numerical results would, of course, change quantitatively (and perhaps qualitatively). For other work on Remez methods with coefficient constraints, see Pinkus & Strauss (1988) and the references therein.

Method 6.1 gives a very efficient means of computing a numerical value for the lower bound (75) for a given $\boldsymbol{\rho} \in \widetilde{\Omega}$. However, it remains to find a way to obtain the greatest lower bound of this type (76) by taking the maximum over $\boldsymbol{\rho} \in \widetilde{\Omega}$ of $B_c(\boldsymbol{\rho})$. This would appear to be a computationally intractable maximization problem; inspecting the definition of $B_c(\boldsymbol{\rho})$ leads to the suspicion, confirmed by numerical experiments, that it is not concave as a function of $\boldsymbol{\rho}$ (numerical evidence suggests that it may be log-concave). Fortunately, there is some hope for an efficient solution.

Definition 6.3. For any $\boldsymbol{\rho} \in \Omega_{n+1}$, the discussion of Method 6.1 shows that there are $n+1$ distinct values of $y \in [-1,1]$ which maximize $\left|\sum_{i=1}^{n+1} a_i^* \exp(c\rho_i y)\right|$, where

$$\boldsymbol{a}^* \quad = \quad \underset{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1}{\arg\min} \max_{y\in[-1,1]} \left|\sum_{i=1}^{n+1} a_i \exp(c\rho_i y)\right|. \tag{93}$$

(Note that $\boldsymbol{a}^*$ is defined up to a change in sign, and multiplying $\boldsymbol{a}^*$ by -1 will not alter the associated maximizing values of $y \in [-1,1]$.) Label these $n+1$ values in ascending order as $y_1(\rho), \ldots, y_{n+1}(\rho)$, and let

$$\boldsymbol{Y}_c(\boldsymbol{\rho}) \quad \equiv \quad \begin{pmatrix} y_1(\boldsymbol{\rho}) \\ \vdots \\ y_{n+1}(\boldsymbol{\rho}) \end{pmatrix} \in \Omega_{n+1}. \tag{94}$$

Theorem 6.3. *For any* $\boldsymbol{\rho} \in \Omega_{n+1}$,

$$B_c(\boldsymbol{\rho}) \quad \leq \quad B_c(\boldsymbol{Y}_c(\boldsymbol{\rho})). \tag{95}$$

Furthermore,

$$B_c(\boldsymbol{\rho}) \quad < \quad B_c(\boldsymbol{Y}_c(\boldsymbol{\rho})) \text{ unless } \boldsymbol{\rho} = \boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho})). \tag{96}$$

Proof. Refer to the components of $\boldsymbol{Y}_c(\boldsymbol{\rho})$ as $y_1 < \cdots < y_{n+1}$ (note that $\boldsymbol{Y}_c(\boldsymbol{\rho}) \in \Omega_{n+1}$, so these inequalities are strict) and form the $(n+1)\times(n+1)$ matrix $F(\boldsymbol{\rho}, \boldsymbol{Y}_c(\boldsymbol{\rho}))$ whose (i,j) element is $\exp(c\rho_i y_j)$. Then the equioscillation theorem implies that

$$\boldsymbol{a}^* F(\boldsymbol{\rho}, \boldsymbol{Y}_c(\boldsymbol{\rho})) \quad = \quad B_c(\boldsymbol{\rho})\, b, \tag{97}$$

where $\boldsymbol{a}^*$ is the row vector defined in (93) and $\boldsymbol{b}$ is defined as in Method 6.1 so that $b_j = (-1)^j$, $j = 1, \ldots, n+1$. (Note that, as mentioned in Definition 6.3, the vector $\boldsymbol{a}^*$ is defined up to a change in sign; choosing the sign of $\boldsymbol{a}^*$ appropriately will cause alternation to begin with a negative. Taking the opposite sign will result in alternation beginning with a positive.)

Following the same logic with $\boldsymbol{Y}_c(\boldsymbol{\rho})$ in place of $\boldsymbol{\rho}$ yields

$$\boldsymbol{a}^{**} F(\boldsymbol{Y}_c(\boldsymbol{\rho}), \boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))) = B_c(\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho})))\, \boldsymbol{b}, \tag{98}$$

where $\boldsymbol{a}^{**}$ satisfies (93) with $\boldsymbol{Y}_c(\boldsymbol{\rho})$ in place of $\boldsymbol{\rho}$. Now, $\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))$ is the *unique* ordered vector of $n+1$ elements $\tilde{y}_1 < \cdots < \tilde{y}_{n+1}$ such that

$$\min_{\boldsymbol{a}:\sum_{i=1}^{n+1}|a_i|=1} \max_{\tilde{y}\in\{\tilde{y}_1,\ldots,\tilde{y}_{n+1}\}} \left|\sum_{i=1}^{n+1} a_i \exp(c y_i \tilde{y})\right| = B_c(\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))). \tag{99}$$

Thus, replacing $\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))$ with $\boldsymbol{\rho}$ to get $F(\boldsymbol{Y}_c(\boldsymbol{\rho}), \boldsymbol{\rho})$ and solving for $\widehat{\boldsymbol{a}}$: $\sum_{i=1}^{n+1} |\widehat{a}_i| = 1$ that generates alternation (beginning with a negative) for $\widehat{\boldsymbol{a}} F(\boldsymbol{Y}_c(\boldsymbol{\rho}), \boldsymbol{\rho})$ yields

$$\widehat{\boldsymbol{a}} F(\boldsymbol{Y}_c(\boldsymbol{\rho}), \boldsymbol{\rho}) = \widehat{B}\boldsymbol{b}, \tag{100}$$

where $\widehat{B} \leq B_c(\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho})))$ (since $\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))$ is optimal) and $\widehat{B} < B_c(\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho})))$ unless $\boldsymbol{\rho} = \boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))$ (since $\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))$ is the unique optimizer).

By the symmetry of the exponential product kernel $\exp(cxy)$,

$$F(\boldsymbol{Y}_c(\boldsymbol{\rho}), \boldsymbol{\rho}) = F(\boldsymbol{\rho}, \boldsymbol{Y}_c(\boldsymbol{\rho}))^T. \tag{101}$$

For any vectors $\boldsymbol{u}, \boldsymbol{v} \in \Omega_{n+1}$, the matrix $F(\boldsymbol{u}, \boldsymbol{v})$ is strictly totally positive because the exponential product kernel is strictly totally positive. By the variation-diminishing property of strictly totally positive matrices, if $\boldsymbol{a}F(\boldsymbol{u}, \boldsymbol{v}) = C\boldsymbol{b}$ for some $C > 0$, then $\operatorname{sign}(a_i) = \operatorname{sign}(b_i) = (-1)^i$ for $i = 1, \ldots, n+1$. Applying this fact to all three of the relations developed thus far in this proof, I obtain

$$\begin{aligned}
\operatorname{sign}(a_i^*) &= \operatorname{sign}(a_i^{**}) && (102)\\
&= \operatorname{sign}(\widehat{a}_i) && (103)\\
&= \operatorname{sign}(b_i) && (104)\\
&= (-1)^i && (105)
\end{aligned}$$

for $i = 1, \ldots, n+1$.

Now observe that any row vector $\boldsymbol{a}$ which has $\operatorname{sign}(a_i) = \operatorname{sign}(b_i) = (-1)^i$

for $i = 1, \ldots, n+1$ and $\sum_{i=1}^{n+1} |a_i| = 1$ satisifies

$$\begin{aligned} \boldsymbol{b}\boldsymbol{a}^T &= \sum_{i=1}^{n+1} a_i b_i & (106)\\ &= \sum_{i=1}^{n+1} |a_i b_i| & (107)\\ &= \sum_{i=1}^{n+1} |a_i| \, |b_i| & (108)\\ &= \sum_{i=1}^{n+1} |a_i| & (109)\\ &= 1, & (110) \end{aligned}$$

where the first line is by the definition of matrix multiplication, the second line follows because $\boldsymbol{a}$ and $\boldsymbol{b}$ have the same sign for each entry, the third line is elementary, the fourth line is by $|b_i| = \left|(-1)^i\right| = 1$, and the fifth line is by the assumption that $\sum_{i=1}^{n+1} |a_i| = 1$. Each of $\boldsymbol{a}^*$, $\boldsymbol{a}^{**}$, and $\widehat{\boldsymbol{a}}$ satisfy the conditions used on $\boldsymbol{a}$ in the argument above, so I conclude that

$$\begin{aligned} \boldsymbol{b}\boldsymbol{a}^{*T} &= \boldsymbol{b}\boldsymbol{a}^{**T} & (111)\\ &= \boldsymbol{b}\widehat{\boldsymbol{a}}^T & (112)\\ &= 1. & (113) \end{aligned}$$

Combining the results above,

$$\begin{aligned} \boldsymbol{a}^* F\left(\boldsymbol{\rho}, \boldsymbol{Y}_c\left(\boldsymbol{\rho}\right)\right) \widehat{\boldsymbol{a}}^T &= B_c\left(\boldsymbol{\rho}\right) \boldsymbol{b}\widehat{\boldsymbol{a}}^T & (114)\\ &= B_c\left(\boldsymbol{\rho}\right) & (115) \end{aligned}$$

where the first equality is by (97) and the second equality is by $\boldsymbol{b}\widehat{\boldsymbol{a}}^T = 1$ as shown above. But also

$$\begin{aligned} \boldsymbol{a}^* F\left(\boldsymbol{\rho}, \boldsymbol{Y}_c\left(\boldsymbol{\rho}\right)\right) \widehat{\boldsymbol{a}}^T &= \left(\boldsymbol{a}^* F\left(\boldsymbol{\rho}, \boldsymbol{Y}_c\left(\boldsymbol{\rho}\right)\right) \widehat{\boldsymbol{a}}^T\right)^T & (116)\\ &= \widehat{\boldsymbol{a}} F\left(\boldsymbol{\rho}, \boldsymbol{Y}_c\left(\boldsymbol{\rho}\right)\right)^T \boldsymbol{a}^{*T} & (117)\\ &= \widehat{\boldsymbol{a}} F\left(\boldsymbol{Y}_c\left(\boldsymbol{\rho}\right), \boldsymbol{\rho}\right) \boldsymbol{a}^{*T} & (118)\\ &= \widehat{B}\boldsymbol{b}\boldsymbol{a}^{*T} & (119)\\ &= \widehat{B}, & (120) \end{aligned}$$

where the first equality follows from the fact that a scalar is its own transpose, the second equality follows by applying the rule of matrix multiplication and transposition that $(AB)^T = B^T A^T$ for any conforming matrices A, B, the third equality is by (101), the fourth equality is by (100), and the fifth equality holds because $\boldsymbol{b}\boldsymbol{a}^{*T} = 1$ as shown above.

Combining (115) and (120), I conclude that $B_c(\boldsymbol{\rho}) = \widehat{B}$. But $\widehat{B} \leq B_c(\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho})))$ and $\widehat{B} < B_c(\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho})))$ unless $\boldsymbol{\rho} = \boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))$ as observed in the second paragraph of this proof. Thus, $B_c(\boldsymbol{\rho}) \leq B_c(\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho})))$ and $B_c(\boldsymbol{\rho}) < B_c(\boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho})))$ unless $\boldsymbol{\rho} = \boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))$. □

Method 6.2. The following method computes a useful $\boldsymbol{\rho} \in \widetilde{\Omega}_{n+1}$.

- **Given**:
  1. $n+1$, the number of terms in the approximation
  2. $c > 0$
  3. a tolerance
  4. a maximum allowable number of iterations
- **Output**:
  1. $\boldsymbol{\rho} \in \Omega_{n+1}$
  2. a vector $\boldsymbol{a}$ of coefficients such that $\sum_{i=1}^{n+1} |a_i| = 1$
  3. $\boldsymbol{y} \in \Omega_{n+1}$
  4. a vector $\boldsymbol{\epsilon}$ of errors
  5. a scalar discrepancy representing the change in $B_c(\boldsymbol{\rho})$ at the last iteration prior to return
  6. the number of iterations performed

Invoke Method 6.1 at some starting guess for $\boldsymbol{\rho}$. This will result in a vector $\boldsymbol{y}$ of nodes. Set $\boldsymbol{\rho} = \boldsymbol{y}$ and iterate.

Conjecture 6.1. Method 6.2 converges to the unique $\boldsymbol{\rho}^*(c)$ which maximizes $B_c(\boldsymbol{\rho})$ subject to $\boldsymbol{\rho} \in \widetilde{\Omega}$.

Theorem 6.3 does not directly imply that Conjecture 6.1 holds, but it does suggest that the conjecture may hold: each iteration of Method 6.2 either increases the objective function $B_c(\boldsymbol{\rho})$ or, if $\boldsymbol{\rho} = \boldsymbol{Y}_c(\boldsymbol{Y}_c(\boldsymbol{\rho}))$, leaves it the same. Since a $\boldsymbol{\rho}$ whose entries are symmetric about zero generates a $\boldsymbol{Y}_c(\boldsymbol{\rho})$ whose entries are also symmetric about zero (in exact arithmetic), the enforcement of symmetry about zero in Method 6.2 does not impact the conclusion of Theorem 6.3. Extensive numerical evidence indicates that Method 6.2 does, in fact, rapidly compute the maximum in (76) and provide the $\boldsymbol{\rho}^*(c) \in \Omega_{n+1}$ for which the maximum is attained. Further, if there is a unique maximizing $\boldsymbol{\rho}^*(c)$ and if Method 6.2 converges to it (numerical evidence suggests both are true), then $\boldsymbol{\rho}^*(c)$ must be a fixed point of $\boldsymbol{Y}_c$: $\boldsymbol{\rho}^*(c) = \boldsymbol{Y}_c(\boldsymbol{\rho}^*(c))$ (see Definition 6.3).

Remark 6.1. Applying a standard interior-point optimization routine (with numerical computation of gradient and Hessian) to maximize $\log(B_c(\boldsymbol{\rho}))$ subject to the constraint $\boldsymbol{\rho} \in \widetilde{\Omega}$ computes a value for $B_c(\boldsymbol{\rho}^*)$ which is below, but quite close to, the value for $B_c(\boldsymbol{\rho}^*)$ computed by Method 6.2.

Remark 6.2. My later results do *not* depend on Method 6.2 finding the maximizing $\boldsymbol{\rho}^*(c) \in \Omega_{n+1}$. Since any $\boldsymbol{\rho} \in \Omega_{n+1}$ provides a lower bound through Method 6.1, one could view Method 6.2 as no more than a mechanism for selecting a potentially interesting $\boldsymbol{\rho}$ with which to construct a lower bound via Method 6.1. In every numerical experiment (that is, for every c and n I have examined), the lower bound thus constructed is within rounding error of the upper bound provided by the actual worst-case error of the approximation developed in Section 7. It was this numerical evidence, combined with the observation of Remark 6.1, that led me to Conjecture 6.1.

7 An Approximating Operator That Is Numerically Optimal

In this section, I construct a novel, numerically-optimal rank-n linear approximation to a given exponential product operator. Optimality is with respect to worst-case error, represented as usual by the same norms on cashflow streams (the total variation norm, treating normalized cashflows as signed measures) and on scenario profiles (the max norm, treating scenario profiles as continuous functions) used in Section 6.

As I show in Subsection 7.1, the nature of this numerically-optimal approximation is especially financially appealing: the n-dimensional image space of the approximation is spanned by n cashflow streams, all having their cashflows on the same $n+1$ times. Thus, a bulky and unwieldy stream of cashflows can be reduced, to within a (typically very small) approximation error, to a portfolio of n very simple cashflow streams for the purposes of a scenario profile. Further, any cashflow stream is approximated using the same n approximating streams, and the linear nature of the approximation means that the holdings of the n approximating streams in what may be thought of as the approximating portfolio aggregate naturally: if two target cashflow streams are combined into a target portfolio, one simply adds the holdings in their individual approximating portfolios together to get the new approximating portfolio. Since all approximating portfolios are composed of the same n approximating cashflow streams, the aggregation is just a sum of vectors.

In the search for a numerically-optimal rank-n linear approximation to a given exponential product operator, I make two key conjectures.

First, I make the educated guess that the optimal rank-n linear approximation to a given exponential product operator is also optimal in approximating a particular rank-$(n+1)$ operator formed from the exponential product operator using the $(n+1)$-dimensional vector $\boldsymbol{\rho}^*$ which maximizes the lower bound (76) in the preceding Section 6. This allows me to exploit the powerful and general results obtained by Micchelli & Pinkus (1979) (and summarized in Pinkus (1985a), Chaper V, Section 5) on the n-widths of rank-$(n+1)$ kernels. My educated guess is plausible due to the extended total positivity of the kernel of any exponential product operator. My guess is also optimistic: if it is correct, then the approximating rank-n linear operator I construct will achieve the lower bound on the Bernstein n-width given by the right-hand side of (76). It will then be the case that the linear, Kolmogorov, Gel'fand, and Bernstein n-widths are all equal (and it will thus be possible to numerically verify their equality to within rounding error).

Second, I conjecture that the optimization problem which Micchelli & Pinkus (1979) show provides the best rank-n approximation to this rank-$(n+1)$ operator is solvable using Theorem 5.5 (along lines similar to Theorem 5.6).

Of course, these conjectures only provide motivation; for each c (that is, each particular exponential product operator) and each n, I use interval analysis in a global optimization to demonstrate that the worst-case error of the approximating rank-n linear operator that I construct is, in fact, within rounding error of the lower bound computed by Method 6.2. A method for using interval analysis to find the worst-case approximation error of the rank-n linear operator that I build is given in Subsection 7.2.

Definition 7.1. The statement that a kernel $K(x,y) \in C([-1,1] \times [-1,1])$ is of rank $n+1$ means that

$$K(x,y) = \sum_{i=1}^{n+1} f_i(x)\, g_i(y), \qquad (121)$$

where $f_1, f_2, \ldots, f_{n+1} \in C([-1,1])$ are linearly independent and $g_1, g_2, \ldots, g_{n+1} \in C([-1,1])$ are also linearly independent.

This definition appears, with cosmetic differences, on page 115 of Micchelli & Pinkus (1979) (just above their Theorem 2.1) and at the top of page 188 in Section 5 of Chapter V in Pinkus (1985a).

Theorem 7.1. *If $K(x,y) \in C([-1,1] \times [-1,1])$ is of rank $n+1$, then let*

$$\sigma = \inf_{\sum_{i=1}^{n+1} a_i b_i = 1} \left\{ \max_{x\in[-1,1]} \left| \sum_{i=1}^{n+1} a_i f_i(x) \right| \max_{y\in[-1,1]} \left| \sum_{i=1}^{n+1} b_i g_i(y) \right| \right\}. \qquad (122)$$

There exist $\overline{\boldsymbol{a}}$ and $\overline{\boldsymbol{b}}$ that achieve this infimum. With this definition of σ,

$$\begin{aligned} d_n(K(M_{TV}); C([-1,1])) &= d^n(K(M_{TV}); C([-1,1])) && (123)\\ &= \delta_n(K(M_{TV}); C([-1,1])) && (124)\\ &= \sigma. && (125) \end{aligned}$$

Proof. This is a special case of a central result of Micchelli & Pinkus (1979) (where it appears as Theorem 2.1). A slightly different form of the general result is also provided in Pinkus (1985a) (as Theorem 5.2 on pages 188 and 189 in Section 5 of Chapter V). □

The following corollary provides a useful and concrete characterization of an optimal linear operator in the setting of Theorem 7.1. For a more general but less concrete result in the same direction, see the latter portion of Theorem 5.2 on pages 188 and 189 in Section 5 of Chapter V in Pinkus (1985a).

Corollary 7.1. *If $K(x,y) \in C([-1,1]\times[-1,1])$ is of rank $n+1$, then an optimal rank-n linear operator for $\delta_n(K(M_{TV});C([-1,1]))$ is*

$$\begin{aligned}
P_n^*(x,y) &= K(x,y) - \left(\sum_{i=1}^{n+1} \overline{a}_i f_i(x)\right)\left(\sum_{i=1}^{n+1} \overline{b}_i g_i(y)\right) &(126)\\
&= \sum_{i=1}^{n+1} f_i(x) g_i(y) - \left(\sum_{i=1}^{n+1} \overline{a}_i f_i(x)\right)\left(\sum_{i=1}^{n+1} \overline{b}_i g_i(y)\right) &(127)\\
&= \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_{n+1}(x) \end{pmatrix}^T \left(I_{n+1} - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T\right) \begin{pmatrix} g_1(y) \\ g_2(y) \\ \vdots \\ g_{n+1}(y) \end{pmatrix}, &(128)
\end{aligned}$$

where $\overline{\boldsymbol{a}}$ and $\overline{\boldsymbol{b}}$ are $(n+1)\times 1$ column vectors achieving the infimum of (122) (as in Theorem 7.1) and I_{n+1} is the $(n+1)\times(n+1)$ identity matrix.

Proof. To prove that P_n^* is an optimal linear operator for $\delta_n(K(M_{TV});C([-1,1]))$, I must show that its error achieves the bound σ (as defined in Theorem 7.1) and that its rank is n.

First, I show that the worst-case error in the approximation of $K(x,y)$ by P_n^* is σ:

$$\begin{aligned}
&\sup_{h\in M_{TV}:\|h\|_{TV}\le 1} \max_{y\in[-1,1]} \left| \int_{-1}^{1} K(x,y)\,dh(x) - \int_{-1}^{1} P_n^*(x,y)\,dh(x)\right| \\
= &\max_{x,y\in[-1,1]} |K(x,y) - P_n^*(x,y)| &(129)\\
= &\max_{x,y\in[-1,1]} \left| \left(\sum_{i=1}^{n+1} \overline{a}_i f_i(x)\right)\left(\sum_{i=1}^{n+1} \overline{b}_i g_i(y)\right)\right| &(130)\\
= &\max_{x\in[-1,1]} \left|\sum_{i=1}^{n+1} \overline{a}_i f_i(x)\right| \max_{y\in[-1,1]} \left|\sum_{i=1}^{n+1} \overline{b}_i g_i(y)\right| &(131)\\
= &\ \sigma, &(132)
\end{aligned}$$

where the first equality holds because the supremum over h is achieved by a measure which places mass of 1 or -1 on the $x \in [-1,1]$ that generates the largest possible value of $\max_{y\in[-1,1]} |K(x,y) - P_n^*(x,y)|$, the second equality is by (126), the third equality recognizes that $\max_{x,y} |p(x)q(y)| = \max_x |p(x)| \max_y |q(y)|$, and the fourth equality is by (122), since $\overline{\boldsymbol{a}}$ and $\overline{\boldsymbol{b}}$ achieve the infimum in (122).

Now I demonstrate that the rank of P_n^* is n by examining $I_{n+1} - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T$. Let $\boldsymbol{u}_1 \equiv \overline{\boldsymbol{a}}/\sqrt{\sum_{i=1}^{n+1} \overline{a}_i^2}$, so that $\sqrt{\sum_{i=1}^{n+1} u_{1i}} = 1$. Then

$$\begin{aligned}
\left(I_{n+1} - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T\right)\boldsymbol{u}_1 &= I_{n+1}\boldsymbol{u}_1 - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T\boldsymbol{u}_1 &(133)\\
&= \boldsymbol{u}_1 - \frac{1}{\sqrt{\sum_{i=1}^{n+1} \overline{a}_i^2}} \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T\overline{\boldsymbol{a}} &(134)
\end{aligned}$$

$$\begin{aligned} &= \boldsymbol{u}_1 - \frac{1}{\sqrt{\sum_{i=1}^{n+1} \overline{a}_i^2}} \overline{\boldsymbol{a}} && (135) \\ &= \boldsymbol{u}_1 - \boldsymbol{u}_1 && (136) \\ &= 0, && (137) \end{aligned}$$

where the first equality is by the distributive property of matrix multiplication, the second equality is by the fact that $I_{n+1}\boldsymbol{x} = \boldsymbol{x}$ for any column vector $\boldsymbol{x}$ of length $n+1$ and by the definition of $\boldsymbol{u}_1$, the third equality is by the fact that $\overline{\boldsymbol{b}}^T \overline{\boldsymbol{a}} = \sum_{i=1}^{n+1} \overline{a}_i \overline{b}_i = 1$, the fourth equality is by the definition of $\boldsymbol{u}_1$, and the fifth equality is obvious. This implies that $\boldsymbol{u}_1$ is an eigenvector of $I_{n+1} - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T$ with eigenvalue zero, so $I_{n+1} - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T$ is of rank no greater than n. There are certainly n linearly independent vectors $\boldsymbol{u}_2, \ldots, \boldsymbol{u}_{n+1}$ in $\mathbb{R}^{n+1}$ which are orthogonal to $\overline{\boldsymbol{b}}$ and which have unit Euclidean norm (use the Gram-Schmidt process, for example). Each of these $\boldsymbol{u}_2, \ldots, \boldsymbol{u}_{n+1}$ is an eigenvector of $I_{n+1} - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T$ with eigenvalue one:

$$\begin{aligned} &\text{For } j \geq 2, \\ \left(I_{n+1} - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T\right) \boldsymbol{u}_j &= I_{n+1}\boldsymbol{u}_j - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T \boldsymbol{u}_j && (138) \\ &= \boldsymbol{u}_j - \overline{\boldsymbol{a}} \times 0 && (139) \\ &= \boldsymbol{u}_j. && (140) \end{aligned}$$

Thus, $I_{n+1} - \overline{\boldsymbol{a}}\overline{\boldsymbol{b}}^T$ is of rank n (its range is spanned by the n vectors $\boldsymbol{u}_2, \ldots, \boldsymbol{u}_{n+1}$). $\square$

I now analyze the following kernel of rank $n+1$:

$$K_c^{(n+1)}(x, y) \equiv \exp\left(cx\boldsymbol{\rho}^\dagger\right)^T \left(\exp\left(c\boldsymbol{\rho}^\dagger \boldsymbol{\rho}^{\dagger T}\right)\right)^{-1} \exp\left(c\boldsymbol{\rho}^\dagger y\right), \quad (141)$$

where

$$\begin{aligned} \exp\left(cx\boldsymbol{\rho}^\dagger\right) &= \begin{pmatrix} \exp\left(cx\rho_1^\dagger\right) \\ \vdots \\ \exp\left(cx\rho_{n+1}^\dagger\right) \end{pmatrix} && (142) \\ \exp\left(c\boldsymbol{\rho}^\dagger y\right) &= \begin{pmatrix} \exp\left(c\rho_1^\dagger y\right) \\ \vdots \\ \exp\left(c\rho_{n+1}^\dagger y\right) \end{pmatrix} && (143) \\ \exp\left(c\boldsymbol{\rho}^\dagger \boldsymbol{\rho}^{\dagger T}\right) &= \begin{pmatrix} \exp\left(c\rho_1^\dagger \rho_1^\dagger\right) & \cdots & \exp\left(c\rho_1^\dagger \rho_{n+1}^\dagger\right) \\ \vdots & \ddots & \vdots \\ \exp\left(c\rho_{n+1}^\dagger \rho_1^\dagger\right) & \cdots & \exp\left(c\rho_{n+1}^\dagger \rho_{n+1}^\dagger\right) \end{pmatrix}, && (144) \end{aligned}$$

$x, y \in [-1, 1]$ and $\boldsymbol{\rho}^\dagger$ is the vector produced by Method 6.2. This can be put into the "standard form" showing that it is of rank $n+1$:

$$K_c^{(n+1)}(x, y) = \sum_{i=1}^{n+1} f_i(x) f_i(y), \tag{145}$$

where $x, y \in [-1, 1]$ and the f_i are defined by

$$\begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_{n+1}(x) \end{pmatrix} \equiv A \exp\left(cx\boldsymbol{\rho}^\dagger\right), \tag{146}$$

where

$$A \equiv \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{\lambda_{n+1}}} \end{pmatrix} U^T, \tag{147}$$

with U being the matrix whose columns are the eigenvectors of $\exp\left(c\boldsymbol{\rho}^\dagger \boldsymbol{\rho}^{\dagger T}\right)$ (the leftmost column is the eigenvector corresponding to the largest eigenvalue, the second column from the left is the eigenvector corresponding to the second largest eigenvalue, and so on) and λ_i is the i^{th}-largest eigenvalue of $\exp\left(c\boldsymbol{\rho}^\dagger \boldsymbol{\rho}^{\dagger T}\right)$. The standard spectral decomposition $\exp\left(c\boldsymbol{\rho}^\dagger \boldsymbol{\rho}^{\dagger T}\right) = U\Lambda U^T$, where Λ is a diagonal matrix whose diagonal is $(\lambda_1, \lambda_2, \cdots, \lambda_{n+1})$, shows that the above definition of f_i leads the right-hand sides of (141) and (145) to be equal. Note that the strict total positivity of $\exp\left(c\boldsymbol{\rho}^\dagger \boldsymbol{\rho}^{\dagger T}\right)$ implies that each λ_i is positive.

A third expression for $K_c^{(n+1)}(x, y)$ can be helpful in understanding the difference between the exponential product kernel $\exp(cxy)$ and $K_c^{(n+1)}(x, y)$ for $x, y \in [-1, 1]$:

$$\exp(cxy) - K_c^{(n+1)}(x, y)$$

$$= \frac{\begin{vmatrix} \exp\left(c\rho_1^\dagger\rho_1^\dagger\right) & \cdots & \exp\left(c\rho_1^\dagger\rho_{n+1}^\dagger\right) & \exp\left(c\rho_1^\dagger y\right) \\ \vdots & \cdots & \vdots & \vdots \\ \exp\left(c\rho_{n+1}^\dagger\rho_1^\dagger\right) & \cdots & \exp\left(c\rho_{n+1}^\dagger\rho_{n+1}^\dagger\right) & \exp\left(c\rho_{n+1}^\dagger y\right) \\ \exp\left(cx\rho_1^\dagger\right) & \cdots & \exp\left(cx\rho_{n+1}^\dagger\right) & \exp(cxy) \end{vmatrix}}{\begin{vmatrix} \exp\left(c\rho_1^\dagger\rho_1^\dagger\right) & \cdots & \exp\left(c\rho_1^\dagger\rho_{n+1}^\dagger\right) \\ \vdots & \cdots & \vdots \\ \exp\left(c\rho_{n+1}^\dagger\rho_1^\dagger\right) & \cdots & \exp\left(c\rho_{n+1}^\dagger\rho_{n+1}^\dagger\right) \end{vmatrix}}. \tag{148}$$

The consistency of this expression with (141) follows from Micchelli & Pinkus (1977b), equation (3) and the discussion following it. Pinkus (1985a) also uses a similar expression in the proof of his Theorem 2.6 on pages 146 and 147 in Section 2 of Chapter V. Briefly, use Laplace's formula to expand the determinant in the numerator of (148) along the last row, then do the same for each of the resulting cofactors (with the exception of the last one, which is multiplied by $\exp(cxy)$) along their last columns. As in Micchelli & Pinkus (1977b), this process leads to $\exp(cxy)$ minus a quadratic form in the vectors $\exp\left(cx\boldsymbol{\rho}^{\dagger}\right)$ and $\exp\left(c\boldsymbol{\rho}^{\dagger}y\right)$ whose coefficient matrix is identifiable as $\left(\exp\left(c\boldsymbol{\rho}^{\dagger}\boldsymbol{\rho}^{\dagger T}\right)\right)^{-1}$ via the usual expression of a matrix inverse in terms of cofactors.

I now guess that the solution of the problem (122) for the kernel $K_c^{(n+1)}(x,y)$ of rank $n+1$ is:

$$
\begin{aligned}
\overline{\boldsymbol{a}}_c &= \overline{\boldsymbol{b}}_c && (149)\\
&= \boldsymbol{z}^* && (150)\\
&= \underset{\boldsymbol{z}:\sum_{i=1}^{n+1} z_i^2=1}{\arg\min}\left\{\max_{x\in[-1,1]}\left|\sum_{i=1}^{n+1} z_i f_i(x)\right|\right\}, && (151)
\end{aligned}
$$

where the f_i are as in (146). I further conjecture that the linear combination $f^{LD} \equiv \sum_{i=1}^{n+1} z_i^* f_i(x)$ based on the solution of the problem (151) satisfies the conditions met by u^{LD} in Theorem 5.6, so that it is unique up to sign (and thus $\boldsymbol{z}^*$ is unique up to sign) and there exist $n+1$ points $\left\{x_i^{LD}\right\}_{i=1}^{n+1}$ with $-1 \le x_1^{LD} < x_2^{LD} < \cdots < x_{n+1}^{LD} \le 1$ such that

$$
(-1)^i \epsilon f^{LD}\left(x_i^{LD}\right) = \max_{x\in[-1,1]}\left|f^{LD}(x)\right|, \qquad (152)
$$

where ϵ is either +1 or -1.

Note that (151) is constructed so that the (equal) quantities $\overline{\boldsymbol{a}}_c$ and $\overline{\boldsymbol{b}}_c$ both have unit Euclidean norm, so they satisfy the sum-of-products-equals-one constraint in (122). The ultimate usefulness of this guess in approximating the exponential product operator will be rigorously verified through interval analysis in Subsection 7.2 below.

Remark 7.1. If the matrix A of (147) were totally positive, then the system $f_1(x), \ldots, f_{n+1}(x)$ defined by (146) would inherit the Descartes nature of the system $\exp\left(cx\rho_1^{\dagger}\right), \ldots, \exp\left(cx\rho_{n+1}^{\dagger}\right)$. If this were true, then a straightforward generalization of Theorem 5.6 would apply and (152) would be a result rather than a guess.

However, this matrix is not totally positive, so (152) remains a conjecture, though I do rigorously verify its utility below using interval analysis in Subsection 7.2.

I now consider the rank-one update of the identity $I_{n+1} - \boldsymbol{z}^*\boldsymbol{z}^{*T}$, where $\boldsymbol{z}^*$ comes from (151). Perform an eigenvalue decomposition of $I_{n+1} - \boldsymbol{z}^*\boldsymbol{z}^{*T}$ to get $I_{n+1} - \boldsymbol{z}^*\boldsymbol{z}^{*T} = QDQ^T$, where Q is an orthogonal $(n+1)\times(n+1)$

matrix whose columns are the eigenvectors of $I_{n+1} - \boldsymbol{z}^* \boldsymbol{z}^{*T}$ and D is a diagonal $(n+1) \times (n+1)$ matrix whose diagonal elements are the eigenvalues of $I_{n+1} - \boldsymbol{z}^* \boldsymbol{z}^{*T}$. Permute if necessary so that the smallest diagonal element of D is in the upper left corner, the second-smallest diagonal element of D is in the second diagonal position, and so on; thus, the leftmost column of Q is the eigenvector corresponding to the smallest eigenvalue, the second column of Q (from the left) is the eigenvector corresponding to the second-smallest eigenvalue, and so on through the rightmost column of Q (which corresponds to the largest eigenvalue). Let the diagonal elements of D be labeled $\widetilde{d}_1 \leq \cdots \leq \widetilde{d}_{n+1}$. The interlacing portion of Theorem 1 of Bunch *et al.* (1978) implies that $\widetilde{d}_i = 1$ for $i \geq 2$. But since $\boldsymbol{z}^*$ satisfies $\boldsymbol{z}^{*T} \boldsymbol{z}^* = \sum_{i=1}^{n+1} z_i^2 = 1$, I also have that

$$\begin{aligned} \left(I_{n+1} - \boldsymbol{z}^* \boldsymbol{z}^{*T}\right) \boldsymbol{z}^* &= I_{n+1} \boldsymbol{z}^* - \boldsymbol{z}^* \boldsymbol{z}^{*T} \boldsymbol{z}^* \\ &= \boldsymbol{z}^* - \boldsymbol{z}^* \left(\boldsymbol{z}^{*T} \boldsymbol{z}^*\right) \\ &= \boldsymbol{z}^* - \boldsymbol{z}^* \\ &= 0, \end{aligned}$$

so $\boldsymbol{z}^*$ is an eigenvector of $I_{n+1} - \boldsymbol{z}^* \boldsymbol{z}^{*T}$ with eigenvalue zero, and $\widetilde{d}_1 = 0$. Thus, $I_{n+1} - \boldsymbol{z}^* \boldsymbol{z}^{*T}$ is of rank n and I have the decomposition

$$I_{n+1} - \boldsymbol{z}^* \boldsymbol{z}^{*T} = \widetilde{Q} \widetilde{Q}^T, \tag{153}$$

where $\widetilde{Q}$ is the $(n+1) \times n$ matrix formed from Q by omitting the first column of Q (the first column of Q is the eigenvector of $I_{n+1} - \boldsymbol{z}^* \boldsymbol{z}^{*T}$ that corresponds to the smallest eigenvalue $\widetilde{d}_1$, which I have shown is zero).

Assuming that the guess (151) is correct, substituting (141), (145), (146), and (151) into (128) of Corollary 7.1 leads to the following expressions for an optimal linear operator for $\delta_n \left(K_c^{(n+1)} \left(M_{TV}\right); C\left([-1, 1]\right) \right)$:

$$\begin{aligned} & P_{c,n}^* (x, y) \\ = \; & \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_{n+1}(x) \end{pmatrix}^T \left(I_{n+1} - \boldsymbol{z}^* \boldsymbol{z}^{*T}\right) \begin{pmatrix} f_1(y) \\ f_2(y) \\ \vdots \\ f_{n+1}(y) \end{pmatrix} && (154) \\ = \; & \exp\left(cx\boldsymbol{\rho}^\dagger\right)^T A^T \left(I_{n+1} - \boldsymbol{z}^* \boldsymbol{z}^{*T}\right) A \exp\left(c\boldsymbol{\rho}^\dagger y\right) && (155) \\ = \; & \exp\left(cx\boldsymbol{\rho}^\dagger\right)^T V^* V^{*T} \exp\left(c\boldsymbol{\rho}^\dagger y\right) && (156) \\ = \; & K_c^{(n+1)}(x, y) - \exp\left(cx\boldsymbol{\rho}^\dagger\right)^T A^T \boldsymbol{z}^* \boldsymbol{z}^{*T} A \exp\left(c\boldsymbol{\rho}^\dagger y\right), && (157) \end{aligned}$$

where the $(n+1) \times (n+1)$ matrix A is defined by (147), the $(n+1) \times n$ matrix V^* is defined as

$$V^* \equiv A^T \widetilde{Q}, \tag{158}$$

and the $(n+1)\times n$ matrix $\widetilde{Q}$ satisfies (153). See the discussion below (153) for how to compute $\widetilde{Q}$.

Since V^* is $(n+1)\times n$, it is clear that $P^*_{c,n}(x,y)$ has rank n. To make this more concrete, I define the system of functions

$$\begin{pmatrix} p^*_1(x) \\ p^*_2(x) \\ \vdots \\ p^*_n(x) \end{pmatrix} \equiv V^{*T}\exp\left(cx\boldsymbol{\rho}^\dagger\right), \tag{159}$$

Then

$$P^*_{c,n}(x,y) = \sum_{i=1}^{n} p^*_i(x)\, p^*_i(y). \tag{160}$$

Remark 7.2. The rank-n operator

$$P^*_{c,n}(x,y) = \exp\left(cx\boldsymbol{\rho}^\dagger\right)^T V^* V^{*T}\exp\left(c\boldsymbol{\rho}^\dagger y\right) \tag{161}$$

given in (156) above, where V^* is defined by (158), is **numerically optimal** for the linear n-width $\delta_n\left(P_c\left(M_{TV}\right); C\left([-1,1]\right)\right)$ of the exponential product operator with parameter c (for each c and n I have examined): its worst-case approximation error is within rounding error of the lower bound computed by Method 6.2. This will be demonstrated rigorously using interval analysis in Subsection 7.2 below.

In light of (157) and (148), the approximation error $\exp(cxy) - P^*_{c,n}(x,y)$ can be expressed as

$$\begin{aligned}
& \exp(cxy) - P^*_{c,n}(x,y) \\
= \ & \exp(cxy) - K_c^{(n+1)}(x,y) + \exp\left(cx\boldsymbol{\rho}^\dagger\right)^T A^T \boldsymbol{z}^*\boldsymbol{z}^{*T} A \exp\left(c\boldsymbol{\rho}^\dagger y\right) \quad (162) \\
= \ & \frac{\begin{vmatrix} \exp\left(c\rho_1^\dagger\rho_1^\dagger\right) & \cdots & \exp\left(c\rho_1^\dagger\rho_{n+1}^\dagger\right) & \exp\left(c\rho_1^\dagger y\right) \\ \vdots & \cdots & \vdots & \vdots \\ \exp\left(c\rho_{n+1}^\dagger\rho_1^\dagger\right) & \cdots & \exp\left(c\rho_{n+1}^\dagger\rho_{n+1}^\dagger\right) & \exp\left(c\rho_{n+1}^\dagger y\right) \\ \exp\left(cx\rho_1^\dagger\right) & \cdots & \exp\left(cx\rho_{n+1}^\dagger\right) & \exp(cxy) \end{vmatrix}}{\begin{vmatrix} \exp\left(c\rho_1^\dagger\rho_1^\dagger\right) & \cdots & \exp\left(c\rho_1^\dagger\rho_{n+1}^\dagger\right) \\ \vdots & \cdots & \vdots \\ \exp\left(c\rho_{n+1}^\dagger\rho_1^\dagger\right) & \cdots & \exp\left(c\rho_{n+1}^\dagger\rho_{n+1}^\dagger\right) \end{vmatrix}} \\
+ \ & \exp\left(cx\boldsymbol{\rho}^\dagger\right)^T A^T \boldsymbol{z}^*\boldsymbol{z}^{*T} A \exp\left(c\boldsymbol{\rho}^\dagger y\right). \quad (163)
\end{aligned}$$

This expression is useful because the first (ratio-of-determinants) term in (163) vanishes if either $x = \rho_i^\dagger$ or $y = \rho_j^\dagger$ (or both) for some $i \in \{1,\ldots,n+1\}$ or

$j \in \{1, \ldots, n+1\}$ (in these cases, the numerator of the ratio of determinants is zero because the determinant of any matrix which has two identical rows, two identical columns, or both must be zero). The constancy at zero of the first (ratio-of-determinants) term in (163) if either $x = \rho_i^\dagger$ or $y = \rho_j^\dagger$ (or both) for some $i \in \{1, \ldots, n+1\}$ or $j \in \{1, \ldots, n+1\}$ means that the gradient of this first (ratio-of-determinants) term must vanish at any pair of coordinates $\left(\rho_i^\dagger, \rho_j^\dagger\right) \in [-1,1]$, $i, j \in \{1, \ldots, n+1\}$. (More detailed consideration of the ratio-of-determinants term shows that each such pair of coordinates is a saddle point of this term.) I revisit the expression (163) in Subsection 7.2, where it is very useful in bounding the error incurred when approximating $\exp(cxy)$ by $P_{c,n}^*(x,y)$.

I now provide a method which implements the construction of $P_{c,n}^*(x,y)$ that is described above.

Method 7.1. The following method builds the numerically-optimal rank-n approximating operator $P_{c,n}^*(x,y)$ of (156) by using results produced by Method 6.2.

- **Given**:
  1. $n+1$, the number of different ρ_i used in the rank-n approximation
  2. $c > 0$
  3. a tolerance
  4. a maximum allowable number of iterations
- **Output**:
  1. $\boldsymbol{\rho}^\dagger \in \Omega_{n+1}$, the output of Method 6.2
  2. a vector $\boldsymbol{z}^*$ of coefficients such that $\sum_{i=1}^{n+1} z_i^{*2} = 1$
  3. a $(n+1) \times n$ matrix V^* such that $P_{c,n}^*(x,y) = \exp\left(cx\boldsymbol{\rho}^\dagger\right)^T V^* V^{*T} \exp\left(c\boldsymbol{\rho}^\dagger y\right)$.
  4. a scalar discrepancy representing the change in $B_c(\boldsymbol{\rho})$ at the last iteration prior to return of Method 6.2
  5. the number of iterations performed in Method 6.2

Step 1 Check the arguments.

Step 2 Call Method 6.2 with the inputs given; label the $\boldsymbol{\rho}$ output of this call $\boldsymbol{\rho}^\dagger$.

Step 3 Compute the eigenvalues and eigenvectors of $\exp\left(c\boldsymbol{\rho}^\dagger \boldsymbol{\rho}^{\dagger T}\right)$. Store the eigenvectors as the columns of a $(n+1) \times (n+1)$ matrix U, with the eigenvector of the largest eigenvalue being the leftmost column, the eigenvector of the second largest eigenvalue being the second column from the left, and so on through the rightmost column, which is the eigenvector of the smallest eigenvalue. Label the eigenvalues $\lambda_1, \ldots, \lambda_{n+1}$ from largest to smallest. Because $\exp\left(c\boldsymbol{\rho}^\dagger \boldsymbol{\rho}^{\dagger T}\right)$ is a real symmetric matrix, U and $(\lambda_1, \ldots, \lambda_{n+1})$ must be real also if all calculations are performed in exact arithmetic. When n is large enough relative to c that Method 6.2 produces a lower bound which is on the order of machine epsilon, the matrix $\exp\left(c\boldsymbol{\rho}^{\dagger}\boldsymbol{\rho}^{\dagger T}\right)$ typically has a very large condition number. To avoid corruption by rounding error in these situations, let U and $(\lambda_1, \ldots, \lambda_{n+1})$ be equal to only the real parts of the eigenvectors and eigenvalues (respectively) produced by a standard eigenvalue method.

Step 4 Construct the vector $\boldsymbol{\gamma}$ so that γ_i is λ_i when λ_i is at least machine epsilon and otherwise is machine epsilon. This avoids numerical difficulties in the steps to follow. Unless the original matrix $\exp\left(c\boldsymbol{\rho}^{\dagger}\boldsymbol{\rho}^{\dagger T}\right)$ is very poorly conditioned (which typically occurs when n is large enough relative to c that Method 6.2 produces a lower bound which is on the order of machine epsilon), $\boldsymbol{\gamma}$ will be $\boldsymbol{\lambda}$.

Step 5 Compute the matrix

$$A \equiv \begin{pmatrix} \frac{1}{\sqrt{\gamma_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{\gamma_2}} & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{\gamma_{n+1}}} \end{pmatrix} U^T. \tag{164}$$

This is just a version of (147) that makes numerical allowances for very small λ_i.

Step 6 Initialize the grid for root-finding to the Chebyshev points for $m = 200$, see Step 8 for the construction of m Chebyshev points.

Step 7 Check the arguments to be sure that dimensions and signs are appropriate.

Step 8 Make an initial guess that the nodes are Chebyshev points (the $n+1$ points at which the $(n+1)^{\text{st}}$ Chebyshev polynomial attains its maximum absolute value on $[-1,1]$):

$$(y_1, y_2, y_3, \ldots, y_{n+1}) = \left(\sin\left(-\frac{\pi}{2}\right), \sin\left(-\frac{\pi}{2}+s\right), \sin\left(-\frac{\pi}{2}+2s\right), \ldots, \sin\left(\frac{\pi}{2}\right)\right),$$

where $s = \pi/n$.

Step 9 Form the matrix $G = A\exp\left(c\boldsymbol{\rho}\boldsymbol{y}^T\right)$, where $\boldsymbol{y}$ is the column vector of the guesses from Step 8.

Step 10 Form the alternating row vector $\boldsymbol{b}$ of length $n+1$ such that $b_i = (-1)^i$.

Step 11 Use the Moore-Penrose pseudoinverse of G to solve the $(n+1)\times(n+1)$ system of linear equations $\boldsymbol{w}G = \boldsymbol{b}$ for the row vector $\boldsymbol{w}$ of coefficients (which has $n+1$ elements). Thus, $\boldsymbol{w} = \boldsymbol{b}G^{+}$, where G^{+} is the Moore-Penrose pseudoinverse of G. It is important to use the Moore-Penrose pseudoinverse in this step because G is nonsingular by construction but may be very poorly conditioned (this typically occurs when n is large enough relative to c that Method 6.2 produces a lower bound which is on the order of machine epsilon). The computation of the Moore-Penrose pseudoinverse can be accomplished via the singular value decomposition; see Golub & Van Loan (2013), page 290 in Section 5 of Chapter 5, for a discussion of this calculation.

Step 12 Set $\boldsymbol{z} = \boldsymbol{w} / \sum_{i=1}^{n+1} w_i^2$, so that $\sum_{i=1}^{n+1} z_i^2 = 1$.

Step 13 Compute the row vector of errors $\boldsymbol{\eta} = \boldsymbol{z}G$. By the definition of $\boldsymbol{b}$ and the construction of $\boldsymbol{w}$ and of $\boldsymbol{z}$, $\eta_i = (-1)^i / \sum_{i=1}^{n+1} w_i^2$.

Step 14 Set the discrepancy to the given tolerance plus one.

Step 15 Set the number of iterations to zero.

Step 16 If the discrepancy is greater than the given tolerance and the number of iterations is less than or equal to the given maximum allowable number of iterations, proceed. Otherwise, compute the eigenvalues and eigenvectors of the matrix $I_{n+1} - \boldsymbol{z}\boldsymbol{z}^T$; label the eigenvalues $d_1, \ldots, d_{n+1}$ from smallest to largest and form the matrix Q of eigenvectors (each column of Q is an eigenvector, and the eigenvector associated with the smallest eigenvalue d_1 is the leftmost column of Q, the eigenvector associated with d_2 is the second column from the left, and so on). Let $\widetilde{Q}$ be the $(n+1) \times n$ matrix formed from Q by dropping the leftmost column of Q (the one corresponding to the smallest eigenvalue d_1) and set $V^* = A^T \widetilde{Q}$. Set $\boldsymbol{z}^* = \boldsymbol{z}$ and return the outputs as described above.

Step 17 Compute the maximum of the absolute values of the entries of $\boldsymbol{\eta}$, call it $\max_{\text{old}}$.

Step 18 Find all of the zeros of $h(y) \equiv \sum_{i=1}^{n+1} a_i \exp(c\rho_i y)$ on $[-1, 1]$. The fact that $\exp(c\rho_i y)$ form a Chebyshev system on $[-1, 1]$ implies that h can have at most n zeros, and h has at least n zeros because $\boldsymbol{a}G$ is a vector of length $n+1$ that represents the evaluation of g at the $n+1$ points y_j, and $\boldsymbol{a}G$ alternates sign between each of its $n+1$ entries, so h has exactly n zeros. Thus, I can search for the zeros using a fine grid; if less than n zeros are found, the grid must be refined and the search performed again. If the grid size becomes implausibly large, generate an error message and stop: this is evidence of numerical instability in the computation of $\boldsymbol{a}$. After finding at least n zeros, if the number of zeros found is too large ($n+1$ or greater) this is evidence that the function $h(y)$ may be within numerical tolerance of zero on a fine grid. If so, let $(\widehat{y}_1, \ldots, \widehat{y}_{n+1})$ be the vector $\boldsymbol{\rho}$ and go directly to Step 20. If the number of zeros found is too large ($n+1$ or greater) and the function $h(y)$ is not within numerical tolerance of zero on a fine grid, this is evidence of numerical instability: generate an error message and stop. Even if the number of zeros found is exactly n, if the function is within numerical tolerance of zero on a fine grid, let $(\widehat{y}_1, \ldots, \widehat{y}_{n+1})$ be the vector $\boldsymbol{\rho}$ and go directly to Step 20.

Step 19 Find all of the relative optima of $h(y)$ by searching between the n zeros (and between -1 and the smallest zero as well as between the largest zero and 1). There will be $n+1$ relative optima found by this procedure, since there are $n+1$ segments defined by the n zeros and the interval endpoints -1 and 1; assemble them into a vector $(\widehat{y}_1, \ldots, \widehat{y}_{n+1})$.

Step 20 Let $(y_1, \ldots, y_{n+1}) = (\widehat{y}_1, \ldots, \widehat{y}_{n+1})$, form a new matrix G using the new y_j and, using the same definition of $\boldsymbol{b}$, use the Moore-Penrose inverse of the new G to solve for a new vector $\boldsymbol{w}$ as above and use the new $\boldsymbol{w}$ and the new G to obtain new vectors $\boldsymbol{z}$ and $\boldsymbol{\eta}$.

Step 21 Set the discrepancy to $\max_{i\in\{1,\ldots,n+1\}} |\eta_i| - \max_{\text{old}}$. As a safeguard against rounding errors, if this results in a negative discrepancy set the discrepancy to zero.

Step 22 Increase the number of iterations by one.

Step 23 Go to Step 16.

Figure 3 and Figure 4 show the error surfaces generated by approximating the exponential product operator using the numerically-optimal approximation. The "fuzzy" or "jagged" appearance of the bottom left panel of Figure 3 arises because the approximation error in that panel has been reduced to the point at which it is being impacted by rounding error (note the scale of the vertical axis in that panel).

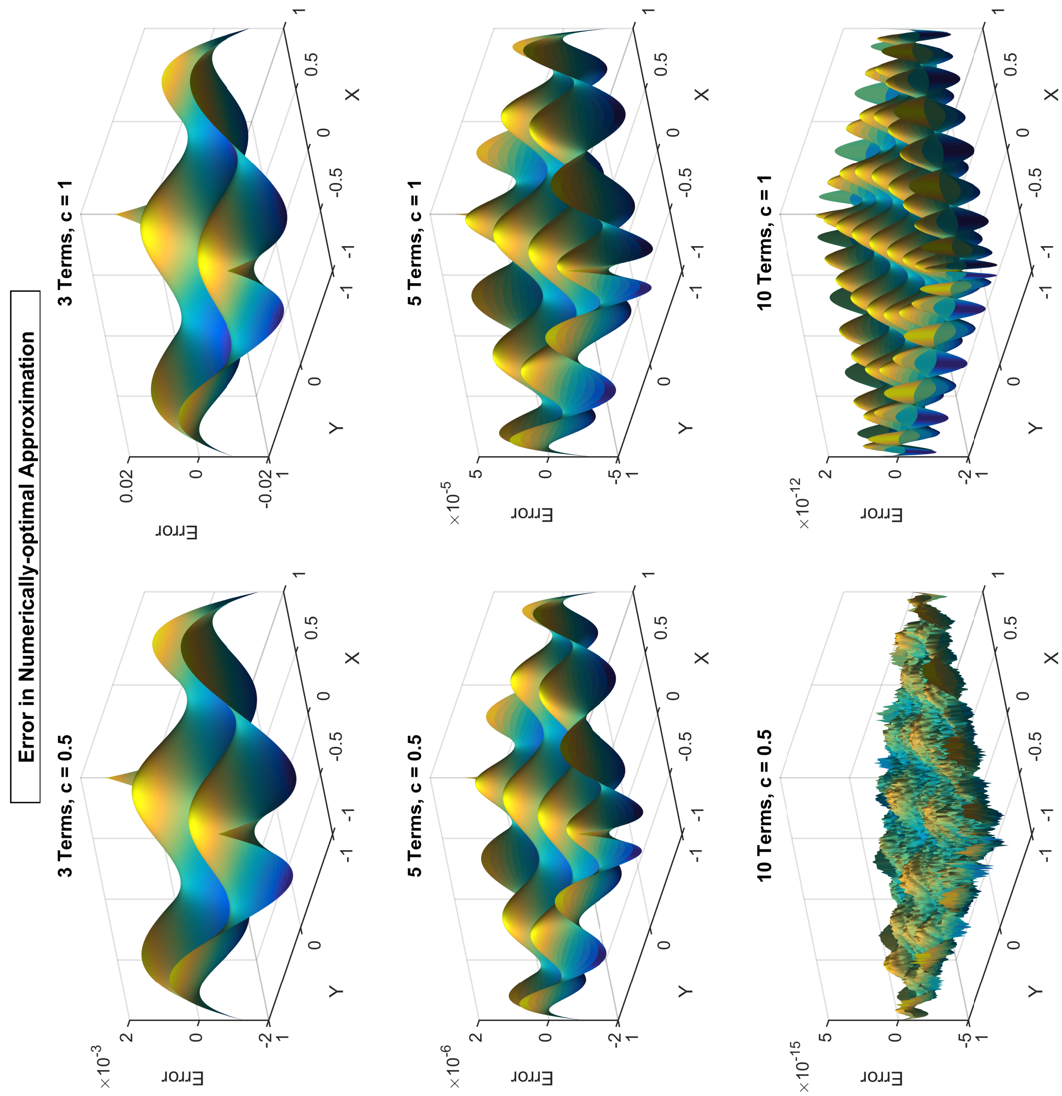

Figure 3: Error Surfaces in the Numerically-optimal Approximation to exp (cxy)

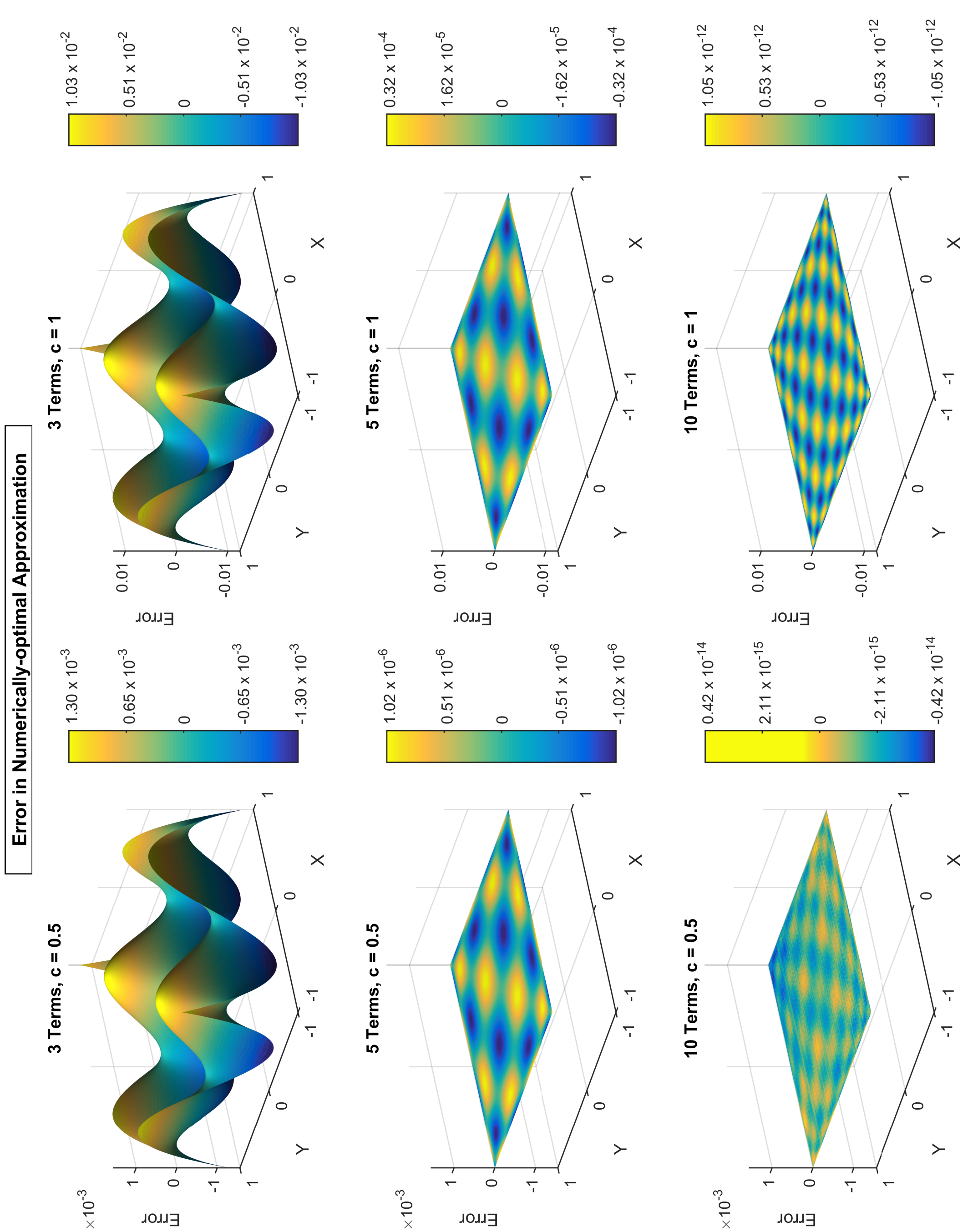

Figure 4: Error Surfaces in the Numerically-optimal Approximation to exp (cxy)

7.1 The Interpretation of the Numerically-optimal Approximation

The numerically-optimal approximation $P^*_{c,n}(x,y)$ can be interpreted as giving a (linear) calculation that provides an approximation, across scenarios, of any given stream of cashflows. Because the numerically-optimal approximation $P^*_{c,n}(x,y)$ is of the form (156), the rank-n approximation it provides for any given cashflow stream and scenario set is a portfolio of n different cashflow streams, each having cashflows at the same $n+1$ times. In approximating the values of different cashflow streams across the same scenario set, the rank-n numerically-optimal approximation uses the same n approximating cashflow streams but weights them differently to produce distinct approximating portfolios. Because the approximating portfolio weights are linear in the cashflow stream being approximated, the approximating weights for a portfolio of cashflow streams are the sums of the approximating weights for the individual cashflow streams in the portfolio.

Please see Subsection 11.1 for a simple example that illustrates these properties.

The numerically-optimal choice of $n+1$ approximating cashflow dates is directly linked to the vector $\boldsymbol{\rho}^\dagger$, while the choices of specific cashflow streams in the approximating portfolio and of the approximating portfolio's holdings of each cashflow stream are determined by the particular form of (156). Choosing any of these aspects of the approximating portfolio differently would correspond to a different approximation to the underlying exponential product operator; a different approximation would not, in general, be optimal and could perform very poorly.

To demonstrate the mechanics of the interpretation, combine (156) with (32) from Section 3 to obtain the numerically-optimal approximation that I use in practice. Note that $\tilde{\boldsymbol{g}}^T$ is just a vector version of the notation of (32) and let $\boldsymbol{x}$ be the normalized products of shift direction S_i and time t_i, so that $S_i t_i = m_{St} + r_{St} x_i$ with $x_i \in [-1,1]$. Similarly, consider y as the normalized version of $z = m_y + r_y y$ (so that $y \in [-1,1]$) and recall that $c = r_y r_{St}$. Then my approach to approximation substitutes (156) into (32) in place of the exponential product kernel to get the approximation:

$$\begin{aligned}
& \tilde{\boldsymbol{g}}^T \exp\left(c\boldsymbol{x}\boldsymbol{\rho}^{\dagger T}\right) V^* V^{*T} \exp\left(c\boldsymbol{\rho}^\dagger y\right) \exp\left(m_{St}\left(y r_y + m_y\right)\right) && (165)\\
= \; & \boldsymbol{H}_{\text{approx}} V^{*T} \exp\left(c\boldsymbol{\rho}^\dagger y + m_{St}\left(y r_y + m_y\right)\right) && (166)\\
= \; & \boldsymbol{H}_{\text{approx}} V^{*T} \exp\left(r_{St}\boldsymbol{\rho}^\dagger r_y y + m_{St}\left(y r_y + m_y\right)\right) && (167)\\
= \; & \boldsymbol{H}_{\text{approx}} V^{*T} \exp\left(\left(m_{St} + r_{St}\boldsymbol{\rho}^\dagger\right) r_y y + m_{St} m_y\right) && (168)\\
= \; & \boldsymbol{H}_{\text{approx}} V^{*T} \exp\left(-m_y r_{St}\boldsymbol{\rho}^\dagger + \left(m_{St} + r_{St}\boldsymbol{\rho}^\dagger\right)\left(m_y + r_y y\right)\right) && (169)\\
= \; & \boldsymbol{H}_{\text{approx}} F_{\text{approx}} \exp\left(\boldsymbol{S}^{\text{approx}} \circ \boldsymbol{t}^{\text{approx}} z\right), && (170)
\end{aligned}$$

where $\boldsymbol{H}_{\text{approx}}$ is the $1 \times n$ row vector of approximating portfolio holdings, F_{approx} is the $n \times (n+1)$ matrix expressing the approximating cashflow streams (its (i,j) element is the present value of approximating cashflow j held in approximating stream i), and $\boldsymbol{S}^{\text{approx}} \circ \boldsymbol{t}^{\text{approx}}$ is the elementwise product of a vector of times $\boldsymbol{t}^{\text{approx}}$ that represents the times at which the approximating cashflows occur (there are $n+1$ of them) and a shift direction vector $\boldsymbol{S}^{\text{approx}}$. The variable z simply represents the multiple of the shift direction (without any normalization). To obtain this expression, let $\boldsymbol{H}_{\text{approx}} \equiv \tilde{\boldsymbol{g}}^T \exp\left(c\boldsymbol{x}\boldsymbol{\rho}^\dagger\right)^T V^*$, $F_{\text{approx}} \equiv V^{*T} \exp\left(-m_y r_{St} \boldsymbol{\rho}^\dagger\right)$, and $z \equiv m_y + r_y y$ (removing the normalization of the normalized shift multiplier y). To separate $m_{St} + r_{St}\boldsymbol{\rho}^\dagger$ into $\boldsymbol{S}^{\text{approx}}$ and $\boldsymbol{t}^{\text{approx}}$ use a zero-finder on an interpolation capturing the relationship between the vector of times $\boldsymbol{t}$ and its elementwise product with the shift direction vector, $\boldsymbol{S} \circ \boldsymbol{t}$, in the original data. This will provide a vector of $\boldsymbol{t}^{\text{approx}}$ at which the interpolation delivers $m_{St} + r_{St}\boldsymbol{\rho}^\dagger$ as the (interpolated) product $\boldsymbol{S}^{\text{approx}} \circ \boldsymbol{t}^{\text{approx}}$.

The approximation $P^*_{c,n}(x, y)$ does not just reduce dimension numerically-optimally in worst-case error: the n-dimensional approximating space it provides is spanned by n clearly-defined, simple financial instruments. Each of these instruments is a cashflow stream, and all n of these streams have their cashflows on a common set of times, the $n+1$ elements of the vector $\boldsymbol{t}^{\text{approx}}$. The linearity of aggregation follows immediately from the linearity of $\boldsymbol{H}_{\text{approx}}$ in $\tilde{\boldsymbol{g}}^T$.

7.2 Interval Analysis for Global Optimization of Approximation Error

I now provide a method that uses interval analysis to verify, for a given c and n, that Method 7.1 produces a rank-n approximation

$$P^*_{c,n}(x, y) = \exp\left(cx\boldsymbol{\rho}^\dagger\right)^T V^* V^{*T} \exp\left(c\boldsymbol{\rho}^\dagger y\right) \tag{171}$$

whose worst-case approximation error $\max_{x,y\in[-1,1]} \left|\exp(cxy) - P^*_{c,n}(x, y)\right|$ is within rounding error of the lower bound $B_c\left(\boldsymbol{\rho}^\dagger\right)$ computed by Method 6.2. Hansen & Walster (2004) is a helpful general reference on interval analysis for global optimization. For more on interval analysis, with applications to optimization, the early work of Moore (1979) (including, interestingly, a simple finance application in Chapter 9), the text of Moore *et al.* (2009), and the applications-focused monograph of Jaulin *et al.* (2001) could be consulted. A very nice application is discussed in Sections 3, 5, and 6 of Chapter 4 in Bornemann *et al.* (2004).

To perform the interval analysis described below, I implemented some standard interval operations (these interval operations are described at length in the texts mentioned above). For a scalar interval $I = [l, u]$,

$$\begin{aligned} \inf(I) &= l && (172)\\ \sup(I) &= u && (173)\\ \operatorname{rad}(I) &= 0.5 \times (u - l) && (174)\\ \operatorname{mid}(I) &= l + \operatorname{rad}(I). && (175) \end{aligned}$$

In many applications of interval analysis, rounding error is a focus; that is not the case here, but my implementation of interval arithmetic still respects the need to round correctly to ensure interval enclosure. To implement rad, I take advantage of IEEE 754 standard's rounding modes for floating point computations and set rounding to be toward $+\infty$ prior to performing the calculation noted above, then set rounding back to the default "to-nearest" mode. Likewise, I set rounding to be toward $+\infty$ prior to performing the calculation for the mid noted above, then set rounding back to the default "to-nearest" mode. Each of the operations described above has an elementwise matrix (and thus vector) analog. Each of these operations is very much standard in interval analysis research and practice.

Given two scalar intervals $I_1 = [l_1, u_1]$ and $I_2 = [l_2, u_2]$, I implemented the following binary operations:

$$I_1 + I_2 = [l_1 + l_2, u_1 + u_2] \tag{176}$$

$$I_1 - I_2 = [l_1 - u_2, u_1 - l_2] \tag{177}$$

$$I_1 \times I_2 = [\min(l_1 l_2, l_1 u_2, u_1 l_2, u_1 u_2), \max(l_1 l_2, l_1 u_2, u_1 l_2, u_1 u_2)]\,. \tag{178}$$

In each case, I exploit the rounding modes of the IEEE 754 standard: I set rounding to be toward $-\infty$, compute the lower bound of the result interval as shown above, set rounding to be toward $+\infty$, compute the upper bound of the result interval as shown above, and then set rounding back to the default "to-nearest" mode. All of these operations are ubiquitous in interval analysis research and practice. All of these operations also have straightforward elementwise matrix (and thus vector) analogs.

I also implemented an interval version of matrix multiplication (again, this is standard in interval analysis research and practice, though there are ongoing efforts to tighten the intervals that result from interval matrix multiplication and the approach I implemented is only one of a number of popular approaches for this problem). This is the approach introduced by Rump (1999) and used in his INTLAB package (I use my own interval analysis software here and do not employ INTLAB, but INTLAB is widely known and appreciated by interval analysts). Given two interval matrices M_1 and M_2 (where an interval matrix is just a matrix of intervals), first compute (using $\mathrm{abs}(M)$ to mean the matrix resulting from taking the elementwise absolute value of the matrix M)

$$\underline{C} = \mathrm{mid}(M_1) \times \mathrm{mid}(M_2) \tag{179}$$

Ordinary matrix multiplication, rounding toward $-\infty$.

$$\overline{C} = \mathrm{mid}(M_1) \times \mathrm{mid}(M_2) \tag{180}$$

Ordinary matrix multiplication, rounding toward $+\infty$.

$$C = \underline{C} + 0.5 \times \left(\overline{C} - \underline{C}\right) \tag{181}$$

Rounding toward $+\infty$.

$$S = C - \underline{C} + \mathrm{rad}(M_1) \times (\mathrm{abs}(\mathrm{mid}(M_2)) + \mathrm{rad}(M_2)) + \mathrm{abs}(\mathrm{mid}(M_1)) \times \mathrm{rad}(M_2) \tag{182}$$

Ordinary matrix multiplication, rounding toward $+\infty$.

With these intermediate results in hand, calculate

$$M_I \times M_J \quad = \quad [C - S, C + S], \tag{183}$$

where (as usual) the sum giving the matrix of lower interval bounds is computed with rounding toward $-\infty$ and the sum giving the matrix of upper interval bounds is computed with rounding toward $+\infty$.

As (177) shows, $I - I = [l - u, u - l] \neq [0, 0]$. This phenomenon, in which subtraction widens intervals which one might hope would collapse to $[0, 0]$, is known as *dependence* in the interval analysis literature (see Hansen & Walster (2004), Section 4 of Chapter 2). It arises because the interval subtraction operation treats $I - I$ as though it were actually $I - J$, where J is a separate, independent interval which happens to have the same upper and lower bounds as I (Hansen & Walster (2004) provide a helpful discussion of this on pages 18-19 in Section 4 of Chapter 2).

The dependence phenomenon is especially problematic when using interval analysis to bound approximation errors, since a direct interval implementation (which I confess to having constructed prior to developing the method below) results in interval operations which subtract an interval of values for the approximation (for me, the numerically-optimal approximation $P^*_{c,n}(x, y)$) from an interval of values for the target of approximation (for me, the exponential product operator $\exp(cxy)$). Unless very small intervals for x and y are employed, the intervals being differenced are much wider than the actual approximation error, and the dependence phenomenon leads to very wide bounds unless great computational effort is expended.

The method I provide is especially tailored to the problem of efficiently locating the global maximum or global minimum of the approximation error $\exp(cxy) - P^*_{c,n}(x, y)$. It largely avoids the problems caused by the dependence phenomenon through a very careful selection of the intervals to be used. By choosing sub-boxes on which all of the relevant functions are monotonic in both the x direction and the y direction, my method removes the especially pernicious problem of dependence in the computation of the approximation itself (since $P^*_{c,n}(x, y)$ is computed as the sum of a number of terms, some of which are positive and some of which are negative, interval computing with $P^*_{c,n}(x, y)$ can result in very wide intervals if not approached carefully). By decomposing the approximation error along the lines of (163), my method enables separate treatment of the (typically very small) interpolation component $\exp(cxy) - K_c^{(n+1)}(x, y)$ of the approximation error and the (crucial) rank-one component $\exp\left(cx\boldsymbol{\rho}^\dagger\right)^T A^T \boldsymbol{z}^* \boldsymbol{z}^{*T} A \exp\left(c\boldsymbol{\rho}^\dagger y\right)$ of the approximation error.

In order to handle the interpolation component $\exp(cxy) - K_c^{(n+1)}(x, y)$ of the approximation error efficiently, it is useful to record some of its properties.

Theorem 7.2. *Let $\mathcal{I}(x, y) \equiv \exp(cxy) - K_c^{(n+1)}(x, y)$, the interpolation component of the approximation error exp$(cxy) - P^*_{c,n}(x, y)$. Then:*

1. $\mathcal{I}\left(\rho_i^\dagger, y\right) = 0$ *for any $i = 1, \ldots, n + 1$ and for any $y \in [-1, 1]$.*

2. $\mathcal{I}\left(x, \rho_j^{\dagger}\right) = 0$ *for any* $j = 1, \ldots, n+1$ *and for any* $x \in [-1, 1]$.

3. *On any rectangle of the form* $(x, y) \in \left[\rho_i^{\dagger}, \rho_{i+1}^{\dagger}\right] \times \left[\rho_j^{\dagger}, \rho_{j+1}^{\dagger}\right]$, $\mathcal{I}(x, y)$ *is nonnegative (and positive on the interior of the rectangle) if* $i+j$ *is even. On any such rectangle, it is nonpositive (and negative on the interior of the rectangle) if* $i+j$ *is odd.*

4. *On any rectangle of the form* $(x, y) \in \left[\rho_i^{\dagger}, \rho_{i+1}^{\dagger}\right] \times \left[\rho_j^{\dagger}, \rho_{j+1}^{\dagger}\right]$, $\mathcal{I}(x, y)$ *is a sign regular kernel. On the interior of any such rectangle, it is strictly sign regular.*

5. *On any rectangle of the form* $(x, y) \in \left[\rho_i^{\dagger}, \rho_{i+1}^{\dagger}\right] \times \left[\rho_j^{\dagger}, \rho_{j+1}^{\dagger}\right]$, *there is exactly one* $y^*(x) \in \left(\rho_j^{\dagger}, \rho_{j+1}^{\dagger}\right)$ *for each for each* $x \in \left(\rho_i^{\dagger}, \rho_{i+1}^{\dagger}\right)$ *such that* $\frac{d\mathcal{I}}{dy} = 0$ *when evaluated at* $(x, y^*(x))$. *There is exactly one* $x^*(y) \in \left(\rho_i^{\dagger}, \rho_{i+1}^{\dagger}\right)$ *for each for each* $y \in \left(\rho_j^{\dagger}, \rho_{j+1}^{\dagger}\right)$ *such that* $\frac{d\mathcal{I}}{dx} = 0$ *when evaluated at* $(x^*(y), y)$.

6. *On any rectangle of the form* $(x, y) \in \left[\rho_i^{\dagger}, \rho_{i+1}^{\dagger}\right] \times \left[\rho_j^{\dagger}, \rho_{j+1}^{\dagger}\right]$, *there is exactly one point* (x^{**}, y^{**}) *that maximizes* $|\mathcal{I}(x, y)|$ *over the rectangle.*

Proof. Conclusions 1 and 2 are direct consequences of the determinantal expression (148), since the determinant of any matrix having two identical rows or two identical columns is zero (this was briefly mentioned above).

Conclusion 3 also follows from the determinantal expression (148), since exchanging the places of any two rows or of any two columns of a matrix results in a matrix whose determinant has the opposite sign from the original matrix. Due to the (strict) total positivity of the exponential product kernel $\exp(cxy)$ (see Theorem 5.8), the numerator determinant is positive if neither x nor y is equal to one of the $\rho_k^{\dagger}$ and if the last row (involving x) is moved to be between $\rho_i^{\dagger}$ and $\rho_{i+1}^{\dagger}$, where $x \in \left(\rho_i^{\dagger}, \rho_{i+1}^{\dagger}\right)$ from the statement of the theorem and the assumption that x is not equal to any $\rho_k^{\dagger}$, and if the last column (involving y) is moved to be between $\rho_j^{\dagger}$ and $\rho_{j+1}^{\dagger}$, where $y \in \left(\rho_j^{\dagger}, \rho_{j+1}^{\dagger}\right)$ from the statement of the theorem and the assumption that y is not equal to any $\rho_k^{\dagger}$. It requires $i+j$ exchanges of columns and rows to achieve this placement, so the determinant in the numerator of (148) is $(-1)^{i+j}$ times a positive number; that is, it is positive if $i+j$ is even and negative if $i+j$ is odd. Of course, from the already-proven conclusions 1 and 2, if $x = \rho_k^{\dagger}$ or $y = \rho_k^{\dagger}$ for some $k = 1, \ldots, n+1$, then $\mathcal{I}(x, y)$ is zero.

Conclusion 4 follows from Sylvester's Determinant Identity, which is stated on pages 3-5 in Chapter 0 of Karlin (1968) and on pages 3 and 4 in Section 1 of Chapter 1 of Pinkus (2010). A simplified version of this result is also stated on page 52 in Section 2 of Chapter III of Pinkus (1985a). To see this, consider the minor $\mathcal{I}\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ y_1 & y_2 & \cdots & y_m \end{pmatrix}$ for $\rho_i^\dagger < x_1 < x_2 < \cdots < x_m < \rho_{i+1}^\dagger$ and $\rho_j^\dagger < y_1 < y_2 < \cdots < y_m < \rho_{j+1}^\dagger$. If any such minor is of the same strict sign for all such $x_1, x_2, \ldots, x_m$ and $y_1, y_2, \ldots, y_m$ (but may have different signs for different numbers m of points), then $\mathcal{I}(x,y)$ is strictly sign regular (of course, by conclusions 1 and 2, already proven, if any x_k or y_k is on the boundary of the rectangle $\left[\rho_i^\dagger, \rho_{i+1}^\dagger\right] \times \left[\rho_j^\dagger, \rho_{j+1}^\dagger\right]$, then the minor is zero because an entire row or column is zero). To see that the minor has the desired strict sign regularity, apply Sylvester's Determinant Identity to the expression (148), letting $K(x,y) = \exp(cxy)$. First observe that

$$\mathcal{I}(x,y) = \frac{K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & x \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & y \end{pmatrix}}{K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \end{pmatrix}} \tag{184}$$

(which is just a re-expression of (148) in different notation). Let $\mathcal{J}(x,y)$ be defined by

$$\mathcal{J}(x,y) \equiv K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & x \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & y \end{pmatrix}. \tag{185}$$

Then I have

$$\mathcal{I}\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ y_1 & y_2 & \cdots & y_m \end{pmatrix} = \frac{\mathcal{J}\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ y_1 & y_2 & \cdots & y_m \end{pmatrix}}{K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \end{pmatrix}^m} \tag{186}$$

by the multilinearity of the determinant. Applying Sylvester's Determinant Identity to evaluate $\mathcal{J}\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ y_1 & y_2 & \cdots & y_m \end{pmatrix}$ I get

$$\begin{aligned} & \mathcal{J}\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ y_1 & y_2 & \cdots & y_m \end{pmatrix} \\ = \; & K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \end{pmatrix}^{m-1} \\ & \times K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & x_1 & \cdots & x_m \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & y_1 & \cdots & y_m \end{pmatrix}. \end{aligned} \tag{187}$$

Thus

$$\mathcal{I}\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ y_1 & y_2 & \cdots & y_m \end{pmatrix}$$

$$= \frac{K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & x_1 & \cdots & x_m \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & y_1 & \cdots & y_m \end{pmatrix}}{K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \end{pmatrix}}. \tag{188}$$

The term $K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger \end{pmatrix}$ is always positive by the strict total positivity of $\exp(cxy)$ (see Theorem 5.8). By the same logic, the term

$$K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & x_1 & \cdots & x_m \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & y_1 & \cdots & y_m \end{pmatrix}$$

requires $m(i+j)$ exchanges of rows and columns to arrange it into a minor of $K(x,y)$ which has all arguments sorted in increasing order and is thus positive (recall that $(x_p, y_q) \in \left[\rho_i^\dagger, \rho_{i+1}^\dagger\right] \times \left[\rho_j^\dagger, \rho_{j+1}^\dagger\right]$ for all the x_p and y_q involved in the minor). This implies that

$$K\begin{pmatrix} \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & x_1 & \cdots & x_m \\ \rho_1^\dagger & \rho_2^\dagger & \cdots & \rho_{n+1}^\dagger & y_1 & \cdots & y_m \end{pmatrix}$$

is positive if $i+j$ is even, is negative if $i+j$ is odd and m is odd, and is positive if $i+j$ is odd and m is even. Thus, $\mathcal{I}\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ y_1 & y_2 & \cdots & y_m \end{pmatrix}$ is positive if $i+j$ is even, negative if $i+j$ and m are both odd, and positive if $i+j$ is odd and m is even. This implies that $\mathcal{I}(x,y)$ is strictly totally positive on the interior of any rectangle $\left[\rho_i^\dagger, \rho_{i+1}^\dagger\right] \times \left[\rho_j^\dagger, \rho_{j+1}^\dagger\right]$ if $i+j$ is even, and is strictly sign regular (with negative sign for m odd and positive sign for m even) on any such rectangle if $i+j$ is odd.

As a note on conclusion 4, Pinkus (1985a) uses similar conclusions (also based on Sylvester's Determinant Identity) in Section 5 of Chapter IV.

Conclusion 5 follows directly from zero-counting arguments. Fixing x, $\mathcal{I}(x,y)$ has $n+1$ zeros as a function of y (these zeros are just $\rho_j^\dagger$, $j = 1, \ldots, n+1$). But for any fixed x, $\mathcal{I}(x,y)$ is a linear combination of $n+2$ terms of the form $\exp(c\tau_k y)$, where $\tau_k = \rho_k^\dagger$ for $k = 1, \ldots, n+1$ and $\tau_{n+2} = x$. Since these exponential terms form a Chebyshev system, $\mathcal{I}(x,y)$ (as a function of y for fixed x) can have no other zeros unless it is identically zero (which occurs only if x is on the boundary of the given rectangle). Rolle's theorem implies that there is at least one local optimum of $\mathcal{I}(x,y)$ (as a function of y for fixed x) between each of its $n+1$ zeros (so, on each interval $\left(\rho_j^\dagger, \rho_{j+1}^\dagger\right)$). This gives at least n local optima (zeros of the derivative with respect to y). If any additional local optimum existed on some interval $\left(\rho_j^\dagger, \rho_{j+1}^\dagger\right)$, there would have to be at least two additional local optima; $\mathcal{I}(x,y)$ must get from zero at $\rho_j^\dagger$ to zero again at $\rho_{j+1}^\dagger$ without any other zeros in between, so the local optimum given by Rolle's theorem must be a maximum if $\mathcal{I}(x, y)$ is positive on the interior of the interval and a minimum if $\mathcal{I}(x, y)$ is negative on the interior of the interval, and if the derivative changes sign again it must do so twice to be consistent with the prescribed zeros. Thus, there are either n local optima or at least $n + 2$ local optima. Since the derivative with respect to y of $\mathcal{I}(x, y)$, where x is held fixed, can also be expressed as a linear combination of $n+2$ functions forming a Chebyshev system (the same exponential terms noted above), there can be only n local optima (one between each pair of neighboring zeros) unless $\mathcal{I}(x, y)$ is identically zero (which it is not on the interior of the rectangle). Identical logic applies when fixing y and regarding $\mathcal{I}(x, y)$ as a function of x.

Conclusion 5 is similar to some results given by Karlin (1968) (Theorem 1.2 and the remark below it, Theorem 1.3, and Theorem 1.4 on pages 157 and 158 in Section 1 of Chapter 4); I could have used the sign regularity already proven in conclusion 4 of $\mathcal{I}(x, y)$ on each rectangle of the given type in conjunction with Karlin's arguments to obtain the same conclusion.

Conclusion 6 follows from the smoothness of $\mathcal{I}(x, y)$ and conclusion 5. Using the notation of conclusion 5, suppose that the curves $x^*(y)$ and $y^*(x)$ coincide on any open interval. This would then be an open interval of critical points. Since both of these curves must be analytic on all of $\mathbb{R}$ (due to the exponential form of $\mathcal{I}(x, y)$ throughout $\mathbb{R} \times \mathbb{R}$ and the analytic implicit function theorem), and since $\mathcal{I}(x, y)$ itself is analytic, this would imply that $\mathcal{I}(x, y^*(x))$ was an analytic function that was constant on an open interval, and thus constant on the whole curve $(x, y^*(x))$. Such a situation occurs, for example, with the Gaussian kernel $\exp\left(-(y-x)^2\right)$, which attains its maximum on the line $y = x$ and is extended totally positive. For $\mathcal{I}(x, y)$, however, this is impossible, as the curve $(x, y^*(x))$ would intersect one of the horizontal or vertical lines on which $\mathcal{I}(x, y)$ is zero (the lines of the form $x = \rho_i^\dagger$ or $y = \rho_j^\dagger$). Thus, if there are multiple local optima of $\mathcal{I}(x, y)$ on a rectange of the given form, they must be isolated.

Any point in the given rectangle at which the gradient of $\mathcal{I}(x, y)$ is zero must be a local maximum if $\mathcal{I}(x, y)$ is positive on the given rectangle or a local minimum if $\mathcal{I}(x, y)$ is negative on the given rectangle. To see this, suppose a local minimum occurred on a rectangle on which $\mathcal{I}(x, y)$ was positive. Then the second derivatives in the x and y directions would be positive at that local minimum (since the Hessian would be positive definite and the minimum would, as shown above, be isolated), but then there would need to be another zero of the derivative with respect to x along the horizontal line through that minimum in order to make the derivative with respect to x negative (which is necessary to bring the value of $\mathcal{I}(x, y)$ to zero at the boundary of the rectangle) and similarly for y, which would contradict the already-established conclusion 5. It is similarly impossible for a zero of the gradient to be a local maximum on a rectangle (of the given form) on which $\mathcal{I}(x, y)$ is negative. Finally, saddle points are also impossible: the direction for which the second derivative at the saddle point has the "wrong" sign (positive if $\mathcal{I}(x, y)$ is positive and negative if $\mathcal{I}(x, y)$ is negative) again must have one too many zeros of the derivative to bring the value of $\mathcal{I}(x, y)$ to zero at the boundary.

Now suppose that there exist two local maxima (two zeros of the gradient) of $\mathcal{I}(x, y)$ on a rectangle of the given form on which $\mathcal{I}(x, y)$ is positive. Then the curve $(x, y^*(x))$ must pass through both points, call them (x_1, y_1) and (x_2, y_2). The derivative $\frac{d\mathcal{I}}{dy} = 0$ for each point on this curve, by the curve's definition. Consider tracing the curve from (x_1, y_1) to (x_2, y_2), supposing without loss of generality that $x_1 < x_2$. Since (x_1, y_1) is a local maximum, there is a point in its neighborhood that is on the curve $(x, y^*(x))$ and at which $\frac{d\mathcal{I}}{dx} < 0$ (the "hilltop" at (x_1, y_1) must be descended from). But since (x_2, y_2) is a also a local maximum, there is a point in its neighborhood that is on the curve $(x, y^*(x))$ and at which $\frac{d\mathcal{I}}{dx} > 0$ (the "hill" must be climbed up to the "local top" (x_2, y_2)). Then Rolle's theorem implies that there is a point on the curve $(x, y^*(x))$ at which $\frac{d\mathcal{I}}{dx} = 0$ and $\frac{d^2\mathcal{I}}{dx^2} = 0$ (since the derivative with respect to x must go from negative to positive at least once as the curve is traced from (x_1, y_1) to (x_2, y_2)). But this point is then a saddle point (a local maximum with respect to y and a local minimum with respect to x), which the previous paragraph showed is impossible. Thus, there can be only one local maximum of $\mathcal{I}(x, y)$ on a rectangle of the given form on which $\mathcal{I}(x, y)$ is positive. An entirely similar argument shows that there can be only one local minimum of $\mathcal{I}(x, y)$ on a rectangle of the given form on which $\mathcal{I}(x, y)$ is negative.

□

With the results of Theorem 7.2 in hand, I now describe my method of bounding the global maximum or minimum of the approximation error $\exp(cxy) - P^*_{c,n}(x, y)$ over $x, y \in [-1, 1]$.

Method 7.2. The following method finds an upper or lower bound, over $x, y \in [-1, 1]$, for the approximation error $\exp(cxy) - P^*_{c,n}(x, y)$ of the rank-n approximating operator $P^*_{c,n}(x, y)$ constructed by Method 7.1.

- **Given**:
  1. m, a scalar which is 1 if approximation error is to be minimized and -1 if it is to be maximized
  2. $c > 0$
  3. a vector $\boldsymbol{\rho}^\dagger$, from the outputs of Method 7.1
  4. a vector $\boldsymbol{z}^*$, from the outputs of Method 7.1
  5. a number d; when a box is divided into sub-boxes, each of the x interval defining the box and the y interval defining the box are divided into d intervals, so there will be d^2 sub-boxes
  6. a tolerance
  7. a maximum allowable number of iterations
- **Output**:

1. a global lower bound (on the minimum approximation error if $m = 1$ and on the maximum approximation error if $m = -1$)
2. a global upper bound (on the minimum approximation error if $m = 1$ and on the maximum approximation error if $m = -1$)

Carefully subdivide the unit square to make all functions of interest monotonic on sub-boxes and use a branch-and-bound approach, employing the results of Theorem 7.2 to control the interpolation component of the approximation error and using the simple rank-one nature of the remaining component of the approximation error.

To use Method 7.2 to find $\max_{x,y\in[-1,1]} \left|\exp(cxy) - P^*_{c,n}(x,y)\right|$, simply run it twice: once to find the global maximum and once to find the global minimum. Taking the larger of the absolute values of these two provides the desired quantity. This simple approach is summarized in the method below.

Method 7.3. The following method finds an upper bound, over $x, y \in [-1, 1]$, for $\left|\exp(cxy) - P^*_{c,n}(x,y)\right|$.

- **Given**:
  1. $c > 0$
  2. a vector $\boldsymbol{\rho}^\dagger$, from the outputs of Method 7.1
  3. a vector $\boldsymbol{z}^*$, from the outputs of Method 7.1
  4. a number d; when a box is divided into sub-boxes, each of the x interval defining the box and the y interval defining the box are divided into d intervals, so there will be d^2 sub-boxes
  5. a tolerance
  6. a maximum allowable number of iterations
- **Output**:
  1. a global lower bound on $\max_{x,y\in[-1,1]} \left|\exp(cxy) - P^*_{c,n}(x,y)\right|$
  2. a global upper bound on $\max_{x,y\in[-1,1]} \left|\exp(cxy) - P^*_{c,n}(x,y)\right|$

Step 1 Run Method 7.2 with $m = 1$ and with the given inputs. Label the resulting lower bound l_{min} and the resulting upper bound u_{min}.

Step 2 Run Method 7.2 with $m = -1$ and with the given inputs. Label the resulting lower bound l_{max} and the resulting upper bound u_{max}.

Step 3 Return $\max(|u_{max}|, |l_{min}|)$ as the global upper bound on $\max_{x,y\in[-1,1]} \left|\exp(cxy) - P^*_{c,n}(x,y)\right|$, and return $\min(|l_{max}|, |u_{min}|)$ as the corresponding global lower bound.

7.3 Numerical Comparison to the Lower Bound

In this subsection, I present a lengthy table comparing the upper bound on the worst-case error of the numerically-optimal approximation $P^*_{c,n}(x, y)$, as found using Method 7.3, to the lower bound for the same c value and approximation rank n, as computed using Method 6.2. The purpose of including this table is to permit the verification that the numerically-optimal approximation has, as claimed, a worst-case approximation error which is within rounding error of the lower bound computed using Method 6.2. The typical difference is in the 14^{th} decimal place. Though it is very rare, the numerically-optimal upper bound is occassionally below the lower bound, but only ever by an amount that is clearly rounding error (for example, this occurs for rank-1 approximation with $c = 2.75$ below, where the difference is 3×10^{-15}, or 3 parts per quadrillion).

Table 1 was generated with maximum iteration parameters of each method set to 100 and with error tolerances set to 10^{-15} in all methods except in the interval-analysis upper bounding of the numerically-optimal approximation's error (through Method 7.2), where the error tolerance was relaxed to 2×10^{-14}.

When interpreting these approximation errors, it is helpful to keep in mind that the numbers being approximated lie between $\exp(-c)$ and $\exp(c)$, so that they bracket the number 1. The absolute, rather than relative, error is important in financial applications because it translates directly into approximation in currency terms.

Table 1: Numerically-optimal Worst-case Approximation Error versus the Lower Bound

	c	n	Lower Bound	Numerically-optimal
	0.050	1	0.050020835937655	0.050020835937655
	0.050	2	0.000312467443172	0.000312467443172
	0.050	3	0.000001302032475	0.000001302032485
	0.050	4	0.000000004068926	0.000000004068936
	0.050	5	0.000000000010172	0.000000000010180
	0.050	6	0.000000000000021	0.000000000000023
	0.050	7	0.000000000000000	0.000000000000002
	0.050	8	0.000000000000000	0.000000000000002
	0.050	9	0.000000000000000	0.000000000000002
	0.050	10	0.000000000000000	0.000000000000008
	0.050	11	0.000000000000000	0.000000000000005
	0.050	12	0.000000000000000	0.000000000000011
	0.050	13	0.000000000000000	0.000000000000022
	0.050	14	0.000000000000000	0.000000000000027
	0.050	15	0.000000000000000	0.000000000000006
	0.100	1	0.100166750019844	0.100166750019844
	0.100	2	0.001249478863397	0.001249478863397
	0.100	3	0.000010415039673	0.000010415039675

Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	0.100	4	0.000000065098742	0.000000065098748
	0.100	5	0.000000000325504	0.000000000325519
	0.100	6	0.000000000001356	0.000000000001362
	0.100	7	0.000000000000005	0.000000000000007
	0.100	8	0.000000000000000	0.000000000000001
	0.100	9	0.000000000000000	0.000000000000010
	0.100	10	0.000000000000000	0.000000000000002
	0.100	11	0.000000000000000	0.000000000000009
	0.100	12	0.000000000000000	0.000000000000004
	0.100	13	0.000000000000000	0.000000000000008
	0.100	14	0.000000000000000	0.000000000000012
	0.100	15	0.000000000000000	0.000000000000003
	0.150	1	0.150563133151613	0.150563133151613
	0.150	2	0.002809859834645	0.002809859834645
	0.150	3	0.000035143900815	0.000035143900815
	0.150	4	0.000000329528062	0.000000329528074
	0.150	5	0.000000002471634	0.000000002471646
	0.150	6	0.000000000015448	0.000000000015461
	0.150	7	0.000000000000083	0.000000000000085
	0.150	8	0.000000000000000	0.000000000000003
	0.150	9	0.000000000000000	0.000000000000002
	0.150	10	0.000000000000000	0.000000000000003
	0.150	11	0.000000000000000	0.000000000000004
	0.150	12	0.000000000000000	0.000000000000004
	0.150	13	0.000000000000000	0.000000000000010
	0.150	14	0.000000000000000	0.000000000000004
	0.150	15	0.000000000000000	0.000000000000030
	0.200	1	0.201336002541094	0.201336002541094
	0.200	2	0.004991647362827	0.004991647362828
	0.200	3	0.000083281328202	0.000083281328203
	0.200	4	0.000001041319603	0.000001041319615
	0.200	5	0.000000010414497	0.000000010414513
	0.200	6	0.000000000086793	0.000000000086805
	0.200	7	0.000000000000620	0.000000000000627
	0.200	8	0.000000000000004	0.000000000000007
	0.200	9	0.000000000000000	0.000000000000003
	0.200	10	0.000000000000000	0.000000000000006
	0.200	11	0.000000000000000	0.000000000000008
	0.200	12	0.000000000000000	0.000000000000012
	0.200	13	0.000000000000000	0.000000000000004
	0.200	14	0.000000000000000	0.000000000000024

Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	0.200	15	0.000000000000000	0.000000000000010
	0.250	1	0.252612316808168	0.252612316808168
	0.250	2	0.007792081610419	0.007792081610420
	0.250	3	0.000162601844046	0.000162601844047
	0.250	4	0.000002541807908	0.000002541807910
	0.250	5	0.000000031778798	0.000000031778814
	0.250	6	0.000000000331063	0.000000000331079
	0.250	7	0.000000000002956	0.000000000002978
	0.250	8	0.000000000000023	0.000000000000025
	0.250	9	0.000000000000000	0.000000000000002
	0.250	10	0.000000000000000	0.000000000000008
	0.250	11	0.000000000000000	0.000000000000010
	0.250	12	0.000000000000000	0.000000000000005
	0.250	13	0.000000000000000	0.000000000000013
	0.250	14	0.000000000000000	0.000000000000017
	0.250	15	0.000000000000000	0.000000000000013
	0.300	1	0.304520293447143	0.304520293447143
	0.300	2	0.011207594610565	0.011207594610567
	0.300	3	0.000280855829961	0.000280855829963
	0.300	4	0.000005269486480	0.000005269486481
	0.300	5	0.000000079064496	0.000000079064508
	0.300	6	0.000000000988452	0.000000000988465
	0.300	7	0.000000000010591	0.000000000010609
	0.300	8	0.000000000000099	0.000000000000102
	0.300	9	0.000000000000001	0.000000000000004
	0.300	10	0.000000000000000	0.000000000000004
	0.300	11	0.000000000000000	0.000000000000004
	0.300	12	0.000000000000000	0.000000000000003
	0.300	13	0.000000000000000	0.000000000000008
	0.300	14	0.000000000000000	0.000000000000019
	0.300	15	0.000000000000000	0.000000000000008
	0.350	1	0.357189729437272	0.357189729437272
	0.350	2	0.015233796266389	0.015233796266390
	0.350	3	0.000445763673789	0.000445763673790
	0.350	4	0.000009759734683	0.000009759734685
	0.350	5	0.000000170860612	0.000000170860621
	0.350	6	0.000000002492217	0.000000002492232
	0.350	7	0.000000000031156	0.000000000031173
	0.350	8	0.000000000000341	0.000000000000348
	0.350	9	0.000000000000003	0.000000000000005
	0.350	10	0.000000000000000	0.000000000000005

Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	0.350	11	0.000000000000000	0.000000000000007
	0.350	12	0.000000000000000	0.000000000000021
	0.350	13	0.000000000000000	0.000000000000006
	0.350	14	0.000000000000000	0.000000000000030
	0.350	15	0.000000000000000	0.000000000000017
	0.400	1	0.410752325802816	0.410752325802816
	0.400	2	0.019865458009982	0.019865458009982
	0.400	3	0.000665010038878	0.000665010038879
	0.400	4	0.000016644484816	0.000016644484817
	0.400	5	0.000000333055720	0.000000333055737
	0.400	6	0.000000005552381	0.000000005552392
	0.400	7	0.000000000079332	0.000000000079345
	0.400	8	0.000000000000992	0.000000000001008
	0.400	9	0.000000000000011	0.000000000000013
	0.400	10	0.000000000000000	0.000000000000003
	0.400	11	0.000000000000000	0.000000000000003
	0.400	12	0.000000000000000	0.000000000000012
	0.400	13	0.000000000000000	0.000000000000004
	0.400	14	0.000000000000000	0.000000000000031
	0.400	15	0.000000000000000	0.000000000000012
	0.450	1	0.465342016934197	0.465342016934198
	0.450	2	0.025096494569984	0.025096494569984
	0.450	3	0.000946238281376	0.000946238281378
	0.450	4	0.000026651829865	0.000026651829868
	0.450	5	0.000000600044436	0.000000600044445
	0.450	6	0.000000011254558	0.000000011254574
	0.450	7	0.000000000180912	0.000000000180931
	0.450	8	0.000000000002544	0.000000000002561
	0.450	9	0.000000000000032	0.000000000000034
	0.450	10	0.000000000000000	0.000000000000003
	0.450	11	0.000000000000000	0.000000000000005
	0.450	12	0.000000000000000	0.000000000000007
	0.450	13	0.000000000000000	0.000000000000022
	0.450	14	0.000000000000000	0.000000000000011
	0.450	15	0.000000000000000	0.000000000000002
	0.500	1	0.521095305493747	0.521095305493747
	0.500	2	0.030919944365647	0.030919944365649
	0.500	3	0.001297045041688	0.001297045041690
	0.500	4	0.000040605572511	0.000040605572514
	0.500	5	0.000001015929275	0.000001015929291
	0.500	6	0.000000021173843	0.000000021173856

Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	0.500	7	0.000000000378196	0.000000000378212
	0.500	8	0.000000000005910	0.000000000005921
	0.500	9	0.000000000000082	0.000000000000088
	0.500	10	0.000000000000001	0.000000000000005
	0.500	11	0.000000000000000	0.000000000000007
	0.500	12	0.000000000000000	0.000000000000003
	0.500	13	0.000000000000000	0.000000000000002
	0.500	14	0.000000000000000	0.000000000000025
	0.500	15	0.000000000000000	0.000000000000010
	0.550	1	0.578151603743454	0.578151603743454
	0.550	2	0.037327949107746	0.037327949107747
	0.550	3	0.001724975036832	0.001724975036834
	0.550	4	0.000059424716036	0.000059424716042
	0.550	5	0.000001635717189	0.000001635717198
	0.550	6	0.000000037503716	0.000000037503733
	0.550	7	0.000000000736896	0.000000000736910
	0.550	8	0.000000000012668	0.000000000012688
	0.550	9	0.000000000000193	0.000000000000201
	0.550	10	0.000000000000002	0.000000000000006
	0.550	11	0.000000000000000	0.000000000000003
	0.550	12	0.000000000000000	0.000000000000006
	0.550	13	0.000000000000000	0.000000000000006
	0.550	14	0.000000000000000	0.000000000000007
	0.550	15	0.000000000000000	0.000000000000007
	0.600	1	0.636653582148241	0.636653582148241
	0.600	2	0.044311733241486	0.044311733241487
	0.600	3	0.002237516081130	0.002237516081131
	0.600	4	0.000084122897828	0.000084122897830
	0.600	5	0.000002526510163	0.000002526510175
	0.600	6	0.000000063199902	0.000000063199914
	0.600	7	0.000000001354755	0.000000001354769
	0.600	8	0.000000000025407	0.000000000025421
	0.600	9	0.000000000000423	0.000000000000436
	0.600	10	0.000000000000006	0.000000000000009
	0.600	11	0.000000000000000	0.000000000000004
	0.600	12	0.000000000000000	0.000000000000004
	0.600	13	0.000000000000000	0.000000000000005
	0.600	14	0.000000000000000	0.000000000000007
	0.600	15	0.000000000000000	0.000000000000020
	0.650	1	0.696747526126440	0.696747526126440
	0.650	2	0.051861583912563	0.051861583912566

Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	0.650	3	0.002842094363336	0.002842094363338
	0.650	4	0.000115807766230	0.000115807766232
	0.650	5	0.000003768689195	0.000003768689214
	0.650	6	0.000000102139095	0.000000102139109
	0.650	7	0.000000002372055	0.000000002372067
	0.650	8	0.000000000048195	0.000000000048208
	0.650	9	0.000000000000870	0.000000000000882
	0.650	10	0.000000000000014	0.000000000000017
	0.650	11	0.000000000000000	0.000000000000003
	0.650	12	0.000000000000000	0.000000000000003
	0.650	13	0.000000000000000	0.000000000000003
	0.650	14	0.000000000000000	0.000000000000011
	0.650	15	0.000000000000000	0.000000000000016
	0.700	1	0.758583701839534	0.758583701839534
	0.700	2	0.059966832173592	0.059966832173596
	0.700	3	0.003546070009045	0.003546070009048
	0.700	4	0.000155680301515	0.000155680301517
	0.700	5	0.000005457091036	0.000005457091043
	0.700	6	0.000000159292483	0.000000159292492
	0.700	7	0.000000003984199	0.000000003984218
	0.700	8	0.000000000087181	0.000000000087200
	0.700	9	0.000000000001696	0.000000000001709
	0.700	10	0.000000000000030	0.000000000000034
	0.700	11	0.000000000000000	0.000000000000005
	0.700	12	0.000000000000000	0.000000000000004
	0.700	13	0.000000000000000	0.000000000000004
	0.700	14	0.000000000000000	0.000000000000007
	0.700	15	0.000000000000000	0.000000000000028
	0.750	1	0.822316731935830	0.822316731935830
	0.750	2	0.068615836173125	0.068615836173129
	0.750	3	0.004356732957934	0.004356732957937
	0.750	4	0.000205034081782	0.000205034081788
	0.750	5	0.000007702177030	0.000007702177038
	0.750	6	0.000000240913986	0.000000240914003
	0.750	7	0.000000006456562	0.000000006456582
	0.750	8	0.000000000151378	0.000000000151399
	0.750	9	0.000000000003154	0.000000000003169
	0.750	10	0.000000000000059	0.000000000000066
	0.750	11	0.000000000000001	0.000000000000005
	0.750	12	0.000000000000000	0.000000000000002
	0.750	13	0.000000000000000	0.000000000000014

Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	0.750	14	0.000000000000000	0.000000000000067
	0.750	15	0.000000000000000	0.000000000000008
	0.800	1	0.888105982187623	0.888105982187623
	0.800	2	0.077795967082414	0.077795967082416
	0.800	3	0.005281299186109	0.005281299186110
	0.800	4	0.000265254494586	0.000265254494588
	0.800	5	0.000010631193429	0.000010631193443
	0.800	6	0.000000354743099	0.000000354743115
	0.800	7	0.000000010141787	0.000000010141803
	0.800	8	0.000000000253645	0.000000000253661
	0.800	9	0.000000000005638	0.000000000005652
	0.800	10	0.000000000000113	0.000000000000119
	0.800	11	0.000000000000002	0.000000000000006
	0.800	12	0.000000000000000	0.000000000000006
	0.800	13	0.000000000000000	0.000000000000008
	0.800	14	0.000000000000000	0.000000000000006
	0.800	15	0.000000000000000	0.000000000000013
	0.850	1	0.956115959988632	0.956115959988632
	0.850	2	0.087493598515021	0.087493598515022
	0.850	3	0.006326907304504	0.006326907304507
	0.850	4	0.000337817895077	0.000337817895080
	0.850	5	0.000014389322539	0.000014389322556
	0.850	6	0.000000510222248	0.000000510222263
	0.850	7	0.000000015499706	0.000000015499726
	0.850	8	0.000000000411895	0.000000000411911
	0.850	9	0.000000000009728	0.000000000009744
	0.850	10	0.000000000000207	0.000000000000215
	0.850	11	0.000000000000004	0.000000000000007
	0.850	12	0.000000000000000	0.000000000000005
	0.850	13	0.000000000000000	0.000000000000021
	0.850	14	0.000000000000000	0.000000000000025
	0.850	15	0.000000000000000	0.000000000000017
	0.900	1	1.026516725708175	1.026516725708176
	0.900	2	0.097694100180484	0.097694100180486
	0.900	3	0.007500615564943	0.007500615564952
	0.900	4	0.000424290711434	0.000424290711437
	0.900	5	0.000019140824085	0.000019140824093
	0.900	6	0.000000718728548	0.000000718728562
	0.900	7	0.000000023120075	0.000000023120094
	0.900	8	0.000000000650579	0.000000000650595
	0.900	9	0.000000000016270	0.000000000016291
				Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	0.900	10	0.000000000000366	0.000000000000377
	0.900	11	0.000000000000007	0.000000000000012
	0.900	12	0.000000000000000	0.000000000000007
	0.900	13	0.000000000000000	0.000000000000005
	0.900	14	0.000000000000000	0.000000000000009
	0.900	15	0.000000000000000	0.000000000000019
	0.950	1	1.099484317930672	1.099484317930673
	0.950	2	0.108381836485067	0.108381836485069
	0.950	3	0.008809399306022	0.008809399306024
	0.950	4	0.000526328498284	0.000526328498289
	0.950	5	0.000025070166141	0.000025070166151
	0.950	6	0.000000993819843	0.000000993819862
	0.950	7	0.000000033748297	0.000000033748313
	0.950	8	0.000000001002464	0.000000001002482
	0.950	9	0.000000000026464	0.000000000026481
	0.950	10	0.000000000000629	0.000000000000644
	0.950	11	0.000000000000013	0.000000000000017
	0.950	12	0.000000000000000	0.000000000000004
	0.950	13	0.000000000000000	0.000000000000004
	0.950	14	0.000000000000000	0.000000000000014
	0.950	15	0.000000000000000	0.000000000000014
	1.000	1	1.175201193643802	1.175201193643802
	1.000	2	0.119540170749650	0.119540170749652
	1.000	3	0.010260148871540	0.010260148871543
	1.000	4	0.000645674938759	0.000645674938763
	1.000	5	0.000032383145029	0.000032383145042
	1.000	6	0.000001351494898	0.000001351494913
	1.000	7	0.000000048314318	0.000000048314336
	1.000	8	0.000000001510760	0.000000001510777
	1.000	9	0.000000000041983	0.000000000042000
	1.000	10	0.000000000001050	0.000000000001062
	1.000	11	0.000000000000024	0.000000000000028
	1.000	12	0.000000000000000	0.000000000000006
	1.000	13	0.000000000000000	0.000000000000021
	1.000	14	0.000000000000000	0.000000000000007
	1.000	15	0.000000000000000	0.000000000000047
	1.250	1	1.601919080300825	1.601919080300827
	1.250	2	0.181740352204681	0.181740352204688
	1.250	3	0.019878444413995	0.019878444413997
	1.250	4	0.001569098482373	0.001569098482380
	1.250	5	0.000098537025810	0.000098537025823

Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	1.250	6	0.000005145160079	0.000005145160092
	1.250	7	0.000000230041324	0.000000230041340
	1.250	8	0.000000008994711	0.000000008994732
	1.250	9	0.000000000312523	0.000000000312543
	1.250	10	0.000000000009771	0.000000000009789
	1.250	11	0.000000000000277	0.000000000000286
	1.250	12	0.000000000000007	0.000000000000014
	1.250	13	0.000000000000000	0.000000000000005
	1.250	14	0.000000000000000	0.000000000000014
	1.250	15	0.000000000000000	0.000000000000046
	1.500	1	2.129279455094817	2.129279455094819
	1.500	2	0.252673041202170	0.252673041202173
	1.500	3	0.034029449823924	0.034029449823933
	1.500	4	0.003235486930446	0.003235486930458
	1.500	5	0.000244317600428	0.000244317600436
	1.500	6	0.000015325480655	0.000015325480668
	1.500	7	0.000000822791581	0.000000822791596
	1.500	8	0.000000038622315	0.000000038622334
	1.500	9	0.000000001610798	0.000000001610813
	1.500	10	0.000000000060446	0.000000000060464
	1.500	11	0.000000000002062	0.000000000002078
	1.500	12	0.000000000000064	0.000000000000075
	1.500	13	0.000000000000002	0.000000000000009
	1.500	14	0.000000000000000	0.000000000000011
	1.500	15	0.000000000000000	0.000000000000017
	1.750	1	2.790414366277642	2.790414366277642
	1.750	2	0.329330364456188	0.329330364456195
	1.750	3	0.053480652887234	0.053480652887238
	1.750	4	0.005954724293127	0.005954724293130
	1.750	5	0.000525843263792	0.000525843263798
	1.750	6	0.000038532256332	0.000038532256344
	1.750	7	0.000002415385419	0.000002415385436
	1.750	8	0.000000132343231	0.000000132343251
	1.750	9	0.000000006441719	0.000000006441745
	1.750	10	0.000000000282086	0.000000000282103
	1.750	11	0.000000000011227	0.000000000011248
	1.750	12	0.000000000000409	0.000000000000428
	1.750	13	0.000000000000014	0.000000000000020
	1.750	14	0.000000000000000	0.000000000000010
	1.750	15	0.000000000000000	0.000000000000019
	2.000	1	3.626860407847019	3.626860407847020
				Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	2.000	2	0.408476154383665	0.408476154383678
	2.000	3	0.078963756663117	0.078963756663127
	2.000	4	0.010081645628606	0.010081645628616
	2.000	5	0.001020237963585	0.001020237963597
	2.000	6	0.000085566871609	0.000085566871628
	2.000	7	0.000006135515856	0.000006135515873
	2.000	8	0.000000384426093	0.000000384426112
	2.000	9	0.000000021393375	0.000000021393393
	2.000	10	0.000000001070967	0.000000001070987
	2.000	11	0.000000000048724	0.000000000048745
	2.000	12	0.000000000002031	0.000000000002050
	2.000	13	0.000000000000078	0.000000000000086
	2.000	14	0.000000000000002	0.000000000000020
	2.000	15	0.000000000000000	0.000000000000013
	2.250	1	4.691168305898331	4.691168305898329
	2.250	2	0.486953715930439	0.486953715930452
	2.250	3	0.111203348536536	0.111203348536543
	2.250	4	0.016010273929693	0.016010273929705
	2.250	5	0.001828377899158	0.001828377899175
	2.250	6	0.000172801608577	0.000172801608591
	2.250	7	0.000013953701764	0.000013953701785
	2.250	8	0.000000984222139	0.000000984222158
	2.250	9	0.000000061646831	0.000000061646855
	2.250	10	0.000000003472979	0.000000003472999
	2.250	11	0.000000000177797	0.000000000177817
	2.250	12	0.000000000008341	0.000000000008365
	2.250	13	0.000000000000361	0.000000000000379
	2.250	14	0.000000000000014	0.000000000000046
	2.250	15	0.000000000000000	0.000000000000018
	2.500	1	6.050204481039786	6.050204481039788
	2.500	2	0.561993735644557	0.561993735644561
	2.500	3	0.150962080321480	0.150962080321487
	2.500	4	0.024166676406501	0.024166676406513
	2.500	5	0.003077298895579	0.003077298895590
	2.500	6	0.000323753197065	0.000323753197076
	2.500	7	0.000029081016043	0.000029081016056
	2.500	8	0.000002280843620	0.000002280843644
	2.500	9	0.000000158815894	0.000000158815915
	2.500	10	0.000000009944953	0.000000009944975
	2.500	11	0.000000000565851	0.000000000565869
	2.500	12	0.000000000029503	0.000000000029533
				Cont'd. on next page

Table 1, cont'd.

	c	n	Lower Bound	Numerically-optimal
	2.500	13	0.000000000001420	0.000000000001442
	2.500	14	0.000000000000063	0.000000000000092
	2.500	15	0.000000000000002	0.000000000000016
	2.750	1	7.789352011490732	7.789352011490729
	2.750	2	0.631447816059510	0.631447816059517
	2.750	3	0.199104696568384	0.199104696568392
	2.750	4	0.034999814456545	0.034999814456554
	2.750	5	0.004922221915624	0.004922221915644
	2.750	6	0.000570797431058	0.000570797431077
	2.750	7	0.000056470282176	0.000056470282207
	2.750	8	0.000004875950727	0.000004875950751
	2.750	9	0.000000373679230	0.000000373679261
	2.750	10	0.000000025750075	0.000000025750093
	2.750	11	0.000000001612138	0.000000001612168
	2.750	12	0.000000000092482	0.000000000092519
	2.750	13	0.000000000004895	0.000000000004929
	2.750	14	0.000000000000240	0.000000000000262
	2.750	15	0.000000000000010	0.000000000000051
	3.000	1	10.017874927409901	10.017874927409904
	3.000	2	0.693900867554392	0.693900867554400
	3.000	3	0.256683305023214	0.256683305023225
	3.000	4	0.048969470769690	0.048969470769698
	3.000	5	0.007548160818644	0.007548160818657
	3.000	6	0.000956997870356	0.000956997870372
	3.000	7	0.000103428857152	0.000103428857183
	3.000	8	0.000009751403624	0.000009751403656
	3.000	9	0.000000815773580	0.000000815773615
	3.000	10	0.000000061352806	0.000000061352825
	3.000	11	0.000000004191719	0.000000004191777
	3.000	12	0.000000000262390	0.000000000262435
	3.000	13	0.000000000015155	0.000000000015209
	3.000	14	0.000000000000812	0.000000000000856
	3.000	15	0.000000000000040	0.000000000000067

8 Optimal Approximations for Average Squared Error: Oblates and Interpolations

In this section, I characterize the n-widths and optimal approximations of exponential product operators when the L^2 norm on approximation error is used. The goal in this case is to minimize average squared approximation error. These approximations are generally not appropriate for risk measurement, where extremes, rather than averages, are crucial. The approximations introduced in Section 7 (the numerically-optimal worst-case-error approximations $P^*_{c,n}(x,y)$) or the Mathieu approximations of Section 9 are more suitable when assessing risk. However, approximations that are optimal for average squared error may be very interesting when averaging over scenarios (as in Monte Carlo simulation or quadrature to compute an expectation) or when cashflow streams are "smooth" (imagine a level annuity) rather than "jagged" (imagine a principal strip).

Theorem 8.1. *Recall that the exponential product operator P_c is positive definite with simple eigenvalues and Markov eigenfunctions (as shown in Corollary 5.1). Label the eigenvalues of P_c in decreasing order: $\lambda_0 > \lambda_1 > \lambda_2 > \cdots$, and label the corresponding eigenfunctions $\psi^c_0, \psi^c_1, \psi^c_2, \ldots$. Let $H = L^2[-1,1]$ with the usual inner product $(f,g) = \int_{-1}^{1} f(x)\, g(x)\, dx$. Then the n-widths of the exponential product operator P_c are characterized by*

$$\begin{aligned} d_n(P_c(H);H) &= d^n(P_c(H);H) \\ &= \delta_n(P_c(H);H) \\ &= b_n(P_c(H);H) \\ &= \lambda_n, \end{aligned} \tag{189}$$

and

1. *the subspace $X_n = span\left\{\psi^c_0, \ldots, \psi^c_{n-1}\right\}$ is optimal for $d_n(P_c(H);H)$,*
2. *the subspace*

   $$L^n = \{\phi : \phi \in H, (\phi, \psi^c_i) = 0, i = 0, \ldots, n-1\}$$

   is optimal for $d^n(P_c(H);H)$,
3. *the linear operator*

   $$Q_n[f](x) = \sum_{i=0}^{n-1} \lambda_i (f, \psi^c_i)\, \psi^c_i(x)$$

   is optimal for $\delta_n(P_c(H);H)$, and
4. *the subspace $X_{n+1} = span\{\psi^c_0, \ldots, \psi^c_n\}$ is optimal for $b_n(P_c(H);H)$.*

Proof. This theorem is a direct result of the much more general Theorem 2.2 on pages 65-66 in Section 2 of Chapter IV in Pinkus (1985a). On page 136, in the Notes and References section at the end of his Chapter IV, Pinkus (1985a) comments that "Theorem 2.2 is one of those known results whose authorship is unclear." Note that Pinkus (1985a) indexes eigenvalues and eigenfunctions starting with 1; as noted in Remark 5.4, I generally prefer to index starting with 1 as well, but I follow the convention of the literature on special functions by indexing eigenvalues and eigenfunctions starting with 0. Thus, my indices are offset from those of Pinkus (1985a) by 1. □

With this result in hand, I need only determine the eigenvalues and eigenfunctions of the exponential product operators (16). Zayed (2007) has noted that these eigenfunctions are the *oblate spheroidal wave functions of order zero*, which are special functions of classical mathematical physics.

Spheroidal wave functions were first studied because they arise when separating variables in the Helmholtz equation in spheroidal coordinates (coordinates that are special cases of three-dimensional ellipsoidal coordinates in which two of the three axes are equal in length). Later, they were found useful in problems of quantum mechanics, acoustics, and electromagnetics. Flammer (1957) gives a brief and helpful history up to the time of his writing in Section 1 of his Chapter 1. A number of monographs and reference texts consider spheroidal wave functions, including Abramowitz & Stegun (1964) (Chapter 21), Arscott (1964) (Chapter VIII), Erdélyi *et al.* (1955) (pages 134-158 in Chapter 16), Flammer (1957), Meixner & Schäfke (1954), Meixner *et al.* (1980), Morse & Feshbach (1953) (pages 1502-1513 in Section 3 of Chapter 11), Olver *et al.* (2010) (Chapter 30), Stratton *et al.* (1956), and Zhang & Jin (1996) (Chapter 15). I will use the abbreviation "oblates" to refer to oblate spheroidal wave functions of order zero.

The notation surrounding spheroidal wave functions is varied and has a propensity to become confusing. I use a notation for the oblates which is close to that employed by Osipov *et al.* (2013) in the prolate case.

There is a clear parallel in analysis of the oblates to the prolate spheroidal wave functions of order zero which have generated great interest in signal processing and numerical analysis since the classic papers of Slepian and his coauthors (Slepian & Pollak (1961), Landau & Pollak (1961), Landau & Pollak (1962)); indeed, one of the characteristics of these prolates is the motivation for the work of Zayed (2007). In the prolate case, Osipov *et al.* (2013) summarize many of the key results in numerical analysis.

As explained in Slepian (1983), the motivating question behind the classic work of Slepian and his coauthors was to find out how large, for a given band limit, "the fraction of the signal's energy that lies in a given time slot" can be (page 382). The work of Melkman (1977) nicely links a related problem to the determination of certain n-widths and their associated optimal subspaces. Melkman (1977) is interested in how well it is possible (in an L^2 sense) to approximate the space of "essentially" time-limited and band-limited functions (of course, only the zero function can be exactly time- and band-limited) using an n-dimensional subspace, and which such subspaces are optimal. See also Pinkus (1985a), whose Example 3.5 on pages 86-93 in Section 3 of Chapter IV summarizes the results of Melkman (1977).

Because the treatment in Zayed (2007) is necessarily brief (the exponential product operator is an example in his concise article), I provide the details of the proof that the oblates with parameter c are the eigenfunctions of the exponential product operator with parameter c in Subsection 8.1 and Subsection 8.2. I also provide an extensive characterization of the oblates and of the eigenvalues of the exponential product operators, which parallels work in the prolate case by Rokhlin and his coauthors (see Osipov *et al.* (2013)) but which seems novel for the oblates, in Subsection 8.3. Finally, I provide a method for evaluating the oblates and another method for the construction of an average-squared-error-optimal approximation to any exponential product operator using a truncated expansion in oblates in Subsection 8.4. My method for evaluating the oblates goes back to Hodge (1970), but my use of them for approximation is new and requires additional theory and computation (the prolates have been much more extensively investigated from this perspective).

Remark 8.1. In this section, I frequently use x and t as the variables names of the arguments to the exponential product operator rather than using y and x as in most of this paper. Of course, this is a completely cosmetic change; it simply allows more appealing notation in many cases in this section.

As I observed in Subsection 5.4, the optimal approximation is not unique in general. Melkman & Micchelli (1978) proved that each Kolmogorov n-width ($n = 1, 2, \ldots$) of an integral operator with a nondegenerate totally positive kernel (a condition weaker than strict total positivity, see Pinkus (1985a), Definition 5.1 on page 108 in Section 5 of Chapter IV) that maps $L^2[0,1]$ to $L^2[0,1]$ is achieved by two optimal interpolatory subspaces as well as by the classical eigenfunction subspace. As usual, replacing $L^2[0,1]$ with $L^2[-1,1]$ does not alter this result, and the exponential product kernel is strictly totally positive (see Theorem 5.8) and thus certainly nondegenerate totally positive, so instances of the general interpolatory subspaces described by Melkman & Micchelli (1978) are also optimal for each Kolmogorov n-width ($n = 1, 2, \ldots$) of the exponential product operator viewed as an operator from $L^2[-1,1]$ to $L^2[-1,1]$.

Pinkus (1985a) (page 115 in Section 5 of Chapter IV) points out that these average-squared-error optimal interpolatory subspaces are constructed using the zeros of the eigenfunctions of the kernel being approximated, so these zeros must be found before an interpolatory approximation can be built. Since I provide a method for the evaluation of the oblates, I can determine their zeros using standard techniques. This immediately makes a linear interpolatory approximation available (I disregard the approximation that involves the kernel and its transpose as being unnecessarily complicated, particularly since the exponential product kernel is already symmetric).

I provide a method for the construction of an average-squared-error-optimal interpolatory approximation to any exponential product operator in Subsection 8.5. The average-squared-error optimal approximations built by the method of Subsection 8.5 are of financial interest because they admit an interpretation that is similar to that given in Subsection 7.1 for the numerically-optimal (for worst-case error) approximation: for a given range of scenarios, the rank-n average-squared-error-optimal interpolatory approximation uses a portfolio of n cashflow streams, each with cashflows at the same n times, to approximate the scenario valuations of any given cashflow stream. Different cashflow streams are approximated using different approximating portfolio weights but the same n approximating cashflow streams.

8.1 A Commuting Differential Operator

A version of Slepian's "lucky accident" (page 379 of Slepian (1983)) occurs for the exponential product operator (16): the oblate spheroidal wave operator (a self-adjoint linear second-order differential operator) commutes with the exponential product operator, just as Slepian and his coauthors found that the prolate spheroidal wave operator commutes with an integral operator that is the key to their problem (Slepian & Pollak (1961), Landau & Pollak (1961), Landau & Pollak (1962)).

The commutation relation between the oblate spheroidal wave operator and the exponential product operator was noted by Zayed (2007) specifically, but can also be derived from some of the classical integral relations in the literature on spheroidal wave functions (see Flammer (1957), for instance). Moreover, such commutation relations have been extensively explored by Grünbaum and his coauthors (Bertero & Grünbaum (1985) and Bertero *et al.* (1986) are especially pertinent, but see also Grünbaum (1981), Grünbaum (1982), and Grünbaum (1983), as well as Duistermaat & Grünbaum (1986)). However, Zayed (2007) is quite concise (as I noted), so I spell out the commutation relation in full.

Because the oblate spheroidal wave operator is a differential operator, it must be defined on a domain that is not $L^2[-1,1]$ itself, but is a dense subset of $L^2[-1,1]$; this is the standard situation when considering unbounded linear operators on Hilbert spaces, see the Hellinger-Toeplitz theorem, provided as a corollary by Reed & Simon (1980), Section III.5, page 84 (and the remarks following this theorem). I first define the appropriate domain for which the oblate spheroidal wave operator is self-adjoint (as I will demonstrate below) and then define the operator itself.

Definition 8.1. The domain D_L consists of all $f(x) \in L^2(-1,1)$ such that $f(x)$ is absolutely continuous on $(-1,1)$, $\left(1-x^2\right)\frac{df}{dx}(x)$ is absolutely continuous on $(-1,1)$, $-\frac{d}{dx}\left(\left(1-x^2\right)\frac{df}{dx}(x)\right) \in L^2(-1,1)$, and $\lim_{x\to 1}\left\{\left(1-x^2\right)\frac{df}{dx}(x)\right\} = \lim_{x\to -1}\left\{\left(1-x^2\right)\frac{df}{dx}(x)\right\} = 0$.

The domain D_L is, of course, a linear subspace of $L^2(-1,1)$; it is also dense in $L^2(-1,1)$ (since D_L contains all polynomials). Less obvious is that the final condition concerning the limits as $x \to 1$ and $x \to -1$ is equivalent to: f has a finite limit as $x \to 1$ and a finite limit as $x \to -1$. Both of these equivalent limiting conditions are also equivalent to: $\sqrt{1-x^2}f(x) \in L^2(-1,1)$. Akhiezer & Glazman (1963), pages 206 - 210, provide a proof that these three conditions are equivalent in their very helpful analysis of the domain on which the Legendre differential operation becomes a self-adjoint differential operator. For a more detailed discussion of self-adjoint extensions, see Akhiezer & Glazman (1963), Lax (2002), Reed & Simon (1980), Reed & Simon (1975), and Riesz & Szökefalvi-Nagy (1955).

Definition 8.2. The *oblate spheroidal wave operator* with parameter $c > 0$ and defined on the domain D_L is

$$L_c[f](x) \equiv -\frac{d}{dx}\left(\left(1-x^2\right)\frac{df}{dx}(x)\right) - c^2x^2f(x). \tag{190}$$

I chose the domain D_L with the help of the results of Akhiezer & Glazman (1963), pages 206 - 210, on the Legendre differential operator; as an operator on many natural smaller domains (e. g., the space of twice-continuously-differentiable functions on $[-1,1]$ whose values and first derivatives equal zero at the endpoints), L_c would be symmetric but *not* self-adjoint. To be self-adjoint, L_c must have a domain that equals the domain of its adjoint operator; the domain of the adjoint is always a superset of the domain of the original operator, and for a symmetric but not self-adjoint operator, it is a strict superset. See Reed & Simon (1980) (page 255 and following) for a detailed discussion of the difference between the symmetry property and self-adjointness. This is essentially an issue of appropriate boundary conditions; if the conditions defining D_L were too restrictive, then the domain of the adjoint of L_c would be larger than D_L. In general, one would begin with restrictive conditions defining the domain of the unbounded operator in question and then seek *self-adjoint extensions* of the operator (which are simply enlargements of the domain, not changes in the actual differential operation to be performed on the members of the domain) – these extensions would correspond to weakening the restrictions defining the domain of the operator until the domain of the operator equaled the domain of its adjoint. There are systematic methods for finding such self-adjoint extensions, which are described in Akhiezer & Glazman (1963) Chapter VII (and, for differential operators in particular, Appendix II), Reed & Simon (1975), Chapter X, and Riesz & Szökefalvi-Nagy (1955), Chapter VIII, Sections 123 - 125. In my case, I use the fact that D_L is the domain of a related self-adjoint operator (the Legendre differential operator) as proven by Akhiezer & Glazman (1963), pages 206 - 210.

Theorem 8.2. *L_c is self-adjoint and is bounded below by $-c^2$.*

Proof. Akhiezer & Glazman (1963), pages 206 - 210, demonstrate that the Legendre differential operator $A \equiv -\frac{d}{dx}\left(\left(1-x^2\right)\frac{df}{dx}(x)\right)$ with domain of definition D_L is self-adjoint. The operator $B \equiv -c^2x^2$ (a multiplication operator with domain all of $L^2(-1,1)$) is clearly bounded with operator norm not exceeding c^2 (if $\left(\int_{-1}^{1}(f(x))^2\,dx\right)^{\frac{1}{2}} = C_0$ then $\left(\int_{-1}^{1}\left(-c^2x^2f(x)\right)^2dx\right)^{\frac{1}{2}} \leq c^2\left(\int_{-1}^{1}(f(x))^2\,dx\right)^{\frac{1}{2}} = c^2C_0$) and is obviously symmetric. In the language of Reed & Simon (1975), Chapter X, Section 2, page 162, B is *infinitesimally small* with respect to A, so that the Reed & Simon (1975) parameter a on page 162 of Chapter X, Section 2 is zero and the parameter b is no greater than c^2. Thus, by the Kato-Rellich theorem (Reed & Simon (1975), Chapter X, Section 2, Theorem X.12 on page 162), the operator $A+B = -\frac{d}{dx}\left(\left(1-x^2\right)\frac{df}{dx}(x)\right) - c^2x^2 = L_c$ is self-adjoint with domain of definition D_L and is bounded below by $-c^2$ (since the Legendre differential operator A is positive definite, and thus bounded below by zero; see Akhiezer & Glazman (1963), pages 206 - 210).

$\square$

Corollary 8.1. *L_c is symmetric, that is, for all $f, g \in D_L$,*

$$\int_{-1}^{1} L_c[f](x)\, g(x)\, dx = \int_{-1}^{1} f(x)\, L_c[g](x)\, dx.$$

Proof. Every self-adjoint operator is, by definition, symmetric; see Reed & Simon (1980), Chapter VIII, Section 2, the relevant definitions on page 255. $\square$

Akhiezer & Glazman (1961), Section 14, page 32 provide a helpful definition of what it means for two operators, at least one of which is bounded and thus defined everywhere, to commute. I use this definition below.

Theorem 8.3. *P_c and L_c commute; that is, $f \in D_L$ implies that $P_c[f] \in D_L$ and that*

$$L_c[P_c[f]](x) = P_c[L_c[f]](x).$$

Proof. It is clear that $f \in D_L$ implies that $P_c[f] \in D_L$, given the form of P_c.

Under the assumption that $f \in D_L$, I compute the functions $L_c[P_c[f]](x)$ and $P_c[L_c[f]](x)$, which demonstrates that these two functions are equal.

$$\begin{aligned}
& L_c[P_c[f]](x) \\
= \; & -\frac{d}{dx}\left(\left(1-x^2\right)\frac{dP_c[f]}{dx}(x)\right) - c^2x^2P_c[f](x) && (191)\\
= \; & -\left(1-x^2\right)\frac{d^2}{dx^2}\int_{-1}^{1}\exp(cxt)f(t)\,dt && (192)\\
& +2x\frac{d}{dx}\int_{-1}^{1}\exp(cxt)f(t)\,dt \\
& -c^2x^2\int_{-1}^{1}\exp(cxt)f(t)\,dt \\
= \; & -\left(1-x^2\right)\int_{-1}^{1}\frac{d^2}{dx^2}\exp(cxt)f(t)\,dt && (193)\\
& +2x\int_{-1}^{1}\frac{d}{dx}\exp(cxt)f(t)\,dt \\
& -c^2x^2\int_{-1}^{1}\exp(cxt)f(t)\,dt \\
= \; & -\left(1-x^2\right)\int_{-1}^{1}c^2t^2\exp(cxt)f(t)\,dt && (194)\\
& +2x\int_{-1}^{1}ct\exp(cxt)f(t)\,dt
\end{aligned}$$

$$-c^2x^2 \int_{-1}^{1} \exp(cxt) f(t)\, dt,$$

where the first equality is by the definition of L_c, the second equality is by the definition of P_c and elementary manipulation, the third equality interchanges the order of differentiation and integration, and the fourth equality computes the derivatives shown in the prior expression.

$$\begin{aligned}
& P_c[L_c[f]](x) && (195)\\
= & \int_{-1}^{1} \exp(cxt)\left(-\frac{d}{dt}\left(\left(1-t^2\right)\frac{df}{dt}(t)\right) - c^2t^2 f(t)\right) dt && (196)\\
= & \int_{-1}^{1} \left(-\frac{d}{dt}\left(\left(1-t^2\right)\frac{d}{dt}\exp(cxt)\right) - c^2t^2\exp(cxt)\right) f(t)\, dt && (197)\\
= & \int_{-1}^{1} \left(-\frac{d}{dt}\left(\left(1-t^2\right)cx\exp(cxt)\right) - c^2t^2\exp(cxt)\right) f(t)\, dt && (198)\\
= & \int_{-1}^{1} \left(\begin{array}{c} -\left(\left(1-t^2\right)cx\frac{d}{dt}\exp(cxt) + 2tcx\exp(cxt)\right) \\ -c^2t^2\exp(cxt) \end{array}\right) f(t)\, dt && (199)\\
= & \int_{-1}^{1} \left(\begin{array}{c} -\left(\left(1-t^2\right)c^2x^2\exp(cxt) + 2tcx\exp(cxt)\right) \\ -c^2t^2\exp(cxt) \end{array}\right) f(t)\, dt && (200)\\
= & \int_{-1}^{1} \left(\begin{array}{c} -c^2x^2\exp(cxt) + c^2x^2t^2\exp(cxt) \\ +2tcx\exp(cxt) - c^2t^2\exp(cxt) \end{array}\right) f(t)\, dt && (201)\\
= & \int_{-1}^{1} \left(\begin{array}{c} -c^2x^2\exp(cxt) - \left(1-x^2\right)c^2t^2\exp(cxt) \\ +2tcx\exp(cxt) \end{array}\right) f(t)\, dt && (202)\\
= & -\left(1-x^2\right)\int_{-1}^{1} c^2t^2\exp(cxt)f(t)\, dt + 2x\int_{-1}^{1} ct\exp(cxt)f(t)\, dt && (203)\\
& -c^2x^2\int_{-1}^{1}\exp(cxt)f(t)\, dt,
\end{aligned}$$

where the first equality is by the definitions of L_c and P_c, the second equality is by Corollary 8.1 (the symmetry of L_c, where I regard $\exp(cxt)$ as a function of t, which applies because $f \in D_L$ by assumption), the third equality computes $\frac{d}{dt}\exp(cxt)$, the fourth equality employs the product rule of differentiation, the fifth equality computes $\frac{d}{dt}\exp(cxt)$, the sixth equality expands the terms of the integrand, the seventh equality regroups terms within the integrand, and the final equality exploits the linearity of the integral.

Since (194) and (203) are identical, $L_c[P_c[f]](x) = P_c[L_c[f]](x)$, that is, P_c and L_c commute as claimed.

$\square$

8.2 The Eigenfunctions of Exponential Product Operators

In this subsection, I characterize the eigenfunctions of the exponential product operators P_c, $c > 0$.

Theorem 8.4. *The eigenfunctions $\psi_i^c(x)$, $i = 0, 1, 2, \dots$ of P_c are analytic in $x \in \mathbb{R}$. They can be analytically continued to entire functions of $x \in \mathbb{C}$.*

Proof. Because

$$\lambda_i(c)\,\psi_i^c(x) = \int_{-1}^{1} \exp(cxt)\psi_i^c(t)\,dt, \tag{204}$$

where $\lambda_i(c) > 0$ by Corollary 5.1,

$$\psi_i^c(x) = \frac{1}{\lambda_i(c)} \int_{-1}^{1} \exp(cxt) f(t)\,dt \tag{205}$$

for some continuous $f(t)$ (indeed, $f(t) = \psi_i^c(t)$ satisfies (205)). But any function that can be written as the expression on the right-hand side of (205) is clearly analytic. Since the right-hand side of (204) depends only on the values of $\psi_i^c(t)$ for $t \in [-1, 1]$, I can use (204) to analytically continue ψ_i^c by simply substituting $x \in \mathbb{C}$ into (204). Note the similarity between this result and Remark 2 of Section 2.4 in Osipov *et al.* (2013). □

In the remainder of this section, I demonstrate that the key integral operator and the oblate spheroidal wave operator have the same eigenfunctions. Although the key results I use here are due to Gantmacher & Krein (2002), there is a very helpful discussion and extension of these results available in Karlin (1968). Further, as Gantmacher & Krein (2002) note, Kellogg (1918) was the first to produce results of the type I use below (with relevant prior work in Kellogg (1916)); Gantmacher & Krein (2002) extended Kellogg's results.

Having already characterized the eigenvalues and eigenvectors of the key integral operator P_c through Corollary 5.1, I must now describe the eigenvalues and eigenvectors of the differential operator L_c that, as I showed in Theorem 8.3, commutes with P_c. Fortunately, the oblate spheroidal wave functions of order zero are classical special functions of mathematical physics; as a result, they have been extensively characterized. The next theorem summarizes the results available in the significant literature on these functions.

Theorem 8.5. *The operator L_c has a countable set of eigenfunctions which form a complete orthonormal basis for $L^2[-1, 1]$. Each eigenvalue $\chi_n(c)$ of L_c is simple. The eigenvalues $\chi_n(c)$ are bounded below but not above: $\chi_n(c) \to \infty$ as $n \to \infty$. Ordering these eigenvalues $\chi_0(c) < \chi_1(c) < \chi_2(c) < \cdots$, the eigenfunction $\phi_n^c(x)$ of L_c that is associated with the eigenvalue $\chi_n(c)$ has exactly n zeros in the interval $[-1, 1]$.*

Proof. These results are proven in Meixner & Schäfke (1954), pages 235 through 238 spanning sections 3.22 and 3.23: Satz 4 on page 238 shows that the countable set of eigenfunctions forms a complete orthonormal basis in $L^2[-1,1]$; Satz 1 on page 235 shows that each eigenvalue is simple, that the eigenvalues are countable, and that they must be bounded below (due to the result for $c^2=0$ and the analyticity of each eigenvalue as a function of c^2); Satz 5 on page 238 shows that the n^{th} eigenfunction has exactly n zeros in the interval $(-1,1)$. This leaves open the possibility of zeros at -1 or 1 (or both), but the prolate-case argument of Hogan & Lakey (2012) (Lemma 2.6.8 on page 82 in section 2.6.1) extends immediately to the oblate case to show that no eigenfunction can have a zero at 1. (Osipov *et al.* (2013) give the same argument, for the prolate case, as part of their Lemma 7.2 on pages 232 and 233 of Section 3 in Chapter 7.) Because the eigenfunctions are even functions for n even and odd functions for n odd (Olver *et al.* (2010), equation 30.4.3 on page 699 in Section 4 of Chapter 30; this is also clear from the expansion of an oblate in Legendre polynomials), this additionally implies that no eigenfunction can have a zero at -1. Thus, the n^{th} eigenfunction has exactly n zeros in the closed interval $[-1,1]$. □

Statements of most of Theorem 8.5, and many other results, can be found in Arscott (1964) (Chapter VIII), Meixner *et al.* (1980), and Olver *et al.* (2010) (Chapter 30).

With characterizations of P_c and L_c both in hand, I can now state the key result of this section. When considering the uniqueness of an eigenfunction corresponding to a given eigenvalue, recall that my definition of an eigenfunction $\psi(x)$, $x\in(-1,1)$ includes the normalization $\int_{-1}^{1}(\psi(x))^2\,dx=1$. Thus, if an eigenvalue is simple there is a unique eigenfunction which has that eigenvalue. If my definition did not include normalization, the uniqueness would only be up to multiplication by a scalar.

Theorem 8.6. *An eigenfunction $\psi_n^c(x)$ of P_c whose eigenvalue is $\lambda_n(c)$ is the unique eigenfunction of P_c that has the eigenvalue $\lambda_n(c)$. An eigenfunction $\phi_n^c(x)$ of L_c whose eigenvalue is $\chi_n(c)$ is the unique eigenfunction of L_c that has the eigenvalue $\chi_n(c)$. $\psi_n^c(x)$ is the unique eigenfunction of P_c associated with the eigenvalue $\lambda_n(c)$ of P_c (where $\lambda_0(c)>\lambda_1(c)>\lambda_2(c)>\cdots$ and $\lambda_m(c)>0$ for every $m=0,1,2,\ldots$) if and only if $\psi_n^c(x)$ is the unique eigenfunction of L_c corresponding to the eigenvalue $\chi_n(c)$ of L_c (where $\chi_0(c)<\chi_1(c)<\chi_2(c)<\cdots$). Further, $\lambda_n(c)\to 0$ as $n\to\infty$ and $\chi_n(c)\to\infty$ as $n\to\infty$.*

Proof. Suppose $\psi_n^c(x)$ is an eigenfunction of P_c with eigenvalue $\lambda_n(c)$. By Corollary 5.1, the eigenvalue $\lambda_n(c)$ is simple, so the space of $f(x)$ (that are not identically zero) such that $\lambda_n(c)f(x)=P_c[f](x)$ is one-dimensional: if f and g are both in this space, then there is a non-zero constant a such that $f(x)=ag(x)$. Since my definition of an eigenfunction $\psi(x)$, $x\in(-1,1)$ includes the normalization $\int_{-1}^{1}(\psi(x))^2\,dx=1$, I conclude that $\psi_n^c(x)$ is the unique eigenfunction of P_c that has the eigenvalue $\lambda_n(c)$.

By Corollary 5.1, P_c is positive definite with simple eigenvalues and Markov eigenfunctions. By Definition 5.28, this implies that its eigenvalues are countable, all positive, and can be ordered as $\lambda_0(c)>\lambda_1(c)>\lambda_2(c)>\cdots$. Because its kernel and the measure of its domain of integration are both bounded, P_c is certainly Hilbert-Schmidt (and thus compact); it is also self-adjoint, so Reed & Simon (1980), Theorem VI.16 on page 203 implies that $\lambda_n(c) \to 0$ as $n \to \infty$.

The uniqueness of $\phi_n^c(x)$ as the only eigenfunction of L_c with eigenvalue $\chi_n(c)$ is assured by Theorem 8.5. Another part of the conclusion of Theorem 8.5 is that the eigenvalues of L_c can be ordered as $\chi_0(c) < \chi_1(c) < \chi_2(c) < \cdots$ and that $\chi_n(c) \to \infty$ as $n \to \infty$.

If $\psi_n^c(x)$ is the unique eigenfunction of P_c associated with the eigenvalue $\lambda_n(c)$ of P_c (where $\lambda_n(c)$ has index n in the ordered sequence $\lambda_0(c) > \lambda_1(c) > \lambda_2(c) > \cdots$), then

$$\lambda_n(c)\,\psi_n^c(x) = P_c[\psi_n^c](x). \tag{206}$$

Apply L_c to each side of (206) to obtain

$$\begin{aligned} \lambda_n(c)\,L_c[\psi_n^c](x) &= L_c[P_c[\psi_n^c]](x) &(207)\\ &= P_c[L_c[\psi_n^c]](x), &(208) \end{aligned}$$

where the first equality is by the linearity of L_c and by (206) and the second equality is by Theorem 8.3 (which shows that L_c and P_c commute). But since $\lambda_n(c)$ is a simple eigenvalue of P_c (by Corollary 5.1), the space of $f(x)$ such that $\lambda_n(c)\,f(x) = P_c[f](x)$ is one-dimensional: all such f are scalar multiples of $\psi_n^c(x)$, the unique eigenfunction that has the eigenvalue $\lambda_n(c)$. Thus, (208) shows that $L_c[\psi_n^c](x)$ must be a scalar multiple of $\psi_n^c(x)$; that is, there exists some constant (with respect to x) χ such that

$$\chi\psi_n^c(x) = L_c[\psi_n^c](x), \tag{209}$$

so $\psi_n^c(x)$ is an eigenfunction of L_c with eigenvalue χ (since $\psi_n^c(x)$ is an eigenfunction of P_c, it satisfies $\int_{-1}^{1}(\psi_n^c(x))^2\,dx = 1$, so it must be an eigenfunction of L_c and not just a scalar multiple of an eigenfunction of L_c). By Corollary 5.1 and by Definition 5.28, since the eigenvalue $\lambda_n(c)$ has index n in the ordered sequence $\lambda_0(c) > \lambda_1(c) > \lambda_2(c) > \cdots$, the eigenfunction $\psi_n^c(x)$ has n zeros in the interval $[-1,1]$ (and is the only eigenfunction of P_c that has n zeros in this interval). By Theorem 8.5, the only eigenfunction of L_c that has n zeros in the interval $[-1,1]$ is associated with the eigenvalue $\chi_n(c)$ that has index n in the ordered sequence $\chi_0(c) < \chi_1(c) < \chi_2(c) < \cdots$. Thus, $\psi_n^c(x)$ is an eigenfunction of L_c with eigenvalue $\chi_n(c)$.

If $\psi_n^c(x)$ is the unique eigenfunction of L_c associated with the eigenvalue $\chi_n(c)$ of L_c (where $\chi_n(c)$ has index n in the ordered sequence $\chi_0(c) < \chi_1(c) < \chi_2(c) < \cdots$), then

$$\chi_n(c)\,\psi_n^c(x) = L_c[\psi_n^c](x). \tag{210}$$

Apply P_c to each side of (210) to obtain

$$\begin{aligned} \chi_n(c)\,P_c[\psi_n^c](x) &= P_c[L_c[\psi_n^c]](x) &(211)\\ &= L_c[P_c[\psi_n^c]](x), &(212) \end{aligned}$$

where the first equality is by the linearity of P_c and by (210) and the second equality is by Theorem 8.3 (which shows that L_c and P_c commute). But since $\chi_n(c)$ is a simple eigenvalue of L_c (by Theorem 8.5), the space of $f(x)$ such that $\chi_n(c) f(x) = L_c[f](x)$ is one-dimensional: all such f are scalar multiples of $\psi_n^c(x)$, the unique eigenfunction that has the eigenvalue $\chi_n(c)$. Thus, (212) shows that $P_c[\psi_n^c](x)$ must be a scalar multiple of $\psi_n^c(x)$; that is, there exists some constant (with respect to x) λ such that

$$\lambda\psi_n^c(x) \quad = \quad P_c[\psi_n^c](x), \tag{213}$$

so $\psi_n^c(x)$ is an eigenfunction of P_c with eigenvalue λ (since $\psi_n^c(x)$ is an eigenfunction of L_c, it satisfies $\int_{-1}^{1}(\psi_n^c(x))^2\,dx = 1$, so it must be an eigenfunction of P_c and not just a scalar multiple of an eigenfunction of P_c). By Theorem 8.5, since the eigenvalue $\chi_n(c)$ has index n in the ordered sequence $\chi_0(c) < \chi_1(c) < \chi_2(c) < \cdots$, the eigenfunction $\psi_n^c(x)$ has n zeros in the interval $[-1,1]$ (and is the only eigenfunction of L_c that has n zeros in this interval). By Corollary 5.1 and by Definition 5.28, the only eigenfunction of P_c that has n zeros in the interval $[-1,1]$ is associated with the eigenvalue $\lambda_n(c)$ that has index n in the ordered sequence $\lambda_0(c) > \lambda_1(c) > \lambda_2(c) > \cdots$. Thus, $\psi_n^c(x)$ is an eigenfunction of P_c with eigenvalue $\lambda_n(c)$.

□

Theorem 8.6 allows me to use the differential operator L_c to characterize the eigenstructure of the integral operator P_c. The interplay between these operators yields a rich set of results for the oblate case, as was noted in the prolate case by Osipov *et al.* (2013). In particular, it is not obvious how one might obtain the approximation bounds that I prove in Section 8.3 without using the fact that the eigenfunctions of P_c are also eigenfunctions of L_c.

8.3 Approximation of the Exponential Product Operator Using Oblates: Theoretical Results

In this subsection, I show how well truncated sums $\sum_{n=0}^{N}\lambda_n(c)\,\psi_n^c(x)\,\psi_n^c(t)$ approximate $K(x,t) = \exp(cxt)$, where $\lambda_n(c)$ and $\psi_n^c(x)$ have the same meanings as in Theorem 8.6. Of course, the results above imply that $K(x,t) = \sum_{n=0}^{\infty}\lambda_n(c)\,\psi_n^c(x)\,\psi_n^c(t)$ with convergence in an L^2 sense; with the benefit of the continuity of K and Mercer's theorem, it is clear that convergence in this expression is also uniform. I will provide methods for the measurement of both L^2 approximation error and worst-case (uniform) approximation error. To do so, I will first derive differential equations satisfied by $\lambda_n(c)$ and $\chi_n(c)$ (both with the same meanings as in Theorem 8.6) as functions of c.

The differential equation for $\lambda_n(c)$ as a function of c has its origins, in the prolate case, in the work of Fuchs (1964), as noted by Osipov *et al.* (2013). The prolate version of this theorem is given, with a slightly modified version of the proof from Fuchs (1964), by Osipov *et al.* (2013) (Theorem 7.9 on page 238). I am not aware of any prior statement or proof of this theorem in the oblate case of interest here, but the proof is a straightforward adaptation of the proof for the prolate case.

Theorem 8.7. *For all $c > 0$ and for all $n = 0, 1, 2, \ldots$,*

$$\frac{d\lambda_n}{dc}(c) \quad = \quad \lambda_n(c)\,\frac{2\left(\psi_n^c(1)\right)^2 - 1}{2c}$$

Proof. I closely follow the proof of Theorem 7.9 of Osipov *et al.* (2013) (pages 238 - 239), with appropriate modifications to account for the fact that I deal with the oblate, rather than the prolate, case. Let $c, a > 0$ and define $\beta = \sqrt{c/a}$. By multiplying both sides of the eigenvalue-eigenfunction identity for $\lambda_n(c)$ by the function $\psi_n^a(\beta x)$ I obtain that

$$\lambda_n(c)\,\psi_n^c(x)\,\psi_n^a(\beta x) \quad = \quad \psi_n^a(\beta x)\int_{-1}^{1}\exp(cxt)\,\psi_n^c(t)\,dt. \tag{214}$$

Integrate both sides of (214) over $[-1, 1]$ with respect to x to obtain

$$\begin{aligned} & \lambda_n(c)\int_{-1}^{1}\psi_n^c(x)\,\psi_n^a(\beta x)\,dx \\ = \quad & \int_{-1}^{1}\psi_n^a(\beta x)\int_{-1}^{1}\exp(cxt)\,\psi_n^c(t)\,dtdx \qquad (215)\\ = \quad & \frac{1}{\beta}\int_{-1}^{1}\psi_n^a(s)\int_{-\beta}^{\beta}\exp(a\beta st)\,\psi_n^c(t)\,dtds, \qquad (216)\end{aligned}$$

where the second equality follows from the change of variable $s = \beta x$ in the integral over x (and the observation that $cx = a\beta^2 x = a\beta s$ due to the definition of β). Now note that the eigenvalue-eigenfunction identity $\lambda_n(a)\,\psi_n^a(\beta t) = \int_{-1}^{1}\exp(a\beta st)\,\psi_n^a(s)\,ds$ implies that

$$\begin{aligned} & \int_{-\beta}^{\beta}\exp(a\beta st)\,\psi_n^a(s)\,ds \\ = \quad & \lambda_n(a)\,\psi_n^a(\beta t) \qquad (217)\\ & +\int_{-\beta}^{-1}\exp(a\beta st)\,\psi_n^a(s)\,ds \\ & +\int_{1}^{\beta}\exp(a\beta st)\,\psi_n^a(s)\,ds.\end{aligned}$$

Note that the order of integration is most natural in the above expression if $\beta > 1$, but with the usual rule that $\int_y^z f(x)\,dx = -\int_z^y f(x)\,dx$ it makes sense for $\beta < 1$ as well. Use Fubini's theorem (Stroock (1999), Theorem 4.1.6) and substitute (217) into (216) to get

$$\begin{aligned} & \left(\lambda_n(c) - \tfrac{\lambda_n(a)}{\beta}\right)\int_{-1}^{1}\psi_n^c(x)\,\psi_n^a(\beta x)\,dx = \\ & \tfrac{1}{\beta}\int_{-1}^{1}\psi_n^c(t)\,dt\left(\int_{-\beta}^{-1}\exp(a\beta st)\,\psi_n^a(s)\,ds + \int_{1}^{\beta}\exp(a\beta st)\,\psi_n^a(s)\,ds\right). \qquad (218)\end{aligned}$$

Divide the left-hand side of (218) by $\beta - 1$ and take the limit as $\beta \to 1$:

$$
\begin{aligned}
& \lim_{\beta\to 1} \frac{1}{\beta - 1}\left(\lambda_n(c) - \frac{\lambda_n(a)}{\beta}\right)\int_{-1}^{1}\psi_n^c(x)\,\psi_n^a(\beta x)\,dx \\
= \quad & \lim_{a\to c} \frac{1}{\sqrt{\frac{c}{a}} - 1}\left(\lambda_n(c) - \frac{\lambda_n(a)}{\sqrt{\frac{c}{a}}}\right)\int_{-1}^{1}\psi_n^c(x)\,\psi_n^a\left(\sqrt{\frac{c}{a}}x\right)dx \\
= \quad & \lim_{a\to c} \frac{\sqrt{a}}{\sqrt{c} - \sqrt{a}}\left(\lambda_n(c) - \frac{\sqrt{a}}{\sqrt{c}}\lambda_n(a)\right)\int_{-1}^{1}\psi_n^c(x)\,\psi_n^a\left(\sqrt{\frac{c}{a}}x\right)dx \\
= \quad & \lim_{a\to c} \frac{\sqrt{a}}{\sqrt{c} - \sqrt{a}}\left(\frac{\sqrt{c}}{\sqrt{c}}\lambda_n(c) - \frac{\sqrt{a}}{\sqrt{c}}\lambda_n(a)\right)\int_{-1}^{1}\psi_n^c(x)\,\psi_n^a\left(\sqrt{\frac{c}{a}}x\right)dx \\
= \quad & \lim_{a\to c} \frac{\sqrt{a}}{\sqrt{c}\left(\sqrt{c} - \sqrt{a}\right)}\left(\sqrt{c}\lambda_n(c) - \sqrt{a}\lambda_n(a)\right)\int_{-1}^{1}\psi_n^c(x)\,\psi_n^a\left(\sqrt{\frac{c}{a}}x\right)dx \\
= \quad & \lim_{a\to c} \frac{\sqrt{a}}{\sqrt{c}}\,\frac{\sqrt{c}\lambda_n(c) - \sqrt{a}\lambda_n(a)}{\sqrt{c} - \sqrt{a}}\int_{-1}^{1}\psi_n^c(x)\,\psi_n^a\left(\sqrt{\frac{c}{a}}x\right)dx \\
= \quad & \lim_{a\to c} \frac{\sqrt{a}}{\sqrt{c}}\,\frac{\sqrt{c}\lambda_n(c) - \sqrt{a}\lambda_n(a)}{\sqrt{c} - \sqrt{a}}\int_{-1}^{1}\left(\psi_n^c(x)\right)^2 dx \\
= \quad & \lim_{a\to c} \frac{\sqrt{c}\lambda_n(c) - \sqrt{a}\lambda_n(a)}{\sqrt{c} - \sqrt{a}} \\
= \quad & \frac{d\left(\sqrt{c}\lambda_n(c)\right)}{d\sqrt{c}} \\
= \quad & \frac{d\left(\sqrt{c}\lambda_n(c)\right)}{dc}\,\frac{dc}{d\sqrt{c}} \\
= \quad & \lambda_n(c) + 2c\frac{d\lambda_n}{dc}(c),
\end{aligned}
\tag{219}
$$

where the first equality follows from $\beta \to 1$ (which implies $a \to c$ by the definition of β) and substitution of the definition of β into the preceding expression, the second equality is by elementary algebraic manipulation, the third equality recognizes that multiplying by $\sqrt{c}/\sqrt{c}$ is just multiplication by 1, the fourth equality is by elementary algebra, the fifth equality rearranges numerator and denominator terms in the prior expression, the sixth equality uses the fact that the limit of a product is the product of the limits if all terms in the product have finite limits and interchanges limit and integration, the seventh equality recognizes that the limit of $\sqrt{a}/\sqrt{c}$ as $a \to c$ is 1 and that $\int_{-1}^{1}\left(\psi_n^c(x)\right)^2 dx = 1$ by the orthonormality of the eigenfunctions $\psi_n^c(x)$, the eighth equality is by the definition of the derivative, the ninth equality is by the chain rule of differentiation, and the tenth equality simply evaluates the prior expression using elementary rules of differentiation (e. g., the product rule).

Focus now on the right-hand side of (218). Divide this expression by $\beta - 1$

and take the limit as $\beta \to 1$:

$$\begin{aligned}
& \lim_{\beta \to 1} \frac{1}{\beta - 1} \frac{1}{\beta} \int_{-1}^{1} \psi_n^c(t)\, dt \begin{pmatrix} \int_{-\beta}^{-1} \exp(a\beta st)\, \psi_n^a(s)\, ds \\ + \int_{1}^{\beta} \exp(a\beta st)\, \psi_n^a(s)\, ds \end{pmatrix} \\
= & \int_{-1}^{1} \psi_n^c(t) \left(\psi_n^c(-1) \exp(-ct) + \psi_n^c(1) \exp(ct) \right) dt \\
= & \psi_n^c(-1) \int_{-1}^{1} \psi_n^c(t) \exp(-ct)\, dt \\
& + \psi_n^c(1) \int_{-1}^{1} \psi_n^c(t) \exp(ct)\, dt \\
= & \lambda_n(c) \left(\psi_n^c(-1) \right)^2 + \lambda_n(c) \left(\psi_n^c(1) \right)^2 \\
= & 2\lambda_n(c) \left(\psi_n^c(1) \right)^2, \qquad (220)
\end{aligned}$$

where the first equality takes the limit as $\beta \to 1$, the second equality uses the linearity of integration, the third equality uses the facts that $\lambda_n(c)\, \psi_n^c(-1) = \int_{-1}^{1} \psi_n^c(t) \exp(-ct)\, dt$ (by the eigenvalue-eigenfunction identity evaluated at $x = -1$) and that $\lambda_n(c)\, \psi_n^c(1) = \int_{-1}^{1} \psi_n^c(t) \exp(ct)\, dt$ (by the eigenvalue-eigenfunction identity evaluated at $x = 1$), and the fourth equality follows from the fact that $|\psi_n^c(-1)| = |\psi_n^c(1)|$ (since $\psi_n^c(x)$ is even if n is even and odd if n is odd).

□

In the prolate case, Osipov *et al.* (2013) provide a differential equation (in c) for the eigenvalues of (the prolate version of) L_c (this is their Theorem 7.12 on page 240); I show below that an analogous differential equation, which involves a different sign, holds in the oblate case as well.

Theorem 8.8. *For $c > 0$ and $n \geq 0$,*

$$\frac{d\chi_n}{dc}(c) = -2c \int_{-1}^{1} x^2 \left(\psi_n^c(x) \right)^2 dx.$$

Proof. Differentiate the eigenvalue-eigenfunction identity $L_c[\psi_n^c](x) - \chi_n(c)\, \psi_n^c(x) = 0$ with respect to c to obtain:

$$\begin{aligned}
& 0 \\
= & \frac{d}{dc} \left\{ -\frac{d}{dx} \left(\left(1 - x^2\right) \frac{d\psi_n^c}{dx}(x) \right) - c^2 x^2 \psi_n^c(x) - \chi_n(c)\, \psi_n^c(x) \right\} \qquad (221) \\
= & -\frac{d}{dx} \left(\left(1 - x^2\right) \frac{d}{dx} \left(\frac{d\psi_n^c}{dc}(x) \right) \right) \\
& -2cx^2 \psi_n^c(x) - c^2 x^2 \frac{d\psi_n^c}{dc}(x) \\
& -\frac{d\chi_n}{dc}(c)\, \psi_n^c(x) - \chi_n(c) \frac{d\psi_n^c}{dc}(x) \qquad (222)
\end{aligned}$$

$$
\begin{aligned}
&= \left(L_c\left[\frac{d\psi_n^c}{dc}\right](x) - \chi_n(c)\frac{d\psi_n^c}{dc}(x)\right) \\
&\quad - \left(2cx^2 + \frac{d\chi_n}{dc}(c) + \chi_n(c)\right)\psi_n^c(x). \qquad (223)
\end{aligned}
$$

Thus, multiplying (223) by $\psi_n^c(x)$ is equal to zero for all $x \in (-1, 1)$ and integrating this from -1 to 1 results in zero, so that

$$
\begin{aligned}
&0 \\
&= \int_{-1}^{1} \left\{ \begin{array}{c} \left(L_c\left[\frac{d\psi_n^c}{dc}\right](x) - \chi_n(c)\frac{d\psi_n^c}{dc}(x)\right) \\ -\left(2cx^2 + \frac{d\chi_n}{dc}(c) + \chi_n(c)\right)\psi_n^c(x) \end{array} \right\} \psi_n^c(x)\,dx \qquad (224) \\
&= \int_{-1}^{1} \left(L_c\left[\frac{d\psi_n^c}{dc}\right](x) - \chi_n(c)\frac{d\psi_n^c}{dc}(x)\right)\psi_n^c(x)\,dx \\
&\quad -2c\int_{-1}^{1} x^2\psi_n^c(x)\,\psi_n^c(x)\,dx - \frac{d\chi_n}{dc}(c)\int_{-1}^{1}\psi_n^c(x)\,\psi_n^c(x)\,dx \qquad (225) \\
&= \int_{-1}^{1} \left(L_c\left[\frac{d\psi_n^c}{dc}\right](x) - \chi_n(c)\frac{d\psi_n^c}{dc}(x)\right)\psi_n^c(x)\,dx \\
&\quad -2c\int_{-1}^{1} x^2\left(\psi_n^c(x)\right)^2 dx - \frac{d\chi_n}{dc}(c)\int_{-1}^{1}\left(\psi_n^c(x)\right)^2 dx \qquad (226) \\
&= \int_{-1}^{1} \left(L_c\left[\frac{d\psi_n^c}{dc}\right](x) - \chi_n(c)\frac{d\psi_n^c}{dc}(x)\right)\psi_n^c(x)\,dx \\
&\quad -2c\int_{-1}^{1} x^2\left(\psi_n^c(x)\right)^2 dx - \frac{d\chi_n}{dc}(c) \qquad (227) \\
&= \int_{-1}^{1} \frac{d\psi_n^c}{dc}(x)\left(L_c\left[\psi_n^c\right](x) + \chi_n(c)\,\psi_n^c(x)\right)dx \\
&\quad -2c\int_{-1}^{1} x^2\left(\psi_n^c(x)\right)^2 dx - \frac{d\chi_n}{dc}(c) \qquad (228) \\
&= -2c\int_{-1}^{1} x^2\left(\psi_n^c(x)\right)^2 dx - \frac{d\chi_n}{dc}(c), \qquad (229)
\end{aligned}
$$

where the first equality is by the argument immediately above this display, the second equality is by a regrouping of terms and the linearity of integration, the third equality simply recognizes that $A \times A = A^2$, the fourth equality is by the orthonormality of the $\psi_n^c(x)$, the fifth equality is by the self-adjointness (and thus symmetry) of L_c (see Theorem 8.2 and Corollary 8.1; the symmetry of L_c leads immediately to the symmetry of $L_c - \chi_n(c)$), and the sixth equality holds because $L_c[\psi_n^c](x) - \chi_n(c)\,\psi_n^c(x) = 0$ (this is simply the eigenvalue-eigenfunction identity for $\psi_n^c(x)$ and $\chi_n(c)$, as noted above). □

I now provide an oblate analog of the prolate-case bounds for $\chi_n(c)$ developed by Osipov *et al.* (2013) (Corollary 7.3 on page 241). Note that the interval in the oblate case is to the left of $n(n+1)$, while in the prolate case it is to the right of $n(n+1)$.

Corollary 8.2.

$$\chi_n(c) \in \left[n(n+1) - c^2, n(n+1)\right]$$

Proof. As $c > 0$ approaches zero (from above), the oblate spheroidal wave equation of order zero approaches the Legendre differential equation, so $\chi_n(c) \to n(n+1)$. Thus,

$$\chi_n(c) = n(n+1) + \int_0^c \frac{d\chi_n}{da}(a)\,da \tag{230}$$

$$= n(n+1) - \int_0^c 2a \int_{-1}^1 x^2 \left(\psi_n^a(x)\right)^2 dxda, \tag{231}$$

where the second equality follows from Theorem 8.8. Since $\psi_m^a(x)$ satisfies $\int_{-1}^1 \left(\psi_m^a(x)\right)^2 dx = 1$ for any $a > 0$ and for any $m = 0, 1, 2, \ldots$ by the orthonormality of the $\psi_m^a(x)$, I conclude that

$$0 \le \int_{-1}^1 x^2 \left(\psi_n^a(x)\right)^2 dx \le \int_{-1}^1 \left(\psi_m^a(x)\right)^2 dx = 1. \tag{232}$$

Substituting these inequalities into the integral in (231), I obtain the following inequalities:

$$\int_0^c 2a \int_{-1}^1 x^2 \left(\psi_n^a(x)\right)^2 dxda \tag{233}$$

$$\le \int_0^c 2ada \tag{234}$$

$$= c^2 \tag{235}$$

and

$$\int_0^c 2a \int_{-1}^1 x^2 \left(\psi_n^a(x)\right)^2 dxda \tag{236}$$

$$\ge \int_0^c 0da \tag{237}$$

$$= 0. \tag{238}$$

Substituting these bounds for the integral $\int_0^c 2a \int_{-1}^1 x^2 \left(\psi_n^a(x)\right)^2 dxda$ into the expression (231), I complete the proof by concluding that

$$\chi_n(c) \in \left[n(n+1) - c^2, n(n+1)\right]. \tag{239}$$

□

The lower bound provided in Corollary 8.2 allows me to give a simple bound (in terms of c) such that $\chi_n(c) \geq c^2$ for any n at or above this bound.

Corollary 8.3. *If $n \geq \sqrt{2}c$, then $\chi_n(c) \geq c^2$.*

Proof. If $n \geq \sqrt{2}c$, then $n(n+1) \geq n^2 \geq 2c^2$. By Corollary 8.2, $\chi_n(c) \geq n(n+1) - c^2$, so $n \geq \sqrt{2}c$ implies that $\chi_n(c) \geq n(n+1) - c^2 \geq n^2 - c^2 \geq c^2$. □

In measuring the accuracy with which I approximate the kernel $K(x,t) = \exp(cxy)$, it will be helpful to localize the maxima of the $\psi_n^c(x)$. In the prolate case, Osipov *et al.* (2013) provide (in their Lemma 4.14 on page 119 and Theorem 4.39 on page 121) very useful results in this direction. As noted by Szegö (1975) (page 166, including footnote 43 on that page), the general version of the key result given by Osipov *et al.* (2013) goes all the way back to Sonin and, in even greater generality, to Pólya; it is often termed the Sonin-Pólya theorem or the Sonin-Pólya-Butlewski theorem. I show below that an oblate analog of this key result also holds.

Lemma 8.1. *For $c > 0$, $n = 0, 1, 2, \ldots$, define the function Q as*

$$Q(x) \equiv (\psi_n^c(x))^2 + \frac{1-x^2}{\chi_n(c) + c^2x^2}\left(\frac{d\psi_n^c}{dx}(x)\right)^2.$$

Then Q is increasing for $x \in \left(\max\left\{0, \sqrt{\frac{1}{2}\left(1 - \frac{\chi_n(c)}{c^2}\right)}\right\}, 1\right]$, so that $y > x$ and $x, y \in \left[\max\left\{0, \sqrt{\frac{1}{2}\left(1 - \frac{\chi_n(c)}{c^2}\right)}\right\}, 1\right]$ together imply that $Q(y) > Q(x)$.

Proof. Differentiate Q to obtain, for all $x \in (-1, 1)$,

$$\begin{aligned}
&\frac{dQ}{dx}(x) \\
= \;& 2\psi_n^c(x)\frac{d\psi_n^c}{dx}(x) \\
&-2\frac{c^2x\left(1-x^2\right) + x\left(\chi_n(c) + c^2x^2\right)}{\left(\chi_n(c) + c^2x^2\right)^2}\left(\frac{d\psi_n^c}{dx}(x)\right)^2 \\
&+2\frac{\left(1-x^2\right)}{\chi_n(c)+c^2x^2}\frac{d\psi_n^c}{dx}(x)\frac{d^2\psi_n^c}{dx^2}(x) \qquad (240)\\
= \;& \frac{2}{\chi_n(c)+c^2x^2}\frac{d\psi_n^c}{dx}(x)\left\{\begin{array}{c} \psi_n^c(x)\left(\chi_n(c)+c^2x^2\right) \\ -\frac{c^2x\left(1-x^2\right)}{\chi_n(c)+c^2x^2}\frac{d\psi_n^c}{dx}(x) \\ -x\frac{d\psi_n^c}{dx}(x) + \left(1-x^2\right)\frac{d^2\psi_n^c}{dx^2}(x)\end{array}\right\} \qquad (241)\\
= \;& \frac{2}{\chi_n(c)+c^2x^2}\frac{d\psi_n^c}{dx}(x)\left\{\begin{array}{c} \psi_n^c(x)\left(\chi_n(c)+c^2x^2\right) \\ -\frac{c^2x\left(1-x^2\right)}{\chi_n(c)+c^2x^2}\frac{d\psi_n^c}{dx}(x) + x\frac{d\psi_n^c}{dx}(x) \\ -2x\frac{d\psi_n^c}{dx}(x) + \left(1-x^2\right)\frac{d^2\psi_n^c}{dx^2}(x)\end{array}\right\} \qquad (242)
\end{aligned}$$

$$= \frac{2}{\chi_n(c)+c^2x^2}\frac{d\psi_n^c}{dx}(x)\left\{\begin{array}{c} -\left(L_c\left[\psi_n^c\right](x)-\chi_n(c)\,\psi_n^c(x)\right) \\ -\frac{c^2x\left(1-x^2\right)}{\chi_n(c)+c^2x^2}\frac{d\psi_n^c}{dx}(x)+x\frac{d\psi_n^c}{dx}(x) \end{array}\right\} \tag{243}$$

$$= \frac{2\left(x\left(\chi_n(c)+c^2x^2\right)-c^2x\left(1-x^2\right)\right)}{\left(\chi_n(c)+c^2x^2\right)^2}\left(\frac{d\psi_n^c}{dx}(x)\right)^2 \tag{244}$$

$$= \frac{2x}{\left(\chi_n(c)+c^2x^2\right)^2}\left(\frac{d\psi_n^c}{dx}(x)\right)^2\left(\chi_n(c)+2c^2x^2-c^2\right), \tag{245}$$

where the first equality follows from the quotient rule of differentiation, the second equality follows by recognizing the common factor $\frac{2}{\chi_n(c)+c^2x^2}\frac{d\psi_n^c}{dx}(x)$, the third equality holds by addition and subtraction of $x\frac{d\psi_n^c}{dx}(x)$ within the curly braces, the fourth equality follows from regrouping terms and the definition of L_c (see Definition 8.2), the fifth equality follows because $L_c\left[\psi_n^c\right](x)-\chi_n(c)\,\psi_n^c(x)=0$ (this is simply the eigenvalue-eigenfunction identity for $\psi_n^c(x)$ and $\chi_n(c)$), and the sixth equality simply regroups terms.

The argument of the preceding paragraph shows that $\frac{dQ}{dx}(x)>0$ for $x\in(a,1]$ (with $a\geq 0$) if and only if $\chi_n(c)+2c^2x^2-c^2>0$ for $x\in(a,1]$. The latter inequality certainly holds if $a=\max\left\{0,\sqrt{\frac{1}{2}\left(1-\frac{\chi_n(c)}{c^2}\right)}\right\}$, since then

$$\begin{aligned} & \chi_n(c)+2c^2a^2-c^2 \\ = \;& \chi_n(c)+2c^2\max\left\{0,\frac{1}{2}\left(1-\frac{\chi_n(c)}{c^2}\right)\right\}-c^2 && (246) \\ = \;& \chi_n(c)+\max\left\{0,c^2-\chi_n(c)\right\}-c^2 && (247) \\ = \;& \max\left\{\chi_n(c)-c^2,0\right\}, && (248) \end{aligned}$$

so $\chi_n(c)+2c^2x^2-c^2>0$ for any $x>a$. Finally, $\frac{dQ}{dx}(a)\geq 0$ with a defined as above, as is clear from the argument just given.

□

Theorem 8.9. *If $\chi_n(c)\geq c^2$, then for any $x\in(-1,1)$, $\left|\psi_n^c(x)\right|<\left|\psi_n^c(1)\right|=\left|\psi_n^c(-1)\right|$. Thus, $\left|\psi_n^c(x)\right|$ attains its maximum over $x\in[-1,1]$ at $x=1$ and also at $x=-1$, and is strictly less than this maximum for any $x\in(-1,1)$.*

Proof. According to Lemma 8.1, if $\max\left\{0,\sqrt{\frac{1}{2}\left(1-\frac{\chi_n(c)}{c^2}\right)}\right\}=0$, that is, if $\chi_n(c)\geq c^2$, then $Q(x)$ is increasing on $(0,1]$. By the definition of Q in Lemma 8.1, $Q(x)\geq\left(\psi_n^c(x)\right)^2$ for all $x\in[-1,1]$ and $Q(1)=\left(\psi_n^c(1)\right)^2$. If $\chi_n(c)\geq c^2$, so that $x,y\in[0,1]$ and $y>x$ together imply that $Q(y)>Q(x)$ by Lemma 8.1, then $\left(\psi_n^c(x)\right)^2\leq Q(x)<Q(1)=\left(\psi_n^c(1)\right)^2$ for all $x\in[0,1)$. Because $\psi_n^c(x)$ is an even function for n even and is an odd function for n odd, $\left(\psi_n^c(x)\right)^2$ is an even function for $n=0,1,2,\ldots$. Thus, $\left(\psi_n^c(x)\right)^2<\left(\psi_n^c(1)\right)^2=\left(\psi_n^c(-1)\right)^2$ for all $x\in(-1,1)$; taking the (postive) square root of the inequality and equality in this expression, I obtain the desired conclusion. □

The following corollary provides a simple lower bound (in terms of c) on n such that, for any n above this bound, $\chi_n(c) \geq c^2$.

Corollary 8.4. *If $n \geq \sqrt{2}c$, then $|\psi_n^c(x)|$ attains its maximum over $x \in [-1,1]$ at $x = 1$ and also at $x = -1$, and is strictly less than this maximum for any $x \in (-1,1)$.*

Proof. If $n \geq \sqrt{2}c$, then Corollary 8.3 implies that $\chi_n(c) \geq c^2$; this is the hypothesis of Theorem 8.9, which then delivers the desired conclusion. $\square$

In the next theorem, I provide a means of measuring the worst-case approximation error incurred by approximating the kernel $\exp(cxt)$ using its first $N+1$ eigenfunctions. Although I suspect that some form of this result must exist in the literature, I have not been able to find it; in the extensive literature on the prolate case, emphasis is instead placed on proving decay rates for the $\lambda_n(c)$ (see, for example, Landau & Widom (1980) and Osipov *et al.* (2013)).

Theorem 8.10. *The error $\left|\exp(cxt) - \sum_{n=0}^{N} \lambda_n(c)\, \psi_n^c(x)\, \psi_n^c(t)\right|$ converges uniformly (over $x, t \in [-1,1]$) to zero as $N \to \infty$. If $\chi_m(c) \geq c^2$ for some integer $m \geq 0$, then for any $N = m, m+1, m+2, \ldots$ this error is maximized over $x, t \in [-1,1]$ at $x = t = -1$ and also at $x = t = 1$.*

Proof. The uniform convergence of the error $\left|\exp(cxt) - \sum_{n=0}^{N} \lambda_n(c)\, \psi_n^c(x)\, \psi_n^c(t)\right|$ to zero as $N \to \infty$ is a direct consequence of Mercer's theorem, since the kernel $\exp(cxt)$ is continuous on $[-1,1] \times [-1,1]$ and is strictly positive definite (see Corollary 5.1).

Suppose that $\chi_m(c) \geq c^2$ for some integer $m \geq 0$, and assume that N is an integer with $N \geq m$. For any $x, t \in [-1,1]$,

$$
\begin{aligned}
&\left|\exp(cxt) - \sum_{n=0}^{N} \lambda_n(c)\, \psi_n^c(x)\, \psi_n^c(t)\right| \\
&= \left|\sum_{n=0}^{\infty} \lambda_n(c)\, \psi_n^c(x)\, \psi_n^c(t) - \sum_{n=0}^{N} \lambda_n(c)\, \psi_n^c(x)\, \psi_n^c(t)\right| && (249) \\
&= \left|\sum_{n=N+1}^{\infty} \lambda_n(c)\, \psi_n^c(x)\, \psi_n^c(t)\right| && (250) \\
&\leq \sum_{n=N+1}^{\infty} \left|\lambda_n(c)\, \psi_n^c(x)\, \psi_n^c(t)\right| && (251) \\
&= \sum_{n=N+1}^{\infty} \lambda_n(c)\, \left|\psi_n^c(x)\right| \left|\psi_n^c(t)\right| && (252) \\
&\leq \sum_{n=N+1}^{\infty} \lambda_n(c)\, \left|\psi_n^c(1)\right| \left|\psi_n^c(1)\right| && (253)
\end{aligned}
$$

$$= \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(1)\right)^2 \tag{254}$$

$$= \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(-1)\right)^2, \tag{255}$$

where the first equality is by Mercer's theorem (the infinite sum converges absolutely and uniformly), the second equality is clear (since $A + B - A = B$), the first inequality is by absolute convergence and the triangle inequality, the third equality is by $\lambda_n(c) > 0$ for $n = 0, 1, 2, \ldots$ (see Corollary 5.1), the second inequality follows from applying Theorem 8.9 to each term of the series in the prior expression (using the facts that $\lambda_n(c) > 0$ for $n = 0, 1, 2, \ldots$ and that $\chi_m(c) > c^2$ implies $\chi_k(c) > c^2$ for every integer $k \geq m$, since the $\chi_k(c)$ form an increasing sequence in k by Theorem 8.6), the fourth equality is clear (for any real number A, $|A|\,|A| = A^2$), and the fifth equality is by Theorem 8.9.

Now consider the error term evaluated at $x = t = 1$:

$$\left|\exp(c) - \sum_{n=0}^{N} \lambda_n(c)\,\psi_n^c(1)\,\psi_n^c(1)\right|$$

$$= \left|\sum_{n=0}^{\infty} \lambda_n(c)\,\psi_n^c(1)\,\psi_n^c(1) - \sum_{n=0}^{N} \lambda_n(c)\,\psi_n^c(1)\,\psi_n^c(1)\right| \tag{256}$$

$$= \left|\sum_{n=N+1}^{\infty} \lambda_n(c)\,\psi_n^c(1)\,\psi_n^c(1)\right| \tag{257}$$

$$= \left|\sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(1)\right)^2\right| \tag{258}$$

$$= \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(1)\right)^2, \tag{259}$$

where the first equality is by Mercer's theorem (the infinite sum converges absolutely and uniformly), the second equality is clear (since $A+B-A = B$), the third equality simply recognizes that $A \times A = A^2$, and the fourth equality holds because every term of the series in the preceding expression is nonnegative. An identical argument shows that the error evaluated at $x = t = -1$ satisifies

$$\left|\exp(c) - \sum_{n=0}^{N} \lambda_n(c)\,\psi_n^c(-1)\,\psi_n^c(-1)\right|$$

$$= \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(-1)\right)^2. \tag{260}$$

Combining the results of the two preceding paragraphs (which requires the assumption that $\chi_m(c) \geq c^2$ for some integer $m \geq 0$ and that N is an integer with $N \geq m)$, I obtain

$$
\begin{aligned}
& \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(1)\right)^2 \\
= & \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(-1)\right)^2 && (261)\\
\geq & \sup_{x,t\in[-1,1]}\left\{\left|\exp(cxt)-\sum_{n=0}^{N}\lambda_n(c)\,\psi_n^c(x)\,\psi_n^c(t)\right|\right\} && (262)\\
\geq & \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(1)\right)^2 && (263)\\
= & \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(-1)\right)^2, && (264)
\end{aligned}
$$

where the first equality is a restatement of (255), the first inequality is from the display beginning with (249), the second inequality is by the fact that a supremum must (by definition) be greater than or equal to the value assumed at any point in the set over which the supremum is taken (and the values at the points $x = t = 1$ and $x = t = -1$ are given by (259) and (260), respectively), and the final equality simply restates (255) yet again. Thus,

$$
\begin{aligned}
& \sup_{x,t\in[-1,1]}\left\{\left|\exp(cxt)-\sum_{n=0}^{N}\lambda_n(c)\,\psi_n^c(x)\,\psi_n^c(t)\right|\right\} \\
= & \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(1)\right)^2 && (265)\\
= & \sum_{n=N+1}^{\infty} \lambda_n(c)\left(\psi_n^c(-1)\right)^2, && (266)
\end{aligned}
$$

which completes the proof of the theorem. □

As I did in obtaining Corollary 8.4 from Theorem 8.9, I can use Corollary 8.3 to provide a simple sufficient condition assuring that $\chi_n(c) \geq c^2$ and thus making the conclusion of Theorem 8.10 available.

Corollary 8.5. *If $N \geq \sqrt{2}c$, then the error $\left|\exp(cxt) - \sum_{n=0}^{N} \lambda_n(c)\,\psi_n^c(x)\,\psi_n^c(t)\right|$ is maximized over $x, t \in [-1, 1]$ at $x = t = -1$ and also at $x = t = 1$.*

Proof. By Corollary 8.3, $N \geq \sqrt{2}c$ implies that $\chi_N(c) \geq c^2$; Theorem 8.10 then delivers the desired conclusion. □

8.4 Numerical Methods for Oblates and for the Oblate Approximation to the Exponential Product Operator

Numerical and computational approaches to spheroidal wave functions have a long history; the similarity of L_c to the Legendre differential operator, which has already been noted above, makes it natural to consider representing spheroidal wave functions of order zero (such as $\psi_n^c(x)$, the eigenfunctions of L_c) as series of Legendre polynomials. (When the order of the spheroidal wave function is not zero, associated Legendre functions, sometimes called "associated Legendre polynomials" despite the fact that they are not, in general, polynomials, are employed.) This is the typical method in the literature, with a continued-fraction approach to determining the series coefficients dating back to Bouwkamp (1947) (as described in Flammer (1957)). A more modern approach, introduced by Hodge (1970), recasts the problem of finding the series coefficients as two symmetric tridiagonal eigenproblems (one for even indices, the other for odd indices). Since methods to obtain the eigenvalues and eigenvectors of symmetric tridiagonal matrices form a well-developed area within numerical linear algebra, the problem as recast is numerically quite tractable.

Osipov *et al.* (2013) provide a helpful description of the tridiagonal-eigensolver method for Legendre polynomial expansions of prolate spheroidal wave functions of order zero (similar descriptions appear in a number of papers of these authors and their coauthors). With a simple sign change, their description can be used in the oblate case as well; I follow them in laying out the tridiagonal-eigensolver method for Legendre polynomial expansions of oblate spheroidal wave functions of order zero (the $\psi_n^c(x)$) below. Of course, Hodge (1970) also gives a (more concise) treatment which covers the oblate case as well.

To emphasize, the material in this subsection that deals with constructing series expansions for the oblates is not new, though I try to lay it out clearly and completely; in contrast, the material concerning the use of oblates to construct an approximation to the exponential product operator is novel.

As in Osipov *et al.* (2013), it is useful to describe the Legendre polynomials before using them as the building blocks in series representations of the $\psi_n^c(x)$. The Legendre polynomials P_k, $k = 0, 1, 2, \ldots$ begin with $P_0(x) = 1$ and $P_1(x) = x$; with these initial polynomials fixed, the remaining of the Legendre polynomials are defined via a three-term recurrence relation that can be used to express P_{k+1} in terms of P_k and P_{k-1}:

$$(k+1)\, P_{k+1}(x) - (2k+1)\, x P_k(x) + k P_{k-1}(x) \quad = \quad 0. \tag{267}$$

P_k is an even function when k is even and is an odd function when k is odd; further,

$$P_k(1) \quad = \quad 1 \quad (k = 0, 1, 2, \ldots). \tag{268}$$

Although $\{P_k\}_{k=0}^{\infty}$ make up a complete orthogonal system in $L^2[-1,1]$, they satisfy

$$\int_{-1}^{1} (P_k(x))^2\, dx \quad = \quad \frac{1}{k+\frac{1}{2}}, \tag{269}$$

and are thus not orthonormal.

I follow the notation of Osipov *et al.* (2013) in defining the normalized Legendre polynomials:

$$\overline{P_k}\left(x\right) \quad \equiv \quad P_k\left(x\right)\sqrt{k+\frac{1}{2}}. \tag{270}$$

This normalization yields

$$\int_{-1}^{1}\left(\overline{P_k}\left(x\right)\right)^2 dx \quad = \quad 1, \tag{271}$$

so the $\left\{\overline{P_k}\right\}_{k=0}^{\infty}$ make up a complete orthonormal system in $L^2\left[-1,1\right]$.

To make use of these normalized Legendre polynomials, I require a method to evaluate them and to evaluate their first derivatives.

Method 8.1. This method computes the result of evaluating the normalized Legendre polynomials $\overline{P_j}$ for $j = 0, 1, \ldots n$ at each of the elements of a column vector $\boldsymbol{x}$.

- **Given**:
  1. n; the normalized Legendre polynomials $\overline{P_k}$ for $k = 0, 1, \ldots, n$ will be evaluated at the given points
  2. $\boldsymbol{x}$, a column vector of points at which to evaluate the normalized Legendre polynomials
- **Output**:
  1. V_{Leg}, a matrix with the same number of rows as there are elements in $\boldsymbol{x}$ and with $n+1$ columns whose (i, j) element (for $i, j = 0, 1, 2, \ldots$, so that indexing begins with 0) is $\overline{P_j}\left(x_i\right)$

Simply use the three-term recurrence and then normalize.

Remark 8.2. The use of three-term recurrences for computational purposes is sometimes problematic (Gautschi (1967)), but "the basic three-term recurrence relation for orthogonal polynomials is generally an excellent means of computing these polynomials" (Gautschi (1993)). Zhang & Jin (1996) present a method similar to Method 8.1 in Section 2 of their Chapter 4.

Method 8.2. This method computes the result of evaluating the first derivatives of the normalized Legendre polynomials $\overline{P_j}$ for $j = 0, 1, \ldots n$ at each of the elements of a column vector $\boldsymbol{x}$.

- **Given**:
  1. n; the first derivatives of the normalized Legendre polynomials $\overline{P_k}$ for $k = 0, 1, \ldots, n$ will be evaluated at the given points

2. $\boldsymbol{x}$, a column vector of points at which to evaluate the first derivatives of the normalized Legendre polynomials

- **Output**:

1. V_{LegDeriv}, a matrix with the same number of rows as there are elements in $\boldsymbol{x}$ and with $n+1$ columns whose (i,j) element (for $i,j = 0,1,2,\ldots$, so that indexing begins with 0) is $\frac{d\overline{P_j}}{dx}(x_i)$

Differentiate the three-term recurrence and use the resulting equation, then normalize.

Having in hand methods to evaluate the normalized Legendre polynomials and their derivatives, I now consider how to use them to represent the oblates. Because of the completeness of the normalized Legendre polynomials in $L^2[-1,1]$, I can expand (for any nonnegative integer n and for any $c > 0$) an oblate in terms of $\left\{\overline{P_k}\right\}_{k=0}^{\infty}$:

$$\psi_n^c(x) = \sum_{k=0}^{\infty} \beta_k^{(n)} \overline{P_k}(x). \tag{272}$$

The fact that $\left\{\overline{P_k}\right\}_{k=0}^{\infty}$ form a complete orthonormal system in $L^2[-1,1]$ would deliver only convergence in the $L^2[-1,1]$ norm in (272) (that is, the partial sums of terms from $k=0$ to $k=N$ of the infinite sum on the righthand-side would converge to the lefthand-side only in the $L^2[-1,1]$ norm as $N \to \infty$). However, because $\psi_n^c(x)$ is an analytic function of $x \in \mathbb{R}$ (by Theorem 8.4), I conclude that the convergence in (272) is pointwise and, indeed, uniform on $[-1,1]$. Further, the orthonormality of $\left\{\overline{P_k}\right\}_{k=0}^{\infty}$ combines with the fact that $\int_{-1}^{1} (\psi_n^c(x))^2\, dx = 1$ (by definition, since ψ_n^c is an eigenfunction of a linear operator whose domain is contained in $L^2[-1,1]$) to imply that

$$\sum_{k=0}^{\infty} \left(\beta_k^{(n)}\right)^2 = 1 \quad (n = 0,1,2,\ldots). \tag{273}$$

By Theorem 8.6, $\psi_n^c(x)$ is the eigenfunction of L_c associated with the n^{th}-smallest eigenvalue of L_c, $\chi_n(c)$ (where my numbering begins with $n=0$ for the smallest eigenvalue). Substitute (272) into the eigenvalue-eigenfunction identity for $\psi_n^c(x)$ as an eigenfunction of the operator L_c ,

$$\chi_n(c)\,\psi_n^c(x) = L_c[\psi_n^c](x), \tag{274}$$

to obtain:

$$
\begin{aligned}
& \chi_n(c) \sum_{k=0}^{\infty} \beta_k^{(n)} \overline{P_k}(x) \\
= \; & L_c\left[\sum_{k=0}^{\infty} \beta_k^{(n)} \overline{P_k}\right](x) && (275) \\
= \; & \sum_{k=0}^{\infty} \beta_k^{(n)} L_c\left[\overline{P_k}\right](x) && (276) \\
= \; & \sum_{k=0}^{\infty} \beta_k^{(n)} \sqrt{k+\frac{1}{2}} L_c\left[P_k\right](x) && (277) \\
= \; & \sum_{k=0}^{\infty} \beta_k^{(n)} \sqrt{k+\frac{1}{2}}\left\{-\frac{d}{dx}\left(\left(1-x^2\right) \frac{dP_k}{dx}(x)\right)-c^2 x^2 P_k(x)\right\}, && (278)
\end{aligned}
$$

where the first equality follows from substitution of (272) into (274) (and the uniform convergence of the series expansion as noted above), the second equality is by the linearity of L_c and the uniform convergence of the series expansion, the third equality is by the normalization (270) and again uses the linearity of L_c, and the final equality is by the definition of L_c (see Definition 8.2).

The term inside the curly braces for index $k \geq 2$ in (278) is:

$$
\begin{aligned}
& -\frac{d}{dx}\left(\left(1-x^2\right) \frac{dP_k}{dx}(x)\right)-c^2 x^2 P_k(x) \\
= \; & k(k+1) P_k(x)-c^2 x^2 P_k(x) && (279) \\
= \; & k(k+1) P_k(x)-c^2 x \frac{1}{2k+1}\left((k+1) P_{k+1}(x)+k P_{k-1}(x)\right) && (280) \\
= \; & k(k+1) P_k(x)-\frac{c^2(k+1)}{2k+1} x P_{k+1}(x)-\frac{c^2 k}{2k+1} x P_{k-1}(x) && (281) \\
= \; & k(k+1) P_k(x) \\
& -\frac{c^2(k+1)}{(2k+1)(2(k+1)+1)}\left((k+2) P_{k+2}(x)+(k+1) P_k(x)\right) \\
& -\frac{c^2 k}{(2k+1)(2(k-1)+1)}\left(k P_k(x)+(k-1) P_{k-2}(x)\right) && (282) \\
= \; & \left(k(k+1)-\frac{c^2}{2k+1}\left(\frac{(k+1)^2}{2k+3}+\frac{k^2}{2k-1}\right)\right) P_k(x) \\
& -\frac{c^2(k+1)(k+2)}{(2k+1)(2k+3)} P_{k+2}(x) \\
& -\frac{c^2 k(k-1)}{(2k+1)(2k-1)} P_{k-2}(x) && (283) \\
= \; & \left(k(k+1)-c^2 \frac{2k(k+1)-1}{(2k+3)(2k-1)}\right) P_k(x)
\end{aligned}
$$

$$
\begin{aligned}
&-c^2\frac{(k+1)(k+2)}{(2k+1)(2k+3)}P_{k+2}(x) \\
&-c^2\frac{k(k-1)}{(2k+1)(2k-1)}P_{k-2}(x), \qquad (284)
\end{aligned}
$$

where the first equality follows from the fact that P_k is an eigenfunction of the operator $-\frac{d}{dx}\left(\left(1-x^2\right)\frac{df}{dx}(x)\right)$ (which has domain D_L, as does L_c) with eigenvalue $k(k+1)$ (for any $k = 0, 1, 2, \ldots$), the second equality is by the three-term recurrence relation (267) (recalling that $k \geq 2$), the third equality simply rearranges terms, the fourth equality applies the three term recurrence relation (267) twice (once to P_{k+1} and once to P_{k-1}) (again, recall that $k \geq 2$ here), the fifth equality regroups terms, and the sixth equality is by elementary algebra.

Since the term inside the curly braces for index k in (278) is multiplied by $\sqrt{k+\frac{1}{2}}$, I now analyze (for $k \geq 2$):

$$
\begin{aligned}
&\sqrt{k+\frac{1}{2}}\left\{-\frac{d}{dx}\left(\left(1-x^2\right)\frac{dP_k}{dx}(x)\right)-c^2x^2P_k(x)\right\} \\
= \ &\sqrt{k+\frac{1}{2}}\left(k(k+1)-c^2\frac{2k(k+1)-1}{(2k+3)(2k-1)}\right)P_k(x) \\
&-\sqrt{k+\frac{1}{2}}c^2\frac{(k+1)(k+2)}{(2k+1)(2k+3)}P_{k+2}(x) \\
&-\sqrt{k+\frac{1}{2}}c^2\frac{k(k-1)}{(2k+1)(2k-1)}P_{k-2}(x) \qquad (285) \\
= \ &\left(k(k+1)-c^2\frac{2k(k+1)-1}{(2k+3)(2k-1)}\right)\sqrt{k+\frac{1}{2}}P_k(x) \\
&-c^2\frac{(k+1)(k+2)}{(2k+1)(2k+3)}\frac{\sqrt{k+\frac{1}{2}}}{\sqrt{k+\frac{5}{2}}}\sqrt{k+\frac{5}{2}}P_{k+2}(x) \\
&-c^2\frac{k(k-1)}{(2k+1)(2k-1)}\frac{\sqrt{k+\frac{1}{2}}}{\sqrt{k-\frac{3}{2}}}\sqrt{k-\frac{3}{2}}P_{k-2}(x) \qquad (286) \\
= \ &\left(k(k+1)-c^2\frac{2k(k+1)-1}{(2k+3)(2k-1)}\right)\overline{P_k}(x) \\
&-c^2\frac{(k+1)(k+2)}{(2k+1)(2k+3)}\frac{\sqrt{k+\frac{1}{2}}}{\sqrt{k+\frac{5}{2}}}\overline{P_{k+2}}(x) \\
&-c^2\frac{k(k-1)}{(2k+1)(2k-1)}\frac{\sqrt{k+\frac{1}{2}}}{\sqrt{k-\frac{3}{2}}}\overline{P_{k-2}}(x) \qquad (287)
\end{aligned}
$$

$$
\begin{aligned}
&= \left(k(k+1) - c^2 \frac{2k(k+1)-1}{(2k+3)(2k-1)}\right) \overline{P_k}(x) \\
&\quad -c^2 \frac{(k+1)(k+2)}{(2k+1)(2k+3)} \frac{\sqrt{2k+1}}{\sqrt{2k+5}} \overline{P_{k+2}}(x) \\
&\quad -c^2 \frac{k(k-1)}{(2k+1)(2k-1)} \frac{\sqrt{2k+1}}{\sqrt{2k-3}} \overline{P_{k-2}}(x) \qquad (288) \\
&= \left(k(k+1) - c^2 \frac{2k(k+1)-1}{(2k+3)(2k-1)}\right) \overline{P_k}(x) \\
&\quad -c^2 \frac{(k+1)(k+2)}{(2k+3)\sqrt{(2k+1)(2k+5)}} \overline{P_{k+2}}(x) \\
&\quad -c^2 \frac{k(k-1)}{(2k-1)\sqrt{(2k-3)(2k+1)}} \overline{P_{k-2}}(x), \qquad (289)
\end{aligned}
$$

where the first equality follows from (284) and the supposition in this paragraph that $k \geq 2$, the second equality just reorders terms and multiplies and divides by the same quantity twice (by $\sqrt{k+\frac{5}{2}}$ in the term multiplying P_{k+2} and by $\sqrt{k-\frac{3}{2}}$ in the term multiplying P_{k-2}), the third equality recognizes that $\sqrt{k+\frac{5}{2}} = \sqrt{k+2+\frac{1}{2}}$ and that $\sqrt{k-\frac{3}{2}} = \sqrt{k-2+\frac{1}{2}}$ and uses the normalization (270), the fourth equality multiplies and divides by $\sqrt{2}$ in both the term multiplying P_{k+2} and the term multiplying P_{k-2}, and the fifth equality is by elementary algebra.

For $k = 0$, $P_0(x) = 1$, so

$$
\begin{aligned}
&\sqrt{\frac{1}{2}} \left\{ -\frac{d}{dx}\left(\left(1-x^2\right)\frac{dP_0}{dx}(x)\right) - c^2 x^2 P_0(x) \right\} \\
&= \sqrt{\frac{1}{2}} \left\{-c^2 x^2\right\} \qquad (290) \\
&= -c^2 \sqrt{\frac{1}{2}} x^2 \qquad (291) \\
&= -c^2 \sqrt{\frac{1}{2}} \left(\frac{2}{3} P_2(x) + \frac{1}{3}\right) \qquad (292) \\
&= -c^2 \sqrt{\frac{1}{2}} \left(\frac{2}{3} P_2(x) + \frac{1}{3} P_0(x)\right) \qquad (293) \\
&= -c^2 \frac{2}{3} \sqrt{\frac{1}{2}} P_2(x) - c^2 \frac{1}{3} \overline{P_0}(x) \qquad (294) \\
&= -c^2 \frac{2}{3} \frac{\sqrt{\frac{1}{2}}}{\sqrt{2+\frac{1}{2}}} \sqrt{2+\frac{1}{2}} P_2(x) - c^2 \frac{1}{3} \overline{P_0}(x) \qquad (295) \\
&= -c^2 \frac{1}{3} \overline{P_0}(x) - c^2 \frac{2}{3\sqrt{5}} \overline{P_2}(x), \qquad (296)
\end{aligned}
$$

where the first equality follows from $P_0(x) = 1$, the second equality simply rearranges terms, the third equality uses the fact that $P_2(x) = \frac{1}{2}\left(3x^2 - 1\right)$ (use the three term recurrence relation (267) or see Gradshteyn & Ryzhik (2000), section 8.912, page 975), the fourth equality follows from $P_0(x) = 1$, the fifth equality uses the normalization (270), the sixth equality multiplies and divides by $\sqrt{2+\frac{1}{2}}$ in the term multiplying P_2, and the seventh equality reorders terms, uses the normalization (270), and simplifies the fraction in the first term of the preceding expression.

For $k = 1$, $P_1(x) = x$, so

$$\begin{aligned}
& \sqrt{1+\frac{1}{2}}\left\{-\frac{d}{dx}\left(\left(1-x^2\right)\frac{dP_1}{dx}(x)\right) - c^2x^2P_1(x)\right\} \\
= \; & \sqrt{1+\frac{1}{2}}\left\{2x - c^2x^3\right\} && (297) \\
= \; & \sqrt{1+\frac{1}{2}}\left\{2P_1(x) - c^2x^3\right\} && (298) \\
= \; & \sqrt{1+\frac{1}{2}}\left\{2P_1(x) - c^2\left(\frac{2}{5}P_3(x) + \frac{3}{5}P_1(x)\right)\right\} && (299) \\
= \; & 2\overline{P_1}(x) - c^2\left(\frac{2}{5}\sqrt{1+\frac{1}{2}}P_3(x) + \frac{3}{5}\overline{P_1}(x)\right) && (300) \\
= \; & 2\overline{P_1}(x) - c^2\left(\frac{2}{5}\frac{\sqrt{1+\frac{1}{2}}}{\sqrt{3+\frac{1}{2}}}\sqrt{3+\frac{1}{2}}P_3(x) + \frac{3}{5}\overline{P_1}(x)\right) && (301) \\
= \; & 2\overline{P_1}(x) - c^2\left(\frac{2}{5}\frac{\sqrt{1+\frac{1}{2}}}{\sqrt{3+\frac{1}{2}}}\overline{P_3}(x) + \frac{3}{5}\overline{P_1}(x)\right) && (302) \\
= \; & \left(2 - c^2\frac{3}{5}\right)\overline{P_1}(x) - c^2\frac{2\sqrt{3}}{5\sqrt{7}}\overline{P_3}(x) && (303) \\
= \; & \left(2 - c^2\frac{3}{5}\right)\overline{P_1}(x) - c^2\frac{6}{5\sqrt{21}}\overline{P_3}(x), && (304)
\end{aligned}$$

where the first equality follows from $P_1(x) = x$, the second equality uses the same fact ($P_1(x) = x$), the third equality uses the fact that $P_3(x) = \frac{1}{2}\left(5x^3 - 3x\right)$ (use the three term recurrence relation (267) or see Gradshteyn & Ryzhik (2000), section 8.912, page 975), the fourth equality uses the normalization (270), the fifth equality multiplies and divides by $\sqrt{3+\frac{1}{2}}$ in the term multiplying P_3, the sixth equality uses the normalization (270), the seventh equality rearranges terms and simplifies, and the eighth equality multiplies and divides by $\sqrt{3}$ in the term multiplying P_3 in order to make the final expression (304) more directly comparable to the $k \geq 2$ case given in (289).

Define

$$A_{k,k}^{c} \equiv k\left(k+1\right)-c^{2}\frac{2k\left(k+1\right)-1}{\left(2k+3\right)\left(2k-1\right)} \quad \left(k=0,1,2,\ldots\right) \tag{305}$$

$$A_{k,k+2}^{c} \equiv -c^{2}\frac{\left(k+1\right)\left(k+2\right)}{\left(2k+3\right)\sqrt{\left(2k+1\right)\left(2k+5\right)}} \quad \left(k=0,1,2,\ldots\right) \tag{306}$$

$$A_{k,k-2}^{c} \equiv -c^{2}\frac{k\left(k-1\right)}{\left(2k-1\right)\sqrt{\left(2k-3\right)\left(2k+1\right)}} \quad \left(k=2,3,4,\ldots\right), \tag{307}$$

and note that these definitions imply that

$$A_{k,k+2}^{c} = A_{k+2,k}^{c} \quad \left(k=0,1,2,\ldots\right). \tag{308}$$

Now observe that the final expressions in (289), (296), and (304) take the form

$$A_{0,0}^{c}\overline{P_{0}}\left(x\right)+A_{0,2}^{c}\overline{P_{2}}\left(x\right) \tag{309}$$

$$A_{1,1}^{c}\overline{P_{1}}\left(x\right)+A_{1,3}^{c}\overline{P_{3}}\left(x\right) \tag{310}$$

$$A_{k,k}^{c}\overline{P_{k}}\left(x\right)+A_{k,k-2}^{c}\overline{P_{k-2}}\left(x\right)+A_{k,k+2}^{c}\overline{P_{k+2}}\left(x\right) \quad \left(k=2,3,4,\ldots\right), \tag{311}$$

where the (309) equals the final expression in (296), (310) the final expression in (304), and (311) equals the final expression in (289).

I substitute the expressions (309), (310), and (311) into (278) to obtain:

$$\begin{aligned}
&\chi_{n}\left(c\right)\sum_{k=0}^{\infty}\beta_{k}^{(n)}\overline{P_{k}}\left(x\right)\\
= \; &\beta_{0}^{(n)}\left\{A_{0,0}^{c}\overline{P_{0}}\left(x\right)+A_{0,2}^{c}\overline{P_{2}}\left(x\right)\right\}\\
&+\beta_{1}^{(n)}\left\{A_{1,1}^{c}\overline{P_{1}}\left(x\right)+A_{1,3}^{c}\overline{P_{3}}\left(x\right)\right\}\\
&+\sum_{k=2}^{\infty}\beta_{k}^{(n)}\left\{A_{k,k}^{c}\overline{P_{k}}\left(x\right)+A_{k,k-2}^{c}\overline{P_{k-2}}\left(x\right)+A_{k,k+2}^{c}\overline{P_{k+2}}\left(x\right)\right\}.
\end{aligned} \tag{312}$$

I can now rearrange the terms on the right-hand side of the equality (312) to group together all terms that involve $\overline{P_{m}}$ for each $m=0,1,2,\ldots$, yielding

$$\begin{aligned}
&\chi_{n}\left(c\right)\sum_{m=0}^{\infty}\beta_{m}^{(n)}\overline{P_{m}}\left(x\right)\\
= \; &\left(A_{0,0}^{c}\beta_{0}^{(n)}+A_{2,0}^{c}\beta_{2}^{(n)}\right)\overline{P_{0}}\left(x\right)\\
&+\left(A_{1,1}^{c}\beta_{1}^{(n)}+A_{3,1}^{c}\beta_{3}^{(n)}\right)\overline{P_{1}}\left(x\right)\\
&+\sum_{m=2}^{\infty}\left(A_{m,m}^{c}\beta_{m}^{(n)}+A_{m+2,m}^{c}\beta_{m+2}^{(n)}+A_{m-2,m}^{c}\beta_{m}^{(n)}\right)\overline{P_{m}}\left(x\right) \qquad (313)\\
= \; &\left(A_{0,0}^{c}\beta_{0}^{(n)}+A_{0,2}^{c}\beta_{2}^{(n)}\right)\overline{P_{0}}\left(x\right)
\end{aligned}$$

$$+\left(A^c_{1,1}\beta^{(n)}_1+A^c_{1,3}\beta^{(n)}_3\right)\overline{P_1}(x)$$
$$+\sum_{m=2}^{\infty}\left(A^c_{m,m}\beta^{(n)}_m+A^c_{m,m+2}\beta^{(n)}_{m+2}+A^c_{m,m-2}\beta^{(n)}_m\right)\overline{P_m}(x), \quad (314)$$

where the first equality simply regroups terms on the right-hand side of (312) as noted above and the second equality follows from the symmetry noted in (308). Because $\{\overline{P_m}\}_{m=0}^{\infty}$ form a complete orthonormal system in $L^2[-1,1]$, the equality (314) must hold not only in the aggregate, but for each individual P_m:

$$\chi_0(c)\,\beta^{(n)}_0 = A^c_{0,0}\beta^{(n)}_0+A^c_{0,2}\beta^{(n)}_2 \quad (315)$$
$$\chi_1(c)\,\beta^{(n)}_1 = A^c_{1,1}\beta^{(n)}_1+A^c_{1,3}\beta^{(n)}_3 \quad (316)$$
$$\chi_n(c)\,\beta^{(n)}_m = A^c_{m,m}\beta^{(n)}_m+A^c_{m,m+2}\beta^{(n)}_{m+2}+A^c_{m,m-2}\beta^{(n)}_m \quad (m\geq 2)\,. \quad (317)$$

These equations form an (infinite) eigenvalue-eigenvector relationship for $\left\{\beta^{(n)}_m\right\}_{m=0}^{\infty}$ (the eigenvector) and $\chi_n(c)$ (the eigenvalue); recall that the (infinite) vector $\left\{\beta^{(n)}_m\right\}_{m=0}^{\infty}$ is also properly normalized by (273). The (infinite) matrix for which $\left\{\beta^{(n)}_m\right\}_{m=0}^{\infty}$ is an eigenvector with eigenvalue $\chi_n(c)$ is:

$$A^c \equiv \begin{pmatrix} A^c_{0,0} & 0 & A^c_{0,2} & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & A^c_{1,1} & 0 & A^c_{1,3} & 0 & 0 & 0 & 0 & \cdots \\ A^c_{0,2} & 0 & A^c_{2,2} & 0 & A^c_{2,4} & 0 & 0 & 0 & \cdots \\ 0 & A^c_{1,3} & 0 & A^c_{3,3} & 0 & A^c_{3,5} & 0 & 0 & \cdots \\ 0 & 0 & A^c_{2,4} & 0 & A^c_{4,4} & 0 & A^c_{4,6} & 0 & \cdots \\ \vdots & & & \ddots & & \ddots & & \ddots & \end{pmatrix}. \quad (318)$$

The matrix A^c is pentadiagonal, but by recognizing that only indices offset by two appear in the equations (315), (316), and (317), it can naturally be separated into two tridiagonal matrices: a matrix with only even indices and a matrix with only odd indices.

$$A^c_{\text{even}} \equiv \begin{pmatrix} A^c_{0,0} & A^c_{0,2} & 0 & 0 & 0 & \cdots \\ A^c_{0,2} & A^c_{2,2} & A^c_{2,4} & 0 & 0 & \cdots \\ 0 & A^c_{2,4} & A^c_{4,4} & A^c_{4,6} & 0 & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \end{pmatrix} \quad (319)$$

$$A^c_{\text{odd}} \equiv \begin{pmatrix} A^c_{1,1} & A^c_{1,3} & 0 & 0 & 0 & \cdots \\ A^c_{1,3} & A^c_{3,3} & A^c_{3,5} & 0 & 0 & \cdots \\ 0 & A^c_{3,5} & A^c_{5,5} & A^c_{5,7} & 0 & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \end{pmatrix} \quad (320)$$

Because $\psi^c_n(x)$ is an even function for n even and an odd function for n odd, and because the same is true of $\overline{P_n}(x)$, the (infinite) vector $\left\{\beta^{(n)}_m\right\}_{m=0}^{\infty}$ has zero entries for all odd m when n is even and has zero entries for all even m when n is odd. In other words, $\beta_m^{(n)} = 0$ unless m and n share the same parity (either both even or both odd). Using this observation, define

$$\beta_{\text{even}}^{(n)} \equiv \begin{pmatrix} \beta_0^{(n)} & \beta_2^{(n)} & \beta_4^{(n)} & \cdots \end{pmatrix}^T \tag{321}$$

$$\beta_{\text{odd}}^{(n)} \equiv \begin{pmatrix} \beta_1^{(n)} & \beta_3^{(n)} & \beta_5^{(n)} & \cdots \end{pmatrix}^T, \tag{322}$$

where the T superscript denotes transposition, so that $\beta_{\text{even}}^{(n)}$ and $\beta_{\text{odd}}^{(n)}$ are infinite column vectors. Then for n even, $\beta_{\text{odd}}^{(n)}$ is the zero vector and

$$\chi_n(c)\,\beta_{\text{even}}^{(n)} = A_{\text{even}}^c \beta_{\text{even}}^{(n)}. \tag{323}$$

If n is odd, then $\beta_{\text{even}}^{(n)}$ is the zero vector and

$$\chi_n(c)\,\beta_{\text{odd}}^{(n)} = A_{\text{odd}}^c \beta_{\text{odd}}^{(n)}. \tag{324}$$

The equations (323) and (324) are (infinite) symmetric tridiagonal eigenvalue-eigenvector equations.

To obtain approximations to $\left\{\beta_m^{(n)}\right\}_{m=0}^{\infty}$ for $n = 0, 1, \ldots, N$ (for any $N \geq 0$), and thus to obtain approximations to $\psi_n^c(x)$ for $n = 0, 1, \ldots, N$, simply truncate the infinite symmetric tridiagonal matrices A_{even}^c and A_{odd}^c at some sizes K_{even} and K_{odd} respectively, each substantially larger than N and differing by at most one, to obtain the $K_{\text{even}} \times K_{\text{even}}$ symmetric tridiagonal matrix $\widetilde{A}_{\text{even}}^c$ and the $K_{\text{odd}} \times K_{\text{odd}}$ symmetric tridiagonal matrix $\widetilde{A}_{\text{odd}}^c$. In general, the appropriate truncation sizes depend on both N and c (and are increasing in both). An overall diagonal length of 150 (including both evens and odds) seems sufficient in all cases of practical financial interest. With these truncated, finite matrices, two applications of any standard numerical method for the computation of eigenvalues and eigenvectors of a symmetric tridiagonal matrix (one application of such a method for $\widetilde{A}_{\text{even}}^c$ and one for $\widetilde{A}_{\text{odd}}^c$) will deliver the truncated approximations to $\chi_n(c)$, $\left\{\beta_{m,\text{ even}}^{(n)}\right\}_{m=0}^{K_{\text{even}}}$, and $\left\{\beta_{m,\text{ odd}}^{(n)}\right\}_{m=0}^{K_{\text{odd}}}$. These, in turn, deliver approximations to $\left\{\beta_m^{(n)}\right\}_{m=0}^{M}$ (where $M = K_{\text{even}} + K_{\text{odd}}$) and thus to $\psi_n^c(x)$ for $n = 0, 1, \ldots, N$.

The computational cost of computing the eigenvalues and eigenvectors of a symmetric tridiagonal matrix is quadratic in the number of rows (or columns) of that matrix. This will not dominate the computational cost of the method I propose, since the symmetric tridiagonal eigensolver calls can be made upfront, prior to any invocations of the method I propose to approximate $\exp(cxt)$.

The discussion above provides the foundation for the following methods.

Method 8.3. The following method computes a given number of the eigenvalues of the oblate spheroidal wave operator L_c and the coefficients of series in (normalized) Legendre polynomials for the corresponding eigenfunctions (the oblate spheroidal wave functions of order zero with parameter c).

- **Given**:
  1. n, the number of eigenvalues to compute (and also the number of corresponding eigenfunctions for which the coefficients of series in (normalized) Legendre polynomials are to be calculated)
  2. $c > 0$
  3. a truncation size M, which must be greater than n
- **Output**:
  1. $\boldsymbol{\chi}$, a $n \times 1$ vector of the n smallest eigenvalues of L_c, sorted in ascending order
  2. $\mathcal{B}$, a matrix with n columns and number of rows equal to M (if M is odd) or $M+1$ (if M is even); column j, indexed starting at 0, of $\mathcal{B}$ is a computation of the column vector $\left(\beta_m^{(j)}\right)_{m=0}^{M_B}$, so column j of $\mathcal{B}$ is a computation of the initial M_B coefficients in an expansion of ψ_j^c in (normalized) Legendre polynomials, and M_B is either M (for M odd) or $M+1$ (for M even)

Step 1 Check the arguments.

Step 2 If the truncation size M is even, set the truncation size for evens $M_{\text{even}} = M$ and set the truncation size for odds $M_{\text{odd}} = M - 1$. If the truncation size M is odd, set the truncation size for evens $M_{\text{even}} = M - 1$ and set the truncation size for odds $M_{\text{odd}} = M$. This assures that, overall, indices from 0 through M will be included in subsequent computations.

Step 3 Generate a vector of the even numbers from 0 through M_{even} (including 0 and M_{even}), in order. Label this vector I_{even}; it contains $\frac{1}{2}\left(M_{\text{even}} + 2\right)$ elements.

Step 4 Generate a vector of the odd numbers from 1 through M_{odd} (including 1 and M_{odd}), in order. Label this vector I_{odd}; it contains $\frac{1}{2}\left(M_{\text{odd}} + 1\right)$ elements.

Step 5 For each index k in I_{even}, compute $A^c_{k,k}$ using (305). Assemble the $A^c_{k,k}$, ordered by k, into a vector $\boldsymbol{d}^c_{\text{even}}$. The vector $\boldsymbol{d}^c_{\text{even}}$ thus contains, in order, the initial $\frac{1}{2}\left(M_{\text{even}} + 2\right)$ diagonal elements of A^c_{even} (as defined in (319)).

Step 6 For each index k in I_{even}, compute $A^c_{k,k+2}$ using (306). Assemble the $A^c_{k,k+2}$, ordered by k, into a vector $\boldsymbol{s}^c_{\text{even}}$. The vector $\boldsymbol{s}^c_{\text{even}}$ thus contains, in order, the initial $\frac{1}{2}\left(M_{\text{even}} + 2\right)$ subdiagonal elements of A^c_{even} (as defined in (319)).

Step 7 For each index k in I_{odd}, compute $A^c_{k,k}$ using (305). Assemble the $A^c_{k,k}$, ordered by k, into a vector $\boldsymbol{d}^c_{\text{odd}}$. The vector $\boldsymbol{d}^c_{\text{odd}}$ thus contains, in order, the initial $\frac{1}{2}\left(M_{\text{odd}} + 1\right)$ diagonal elements of A^c_{odd} (as defined in (320)).

Step 8 For each index k in I_{odd}, compute $A^c_{k,k+2}$ using (306). Assemble the $A^c_{k,k+2}$, ordered by k, into a vector $\boldsymbol{s}^c_{\mathrm{odd}}$. The vector $\boldsymbol{s}^c_{\mathrm{odd}}$ thus contains, in order, the initial $\frac{1}{2}(M_{\mathrm{odd}}+1)$ subdiagonal elements of A^c_{odd} (as defined in (320)).

Step 9 Discard the final entry of $\boldsymbol{s}^c_{\mathrm{even}}$ and the final entry of $\boldsymbol{s}^c_{\mathrm{odd}}$.

Step 10 Form, as a sparse matrix to avoid unneccessary storage of zeros, the $\frac{1}{2}(M_{\mathrm{even}}+2) \times \frac{1}{2}(M_{\mathrm{even}}+2)$ symmetric tridiagonal matrix $A^{\mathrm{trunc}}_{\mathrm{even}}$ that represents the top left corner of the infinite symmetric tridiagonal matrix A^c_{even} given by (319). Only three vectors need be stored: $\boldsymbol{d}^c_{\mathrm{even}}$ (the diagonal), $\boldsymbol{s}^c_{\mathrm{even}}$ (the subdiagonal), and a copy of $\boldsymbol{s}^c_{\mathrm{even}}$ (the superdiagonal, which equals the subdiagonal by symmetry).

Step 11 Form, as a sparse matrix to avoid unneccessary storage of zeros, the $\frac{1}{2}(M_{\mathrm{odd}}+1) \times \frac{1}{2}(M_{\mathrm{odd}}+1)$ symmetric tridiagonal matrix $A^{\mathrm{trunc}}_{\mathrm{odd}}$ that represents the top left corner of the infinite symmetric tridiagonal matrix A^c_{odd} given by (320). Only three vectors need be stored: $\boldsymbol{d}^c_{\mathrm{odd}}$ (the diagonal), $\boldsymbol{s}^c_{\mathrm{odd}}$ (the subdiagonal), and a copy of $\boldsymbol{s}^c_{\mathrm{odd}}$ (the superdiagonal, which equals the subdiagonal by symmetry).

Step 12 Using a standard eigenvalue and eigenvector routine, compute the eigenvalues and eigenvectors of $A^{\mathrm{trunc}}_{\mathrm{even}}$. It is more efficient, but not necessary, to use a routine specifically designed for symmetric tridiagonal matrices. Store the eigenvalues in a $\frac{1}{2}(M_{\mathrm{even}}+2) \times \frac{1}{2}(M_{\mathrm{even}}+2)$ diagonal matrix χ_{even}, and store the eigenvectors as the columns of a matrix $\mathcal{B}_{\mathrm{even}}$ of the same size.

Step 13 Using a standard eigenvalue and eigenvector routine, compute the eigenvalues and eigenvectors of $A^{\mathrm{trunc}}_{\mathrm{odd}}$. It is more efficient, but not necessary, to use a routine specifically designed for symmetric tridiagonal matrices. Store the eigenvalues in a $\frac{1}{2}(M_{\mathrm{odd}}+1) \times \frac{1}{2}(M_{\mathrm{odd}}+1)$ diagonal matrix χ_{odd}, and store the eigenvectors as the columns of a matrix $\mathcal{B}_{\mathrm{odd}}$ of the same size.

Step 14 If $\frac{1}{2}(M_{\mathrm{even}}+2) > \frac{1}{2}(M_{\mathrm{odd}}+1)$, then increase the number of rows and the number of columns of both χ_{odd} and $\mathcal{B}_{\mathrm{odd}}$ by 1; set all of the entries in the last row and in the last column of each of these newly-enlarged matrices to 0. If $\frac{1}{2}(M_{\mathrm{even}}+2) < \frac{1}{2}(M_{\mathrm{odd}}+1)$, do the same but for the corresponding "even" matrices (this will never occur given the steps above, but is included as a robustness feature if those steps were ever to change). Note that the two sizes being compared can never differ by more than 1.

Step 15 Set

$$\widetilde{\boldsymbol{\chi}} \quad = \quad \mathrm{diag}\left(\chi_{\mathrm{even}} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \chi_{\mathrm{odd}} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right), \qquad (325)$$

where the symbol $\otimes$ denotes, as usual, the Kronecker product of two matrices and the function "diag" extracts the diagonal of a matrix and returns it as a column vector. Then set the output variable $\boldsymbol{\chi}$ to be equal to the first n elements of $\widetilde{\boldsymbol{\chi}}$.

Step 16 Set

$$\widetilde{\mathcal{B}} = \mathcal{B}_{\text{even}} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \mathcal{B}_{\text{odd}} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{326}$$

where the symbol $\otimes$ denotes, as in Step 15, the Kronecker product of two matrices. Then set the output variable $\mathcal{B}$ to be equal to the first n columns of $\widetilde{\mathcal{B}}$, but with the final row deleted.

Remark 8.3. I could make Method 8.3 more efficient in a number of places, since the output variable $\mathcal{B}$ has roughly half of its elements equal to zero in a pattern that is known in advance (each column has an alternating non-zero / zero pattern, with the first element being zero for odd-indexed columns and non-zero for even-indexed columns when indexing is started at 0). Because Method 8.3 can be used for precalculation in approximation of the exponential product operator by oblates or by the method of Subsection 8.5, I do not pursue further refinements here.

I now require methods to evaluate the oblates and their first derivatives.

Method 8.4. This method computes the result of evaluating the oblates ψ_j^c for $j = 0, 1, \ldots n-1$ at each of the elements of a column vector $\boldsymbol{x}$, given a matrix of (normalized) Legendre series coefficients such as the one produced by Method 8.3.

- **Given**:

  1. $\mathcal{B}$, a matrix with n columns and number of rows equal to some truncation size M; column j, indexed starting at 0, of $\mathcal{B}$ is a computation of the column vector $\left(\beta_m^{(j)}\right)_{m=0}^{M_B}$, so column j of $\mathcal{B}$ is a computation of the initial M_B coefficients in an expansion of ψ_j^c in (normalized) Legendre polynomials.
  2. $\boldsymbol{x}$, a column vector of points at which to evaluate the oblates.

- **Output**:

  1. $\mathcal{F}$, a matrix with n columns and number of rows equal to the length of $\boldsymbol{x}$; the (i, j) element of $\mathcal{F}$, indexed starting at 0, is $\psi_j^c(x_i)$.

Step 1 Check the arguments.

Step 2 Invoke Method 8.1 with inputs $M_B - 1$ and $\boldsymbol{x}$ to obtain the matrix V_{Leg} which has M_B columns and a number of rows equal to the length of $\boldsymbol{x}$.

Step 3 Set the output variable $\mathcal{F} = V_{\text{Leg}}\mathcal{B}$. Note that this results in $\mathcal{F}$ having the proper size.

Method 8.5. This method computes the result of evaluating the first derivatives of the oblates ψ_j^c for $j = 0, 1, \ldots n-1$ at each of the elements of a column vector $\boldsymbol{x}$, given a matrix of (normalized) Legendre series coefficients such as the one produced by Method 8.3.

- **Given**:

  1. $\mathcal{B}$, a matrix with n columns and number of rows equal to some truncation size M; column j, indexed starting at 0, of $\mathcal{B}$ is a computation of the column vector $\left(\beta_m^{(j)}\right)_{m=0}^{M_B}$, so column j of $\mathcal{B}$ is a computation of the initial M_B coefficients in an expansion of ψ_j^c in (normalized) Legendre polynomials.
  2. $\boldsymbol{x}$, a column vector of points at which to evaluate the first derivatives of the oblates.

- **Output**:

  1. $\mathcal{G}$, a matrix with n columns and number of rows equal to the length of $\boldsymbol{x}$; the (i, j) element of $\mathcal{G}$, indexed starting at 0, is the first derivative of $\psi_j^c(x_i)$.

Step 1 Check the arguments.

Step 2 Invoke Method 8.2 with inputs $M_B - 1$ and $\boldsymbol{x}$ to obtain the matrix V_{LegDeriv} which has M_B columns and a number of rows equal to the length of $\boldsymbol{x}$.

Step 3 Set the output variable $\mathcal{G} = V_{\text{LegDeriv}}\mathcal{B}$. Note that this results in $\mathcal{G}$ having the proper size.

With all of the preceding methods in place, it is now possible to provide a method for constructing the approximation Q_n of the exponential product operator P_c using its eigenfunctions, the oblates ψ_n^c, $n = 0, 1, 2, \ldots$. I have already provided computational methods for the construction and evaluation of the oblates (the eigenfunctions of P_c), but I have not yet discussed the calculation of the eigenvalues $\lambda_n(c)$, $n = 0, 1, 2, \ldots$, of P_c.

By evaluating the basic eigenvalue-eigenfunction identity

$$\lambda_n(c)\,\psi_n^c(x) \quad = \quad \int_{-1}^{1} \exp(cxt)\psi_n^c(t)\,dt \tag{327}$$

at $x = 0$, I obtain (following Osipov *et al.* (2013), Chapter 10, where the prolate case is treated in a very similar way)

$$\begin{aligned}
\lambda_n(c)\,\psi_n^c(0) &= \int_{-1}^{1} \psi_n^c(t)\,dt && (328)\\
&= \int_{-1}^{1} \sum_{k=0}^{\infty} \beta_k^{(n)} \overline{P_k}(t)\,dt && (329)\\
&= \sum_{k=0}^{\infty} \beta_k^{(n)} \int_{-1}^{1} \overline{P_k}(t)\,dt && (330)\\
&= \beta_0^{(n)}\sqrt{2}, && (331)
\end{aligned}$$

where the first equality follows from evaluating (327) at $x = 0$ and noting that $\exp(0) = 1$, the second equality follows from substituting the series expression (272) into the prior expression, the third equality employs a semidiscrete form of Fubini's theorem, and the fourth equality follows from the fact that $\overline{P_0}(t) = \sqrt{\frac{1}{2}}$, so $\int_{-1}^{1} \overline{P_k}(t)\,dt = \int_{-1}^{1} \overline{P_k}(t)\sqrt{2}\overline{P_0}(t)\,dt$, which is zero (by orthogonality) if $k \neq 0$ and is $\sqrt{2}$ if $k = 0$ (since the $\overline{P_k}(t)$ are not just orthogonal, but orthonormal).

For n even, $\psi_n^c(0) \neq 0$, so a valid calculation for $\lambda_n(c)$ for n even comes from dividing (331) by $\psi_n^c(0)$:

$$\lambda_n(c) = \frac{\beta_0^{(n)}}{\psi_n^c(0)}\sqrt{2} \quad \text{for } n \text{ even.} \tag{332}$$

When n is odd, $\psi_n^c(0) = 0$ and (332) is not helpful in determining $\lambda_n(c)$. Instead, consider differentiating both sides of (327) with respect to x and then evaluating the resulting expression at $x = 0$ (again, following the very similar treatment given by Osipov *et al.* (2013) in the prolate case):

$$\begin{aligned}
\lambda_n(c)\,\frac{d\psi_n^c}{dx}(0) &= \int_{-1}^{1} ct\psi_n^c(t)\,dt && (333)\\
&= \int_{-1}^{1} ct\sum_{k=0}^{\infty} \beta_k^{(n)} \overline{P_k}(t)\,dt && (334)\\
&= c\sum_{k=0}^{\infty} \beta_k^{(n)} \int_{-1}^{1} t\overline{P_k}(t)\,dt && (335)\\
&= \beta_1^{(n)} c\sqrt{\frac{2}{3}}, && (336)
\end{aligned}$$

where the first equality follows from differentiation with respect to x and then evaluation at $x = 0$ of (327), the second equality follows from substituting the series expression (272) into the prior expression, the third equality employs a semidiscrete form of Fubini's theorem (and factoring the constant c out of the integration and the summation), and the fourth equality follows from the fact that $\overline{P_1}(t) = t\sqrt{\frac{3}{2}}$, so $\int_{-1}^{1} t\overline{P_k}(t)\,dt = \int_{-1}^{1} \overline{P_k}(t)\sqrt{\frac{2}{3}}\overline{P_1}(t)\,dt$, which is zero (by orthogonality) if $k \neq 1$ and is $\sqrt{\frac{2}{3}}$ if $k = 1$ (since the $\overline{P_k}(t)$ are not just orthogonal, but orthonormal).

For n odd, $\frac{d\psi_n^c}{dx}(0) \neq 0$, so a valid calculation for $\lambda_n(c)$ for n odd comes from dividing (336) by $\psi_n^c(0)$:

$$\lambda_n(c) = \frac{\beta_1^{(n)}}{\frac{d\psi_n^c}{dx}(0)} c\sqrt{\frac{2}{3}} \quad \text{for } n \text{ odd.} \tag{337}$$

I can, finally, provide the following method to construct the components of the truncated eigenfunction approximation Q_n of Theorem 8.1.

Method 8.6. This method computes the first (largest) n eigenvalues of the exponential product operator P_c and the coefficients of the (normalized) Legendre series expansions of the corresponding first n eigenfunctions (which are oblates); Method 8.3 is invoked to obtain the series expansion coefficients.

- **Given**:
  1. n, the number of eigenvalues to compute (and also the number of corresponding eigenfunctions for which the coefficients of series in (normalized) Legendre polynomials are to be calculated)
  2. $c > 0$
  3. a truncation size M, which must be greater than n
- **Output**:
  1. $\boldsymbol{\lambda}$, a $n \times 1$ vector of the n largest eigenvalues of P_c, sorted in descending order
  2. $\mathcal{B}$, a matrix with n columns and number of rows equal to M (if M is odd) or $M+1$ (if M is even); column j, indexed starting at 0, of $\mathcal{B}$ is a computation of the column vector $\left(\beta_m^{(j)}\right)_{m=0}^{M_B}$, so column j of $\mathcal{B}$ is a computation of the initial M_B coefficients in an expansion of ψ_j^c in (normalized) Legendre polynomials, and M_B is either M (for M odd) or $M+1$ (for M even)

Simply combine the methods and logic above.

Recalling Theorem 8.1 at the beginning of this section, the n-term truncated eigenfunction expansion Q_n whose components are constructed by invoking Method 8.6 approximates the exponential product operator P_c optimally in terms of its average squared error (in the sense that this approximation achieves the linear and Kolmogorov n-widths when the exponential product operator is viewed as mapping $L^2[-1,1]$ to itself).

Figure 5 and Figure 6 show the error surfaces generated by approximating the exponential product operator using the oblate approximation. The "fuzzy" or "jagged" appearance of the bottom left panel of Figure 5 arises because the

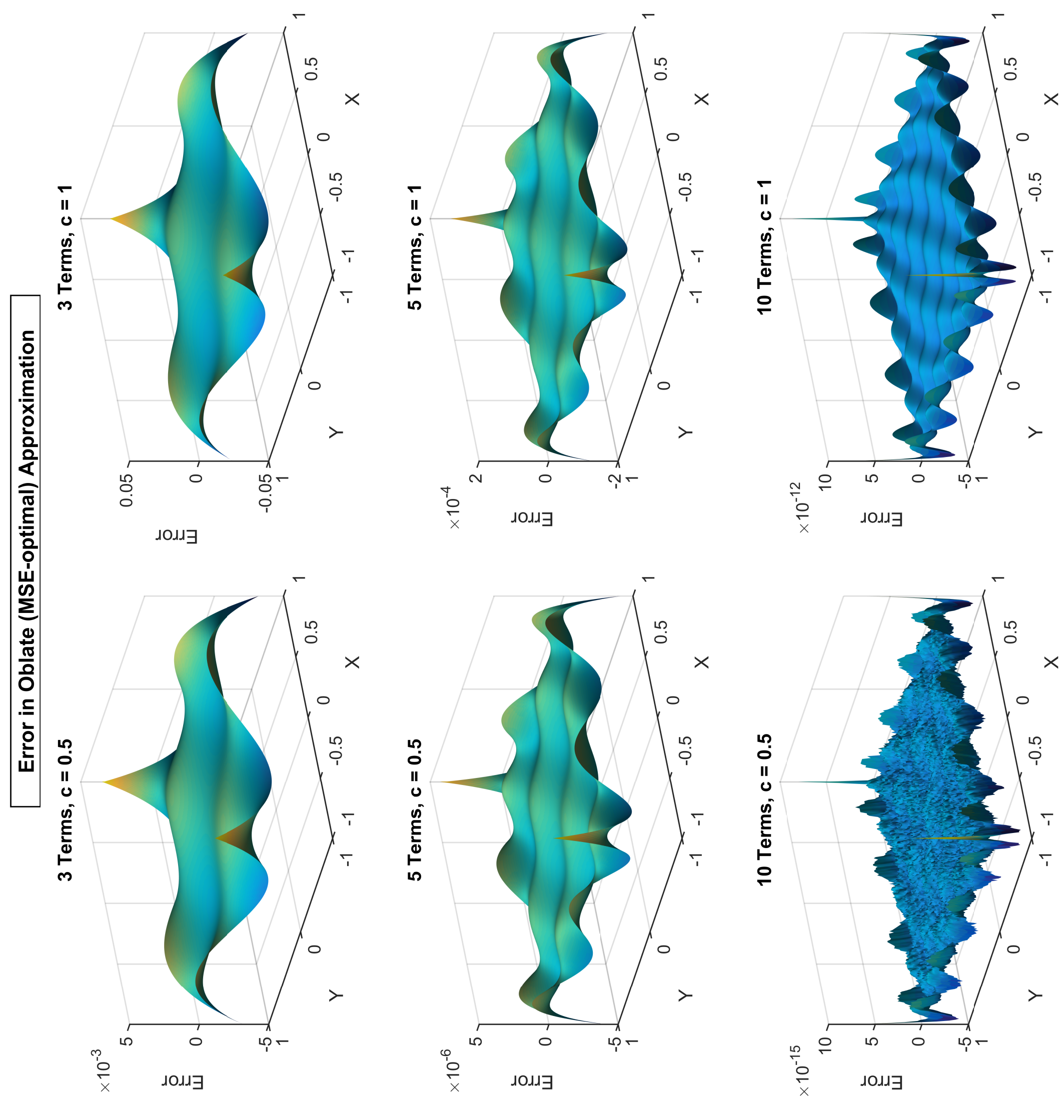

Figure 5: Error Surfaces in the Oblate (MSE-optimal) Approximation to $\exp(cxy)$

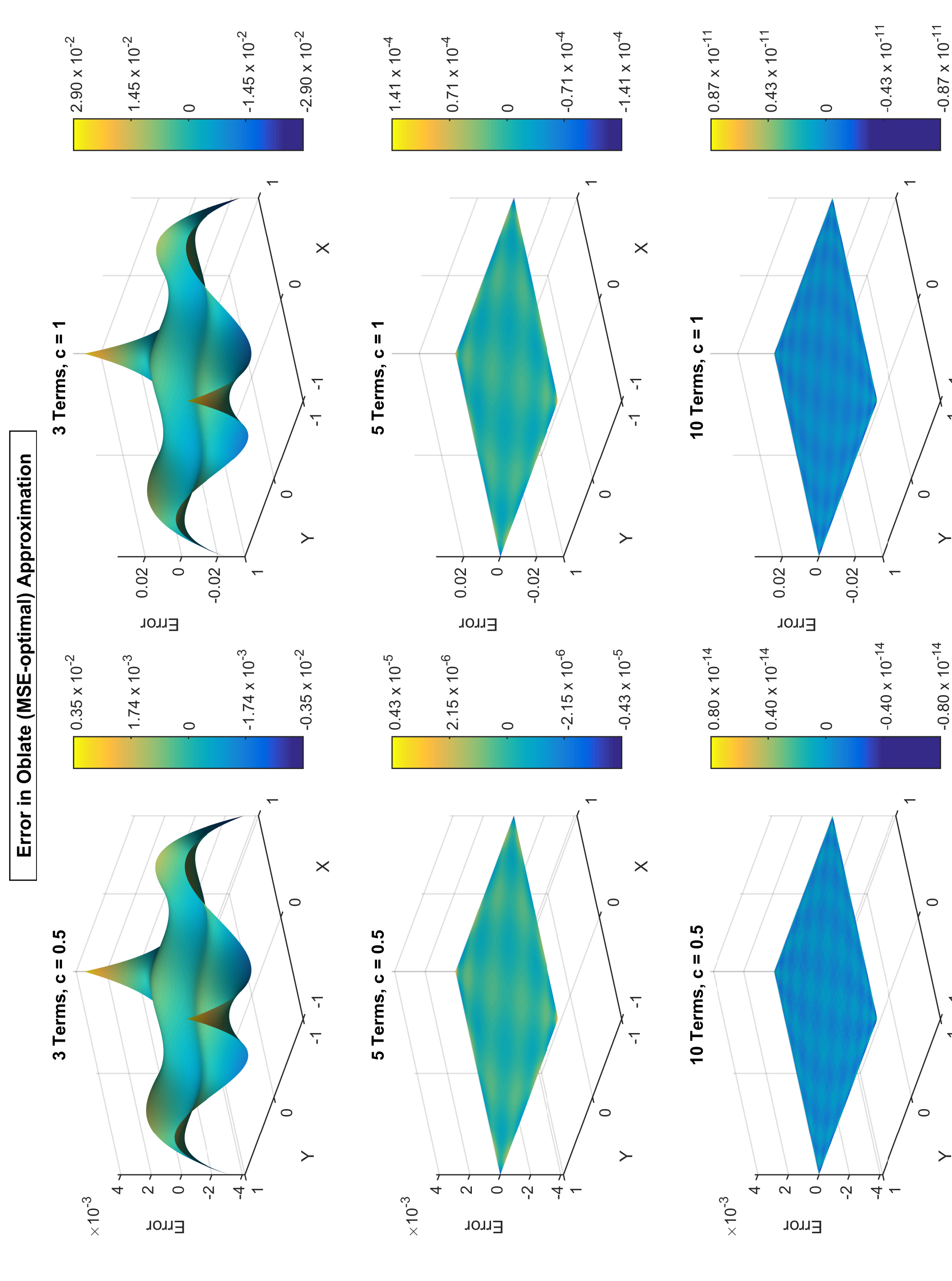

Figure 6: Error Surfaces in the Oblate (MSE-optimal) Approximation to $\exp{(cxy)}$

approximation error in that panel has been reduced to the point at which it is being impacted by rounding error (note the scale of the vertical axis in that panel).

8.5 An Interpolatory Average-squared-error Optimal Approximation and Its Interpretation

In this subsection, I construct new interpolatory linear approximations to the exponential product operator that are optimal in an average-squared-error sense (they achieve the linear and Kolmogorov n-widths when the exponential product operator is viewed as mapping $L^2[-1,1]$ to itself, just as the truncated eigenfunction expansions Q_n constructed in Subsection 8.4 do). I apply the remarkable results of Melkman & Micchelli (1978) (nicely summarized by Pinkus (1985a) in Section 5 of his Chapter 4) to build the approximations of this subsection.

The original motivation of Melkman & Micchelli (1978) came from consideration of spline spaces. My work is the first to apply their theory to the approximation of the exponential product operator.

The interpolatory character of the approximations built in this subsection leads to a helpful financial interpretation that is very similar to the one offered for the worst-case-error numerically-optimal approximations $P^*_{c,n}(x,y)$ in Subsection 7.1: the rank-n approximation constructed by the method of this subsection approximates any given cashflow stream using an approximating portfolio of a fixed set of n approximating cashflow streams, each having cashflows at the same n times. The holdings of the approximating portfolio are linear in the cashflow stream to be approximated: different cashflow streams have distinct approximating portfolio holdings of the same n approximating cashflow streams, and the approximating holdings for a portfolio of cashflow streams are found by summing the approximating holdings for each individual cashflow stream in the portfolio.

As a consequence of the representation (339) below, the logic of (170) in Subsection 7.1 applies directly to demonstrate my claims regarding the interpretation of this subsection's approximation. Only a single dimension changes, as V^* in Subsection 7.1 is $(n+1)\times n$, while U is $n\times n$; this change leads to the smaller set of common cashflow times in the approximating cashflow streams (there are n approximating cashflow times here, while for the numerically-optimal approximation there are still n approximating cashflow streams, but they have cashflows on $n+1$ common times).

To construct the interpolatory approximations of this section, I use the oblate-evaluation functionality provided by Method 8.4 in Subsection 8.4 to find all of the zeros of a given oblate. Once these zeros are found, recognizing the form of the Melkman & Micchelli (1978) interpolatory approximation leads quickly to a method of constructing it.

Theorem 8.11. *ψ^c_n has exactly n zeros in the interval $[-1,1]$. Label these zeros in ascending order as $\eta_1 < \eta_2 < \cdots < \eta_n$. Perform the eigenvalue-eigenvector decomposition*

$$\begin{pmatrix} \exp(c\eta_1\eta_1) & \cdots & \exp(c\eta_1\eta_n) \\ \vdots & \ddots & \vdots \\ \exp(c\eta_n\eta_1) & \cdots & \exp(c\eta_n\eta_n) \end{pmatrix} = U^T \Lambda U, \tag{338}$$

so that Λ is a diagonal matrix with the eigenvalues of the lefthand side of (338), in descending order, on its diagonal and U is the orthogonal matrix whose columns are the corresponding eigenvectors (the leftmost column of U is the eigenvector associated with the largest eigenvalue of the matrix, the second-from-left column of U is the eigenvector associated with the second largest eigenvalue, and so on). The approximation

$$P_{c,n}^{\mathrm{Interp}}(x,y) = \begin{pmatrix} \exp(cx\eta_1) \\ \vdots \\ \exp(cx\eta_n) \end{pmatrix}^T U^T \Lambda^{-1} U \begin{pmatrix} \exp(c\eta_1 y) \\ \vdots \\ \exp(c\eta_n y) \end{pmatrix} \tag{339}$$

can be expressed as

$$P_{c,n}^{\mathrm{Interp}}(x,y) = K(x,y) - \frac{K\begin{pmatrix} \eta_1 & \cdots & \eta_n & x \\ \eta_1 & \cdots & \eta_n & y \end{pmatrix}}{K\begin{pmatrix} \eta_1 & \cdots & \eta_n \\ \eta_1 & \cdots & \eta_n \end{pmatrix}}, \tag{340}$$

where the kernel $K(x,y)$ is the exponential product kernel $\exp(cxy)$. $P_{c,n}^{\mathrm{Interp}}$ is interpolatory at the η_i:

$$\exp(c\eta_i y) = P_{c,n}^{\mathrm{Interp}}(\eta_i, y) \ \forall y \in [-1,1], \ i = 1,2,\ldots,n \tag{341}$$

$$\exp(cx\eta_j) = P_{c,n}^{\mathrm{Interp}}(x, \eta_j) \ \forall x \in [-1,1], \ j = 1,2,\ldots,n. \tag{342}$$

$P_{c,n}^{\mathrm{Interp}}$ is also achieves the linear and Kolmogorov n-widths of the exponential product operator P_c when the exponential product operator is viewed as mapping $L^2[-1,1]$ to itself.

Proof. The fact that ψ_n^c has exactly n zeros on the interval $[-1,1]$ is a consequence of Corollary 5.1 (which shows that this property is true for the eigenfunctions of P_c) and Theorem 8.6 (which shows that ψ_n^c is the eigenfunction of P_c associated with the eigenvalue $\lambda_n(c)$ in the ordered sequence $\lambda_0(c) > \lambda_1(c) > \lambda_2(c) > \cdots$).

The equivalence of (339) and (340) follows from the standard Laplace formula for the expansion of a determinant in its minors and the usual expression of an inverse in terms of minors, since the expression (338) defining U and Λ implies that

$$\begin{pmatrix} \exp(c\eta_1\eta_1) & \cdots & \exp(c\eta_1\eta_n) \\ \vdots & \ddots & \vdots \\ \exp(c\eta_n\eta_1) & \cdots & \exp(c\eta_n\eta_n) \end{pmatrix}^{-1} = U^T \Lambda^{-1} U, \tag{343}$$

because U is orthogonal, so $U^{-1} = U^T$, and $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$ for any conforming invertible matrices A, B, and C. This is just a relabeling of the same equivalence of expressions that I observed in Section 7 for the interpolatory component of $P_{c,n}^*(x,y)$ (the equivalence between (141) and (148)).

The interpolatory equalities (341) and (342) follow directly from (340), since the determinant of a matrix having two identical rows or two identical columns is zero. This is, again, similar to the situation I discussed in Section 7 for the interpolatory component $\exp(cxy) - K_c^{(n+1)}(x, y)$ of the approximation error $\exp(cxy) - P_{c,n}^{*}(x, y)$.

Finally, the optimality of $P_{c,n}^{\mathrm{Interp}}$ for the linear and Kolmogorov n-widths of the exponential product operator P_c when P_c is viewed as mapping $L^2[-1, 1]$ to itself is a consequence of Melkman & Micchelli (1978), Theorem 2.4 on page 549. Note that the "λ_{n+1}" of Melkman & Micchelli (1978) is the $(n+1)^{\mathrm{st}}$ largest eigenvalue of the operator P_c^2 (the iterated operator), so their "$\lambda_{n+1}^{1/2}$" is equivalent to my λ_n. Pinkus (1985a) also states and proves this result, referencing Melkman & Micchelli (1978), in his Theorem 5.5 on pages 114-115 in Section 5 of Chapter IV. The Pinkus (1985a) proof notes that this interpolatory approximating operator has the determinantal form shown in (340), and his statement includes the fact that this approximation achieves the Kolmogorov n-width in the Hilbert-space case assumed here. Since the approximating operator is linear, it is obviously optimal for the linear n-width as well. □

Having established the properties of $P_{c,n}^{\mathrm{Interp}}$, I now describe its construction in the following method.

Method 8.7. This method computes $\eta_1 < \eta_2 < \cdots < \eta_n$, all in the interval $[-1, 1]$, an $n \times n$ orthogonal matrix U, and the $n \times 1$ vector $\boldsymbol{\lambda}_{\mathrm{Interp}}^{-1}$ of the diagonal elements of the $n \times n$ matrix Λ^{-1}, where the η_i, U, and Λ are as described in Theorem 8.11.

- **Given**:
  1. n, the rank with which to approximate the exponential product operator
  2. $c > 0$
  3. a truncation size M, which must be greater than n and is used in computing the η_i
- **Output**:
  1. $\eta_1 < \eta_2 < \cdots < \eta_n$, all in the interval $[-1, 1]$, the n roots of ψ_n^c on $[-1, 1]$
  2. $\boldsymbol{\lambda}_{\mathrm{Interp}}^{-1}$, a $n \times 1$ vector of the diagonal elements, sorted in descending order, of Λ^{-1}, where Λ is the diagonal $n \times n$ matrix defined by the decomposition (338)
  3. U, an orthogonal $n \times n$ matrix defined by the decomposition (338)

Step 1 Check the arguments

Step 2 Invoke Method 8.3 with the arguments n, c, and M to obtain the output $\mathcal{B}$ of coefficients in the (truncated) Legendre polynomial series for the first $n+1$ oblates (the ψ_j^c for $j = 0, 1, \ldots, n$). Discard the output $\boldsymbol{\chi}$.

Step 3 Using the rightmost column of $\mathcal{B}$ as found in Step 2 (the column which represents the coefficients in the series for ψ_n^c) invoke Method 8.4 as needed to evaluate ψ_n^c within a standard zero-finding routine run between each neighboring pair of points in a fine grid in order to find all of the zeros of ψ_n^c on the interval $[-1, 1]$. Since ψ_n^c is known to have exactly n zeros in the interval $[-1, 1]$ (as noted in Theorem 8.11), if fewer than n zeros are found the search grid must be refined and the search repeated. Continue this process until n zeros are located or until the search grid becomes implausibly large (in the latter case, generate an error message and stop). If more than n zeros are found, either ψ_n^c is within numerical tolerance of the zero function across part of $[-1, 1]$ or there is a numerical problem. If more than n zeros are found and ψ_n^c is within numerical tolerance of zero, set the zeros to n Chebyshev points; if more than n zeros are found and ψ_n^c is not within numerical tolerance of zero, generate an error message and stop. Label the zeros that are found in order as $\eta_1 < \eta_2 < \cdots < \eta_n$.

Step 4 Use a standard eigenvalue-eigenvector decomposition routine to find U and Λ in the decomposition (338), where the $\eta_1 < \eta_2 < \cdots < \eta_n$ found in Step 3 are employed in forming the matrix on the lefthand side of (338).

Step 5 Extract the diagonal of the diagonal matrix Λ computed in Step 4. Form the vector $\boldsymbol{\lambda}_{\text{Interp}}^{-1}$ which has, for each λ_i on the diagonal of Λ, the entry $1/\lambda_i$ at index i, $i = 1, \ldots, n$.

Figure 7 and Figure 8 show the error surfaces generated by approximating the exponential product operator using the interpolatory approximation. The "fuzzy" or "jagged" appearance of the bottom left panel of Figure 7 arises because the approximation error in that panel has been reduced to the point at which it is being impacted by rounding error (note the scale of the vertical axis in that panel).

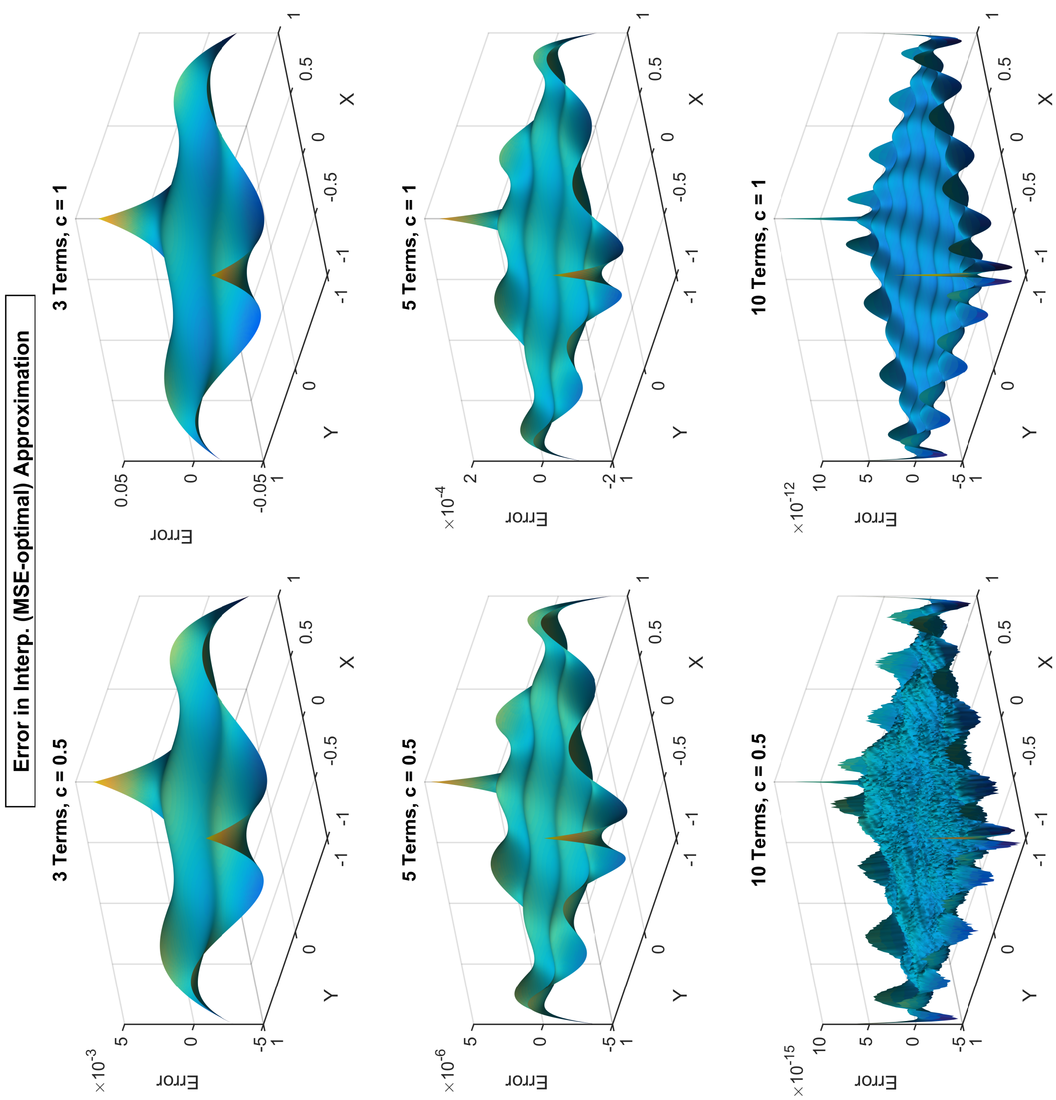

Figure 7: Error Surfaces in the Interpolatory (MSE-optimal) Approximation to $\exp(cxy)$

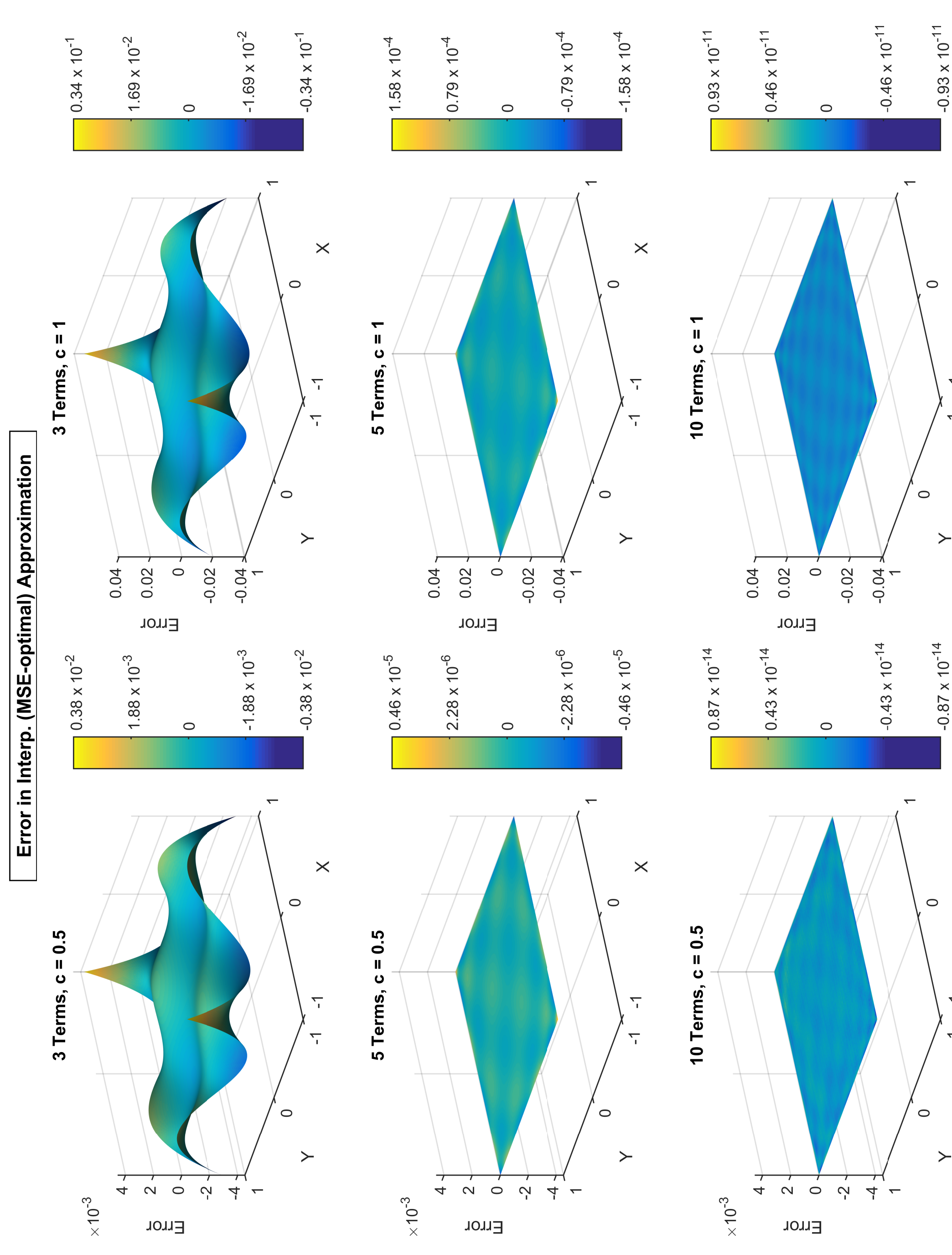

Figure 8: Error Surfaces in the Interpolatory (MSE-optimal) Approximation to $\exp(cxy)$

9 A Nearly-optimal Approximation for Worst-case Error

In this section, I use weightings of exponential product operators to motivate novel approximations based on the eigenfunctions of these weighted operators. Chebyshev polynomials are well-known to produce near-optimal max-norm (worst-case error) approximations in the univariate case (see Trefethen (2013)), so I use a weighting linked to the Chebyshev weight function. The eigenfunctions of the resulting weighted exponential product operators are the Mathieu functions of classical mathematical physics. I show that truncated eigenfunction expansions of a weighted exponential product operator lead directly to Mathieu expansions for the corresponding (unweighted) exponential product operator. Further, these Mathieu expansions are nearly optimal for worst-case error approximation: they are close to achieving the lower bound on the linear n-width, though they are not quite as accurate as the numerically-optimal approximations introduced in Section 7.

The Mathieu expansions are slightly less accurate than the numerically-optimal approximations $P^*_{c,n}(x,y)$, and they lack the remarkable interpretation provided for $P^*_{c,n}(x,y)$ in Subsection 7.1. However, they have some advantages over the numerically-optimal approximations: first, they provided nested approximations (the rank-$(n+1)$ approximation is obtained just by adding a new term to the rank-n approximation); second, they are less computationally intensive to construct than the $P^*_{c,n}(x,y)$; third, computing a (loose) upper bound on the global maximum of the absolute value of their approximation errors is very straightforward.

The motivation for the Mathieu approximation comes from the success of Chebyshev polynomials in univariate max-norm (worst-case error) approximation; the Chebyshev polynomials are orthogonal over $[-1,1]$ with respect to the weight function $\left(1-x^2\right)^{-\frac{1}{2}}$, so they would be orthogonal (in an unweighted sense) when scaled by the square root of that function, $\left(1-x^2\right)^{-\frac{1}{4}}$. This leads me to consider a similar scaling of the exponential product operators.

Definition 9.1. The weighted exponential product operators are the family of integral operators

$$P_c^{\text{weighted}}[f](y) \equiv \int_{-1}^{1}\left(1-y^2\right)^{-\frac{1}{4}}\exp(cxy)\left(1-x^2\right)^{-\frac{1}{4}}f(x)\,dx, \quad (344)$$

where the family is indexed by the real parameter $c>0$ and f satisfies

$$\int_{-1}^{1}\left(1-x^2\right)^{-1/2}(f(x))^2\,dx<\infty,$$

and $y\in[-1,1]$.

Since $\int_{-1}^{1}\left(1-x^2\right)^{-1/2}dx=\pi$, and $\left(1-x^2\right)^{-1/2}$ is obviously positive on $(-1,1)$, functions bounded on $[-1,1]$ are certainly in the domain of P_c^{weighted}.

For the same reason, the boundedness of $\exp(cxy)$ on $[-1,1] \times [-1,1]$ immediately yields that P_c^{weighted} is Hilbert-Schmidt (and thus compact). Since P_c^{weighted} is also self-adjoint on the stated domain, it has a convergent eigenfunction expansion (where convergence is in the weighted L^2 norm that is specified in the description of the domain in (9.1)).

An eigenfunction of P_c^{weighted} satisfies (from the (68), with relabeling)

$$\gamma\phi(y) = \int_{-1}^{1} \left(1-y^2\right)^{-\frac{1}{4}} \exp(cxy) \left(1-x^2\right)^{-\frac{1}{4}} \phi(x)\, dx, \tag{345}$$

where γ is the corresponding eigenvalue. Any such eigenfunction ϕ must have $\int_{-1}^{1} \left(1-x^2\right)^{-1/2} (\phi(x))^2\, dx < \infty$, so the function $\vartheta(x) = \left(1-x^2\right)^{\frac{1}{4}} \phi(x)$ is clearly in $L^2[-1,1]$. Express (345) in terms of ϑ to get:

$$\gamma\vartheta(y) = \int_{-1}^{1} \exp(cxy) \left(1-x^2\right)^{-\frac{1}{2}} \vartheta(x)\, dx, \tag{346}$$

where I have cancelled the term $\left(1-y^2\right)^{-\frac{1}{4}}$ that appears on both the left and the right of the equation and aggregated the product of the two identical $\left(1-x^2\right)^{-\frac{1}{4}}$ terms to get $\left(1-x^2\right)^{-\frac{1}{2}}$.

Now make the substitutions $x = \cos(t)$ and $y = \cos(u)$ in (346) to obtain

$$\gamma\vartheta(\cos(u)) = \int_0^{\pi} \exp(c\cos(t)\cos(u))\vartheta(\cos(t))\, dt, \tag{347}$$

where the disappearance of the factor $\left(1-x^2\right)^{-\frac{1}{2}}$ comes from the usual impact of trigonometric substitution in an integral and I have exploited the evenness of $\cos(t)$ to make the bounds of integration more natural (integrating from 0 to π rather than from $-\pi$ to 0).

The trigonometric form of (347) leads naturally to the following definition of the exponential cosine product operators.

Definition 9.2. The exponential cosine product operators are the family of integral operators

$$P_c^{\cos}[f](y) \equiv \int_0^{\pi} \exp(c\cos(t)\cos(u)) f(t)\, dt, \tag{348}$$

where the family is indexed by the real parameter $c > 0$ and f satisfies

$$\int_0^{\pi} (f(t))^2\, dt < \infty,$$

and $y \in [-1,1]$.

From this point, the analysis of the problem closely parallels that of Section 8 (though the resulting eigenfunctions and approximations are qualitatively different). I indicate only the key results, proceeding more rapidly than in Section 8.

As in Section 8, there is a commuting second-order differential operator for $P_c^{\cos}$, and it is one of the classical differential operators of mathematical physics: the Mathieu differential operator.

Definition 9.3. The domain D_M consists of all $f(t) \in L^2[0,\pi]$ such that $f(t)$ is absolutely continuous on $[0,\pi]$, $\frac{df}{dt}(t)$ is absolutely continuous on $[0,\pi]$, $\frac{d^2f}{dt^2}(t) \in L^2[0,\pi]$, and $\frac{df}{dt}(0) = \frac{df}{dt}(\pi) = 0$.

The domain D_M is, of course, a linear subspace of $L^2[0,\pi]$. It is also dense in $L^2[0,\pi]$ (it contains all Fourier cosine series). The boundary conditions are of the standard separated type.

Definition 9.4. The *Mathieu operator* with parameter $c > 0$ and defined on the domain D_M is

$$M_c[f](t) \equiv -\frac{d^2f}{dt^2}(t) - \frac{c^2}{2}\cos(2t)f(t). \tag{349}$$

Typically, the Mathieu differential equation is written in the form

$$w''(z) + (a - 2q\cos(2z))\,w(z) = 0, \tag{350}$$

where a plays the role of an eigenvalue; to pass from the lefthand side of (350) to Definition 9.4, let $a = 0$ (to get an operator rather than an equation) and let $q = -\frac{c^2}{4}$, then take the negative of the expression.

The eigenfunctions of the Mathieu operator are known as Mathieu functions, and have been extensively investigated since their namesake initially explored them in analyzing the vibrations of an elliptical drumhead. Helpful references include Abramowitz & Stegun (1964) (Chapter 20), Arscott (1964), Erdélyi *et al.* (1955) (Chapter XVI), McLachlan (1947) (a monograph on Mathieu functions), Meixner & Schäfke (1954), and Olver *et al.* (2010) (Chapter 28). The relationship between spheroidal wave functions and Mathieu functions that is suggested by the discussion above is noted specifically in Chapter 30 of **?** (see the subsection "Special Cases" on page 698): the "half-order" spheroidal wave equation reduces to the Mathieu equation (this is also noted on page 134 in Section 9 of Chapter 16 in Erdélyi *et al.* (1955)).

Theorem 9.1. *M_c is self-adjoint and is bounded below by $-\frac{c^2}{2}$.*

Proof. One proof consists of observing that the boundary conditions imposed through the definition of D_M are of the standard separated, self-adjoint type (see Zettl (2005), page 71 in Section 2 of Chapter 4) and the operator M_c itself is of the self-adjoint variety detailed on the lefthand side of equation (4.1.1) in Section 1 of Chapter 4 of Zettl (2005).

Another proof mirrors that of Theorem 8.2, with an adjustment for the regularity of the operator $-\frac{d^2f}{dt^2}$ where $t \in [0,\pi]$.

Akhiezer & Glazman (1963), pages 168 - 170, provide criteria for boundary conditions that define the domain of a self-adjoint extension of a regular differential operator. The boundary conditions defining D_M satisfy these criteria for the operator $A \equiv -\frac{d^2 f}{dt^2}$: to see this, note that for A with domain D_M, $n = 1$, $a = -1$, $b = 1$, $\alpha_{11} = 0$, $\alpha_{12} = 1$, $\beta_{11} = 0$, $\beta_{12} = 0$, $\alpha_{21} = 0$, $\alpha_{22} = 0$, $\beta_{21} = 0$, and $\beta_{22} = 1$ in the notation of Akhiezer & Glazman (1963), equations (2') and (3'), which are thus satisfied. The operator $B \equiv -\frac{c^2}{2}\cos(2t)$ (a multiplication operator with domain all of $L^2[0,\pi]$) is clearly bounded with operator norm not exceeding $\frac{c^2}{2}$ (if $\left(\int_0^\pi (f(t))^2\,dt\right)^{\frac{1}{2}} = C_0$ then $\left(\int_0^\pi \left(-\frac{c^2}{2}\cos(2t) f(t)\right)^2 dt\right)^{\frac{1}{2}} \leq \frac{c^2}{2}\left(\int_0^\pi (f(t))^2\,dt\right)^{\frac{1}{2}} = \frac{c^2}{2}C_0$) and is obviously symmetric. In the language of Reed & Simon (1975), Chapter X, Section 2, page 162, B is *infinitesimally small* with respect to A, so that the Reed & Simon (1975) parameter a on page 162 of Chapter X, Section 2 is zero and the parameter b is no greater than $\frac{c^2}{2}$. Thus, by the Kato-Rellich theorem (Reed & Simon (1975), Chapter X, Section 2, Theorem X.12 on page 162), the operator $A+B = -\frac{d^2 f}{dt^2}(t) - \frac{c^2}{2}\cos(2t) f(t) = M_c$ is self-adjoint with domain of definition D_M and is bounded below by $-\frac{c^2}{2}$ (since the differential operator A is positive definite, and thus bounded below by zero).

□

Corollary 9.1. *M_c is symmetric, that is, for all $f, g \in D_M$,*

$$\int_0^\pi M_c[f](t)\, g(t)\, dt \quad = \quad \int_0^\pi f(t)\, M_c[g](t)\, dt.$$

Proof. The logic is identical to that of Corollary 8.1. □

As mentioned earlier, $P_c^{\cos}$ and M_c commute. This could be inferred from some of the classical integral relations for the Mathieu functions, see Arscott (1964) (Chapter IV), McLachlan (1947) (Chapter X), or Olver *et al.* (2010) (Section 10 of Chapter 28), but none of these references seem to provide the result in precisely the needed form.

Theorem 9.2. *$P_c^{\cos}$ and L_c commute; that is, $f \in D_M$ implies that $P_c^{\cos}[f] \in D_M$ and that*

$$M_c\left[P_c^{\cos}[f]\right](u) \quad = \quad P_c^{\cos}\left[M_c[f]\right](u).$$

Proof. Differentiating makes it clear that $f \in D_M$ implies that $P_c^{\cos}[f] \in D_M$:

$$\begin{aligned} & \frac{d}{du}\int_0^\pi \exp(c\cos(t)\cos(u)) f(t)\, dt \\ = \quad & -c\sin(u)\int_0^\pi \cos(t)\exp(c\cos(t)\cos(u)) f(t)\, dt, \end{aligned} \tag{351}$$

which is zero when evaluated at $u = 0$ or at $u = \pi$.

Assuming that $f \in D_M$, I show that the functions $M_c\left[P_c^{\cos}[f]\right](u)$ and $P_c^{\cos}\left[M_c[f]\right](u)$ are equal by computing each of them.

$$
\begin{aligned}
& M_c\left[P_c^{\cos}[f]\right](u) \\
= \quad & -\frac{d^2 P_c[f]}{du^2}(u) - \frac{c^2}{2}\cos(2u) P_c[f](u) && (352) \\
= \quad & -\frac{d^2}{du^2}\int_0^\pi \exp(c\cos(t)\cos(u)) f(t)\,dt && (353) \\
& -\frac{c^2}{2}\cos(2u)\int_0^\pi \exp(c\cos(t)\cos(u)) f(t)\,dt \\
= \quad & -\int_0^\pi \frac{d^2}{du^2}\exp(c\cos(t)\cos(u)) f(t)\,dt && (354) \\
& -\frac{c^2}{2}\cos(2u)\int_0^\pi \exp(c\cos(t)\cos(u)) f(t)\,dt
\end{aligned}
$$

$$
\begin{aligned}
= \quad & -\int_0^\pi \begin{pmatrix} c^2\cos^2(t)\sin^2(u) \\ -c\cos(t)\cos(u) \end{pmatrix} \exp(c\cos(t)\cos(u)) f(t)\,dt && (355) \\
& -\frac{c^2}{2}\cos(2u)\int_0^\pi \exp(c\cos(t)\cos(u)) f(t)\,dt,
\end{aligned}
$$

where the first equality is by the definition of M_c, the second equality is by the definition of $P_c^{\cos}$ and elementary manipulation, the third equality interchanges the order of differentiation and integration, and the fourth equality computes the second derivative shown in the prior expression.

$$
\begin{aligned}
& P_c^{\cos}\left[L_c[f]\right](u) && (356) \\
= \quad & \int_0^\pi \exp(c\cos(t)\cos(u))\left(-\frac{d^2 f}{dt^2}(t) - \frac{c^2}{2}\cos(2t) f(t)\right) dt && (357) \\
= \quad & \int_0^\pi \begin{pmatrix} -\frac{d^2}{dt^2}\exp(c\cos(t)\cos(u)) \\ -\frac{c^2}{2}\cos(2t)\exp(c\cos(t)\cos(u)) \end{pmatrix} f(t)\,dt && (358) \\
= \quad & -\int_0^\pi \begin{pmatrix} c^2\cos^2(u)\sin^2(t) \\ -c\cos(t)\cos(u) \\ +\frac{c^2}{2}\cos(2t) \end{pmatrix} \exp(c\cos(t)\cos(u)) f(t)\,dt && (359) \\
= \quad & -\int_0^\pi \begin{pmatrix} \frac{c^2}{2}(1+\cos(2u)) \\ \times\left(1-\cos^2(t)\right) \\ -c\cos(t)\cos(u) \\ +\frac{c^2}{2}\left(2\cos^2(t)-1\right) \end{pmatrix} \exp(c\cos(t)\cos(u)) f(t)\,dt && (360) \\
= \quad & -\int_0^\pi \begin{pmatrix} \frac{c^2}{2}\cos(2u) + \frac{c^2}{2} \\ -\frac{c^2}{2}\cos^2(t) \\ -\frac{c^2}{2}\cos(2u)\cos^2(t) \\ -c\cos(t)\cos(u) \\ +c^2\cos^2(t) - \frac{c^2}{2} \end{pmatrix} \exp(c\cos(t)\cos(u)) f(t)\,dt && (361)
\end{aligned}
$$

$$= -\int_0^\pi \begin{pmatrix} \frac{c^2}{2}\cos(2u) \\ +\frac{c^2}{2}(1-\cos(2u)) \\ \times\cos^2(t) \\ -c\cos(t)\cos(u) \end{pmatrix} \exp(c\cos(t)\cos(u)) f(t)\,dt \quad (362)$$

$$= -\int_0^\pi \begin{pmatrix} \frac{c^2}{2}\cos(2u) \\ +\frac{c^2}{2}2\sin^2(u)\cos^2(t) \\ -c\cos(t)\cos(u) \end{pmatrix} \exp(c\cos(t)\cos(u)) f(t)\,dt \quad (363)$$

$$= -\int_0^\pi \begin{pmatrix} c^2\cos^2(t)\sin^2(u) \\ -c\cos(t)\cos(u) \end{pmatrix} \exp(c\cos(t)\cos(u)) f(t)\,dt \quad (364)$$
$$-\frac{c^2}{2}\cos(2u)\int_0^\pi \exp(c\cos(t)\cos(u)) f(t)\,dt,$$

where the first equality is by the definitions of M_c and $P_c^{\cos}$, the second equality is by Corollary 9.1 (the symmetry of M_c, where I regard $\exp(c\cos(t)\cos(u))$ as a function of t, which applies because $f \in D_M$ by assumption), the third equality computes the second derivative, the fourth equality employs trigonometric identities ($\cos^2(u) = \frac{1}{2}(1+\cos(2u))$, $\sin^2(t) = 1-\cos^2(t)$, and $\cos(2t) = 2\cos^2(t)-1$), the fifth equality expands the terms of the prior line, the sixth equality cancels and regroups, the seventh equality uses the trigonometric identity $1-\cos(2u) = 2\sin^2(u)$, and the final equality exploits the linearity of the integral.

Since (355) and (364) are identical, $M_c[P_c^{\cos}[f]](u) = P_c^{\cos}[M_c[f]](u)$, that is, $P_c^{\cos}$ and M_c commute as claimed.

□

I now need a number of results characterizing the exponential cosine product kernel; these closely parallel the analogous results for the exponential product operator.

The exponential cosine product kernel, like the exponential product kernel, is a strictly totally positive kernel, though the appropriate domain is more restricted (compare Theorem 9.3 below to Theorem 5.8).

Theorem 9.3. *For any $c > 0$, the exponential cosine product kernel*

$$\exp(c\cos(t)\cos(u))$$

is STP on $[0,\pi]\times[0,\pi]$.

Proof. This follows from Theorem 5.8, $\cos(t)\cos(u) = (-\cos(t))(-\cos(u))$, and the fact that $-\cos(t)$ and $-\cos(u)$ are both strictly increasing and continuous on $[0,\pi]$, given Definition 5.19; see observation 2) on page 56 of Pinkus (1985a). Substitution into Definition 5.19 makes observation 2) on page 56 of Pinkus (1985a) transparent. □

The lemma below can be compared to Lemma 5.1.

Lemma 9.1. *For any $c > 0$, the exponential cosine product kernel*

$$\exp\left(c \cos(t) \cos(u)\right),$$

defined on $t, u \in [0, \pi]$, is symmetric, continuous, Kellogg, and strictly positive.

Proof. The kernel $\exp\left(c \cos(t) \cos(u)\right)$ is clearly symmetric, continuous and strictly positive for $t, u \in [0, \pi]$. For any $c > 0$, it is STP by Theorem 9.3; thus, it is Kellogg for any $c > 0$ by Definition 5.26. □

The following corollary is the exponential cosine product analog of Corollary 5.1 for the exponential product operator.

Corollary 9.2 (of Theorem 5.9)**.** *For any $c > 0$, the exponential cosine product operator is positive definite with simple eigenvalues and Markov eigenfunctions (as defined in Definition 5.28).*

Proof. The exponential cosine product operator has the kernel

$$K(t, u) = \exp\left(c \cos(t) \cos(u)\right),$$

defined on $[0, \pi] \times [0, \pi]$, and for any $c > 0$ this kernel satisfies all of the assumptions of Theorem 5.9 by Lemma 9.1. □

Having characterized the exponential cosine product operator $P_c^{\cos}$, I now refer to classical results for the eigenfunctions and eigenvalues of the Mathieu operator M_c (defined on the domain D_M). The form of the boundary conditions embedded in Definition 9.3 implies that the eigenfunctions of M_c (defined on the domain D_M) are just constant multiples of the Mathieu functions commonly referred to as $\mathrm{ce}_n\left(t, -\frac{c^2}{4}\right)$, $n = 0, 1, 2, \ldots$. Indeed, this is the content of equation (3) on page 112 in Section 4 of Chapter 16 of Erdélyi *et al.* (1955). The notation ce_n comes from an abbreviation of "cosine elliptic."

Theorem 9.4. *The operator M_c has a countable set of eigenfunctions which form a complete orthonormal basis for $L^2[0, \pi]$. Each eigenvalue $a_n(c)$ of M_c is simple. The eigenvalues $a_n(c)$ are bounded below but not above: $a_n(c) \to \infty$ as $n \to \infty$. Ordering these eigenvalues $a_0(c) < a_1(c) < a_2(c) < \cdots$, the eigenfunction $\phi_n^c(t)$ of M_c that is associated with the eigenvalue $a_n(c)$ has exactly n zeros in the interval $[0, \pi]$.*

Proof. The eigenproblem for M_c is, given its domain of definition D_M, a self-adjoint, regular Sturm-Liouville problem with separated boundary conditions. Theorem 4.6.2 on page 87 in Section 6 of Chapter 4 of Zettl (2005) gives all of the results in the theorem's statement (letting $p(t) \equiv 1$, $q(t) = -\frac{c^2}{2}\cos(2t)$, $w(t) \equiv 1$, $a = 0$, $b = \pi$, $A_1 = 0$, $A_2 = 1$, $B_1 = 0$, and $B_2 = 1$ in the notation of Zettl (2005)), but counts zeros only on the open interval $(0, \pi)$ rather than on the closed interval $[0, \pi]$. This leaves open the possibility of zeros at $t = 0$ or

$t = \pi$ (or both). To see that no eigenfunction can have a zero at $t = 0$, observe that if one did, the defining equation

$$-\frac{d^2\phi_n^c}{dt^2}(0) - \frac{c^2}{2}\cos(2t)\phi_n^c(0) = a_n(c)\,\phi_n^c(0) \tag{365}$$

implies that the second derivative of ϕ_n^c also has a zero at $t = 0$. Successively differentiating and using the fact (from the boundary conditions embedded in D_M) that the first derivative of ϕ_n^c has a zero at $t = 0$, every derivative (third, fourth, and so on) of ϕ_n^c must have a zero at $t = 0$. But each ϕ_n^c, $n = 0, 1, 2, \ldots$ is analytic in t (see Olver *et al.* (2010), the discussion above equation 28.2.4 on page 653 in Section 2 of Chapter 28), so ϕ_n^c must be identically zero. This is a contradiction, so none of the ϕ_n^c, $n = 0, 1, 2, \ldots$, can have a zero at $t = 0$. An entirely similar argument shows that no ϕ_n^c, $n = 0, 1, 2, \ldots$, can have a zero at $t = \pi$. □

Statements of most of Theorem 9.4, and many other results, can be found in the standard references on Mathieu functions cited above.

With characterizations of $P_c^{\cos}$ and M_c both in hand, I can now state the key result of this section. When considering the uniqueness of an eigenfunction corresponding to a given eigenvalue, recall that my definition of an eigenfunction $\phi(t)$, $x \in [0, \pi]$ includes the normalization $\int_0^\pi (\phi(t))^2\, dt = 1$. Thus, if an eigenvalue is simple there is a unique eigenfunction which has that eigenvalue. If my definition did not include normalization, the uniqueness would only be up to multiplication by a scalar.

Theorem 9.5. *An eigenfunction $\xi_n^c(t)$ of $P_c^{\cos}$ whose eigenvalue is $\gamma_n(c)$ is the unique eigenfunction of $P_c^{\cos}$ that has the eigenvalue $\gamma_n(c)$. An eigenfunction $\phi_n^c(t)$ of M_c whose eigenvalue is $a_n(c)$ is the unique eigenfunction of M_c that has the eigenvalue $a_n(c)$. $\xi_n^c(t)$ is the unique eigenfunction of $P_c^{\cos}$ associated with the eigenvalue $\gamma_n(c)$ of $P_c^{\cos}$ (where $\gamma_0(c) > \gamma_1(c) > \gamma_2(c) > \cdots$ and $\gamma_m(c) > 0$ for every $m = 0, 1, 2, \ldots$) if and only if $\xi_n^c(t)$ is the unique eigenfunction of M_c corresponding to the eigenvalue $a_n(c)$ of M_c (where $a_0(c) < a_1(c) < a_2(c) < \cdots$). Further, $\gamma_n(c) \to 0$ as $n \to \infty$ and $a_n(c) \to \infty$ as $n \to \infty$.*

Proof. Renaming objects in the proof of Theorem 8.6 appropriately is enough to demonstrate the statement of this theorem, since Corollary 9.2 and Theorem 9.4 provide the necessary characterizations here (as Corollary 5.1 and Theorem 8.5 did in the proof of Theorem 8.6). □

Theorem 9.5 shows that I can find the eigenfunctions of $P_c^{\cos}$ by solving the eigenproblem for M_c, in a clear parallel to the way in which L_c allowed me to characterize the eigenfunctions of P_c in Section 8. The ce_n are the Mathieu functions whose Fourier expansions have only cosine terms, which is consistent with the expression (347). Each ce_n must be multiplied by a constant to produce my eigenfunctions because the ce_n are orthogonal, but not orthonormal, as typically defined (I require eigenfunctions to be normalized, see Definition 5.27). The usual orthogonality relation for the Mathieu functions involves the integral from 0 to 2π (see, for example, Arscott (1964), equation (1a) on page 57 in Section 4 of Chapter 3 or Olver *et al.* (2010), equation 28.2.31 on page 654 in Section 2 of Chapter 28). The orthogonality I require, which involves integration from 0 to π, is given by Erdélyi *et al.* (1955), equation (16) on page 114 in Section 4 of Chapter XVI.

The ce_n are normalized such that

$$\int_0^\pi \left(\mathrm{ce}_{2n}\left(t, -\frac{c^2}{4}\right)\right)^2 dt \quad = \quad \frac{\pi}{2}, \tag{366}$$

where the usual normalization is stated on the interval $[0, 2\pi]$ but symmetry in combination with periodicity imply the statement (366). The eigenfunctions I consider (which are, in light of Theorem 9.5, eigenfunctions of both M_c and $P_c^{\cos}$) will therefore be

$$\xi_n^c(t) \quad \equiv \quad \sqrt{\frac{2}{\pi}}\,\mathrm{ce}_{2n}\left(t, -\frac{c^2}{4}\right), \tag{367}$$

which satisfy

$$\int_0^\pi \left(\xi_n^c(t)\right)^2 dt \quad = \quad 1 \tag{368}$$

by construction.

The Fourier series expansions of the ce_n, $n = 0, 1, 2, \ldots$, are well-known; the coefficients satisfy recurrence relations which differ depending on whether n is even or odd. For $n = 0, 1, 2, \ldots$,

$$\mathrm{ce}_{2n}\left(t, -\frac{c^2}{4}\right) \quad = \quad \sum_{m=0}^{\infty} A_{2m}^{2n}(c)\cos(2mt) \tag{369}$$

$$\mathrm{ce}_{2n+1}\left(t, -\frac{c^2}{4}\right) \quad = \quad \sum_{m=0}^{\infty} A_{2m+1}^{2n+1}(c)\cos((2m+1)t), \tag{370}$$

which are given in all of the standard references on Mathieu functions listed above (albeit with potential variation in notation; I express the dependence of A_j^k on c directly, while most texts would favor expressing dependence on some quantity such as $q = -\frac{c^2}{4}$); for example, Olver *et al.* (2010), equations 28.4.1 and 28.4.2 on page 656 in Section 4 of Chapter 28.

The coefficients of the Fourier series (369) and (370) satisfy

$$a_{2n}(c)\,A_0^{2n}(c) + \frac{c^2}{4}A_2^{2n}(c) \quad = \quad 0 \tag{371}$$

$$\left(a_{2n}(c) - 4\right)A_2^{2n}(c) + \frac{c^2}{4}\left(2A_0^{2n}(c) + A_4^{2n}(c)\right) \quad = \quad 0 \tag{372}$$

and, for $m = 2, 3, 4, \ldots$,

$$\left(a_{2n}(c) - 4m^2\right)A_{2m}^{2n}(c) + \frac{c^2}{4}\left(A_{2m-2}^{2n}(c) + A_{2m+2}^{2n}(c)\right) \quad = \quad 0 \tag{373}$$

for the even indices $2n$ of (369). For the odd indices $2n+1$ of (370), they satisfy

$$\left(a_{2n+1}(c) - 1 + \frac{c^2}{4}\right) A_1^{2n+1}(c) + \frac{c^2}{4} A_3^{2n+1}(c) = 0 \tag{374}$$

and, for $m = 1, 2, 3, \ldots$,

$$\left(a_{2n}(c) - (2m+1)^2\right) A_{2m+1}^{2n+1}(c) + \frac{c^2}{4}\left(A_{2m-1}^{2n+1}(c) + A_{2m+3}^{2n+1}(c)\right) = 0. \tag{375}$$

These relations are, as with the Fourier series themselves, present in all of the standard references on Mathieu functions (though sometimes with variations of notation). I simply substituted $q = -\frac{c^2}{4}$ into equations 28.4.5 and 28.4.6 on page 656 in Section 4 of Chapter 28 in Olver *et al.* (2010) and used (hopefully) clearer indications of the dependence of the a_k and A_j^k quantities on the parameter c. The standard normalizations for the coefficients in the literature, corresponding to the normalizations (366), are

$$2\left(A_0^{2n}(c)\right)^2 + \sum_{m=1}^{\infty}\left(A_{2m}^{2n}(c)\right)^2 = 1 \tag{376}$$

$$\sum_{m=0}^{\infty}\left(A_{2m+1}^{2n+1}(c)\right)^2 = 1. \tag{377}$$

These normalizations can be found as, for example, equations 28.4.9 and 28.4.10 on page 657 in Chapter 28 of Olver *et al.* (2010).

Of course, the Fourier series for an eigenfunction ξ_n^c is simply a constant multiple of the Fourier series for the corresponding ce_n:

$$\xi_{2n}^c(t) = \sqrt{\frac{2}{\pi}} \sum_{m=0}^{\infty} A_{2m}^{2n}(c)\cos(2mt) \tag{378}$$

$$\xi_{2n+1}^c(t) = \sqrt{\frac{2}{\pi}} \sum_{m=0}^{\infty} A_{2m+1}^{2n+1}(c)\cos((2m+1)t), \tag{379}$$

so the recurrence relations for the coefficients stated above are also appropriate in computing ξ_n^c.

I wish to rephrase the Fourier series for ξ_n^c in terms of normalized Chebyshev polynomials; the existence of such an equivalent expression has been known since Arscott (1964) (see Example 2 of Chapter III, on page 74), if not earlier. I briefly review the properties of Chebyshev polynomials and discuss the normalization I employ, but the excellent monographs Mason & Handscomb (2003) and Rivlin (1990) cover Chebyshev polynomials in great detail. Trefethen (2013) also gives a very helpful treatment of Chebyshev polynomials, including both theoretical and practical perspectives and emphasizing the use of the software package Chebfun (I use Chebfun, under the BSD license, only to compute the Gauss-Legendre nodes and weights in Section 10, but it is well-known and much-appreciated by numerical analysts; my use of Chebfun to compute the Gauss-Legendre nodes and weights utilized in Section 10 should not be taken to constitute any endorsement of my work by Chebfun's authors).

The Chebyshev polynomials T_k, $k = 0, 1, 2, \ldots$ are defined by the property that

$$T_k (\cos (\theta)) = \cos (k\theta). \tag{380}$$

It can be shown that this implies that $T_k(x)$ is a polynomial in x (see Rivlin (1990), Section 1 of Chapter 1). Their definition leads the Chebyshev polynomials to begin with $T_0(x) = 1$ and $T_1(x) = x$; with these initial polynomials fixed, the remaining of the Chebyshev polynomials are defined via a three-term recurrence relation that can be used to express T_{k+1} in terms of T_k and T_{k-1} (and is the result of an elementary trigonometric identity):

$$T_{k+1}(x) - 2xT_k(x) + T_{k-1}(x) = 0. \tag{381}$$

T_k is an even function when k is even and is an odd function when k is odd; further,

$$T_k(1) = 1 \quad (k = 0, 1, 2, \ldots) \tag{382}$$

and

$$|T_k(x)| \leq 1 \quad (x \in [-1, 1],\ k = 0, 1, 2, \ldots). \tag{383}$$

Although $\{T_k\}_{k=0}^{\infty}$ make up a complete orthogonal system in $L^2[-1, 1]$ with the weight function $\left(1 - x^2\right)^{-1/2}$, they satisfy

$$\int_{-1}^{1} (T_0(x))^2 \left(1 - x^2\right)^{-1/2} dx = \pi \tag{384}$$

$$\int_{-1}^{1} (T_k(x))^2 \left(1 - x^2\right)^{-1/2} dx = \frac{\pi}{2},\ k \geq 1, \tag{385}$$

and are thus not orthonormal in this weighted L^2 space.

I define the normalized Chebyshev polynomials:

$$\overline{T_0}(x) \equiv T_0(x)\sqrt{\frac{1}{\pi}} \tag{386}$$

$$\overline{T_k}(x) \equiv T_k(x)\sqrt{\frac{2}{\pi}},\ k \geq 1. \tag{387}$$

This normalization yields

$$\int_{-1}^{1} \left(\overline{T_k}(x)\right)^2 \left(1 - x^2\right)^{-1/2} dx = 1, \tag{388}$$

so the $\{\overline{T_k}\}_{k=0}^{\infty}$ make up a complete orthonormal system in $L^2[-1,1]$ with the weight function $(1-x^2)^{-1/2}$. The normalized Chebyshev polynomials satisfy the bounds (due to (383) and the normalization constants)

$$\left|\overline{T_0}(x)\right| \leq \sqrt{\frac{1}{\pi}} \quad (x \in [-1,1]) \tag{389}$$

$$\left|\overline{T_k}(x)\right| \leq \sqrt{\frac{2}{\pi}} \quad (x \in [-1,1],\ k = 1,2,3,\ldots). \tag{390}$$

To make use of these normalized Chebyshev polynomials, I require a method to evaluate them and to evaluate their first derivatives.

Method 9.1. This method computes the result of evaluating the normalized Chebyshev polynomials $\overline{T_j}$ for $j = 0, 1, \ldots n$ at each of the elements of a column vector $\boldsymbol{x}$.

- **Given**:
  1. n; the normalized Chebyshev polynomials $\overline{T_k}$ for $k = 0, 1, \ldots, n$ will be evaluated at the given points
  2. $\boldsymbol{x}$, a column vector of points at which to evaluate the normalized Chebyshev polynomials
- **Output**:
  1. V_{Cheb}, a matrix with the same number of rows as there are elements in $\boldsymbol{x}$ and with $n+1$ columns whose (i,j) element (for $i,j = 0,1,2,\ldots$, so that indexing begins with 0) is $\overline{T_j}(x_i)$

Step 1 Check the arguments.

Step 2 Allocate space for V_{Cheb}, sized as described above.

Step 3 Set the leftmost column of V_{Cheb} equal to a conforming column vector of ones multiplied by the scalar $\sqrt{\frac{2}{\pi}}$.

Step 4 If $n > 0$, set the second column (from the left) of V_{Cheb} equal to the column vector $\boldsymbol{x}$.

Step 5 Index the columns of V_{Cheb} starting with 0, so that the third column of V_{Leg} from the left is indexed as 2. For column indices j starting at 2 and ending at n, taken in order, execute the following loop step (if $n < 2$, do not execute the loop step at all).

Loop Step 1 Let $\boldsymbol{V}_{\text{Cheb}}^{(k)}$ denote column k of V_{Cheb}, $k = 0, 1, \ldots, n$. Set

$$\boldsymbol{V}_{\text{Cheb}}^{(j)} = 2\boldsymbol{x} \circ \boldsymbol{V}_{\text{Cheb}}^{(j-1)} - \boldsymbol{V}_{\text{Cheb}}^{(j-2)}, \tag{391}$$

where $\circ$ denotes elementwise multiplication. This is, of course, a matrix version of the functional three term recurrence (381).

Step 6 Divide the leftmost column of V_{Cheb} by $\sqrt{2}$. This normalizes the evaluations using the same scaling as (387).

Method 9.2. This method computes the result of evaluating the first derivatives of the normalized Chebyshev polynomials $\overline{T_j}$ for $j = 0, 1, \ldots n$ at each of the elements of a column vector $\boldsymbol{x}$.

- **Given**:
  1. n; the first derivatives of the normalized Chebyshev polynomials $\overline{T_k}$ for $k = 0, 1, \ldots, n$ will be evaluated at the given points
  2. $\boldsymbol{x}$, a column vector of points at which to evaluate the first derivatives of the normalized Chebyshev polynomials
- **Output**:
  1. $V_{\text{ChebDeriv}}$, a matrix with the same number of rows as there are elements in $\boldsymbol{x}$ and with $n + 1$ columns whose (i, j) element (for $i, j = 0, 1, 2, \ldots$, so that indexing begins with 0) is $\frac{d\overline{T_j}}{dx}(x_i)$

Step 1 Check the arguments.

Step 2 Use Method 9.1 to obtain V_{Cheb}, a matrix with the same number of rows as there are elements in $\boldsymbol{x}$ and with $n + 1$ columns whose (i, j) element (for $i, j = 0, 1, 2, \ldots$, so that indexing begins with 0) is $\overline{T_j}(x_i)$.

Step 3 Allocate space for $V_{\text{ChebDeriv}}$, sized as described above.

Step 4 Set the leftmost column of $V_{\text{ChebDeriv}}$ equal to a conforming column vector of zeros.

Step 5 If $n > 0$, set the second column (from the left) of $V_{\text{ChebDeriv}}$ equal to a conforming column vector of ones multiplied by the scalar $\sqrt{\frac{2}{\pi}}$.

Step 6 Index the columns of $V_{\text{ChebDeriv}}$ starting with 0, so that the third column of $V_{\text{ChebDeriv}}$ from the left is indexed as 2. For column indices j starting at 2 and ending at n, taken in order, execute the following loop step (if $n < 2$, do not execute the loop step at all).

Loop Step 1 Let $\boldsymbol{V}^{(k)}_{\text{ChebDeriv}}$ denote column k of $V_{\text{ChebDeriv}}$, $k = 0, 1, \ldots, n$. Set

$$\boldsymbol{V}^{(j)}_{\text{ChebDeriv}} = 2\boldsymbol{x} \circ \boldsymbol{V}^{(j-1)}_{\text{ChebDeriv}} + 2\boldsymbol{V}^{(j-1)}_{\text{Cheb}} - \boldsymbol{V}^{(j-2)}_{\text{ChebDeriv}}, \tag{392}$$

where $\circ$ denotes elementwise multiplication. This is a matrix version of the equation that results from differentiating the functional three term recurrence (381). Note that the second term on the righthand side of (392) is a column of V_{Cheb}, not a column of $V_{\text{ChebDeriv}}$.

Step 7 There is no need to scale the leftmost column of $V_{\text{ChebDeriv}}$ in order to achieve normalization: it is zero.

I can now write the Fourier series expansions of the eigenfunctions (378) and (379) as normalized Chebyshev polynomial series instead.

$$\xi_{2n}^{c}(t) = \sqrt{2}A_{0}^{2n}(c)\,\overline{T_0}(\cos(t)) + \sum_{m=1}^{\infty} A_{2m}^{2n}(c)\,\overline{T_{2m}}(\cos(t)) \tag{393}$$

$$\xi_{2n+1}^{c}(t) = \sum_{m=0}^{\infty} A_{2m+1}^{2n+1}(c)\,\overline{T_{2m+1}}(\cos(t)), \tag{394}$$

and this suggests defining the coefficients so that the first (subscript zero) coefficient is not multiplied by an additional scale factor of $\sqrt{2}$. I define

$$\widetilde{A}_{0}^{2n} = \sqrt{2}A_{0}^{2n} \tag{395}$$

$$\widetilde{A}_{2m}^{2n} = A_{2m}^{2n} \quad (m = 1, 2, 3, \ldots) \tag{396}$$

$$\widetilde{A}_{2m+1}^{2n+1} = A_{2m+1}^{2n+1} \quad (m = 0, 2, 3, \ldots). \tag{397}$$

With these definitions, I have

$$\xi_{2n}^{c}(t) = \sum_{m=0}^{\infty} \widetilde{A}_{2m}^{2n}(c)\,\overline{T_{2m}}(\cos(t)) \tag{398}$$

$$\xi_{2n+1}^{c}(t) = \sum_{m=0}^{\infty} \widetilde{A}_{2m+1}^{2n+1}(c)\,\overline{T_{2m+1}}(\cos(t)) \tag{399}$$

by subsitution into (393) and (394). The recurrence relations and normalizations of the new coefficients change, for the even indices, in a way which turns out to be helpful:

$$a_{2n}(c)\,\widetilde{A}_{0}^{2n}(c) + \frac{c^2}{2\sqrt{2}}\widetilde{A}_{2}^{2n}(c) = 0 \tag{400}$$

$$(a_{2n}(c) - 4)\,\widetilde{A}_{2}^{2n}(c) + \frac{c^2}{2\sqrt{2}}\widetilde{A}_{0}^{2n}(c) + \frac{c^2}{4}\widetilde{A}_{4}^{2n}(c) = 0 \tag{401}$$

(where (400) comes from multiplying both sides of the equality (371) by $\sqrt{2}$ and (401) comes from simple substitution into (372)) and, for $m = 2, 3, 4, \ldots$,

$$\left(a_{2n}(c) - 4m^2\right)\widetilde{A}_{2m}^{2n}(c) + \frac{c^2}{4}\left(\widetilde{A}_{2m-2}^{2n}(c) + \widetilde{A}_{2m+2}^{2n}(c)\right) = 0 \tag{402}$$

for the even indices $2n$. For the odd indices $2n+1$, nothing changes:

$$\left(a_{2n+1}(c) - 1 + \frac{c^2}{4}\right)\widetilde{A}_{1}^{2n+1}(c) + \frac{c^2}{4}\widetilde{A}_{3}^{2n+1}(c) = 0 \tag{403}$$

and, for $m = 1, 2, 3, \ldots$,

$$\left(a_{2n}(c) - (2m+1)^2\right)\widetilde{A}_{2m+1}^{2n+1}(c) + \frac{c^2}{4}\left(\widetilde{A}_{2m-1}^{2n+1}(c) + \widetilde{A}_{2m+3}^{2n+1}(c)\right) = 0. \tag{404}$$

Finally, the new coefficients are very naturally normalized:

$$\sum_{m=0}^{\infty} \left(\widetilde{A}_{2m}^{2n}(c) \right)^2 = 1 \tag{405}$$

$$\sum_{m=0}^{\infty} \left(\widetilde{A}_{2m+1}^{2n+1}(c) \right)^2 = 1. \tag{406}$$

For the even indices $2n$, the equations (400), (401), and (402) form an (infinite) eigenvalue-eigenvector relationship for the $\left\{ \widetilde{A}_{2m}^{2n}(c) \right\}_{m=0}^{\infty}$ (the eigenvector) and $a_{2n}(c)$ (the eigenvalue); recall that the (infinite) vector $\left\{ \widetilde{A}_{2m}^{2n}(c) \right\}_{m=0}^{\infty}$ is also properly normalized by (405). The (infinite) matrix for which $\left\{ \widetilde{A}_{2m}^{2n}(c) \right\}_{m=0}^{\infty}$ is an eigenvector with eigenvalue $a_{2n}(c)$ is:

$$\mathcal{A}_{\text{MathieuEven}}^{c} \equiv \begin{pmatrix} 0 & -\frac{c^2}{2\sqrt{2}} & 0 & 0 & 0 & \cdots \\ -\frac{c^2}{2\sqrt{2}} & 4 & -\frac{c^2}{4} & 0 & 0 & \cdots \\ 0 & -\frac{c^2}{4} & 16 & -\frac{c^2}{4} & 0 & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \end{pmatrix}. \tag{407}$$

This is a symmetric tridiagonal infinite matrix which, indexing beginning from 0, has $(2i)^2$ as the diagonal element in row i and has $-\frac{c^2}{4}$ along the subdiagonal and the superdiagonal, with the sole exceptions of the $(1,0)$ entry and the $(0,1)$ entry that equals it, both of these being $-\frac{c^2}{2\sqrt{2}}$. It is interesting to compare this symmetric tridiagonal matrix to the one used by Zhang & Jin (1996), who also take the tridiagonal-eigenproblem approach to finding the Mathieu eigenvalues and coefficients, but work with the original coefficients A_{2m}^{2n}, so that the tridiagonal matrix in their eigenproblem is not symmetric (see their equation 14.2.9 on page 479 in Section 2 of Chapter 14).

Applying the same logic to the odd indices $2n+1$, the equations (403) and (404) form an (infinite) eigenvalue-eigenvector relationship for the $\left\{ \widetilde{A}_{2m+1}^{2n+1}(c) \right\}_{m=0}^{\infty}$ (the eigenvector) and $a_{2n+1}(c)$ (the eigenvalue); recall that the (infinite) vector $\left\{ \widetilde{A}_{2m+1}^{2n+1}(c) \right\}_{m=0}^{\infty}$ is also properly normalized by (406). The (infinite) matrix for which $\left\{ \widetilde{A}_{2m+1}^{2n+1}(c) \right\}_{m=0}^{\infty}$ is an eigenvector with eigenvalue $a_{2n+1}(c)$ is:

$$\mathcal{A}_{\text{MathieuOdd}}^{c} \equiv \begin{pmatrix} \left(1 - \frac{c^2}{4}\right) & -\frac{c^2}{4} & 0 & 0 & 0 & \cdots \\ -\frac{c^2}{4} & 9 & -\frac{c^2}{4} & 0 & 0 & \cdots \\ 0 & -\frac{c^2}{4} & 25 & -\frac{c^2}{4} & 0 & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \end{pmatrix}. \tag{408}$$

This is a symmetric tridiagonal infinite matrix which, indexing beginning from 0, has $-\frac{c^2}{4}$ along the subdiagonal and the superdiagonal and has $(2i+1)^2$ as the diagonal element in row i with the sole exception of the $(0,0)$ entry, which is $1-\frac{c^2}{4}$.

First consider the even indices $2n$, $n=0,1,2,\ldots$. To obtain approximations to $\left\{\widetilde{A}_{2m}^{2n}(c)\right\}_{m=0}^{\infty}$ for $n=0,1,\ldots,N$ (for any $N\geq 0$), and thus to obtain approximations to $\xi_{2n}^{c}(t)$ for $n=0,1,\ldots,N$, truncate the infinite symmetric tridiagonal matrix $\mathcal{A}_{\text{MathieuEven}}^{c}$ at some size K_{even} substantially larger than N to obtain the $K_{\text{even}}\times K_{\text{even}}$ symmetric tridiagonal matrix $\mathcal{A}_{\text{MathieuEven}}^{\text{trunc}}$. In general, the appropriate truncation size depends on both N and c (and is increasing in both). A diagonal length of 75 seems sufficient in all cases of practical financial interest. With this truncated, finite symmetric tridiagonal matrix, applying any standard numerical method for the computation of the eigenvalues and eigenvectors of a symmetric tridiagonal matrix will deliver the truncated approximations to $a_{2n}(c)$, and $\left\{\widetilde{A}_{2m}^{2n}(c)\right\}_{m=0}^{K_{\text{even}}}$. These, in turn, deliver approximations to $\xi_{2n}^{c}(t)$ for $n=0,1,\ldots,N$.

Now turn to the odd indices $2n+1$, $n=0,1,2,\ldots$. To obtain approximations to $\left\{\widetilde{A}_{2m+1}^{2n+1}(c)\right\}_{m=0}^{\infty}$ for $n=0,1,\ldots,N$ (for any $N\geq 0$), and thus to obtain approximations to $\xi_{2n+1}^{c}(t)$ for $n=0,1,\ldots,N$, truncate the infinite symmetric tridiagonal matrix $\mathcal{A}_{\text{MathieuOdd}}^{c}$ at some size K_{odd} substantially larger than N (and differing from K_{even} by at most one) to obtain the $K_{\text{odd}}\times K_{\text{odd}}$ symmetric tridiagonal matrix $\mathcal{A}_{\text{MathieuOdd}}^{\text{trunc}}$. As for the even indices, a diagonal length of 75 seems sufficient in all cases of practical financial interest. With this truncated, finite symmetric tridiagonal matrix, applying any standard numerical method for the computation of the eigenvalues and eigenvectors of a symmetric tridiagonal matrix will deliver the truncated approximations to $a_{2n+1}(c)$, and $\left\{\widetilde{A}_{2m+1}^{2n+1}(c)\right\}_{m=0}^{K_{\text{odd}}}$. These, in turn, deliver approximations to $\xi_{2n+1}^{c}(t)$ for $n=0,1,\ldots,N$.

Note that the computational cost of these two invocations of a tridiagonal eigensolver will not dominate the time required to evaluate approximations to the exponential cosine product operator, since the solutions to the eigenproblems can be precalculated.

The discussion above provides the foundation for the following methods.

Method 9.3. The following method computes a given number of the eigenvalues of the Mathieu operator M_c and the coefficients of series in (normalized) Chebyshev polynomials for the corresponding eigenfunctions (the $\xi_n^c(t)$, which are just normalized versions of the Mathieu functions $\mathrm{ce}_n\left(t,-\frac{c^2}{4}\right)$).

- **Given**:
    1. n, the number of eigenvalues to compute (and also the number of corresponding eigenfunctions for which the coefficients of series in (normalized) Chebyshev polynomials are to be calculated)
    2. $c>0$

3. a truncation size M, which must be greater than n

- **Output**:

1. $\boldsymbol{a}$, a $n \times 1$ vector of the n smallest eigenvalues of M_c, sorted in ascending order
2. $\mathcal{B}_{\text{Mathieu}}$, a matrix with n columns and number of rows equal to M (if M is odd) or $M+1$ (if M is even); column j, indexed starting at 0, of $\mathcal{B}_{\text{Mathieu}}$ is a computation of the column vector $\left(\widetilde{A}_m^j(c)\right)_{m=0}^{M_B}$ (where $\widetilde{A}_m^j(c)$ is zero unless j and m are either both even or both odd), so column j of $\mathcal{B}_{\text{Mathieu}}$ is a computation of the initial M_B coefficients in an expansion of ξ_j^c in (normalized) Chebyshev polynomials (about half of which are zero, since odd-index coefficients are zero for j even and even-index coefficients are zero for j odd), and M_B is either M (for M odd) or $M+1$ (for M even)

Step 1 Check the arguments.

Step 2 If the truncation size M is even, set the truncation size for evens $M_{\text{even}} = M$ and set the truncation size for odds $M_{\text{odd}} = M - 1$. If the truncation size M is odd, set the truncation size for evens $M_{\text{even}} = M - 1$ and set the truncation size for odds $M_{\text{odd}} = M$. This assures that, overall, indices from 0 through M will be included in subsequent computations.

Step 3 Generate a vector of the even numbers from 0 through M_{even} (including 0 and M_{even}), in order. Label this vector I_{even}; it contains $\frac{1}{2}(M_{\text{even}} + 2)$ elements.

Step 4 Generate a vector of the odd numbers from 1 through M_{odd} (including 1 and M_{odd}), in order. Label this vector I_{odd}; it contains $\frac{1}{2}(M_{\text{odd}} + 1)$ elements.

Step 5 Form, as a sparse matrix to avoid unneccessary storage of zeros, the $\frac{1}{2}(M_{\text{even}} + 2) \times \frac{1}{2}(M_{\text{even}} + 2)$ symmetric tridiagonal matrix $\mathcal{A}_{\text{MathieuEven}}^{\text{trunc}}$ that represents the top left corner of the infinite symmetric tridiagonal matrix $\mathcal{A}_{\text{MathieuEven}}^{c}$ given by (407). Only three vectors need be stored: the diagonal, the subdiagonal, and the superdiagonal (which equals the subdiagonal by symmetry; if the sparse matrix framework used allowed flagging for symmetry, there would be no need to store the superdiagonal).

Step 6 Form, as a sparse matrix to avoid unneccessary storage of zeros, the $\frac{1}{2}(M_{\text{odd}} + 1) \times \frac{1}{2}(M_{\text{odd}} + 1)$ symmetric tridiagonal matrix $\mathcal{A}_{\text{MathieuOdd}}^{\text{trunc}}$ that represents the top left corner of the infinite symmetric tridiagonal matrix $\mathcal{A}_{\text{MathieuOdd}}^{c}$ given by (408). Only three vectors need be stored: the diagonal, the subdiagonal, and the superdiagonal (which equals the subdiagonal by symmetry; if the sparse matrix framework used allowed flagging for symmetry, there would be no need to store the superdiagonal).

Step 7 Using a standard eigenvalue and eigenvector routine, compute the eigenvalues and eigenvectors of $\mathcal{A}^{\text{trunc}}_{\text{MathieuEven}}$. It is more efficient, but not necessary, to use a routine specifically designed for symmetric tridiagonal matrices. Store the eigenvalues in a $\frac{1}{2}(M_{\text{even}}+2)\times\frac{1}{2}(M_{\text{even}}+2)$ diagonal matrix a_{even}, and store the eigenvectors as the columns of a matrix $\mathcal{B}_{\text{MathieuEven}}$ of the same size.

Step 8 Using a standard eigenvalue and eigenvector routine, compute the eigenvalues and eigenvectors of $\mathcal{A}^{\text{trunc}}_{\text{MathieuOdd}}$. It is more efficient, but not necessary, to use a routine specifically designed for symmetric tridiagonal matrices. Store the eigenvalues in a $\frac{1}{2}(M_{\text{odd}}+1)\times\frac{1}{2}(M_{\text{odd}}+1)$ diagonal matrix a_{odd}, and store the eigenvectors as the columns of a matrix $\mathcal{B}_{\text{MathieuOdd}}$ of the same size.

Step 9 If $\frac{1}{2}(M_{\text{even}}+2) > \frac{1}{2}(M_{\text{odd}}+1)$, then increase the number of rows and the number of columns of both a_{odd} and $\mathcal{B}_{\text{MathieuOdd}}$ by 1; set all of the entries in the last row and in the last column of each of these newly-enlarged matrices to 0. If $\frac{1}{2}(M_{\text{even}}+2) < \frac{1}{2}(M_{\text{odd}}+1)$, do the same but for the corresponding "even" matrices (this will never occur given the steps above, but is included as a robustness feature if those steps were ever to change). Note that the two sizes being compared can never differ by more than 1.

Step 10 Set

$$\widetilde{\boldsymbol{a}} \quad = \quad \operatorname{diag}\left(a_{\text{even}}\otimes\begin{pmatrix}1&0\\0&0\end{pmatrix}+a_{\text{odd}}\otimes\begin{pmatrix}0&0\\0&1\end{pmatrix}\right), \tag{409}$$

where the symbol $\otimes$ denotes, as usual, the Kronecker product of two matrices and the function "diag" extracts the diagonal of a matrix and returns it as a column vector. Then set the output variable $\boldsymbol{a}$ to be equal to the first n elements of $\widetilde{\boldsymbol{a}}$.

Step 11 Set

$$\widetilde{\mathcal{B}}_{\text{Mathieu}} \quad = \quad \mathcal{B}_{\text{MathieuEven}}\otimes\begin{pmatrix}1&0\\0&0\end{pmatrix}+\mathcal{B}_{\text{MathieuOdd}}\otimes\begin{pmatrix}0&0\\0&1\end{pmatrix}, \tag{410}$$

where the symbol $\otimes$ denotes, as in Step 10, the Kronecker product of two matrices. Then set the output variable $\mathcal{B}_{\text{Mathieu}}$ to be equal to the first n columns of $\widetilde{\mathcal{B}}_{\text{Mathieu}}$, but with the final row deleted.

Remark 9.1. I could make Method 9.3 more efficient in a number of places, since the output variable $\mathcal{B}_{\text{Mathieu}}$ has roughly half of its elements equal to zero in a pattern that is known in advance (each column has an alternating non-zero / zero pattern, with the first element being zero for odd-indexed columns and non-zero for even-indexed columns when indexing is started at 0). Because Method 9.3 can be used for precalculation in approximation of the exponential cosine product operator, I do not pursue further refinements here.

I must now calculate the eigenvalues $\gamma_n(c)$, $n = 0, 1, 2, \ldots$, of $P_c^{\cos}$. By evaluating the basic eigenvalue-eigenfunction identity

$$\gamma_n(c)\,\xi_n^c(u) = \int_0^\pi \exp(c\cos(t)\cos(u))\xi_n^c(t)\,dt \tag{411}$$

at $u = \frac{\pi}{2}$, I obtain (proceeding as I did for the exponential product operator in Subsection 8.4)

$$\begin{aligned}
\gamma_n(c)\,\xi_n^c\left(\frac{\pi}{2}\right) &= \int_0^\pi \xi_n^c(\cos(t))\,dt & (412)\\
&= \int_0^\pi \sum_{k=0}^\infty \widetilde{A}_k^n(c)\,\overline{T_k}(\cos(t))\,dt & (413)\\
&= \sum_{k=0}^\infty \widetilde{A}_k^n(c)\int_0^\pi \overline{T_k}(\cos(t))\,dt & (414)\\
&= \widetilde{A}_0^n(c)\sqrt{\pi}, & (415)
\end{aligned}$$

where the first equality follows from evaluating (411) at $u = \frac{\pi}{2}$ and noting that $\exp(0) = 1$, the second equality follows from substituting the series expression (393) or (394) (depending on whether n is even or odd) into the prior expression (where it is understood that $\widetilde{A}_k^n(c)$ is zero if n is even and k is odd or if n is odd and k is even), the third equality employs a semidiscrete form of Fubini's theorem, and the fourth equality follows from the fact that $\overline{T_0}(\cos(t)) = \sqrt{\frac{1}{\pi}}$, so $\int_0^\pi \overline{T_k}(\cos(t))\,dt = \int_0^\pi \overline{T_k}(\cos(t))\sqrt{\pi}\overline{T_0}(\cos(t))\,dt$, which is zero (by orthogonality) if $k \neq 0$ and is $\sqrt{\pi}$ if $k = 0$ (since the $\overline{T_k}(\cos(t))$ are both orthogonal over $[0, \pi]$, because they are normalizations by constants of the usual Chebyshev polynomials T_k and $T_k(\cos(t)) = \cos(kt)$ by definition, and also orthonormal over $[0, \pi]$ because of the normalizations chosen).

For n even, $\xi_n^c\left(\frac{\pi}{2}\right) \neq 0$, so a valid calculation for $\gamma_n(c)$ for n even comes from dividing (415) by $\xi_n^c\left(\frac{\pi}{2}\right)$:

$$\gamma_n(c) = \frac{\widetilde{A}_0^n(c)}{\xi_n^c\left(\frac{\pi}{2}\right)}\sqrt{\pi} \quad \text{for } n \text{ even.} \tag{416}$$

When n is odd, $\xi_n^c\left(\frac{\pi}{2}\right)$ is zero and $\widetilde{A}_0^n(c)$ is zero, so (416) is not helpful in determining $\gamma_n(c)$. Instead, consider differentiating both sides of (411) with respect to u and then evaluating the resulting expression at $u = \frac{\pi}{2}$:

$$\begin{aligned}
\gamma_n(c)\,\frac{d\xi_n^c}{du}\left(\frac{\pi}{2}\right) &= -\int_0^\pi c\sin\left(\frac{\pi}{2}\right)\cos(t)\xi_n^c(t)\,dt & (417)\\
&= -\int_0^\pi c\sin\left(\frac{\pi}{2}\right)\cos(t)\sum_{k=0}^\infty \widetilde{A}_k^n(c)\,\overline{T_k}(\cos(t))\,dt & (418)\\
&= -c\sum_{k=0}^\infty \widetilde{A}_k^n(c)\int_0^\pi \cos(t)\overline{T_k}(\cos(t))\,dt & (419)
\end{aligned}$$

$$= \quad -\widetilde{A}_1^n(c)\, c\sqrt{\frac{\pi}{2}}, \tag{420}$$

where the first equality follows from differentiation with respect to u and then evaluation at $u = \frac{\pi}{2}$ of (411), the second equality follows from substituting the series expression (393) or (394) (depending on whether n is even or odd) into the prior expression, the third equality employs a semidiscrete form of Fubini's theorem (and factoring the constant $-c\sin\left(\frac{\pi}{2}\right) = -c$ out of the integration and the summation), and the fourth equality follows from the fact that $\overline{T_1}(\cos(t)) = \cos(t)\sqrt{\frac{2}{\pi}}$, so $\int_0^\pi \cos(t)\overline{T_k}(\cos(t))\,dt = \int_0^\pi \overline{T_k}(\cos(t))\sqrt{\frac{\pi}{2}}\overline{T_1}(\cos(t))\,dt$, which is zero (by orthogonality) if $k \neq 1$ and is $\sqrt{\frac{\pi}{2}}$ if $k = 1$ (since the $\overline{T_k}(\cos(t))$ are both orthogonal over $[0,\pi]$, because they are normalizations by constants of the usual Chebyshev polynomials T_k and $T_k(\cos(t)) = \cos(kt)$ by definition, and also orthonormal over $[0,\pi]$ because of the normalizations chosen).

For n odd, $\frac{d\xi_n^c}{du}\left(\frac{\pi}{2}\right) \neq 0$, so a valid calculation for $\gamma_n(c)$ for n odd comes from dividing (420) by $-\frac{d\xi_n^c}{du}\left(\frac{\pi}{2}\right)$:

$$\gamma_n(c) \quad = \quad \frac{\widetilde{A}_1^n(c)}{-\frac{d\xi_n^c}{du}\left(\frac{\pi}{2}\right)} c\sqrt{\frac{\pi}{2}} \quad \text{for } n \text{ odd}. \tag{421}$$

With both the eigenfunctions ξ_m^c and the eigenvalues $\gamma_m(c)$ of $P_c^{\cos}$ available, I can construct an L^2-optimal rank-n approximation $Q_n^{\cos}$ (the analog to Q_n of Theorem 8.1, so optimality means achieving the linear n-width) to $P_c^{\cos}$ by building a truncated eigenfunction expansion with n terms.

$$Q_n^{\cos}(t,u) \quad \equiv \quad \sum_{i=0}^{n-1} \gamma_i(c)\,\xi_i^c(t)\,\xi_i^c(u). \tag{422}$$

This truncated eigenfunction expansion converges to $\exp(c\cos(t)\cos(u))$ in L^2; by Mercer's theorem, the continuity of the kernel $\exp(c\cos(t)\cos(u))$ implies that it also converges uniformly to $\exp(c\cos(t)\cos(u))$ for any $u, t \in [0,\pi]$.

If I provided methods to evaluate both the ξ_m^c and their first derivatives, I could then provide a method constructing the L^2-optimal approximation $Q_n^{\cos}$. However, my goal in this section is to obtain a near-optimal approximation to P_c (the exponential product operator) in worst-case terms; my L^2 analysis of $P_c^{\cos}$ is a means to this end. With that in mind, I now show how to return to the original variables $x = \cos(t)$ and $y = \cos(u)$.

Because (425) (for even indices) and (426) (for odd indices) express the $\xi_j^c(t)$ as series in functions that depend on t only through $\cos(t)$, the ξ_j^c themselves depend on t only through $\cos(t)$. This permits me to write

$$\mu_{2n}^c(\cos(t)) \equiv \quad \xi_{2n}^c(t) = \quad \sum_{m=0}^{\infty} \widetilde{A}_{2m}^{2n}(c)\,\overline{T_{2m}}(\cos(t)) \tag{423}$$

$$\mu_{2n+1}^c(\cos(t)) \equiv \quad \xi_{2n+1}^c(t) = \quad \sum_{m=0}^{\infty} \widetilde{A}_{2m+1}^{2n+1}(c)\,\overline{T_{2m+1}}(\cos(t)). \tag{424}$$

I now substitute $x = \cos(t)$ into (425) and (426) to obtain

$$\mu^c_{2n}(x) = \sum_{m=0}^{\infty} \widetilde{A}^{2n}_{2m}(c)\,\overline{T_{2m}}(x) \tag{425}$$

$$\mu^c_{2n+1}(x) = \sum_{m=0}^{\infty} \widetilde{A}^{2n+1}_{2m+1}(c)\,\overline{T_{2m+1}}(x)\,. \tag{426}$$

Since the approximation $Q^{\cos}_n$ defined in (422) is a series in $\xi^c_j(t)$ and $\xi^c_j(u)$, it depends on t and u only through $x = \cos(t)$ and $y = \cos(u)$ and it can be expressed in these terms. I adopt a separate notation for the x, y version of the approximation:

$$\begin{aligned} P^{\text{Mathieu}}_n(x,y) &\equiv Q^{\cos}_n(\arccos(x), \arccos(y)) \\ &= \sum_{i=0}^{n-1} \gamma_i(c)\,\mu^c_i(x)\,\mu^c_i(y)\,. \end{aligned} \tag{427}$$

It remains to rephrase the relations (416) and (421) in terms of the μ^c_j and their first derivatives rather than the ξ^c_j and their first derivatives. To do so, note that $\xi^c_j\left(\frac{\pi}{2}\right) = \mu^c_j\left(\cos\left(\frac{\pi}{2}\right)\right) = \mu^c_j(0)$ and $-\frac{d\xi^c_j}{dt}\left(\frac{\pi}{2}\right) = -\frac{d\mu^c_j}{dx}\left(\cos\left(\frac{\pi}{2}\right)\right)\left(-\sin\left(\frac{\pi}{2}\right)\right) = \frac{d\mu^c_j}{dx}(0)$, where the first equality for the derivative follows from an application of the chain rule of differentiation. Substitution of these results into (416) and (421) yields

$$\gamma_n(c) = \frac{\widetilde{A}^n_0(c)}{\mu^c_n(0)}\sqrt{\pi} \quad \text{for } n \text{ even} \tag{428}$$

$$\gamma_n(c) = \frac{\widetilde{A}^n_1(c)}{\frac{d\mu^c_n}{dx}(0)}c\sqrt{\frac{\pi}{2}} \quad \text{for } n \text{ odd.} \tag{429}$$

I now provide methods to evaluate the μ^c_j and to evaluate their first derivatives. In using these methods, it is helpful to note that the $\widetilde{A}^{2n}_{2m}(c)$ and the $\widetilde{A}^{2n+1}_{2m+1}(c)$ are not changed by passing from t to $x = \cos(t)$, so Method 9.3 gives a computation of them.

Method 9.4. This method computes the result of evaluating the μ^c_j for $j = 0, 1, \ldots n-1$ at each of the elements of a column vector $\boldsymbol{x}$, given a matrix of (normalized) Chebyshev series coefficients such as the one produced by Method 9.3.

- **Given**:
  1. $\mathcal{B}_{\text{Mathieu}}$, a matrix with n columns and number of rows equal to M (if M is odd) or $M+1$ (if M is even); column j, indexed starting at 0, of $\mathcal{B}_{\text{Mathieu}}$ is a computation of the column vector $\left(\widetilde{A}^j_m(c)\right)_{m=0}^{M_B}$ (where $\widetilde{A}^j_m(c)$ is zero unless j and m are either both even or both odd), so

column j of $\mathcal{B}_{\text{Mathieu}}$ is a computation of the initial M_B coefficients in an expansion of μ_j^c in (normalized) Chebyshev polynomials (about half of which are zero, since odd-index coefficients are zero for j even and even-index coefficients are zero for j odd), and M_B is either M (for M odd) or $M+1$ (for M even)

2. $\boldsymbol{x}$, a column vector of points at which to evaluate the ξ_j^c.

- **Output**:

1. $\mathcal{F}$, a matrix with n columns and number of rows equal to the length of $\boldsymbol{x}$; the (i,j) element of $\mathcal{F}$, indexed starting at 0, is $\mu_j^c(x_i)$.

Step 1 Check the arguments.

Step 2 Invoke Method 9.1 with inputs $M_B - 1$ and $\boldsymbol{x}$ to obtain the matrix V_{Cheb} which has M_B columns and a number of rows equal to the length of $\boldsymbol{x}$.

Step 3 Set the output variable $\mathcal{F} = V_{\text{Cheb}}\mathcal{B}_{\text{Mathieu}}$. Note that this results in $\mathcal{F}$ having the proper size.

Method 9.5. This method computes the result of evaluating the first derivatives of the μ_j^c for $j = 0, 1, \ldots n-1$ at each of the elements of a column vector $\boldsymbol{x}$, given a matrix of (normalized) Chebyshev series coefficients such as the one produced by Method 9.3.

- **Given**:

1. $\mathcal{B}_{\text{Mathieu}}$, a matrix with n columns and number of rows equal to M (if M is odd) or $M+1$ (if M is even); column j, indexed starting at 0, of $\mathcal{B}_{\text{Mathieu}}$ is a computation of the column vector $\left(\widetilde{A}_m^j(c)\right)_{m=0}^{M_B}$ (where $\widetilde{A}_m^j(c)$ is zero unless j and m are either both even or both odd), so column j of $\mathcal{B}_{\text{Mathieu}}$ is a computation of the initial M_B coefficients in an expansion of μ_j^c in (normalized) Chebyshev polynomials (about half of which are zero, since odd-index coefficients are zero for j even and even-index coefficients are zero for j odd), and M_B is either M (for M odd) or $M+1$ (for M even)
2. $\boldsymbol{x}$, a column vector of points at which to evaluate the first derivatives of the ξ_j^c.

- **Output**:

1. $\mathcal{G}$, a matrix with n columns and number of rows equal to the length of $\boldsymbol{x}$; the (i,j) element of $\mathcal{G}$, indexed starting at 0, is the first derivative of $\mu_j^c(x_i)$.

Step 1 Check the arguments.

Step 2 Invoke Method 9.2 with inputs $M_B - 1$ and $\boldsymbol{x}$ to obtain the matrix $V_{\text{ChebDeriv}}$ which has M_B columns and a number of rows equal to the length of $\boldsymbol{x}$.

Step 3 Set the output variable $\mathcal{G} = V_{\text{ChebDeriv}} \mathcal{B}_{\text{Mathieu}}$. Note that this results in $\mathcal{G}$ having the proper size.

Remark 9.2. In the evaluation of (truncated) series whose terms satisfy a three-term recurrence relation (as both Chebyshev and Legendre polynomials do), such as those for the μ_j^c, it is typical to use the approach introduced by Clenshaw (1955) (see the helpful discussion of this method in Gautschi (2004), pages 78-82 in Section 2.1.8.1) to increase computational efficiency. When I developed and tested Clenshaw's method, however, it appeared slower than the direct approach laid out above. I believe this is likely due to the fact that I typically wish to evaluate a number of such (truncated) series at many different points simultaneously, and the direct method I describe above allows level-3 BLAS operations (matrix-matrix computations), with their typically excellent performance on modern computers, to be used in such an evaluation. Clenshaw's method allows for simultaneous evaluation of one series at many points, but I have not yet found a way to employ it to efficiently and simultaneously evaluate many series at many points.

All of the necessary methods are finally in place, and I can present a method for constructing the Mathieu approximation P_n^{Mathieu}.

Method 9.6. This method computes the components of the approximation P_n^{Mathieu}: the first (largest) n of the $\gamma_j(c)$ (the eigenvalues of the exponential cosine product operator) and the coefficients of the (normalized) Chebyshev series expansions of the corresponding first n μ_j^c; Method 9.3 is invoked to obtain the series expansion coefficients.

- **Given**:
  1. n, the number of $\gamma_j(c)$ to compute (and also the number of corresponding μ_j^c for which the coefficients of series in (normalized) Chebyshev polynomials are to be calculated)
  2. $c > 0$
  3. a truncation size M, which must be greater than n
- **Output**:
  1. $\boldsymbol{\gamma}$, a $n \times 1$ vector of the n largest $\gamma_j(c)$, sorted in descending order
  2. $\mathcal{B}_{\text{Mathieu}}$, a matrix with n columns and number of rows equal to M (if M is odd) or $M+1$ (if M is even); column j, indexed starting at 0, of $\mathcal{B}_{\text{Mathieu}}$ is a computation of the column vector $\left(\widetilde{A}_m^j(c)\right)_{m=0}^{M_B}$ (where $\widetilde{A}_m^j(c)$ is zero unless j and m are either both even or both odd), so column j of $\mathcal{B}_{\text{Mathieu}}$ is a computation of the initial M_B coefficients

in an expansion of μ_j^c in (normalized) Chebyshev polynomials (about half of which are zero, since odd-index coefficients are zero for j even and even-index coefficients are zero for j odd), and M_B is either M (for M odd) or $M+1$ (for M even)

Simply combine the methods and logic above.

A final note is useful in bounding the worst-case approximation error of the Mathieu approximation. For any $x, y \in [-1, 1]$ the approximation error is:

$$
\begin{aligned}
&\left|\exp(cxy) - \sum_{j=0}^{n-1} \gamma_j(c)\, \mu_j^c(x)\, \mu_j^c(y)\right| \\
&= \left|\sum_{j=0}^{\infty} \gamma_j(c)\, \mu_j^c(x)\, \mu_j^c(y) - \sum_{j=0}^{n-1} \gamma_j(c)\, \mu_j^c(x)\, \mu_j^c(y)\right| && (430)\\
&= \left|\sum_{j=n}^{\infty} \gamma_j(c)\, \mu_j^c(x)\, \mu_j^c(y)\right| && (431)\\
&\leq \sum_{j=n}^{\infty} \gamma_j(c) \left|\mu_j^c(x)\right| \left|\mu_j^c(y)\right|, && (432)
\end{aligned}
$$

where the first equality is by the uniform convergence of the Mathieu expansion to the exponential kernel (see the discussion following (422) for this convergence for the exponential cosine product operator and change variables as usual), the second equality is follows from $A+B-B=A$, and the inequality holds because $\gamma_j(c) > 0$ by Theorem 9.5. Further,

$$
\begin{aligned}
\left|\mu_j^c(x)\right| &= \left|\sum_{m=0}^{\infty} \widetilde{A}_m^j(c)\, \overline{T_m}(x)\right| && (433)\\
&\leq \sum_{m=0}^{\infty} \left|\widetilde{A}_m^j(c)\right| \left|\overline{T_m}(x)\right| && (434)\\
&\leq \sqrt{\frac{2}{\pi}} \sum_{m=0}^{\infty} \left|\widetilde{A}_m^j(c)\right|, && (435)
\end{aligned}
$$

where the expressions (425) and (426) give the first equality, the first inequality is by the triangle inequality, and the final inequality is a consequence of (389) and (390) (I have replaced the tighter bound of (389) with the looser one of (390) for the index-zero term to simplify the expression). As usual, $\widetilde{A}_m^j(c)$ is zero unless j and m are both odd or both even. This is evidently a loose upper bound. Putting all of this together yields the upper bound

$$
\left|\exp(cxy) - \sum_{j=0}^{n-1} \gamma_j(c)\, \mu_j^c(x)\, \mu_j^c(y)\right| \leq \frac{2}{\pi} \sum_{j=n}^{\infty} \gamma_j(c) \left(\sum_{m=0}^{\infty} \left|\widetilde{A}_m^j(c)\right|\right)^2 \quad (436)
$$

for the rank-n Mathieu approximation (as before, $\widetilde{A}^{j}_{m}(c)$ is zero unless j and m are both odd or both even).

Figure 9 and Figure 10 show the error surfaces generated by approximating the exponential product operator using the Mathieu approximation. The "fuzzy" or "jagged" appearance of the bottom left panel of Figure 9 arises because the approximation error in that panel has been reduced to the point at which it is being impacted by rounding error (note the scale of the vertical axis in that panel).

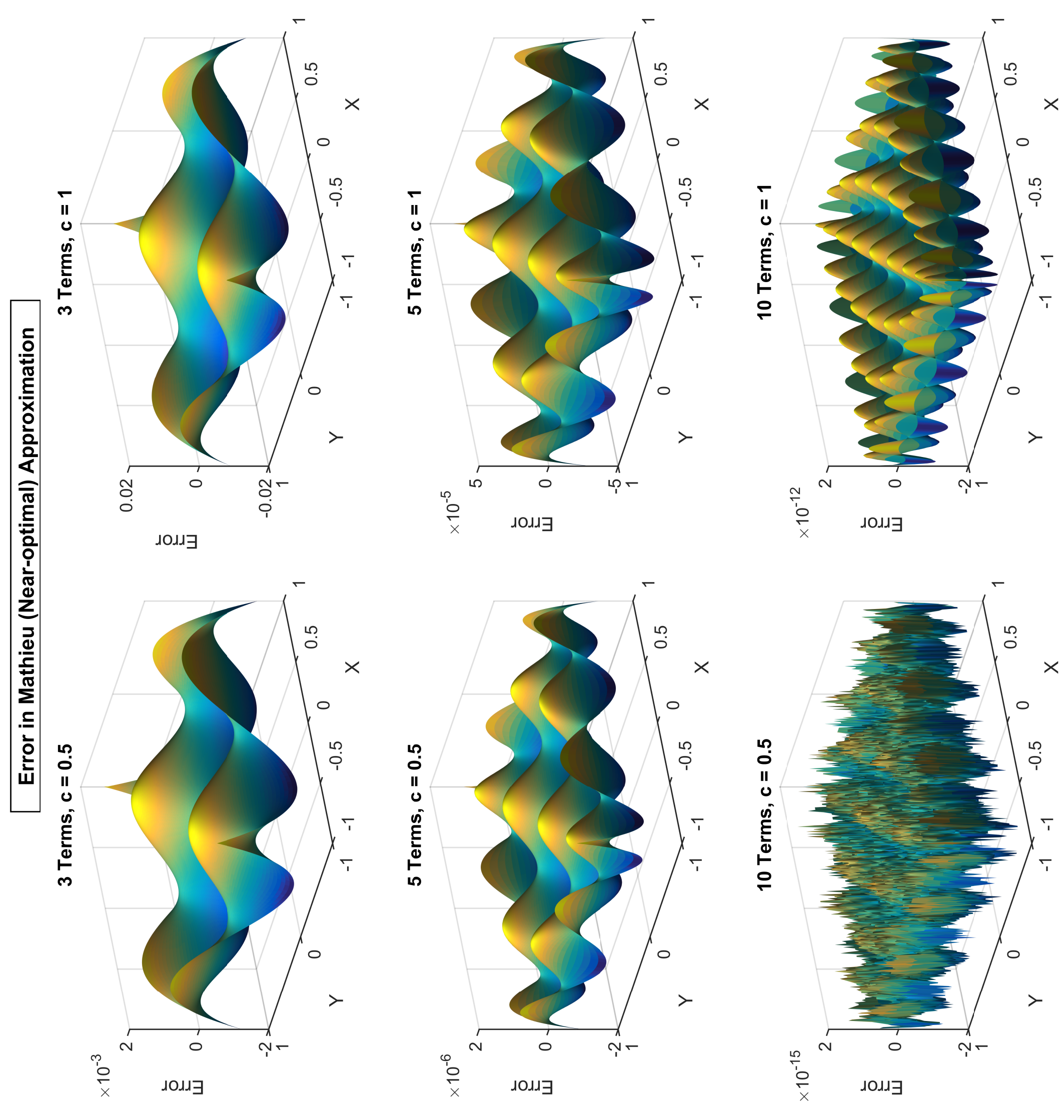

Figure 9: Error Surfaces in the Mathieu (Near-optimal) Approximation to $\exp{(cxy)}$

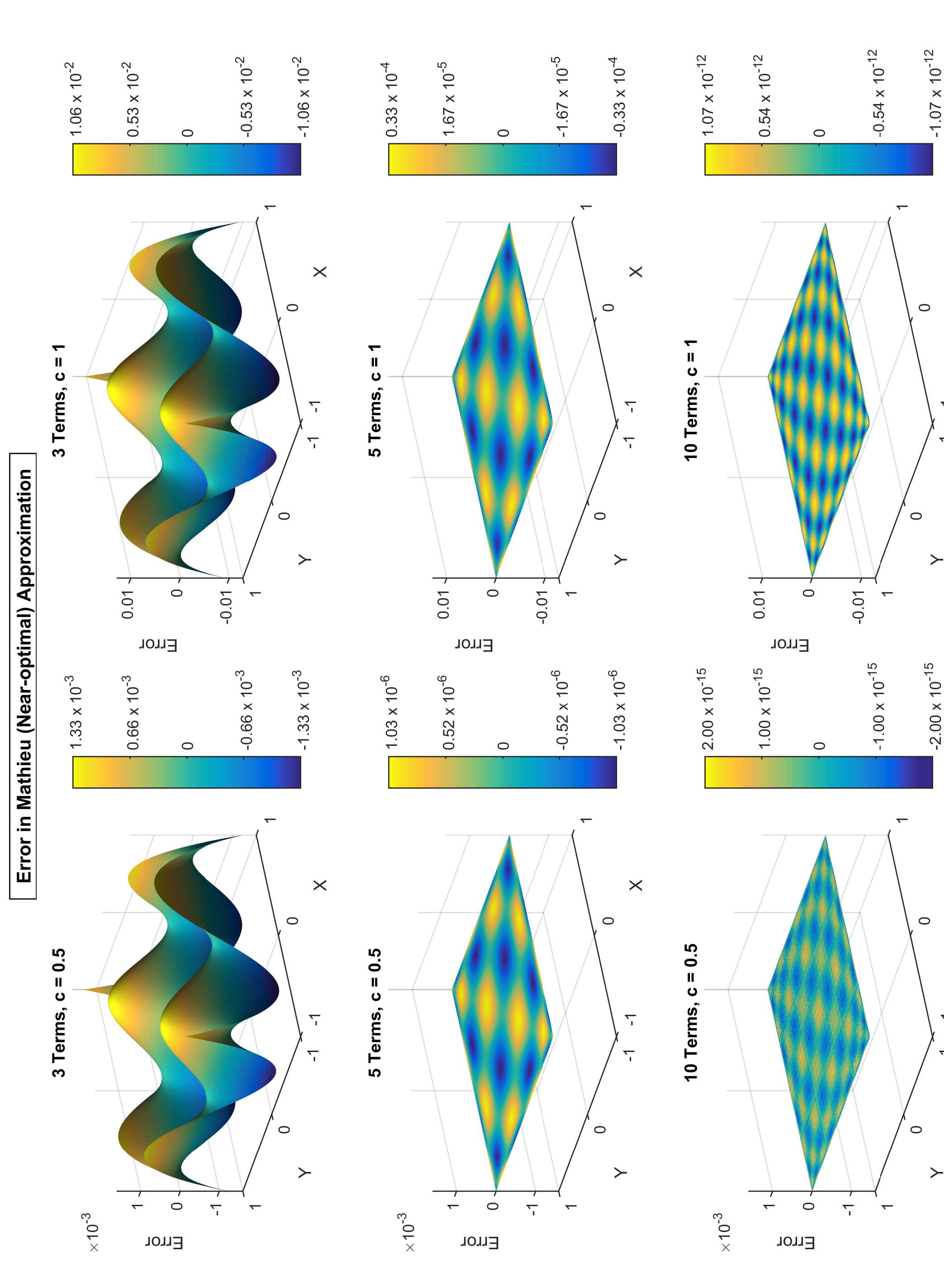

Figure 10: Error Surfaces in the Mathieu (Near-optimal) Approximation to $\exp(cxy)$

10 Numerical Comparisons of Approximating Operators

The main purpose of this section is to compare the worst-case and root-mean-square errors of 6 approximations to the exponential product operator: the numerically-optimal, Mathieu, interpolatory, oblate, univariate-Remez, and Taylor-series approximations. The numerically-optimal approximations were introduced and analyzed in Section 7, the Mathieu approximations were derived and discussed in Section 9, the interpolatory (and root-mean-square error optimal) approximations were the subject of Subsection 8.5, and the oblate approximation was introduced and characterized in Section 8. However, the univariate-Remez and the Taylor-series approximations have not yet been discussed.

The rank-n univariate-Remez approximation is constructed by computing the best (in a worst-case error sense) n-term polynomial approximation to the univariate function $\exp(cz)$ (using the classical Remez method) on the interval $[-1, 1]$ to get a polynomial in z: $\sum_{i=0}^{n-1} r_i z^i$, then simply substituting $z = xy$. The rank-n Taylor-series approximation, of course, comes from the well-known Taylor expansion of $\exp(cz) = \sum_{i=0}^{\infty} \left(c^i z^i\right)/i!$; simply truncate the series after n terms and substitute $z = xy$. The Taylor-series approximation to the exponential product operator corresponds, in a financial setting, to the use of duration and its higher-order analogs in portfolio approximation.

Figures 11 and 12 examine univariate-Remez approximation error surfaces. As usual, Figures 11 and 12 differ only in that the vertical axis varies in the first and is locked in the second. The univariate-Remez approximation is certainly not as accurate as the four approximations I introduce in this work (the numerically-optimal, Mathieu, interpolatory, and oblate approximations), but it is far superior to the Taylor-series approximation's error surfaces as shown in Figures 13 and 14. This accords well with the classical fact that the n-term univariate approximation of the univariate function $\exp(cz)$ is roughly 2^n more accurate than the n-term Taylor-series approximation of this univariate function (see Exercise 10.9 on pages 78-79 of Trefethen (2013)).

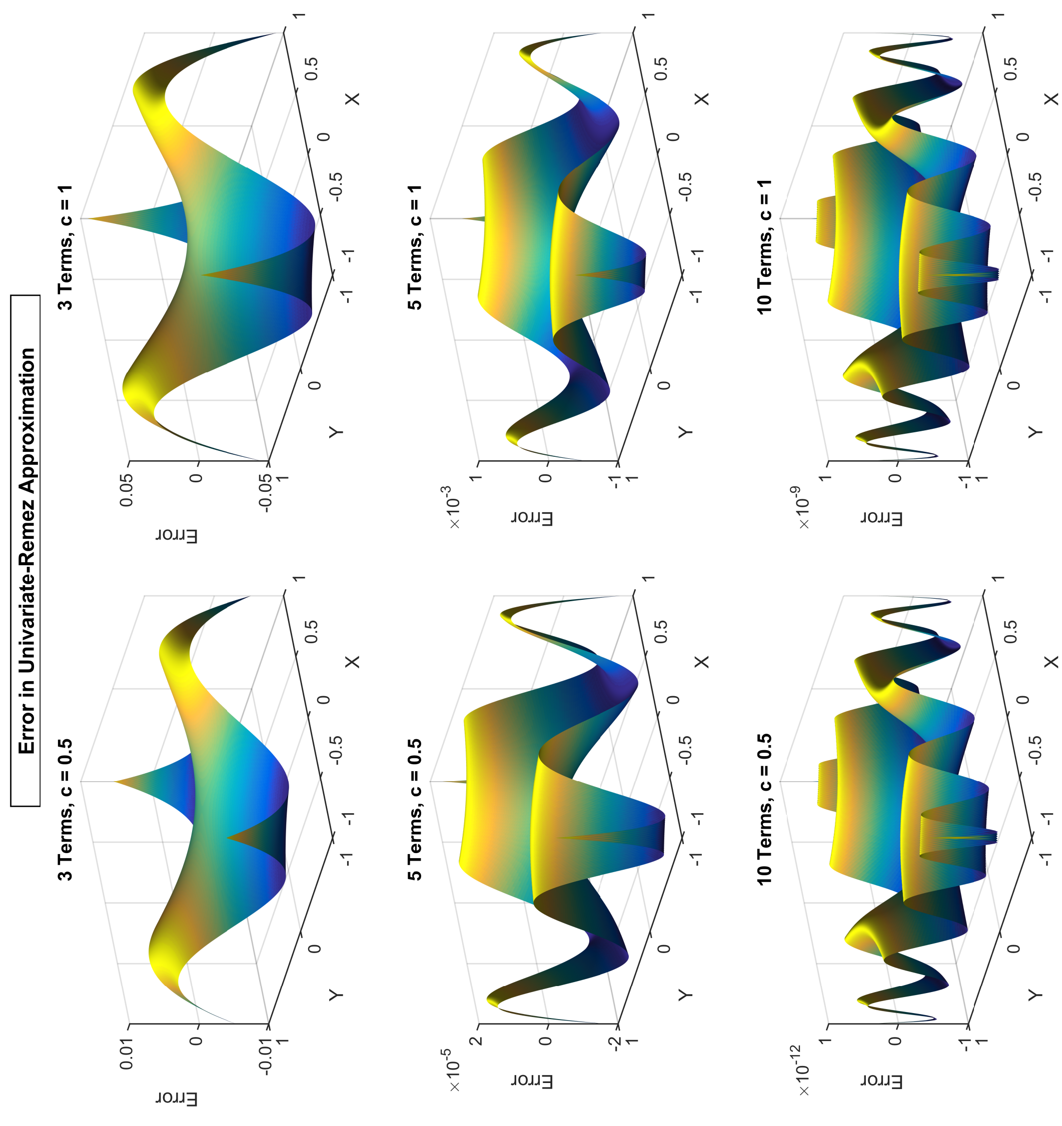

Figure 11: Error Surfaces in the Univariate-Remez Approximation to $\exp(cxy)$

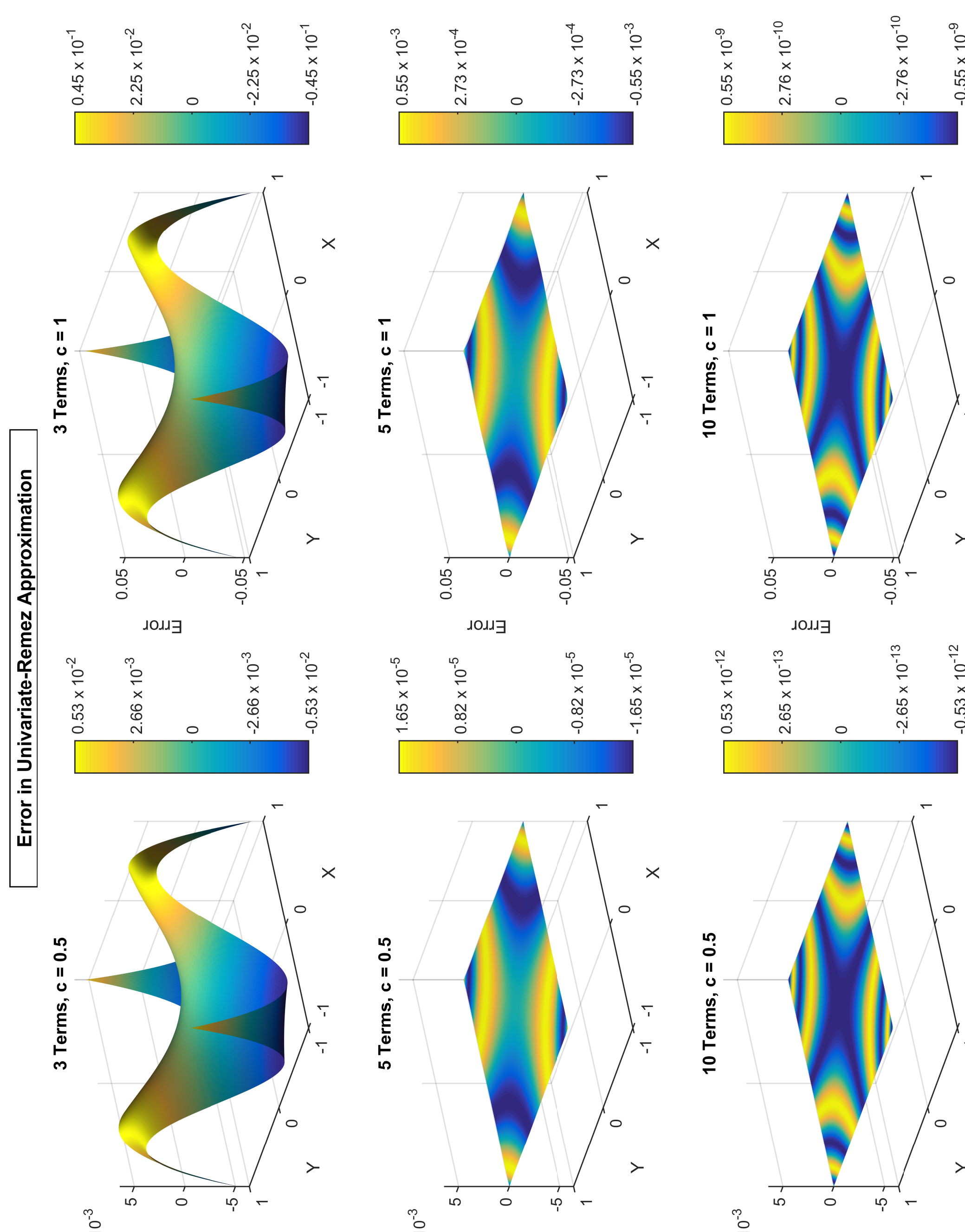

Figure 12: Error Surfaces in the Univariate-Remez Approximation to $\exp(cxy)$

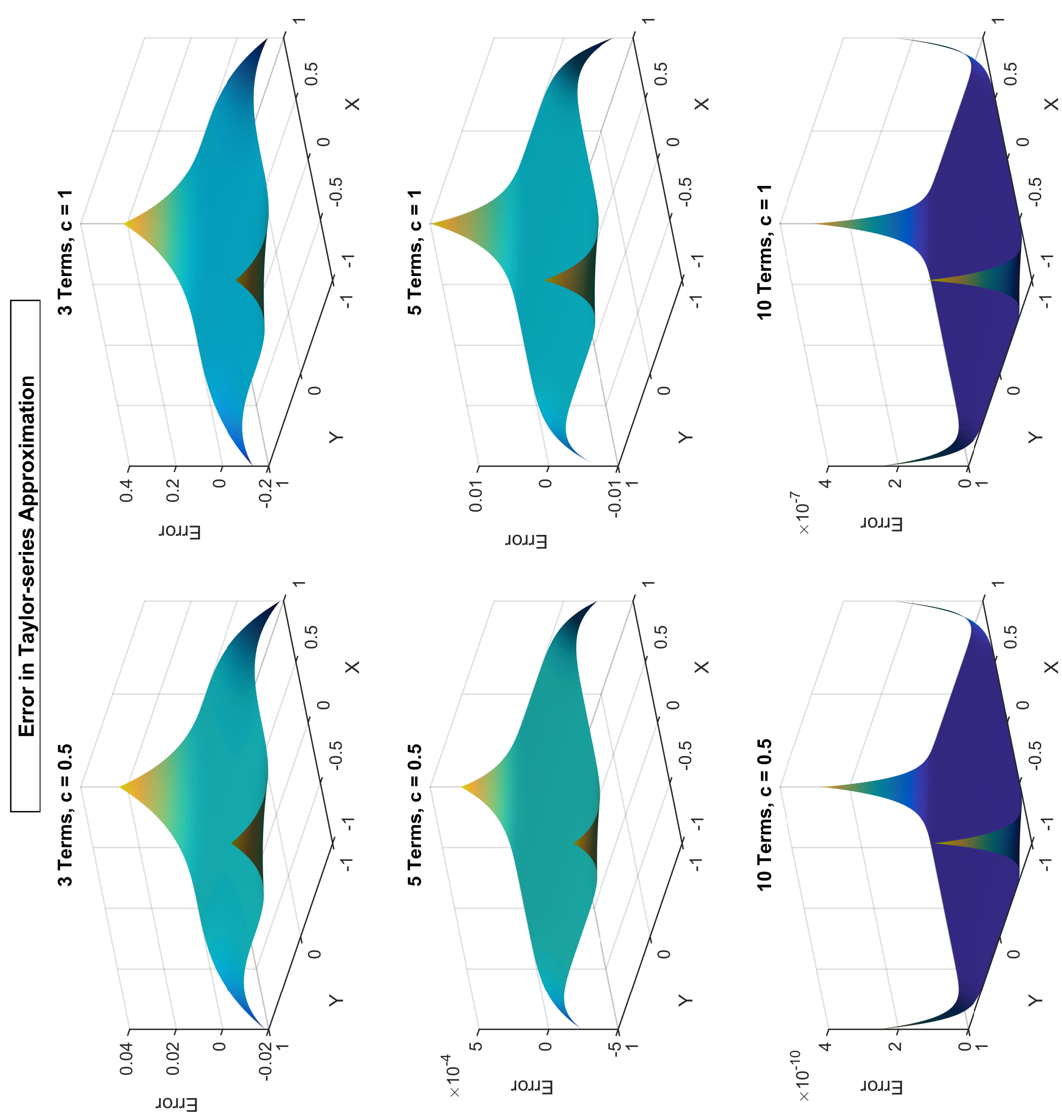

Figure 13: Error Surfaces in the Taylor-series Approximation to $\exp(cxy)$

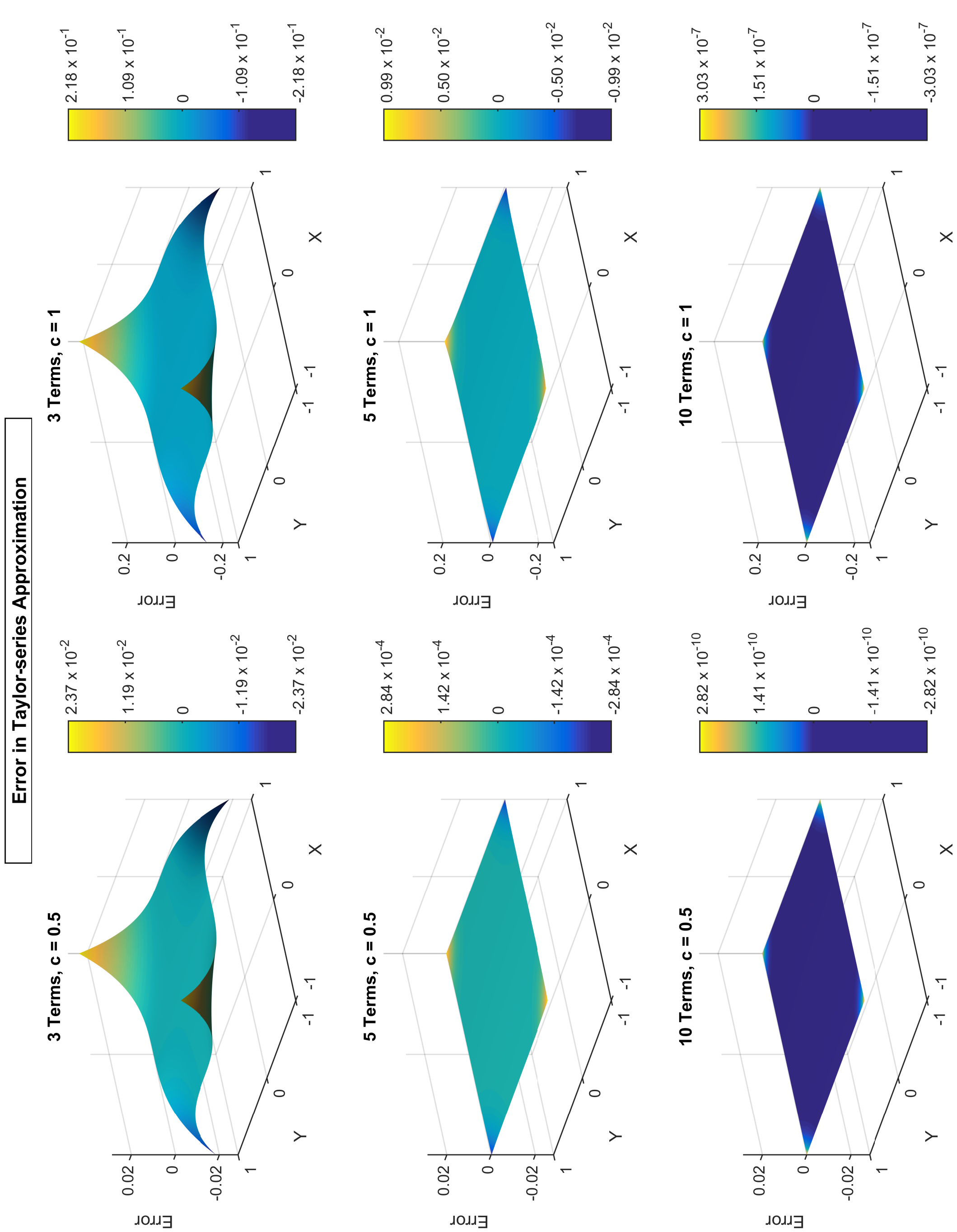

Figure 14: Error Surfaces in the Taylor-series Approximation to $\exp(cxy)$

The following table shows a comparison of worst-case approximation errors across the 6 approximations analyzed here: the four that I introduce in this work (numerically-optimal, Mathieu, interpolatory, and oblate), and the univariate-Remez and Taylor-series approximations. The table was generated with maximum iteration parameters of each method set to 100, with truncation sizes of oblate and Mathieu approximations set to 150 for the relevant methods, and with error tolerances set to 10^{-15} in all methods except in the interval-analysis upper bounding of the numerically-optimal approximation's error (through Method 7.2), where the error tolerance was relaxed to 2×10^{-14}.

The parameter c in Table 2 has its usual meaning as the parameter indexing the family of exponential product operators, and n is the rank of the approximation analyzed. Method 7.3 gives the worst-case absolute error for the numerically-optimal approximation, while the Mathieu approximation's bound comes from (436).

The interpolatory approximation always appears to achieve its largest absolute error at $x = y = 1$, and I enter the evaluations at that point in Table 2. The worst-case error for the oblate approximation occurs at $x = y = 1$ by Corollary 8.5 as long as $n \geq 1 + c\sqrt{2}$; my observation is that the worst-case error for the oblate approximation always occurs there even for n that are smaller than that relative to c, and I enter the evaluations at that point in Table 2. The worst-case error for the univariate-Remez approximation is given by the Remez method that constructs it. The worst-case error of the Taylor-series approximation always occurs at $x = y = 1$, since it is the maximum absolute value of $\sum_{i=n}^{\infty} \left(c^i x^i y^i\right)/i!$ over $x, y \in [-1, 1]$.

The table generally illustrates the excellence of the numerically-optimal and Mathieu approximations, especially in contrast to the Taylor-series approximations. For example, when the approximation problem is difficult, say $c = 3$, and rank is high ($n = 15$, the last row of the table), the worst-case error of the Taylor-series approximation is small in absolute terms but is more than 200,000,000 times larger than the worst-case error of the numerically-optimal approximation. Even for more moderate problems, the numerically-optimal and Mathieu approximations are vastly superior to the Taylor-series approximation; for $c = 1$ and $n = 6$, for instance, the numerically-optimal approximation is over 1,000 times more accurate than the Taylor-series approximation.

Two questions arise upon inspection of Table 2: first, why do the rank-one numerically-optimal and univariate-Remez errors always appear identical to within rounding error? Second, why do the interpolatory and Taylor-series errors also appear identical in the rank-one cases? Neither of these phenomena persist to higher-rank cases, but both initially appear puzzling. The answer to the first question comes from the fact that the rank-one numerically-optimal approximation appears to have the form $\cosh(cx)\cosh(cy)/\cosh(c)$, while the univariate-Remez rank-one approximation is just the constant that best approximates $\exp(cz)$ on $[-1, 1]$, which is $\cosh(c)$. Both of these approximations have worst-case error of $\sinh(c)$, which is the number appearing for both approximations in the $n = 1$ rows of Table 2. The answer to the second question is that the interpolatory approximation interpolates the exponential product kernel at the zeros of the index-n oblate to form a rank-n approximation, where $n = 1$ indicates use of the index-one (that is, second after index-zero) oblate. The oblate with index one is an odd function with just one zero at $x = 0$ (it is odd because its index is odd, implying that all of the terms in its Legendre expansion are odd by (324), and it has only one zero by Corollary 5.1). Interpolating the exponential product kernel at $x = 0$ and $y = 0$ results in the constant one, which is just the approximation used by the rank-one Taylor series, so the rank-one interpolatory approximation and the rank-one Taylor-series approximation are the same.

More interesting than the clear comparison between the four approximations I introduce in this work and Taylor-series approximation (or, as a lesser evil, univariate-Remez approximation) is the comparison between the four high-performance approximations. The numerically-optimal approximation is always within rounding error of being best, as it should be, and the Mathieu approximation is rarely far behind until one reaches, for example, $n = 2$ for large c. Both the numerically-optimal and the Mathieu approximations are notably superior to the interpolatory and oblate approximations in terms of worst-case error, with the interpolatory approximation being the laggard of the quartet. In comparing the oblate approximation to the numerically-optimal approximation, one notes that for intermediate values of n (which may be particularly relevant in financial problems) the numerically-optimal approximation is usually about three to five times more accurate than the oblate approximation in worst-case error.

The abstract made comparisons between the accuracy of the numerically-optimal approximation and the accuracy of Taylor-series approximation; these comparisons come from the $c = 1$ rows of Table 2. The compression-accuracy comment in the abstract can be linked to any of the smaller-c rows with $n = 3$ in Table 2; it holds, for example, for $c = 0.15$.

Table 2: Worst-case Approximation Errors

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.050	1	0.050020835937655	0.050333478501337	0.051271096376024	0.050576419249514	0.050020835937655	0.051271096376024
	0.050	2	0.000312467443172	0.000313944504616	0.000564855375228	0.000558916709805	0.000625173626785	0.001271096376024
	0.050	3	0.000001302032485	0.000001306417430	0.000003375122421	0.000003347003529	0.000005209350660	0.000021096376024
	0.050	4	0.000000004068936	0.000000004079799	0.000000013741812	0.000000013647576	0.000000032556967	0.000000263042691
	0.050	5	0.000000000010180	0.000000000010195	0.000000000042342	0.000000000042097	0.000000000162780	0.000000002626024
	0.050	6	0.000000000000023	0.000000000000021	0.000000000000100	0.000000000000105	0.000000000000679	0.000000000021857
	0.050	7	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000003	0.000000000000156
	0.050	8	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000001
	0.050	9	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	10	0.000000000000008	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	11	0.000000000000005	0.000000000000000	0.000000000000004	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	12	0.000000000000011	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	13	0.000000000000022	0.000000000000000	0.000000000000004	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	14	0.000000000000027	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	15	0.000000000000006	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	1	0.100166750019844	0.101419036630956	0.105170918075648	0.102389415311835	0.100166750019844	0.105170918075648
	0.100	2	0.001249478863397	0.001262698722527	0.002296947020572	0.002249334024946	0.002502778781070	0.005170918075648
	0.100	3	0.000010415039675	0.000010490271889	0.000027337256701	0.000026886429076	0.000041699228151	0.000170918075648
	0.100	4	0.000000065098748	0.000000065469186	0.000000222060908	0.000000219039902	0.000000521145908	0.000004251408981
	0.100	5	0.000000000325519	0.000000000327035	0.000000001366228	0.000000001350446	0.000000005210866	0.000000084742314
	0.100	6	0.000000000001362	0.000000000001362	0.000000000006762	0.000000000006691	0.000000000043421	0.000000001408981
	0.100	7	0.000000000000007	0.000000000000005	0.000000000000025	0.000000000000028	0.000000000000310	0.000000000020092
	0.100	8	0.000000000000001	0.000000000000000	0.000000000000006	0.000000000000000	0.000000000000002	0.000000000000251

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.100	9	0.000000000000010	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000003
	0.100	10	0.000000000000002	0.000000000000000	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	11	0.000000000000009	0.000000000000000	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	12	0.000000000000004	0.000000000000000	0.000000000000014	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	13	0.000000000000008	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	14	0.000000000000012	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	15	0.000000000000003	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000000
	0.150	1	0.150563133151613	0.153387234260513	0.161834242728283	0.155565367487309	0.150563133151613	0.161834242728283
	0.150	2	0.002809859834645	0.002859222786834	0.005253300412931	0.005092250010104	0.005639073931056	0.011834242728283
	0.150	3	0.000035143900815	0.000035550526882	0.000093404833286	0.000091117818189	0.000140872353039	0.000584242728283
	0.150	4	0.000000329528074	0.000000332512683	0.000001135331086	0.000001112348460	0.000002640280227	0.000021742728283
	0.150	5	0.000000002471646	0.000000002490078	0.000000010460656	0.000000010280579	0.000000039594060	0.000000648978283
	0.150	6	0.000000000015461	0.000000000015547	0.000000000077531	0.000000000076373	0.000000000494839	0.000000016165783
	0.150	7	0.000000000000085	0.000000000000083	0.000000000000481	0.000000000000474	0.000000000005301	0.000000000345471
	0.150	8	0.000000000000003	0.000000000000000	0.000000000000003	0.000000000000002	0.000000000000050	0.000000000006464
	0.150	9	0.000000000000002	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000000108
	0.150	10	0.000000000000003	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000002
	0.150	11	0.000000000000004	0.000000000000000	0.000000000000006	0.000000000000000	0.000000000000000	0.000000000000000
	0.150	12	0.000000000000004	0.000000000000000	0.000000000000018	0.000000000000000	0.000000000000000	0.000000000000000
	0.150	13	0.000000000000010	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000000
	0.150	14	0.000000000000004	0.000000000000000	0.000000000000023	0.000000000000000	0.000000000000000	0.000000000000000
	0.150	15	0.000000000000030	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000000
	0.200	1	0.201336002541094	0.206372736910351	0.221402758160170	0.210231914482792	0.201336002541094	0.221402758160170
	0.200	2	0.004991647362828	0.005119922130595	0.009491922691764	0.009109311229877	0.010044508694665	0.021402758160170

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.200	3	0.000083281328203	0.000084648118772	0.000224125868261	0.000216882859389	0.000334376203896	0.001402758160170
	0.200	4	0.000001041319615	0.000001054622732	0.000003623600727	0.000003526575594	0.000008353352381	0.000069424826836
	0.200	5	0.000000010414513	0.000000010523773	0.000000044444274	0.000000043430860	0.000000166991002	0.000002758160170
	0.200	6	0.000000000086805	0.000000000087567	0.000000000438712	0.000000000430008	0.000000002782316	0.000000091493503
	0.200	7	0.000000000000627	0.000000000000625	0.000000000003619	0.000000000003558	0.000000000039739	0.000000002604614
	0.200	8	0.000000000000007	0.000000000000004	0.000000000000024	0.000000000000025	0.000000000000497	0.000000000064932
	0.200	9	0.000000000000003	0.000000000000000	0.000000000000002	-0.000000000000000	0.000000000000006	0.000000000001440
	0.200	10	0.000000000000006	0.000000000000000	0.000000000000004	-0.000000000000000	0.000000000000000	0.000000000000029
	0.200	11	0.000000000000008	0.000000000000000	0.000000000000002	-0.000000000000000	0.000000000000000	0.000000000000001
	0.200	12	0.000000000000012	0.000000000000000	0.000000000000006	-0.000000000000000	0.000000000000000	0.000000000000000
	0.200	13	0.000000000000004	0.000000000000000	0.000000000000000	-0.000000000000000	0.000000000000000	0.000000000000000
	0.200	14	0.000000000000024	0.000000000000000	0.000000000000008	-0.000000000000000	0.000000000000000	0.000000000000000
	0.200	15	0.000000000000010	0.000000000000000	0.000000000000024	-0.000000000000000	0.000000000000000	0.000000000000000
	0.250	1	0.252612316808168	0.260514642951256	0.284025416687741	0.266518230668751	0.252612316808168	0.284025416687741
	0.250	2	0.007792081610420	0.008064642758701	0.015071789741181	0.014322781426025	0.015733752152689	0.034025416687741
	0.250	3	0.000162601844047	0.000166138565124	0.000443092856226	0.000425373169235	0.000654226323747	0.002775416687741
	0.250	4	0.000002541807910	0.000002584630971	0.000008933486892	0.000008636850413	0.000020421459594	0.000171250021075
	0.250	5	0.000000031778814	0.000000032217270	0.000000136745653	0.000000132873560	0.000000510173555	0.000008489604408
	0.250	6	0.000000000331079	0.000000000334937	0.000000001685338	0.000000001643787	0.000000010623443	0.000000351583574
	0.250	7	0.000000000002978	0.000000000002986	0.000000000017373	0.000000000016996	0.000000000189637	0.000000012499373
	0.250	8	0.000000000000025	0.000000000000023	0.000000000000150	0.000000000000150	0.000000000002963	0.000000000389223
	0.250	9	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000042	0.000000000010781
	0.250	10	0.000000000000008	0.000000000000000	0.000000000000002	-0.000000000000001	0.000000000000001	0.000000000000269
	0.250	11	0.000000000000010	0.000000000000000	0.000000000000005	-0.000000000000001	0.000000000000000	0.000000000000006
								Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.250	12	0.000000000000005	0.000000000000000	0.000000000000010	-0.000000000000001	0.000000000000000	-0.000000000000000
	0.250	13	0.000000000000013	0.000000000000000	0.000000000000002	-0.000000000000001	0.000000000000000	-0.000000000000000
	0.250	14	0.000000000000017	0.000000000000000	0.000000000000002	-0.000000000000001	0.000000000000000	-0.000000000000000
	0.250	15	0.000000000000013	0.000000000000000	0.000000000000004	-0.000000000000001	0.000000000000000	-0.000000000000000
	0.300	1	0.304520293447143	0.315956795920647	0.349858807576003	0.324555288932378	0.304520293447143	0.349858807576003
	0.300	2	0.011207594610567	0.011716649523186	0.022052818798955	0.020755494235796	0.022725732600818	0.049858807576003
	0.300	3	0.000280855829963	0.000288604805029	0.000774957673927	0.000738137970899	0.001132930742422	0.004858807576003
	0.300	4	0.000005269486481	0.000005381606146	0.000018705297215	0.000017965827161	0.000042415800987	0.000358807576003
	0.300	5	0.000000079064508	0.000000080438564	0.000000343045677	0.000000331465081	0.000001271172159	0.000021307576003
	0.300	6	0.000000000988465	0.000000001002995	0.000000005067722	0.000000004918625	0.000000031757039	0.000001057576003
	0.300	7	0.000000000010609	0.000000000010727	0.000000000062645	0.000000000061010	0.000000000680162	0.000000045076003
	0.300	8	0.000000000000102	0.000000000000100	0.000000000000665	0.000000000000650	0.000000000012748	0.000000001683146
	0.300	9	0.000000000000004	0.000000000000001	0.000000000000012	0.000000000000006	0.000000000000213	0.000000000055914
	0.300	10	0.000000000000004	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000004	0.000000000001673
	0.300	11	0.000000000000004	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000046
	0.300	12	0.000000000000003	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000001
	0.300	13	0.000000000000008	0.000000000000000	0.000000000000021	0.000000000000000	0.000000000000000	0.000000000000000
	0.300	14	0.000000000000019	0.000000000000000	0.000000000000006	0.000000000000000	0.000000000000000	0.000000000000000
	0.300	15	0.000000000000008	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000000
	0.350	1	0.357189729437272	0.372848080353399	0.419067548593257	0.384476110715720	0.357189729437272	0.419067548593257
	0.350	2	0.015233796266390	0.016102585091228	0.030495825544290	0.028430789683657	0.031043688854379	0.069067548593257
	0.350	3	0.000445763673790	0.000460890203339	0.001245446444143	0.001177091604824	0.001803615891796	0.007817548593257
	0.350	4	0.000009759734685	0.000010014149023	0.000034990379142	0.000033389184870	0.000078733687413	0.000671715259924
	0.350	5	0.000000170860621	0.000000174489534	0.000000747485510	0.000000718236673	0.000002751844187	0.000046454843257
								Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.350	6	0.000000002492232	0.000000002536960	0.000000012868288	0.000000012429018	0.000000080185617	0.000002686614090
	0.350	7	0.000000000031173	0.000000000031641	0.000000000185416	0.000000000179809	0.000000002003252	0.000000133467389
	0.350	8	0.000000000000348	0.000000000000345	0.000000000002296	0.000000000002234	0.000000000043799	0.000000005810054
	0.350	9	0.000000000000005	0.000000000000003	0.000000000000024	0.000000000000024	0.000000000000852	0.000000000225045
	0.350	10	0.000000000000005	0.000000000000000	0.000000000000008	0.000000000000000	0.000000000000015	0.000000000007851
	0.350	11	0.000000000000007	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000001	0.000000000000249
	0.350	12	0.000000000000021	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000007
	0.350	13	0.000000000000006	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000	-0.000000000000000
	0.350	14	0.000000000000030	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000	-0.000000000000000
	0.350	15	0.000000000000017	0.000000000000000	0.000000000000006	0.000000000000000	0.000000000000000	-0.000000000000000
	0.400	1	0.410752325802816	0.431342698972860	0.491824697641270	0.446416003795687	0.410752325802816	0.491824697641270
	0.400	2	0.019865458009982	0.021252406733913	0.040462464587529	0.037372434142185	0.040715233252443	0.091824697641270
	0.400	3	0.000665010038879	0.000692131651865	0.001881374308399	0.001764520808895	0.002700154395186	0.011824697641270
	0.400	4	0.000016644484817	0.000017164175803	0.000060268695675	0.000057141226371	0.000134618217271	0.001158030974604
	0.400	5	0.000000333055737	0.000000341508612	0.000001469143485	0.000001403867057	0.000005374948766	0.000091364307937
	0.400	6	0.000000005552392	0.000000005671326	0.000000028872745	0.000000027752545	0.000000178942019	0.000006030974604
	0.400	7	0.000000000079345	0.000000000080803	0.000000000475052	0.000000000458710	0.000000005108007	0.000000342085715
	0.400	8	0.000000000001008	0.000000000001008	0.000000000006713	0.000000000006512	0.000000000127613	0.000000017006350
	0.400	9	0.000000000000013	0.000000000000011	0.000000000000085	0.000000000000081	0.000000000002835	0.000000000752382
	0.400	10	0.000000000000003	0.000000000000000	0.000000000000004	0.000000000000001	0.000000000000057	0.000000000029983
	0.400	11	0.000000000000003	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000001	0.000000000001087
	0.400	12	0.000000000000012	0.000000000000000	0.000000000000004	0.000000000000000	0.000000000000000	0.000000000000036
	0.400	13	0.000000000000004	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000001
	0.400	14	0.000000000000031	0.000000000000000	0.000000000000008	0.000000000000000	0.000000000000000	0.000000000000000

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.400	15	0.000000000000012	0.000000000000000	0.000000000000008	0.000000000000000	0.000000000000000	0.000000000000000
	0.450	1	0.465342016934198	0.491600428602792	0.568312185490169	0.510512788346942	0.465342016934198	0.568312185490169
	0.450	2	0.025096494569984	0.027199297570874	0.052015152266045	0.047604523436608	0.051772426541026	0.118312185490169
	0.450	3	0.000946238281378	0.000991792701935	0.002710660056968	0.002523091709190	0.003857295117568	0.017062185490169
	0.450	4	0.000026651829868	0.000027631202831	0.000097466626021	0.000091820573950	0.000216181683778	0.001874685490169
	0.450	5	0.000000600044445	0.000000617927928	0.000002668784297	0.000002536238243	0.000009705833388	0.000166091740169
	0.450	6	0.000000011254574	0.000000011537295	0.000000058939661	0.000000056381183	0.000000363395900	0.000012318302669
	0.450	7	0.000000000180931	0.000000000184843	0.000000001090070	0.000000001048069	0.000000011667223	0.000000785294856
	0.450	8	0.000000000002561	0.000000000002593	0.000000000017334	0.000000000016735	0.000000000327856	0.000000043887211
	0.450	9	0.000000000000034	0.000000000000032	0.000000000000239	0.000000000000234	0.000000000008191	0.000000002183031
	0.450	10	0.000000000000003	0.000000000000000	0.000000000000003	0.000000000000003	0.000000000000185	0.000000000097822
	0.450	11	0.000000000000005	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000004	0.000000000003988
	0.450	12	0.000000000000007	0.000000000000000	0.000000000000168	0.000000000000000	0.000000000000000	0.000000000000149
	0.450	13	0.000000000000022	0.000000000000000	0.000000000000068	0.000000000000000	0.000000000000000	0.000000000000005
	0.450	14	0.000000000000011	0.000000000000000	0.000000000000016	0.000000000000000	0.000000000000001	0.000000000000000
	0.450	15	0.000000000000002	0.000000000000000	0.000000000000131	0.000000000000000	0.000000000000000	0.000000000000000
	0.500	1	0.521095305493747	0.553786851770743	0.648721270700128	0.576907011833338	0.521095305493748	0.648721270700128
	0.500	2	0.030919944365649	0.033979548505459	0.065216970645418	0.059151368765508	0.064251864848446	0.148721270700128
	0.500	3	0.001297045041690	0.001369696664260	0.003762340565679	0.003475856467029	0.005310795869779	0.023721270700128
	0.500	4	0.000040605572514	0.000042337146578	0.000149974988843	0.000140395958209	0.000330432791503	0.002887937366795
	0.500	5	0.000001015929291	0.000001050991950	0.000004555883853	0.000004306075884	0.000016474839835	0.000283770700128
	0.500	6	0.000000021173856	0.000000021789026	0.000000111672434	0.000000106315664	0.000000685118718	0.000023354033462
	0.500	7	0.000000000378212	0.000000000387691	0.000000002292905	0.000000002195218	0.000000024433992	0.000001652644573
	0.500	8	0.000000000005921	0.000000000006041	0.000000000040492	0.000000000038938	0.000000000762745	0.000000102545366

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.500	9	0.000000000000088	0.000000000000083	0.000000000000625	0.000000000000605	0.000000000021170	0.000000005664166
	0.500	10	0.000000000000005	0.000000000000001	0.000000000000007	0.000000000000008	0.000000000000529	0.000000000281877
	0.500	11	0.000000000000007	0.000000000000000	0.000000000000004	-0.000000000000000	0.000000000000013	0.000000000012763
	0.500	12	0.000000000000003	0.000000000000000	0.000000000000002	-0.000000000000000	0.000000000000000	0.000000000000530
	0.500	13	0.000000000000002	0.000000000000000	0.000000000000006	-0.000000000000000	0.000000000000000	0.000000000000021
	0.500	14	0.000000000000025	0.000000000000000	0.000000000000008	-0.000000000000000	0.000000000000000	0.000000000000001
	0.500	15	0.000000000000010	0.000000000000000	0.000000000000002	-0.000000000000000	0.000000000000001	0.000000000000000
	0.550	1	0.578151603743454	0.618073560707356	0.733253017867395	0.645742153302884	0.578151603743454	0.733253017867395
	0.550	2	0.037327949107747	0.041632406806893	0.080131551637898	0.072037365131765	0.078194778999656	0.183253017867395
	0.550	3	0.001724975036834	0.001836059614071	0.005066584990673	0.004646259533532	0.007097559167510	0.032003017867395
	0.550	4	0.000059424716042	0.000062331654186	0.000221667285159	0.000206212095931	0.000485307772593	0.004273851200729
	0.550	5	0.000001635717198	0.000001700344816	0.000007395933051	0.000006952675392	0.000026600539527	0.000461090784062
	0.550	6	0.000000037503733	0.000000038749590	0.000000199197790	0.000000188744084	0.000001216328960	0.000041687138228
	0.550	7	0.000000000736910	0.000000000758033	0.000000004495283	0.000000004285619	0.000000047702794	0.000003241804027
	0.550	8	0.000000000012688	0.000000000012988	0.000000000087258	0.000000000083601	0.000000001637660	0.000000221099197
	0.550	9	0.000000000000201	0.000000000000198	0.000000000001487	0.000000000001429	0.000000000049989	0.000000013425740
	0.550	10	0.000000000000006	0.000000000000003	0.000000000000023	0.000000000000022	0.000000000001374	0.000000000734584
	0.550	11	0.000000000000003	0.000000000000000	0.000000000000004	0.000000000000000	0.000000000000035	0.000000000036571
	0.550	12	0.000000000000006	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000001	0.000000000001670
	0.550	13	0.000000000000006	0.000000000000000	0.000000000000004	0.000000000000000	0.000000000000000	0.000000000000070
	0.550	14	0.000000000000007	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000002
	0.550	15	0.000000000000007	0.000000000000000	0.000000000000012	0.000000000000000	0.000000000000000	-0.000000000000000
	0.600	1	0.636653582148241	0.684638330338657	0.822118800390509	0.717164817719996	0.636653582148241	0.822118800390509
	0.600	2	0.044311733241487	0.050199887173993	0.096822940185984	0.086286842027701	0.093647146450463	0.222118800390509

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.600	3	0.002237516081131	0.002401523241887	0.006654708675353	0.006058143469337	0.009255771454665	0.042118800390509
	0.600	4	0.000084122897830	0.000088797981058	0.000316918157228	0.000292995651958	0.000689705505988	0.006118800390509
	0.600	5	0.000002526510175	0.000002639693705	0.000011518023760	0.000010769715021	0.000041213911449	0.000718800390509
	0.600	6	0.000000063199914	0.000000065577609	0.000000338054990	0.000000318806025	0.000002054942481	0.000070800390509
	0.600	7	0.000000001354769	0.000000001398728	0.000000008315547	0.000000007894451	0.000000087890182	0.000006000390509
	0.600	8	0.000000000025421	0.000000000026132	0.000000000175980	0.000000000167961	0.000000003290804	0.000000446104794
	0.600	9	0.000000000000436	0.000000000000434	0.000000000003263	0.000000000003133	0.000000000109561	0.000000029533366
	0.600	10	0.000000000000009	0.000000000000006	0.000000000000048	0.000000000000053	0.000000000003284	0.000000001761937
	0.600	11	0.000000000000004	0.000000000000000	0.000000000000001	0.000000000000002	0.000000000000090	0.000000000095652
	0.600	12	0.000000000000004	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000002	0.000000000004763
	0.600	13	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000219
	0.600	14	0.000000000000007	0.000000000000000	0.000000000000093	0.000000000000001	0.000000000000001	0.000000000000009
	0.600	15	0.000000000000020	0.000000000000000	0.000000000000011	0.000000000000001	0.000000000000000	0.000000000000000
	0.650	1	0.696747526126440	0.753665256958258	0.915540829013896	0.791324921062298	0.696747526126440	0.915540829013896
	0.650	2	0.051861583912566	0.059726541188043	0.115355435523101	0.101923896201725	0.110659816157270	0.265540829013896
	0.650	3	0.002842094363338	0.003077187493425	0.008559186727545	0.007735754294505	0.011825046210599	0.054290829013896
	0.650	4	0.000115807766232	0.000123059433364	0.000440622058707	0.000404861279342	0.000953526756089	0.008519995680563
	0.650	5	0.000003768689214	0.000003958555847	0.000017322720190	0.000016111157052	0.000061683579124	0.001082235263896
	0.650	6	0.000000102139109	0.000000106455685	0.000000550200012	0.000000516443421	0.000003330256080	0.000115326409730
	0.650	7	0.000000002372067	0.000000002458476	0.000000014649791	0.000000013849890	0.000000154250899	0.000010577950528
	0.650	8	0.000000000048208	0.000000000049738	0.000000000335653	0.000000000319148	0.000000006255118	0.000000851307888
	0.650	9	0.000000000000882	0.000000000000895	0.000000000006744	0.000000000006446	0.000000000225560	0.000000061018173
	0.650	10	0.000000000000017	0.000000000000014	0.000000000000121	0.000000000000116	0.000000000007323	0.000000003941694
	0.650	11	0.000000000000003	0.000000000000000	0.000000000000002	0.000000000000002	0.000000000000216	0.000000000231723

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.650	12	0.000000000000003	0.000000000000000	0.000000000000005	0.000000000000000	0.000000000000006	0.000000000012497
	0.650	13	0.000000000000003	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000623
	0.650	14	0.000000000000011	0.000000000000000	0.000000000000008	0.000000000000000	0.000000000000000	0.000000000000029
	0.650	15	0.000000000000016	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000001
	0.700	1	0.758583701839534	0.825344859594989	1.013752707470477	0.868375867043068	0.758583701839534	1.013752707470477
	0.700	2	0.059966832173596	0.070259181371596	0.135793409621531	0.118972206456224	0.129288646733140	0.313752707470477
	0.700	3	0.003546070009048	0.003874642945963	0.010813667230029	0.009703746341908	0.014846571371891	0.068752707470477
	0.700	4	0.000155680301517	0.000166586394509	0.000598212130890	0.000546317731249	0.001287717660901	0.011586040803810
	0.700	5	0.000005457091043	0.000005764096884	0.000025290218773	0.000023397238132	0.000089642227013	0.001581874137143
	0.700	6	0.000000159292492	0.000000166801797	0.000000864128842	0.000000807339477	0.000005209295797	0.000181290803810
	0.700	7	0.000000003984218	0.000000004145992	0.000000024758350	0.000000023309347	0.000000259746074	0.000017889414921
	0.700	8	0.000000000087200	0.000000000090292	0.000000000610502	0.000000000578309	0.000000011340084	0.000001549276032
	0.700	9	0.000000000001709	0.000000000001750	0.000000000013210	0.000000000012576	0.000000000440279	0.000000119513879
	0.700	10	0.000000000000034	0.000000000000031	0.000000000000254	0.000000000000243	0.000000000015390	0.000000008310156
	0.700	11	0.000000000000005	0.000000000000000	0.000000000000004	0.000000000000004	0.000000000000490	0.000000000525895
	0.700	12	0.000000000000004	0.000000000000000	0.000000000000006	0.000000000000000	0.000000000000015	0.000000000030533
	0.700	13	0.000000000000004	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000001637
	0.700	14	0.000000000000007	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000000081
	0.700	15	0.000000000000028	0.000000000000000	0.000000000000006	0.000000000000000	0.000000000000000	0.000000000000004
	0.750	1	0.822316731935830	0.899874141701028	1.117000016612675	0.948474716500688	0.822316731935830	1.117000016612675
	0.750	2	0.068615836173129	0.081846556657364	0.158201102076301	0.137454830594006	0.149594658276974	0.367000016612675
	0.750	3	0.004356732957937	0.004806002873268	0.013452984053855	0.011987186596616	0.018363261511667	0.085750016612675
	0.750	4	0.000205034081788	0.000221003955987	0.000795679278637	0.000722274037265	0.001704317613560	0.015437516612675
	0.750	5	0.000007702177038	0.000008183068601	0.000035988799630	0.000033120549715	0.000127014320877	0.002253922862675
								Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.750	6	0.000000240914003	0.000000253504993	0.000001316124292	0.000001223946887	0.000007903864904	0.000276383800175
	0.750	7	0.000000006456582	0.000000006747026	0.000000040369508	0.000000037849919	0.000000422081097	0.000029191417362
	0.750	8	0.000000000151399	0.000000000157359	0.000000001065869	0.000000001005906	0.000000019737498	0.000002706519203
	0.750	9	0.000000000003169	0.000000000003265	0.000000000024694	0.000000000023432	0.000000000820846	0.000000223560001
	0.750	10	0.000000000000066	0.000000000000061	0.000000000000508	0.000000000000485	0.000000000030735	0.000000016646734
	0.750	11	0.000000000000005	0.000000000000001	0.000000000000012	0.000000000000009	0.000000000001047	0.000000001128239
	0.750	12	0.000000000000002	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000033	0.000000000070160
	0.750	13	0.000000000000014	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000004030
	0.750	14	0.000000000000067	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000215
	0.750	15	0.000000000000008	0.000000000000000	0.000000000000008	0.000000000000000	0.000000000000001	0.000000000000011
	0.800	1	0.888105982187623	0.977456611715655	1.225540928492468	1.031782350729575	0.888105982187623	1.225540928492468
	0.800	2	0.077795967082416	0.094538976976197	0.182642390859196	0.157393984849647	0.171644198299204	0.425540928492468
	0.800	3	0.005281299186110	0.005883934956946	0.016513169251112	0.014611558513895	0.022419915233094	0.105540928492468
	0.800	4	0.000265254494588	0.000288100174315	0.001039591438844	0.000938045736014	0.002216511695546	0.020207595159135
	0.800	5	0.000010631193443	0.000011363854203	0.000050083572534	0.000045852209503	0.000176045261373	0.003140928492468
	0.800	6	0.000000354743115	0.000000375186779	0.000001951630610	0.000001806606690	0.000011678330143	0.000410261825801
	0.800	7	0.000000010141803	0.000000010644523	0.000000063802164	0.000000059574384	0.000000664931768	0.000046172936912
	0.800	8	0.000000000253661	0.000000000264682	0.000000001795734	0.000000001688419	0.000000033155726	0.000004562778182
	0.800	9	0.000000000005652	0.000000000005857	0.000000000044361	0.000000000041946	0.000000001470428	0.000000401762309
	0.800	10	0.000000000000119	0.000000000000117	0.000000000000977	0.000000000000928	0.000000000058716	0.000000031894231
	0.800	11	0.000000000000006	0.000000000000002	0.000000000000021	0.000000000000019	0.000000000002133	0.000000002304785
	0.800	12	0.000000000000006	0.000000000000000	0.000000000000004	0.000000000000001	0.000000000000071	0.000000000152825
	0.800	13	0.000000000000008	0.000000000000000	0.000000000000008	0.000000000000000	0.000000000000002	0.000000000009361
	0.800	14	0.000000000000006	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000000533

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.800	15	0.000000000000013	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000001	0.000000000000028
	0.850	1	0.956115959988632	1.058302262342846	1.339646851925991	1.118463630319488	0.956115959988632	1.339646851925991
	0.850	2	0.087493598515022	0.108387885928017	0.209180538616133	0.178810806418913	0.195509122205022	0.489646851925991
	0.850	3	0.006326907304507	0.007121692609009	0.020031465012630	0.017602765319408	0.027063378248562	0.128396851925991
	0.850	4	0.000337817895080	0.000369834977022	0.001337113033186	0.001199361155122	0.002838687834685	0.026042685259324
	0.850	5	0.000014389322556	0.000015478629555	0.000068345520262	0.000062248124623	0.000239332105264	0.004292424842657
	0.850	6	0.000000510222263	0.000000542489770	0.000002826760395	0.000002604759043	0.000016858188575	0.000594880571824
	0.850	7	0.000000015499726	0.000000016342276	0.000000098109393	0.000000091234032	0.000001019378476	0.000071061800123
	0.850	8	0.000000000411911	0.000000000431540	0.000000002932072	0.000000002746662	0.000000053987507	0.000007455234987
	0.850	9	0.000000000009744	0.000000000010142	0.000000000076914	0.000000000072487	0.000000002543242	0.000000697037442
	0.850	10	0.000000000000215	0.000000000000214	0.000000000001796	0.000000000001702	0.000000000107878	0.000000058763229
	0.850	11	0.000000000000007	0.000000000000004	0.000000000000036	0.000000000000036	0.000000000004162	0.000000004509921
	0.850	12	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000148	0.000000000317620
	0.850	13	0.000000000000021	0.000000000000000	0.000000000000001	-0.000000000000001	0.000000000000006	0.000000000020665
	0.850	14	0.000000000000025	0.000000000000000	0.000000000000001	-0.000000000000001	0.000000000000001	0.000000000001249
	0.850	15	0.000000000000017	0.000000000000000	0.000000000000002	-0.000000000000001	0.000000000000001	0.000000000000070
	0.900	1	1.026516725708176	1.142627510037619	1.459603111156950	1.208687551428468	1.026516725708175	1.459603111156950
	0.900	2	0.097694100180486	0.123445382124493	0.237877914480863	0.201725100040748	0.221266988834628	0.559603111156950
	0.900	3	0.007500615564952	0.008533145878606	0.024046335185800	0.020987132807073	0.032342712632326	0.154603111156950
	0.900	4	0.000424290711437	0.000468349741963	0.001696024595977	0.001512367728751	0.003586498876385	0.033103111156950
	0.900	5	0.000019140824093	0.000020725649064	0.000091660842058	0.000083055347208	0.000319855994245	0.005765611156950
	0.900	6	0.000000718728562	0.000000768395225	0.000004009938017	0.000003678247260	0.000023839461302	0.000844861156950
	0.900	7	0.000000023120094	0.000000024492435	0.000000147244854	0.000000136369705	0.000001525569362	0.000106748656950
	0.900	8	0.000000000650595	0.000000000684441	0.000000004656472	0.000000004345988	0.000000085516850	0.000011848478378

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.900	9	0.000000000016291	0.000000000017024	0.000000000129277	0.000000000121422	0.000000004264250	0.000001172208289
	0.900	10	0.000000000000377	0.000000000000381	0.000000000003200	0.000000000003020	0.000000000191473	0.000000104581280
	0.900	11	0.000000000000012	0.000000000000008	0.000000000000062	0.000000000000068	0.000000000007820	0.000000008494849
	0.900	12	0.000000000000007	0.000000000000000	0.000000000000007	0.000000000000002	0.000000000000293	0.000000000633232
	0.900	13	0.000000000000005	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000011	0.000000000043611
	0.900	14	0.000000000000009	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000002791
	0.900	15	0.000000000000019	0.000000000000000	0.000000000000006	0.000000000000001	0.000000000000000	0.000000000000167
	0.950	1	1.099484317930673	1.230655098228145	1.585709659315846	1.302627401844906	1.099484317930672	1.585709659315846
	0.950	2	0.108381836485069	0.139763691789207	0.268795691744500	0.226155069997696	0.249001271599627	0.635709659315846
	0.950	3	0.008809399306024	0.010132811925343	0.028597476358897	0.024791411662574	0.038309372751932	0.184459659315846
	0.950	4	0.000526328498289	0.000585977575491	0.002124742566698	0.001883638342020	0.004476929772500	0.041563825982513
	0.950	5	0.000025070166151	0.000027331665090	0.000121040599789	0.000109118522867	0.000421016437473	0.007626065565846
	0.950	6	0.000000993819862	0.000001068571293	0.000005583684308	0.000005098716412	0.000033098964425	0.001177891086679
	0.950	7	0.000000033748313	0.000000035927270	0.000000216253962	0.000000199472331	0.000002234634818	0.000156930127478
	0.950	8	0.000000001002482	0.000000001059191	0.000000007214281	0.000000006708612	0.000000132171149	0.000018371140158
	0.950	9	0.000000000026481	0.000000000027797	0.000000000211302	0.000000000197807	0.000000006954640	0.000001917260413
	0.950	10	0.000000000000644	0.000000000000657	0.000000000005514	0.000000000005192	0.000000000329543	0.000000180461996
	0.950	11	0.000000000000017	0.000000000000014	0.000000000000126	0.000000000000123	0.000000000014203	0.000000015466146
	0.950	12	0.000000000000004	0.000000000000000	0.000000000000004	0.000000000000004	0.000000000000562	0.000000001216504
	0.950	13	0.000000000000004	0.000000000000000	0.000000000000006	0.000000000000001	0.000000000000021	0.000000000088408
	0.950	14	0.000000000000014	0.000000000000000	0.000000000000000	0.000000000000001	0.000000000000002	0.000000000005970
	0.950	15	0.000000000000014	0.000000000000000	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000376
	1.000	1	1.175201193643802	1.322613970185320	1.718281828459046	1.400460919701897	1.175201193643802	1.718281828459046
	1.000	2	0.119540170749652	0.157394597546776	0.301993522159932	0.252117039387740	0.278801585795503	0.718281828459046

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	1.000	3	0.010260148871543	0.011935885052451	0.033725828530472	0.029042779353980	0.045017388402819	0.218281828459046
	1.000	4	0.000645674938763	0.000725254316229	0.002632339235752	0.002320177690792	0.005528370108688	0.051615161792379
	1.000	5	0.000032383145042	0.000035554490037	0.000157630669318	0.000141386432461	0.000546667600514	0.009948495125712
	1.000	6	0.000001351494913	0.000001461753834	0.000007646546764	0.000006951107872	0.000045205511927	0.001615161792379
	1.000	7	0.000000048314336	0.000000051695570	0.000000311492087	0.000000286164421	0.000003210877104	0.000226272903490
	1.000	8	0.000000001510777	0.000000001603381	0.000000010931647	0.000000010128384	0.000000199825277	0.000027860205077
	1.000	9	0.000000000042000	0.000000000044274	0.000000000336877	0.000000000314301	0.000000011064290	0.000003058617775
	1.000	10	0.000000000001062	0.000000000001101	0.000000000009251	0.000000000008682	0.000000000551725	0.000000302885853
	1.000	11	0.000000000000028	0.000000000000025	0.000000000000228	0.000000000000216	0.000000000025024	0.000000027312661
	1.000	12	0.000000000000006	0.000000000000001	0.000000000000001	0.000000000000005	0.000000000001042	0.000000002260553
	1.000	13	0.000000000000021	0.000000000000000	0.000000000000004	0.000000000000000	0.000000000000040	0.000000000172877
	1.000	14	0.000000000000007	0.000000000000000	0.000000000000008	0.000000000000000	0.000000000000003	0.000000000012286
	1.000	15	0.000000000000047	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000815
	1.250	1	1.601919080300827	1.849805513367745	2.490342957461841	1.954587688894546	1.601919080300826	2.490342957461841
	1.250	2	0.181740352204688	0.266941205124663	0.504119361688439	0.405309368090282	0.462392346356437	1.240342957461841
	1.250	3	0.019878444413997	0.024585253440759	0.069546963393259	0.057987204516070	0.091773893559269	0.459092957461841
	1.250	4	0.001569098482380	0.001850865491227	0.006719483661928	0.005758096830509	0.013953014345963	0.133572124128508
	1.250	5	0.000098537025823	0.000112474316453	0.000499597017131	0.000437137687868	0.001714053754161	0.031846863711841
	1.250	6	0.000005145160092	0.000005748787507	0.000030143930172	0.000026804483755	0.000176419846872	0.006415548607675
	1.250	7	0.000000230041340	0.000000253119769	0.000001529103990	0.000001377184853	0.000015614691254	0.001117357960974
	1.250	8	0.000000008994732	0.000000009783355	0.000000066876667	0.000000060857535	0.000001211816909	0.000171252488348
	1.250	9	0.000000000312543	0.000000000336864	0.000000002569802	0.000000002358483	0.000000083715875	0.000023423508251
	1.250	10	0.000000000009789	0.000000000010455	0.000000000088042	0.000000000081374	0.000000005210289	0.000002891705460
	1.250	11	0.000000000000286	0.000000000000295	0.000000000002725	0.000000000002529	0.000000000295018	0.000000325230110

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	1.250	12	0.000000000000014	0.000000000000007	0.000000000000095	0.000000000000071	0.000000000015322	0.000000033585184
	1.250	13	0.000000000000005	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000736	0.000000003205505
	1.250	14	0.000000000000014	0.000000000000000	0.000000000000005	-0.000000000000001	0.000000000000034	0.000000000284382
	1.250	15	0.000000000000046	0.000000000000000	0.000000000000043	-0.000000000000001	0.000000000000002	0.000000000023567
	1.500	1	2.129279455094819	2.512952253524951	3.481689070338065	2.635612237449296	2.129279455094817	3.481689070338065
	1.500	2	0.252673041202173	0.415731231956001	0.770664665503095	0.597653873737525	0.715087005755588	1.981689070338065
	1.500	3	0.034029449823933	0.045039085758498	0.126671647992346	0.102432978227066	0.167009307123871	0.856689070338065
	1.500	4	0.003235486930458	0.004037543836967	0.014552004746624	0.012137781685187	0.030123196643835	0.294189070338065
	1.500	5	0.000244317600436	0.000291588086460	0.001290057180890	0.001102078328543	0.004407791358664	0.083251570338065
	1.500	6	0.000015325480668	0.000017775028215	0.000092961667155	0.000080912139975	0.000541609320860	0.019970320338064
	1.500	7	0.000000822791596	0.000000934908357	0.000005638011681	0.000004980685747	0.000057307418440	0.004150007838065
	1.500	8	0.000000038622334	0.000000043211479	0.000000295035598	0.000000263801121	0.000005321596184	0.000759940873778
	1.500	9	0.000000001610813	0.000000001780526	0.000000013572059	0.000000012256791	0.000000440156118	0.000124303317975
	1.500	10	0.000000000060464	0.000000000066168	0.000000000556841	0.000000000507103	0.000000032812934	0.000018363725341
	1.500	11	0.000000000002078	0.000000000002238	0.000000000020610	0.000000000018906	0.000000002226156	0.000002472786446
	1.500	12	0.000000000000075	0.000000000000069	0.000000000000693	0.000000000000640	0.000000000138558	0.000000305840233
	1.500	13	0.000000000000009	0.000000000000002	0.000000000000024	0.000000000000019	0.000000000007967	0.000000034971957
	1.500	14	0.000000000000011	0.000000000000000	0.000000000000007	-0.000000000000001	0.000000000000426	0.000000003717925
	1.500	15	0.000000000000017	0.000000000000000	0.000000000000008	-0.000000000000002	0.000000000000024	0.000000000369278
	1.750	1	2.790414366277642	3.347293689003606	4.754602676005730	3.472190830938276	2.790414366277643	4.754602676005730
	1.750	2	0.329330364456195	0.606299005254604	1.105075612795814	0.827404955328104	1.056814527986845	3.004602676005730
	1.750	3	0.053480652887238	0.076130929243002	0.211675505637431	0.166265276382500	0.281698359989925	1.473352676005730
	1.750	4	0.005954724293130	0.007916526009659	0.028123866165615	0.022857915914260	0.058506042253308	0.580123509339063
	1.750	5	0.000525843263798	0.000659736100237	0.002891216526656	0.002413472239277	0.009902467806397	0.189335748922396
								Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	1.750	6	0.000038532256344	0.000046606599850	0.000241960784786	0.000206266109164	0.001411067835655	0.052560032776563
	1.750	7	0.000002415385436	0.000002845579752	0.000017059866162	0.000014789723505	0.000173420115011	0.012667115567361
	1.750	8	0.000000132343251	0.000000152851522	0.000001038581837	0.000000912803431	0.000018724407505	0.002693886265061
	1.750	9	0.000000006441745	0.000000007325403	0.000000055610034	0.000000049433695	0.000001802024924	0.000512242355183
	1.750	10	0.000000000282103	0.000000000316805	0.000000002656744	0.000000002384347	0.000000156390109	0.000088033817151
	1.750	11	0.000000000011248	0.000000000012480	0.000000000114554	0.000000000103650	0.000000012356464	0.000013797322995
	1.750	12	0.000000000000428	0.000000000000451	0.000000000004500	0.000000000004099	0.000000000895917	0.000001986971652
	1.750	13	0.000000000000020	0.000000000000015	0.000000000000161	0.000000000000147	0.000000000060014	0.000000264628748
	1.750	14	0.000000000000010	0.000000000000001	0.000000000000009	0.000000000000004	0.000000000003739	0.000000032774896
	1.750	15	0.000000000000019	0.000000000000000	0.000000000000003	-0.000000000000001	0.000000000000220	0.000000003793164
	2.000	1	3.626860407847020	4.393474014007037	6.389056098930650	4.498648064977125	3.626860407847019	6.389056098930650
	2.000	2	0.408476154383678	0.838097985545888	1.507160175356232	1.090149315829956	1.514077880708201	4.389056098930650
	2.000	3	0.078963756663127	0.121334900369822	0.331992541700052	0.253654681140236	0.450331365239263	2.389056098930650
	2.000	4	0.010081645628616	0.014373387388424	0.049991694674816	0.039631295193399	0.105342554184843	1.055722765597317
	2.000	5	0.001020237963597	0.001352343294251	0.005840237675182	0.004767417557445	0.020181001242475	0.389056098930650
	2.000	6	0.000085566871628	0.000108407536107	0.000556161029920	0.000464638465251	0.003264232827582	0.122389432263984
	2.000	7	0.000006135515873	0.000007523284525	0.000044662563183	0.000038014972659	0.000456181010543	0.033500543375095
	2.000	8	0.000000384426112	0.000000459911076	0.000003098924200	0.000002678232208	0.000056073670604	0.008103717978270
	2.000	9	0.000000021393393	0.000000025105644	0.000000189207552	0.000000165609354	0.000006148608103	0.001754511629064
	2.000	10	0.000000001070987	0.000000001237472	0.000000010311239	0.000000009122274	0.000000608335792	0.000343576884796
	2.000	11	0.000000000048745	0.000000000055586	0.000000000507328	0.000000000452942	0.000000054819354	0.000061389935943
	2.000	12	0.000000000002050	0.000000000002293	0.000000000022743	0.000000000020469	0.000000004534773	0.000010083217969
	2.000	13	0.000000000000086	0.000000000000087	0.000000000000944	0.000000000000848	0.000000000346656	0.000001532098307
	2.000	14	0.000000000000020	0.000000000000004	0.000000000000025	0.000000000000032	0.000000000024629	0.000000216541435

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	2.000	15	0.000000000000013	0.000000000000000	0.000000000000021	0.000000000000001	0.000000000001638	0.000000028604740
	2.250	1	4.691168305898329	5.701743169409006	8.487735836358526	5.758465121579061	4.691168305898331	8.487735836358526
	2.250	2	0.486953715930452	1.109159784312400	1.972955865407757	1.379724058565655	2.121766957813410	6.237735836358526
	2.250	3	0.111203348536543	0.184873695433645	0.495977117511378	0.369092892969045	0.692093839084089	3.706485836358526
	2.250	4	0.016010273929705	0.024628299406608	0.083335575664451	0.064497786731129	0.179259914334334	1.808048336358526
	2.250	5	0.001828377899175	0.002572240081383	0.010895075442848	0.008703354521021	0.038224676634059	0.740177242608526
	2.250	6	0.000172801608591	0.000230262908980	0.001162414451199	0.000952256089317	0.006903122020574	0.259635250421026
	2.250	7	0.000013953701785	0.000017873014558	0.000104673603756	0.000087512936089	0.001079220352885	0.079432003350712
	2.250	8	0.000000984222158	0.000001223672658	0.000008149073247	0.000006927989341	0.000148594502109	0.021509531078111
	2.250	9	0.000000061646855	0.000000074877432	0.000000558522586	0.000000481500773	0.000018267644554	0.005218835751442
	2.250	10	0.000000003472999	0.000000004139856	0.000000034179978	0.000000029815892	0.000002027651402	0.001146161919776
	2.250	11	0.000000000177817	0.000000000208688	0.000000001888951	0.000000001664475	0.000000205086876	0.000229810307651
	2.250	12	0.000000000008365	0.000000000009665	0.000000000095168	0.000000000084576	0.000000019049058	0.000042374750626
	2.250	13	0.000000000000379	0.000000000000414	0.000000000004421	0.000000000003940	0.000000001635510	0.000007230583684
	2.250	14	0.000000000000046	0.000000000000016	0.000000000000187	0.000000000000167	0.000000000130537	0.000001147939406
	2.250	15	0.000000000000018	0.000000000000000	0.000000000000018	0.000000000000004	0.000000000009738	0.000000170371575
	2.500	1	6.050204481039788	7.338084127957771	11.182493960703473	7.310053891861184	6.050204481039788	11.182493960703473
	2.500	2	0.561993735644561	1.419005182225533	2.496059319419496	1.690775976486195	2.925539606868060	8.682493960703473
	2.500	3	0.150962080321487	0.271860472439069	0.713014360511886	0.517468383568000	1.032394029484328	5.557493960703473
	2.500	4	0.024166676406513	0.040334741080585	0.132012828197940	0.099829669376373	0.292093438796153	2.953327294036807
	2.500	5	0.003077298895590	0.004613838154302	0.019085047399345	0.014929544086725	0.068408850870140	1.325723127370141
	2.500	6	0.000323753197076	0.000455496508319	0.002253661686973	0.001811315164977	0.013613555772578	0.511921044036807
	2.500	7	0.000029081016056	0.000039044145421	0.000224783809953	0.000184677396325	0.002350200238999	0.172836842647918
	2.500	8	0.000002280843644	0.000002956130917	0.000019394859807	0.000016225682256	0.000357830688946	0.051735342151886

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	2.500	9	0.000000158815915	0.000000200220793	0.000001473869114	0.000001251856739	0.000048692858776	0.013891123246877
	2.500	10	0.000000009944975	0.000000012261328	0.000000100040355	0.000000086068633	0.000005986786892	0.003378840217707
	2.500	11	0.000000000565869	0.000000000684960	0.000000006133730	0.000000005335462	0.000000671100142	0.000750769460414
	2.500	12	0.000000000029533	0.000000000035168	0.000000000342869	0.000000000301094	0.000000069111232	0.000153480651939
	2.500	13	0.000000000001442	0.000000000001670	0.000000000017598	0.000000000015593	0.000000006580995	0.000029045483506
	2.500	14	0.000000000000092	0.000000000000075	0.000000000000831	0.000000000000746	0.000000000582695	0.000005115643424
	2.500	15	0.000000000000016	0.000000000000004	0.000000000000043	0.000000000000034	0.000000000048209	0.000000842457695
	2.750	1	7.789352011490729	9.390817475964827	14.642631884188171	9.234402248244397	7.789352011490732	14.642631884188171
	2.750	2	0.631447816059517	1.771141303537963	3.070369088191079	2.022359026162130	3.984928029428717	11.892631884188171
	2.750	3	0.199104696568392	0.388491300854255	0.993708767353025	0.704196229789828	1.504848334467708	8.111381884188171
	2.750	4	0.034999814456554	0.063694338995223	0.200598109838078	0.148324102513659	0.459983344872063	4.645236050854837
	2.750	5	0.004922221915644	0.007893524956689	0.031765772538291	0.024349128733245	0.117041331656493	2.262260790438171
	2.750	6	0.000570797431077	0.000850934798816	0.004111112077821	0.003243393297316	0.025392289384700	0.951624397209004
	2.750	7	0.000056470282207	0.000079724292146	0.000449713286676	0.000363228478029	0.004789683377801	0.350916050312303
	2.750	8	0.000004875950751	0.000006607093280	0.000042578099672	0.000035064706449	0.000798007291306	0.114923485460029
	2.750	9	0.000000373679261	0.000000490289853	0.000003551926424	0.000002973205499	0.000118955231508	0.033801041292058
	2.750	10	0.000000025750093	0.000000032918514	0.000000264742170	0.000000224694787	0.000016033821911	0.009013627796289
	2.750	11	0.000000001612168	0.000000002017231	0.000000017828929	0.000000015312771	0.000001971546528	0.002197089084953
	2.750	12	0.000000000092519	0.000000000113658	0.000000001094826	0.000000000950083	0.000000222811453	0.000492954407120
	2.750	13	0.000000000004929	0.000000000005926	0.000000000061801	0.000000000054099	0.000000023291649	0.000102423543449
	2.750	14	0.000000000000262	0.000000000000288	0.000000000003238	0.000000000002848	0.000000002264581	0.000019811245366
	2.750	15	0.000000000000051	0.000000000000014	0.000000000000131	0.000000000000140	0.000000000205773	0.000003583829670
	3.000	1	10.017874927409904	11.977023867945320	19.085536923187668	11.643106298156651	10.017874927409903	19.085536923187668
	3.000	2	0.693900867554400	2.174213283927795	3.693292940109746	2.380835393150626	5.377375153869116	16.085536923187668

Cont'd. on next page

Table 2: Worst-case Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	3.000	3	0.256683305023225	0.542295718531694	1.350182640858005	0.935408101745530	2.153866444596346	11.585536923187668
	3.000	4	0.048969470769698	0.097578414795777	0.294399527997673	0.212974736774715	0.704832365419478	7.085536923187668
	3.000	5	0.007548160818657	0.012986954925445	0.050677893228578	0.038086259454214	0.193092436683810	3.710536923187668
	3.000	6	0.000956997870372	0.001516666924065	0.007130653267016	0.005524782993167	0.045264675483397	1.685536923187669
	3.000	7	0.000103428857183	0.000153999347049	0.000848535167183	0.000674036748919	0.009247251092084	0.673036923187670
	3.000	8	0.000009751403656	0.000013852442596	0.000087435954480	0.000070906586032	0.001671306515991	0.239108351759100
	3.000	9	0.000000815773615	0.000001116758654	0.000007941461384	0.000006553127580	0.000270563321237	0.076385137473384
	3.000	10	0.000000061352825	0.000000081516102	0.000000644645880	0.000000539877298	0.000039638653924	0.022144066044813
	3.000	11	0.000000004191777	0.000000005433616	0.000000047291945	0.000000040113260	0.000005300978110	0.005871744616243
	3.000	12	0.000000000262435	0.000000000333159	0.000000003164260	0.000000002713733	0.000000651871009	0.001433838772087
	3.000	13	0.000000000015209	0.000000000018908	0.000000000194628	0.000000000168487	0.000000074175876	0.000324362311048
	3.000	14	0.000000000000856	0.000000000000998	0.000000000011116	0.000000000009660	0.000000007852694	0.000068329281579
	3.000	15	0.000000000000067	0.000000000000050	0.000000000000544	0.000000000000508	0.000000000777128	0.000013465060977

I now provide another table, this time detailing the root-mean-square error performance of the 6 approximations analyzed in Table 2. The root-mean-square error of an approximation f_{approx} is in an operator-norm sense, as with the definitions of n-widths:

$$\sup_{h:\int_{-1}^{1} h^2(x)dx=1} \left\{ \sqrt{\int_{-1}^{1} \left(\int_{-1}^{1} \left(\exp\left(cxy\right) - f_{\text{approx}}\left(x,y\right)\right) h\left(x\right) dx \right)^2 dy} \right\}. \quad (437)$$

Each cell of Table 3 shows the maximal absolute value among the eigenvalues of the approximation error for the approximation in that cell's column and the c and n (rank) of that cell's row. This maximal absolute value among the eigenvalues is equal to (437), as consideration of the usual variational principles for eigenvalues shows (in fact, (437) is the square root of the largest eigenvalue of the iterated error kernel, which is the same thing). To compute the largest-magnitude eigenvalues for each approximation error, I use a Nystrom method based on 100 Gauss-Legendre nodes and weights to sample the error surfaces at discrete points and scale the rows and columns of the resulting matrix. I then employ a standard eigenvalue routine on the scaled matrix; see Delves & Mohamed (1984), equation (2a) on page 154 in Section 1.1 of Chapter 6, which applies because the Gauss-Legendre weights are all positive. To obtain the Gauss-Legendre nodes and weights, I used the Chebfun package as described by Trefethen (2013) under the BSD license (my use of the Chebfun package should be not construed as indicating any endorsement of my work by its authors). The problem of computing Gauss-Legendre nodes and weights is classical, but the Chebfun implementation is particularly pleasant and user-friendly.

The construction of the approximations themselves is as in Table 2.

As guaranteed by the theory of Subsection 8.5, the interpolatory and oblate approximations always have the same root-mean-square error (which is also the n-width of the Hilbert-space problem, and is thus optimal for root-mean-square error) to within rounding error. More puzzling is the observation that every root-mean-square error is identical to within rounding error whenever $n = 1$ for all of the c values in the table below 0.75; why would every rank-one approximation have the same root-mean-square error, to within rounding error, for small and moderate values of c? The answer is that the oblate of index one, which is the second oblate because indexing of the oblates begins with zero, is an odd function. All of the rank-one approximations use a product of two *even* univariate functions, one a function of x and the other of y. For the interpolatory, Taylor-series, and univariate-Remez approximations this is actually a constant function, while the numerically-optimal, Mathieu, and oblate approximations use a product of two non-constant even functions, one a function of x and the other of y. Because of this, the oblate of index one, as a function of x, is orthogonal to all of the (even) functions of x used in the rank-one approximations I analyze. Of all the oblates other than the index-zero oblate, the oblate of index one is associated with the largest eigenvalue of the exponential product operator (by definition). The impact of the exponential product operator on the index-zero oblate is captured well by all of the rank-one approximations here for small and moderate c (and always captured perfectly, by definition, by the rank-one oblate approximation), leaving the index-one oblate as the one maximizing L^2 approximation error in these cases. Because the index-one oblate is orthogonal to the functions of x used by each rank-one approximation, the root-mean-square approximation error for rank $n = 1$ is, for all of the approximations here when c is small or moderate and for all but the univariate-Remez approximation even for larger c, the eigenvalue of the exponential product operator associated with the index-one oblate: the index-one eigenvalue (the second eigenvalue, since indexing begins with zero).

As c grows, the rank-one univariate-Remez approximation suffers relative to all of the other rank-one approximations because it approximates using the constant $\cosh(c)$. The growth of this constant as c grows means that the rank-one univariate-Remez approximation has large root-mean-square error for the index-zero oblate, which makes its error in this sense diverge from those of the other approximations considered here.

As in Table 2, the four approximations I introduce in this work (numerically-optimal, Mathieu, interpolatory, and oblate) all have much more attractive root-mean-square error properties than the univariate-Remez and Taylor-series approximations (as long as the approximation rank exceeds one). With that said, the good local properties of the Taylor-series approximation make it less awful in a root-mean-square error comparison than in a worst-case error comparison; for $c = 1$ and $n = 6$, for example, the interpolatory and oblate approximations (which always have equal root-mean-square approximation errors) are only 211 times more accurate than the Taylor-series approximation in root-mean-square error (in worst-case error, the numerically-optimal approximation is more than 1,000 times more accurate than the Taylor-series approximation for $c = 1$ and $n = 6$).

More interesting is the comparison between the numerically-optimal and Mathieu approximations, which are respectively numerically optimal and nearly optimal for worst-case error, and the interpolatory and oblate approximations, which are optimal for root-mean-square error. The root-mean-square error reduction in moving from numerically-optimal or Mathieu approximation to interpolatory or oblate approximation is generally only about 25 percent of the numerically-optimal or Mathieu root-mean-square error. This suggests that the numerically-optimal and Mathieu approximations possess reasonable root-mean-square properties (they are not too far from the best possible error control in root-mean-square). It also contrasts with the gap between the numerically-optimal approximation and the oblate approximation in terms of worst-case error (where the numerically-optimal approximation is often more than 4 times more accurate than the oblate approximation).

Taken together, these results show that all four approximations that I introduce here (numerically-optimal, Mathieu, interpolatory, and oblate) are far better in worst-case error and in root-mean-square error than either Taylor-series approximation or univariate-Remez approximation. Among the four approximations introduced here, the numerically-optimal approximation seems to have the most favorable error properties overall; it is certainly the best (as it should be) for worst-case error approximation, and it does not lag the root-mean-square error optimal approximations (interpolatory and oblate) by too much in root-mean-square error.

Table 3: Root-mean-square Approximation Errors

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.050	1	0.033338333795091	0.033338333795091	0.033338333795090	0.033338333795090	0.033338333795091	0.033338333795090
	0.050	2	0.000291646606443	0.000291690166093	0.000222225368305	0.000222225368305	0.001021996461861	0.000500053149440
	0.050	3	0.000001264836740	0.000001264878670	0.000000952385066	0.000000952385066	0.000006093379679	0.000005952831075
	0.050	4	0.000000004004344	0.000000004004408	0.000000003023437	0.000000003023437	0.000000040351348	0.000000057873599
	0.050	5	0.000000000010070	0.000000000010070	0.000000000007635	0.000000000007635	0.000000000172933	0.000000000473505
	0.050	6	0.000000000000021	0.000000000000021	0.000000000000016	0.000000000000016	0.000000000000692	0.000000000003339
	0.050	7	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000021
	0.050	8	0.000000000000001	0.000000000000001	0.000000000000004	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	9	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	10	0.000000000000005	0.000000000000001	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	11	0.000000000000005	0.000000000000001	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	12	0.000000000000010	0.000000000000001	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	13	0.000000000000007	0.000000000000001	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	14	0.000000000000012	0.000000000000001	0.000000000000014	0.000000000000000	0.000000000000000	0.000000000000000
	0.050	15	0.000000000000003	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	1	0.066706681445055	0.066706681445054	0.066706681445054	0.066706681445054	0.066706681445054	0.066706681445054
	0.100	2	0.001166345519139	0.001167042030347	0.000888939066550	0.000888939066550	0.004089856584380	0.002000850543770
	0.100	3	0.000010117633123	0.000010118973875	0.000007619179070	0.000007619179070	0.000048765855649	0.000047633453434
	0.100	4	0.000000064065730	0.000000064069817	0.000000048375272	0.000000048375272	0.000000645822063	0.000000926132565
	0.100	5	0.000000000322217	0.000000000322229	0.000000000244319	0.000000000244319	0.000000005535345	0.000000015154098
	0.100	6	0.000000000001347	0.000000000001347	0.000000000001025	0.000000000001025	0.000000000044314	0.000000000213704
	0.100	7	0.000000000000005	0.000000000000005	0.000000000000004	0.000000000000004	0.000000000000288	0.000000000002646
	0.100	8	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000029

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.100	9	0.000000000000003	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	10	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	11	0.000000000000007	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	12	0.000000000000001	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	13	0.000000000000007	0.000000000000000	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	14	0.000000000000008	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000000
	0.100	15	0.000000000000000	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000000
	0.150	1	0.100135112250700	0.100135112250700	0.100135112250701	0.100135112250700	0.100135112250701	0.100135112250701
	0.150	2	0.002623372643227	0.002626894973206	0.002000252676686	0.002000252676686	0.009209196649863	0.004504307166872
	0.150	3	0.000034141048156	0.000034151216877	0.000025715281677	0.000025715281677	0.000164690663652	0.000160823703816
	0.150	4	0.000000324300884	0.000000324347430	0.000000244902127	0.000000244902127	0.000003271166477	0.000004689854152
	0.150	5	0.000000002446678	0.000000002446890	0.000000001855304	0.000000001855304	0.000000042052936	0.000000115100960
	0.150	6	0.000000000015340	0.000000000015341	0.000000000011677	0.000000000011677	0.000000000504964	0.000000002434629
	0.150	7	0.000000000000082	0.000000000000082	0.000000000000063	0.000000000000063	0.000000000004919	0.000000000045212
	0.150	8	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000044	0.000000000000748
	0.150	9	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000011
	0.150	10	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000000	0.000000000000000
	0.150	11	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000000
	0.150	12	0.000000000000002	0.000000000000000	0.000000000000066	0.000000000000001	0.000000000000000	0.000000000000000
	0.150	13	0.000000000000013	0.000000000000000	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000000
	0.150	14	0.000000000000003	0.000000000000000	0.000000000000034	0.000000000000001	0.000000000000000	0.000000000000000
	0.150	15	0.000000000000003	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000000	0.000000000000000
	0.200	1	0.133653806518054	0.133653806518054	0.133653806518054	0.133653806518054	0.133653806518054	0.133653806518054
	0.200	2	0.004661516599592	0.004672632150751	0.003556348171385	0.003556348171385	0.016389390706900	0.008013618480362

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.200	3	0.000080907155230	0.000080949931813	0.000060956564727	0.000060956564727	0.000390729544100	0.000381413609264
	0.200	4	0.000001024809959	0.000001025071354	0.000000774021845	0.000000774021845	0.000010345993483	0.000014828045124
	0.200	5	0.000000010309377	0.000000010310964	0.000000007818291	0.000000007818291	0.000000177322341	0.000000485179238
	0.200	6	0.000000000086187	0.000000000086195	0.000000000065608	0.000000000065608	0.000000002838770	0.000000013682553
	0.200	7	0.000000000000617	0.000000000000617	0.000000000000471	0.000000000000471	0.000000000036872	0.000000000338771
	0.200	8	0.000000000000004	0.000000000000004	0.000000000000003	0.000000000000003	0.000000000000440	0.000000000007472
	0.200	9	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000005	0.000000000000149
	0.200	10	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000003
	0.200	11	0.000000000000004	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000000	0.000000000000000
	0.200	12	0.000000000000004	0.000000000000000	0.000000000000003	0.000000000000001	0.000000000000000	0.000000000000000
	0.200	13	0.000000000000002	0.000000000000000	0.000000000000007	0.000000000000001	0.000000000000000	0.000000000000000
	0.200	14	0.000000000000025	0.000000000000000	0.000000000000004	0.000000000000001	0.000000000000000	0.000000000000000
	0.200	15	0.000000000000013	0.000000000000000	0.000000000000047	0.000000000000001	0.000000000000000	0.000000000000000
	0.250	1	0.167293111344445	0.167293111344445	0.167293111344445	0.167293111344445	0.167293111344445	0.167293111344445
	0.250	2	0.007279071994212	0.007306156478188	0.005557471905150	0.005557471905150	0.025643590959858	0.012533266164379
	0.250	3	0.000157972188539	0.000158102441188	0.000119060335224	0.000119060335224	0.000764027559328	0.000745455730945
	0.250	4	0.000002501535070	0.000002502531570	0.000001889733543	0.000001889733543	0.000025282307816	0.000036219466496
	0.250	5	0.000000031458167	0.000000031465733	0.000000023859745	0.000000023859745	0.000000541583062	0.000001481217690
	0.250	6	0.000000000328750	0.000000000328803	0.000000000250275	0.000000000250275	0.000000010836645	0.000000052210556
	0.250	7	0.000000000002941	0.000000000002941	0.000000000002246	0.000000000002246	0.000000000175929	0.000000001615778
	0.250	8	0.000000000000023	0.000000000000023	0.000000000000018	0.000000000000018	0.000000000002622	0.000000000044547
	0.250	9	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000034	0.000000000001107
	0.250	10	0.000000000000003	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000025
	0.250	11	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000001

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.250	12	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000000	0.000000000000000
	0.250	13	0.000000000000005	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000000	0.000000000000000
	0.250	14	0.000000000000003	0.000000000000000	0.000000000000014	0.000000000000001	0.000000000000000	0.000000000000000
	0.250	15	0.000000000000003	0.000000000000000	0.000000000000004	0.000000000000001	0.000000000000000	0.000000000000000
	0.300	1	0.201083596745094	0.201083596745094	0.201083596745094	0.201083596745094	0.201083596745094	0.201083596745094
	0.300	2	0.010473830196131	0.010529856417764	0.008003926176325	0.008003926176325	0.036988756359175	0.018069026170854
	0.300	3	0.000272871336807	0.000273194563578	0.000205745770859	0.000205745770859	0.001322108590195	0.001289219806948
	0.300	4	0.000005186062519	0.000005189035709	0.000003918631386	0.000003918631386	0.000052485095420	0.000075150803642
	0.300	5	0.000000078267180	0.000000078294276	0.000000059371322	0.000000059371322	0.000001348965808	0.000003687472474
	0.300	6	0.000000000981547	0.000000000981777	0.000000000747321	0.000000000747321	0.000000032385793	0.000000155957417
	0.300	7	0.000000000010537	0.000000000010539	0.000000000008048	0.000000000008048	0.000000000630868	0.000000005791349
	0.300	8	0.000000000000099	0.000000000000099	0.000000000000076	0.000000000000076	0.000000000011285	0.000000000191592
	0.300	9	0.000000000000002	0.000000000000001	0.000000000000002	0.000000000000001	0.000000000000176	0.000000000005713
	0.300	10	0.000000000000003	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000003	0.000000000000155
	0.300	11	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000004
	0.300	12	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000000
	0.300	13	0.000000000000003	0.000000000000001	0.000000000000005	0.000000000000001	0.000000000000000	0.000000000000000
	0.300	14	0.000000000000023	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000000
	0.300	15	0.000000000000006	0.000000000000001	0.000000000000003	0.000000000000001	0.000000000000000	0.000000000000000
	0.350	1	0.235056112234223	0.235056112234223	0.235056112234223	0.235056112234223	0.235056112234223	0.235056112234223
	0.350	2	0.014243067823059	0.014346561006362	0.010896058335641	0.010896058335641	0.050445689437764	0.024627979105459
	0.350	3	0.000433113600457	0.000433809969859	0.000326734319361	0.000326734319361	0.002102971740869	0.002049247794663
	0.350	4	0.000009605368954	0.000009612859274	0.000007259921774	0.000007259921774	0.000097365980588	0.000139327197261
	0.350	5	0.000000169138613	0.000000169218287	0.000000128326447	0.000000128326447	0.000002919057104	0.000007974500062

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.350	6	0.000000002474815	0.000000002475603	0.000000001884477	0.000000001884477	0.000000081747618	0.000000393437883
	0.350	7	0.000000000030997	0.000000000031004	0.000000000023677	0.000000000023677	0.000000001857627	0.000000017043545
	0.350	8	0.000000000000339	0.000000000000340	0.000000000000260	0.000000000000260	0.000000000038763	0.000000000657776
	0.350	9	0.000000000000003	0.000000000000003	0.000000000000005	0.000000000000003	0.000000000000704	0.000000000022883
	0.350	10	0.000000000000005	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000012	0.000000000000725
	0.350	11	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000021
	0.350	12	0.000000000000008	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000001
	0.350	13	0.000000000000005	0.000000000000001	0.000000000000012	0.000000000000000	0.000000000000000	0.000000000000000
	0.350	14	0.000000000000034	0.000000000000001	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000000
	0.350	15	0.000000000000018	0.000000000000001	0.000000000000004	0.000000000000000	0.000000000000000	0.000000000000000
	0.400	1	0.269241843975187	0.269241843975188	0.269241843975188	0.269241843975187	0.269241843975187	0.269241843975188
	0.400	2	0.018583533187674	0.018759483432795	0.014234247411758	0.014234247411758	0.066039081452392	0.032218522959105
	0.400	3	0.000646176728652	0.000647529366230	0.000487750053309	0.000487750053309	0.003145188882680	0.003062409414161
	0.400	4	0.000016381511806	0.000016398183828	0.000012385443558	0.000012385443558	0.000166359990518	0.000237885155208
	0.400	5	0.000000329701356	0.000000329904123	0.000000250197052	0.000000250197052	0.000005698860984	0.000015557543061
	0.400	6	0.000000005513632	0.000000005515923	0.000000004199004	0.000000004199004	0.000000182361954	0.000000877094587
	0.400	7	0.000000000078926	0.000000000078950	0.000000000060293	0.000000000060293	0.000000004735386	0.000000043418994
	0.400	8	0.000000000000988	0.000000000000988	0.000000000000757	0.000000000000757	0.000000000112918	0.000000001914956
	0.400	9	0.000000000000011	0.000000000000011	0.000000000000008	0.000000000000008	0.000000000002343	0.000000000076133
	0.400	10	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000045	0.000000000002755
	0.400	11	0.000000000000003	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000091
	0.400	12	0.000000000000008	0.000000000000001	0.000000000000007	0.000000000000000	0.000000000000000	0.000000000000003
	0.400	13	0.000000000000002	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000000
	0.400	14	0.000000000000027	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000000

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.400	15	0.000000000000007	0.000000000000001	0.000000000000018	0.000000000000000	0.000000000000000	0.000000000000000
	0.450	1	0.303672372614690	0.303672372614689	0.303672372614689	0.303672372614689	0.303672372614689	0.303672372614689
	0.450	2	0.023491431009542	0.023772152379420	0.018018888220347	0.018018888220347	0.083797565902771	0.040850388214603
	0.450	3	0.000919502253264	0.000921929474668	0.000694519972543	0.000694519972543	0.004488003570543	0.004365958295008
	0.450	4	0.000026231264377	0.000026265022613	0.000019839690803	0.000019839690803	0.000266945471109	0.000381407869698
	0.450	5	0.000000594005791	0.000000594467930	0.000000450870520	0.000000450870520	0.000010285239718	0.000028055533712
	0.450	6	0.000000011176057	0.000000011181933	0.000000008512667	0.000000008512667	0.000000370189778	0.000001779136734
	0.450	7	0.000000000179986	0.000000000180055	0.000000000137510	0.000000000137510	0.000000010812762	0.000000099071136
	0.450	8	0.000000000002534	0.000000000002535	0.000000000001941	0.000000000001941	0.000000000290031	0.000000004915223
	0.450	9	0.000000000000032	0.000000000000032	0.000000000000024	0.000000000000024	0.000000000006770	0.000000000219827
	0.450	10	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000146	0.000000000008948
	0.450	11	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000000	0.000000000000003	0.000000000000334
	0.450	12	0.000000000000002	0.000000000000000	0.000000000000091	0.000000000000000	0.000000000000000	0.000000000000012
	0.450	13	0.000000000000026	0.000000000000000	0.000000000000019	0.000000000000000	0.000000000000000	0.000000000000000
	0.450	14	0.000000000000003	0.000000000000000	0.000000000000021	0.000000000000000	0.000000000000000	0.000000000000000
	0.450	15	0.000000000000001	0.000000000000000	0.000000000000381	0.000000000000000	0.000000000000000	0.000000000000000
	0.500	1	0.338379731916834	0.338379731916834	0.338379731916834	0.338379731916834	0.338379731916833	0.338379731916834
	0.500	2	0.028962405746725	0.029388330286389	0.022250372921149	0.022250372921149	0.103753780515456	0.050534655358850
	0.500	3	0.001260490636575	0.001264581784352	0.000952774246848	0.000952774246848	0.006171431568525	0.005997588830197
	0.500	4	0.000039965696062	0.000040029132925	0.000030239837545	0.000030239837545	0.000407667206000	0.000581942968968
	0.500	5	0.000001005714163	0.000001006679653	0.000000763566832	0.000000763566832	0.000017447916928	0.000047550821065
	0.500	6	0.000000021026243	0.000000021039886	0.000000016018256	0.000000016018256	0.000000697608865	0.000003349895305
	0.500	7	0.000000000376261	0.000000000376438	0.000000000287501	0.000000000287501	0.000000022636696	0.000000207239648
	0.500	8	0.000000000005887	0.000000000005889	0.000000000004510	0.000000000004510	0.000000000674559	0.000000011423174

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.500	9	0.000000000000082	0.000000000000082	0.000000000000063	0.000000000000063	0.000000000017495	0.000000000567612
	0.500	10	0.000000000000003	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000419	0.000000000025670
	0.500	11	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000009	0.000000000001065
	0.500	12	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000041
	0.500	13	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000001
	0.500	14	0.000000000000004	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000000
	0.500	15	0.000000000000004	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000000	0.000000000000000
	0.550	1	0.373396468313601	0.373396468313601	0.373396468313601	0.373396468313602	0.373396468313601	0.373396468313601
	0.550	2	0.034991523946813	0.035611917690345	0.026929069987379	0.026929069987379	0.125944437794549	0.061283774835724
	0.550	3	0.001676496551028	0.001683051050941	0.001268246392252	0.001268246392252	0.008236363224813	0.007995493842303
	0.550	4	0.000058489698166	0.000058601914566	0.000044275763896	0.000044275763896	0.000598162799182	0.000853022737145
	0.550	5	0.000001619285880	0.000001621165783	0.000001229755950	0.000001229755950	0.000028153128107	0.000076649729214
	0.550	6	0.000000037242457	0.000000037271687	0.000000028377646	0.000000028377646	0.000001237879228	0.000005938729306
	0.550	7	0.000000000733127	0.000000000733544	0.000000000560263	0.000000000560263	0.000000044176991	0.000000404081031
	0.550	8	0.000000000012618	0.000000000012624	0.000000000009667	0.000000000009667	0.000000001447872	0.000000024498038
	0.550	9	0.000000000000193	0.000000000000193	0.000000000000148	0.000000000000148	0.000000000041302	0.000000001338923
	0.550	10	0.000000000000003	0.000000000000003	0.000000000000002	0.000000000000002	0.000000000001087	0.000000000066605
	0.550	11	0.000000000000001	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000026	0.000000000003040
	0.550	12	0.000000000000002	0.000000000000000	0.000000000000003	0.000000000000000	0.000000000000001	0.000000000000128
	0.550	13	0.000000000000004	0.000000000000000	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000005
	0.550	14	0.000000000000004	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000000	0.000000000000000
	0.550	15	0.000000000000002	0.000000000000000	0.000000000000015	0.000000000000000	0.000000000000000	0.000000000000000
	0.600	1	0.408755701489223	0.408755701489224	0.408755701489224	0.408755701489224	0.408755701489224	0.408755701489224
	0.600	2	0.041573256058676	0.042446842844434	0.032055300554336	0.032055300554336	0.150410404256717	0.073111589479796

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.600	3	0.002174824308664	0.002184893524383	0.001646673375040	0.001646673375040	0.010724667940519	0.010398423168623
	0.600	4	0.000082801455920	0.000082990293558	0.000062710082764	0.000062710082764	0.000849192390063	0.001209686840728
	0.600	5	0.000002501157115	0.000002504610619	0.000001900075416	0.000001900075416	0.000043588005411	0.000118544055644
	0.600	6	0.000000062759957	0.000000062818555	0.000000047831408	0.000000047831408	0.000002090197550	0.000010017522382
	0.600	7	0.000000001347831	0.000000001348742	0.000000001030185	0.000000001030185	0.000000081359954	0.000000743462002
	0.600	8	0.000000000025308	0.000000000025321	0.000000000019391	0.000000000019391	0.000000002908451	0.000000049165695
	0.600	9	0.000000000000422	0.000000000000422	0.000000000000324	0.000000000000324	0.000000000090497	0.000000002931155
	0.600	10	0.000000000000006	0.000000000000006	0.000000000000005	0.000000000000005	0.000000000002599	0.000000000159055
	0.600	11	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000067	0.000000000007919
	0.600	12	0.000000000000002	0.000000000000000	0.000000000000001	0.000000000000000	0.000000000000002	0.000000000000364
	0.600	13	0.000000000000002	0.000000000000000	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000016
	0.600	14	0.000000000000004	0.000000000000000	0.000000000000134	0.000000000000000	0.000000000000000	0.000000000000001
	0.600	15	0.000000000000008	0.000000000000000	0.000000000000006	0.000000000000000	0.000000000000000	0.000000000000000
	0.650	1	0.444491186117059	0.444491186117059	0.444491186117059	0.444491186117058	0.444491186117058	0.444491186117059
	0.650	2	0.048701458183004	0.049896935955805	0.037629312119547	0.037629312119547	0.177196788484194	0.086033359475897
	0.650	3	0.002762723459732	0.002777654890180	0.002093795638430	0.002093795638430	0.013679300981532	0.013245743273959
	0.650	4	0.000113991867228	0.000114296590298	0.000086378166387	0.000086378166387	0.001172671778466	0.001668507604928
	0.650	5	0.000003730914096	0.000003736955881	0.000002835248180	0.000002835248180	0.000065185785762	0.000177073620524
	0.650	6	0.000000101428649	0.000000101539743	0.000000077319956	0.000000077319956	0.000003385362306	0.000016206814000
	0.650	7	0.000000002359938	0.000000002361810	0.000000001804071	0.000000001804071	0.000000142725105	0.000001302829220
	0.650	8	0.000000000048007	0.000000000048035	0.000000000036788	0.000000000036788	0.000000005526319	0.000000093325608
	0.650	9	0.000000000000868	0.000000000000868	0.000000000000666	0.000000000000666	0.000000000186255	0.000000006026984
	0.650	10	0.000000000000014	0.000000000000014	0.000000000000011	0.000000000000011	0.000000000005794	0.000000000354275
	0.650	11	0.000000000000001	0.000000000000001	0.000000000000003	0.000000000000000	0.000000000000162	0.000000000019107

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.650	12	0.000000000000001	0.000000000000001	0.000000000000002	0.000000000000000	0.000000000000004	0.000000000000952
	0.650	13	0.000000000000001	0.000000000000001	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000044
	0.650	14	0.000000000000006	0.000000000000001	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000002
	0.650	15	0.000000000000004	0.000000000000001	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000000
	0.700	1	0.480637374868671	0.480637374868671	0.480637374868671	0.480637374868671	0.480637374868671	0.480637374868671
	0.700	2	0.056369354270822	0.057965787593154	0.043651249575747	0.043651249575747	0.206353038146335	0.100065789894657
	0.700	3	0.003447384578931	0.003468867913702	0.002615357047826	0.002615357047826	0.017144412887943	0.016577498001760
	0.700	4	0.000153243908796	0.000153718371268	0.000116188171773	0.000116188171773	0.001581709045949	0.002247617887951
	0.700	5	0.000005402459452	0.000005412598439	0.000004107000635	0.000004107000635	0.000094651933687	0.000256789979898
	0.700	6	0.000000158185442	0.000000158386289	0.000000120616253	0.000000120616253	0.000005292073795	0.000025304614399
	0.700	7	0.000000003963860	0.000000003967506	0.000000003030752	0.000000003030752	0.000000240219558	0.000002190270961
	0.700	8	0.000000000086840	0.000000000086901	0.000000000066556	0.000000000066556	0.000000010014856	0.000000168942725
	0.700	9	0.000000000001690	0.000000000001691	0.000000000001298	0.000000000001298	0.000000000363443	0.000000011748453
	0.700	10	0.000000000000030	0.000000000000030	0.000000000000023	0.000000000000023	0.000000000012175	0.000000000743658
	0.700	11	0.000000000000002	0.000000000000001	0.000000000000003	0.000000000000000	0.000000000000367	0.000000000043190
	0.700	12	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000010	0.000000000002317
	0.700	13	0.000000000000002	0.000000000000001	0.000000000000002	0.000000000000000	0.000000000000000	0.000000000000115
	0.700	14	0.000000000000002	0.000000000000001	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000005
	0.700	15	0.000000000000023	0.000000000000001	0.000000000000005	0.000000000000000	0.000000000000000	0.000000000000000
	0.750	1	0.517229482816229	0.517229482816229	0.517229482816228	0.517229482816228	0.531425069142336	0.517229482816228
	0.750	2	0.064569519303954	0.066656591132543	0.050121123570983	0.050121123570983	0.237933046157911	0.115227060859266
	0.750	3	0.004235935259644	0.004266049782196	0.003217104751581	0.003217104751581	0.021165461741478	0.020434470576699
	0.750	4	0.000201831950260	0.000202548283369	0.000153121064077	0.000153121064077	0.002090644768868	0.002966741606281
	0.750	5	0.000007625171011	0.000007641585969	0.000005798980828	0.000005798980828	0.000133991272978	0.000363021417946
								Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.750	6	0.000000239241229	0.000000239589761	0.000000182470063	0.000000182470063	0.000008023895918	0.000038317957562
	0.750	7	0.000000006423625	0.000000006430405	0.000000004912439	0.000000004912439	0.000000390145300	0.000003552888679
	0.750	8	0.000000000150787	0.000000000150908	0.000000000115584	0.000000000115584	0.000000017423521	0.000000293579617
	0.750	9	0.000000000003145	0.000000000003147	0.000000000002415	0.000000000002415	0.000000000677363	0.000000021871794
	0.750	10	0.000000000000059	0.000000000000059	0.000000000000045	0.000000000000045	0.000000000024307	0.000000001483216
	0.750	11	0.000000000000004	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000785	0.000000000092288
	0.750	12	0.000000000000001	0.000000000000001	0.000000000000002	0.000000000000001	0.000000000000024	0.000000000005304
	0.750	13	0.000000000000002	0.000000000000001	0.000000000000005	0.000000000000001	0.000000000000001	0.000000000000283
	0.750	14	0.000000000000047	0.000000000000001	0.000000000000016	0.000000000000001	0.000000000000000	0.000000000000014
	0.750	15	0.000000000000004	0.000000000000001	0.000000000000008	0.000000000000001	0.000000000000000	0.000000000000001
	0.800	1	0.554303553350709	0.554303553350709	0.554303553350709	0.554303553350708	0.608776704808959	0.554303553350709
	0.800	2	0.073293864008972	0.075971969534895	0.057038776207770	0.057038776207770	0.271995266160804	0.131536860403988
	0.800	3	0.005135436336447	0.005176699144562	0.003904788955388	0.003904788955388	0.025789328557899	0.024858246974282
	0.800	4	0.000261121016950	0.000262173869949	0.000198230636857	0.000198230636857	0.002715095927925	0.003847226969135
	0.800	5	0.000010525052418	0.000010550810270	0.000008007676823	0.000008007676823	0.000185536224187	0.000501939336043
	0.800	6	0.000000352282389	0.000000352866000	0.000000268763751	0.000000268763751	0.000011846909238	0.000056497251664
	0.800	7	0.000000010090088	0.000000010102200	0.000000007717961	0.000000007717961	0.000000614275388	0.000005586598976
	0.800	8	0.000000000252655	0.000000000252886	0.000000000193700	0.000000000193700	0.000000029255341	0.000000492329369
	0.800	9	0.000000000005621	0.000000000005625	0.000000000004318	0.000000000004318	0.000000001212955	0.000000039119555
	0.800	10	0.000000000000112	0.000000000000113	0.000000000000087	0.000000000000087	0.000000000046422	0.000000002829467
	0.800	11	0.000000000000002	0.000000000000002	0.000000000000002	0.000000000000002	0.000000000001599	0.000000000187778
	0.800	12	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000051	0.000000000011512
	0.800	13	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000001	0.000000000000002	0.000000000000656
	0.800	14	0.000000000000004	0.000000000000000	0.000000000000004	0.000000000000001	0.000000000000000	0.000000000000035

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.800	15	0.000000000000012	0.000000000000000	0.000000000000037	0.000000000000001	0.000000000000000	0.000000000000002
	0.850	1	0.591896525740067	0.591896525740068	0.591896525740067	0.591896525740066	0.692245639762573	0.591896525740067
	0.850	2	0.082533621665913	0.085913787289891	0.064403844115991	0.064403844115991	0.308602837534978	0.149016420090278
	0.850	3	0.006152878355848	0.006208292854712	0.004684162609628	0.004684162609628	0.031064436078444	0.029891280775895
	0.850	4	0.000332566001911	0.000334077367556	0.000252643528054	0.000252643528054	0.003472003628513	0.004912082485428
	0.850	5	0.000014245874980	0.000014285198035	0.000010843335148	0.000010843335148	0.000251976248197	0.000680626159044
	0.850	6	0.000000506686719	0.000000507633791	0.000000386679639	0.000000386679639	0.000017088087636	0.000081373491735
	0.850	7	0.000000015420755	0.000000015441643	0.000000011798034	0.000000011798034	0.000000941155837	0.000008547489253
	0.850	8	0.000000000410289	0.000000000410711	0.000000000314604	0.000000000314604	0.000000047613506	0.000000800215136
	0.850	9	0.000000000009698	0.000000000009706	0.000000000007451	0.000000000007451	0.000000002097100	0.000000067549417
	0.850	10	0.000000000000206	0.000000000000206	0.000000000000159	0.000000000000159	0.000000000085263	0.000000005190639
	0.850	11	0.000000000000004	0.000000000000004	0.000000000000003	0.000000000000003	0.000000000003120	0.000000000365981
	0.850	12	0.000000000000001	0.000000000000001	0.000000000000002	0.000000000000000	0.000000000000106	0.000000000023837
	0.850	13	0.000000000000009	0.000000000000001	0.000000000000002	0.000000000000000	0.000000000000003	0.000000000001443
	0.850	14	0.000000000000006	0.000000000000001	0.000000000000003	0.000000000000000	0.000000000000000	0.000000000000082
	0.850	15	0.000000000000010	0.000000000000001	0.000000000000006	0.000000000000000	0.000000000000000	0.000000000000004
	0.900	1	0.630046304453233	0.630046304453231	0.630046304453232	0.630046304453232	0.782053712496280	0.630046304453232
	0.900	2	0.092279337574152	0.096482949019713	0.072215718962917	0.072215718962917	0.347823720164774	0.167688553451906
	0.900	3	0.007295178317417	0.007368282430334	0.005560981010366	0.005560981010366	0.037040871242114	0.035576959630715
	0.900	4	0.000417710827801	0.000419835482454	0.000317559230476	0.000317559230476	0.004379684755981	0.006186015811949
	0.900	5	0.000018950310111	0.000019008898789	0.000014430879283	0.000014430879283	0.000336388599600	0.000907144882218
	0.900	6	0.000000713753681	0.000000715248460	0.000000544878914	0.000000544878914	0.000024144433788	0.000114798404740
	0.900	7	0.000000023002403	0.000000023037319	0.000000017602704	0.000000017602704	0.000001407610932	0.000012764853380
	0.900	8	0.000000000648042	0.000000000648790	0.000000000497000	0.000000000497000	0.000000075381768	0.000001265127833

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	0.900	9	0.000000000016220	0.000000000016235	0.000000000012463	0.000000000012463	0.000000003514751	0.000000113062188
	0.900	10	0.000000000000365	0.000000000000366	0.000000000000281	0.000000000000281	0.000000000151281	0.000000009198080
	0.900	11	0.000000000000007	0.000000000000007	0.000000000000006	0.000000000000006	0.000000000005862	0.000000000686633
	0.900	12	0.000000000000001	0.000000000000001	0.000000000000002	0.000000000000001	0.000000000000211	0.000000000047350
	0.900	13	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000007	0.000000000003034
	0.900	14	0.000000000000005	0.000000000000001	0.000000000000007	0.000000000000001	0.000000000000000	0.000000000000182
	0.900	15	0.000000000000007	0.000000000000001	0.000000000000003	0.000000000000001	0.000000000000000	0.000000000000010
	0.950	1	0.668791830377746	0.668791830377746	0.668791830377746	0.668791830377746	0.878439478828364	0.668791830377746
	0.950	2	0.102520861729734	0.107679187005677	0.080473505498705	0.080473505498705	0.389730839207527	0.187577697346059
	0.950	3	0.008569176704919	0.008664090245110	0.006541001316089	0.006541001316089	0.043770511629071	0.041959673449185
	0.950	4	0.000518187559214	0.000521119145952	0.000394250095476	0.000394250095476	0.005457887700152	0.007695475516895
	0.950	5	0.000024821051780	0.000024906469728	0.000018910828094	0.000018910828094	0.000442270497390	0.001190610385467
	0.950	6	0.000000986948858	0.000000989250346	0.000000753692074	0.000000753692075	0.000033492911669	0.000158987602823
	0.950	7	0.000000033576679	0.000000033633442	0.000000025701106	0.000000025701106	0.000002060470599	0.000018656036990
	0.950	8	0.000000000998558	0.000000000999840	0.000000000765964	0.000000000765964	0.000000116443849	0.000001951379083
	0.950	9	0.000000000026383	0.000000000026409	0.000000000020275	0.000000000020275	0.000000005729766	0.000000184054766
	0.950	10	0.000000000000627	0.000000000000628	0.000000000000483	0.000000000000483	0.000000000260274	0.000000015803837
	0.950	11	0.000000000000014	0.000000000000014	0.000000000000010	0.000000000000010	0.000000000010644	0.000000001245189
	0.950	12	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000404	0.000000000090633
	0.950	13	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000014	0.000000000006129
	0.950	14	0.000000000000005	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000387
	0.950	15	0.000000000000006	0.000000000000001	0.000000000000001	0.000000000000000	0.000000000000001	0.000000000000023
	1.000	1	0.708173154060911	0.708173154060909	0.708173154060910	0.708173154060909	0.981658820988201	0.708173154060910
	1.000	2	0.113247345250785	0.119500840765263	0.089175977277815	0.089175977277815	0.434402240133427	0.208709956293232
Cont'd. on next page								

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	1.000	3	0.009981634833707	0.010103105478807	0.007629981983828	0.007629981983828	0.051307156174414	0.049084884456301
	1.000	4	0.000635715465967	0.000639693247585	0.000484061328413	0.000484061328413	0.006727852293840	0.009468695838582
	1.000	5	0.000032061928360	0.000032184057172	0.000024440214131	0.000024440214131	0.000573572824624	0.001541262644932
	1.000	6	0.000001342162488	0.000001345628104	0.000001025320943	0.000001025320943	0.000045701217444	0.000216566826198
	1.000	7	0.000000048068852	0.000000048158850	0.000000036803697	0.000000036803697	0.000002958539568	0.000026743225553
	1.000	8	0.000000001504878	0.000000001507018	0.000000001154577	0.000000001154577	0.000000175947506	0.000002943952379
	1.000	9	0.000000000041854	0.000000000041901	0.000000000032170	0.000000000032170	0.000000009111420	0.000000292247461
	1.000	10	0.000000000001047	0.000000000001048	0.000000000000806	0.000000000000806	0.000000000435587	0.000000026411599
	1.000	11	0.000000000000024	0.000000000000024	0.000000000000018	0.000000000000018	0.000000000018747	0.000000002190312
	1.000	12	0.000000000000002	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000751	0.000000000167804
	1.000	13	0.000000000000004	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000028	0.000000000011945
	1.000	14	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000794
	1.000	15	0.000000000000006	0.000000000000001	0.000000000000207	0.000000000000001	0.000000000000000	0.000000000000049
	1.250	1	0.916105881552751	0.916105881552752	0.916105881552750	0.916105881552749	1.610514096709184	0.916105881552750
	1.250	2	0.173726018887157	0.187806788270284	0.139286510968215	0.139286510968215	0.702335278581951	0.334054616230189
	1.250	3	0.019351663920754	0.019702040751420	0.014910841671314	0.014910841671314	0.103130369335659	0.097523898746181
	1.250	4	0.001545273243601	0.001560229966690	0.001182102178333	0.001182102178333	0.016810385593074	0.023410699768201
	1.250	5	0.000097569404105	0.000098146647649	0.000074596437045	0.000074596437045	0.001786034629251	0.004749149140058
	1.250	6	0.000005109880690	0.000005130417284	0.000003911632538	0.000003911632538	0.000177409262241	0.000832428472909
	1.250	7	0.000000228878646	0.000000229546395	0.000000175503722	0.000000175503722	0.000014329404679	0.000128303679667
	1.250	8	0.000000008959831	0.000000008979704	0.000000006882085	0.000000006882085	0.000001063526622	0.000017635760136
	1.250	9	0.000000000311564	0.000000000312108	0.000000000239692	0.000000000239692	0.000000068757660	0.000002186586690
	1.250	10	0.000000000009747	0.000000000009760	0.000000000007509	0.000000000007509	0.000000004104418	0.000000246856030
	1.250	11	0.000000000000277	0.000000000000277	0.000000000000214	0.000000000000214	0.000000000220619	0.000000025576813
								Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	1.250	12	0.000000000000007	0.000000000000007	0.000000000000006	0.000000000000006	0.000000000011039	0.000000002448365
	1.250	13	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000001	0.000000000000508	0.000000000217782
	1.250	14	0.000000000000004	0.000000000000001	0.000000000000002	0.000000000000001	0.000000000000022	0.000000000018090
	1.250	15	0.000000000000009	0.000000000000001	0.000000000000216	0.000000000000001	0.000000000000001	0.000000000001409
	1.500	1	1.146718097582056	1.146718097582056	1.146718097582053	1.146718097582056	2.459832639455093	1.146718097582053
	1.500	2	0.244315322611049	0.270330914015072	0.200054486120219	0.200054486120219	1.054348230947220	0.495444070205792
	1.500	3	0.033148879885478	0.033962665729718	0.025778947230398	0.025778947230398	0.184555106572951	0.172093855369611
	1.500	4	0.003187283198640	0.003231208884579	0.002451761266228	0.002451761266228	0.035856551787189	0.049301675506972
	1.500	5	0.000241947450462	0.000243993752837	0.000185643818771	0.000185643818771	0.004554795661705	0.011957947104002
	1.500	6	0.000015221290646	0.000015308968459	0.000011681003224	0.000011681003224	0.000541156640724	0.002508836781121
	1.500	7	0.000000818659281	0.000000822087192	0.000000628894483	0.000000628894483	0.000052332639841	0.000463192231960
	1.500	8	0.000000038473272	0.000000038595849	0.000000029592774	0.000000029592774	0.000004651864873	0.000076298988899
	1.500	9	0.000000001605874	0.000000001609902	0.000000001236791	0.000000001236791	0.000000360348450	0.000011340552072
	1.500	10	0.000000000060296	0.000000000060418	0.000000000046494	0.000000000046494	0.000000025778879	0.000001535157598
	1.500	11	0.000000000002057	0.000000000002061	0.000000000001588	0.000000000001588	0.000000001660994	0.000000190751870
	1.500	12	0.000000000000064	0.000000000000064	0.000000000000050	0.000000000000050	0.000000000099635	0.000000021901028
	1.500	13	0.000000000000002	0.000000000000002	0.000000000000001	0.000000000000001	0.000000000005489	0.000000002336772
	1.500	14	0.000000000000007	0.000000000000000	0.000000000000003	0.000000000000001	0.000000000000284	0.000000000232848
	1.500	15	0.000000000000002	0.000000000000000	0.000000000000002	0.000000000000001	0.000000000000014	0.000000000021763
	1.750	1	1.406764691686597	1.406764691686596	1.406764691686593	1.406764691686596	3.585929581044942	1.406764691686593
	1.750	2	0.322976403886020	0.364246907247344	0.270658924109484	0.270658924109484	1.507166105893815	0.698364813558866
	1.750	3	0.052122833960373	0.053744645583142	0.040950943410898	0.040950943410898	0.305383705823291	0.280161238809559
	1.750	4	0.005867960969906	0.005976729803579	0.004542782814872	0.004542782814872	0.068676433744354	0.093029481662724
	1.750	5	0.000520813882703	0.000526758343857	0.000401283575103	0.000401283575103	0.010134059804372	0.026210741384015

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	1.750	6	0.000038272905016	0.000038571320865	0.000029456747786	0.000029456747786	0.001399344113994	0.006396475349668
	1.750	7	0.000002403344616	0.000002416988348	0.000001850211799	0.000001850211799	0.000157457764914	0.001374818303666
	1.750	8	0.000000131835453	0.000000132405485	0.000000101571372	0.000000101571372	0.000016291783698	0.000263792690076
	1.750	9	0.000000006422120	0.000000006443991	0.000000004952514	0.000000004952514	0.000001469723179	0.000045688285245
	1.750	10	0.000000000281389	0.000000000282164	0.000000000217208	0.000000000217208	0.000000122476204	0.000007208877766
	1.750	11	0.000000000011204	0.000000000011230	0.000000000008657	0.000000000008657	0.000000009194949	0.000001044268852
	1.750	12	0.000000000000409	0.000000000000410	0.000000000000316	0.000000000000316	0.000000000642771	0.000000139797549
	1.750	13	0.000000000000014	0.000000000000014	0.000000000000011	0.000000000000011	0.000000000041275	0.000000017393634
	1.750	14	0.000000000000002	0.000000000000001	0.000000000000003	0.000000000000001	0.000000000002486	0.000000002021260
	1.750	15	0.000000000000005	0.000000000000001	0.000000000000002	0.000000000000001	0.000000000000140	0.000000000220327
	2.000	1	1.704384022237168	1.704384022237164	1.704384022237166	1.704384022237165	5.063195563588776	1.704384022237166
	2.000	2	0.407482959021719	0.465316769736917	0.349731319261152	0.349731319261152	2.082500704070905	0.949859867567970
	2.000	3	0.076975075420392	0.079856833695722	0.061138762077413	0.061138762077413	0.477909181697781	0.430421727892313
	2.000	4	0.009938485867380	0.010176247866988	0.007749515848376	0.007749515848376	0.121728448995156	0.162114827201644
	2.000	5	0.001010635988608	0.001025554896110	0.000782373710140	0.000782373710140	0.020427516890548	0.051938812330543
	2.000	6	0.000084997493026	0.000085857624715	0.000065635867101	0.000065635867101	0.003209576855669	0.014435678804294
	2.000	7	0.000006105189845	0.000006150255008	0.000004711609161	0.000004711609161	0.000411482724917	0.003537156510879
	2.000	8	0.000000382960827	0.000000385116271	0.000000295603958	0.000000295603958	0.000048529283232	0.000774225337396
	2.000	9	0.000000021328625	0.000000021423248	0.000000016472360	0.000000016472360	0.000004993114763	0.000153037743319
	2.000	10	0.000000001068332	0.000000001072166	0.000000000825652	0.000000000825652	0.000000474686734	0.000027566812787
	2.000	11	0.000000000048625	0.000000000048769	0.000000000037606	0.000000000037606	0.000000040668703	0.000004559894342
	2.000	12	0.000000000002028	0.000000000002033	0.000000000001570	0.000000000001570	0.000000003244899	0.000000697169369
	2.000	13	0.000000000000078	0.000000000000078	0.000000000000060	0.000000000000060	0.000000000237876	0.000000099078681
	2.000	14	0.000000000000004	0.000000000000003	0.000000000000002	0.000000000000002	0.000000000016362	0.000000013152409

Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	2.000	15	0.000000000000002	0.000000000000001	0.000000000000002	0.000000000000001	0.000000000001048	0.000000001637861
	2.250	1	2.049480244785901	2.049480244785903	2.049480244785902	2.049480244785895	6.988848153828524	2.049480244785902
	2.250	2	0.495655843411444	0.568835053521580	0.435334985330309	0.435334985330309	2.808185030172211	1.258855813390984
	2.250	3	0.108373023511148	0.113055230783004	0.087049696486887	0.087049696486887	0.717671533993121	0.633268204786377
	2.250	4	0.015789609642188	0.016262315877622	0.012409700124868	0.012409700124868	0.203590965258827	0.266038513701729
	2.250	5	0.001811475773225	0.001844943807289	0.001409708965860	0.001409708965860	0.038222934420611	0.095341802885027
	2.250	6	0.000171666406959	0.000173849365144	0.000133055305828	0.000133055305828	0.006723152170537	0.029693617634219
	2.250	7	0.000013885391266	0.000014014451112	0.000010745409604	0.000010745409604	0.000966348598452	0.008162115429323
	2.250	8	0.000000980498534	0.000000987456349	0.000000758440709	0.000000758440709	0.000127834847001	0.002005677180438
	2.250	9	0.000000061461355	0.000000061805425	0.000000047546945	0.000000047546945	0.000014762639840	0.000445305343735
	2.250	10	0.000000003464476	0.000000003480173	0.000000002681130	0.000000002681130	0.000001575723894	0.000090129472570
	2.250	11	0.000000000177437	0.000000000178101	0.000000000137384	0.000000000137384	0.000000151623729	0.000016755913887
	2.250	12	0.000000000008327	0.000000000008353	0.000000000006451	0.000000000006451	0.000000013590312	0.000002879835045
	2.250	13	0.000000000000361	0.000000000000362	0.000000000000280	0.000000000000280	0.000000001119421	0.000000460137942
	2.250	14	0.000000000000014	0.000000000000015	0.000000000000011	0.000000000000011	0.000000000086526	0.000000068681487
	2.250	15	0.000000000000003	0.000000000000001	0.000000000000002	0.000000000000001	0.000000000006226	0.000000009617792
	2.500	1	2.454185163821871	2.454185163821870	2.454185163821873	2.454185163821868	9.489132622449082	2.454185163821873
	2.500	2	0.585629275823988	0.670871261190140	0.525128225730018	0.525128225730017	3.719624322192464	1.636580072248835
	2.500	3	0.146973065463321	0.154054758337564	0.119391017468873	0.119391017468873	1.044406404229517	0.901239727671028
	2.500	4	0.023845075428935	0.024717535224060	0.018902346522881	0.018902346522881	0.325579345077541	0.416657983374334
	2.500	5	0.003049402681084	0.003118087961153	0.002386712930066	0.002386712930066	0.067501014924178	0.164849699670401
	2.500	6	0.000321656250623	0.000326666067545	0.000250330075624	0.000250330075624	0.013120565907238	0.056792978121632
	2.500	7	0.000028940154819	0.000029270358767	0.000022463962212	0.000022463962212	0.002087425674981	0.017290553941823
	2.500	8	0.000002272285549	0.000002292107689	0.000001761797237	0.000001761797237	0.000305810336598	0.004709822754096
								Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	2.500	9	0.000000158341213	0.000000159431921	0.000000122721951	0.000000122721951	0.000039138702570	0.001159809969886
	2.500	10	0.000000009920735	0.000000009976078	0.000000007689178	0.000000007689178	0.000004631375178	0.000260468860454
	2.500	11	0.000000000564707	0.000000000567311	0.000000000437783	0.000000000437783	0.000000494258746	0.000053745651944
	2.500	12	0.000000000029453	0.000000000029567	0.000000000022841	0.000000000022841	0.000000049144087	0.000010254685210
	2.500	13	0.000000000001417	0.000000000001422	0.000000000001100	0.000000000001100	0.000000004491516	0.000001819250959
	2.500	14	0.000000000000063	0.000000000000064	0.000000000000049	0.000000000000049	0.000000000385275	0.000000301542203
	2.500	15	0.000000000000003	0.000000000000003	0.000000000000005	0.000000000000002	0.000000000030770	0.000000046895345
	2.750	1	2.933419823583611	2.933419823583610	2.933419823583614	2.933419823583613	12.727363779681726	2.933419823583614
	2.750	2	0.676099799838213	0.769100711044218	0.616730931281785	0.616730931281784	4.861641415602579	2.097089157334456
	2.750	3	0.193452990131174	0.203557995870570	0.158881741383581	0.158881741383581	1.483241641802373	1.249573633795149
	2.750	4	0.034553384556415	0.036070184181193	0.027644251384490	0.027644251384490	0.502547985719467	0.628733639595687
	2.750	5	0.004878528290171	0.005009623860124	0.003841994650577	0.003841994650577	0.113842135215208	0.271688755577640
	2.750	6	0.000567157090877	0.000577753046480	0.000443356169695	0.000443356169695	0.024196514077570	0.102452858780760
	2.750	7	0.000056199918655	0.000056970972355	0.000043768654128	0.000043768654128	0.004216802414487	0.034189498076812
	2.750	8	0.000004857820632	0.000004908855872	0.000003776153663	0.000003776153663	0.000677089107758	0.010217535521553
	2.750	9	0.000000372570435	0.000000375664380	0.000000289349706	0.000000289349706	0.000095052611373	0.002762245598819
	2.750	10	0.000000025687738	0.000000025860621	0.000000019942634	0.000000019942634	0.000012342229222	0.000681333237907
	2.750	11	0.000000001608895	0.000000001617849	0.000000001248993	0.000000001248993	0.000001445939433	0.000154459958167
	2.750	12	0.000000000092325	0.000000000092757	0.000000000071681	0.000000000071681	0.000000157864997	0.000032386763963
	2.750	13	0.000000000004889	0.000000000004908	0.000000000003796	0.000000000003796	0.000000015846732	0.000006315224116
	2.750	14	0.000000000000240	0.000000000000241	0.000000000000187	0.000000000000187	0.000000001493214	0.000001150681876
	2.750	15	0.000000000000011	0.000000000000011	0.000000000000009	0.000000000000009	0.000000000131026	0.000000196740950
	3.000	1	3.505581678350746	3.505581678350747	3.505581678350748	3.505581678350751	16.914316866680025	3.505581678350748
	3.000	2	0.766525476672830	0.862937127326289	0.708197236840081	0.708197236840082	6.290815856846200	2.657934118129254
								Cont'd. on next page

Table 3: Root-mean-square Approximation Errors, cont'd.

	c	n	Num.-optimal	Mathieu	Interpolatory	Oblate	Univar.-Remez	Taylor-series
	3.000	3	0.248553453889047	0.262301157077932	0.206273949473054	0.206273949473054	2.066211766959514	1.696888244122908
	3.000	4	0.048377237704455	0.050887377294868	0.039084460497474	0.039084460497474	0.753928564998021	0.920590968058153
	3.000	5	0.007482620913229	0.007718148683242	0.005931710406253	0.005931710406253	0.184961434519945	0.430578723019402
	3.000	6	0.000950995600558	0.000971941113430	0.000746967690762	0.000746967690762	0.042610945678093	0.176168177898430
	3.000	7	0.000102939876458	0.000104609204643	0.000080458070528	0.000080458070528	0.008064226153486	0.063883416292738
	3.000	8	0.000009715500953	0.000009836348105	0.000007573253075	0.000007573253075	0.001407039647048	0.020767216960030
	3.000	9	0.000000813372055	0.000000821378237	0.000000633092285	0.000000633092285	0.000214820473677	0.006111331252045
	3.000	10	0.000000061205241	0.000000061693881	0.000000047602462	0.000000047602462	0.000030348140188	0.001641682481413
	3.000	11	0.000000004183334	0.000000004210968	0.000000003252409	0.000000003252409	0.000003870055546	0.000405468075870
	3.000	12	0.000000000261947	0.000000000263402	0.000000000203632	0.000000000203632	0.000000460042083	0.000092647582720
	3.000	13	0.000000000015134	0.000000000015206	0.000000000011766	0.000000000011765	0.000000050294349	0.000019690986886
	3.000	14	0.000000000000812	0.000000000000815	0.000000000000631	0.000000000000631	0.000000005162373	0.000003911221858
	3.000	15	0.000000000000041	0.000000000000041	0.000000000000032	0.000000000000032	0.000000000493534	0.000000729091557

11 Examples

In this section, I provide three examples: the first, in Subsection 11.1, shows how well my portfolio approximations work in capturing the scenario behavior of an idealized 30-year Treasury bond, even in very large scenario shifts. This example also displays the approximating-cashflow-stream interpretation discussed in Subsection 7.1 (for the numerically-optimal approximation for worst-case error) and in Subsection 8.5 (for the interpolatory, mean-square-error optimal approximation). The second example, in Subsection 11.2, demonstrates how a multidimensional scenario problem can be handled; I use Mathieu approximation in a two-dimensional equity scenario problem that includes both dividend growth and discounting changes. The third example, in Subsection 11.3, shows how a sequence of numerically-optimal approximations can solve a structured, but higher-dimensional, scenario problem in a fraction of a second; the same problem would require more than a day in a direct approach.

11.1 A Simple One-factor Example

I consider an idealized version of a 30-year Treasury bond, which has 60 cashflow payment dates extended over 30 years. The payment dates are *not* separated by exactly the same times (though they are on a hypothetical semiannual schedule), which is both realistic and useful in demonstrating that my approximations do not require any regular spacing in the cashflow streams to be approximated.

Some parts of this example were discussed in the introduction; I repeat some comments and two figures from the introduction here for clarity and convenience.

I focus through most of the example on rank $n = 3$ approximations; these improve upon the traditional Taylor series that uses a constant, a linear term, and a quadratic term (this Taylor series is equivalent to approximation using base value, duration, and convexity).

For ease of exposition, I constructed the whole bond (coupons and principal strip) to be valued at par, with the coupons worth just over 59.34 and the principal strip worth just under 40.66. In Figure 15, I begin by considering a 100-basis-point parallel shift in the rates curve. As Sections 2 and 3 make clear, the shape of any intial curve impacts only the present values of the cashflow stream to be approximated (see the upper right panel of Figure 15 for the present values of the cashflows of this hypothetical Treasury bond), but the shape of the vector of shift directions is very obviously very important in any scenario analysis. The middle left panel of Figure 15 shows the approximating cashflow streams used by the numerically-optimal approximation. The numerically-optimal approximation produces $n = 3$ approximating cashflows streams (because the rank of the approximation being used here is $n = 3$), each having cashflows on the same $n + 1 = 4$ cashflow dates, see Subsection 7.1.

The interpretation of the middle left panel in Figure 15 is especially helpful. The approximating cashflow times chosen for numerically-optimal worst-case error approximation are the two most extreme available (the times corresponding to 6 months and to 30 years) and two intermediate times (one in the belly of the curve, the other in the long end between 20 and 25 years). Although the approximation does not order the streams in any particular way, there are clearly a "level annuity" stream (Stream 2 in the panel, in which the present values of the cashflows across the approximating cashflow times all have the same sign and are similar in magnitude), a "flattener" stream (Stream 3 in the panel, in which the present values of the cashflows at the first two approximating cashflow times are negative and the present values of the cashflows at the second two approximating cashflow times are positive), and a "butterfly" stream (Stream 1 in the panel, in which the present values of the cashflows at the middle two approximating cashflow times are negative and the present values of the cashflows at the first and last approximating cashflow times are positive). Note that the approximation is indifferent to changing the signs of all of the present values in a given approximating cashflow stream, since the approximating portfolio is permitted to have short positions in an approximating cashflow stream, but the shape of the present values over approximating cashflow times is very important.

Unsurprisingly, the numerically-optimal (for worst-case error) approximating portfolio for the hypothetical Treasury bond of Figure 15 in a set of parallel-shift scenarios is focused on Stream 2, the "level annuity" discussed above (matching the coupon stream), and on Stream 3, the "flattener" discussed above (which relates to the principal strip). It is worth noting that the present value of the approximating cashflow stream is always, by construction, equal to the approximation's value for the target portfolio with no scenario movement.

As Subsection 8.5 shows, a useful cashflow interpretation is also available for the interpolatory (and root-mean-square error optimal) approximation of that subsection. The middle right panel of Figure 15 shows this interpretation for the hypothetical Treasury bond considered in the figure and for a set of parallel shifts. The "level annuity" (here, Stream 3), "flattener" (here, Stream 2), and "butterfly" (here, Stream 1) approximating streams are, as in the numerically-optimal case of the middle left panel, clearly identifiable. Interestingly, the interpolatory approximation (with its focus on average squared error) weights these streams rather differently, choosing a much larger butterfly position than the numerically-optimal approximation (with its focus on worst-case error) does.

In the bottom left panel of Figure 15, I show approximation error across parallel shifts. Each approximation is conducted in the recentered form of (32), so the Taylor series here will actually be much better than the standard three-term series based at zero (as often used in financial markets). The shift multiples of -5 to 5 mean that I consider parallel shifts of -500 to 500 basis points. That 1,000 basis-point range in rates is quite wide relative to standard scenario analysis ranges; I wish to show how well my approximations perform even under challenging conditions. The first feature that emerges from the bottom left panel of Figure 15 is the terrible performance of the Taylor series, even after my recentering of it (which greatly improves it). It makes errors of more that 6.62 bond points across the scenario set; since the hypothetical bond is valued at par, this corresponds to 6.62 percent return errors across the scenario set. The univariate-Remez approximation is also deeply problematic; it incurs a nontrivial approximation error even when no scenario shift has happened!

Discarding the obviously inadequate Taylor series and univariate-Remez approximations and rescaling the vertical axis to get a better view of the four approximations I introduce in this work, I come to the bottom right panel of Figure 15. The error incurred by the numerically-optimal and Mathieu approximations is excellent; the magnitude of error is less than 0.42 points across a 1,000 basis-point range of rates! The numerically-optimal and Mathieu approximations are more than 16 times more accurate than the Taylor series. The interpolatory and oblate approximations perform well, but not as well, with maximum absolute errors of 1.107 (to three decimals) and 0.984, respectively; the presence of an important cashflow at an extreme time (the principal strip) in the target cashflow stream creates challenges for them.

The sizes of errors discussed above should be viewed in the context of the wide range of rates considered; for smaller ranges or for a higher approximating rank, the error of the numerically-optimal and Mathieu approximations is reduced to negligible levels. It seemed worth highlighting that even with only 3 degrees of freedom (rank-3 approximation), the numerically-optimal and Mathieu approximations can perform well enough across 1,000 basis-point rate ranges to be useful for high-level risk management.

Figure 16 repeats the analysis of Figure 15, but applied to just the coupons of the hypothetical Treasury bond. As noted in Subsections 7.1 and 8.5, the approximating cashflows used by the numerically-optimal (Subsection 7.1) and interpolatory (Subsection 8.5) approximations do not change across different target cashflow streams; only the amounts of them held in the approximating portfolio change. This fact is evident in the two middle panels of Figure 16, where the approximating cashflow streams are identical to those in Figure 15, but the holdings of them change to emphasize the "level annuity" approximating streams (as expected given the shape of the present values of the coupon stream).

The two bottom panels of Figure 16 show that the coupon stream is much easier to approximate than the whole bond, though there is again a great difference between the four approximations that I introduce in this work and the Taylor and univariate-Remez approximations. The average-squared-error optimal approximations (the interpolatory and oblate approximations) really come into their own with this smooth profile of cashflow present values, and outpeform the numerically-optimal and Mathieu approximations in this case. The interpolatory approximation never exceeds 0.0051 points in error here, and the oblate approximation never exceeds 0.0082 points; this is across a 1,000 basis-point range in rates.

Figure 17 again repeats the analysis of Figure 15, but applied to just the principal strip of the hypothetical Treasury bond. As one would expect, the middle panels of Figure 17 show the same approximating cashflow streams used by the numerically-optimal (middle left panel) and interpolatory (middle right panel) approximations as in Figure 15 and Figure 16, though the numerically-optimal approximating portfolio and the interpolatory approximating portfolio do change their holdings of these approximating cashflows to emphasize the "flattener" and the "butterfly." The two bottom panels of this figure show that the principal strip is much harder to approximate than the coupon stream, and the numerically-optimal and Mathieu approximations really show their value here. The Taylor series maximal absolute error is over 7.08 points (on a base value of 40.6553), while the numerically-optimal approximation achieves a maximal absolute error of 0.3612296 and the Mathieu approximation achieves 0.3723909. The numerically-optimal approximation is more than 19 times more accurate than the Taylor series here, an advantage which only grows (in these relative terms) when higher-rank approximations are considered.

Inspecting the middle panels of Figures 15, 16, and 17 shows not only that the approximating cashflow streams used as the building blocks of the approximating portfolio are the same for every target cashflow stream, but also that aggregation is linear (as noted in Subsections 7.1 and 8.5). To get the approximating portfolio used by the numerically-optimal approximation for the whole bond (shown in the middle left panels), simply add the holdings of the numerically-optimal approximating portfolio for the coupon stream to the holdings of the numerically-optimal approximating portfolio for the principal strip. The same is true for the interpolatory approximation, as shown in the middle right panels.

Figures 18, 19, and 20 repeat the analysis of Figures 15, 16, and 17, but with a set of scenarios that shifts the slope of the rates curve. The most interesting aspect of these results is the approximating cashflow streams used by the numerically-optimal and interpolatory approximations in the middle panels: for a given scenario set, the approximating cashflow streams used are the same for any target cashflow stream, but different scenario sets definitely do lead to different approximating cashflow streams. The approximating cashflow streams for the slope scenarios are much more focused on the long end, understandably.

Figures 21, 22, and 23 repeat the analysis of Figures 15, 16, and 17, but with a set of scenarios that shifts the curvature of the rates curve. With a new scenario set come new approximating cashflow streams for both the numerically-optimal and the interpolatory approximations, and the middle panels of these figures reflect that, with interesting bunching of approximating cashflow times in the long end. The behavior of the principal strip here is particularly difficult to approximate; the Taylor series approximation incurs a maximal absolute error of more than 20.73 points on a base value of 40.6553 points. The numerically-optimal and Mathieu approximations fare much better, with maximal absolute errors of 0.9226813 and 0.9556714 points, respectively. Thus, the numerically-optimal approximation is more than 22 times more accurate than the Taylor series in this example.

Finally, Figure 24 revisits the analysis of Figures 15 but increases the rank of all approximations to 5. All approximations become much more accurate, but the relative advantage of the numerically-optimal and Mathieu approximations grows even greater. The maximal absolute error of the Taylor series approximation here is 0.1688086 points; this seems reasonable until one views the corresponding errors for the numerically-optimal approximation, at 0.0006286 points, and the Mathieu approximation, at 0.0006441 points. The error for the Taylor series is not large, but for the same number of terms one can achieve negligible error with the numerically-optimal or Mathieu approximations. The numerically-optimal approximation is more than 268 times more accurate than the Taylor series approximation here; as mentioned above, this relative advantage only grows as approximation rank increases further. Table 2 shows how extreme the relative advantage can become.

11.2 A Two-factor Example: Dividend Discounting

In this simple two-dimensional example, I consider an equity valuation problem in which both the interest rate curve and the dividend growth rate curve are subjected to parallel-shift scenarios. Figure 25 summarizes this situation. The cashflows are quarterly dividends over the next 100 years. I allow parallel shifts from −100 basis points through 100 basis points in both curves. The Mathieu approximations are well-suited to these multidimensional problems, and I construct a 16-term approximation by taking the product of a 4-term approximation for the interest rate and a 4-term approximation for the dividend growth.

As Figure 25 shows, the stock value goes from above 80 to below 40 across these scenarios, but the multidimensional Mathieu approximation achieves error that is uniformly less than 0.002 in magnitude. By using the mulitdimensional Mathieu approximation, I have summarized 400 cashflows over 100 years using 16 numbers to very high precision.

11.3 An Example of Speed: Segmented Scenarios

One of the motivations for this work was the potential to improve the speed of scenario analysis; here I show how well my approximations can do in speeding the solution of complex scenario problems. I randomly generated cashflow present values at 5,801 times (evenly-spaced, though that is not important to the analysis) over 30 years, and considered a set of scenarios involving parallel shifts of 4 separate segments of the rates curve: from 0 to 2 years, from 2 to 5 years, from 5 to 10 years, and from 10 to 30 years. Figure 26 shows the randomly-generated cashflow present values that I used.

In each segment, 201 parallel shifts were considered, ranging from −500 basis points to 500 basis points in increments of 5 basis points. The total number of scenarios under consideration is therefore 1,632,240,801. I built a segment-wise approximation using 4 numerically-optimal approximations, one on each segment. This allowed me to exploit both the separability of the problem and the worst-case dimension reduction of the numerically-optimal approximations. To speed the analysis, I used automatic loading of appropriate approximation parameters from a precalculated database so that I would not need to construct the numerically-optimal approximations at runtime.

I then timed how long it took my approximation to find the greatest loss across all 1,632,240,801 scenarios, exploiting the linearity of the problem to analyze each segment separately and add up the greatest losses due to each segment. The loss was 62.360859299, and it was calculated in 0.050704517 seconds (averaged over several timing trials).

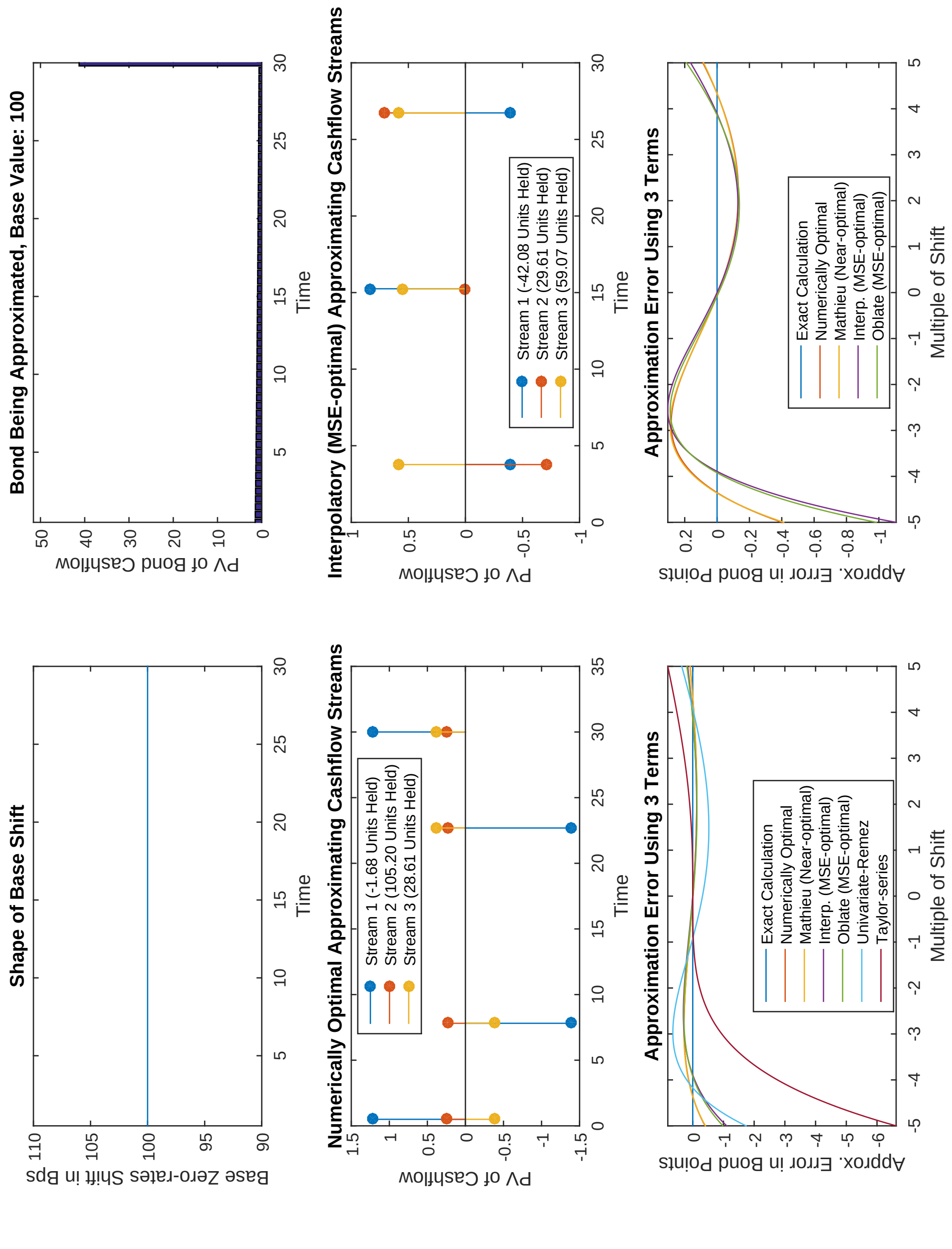

Figure 15: Approximating the Parallel-shift Profile of an Idealized Bullet Bond

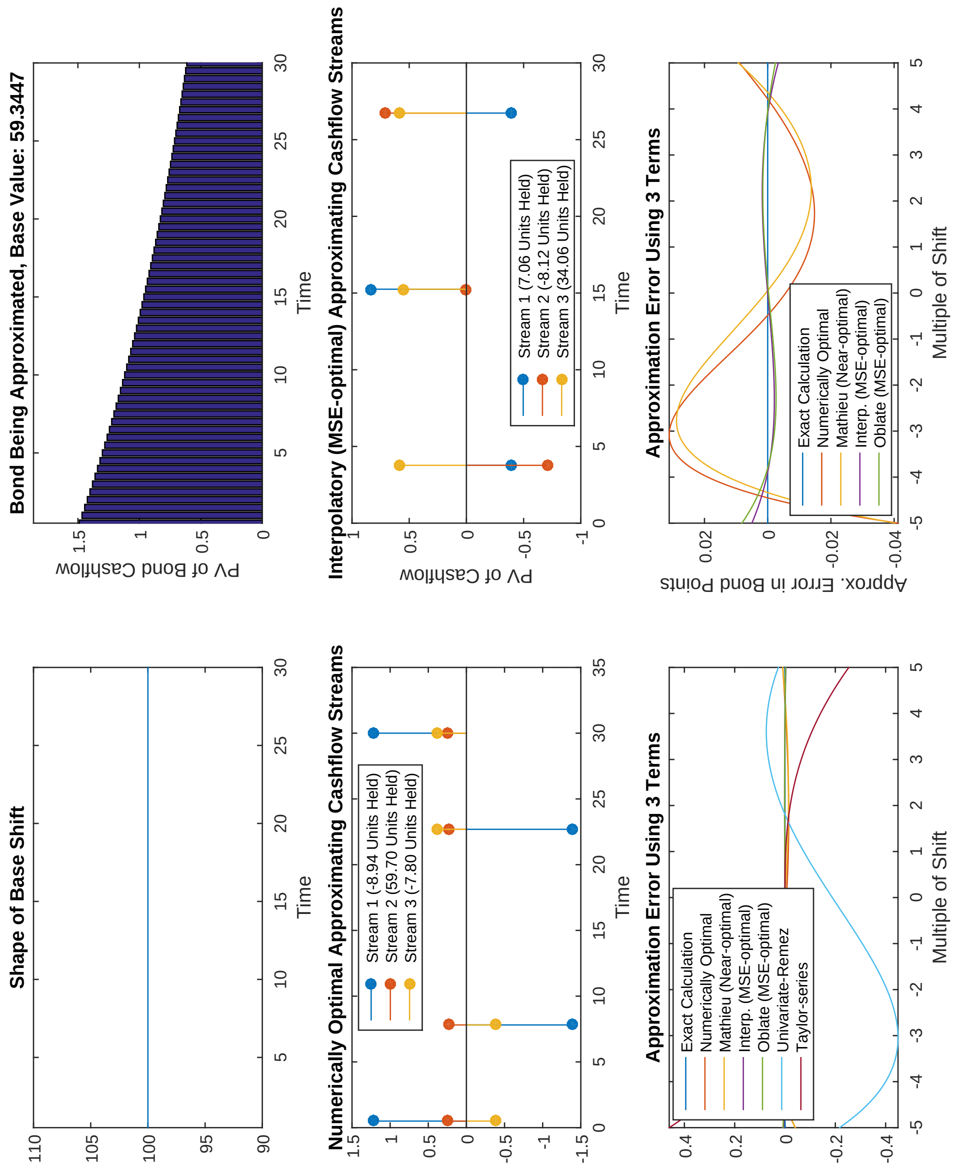

Figure 16: Approximating the Parallel-shift Profile of Coupons

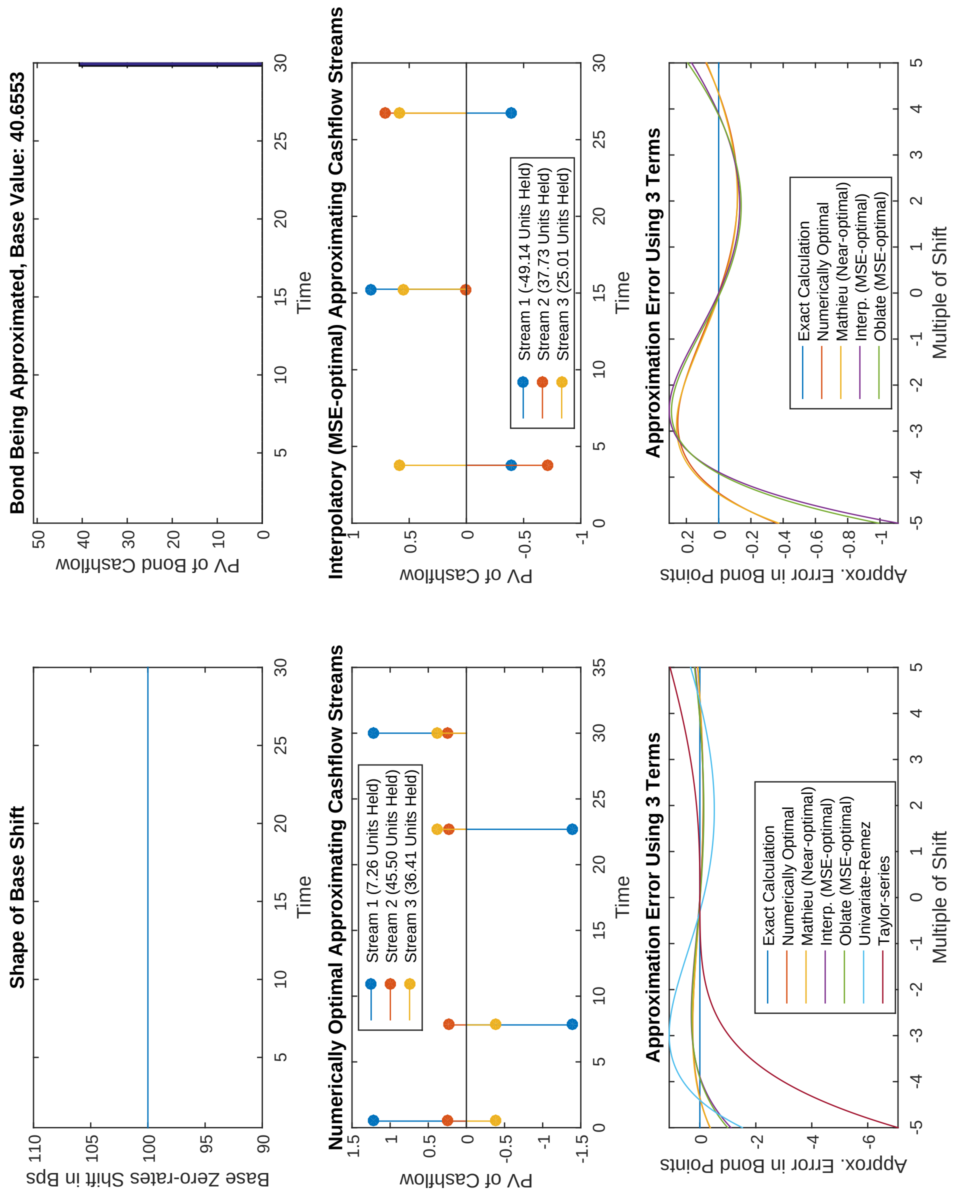

Figure 17: Approximating the Parallel-shift Profile of a Principal Strip

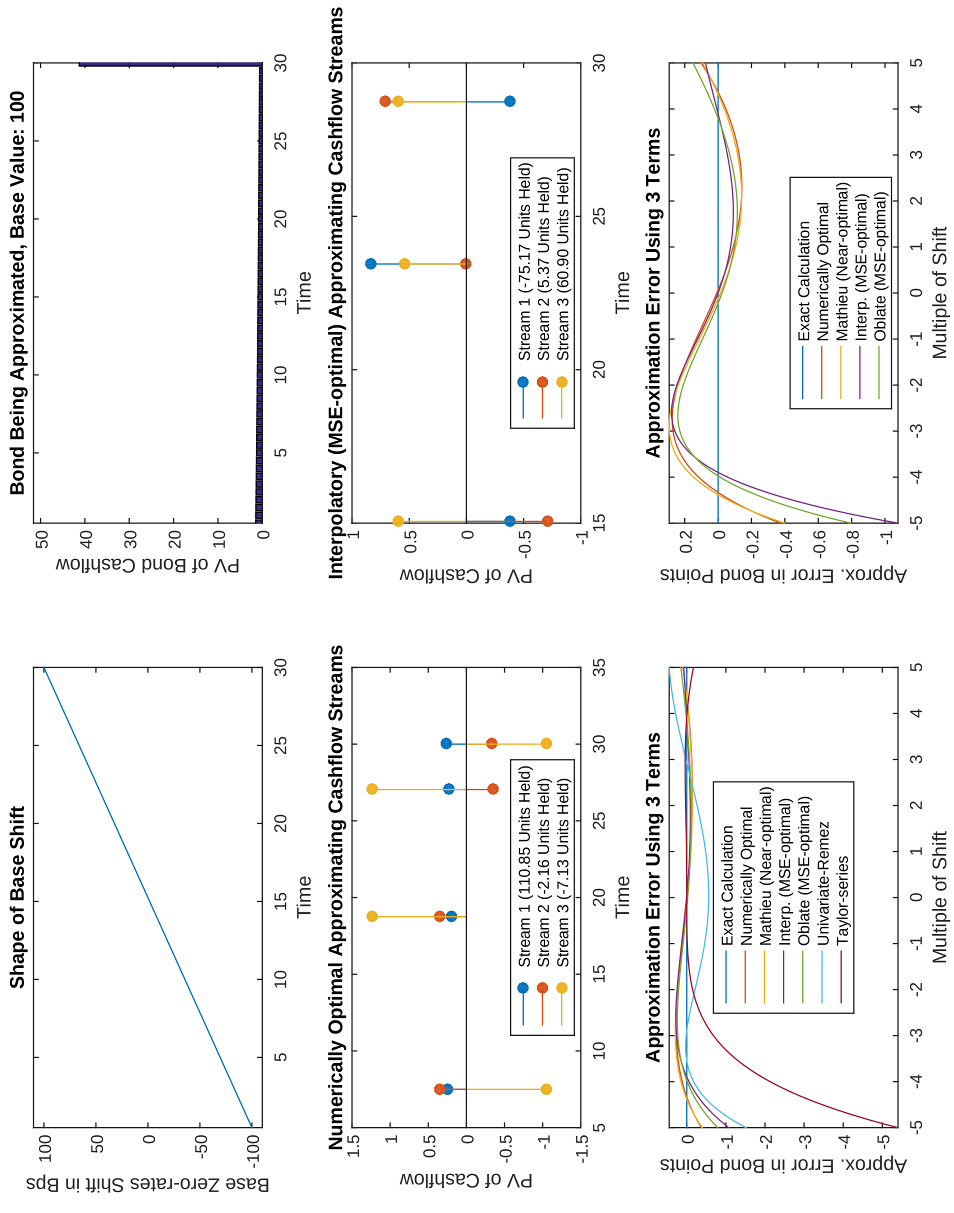

Figure 18: Approximating the Slope-shift Profile of an Idealized Bullet Bond

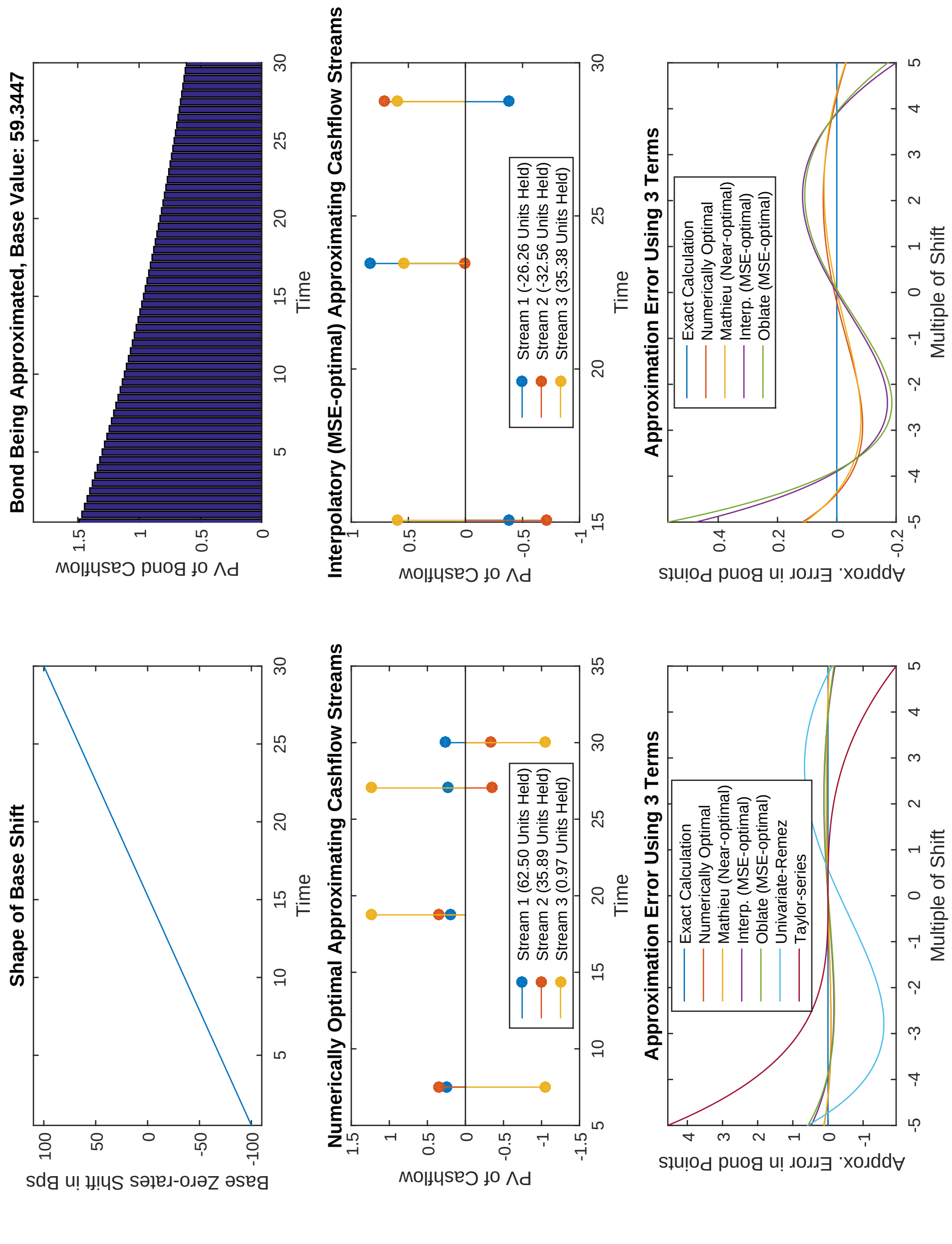

Figure 19: Approximating the Slope-shift Profile of Coupons

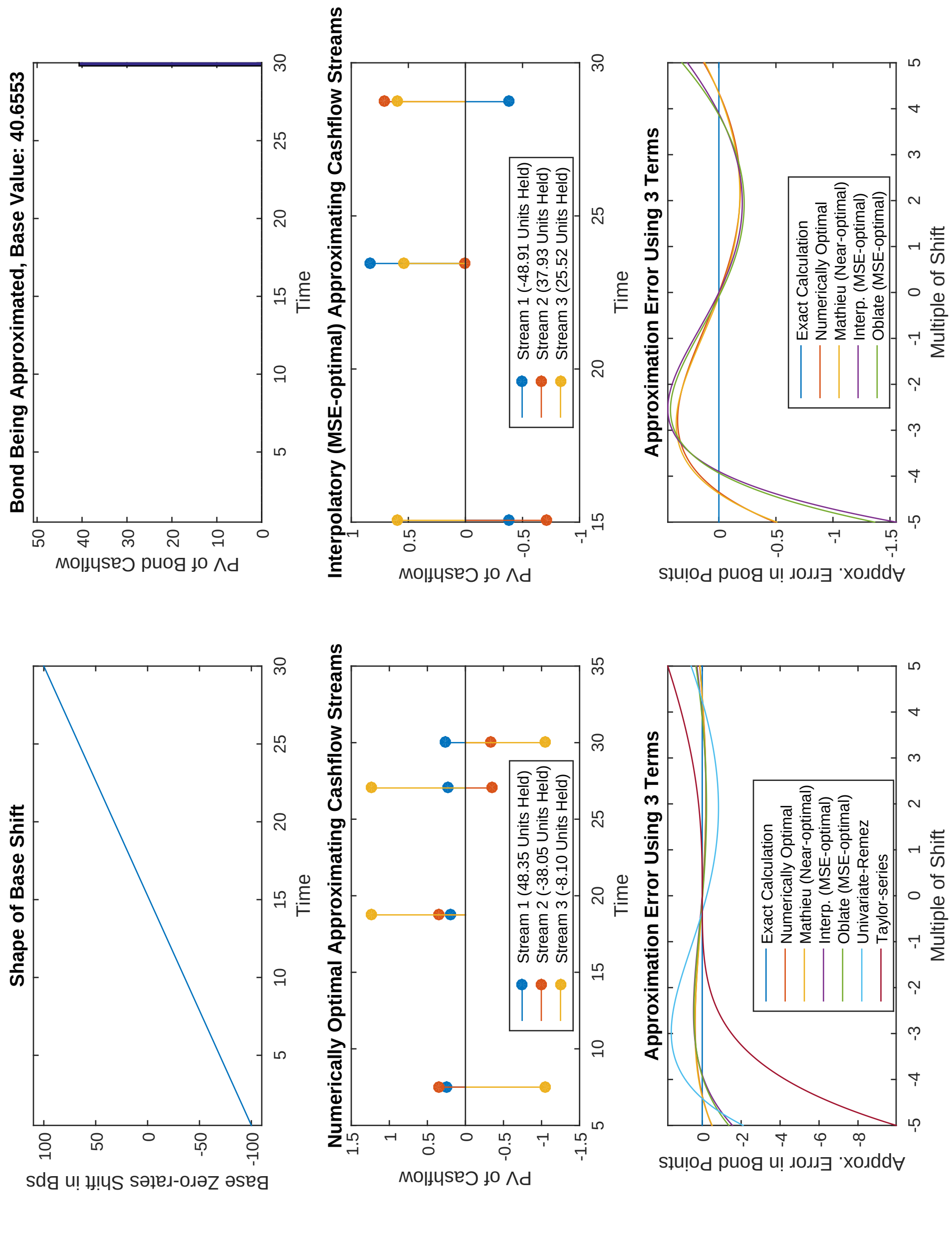

Figure 20: Approximating the Slope-shift Profile of a Principal Strip

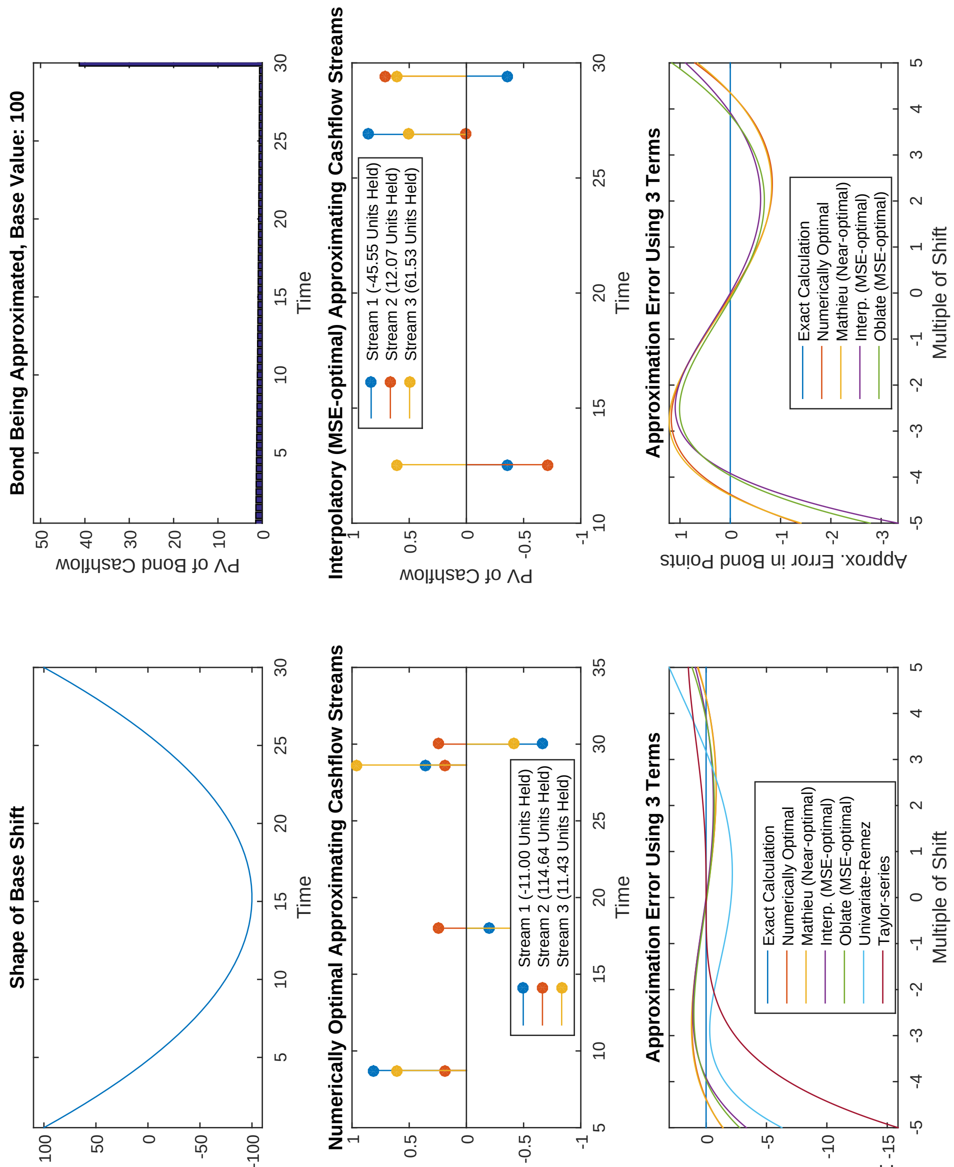

Figure 21: Approximating the Curvature-shift Profile of an Idealized Bullet Bond

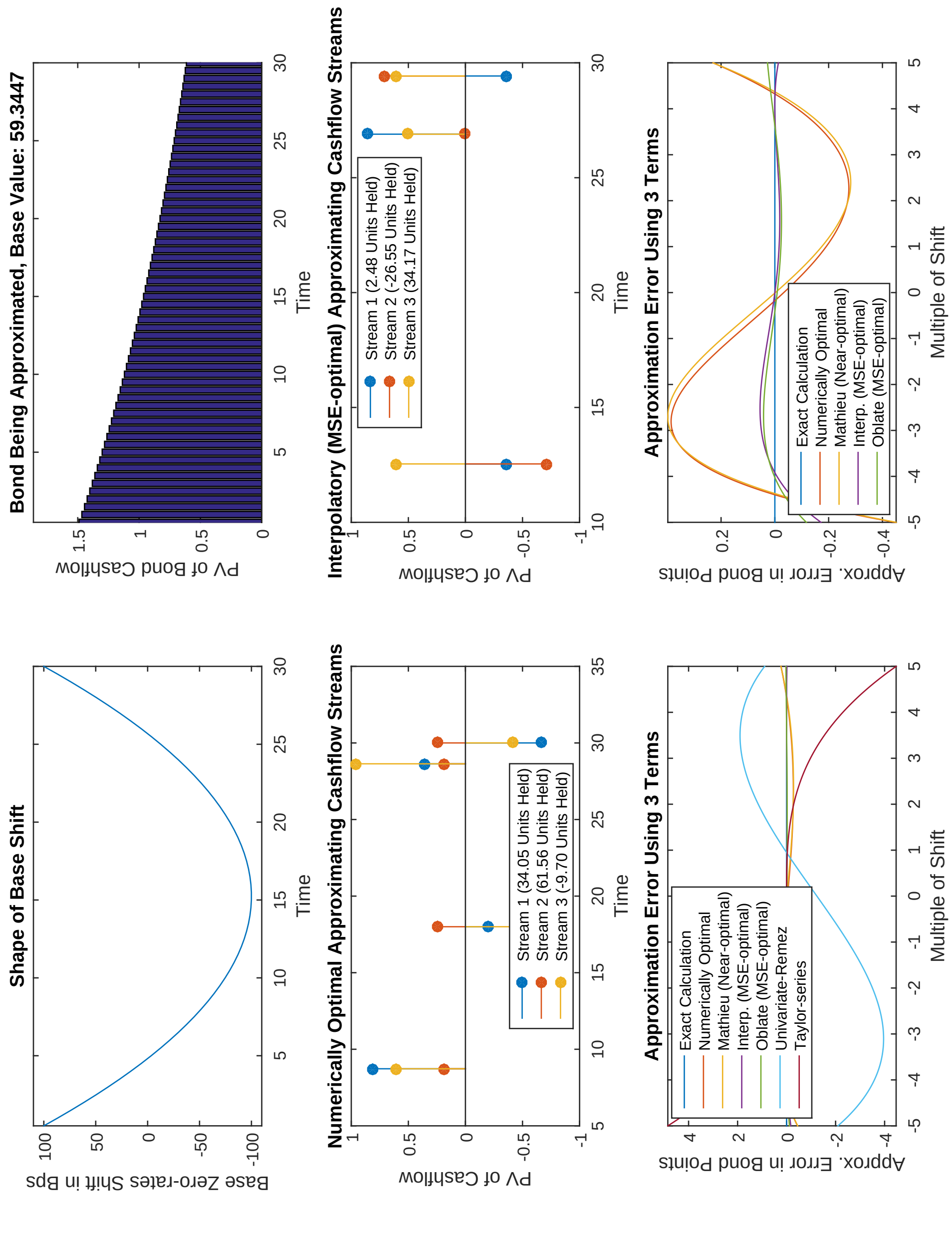

Figure 22: Approximating the Curvature-shift Profile of Coupons

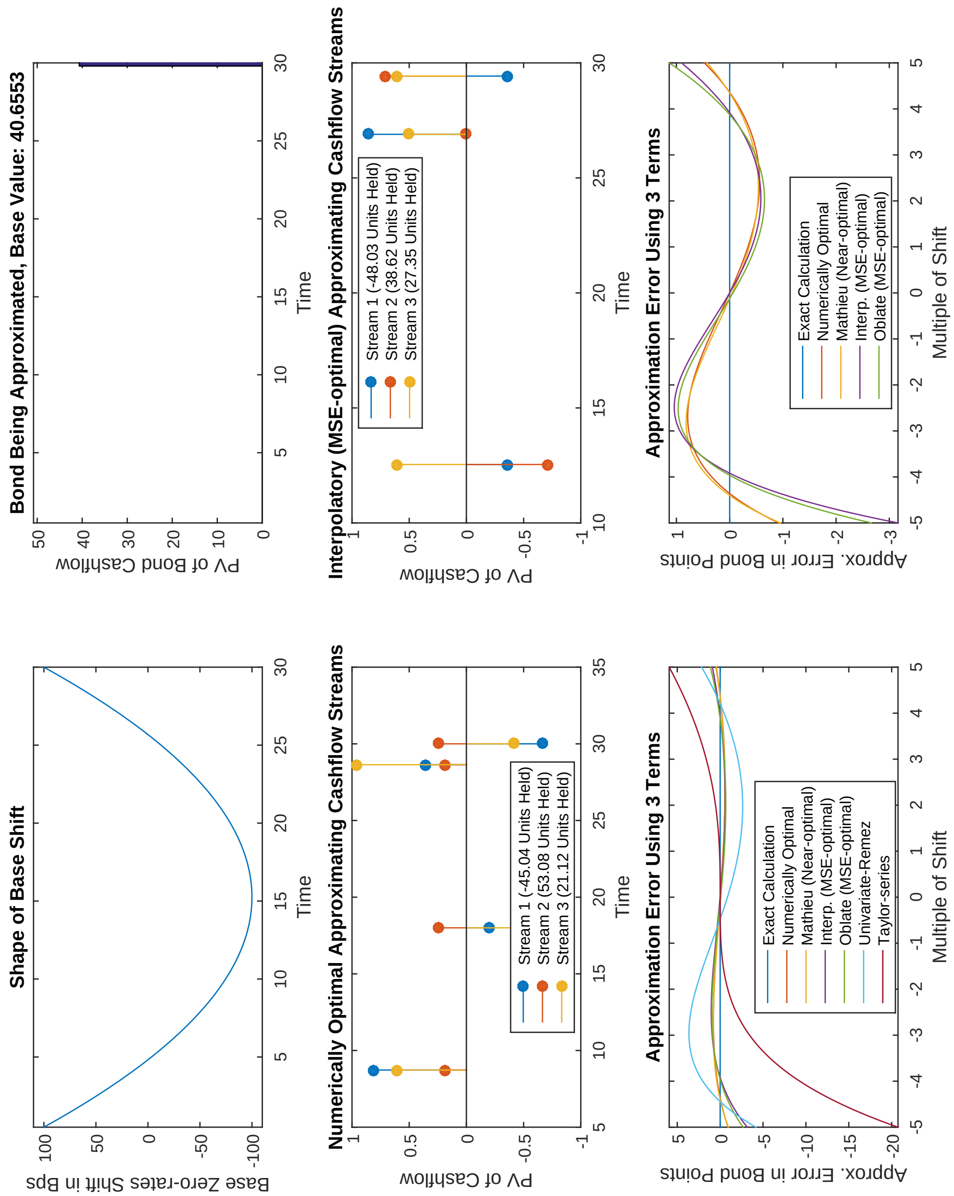

Figure 23: Approximating the Curvature-shift Profile of a Principal Strip

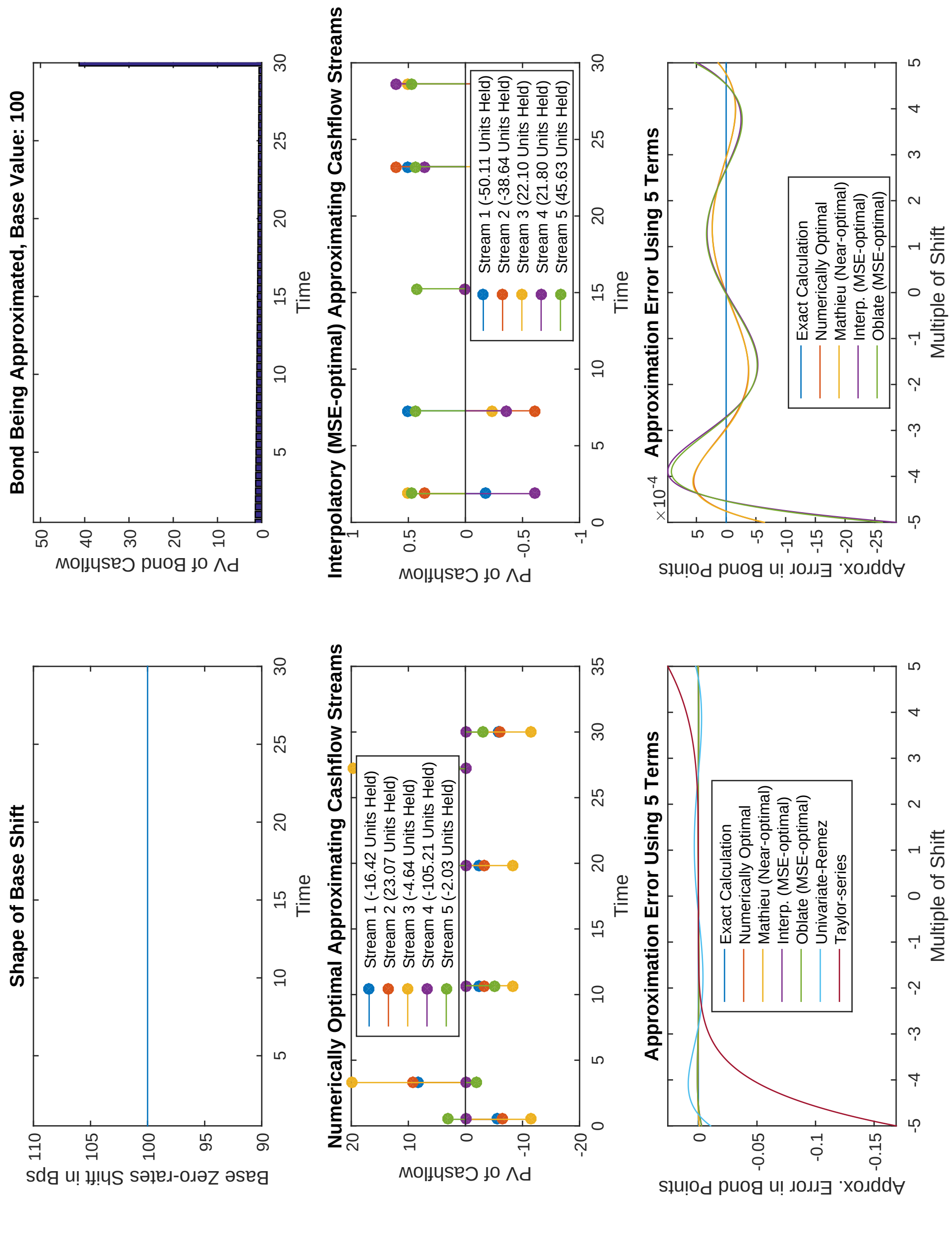

Figure 24: Rank Five Approximation of the Parallel-shift Profile of an Idealized Bullet Bond

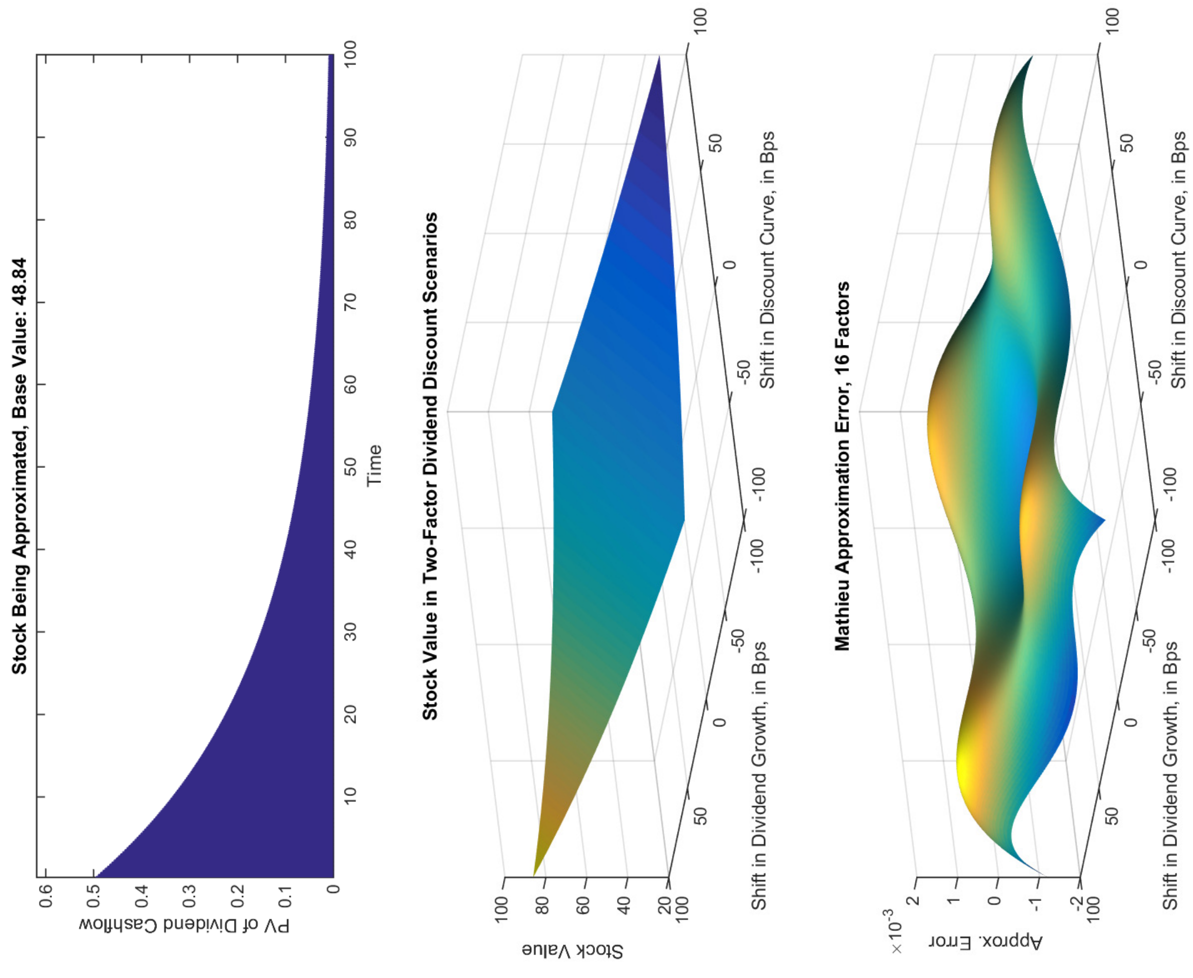

Figure 25: Two-factor Approximation for Equities: Dividend Growth and Discounting Scenarios

To understand the time a direct analysis would require, I revalued the cash-flows 201 times directly, using just one unchanging shift. This is the most favorable arrangement for the direct approach, since no time is taken to change the shift being used. Since there are actually 1,632,240,801 scenarios to consider, I multiplied the resulting time by 8,120,601 to get the implied time for a full analysis. The resulting implied time was 89,585 seconds, so just over 1 day. This substantiates the potential for speed-up that was noted in the abstract.

There are obviously ways to exploit the separability of this problem without using my approximations to reduce dimension within each segment, but they would not be as fast; this example shows that my approximation methods can be used in conjunction with structure-exploiting approaches to further enhance the speed of scenario analyses.

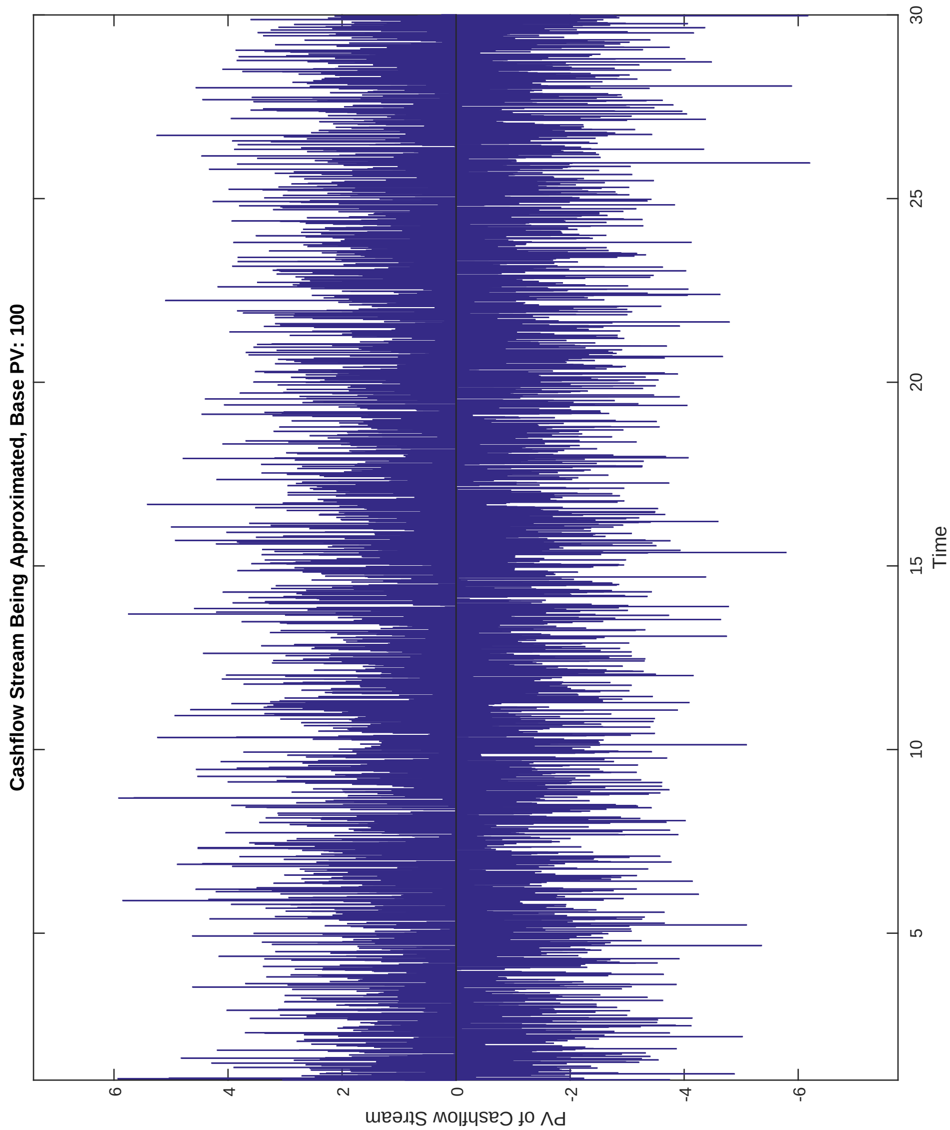

Figure 26: Cashflows, Randomly Drawn, for Segmented Shift Scenarios

12 Conclusion

I introduced four novel approximations of portfolio behavior across a set of scenarios: the numerically-optimal, Mathieu, interpolatory, and oblate approximations. No asset pricing model was assumed; accuracy was measured by worst-case approximation error or average squared approximation error across the scenario set. These new approximations are the first to have verifiable optimality properties for portfolios of bonds, interest rate swaps, credit default swaps, and equities.

The numerically-optimal approximation achieves, to within rounding error, a lower bound that I derived and computed for worst-case approximation error. It also possesses a very helpful interpretation in terms of an approximating portfolio of cashflow streams. The interpolatory approximation has a similar interpretation, but achieves optimal root-mean-square error. The Mathieu approximation is nearly optimal for worst-case error and has a useful nested structure, and the oblate approximation is optimal for root-mean-square error with a nested structure.

The approximations introduced here were shown be to greatly superior to Taylor-series approximations, which are widely used for portfolio approximation.

References

Abramowitz, Milton, & Stegun, Irene A. (eds). 1964. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* National Bureau of Standards Applied Mathematics Series, vol. 55. Washington, D. C.: U. S. Government Printing Office. Sixth Printing, November 1967, with corrections.

Achieser, N. I. 1956. *Theory of Approximation.* New York: Frederick Ungar Publishing Co. Dover edition first published in 1992, an unabridged and unaltered republication of the 1956 edition.

Akhiezer, N. I, & Glazman, I. M. 1961. *Theory of Linear Operators in Hilbert Space, Volume I.* New York: Frederick Ungar Publishing Co. Dover edition first published in 1993, an unabridged and unaltered republication of the 1961 edition, bound with volume II.

Akhiezer, N. I, & Glazman, I. M. 1963. *Theory of Linear Operators in Hilbert Space, Volume II.* New York: Frederick Ungar Publishing Co. Dover edition first published in 1993, an unabridged and unaltered republication of the 1963 edition, bound with volume I.

Anderson, Evan W., Hansen, Lars Peter, & Sargent, Thomas J. 2003. A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection. *Journal of the European Economic Association*, **1**, 68–123.

Arscott, F. M. 1964. *Periodic Differential Equations: An Introduction to Mathieu, Lamé, and Allied Functions.* International Series of Monographs in Pure and Applied Mathematics, vol. 66. New York: Pergamon Press, The Macmillan Co.

Baraniuk, Richard, Davenport, Mark, DeVore, Ronald, & Wakin, Michael. 2008. A Simple Proof of the Restricted Isometry Property for Random Matrices. *Constructive Approximation*, **28**, 253–263.

Barndorff-Nielsen, Ole E., & Shephard, Neil. 2001. Non-Gaussian Ornstein-Uhlenbeck-based Models and Some of Their Uses in Financial Economics. *Journal of the Royal Statistical Society: Series B*, **63**, 167–241.

Beatson, Rick, & Greengard, Leslie. 1997. A Short Course on Fast Multipole Methods. *Pages 1–37 of:* Ainsworth, M., Levesley, J., Light, W. A., & Marletta, M. (eds), *Wavelets, Multilevel Methods and Elliptic PDEs*. Numerical Mathematics and Scientific Computation. Oxford: Oxford University Press.

Berge, Claude. 1963. *Topological Spaces, Including a Treatment of Multi-valued Functions, Vector Spaces and Convexity*. Edinburgh: Oliver and Boyd.

Bernoulli, Jacob. 1690. Qustiones nonnull de usuris, cum solutione Problematis de Sorte Alearum, propositi in Ephem. Gall. A. 1685, artic. 25. *Acta Eruditorum*, 219–223.

Bertero, M., & Grünbaum, F. A. 1985. Commuting Differential Operators for the Finite Laplace Transform. *Inverse Problems*, **1**, 181–192.

Bertero, M., Grünbaum, F. A., & Rebolia, L. 1986. Spectral Properties of a Differential Operator Related to the Inversion of the Finite Laplace Transform. *Inverse Problems*, **2**, 131–139.

Bornemann, Folkmar, Laurie, Dirk, Wagon, Stan, & Waldvogel, Jörg. 2004. *The SIAM 100-Digit Challenge: A Study in High-Accuracy Numerical Computing*. Other Titles in Applied Mathematics, vol. 86. Philadelphia: Society for Industrial and Applied Mathematics. With a foreword by David H. Bailey.

Borovička, Jaroslav, Hansen, Lars Peter, & Scheinkman, José A. 2014. Misspecified Recovery. *working paper*.

Borsuk, K. 1933. Drei Sätze über die n-dimensionale euklidische Sphäre. *Fundamenta Mathematicae*, **20**, 177–191.

Bouwkamp, C. J. 1947. On Spheroidal Wave Functions of Order Zero. *Journal of Mathematics and Physics*, **26**, 79–92.

Bunch, James R., Nielsen, Christopher P., & Sorensen, Danny C. 1978. Rank-One Modification of the Symmetric Eigenproblem. *Numerische Mathematik*, **31**, 31–48.

Campbell, John Y., Sunderam, Adi, & Viceira, Luis M. 2013. Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds. *working paper*.

Campbell, John Y., Pflueger, Carolin, & Viceira, Luis M. 2014. Monetary Policy Drivers of Bond and Equity Risks. *NBER Working Paper*, **20070**.

Candès, Emmanuel, & Tao, Terence. 2006. Near Optimal Signal Recovery from Random Projections: Universal Encoding Strategies? *IEEE Transactions on Information Theory*, **52**, 5406–5425.

Chamberlain, Gary. 2000. Econometric Applications of Maxmin Expected Utility. *Journal of Applied Econometrics*, **15**, 625–644.

Chamberlain, Gary. 2001. Minimax Estimation and Forecasting in a Stationary Autoregression Model. *American Economic Review*, **91**, 55–59.

Chambers, Donald R., Carleton, Willard T., & McEnally, Richard W. 1988. Immunizing Default-Free Bond Portfolios with a Duration Vector. *The Journal of Financial and Quantitative Analysis*, **23**, 89–104.

Cheney, E. W. 1982. *Introduction to Approximation Theory.* Second edn. Providence: AMS Chelsea Publishing. Reprinted by the American Mathematical Society, 1998.

Cipra, Barry A. 2000. The Best of the 20th Century: Editors Name Top 10 Algorithms. *SIAM News*, **33**, 1–2.

Clenshaw, C. W. 1955. A Note on the Summation of Chebyshev Series. *Mathematical Tables and other Aids to Computation*, **9**, 118–120.

Cox, John C., Jonathan E. Ingersoll, Jr., & Ross, Stephen A. 1979. Duration and the Measurement of Basis Risk. *The Journal of Business*, **52**, 51–61.

Cox, John C., Jonathan E. Ingersoll, Jr., & Ross, Stephen A. 1985. A Theory of the Term Structure of Interest Rates. *Econometrica*, **53**, 385–407.

Delves, L. M., & Mohamed, J. L. 1984. *Computational Methods for Integral Equations.* Cambridge: Cambridge University Press.

Donoho, David L. 2006. Compressed Sensing. *IEEE Transactions on Information Theory*, **52**, 1289–1306.

Duffie, Darrell, Pan, Jun, & Singleton, Kenneth. 2000. Transform Analysis and Asset Pricing for Affine Jump-diffusions. *Econometrica*, **68**, 1343–1376.

Duistermaat, J. J., & Grünbaum, F. A. 1986. Differential Equations in the Spectral Parameter. *Communications in Mathematical Physics*, **103**, 177–240.

Erdélyi, Arthur, Magnus, Wilhelm, Oberhettinger, Fritz, & Tricomi, Francesco G. 1955. *Higher Transcendental Functions. Vol. III.* New York: McGraw-Hill Book Company, Inc. Reprinted by Robert E. Krieger Publishing Co. Inc., 1981. Table errata: Math. Comp. v. 41 (1983), no. 164, p. 778.

Fallat, Shaun M., & Johnson, Charles R. 2011. *Totally Nonnegative Matrices*. Princeton Series in Applied Mathematics. Princeton: Princeton University Press.

Fisher, Lawrence, & Weil, Roman L. 1971. Coping with the Risk of Interest-Rate Fluctuations: Returns to Bondholders from Naïve and Optimal Strategies. *The Journal of Business*, **44**, 408–431.

Flammer, Carson. 1957. *Spheroidal Wave Functions*. Stanford: Stanford University Press. Dover Phoenix Edition first published in 2005, an unabridged republication of the 1957 edition.

Fong, H. Gifford, & Vasicek, Oldrich A. 1984. A Risk Minimizing Strategy for Portfolio Immunization. *The Journal of Finance*, **39**, 1541–1546.

Fuchs, W. H. J. 1964. On the Eigenvalues of an Integral Equation Arising in the Theory of Band-Limited Signals. *Journal of Mathematical Analysis and Applications*, **9**, 317–330.

Gantmacher, Feliks R., & Krein, Mark G. 2002. *Oscillation Matrices and Kernels and Small Vibrations of Mechanical Systems*. Revised edn. Providence: American Mathematical Society.

Gautschi, Walter. 1967. Computational Aspects of Three-term Recurrence Relations. *SIAM Review*, **9**, 24–82.

Gautschi, Walter. 1993. Is the Recurrence Relation for Orthogonal Polynomials Always Stable? *BIT Numerical Mathematics*, **33**, 277–284.

Gautschi, Walter. 2004. *Orthogonal Polynomials: Computation and Approximation*. Numerical Mathematics and Scientific Computation. Providence: Oxford University Press.

Goetzmann, William N. 2005. Fibonacci and the Financial Revolution. *Pages 123–143 of:* Goetzmann, William N., & Rouwenhorst, K. Geert (eds), *The Origins of Value: The Financial Innovations That Created Modern Capital Markets*. New York: Oxford University Press.

Gohberg, I. C., & Krein, M. G. 1969. *Introduction to the Theory of Linear Nonselfadjoint Operators*. Translations of Mathematical Monographs, vol. 18. Providence: American Mathematical Society.

Golub, Gene H., & Van Loan, Charles F. 2013. *Matrix Computations*. Fourth edn. Johns Hopkins Studies in the Mathematical Sciences. Baltimore: The Johns Hopkins University Press.

Gradshteyn, I. S., & Ryzhik, I. M. 2000. *Table of Integrals, Series, and Products*. Sixth edn. New York: Academic Press.

Greengard, L., & Rokhlin, V. 1987. A Fast Algorithm for Particle Simulations. *Journal of Computational Physics*, **73**, 325–348.

Greenwood, Robin, & Vayanos, Dimitri. 2014. Bond Supply and Excess Bond Returns. *Review of Financial Studies*, **27**, 663–713.

Grünbaum, F. Alberto. 1981. Toeplitz Matrices Commuting with Tridiagonal Matrices. *Linear Algebra and Its Applications*, **40**, 25–36.

Grünbaum, F. Alberto. 1982. A Remark on Hilbert's Matrix. *Linear Algebra and Its Applications*, **43**, 119–124.

Grünbaum, F. Alberto. 1983. Differential Operators Commuting with Convolution Integral Operators. *Journal of Mathematical Analysis and Applications*, **91**, 80–93.

Haar, Alfréd. 1917. Die Minkowskische Geometrie und die Annäherung an stetige Funktionen. *Mathematishe Annalen*, **78**, 294–311.

Hansen, Eldon, & Walster, G. William. 2004. *Global Optimization Using Interval Analysis: Second Edition, Revised and Expanded*. Pure and Applied Mathematics: A Dekker Series of Monographs and Textbooks, vol. 264. New York: Marcel Dekker, Inc.

Hansen, Lars Peter. 2012. Dynamic Valuation Decomposition within Stochastic Economies. *Econometrica*, **80**, 911–967.

Hansen, Lars Peter, & Sargent, Thomas J. 1995. Discounted Linear Exponential Quadratic Gaussian Control. *IEEE Transactions on Automatic Control*, **40**, 968–971.

Hansen, Lars Peter, & Sargent, Thomas J. 2008. *Robustness*. Princeton, NJ: Princeton University Press.

Hansen, Lars Peter, & Sargent, Thomas J. 2012. Three Types of Ambiguity. *Journal of Monetary Economics*, **59**, 422–445.

Hansen, Lars Peter, & Scheinkman, José A. 2009. Long Term Risk: An Operator Approach. *Econometrica*, **77**, 177–234.

Hansen, Lars Peter, Sargent, Thomas J., & Thomas D. Tallarini, Jr. 1999. Robust Permanent Income and Pricing. *Review of Economic Studies*, **66**, 873–907.

Hansen, Lars Peter, Sargent, Thomas J., Turmuhambetova, Gauhar, & Williams, Noah. 2006. Robust Control and Model Misspecification. *Journal of Economic Theory*, **128**, 45–90.

Heath, David, Jarrow, Robert, & Morton, Andrew. 1992. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, **60**, 77–105.

Hicks, John R. 1939. *Value and Capital*. Oxford: Clarendon Press.

Ho, Thomas S. Y. 1992. Key Rate Durations: Measures of Interest Rate Risks. *The Journal of Fixed Income*, **2**, 29–44.

Hodge, D. B. 1970. Eigenvalues and Eigenfunctions of the Spheroidal Wave Equation. *Journal of Mathematical Physics*, **11**, 2308–2312.

Hogan, Jeffrey A., & Lakey, Joseph D. 2012. *Duration and Bandwidth Limiting: Prolate Functions, Sampling, and Applications.* New York: Birkhäuser.

Ingersoll, Jonathan E., Jr., Skelton, Jeffrey, & Weil, Roman L. 1978. Duration Forty Years Later. *The Journal of Financial and Quantitative Analysis*, **13**, 627–650.

Jaulin, Luc, Kieffer, Michel, Didrit, Olivier, & Walter, Éric. 2001. *Applied Interval Analysis.* London: Springer.

Jeffrey, Andrew. 2000. Duration, Convexity and Higher Order Hedging (Revisited). *Yale School of Management Working Paper.*

Karlin, Samuel. 1963. Representation Theorems for Positive Functions. *Journal of Mathematics and Mechanics*, **12**, 599–618.

Karlin, Samuel. 1968. *Total Positivity, Volume I.* Stanford: Stanford University Press.

Karlin, Samuel, & Studden, William J. 1966. *Tchebycheff Systems: With Applications in Analysis and Statistics.* Pure and Applied Mathematics, vol. XV. New York: John Wiley & Sons.

Keener, Lee L. 1993. The Snake Theorem for Unisolvent Families. *Journal of Approximation Theory*, **74**, 110–121.

Kellogg, O. D. 1916. The Oscillation of Functions of an Orthogonal Set. *American Journal of Mathematics*, **38**, 1–5.

Kellogg, O. D. 1918. Orthogonal Function Sets Arising from Integral Equations. *American Journal of Mathematics*, **40**, 145–154.

Kolmogorov, A. 1936. Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse. *Annals of Mathematics*, **37**, 107–110.

Korneichuk, N. 1991. *Exact Constants in Approximation Theory.* Encyclopedia of Mathematics and Its Applications, vol. 38. Cambridge: Cambridge University Press.

Krein, M. G., Krasnosel'ski, M. A., & Milman, D. P. 1948. On Deficiency Numbers of Linear Operators in Banach Spaces and on Some Geometric Problems. *Sb. Trudov Inst. Mat. Akad. Nauk SSSR*, **11**, 97–112.

Kreĭn, M. G., & Nudel'man, A. A. 1977. *The Markov Moment Problem and Extremal Problems: Ideas and Problems of P. L. Čebyšev and A. A. Markov and Their Further Development*. Translations of Mathematical Monographs, vol. 50. Providence: American Mathematical Society.

Landau, H. J., & Pollak, H. O. 1961. Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty – II. *The Bell System Technical Journal*, **40**, 65–84.

Landau, H. J., & Pollak, H. O. 1962. Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty – III: The Dimension of the Space of Essentially Time- and Band-Limited Signals. *The Bell System Technical Journal*, **41**, 1295–1336.

Landau, H. J., & Widom, H. 1980. Eigenvalue Distribution of Time and Frequency Limiting. *Journal of Mathematical Analysis and Applications*, **77**, 469–481.

Lax, Peter D. 2002. *Functional Analysis*. New York: Wiley-Interscience.

Lorentz, G. G. 1986. *Approximation of Functions*. Second edn. Providence: AMS Chelsea Publishing. Reprinted by the American Mathematical Society, 2005.

Lorentz, George G., von Golitschek, Manfred, & Makovoz, Yuly. 1996. *Constructive Approximation: Advanced Problems*. Grundlehren der mathematischen Wissenschaften. A Series of Comprehensive Studies in Mathematics, vol. 304. Berlin: Springer-Verlag.

Macaulay, Frederick R. 1938. *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856*. New York: Columbia University Press for the National Bureau of Economic Research. Reprinted by Risk Books as part of the Risk Classics Library, 1999.

Mason, J. C., & Handscomb, D. C. 2003. *Chebyshev Polynomials*. Boca Raton: Chapman & Hall/CRC.

McLachlan, N. W. 1947. *Theory and Application of Mathieu Functions*. Oxford: Clarendon Press. Reprinted lithographically at the University Press, Oxford, 1951 from corrected sheets of the first edition.

Meinardus, Günter. 1967. *Approximation of Functions: Theory and Numerical Methods*. Springer Tracts in Natural Philosophy, vol. 13. New York: Springer-Verlag New York Inc.

Meixner, Josef, & Schäfke, Friedrich Wilhelm. 1954. *Mathieusche Funktionen und Sphäroidfunktionen mit Anwendungen auf physicalische und technische Probleme*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Band LXXI. Berlin: Springer-Verlag.

Meixner, Josef, Schäfke, Friedrich W., & Wolf, Gerhard. 1980. *Mathieu Functions and Spheroidal Functions and their Mathematical Foundations: Further Studies.* Lecture Notes in Mathematics, vol. 837. Berlin: Springer-Verlag.

Melkman, Avraham A. 1977. n-Widths and Optimal Interpolation of Time- and Band-limited Functions. *Pages 55–68 of:* Micchelli, Charles A., & Rivlin, Theodore J. (eds), *Optimal Estimation in Approximation Theory.* The IBM Research Symposia Series. New York: Springer Science+Business Media, LLC.

Melkman, Avraham A. 1985. n-Widths and Optimal Interpolation of Time- and Band-limited Functions II. *SIAM Journal of Mathematical Analysis*, **16**, 803–813.

Melkman, Avraham A., & Micchelli, Charles A. 1978. Spline Spaces Are Optimal for L^2 n-Width. *Illinois Journal of Mathematics*, **22**, 541–564.

Micchelli, C. A., & Pinkus, A. 1977a. Best Mean Approximation to a 2-Dimensional Kernel by Tensor Products. *Bulletin of the American Mathematical Society*, **83**, 400–402.

Micchelli, Charles A., & Pinkus, Allan. 1977b. On n-Widths in L^∞. *Transactions of the American Mathematical Society*, **234**, 139–174.

Micchelli, Charles A., & Pinkus, Allan. 1977c. Total Positivity and the Exact n-Width of Certain Sets in L^1. *Pacific Journal of Mathematics*, **71**, 499–515.

Micchelli, Charles A., & Pinkus, Allan. 1978. Some Problems in the Approximation of Functions of Two Variables and n-Widths of Integral Operators. *Journal of Approximation Theory*, **24**, 51–77.

Micchelli, Charles A., & Pinkus, Allan. 1979. The n-Widths of Rank $n+1$ Kernels. *Journal of Integral Equations*, **1**, 111–130.

Moore, Ramon E. 1979. *Methods and Applications of Interval Analysis.* Studies in Applied Mathematics, vol. 2. Philadelphia: Society for Industrial and Applied Mathematics.

Moore, Ramon E., Kearfott, R. Baker, & Cloud, Michael J. 2009. *Introduction to Interval Analysis.* Other Titles in Applied Mathematics, vol. 110. Philadelphia: Society for Industrial and Applied Mathematics.

Morse, Philip M., & Feshbach, Herman. 1953. *Methods of Theoretical Physics.* International Series in Pure and Applied Physics, no. Part II: Chapters 9 to 13. New York: McGraw-Hill Book Company, Inc.

Nawalkha, Sanjay K., & Chambers, Donald R. 1996. An Improved Immunization Strategy: M-Absolute. *Financial Analysts Journal*, **52**, 69–76.

Olver, F. W. J., Lozier, D. W., Boisvert, R. F., & Clark, C. W. (eds). 2010. *NIST Handbook of Mathematical Functions*. New York: Cambridge University Press. Print companion to DLMF (n.d.).

Osipov, Andrei, Rokhlin, Vladimir, & Xiao, Hong. 2013. *Prolate Spheroidal Wave Functions of Order Zero: Mathematical Tools for Bandlimited Approximation*. New York: Springer.

Pachón, Ricardo, & Trefethen, Lloyd N. 2009. Barycentric-Remez Algorithms for Best Polynomial Approximation in the chebfun System. *BIT Numerical Mathematics*, **49**, 721–741.

Piazzesi, Monika. 2010. Affine Term Structure Models. *In:* Aït-Sahalia, Yacine, & Hansen, Lars Peter (eds), *Handbook of Finance*. Oxford: North-Holland.

Pinkus, A., & Strauss, H. 1988. Best Approximation with Coefficient Constraints. *IMA Journal of Numerical Analysis*, **8**, 1–22.

Pinkus, Allan. 1976. Applications of Representation Theorems to Problems of Chebyshev Approximation with Constraints. *In:* Karlin, S., Micchelli, C., Pinkus, A., & Schoenberg, I. J. (eds), *Studies in Spline Functions and Approximation Theory*. New York: Academic Press.

Pinkus, Allan. 1985a. *n-Widths in Approximation Theory*. Ergebnisse der Mathematik und ihrer Grenzgebiete 3. Folge · Band 7. A Series of Modern Surveys in Mathematics. Berlin: Springer-Verlag.

Pinkus, Allan. 1985b. n-Widths of Sobolev Spaces in L^p. *Constructive Approximation*, **1**, 15–62.

Pinkus, Allan. 1996. Spectral Properties of Totally Positive Kernels and Matrices. *Pages 477–511 of:* Gasca, Mariano, & Micchelli, Charles A. (eds), *Total Positivity and Its Applications*. Dordrecht: Kluwer Academic Publishers.

Pinkus, Allan. 2010. *Totally Positive Matrices*. Cambridge Tracts in Mathematics, vol. 181. Cambridge: Cambridge University Press.

Powell, M. J. D. 1981. *Approximation Theory and Methods*. New York: Cambridge University Press. Reprinted 1996.

Redington, F. M. 1952. Review of the Principles of Life-office Valuations. *Journal of the Institute of Actuaries*, **78**, 286–340.

Reed, Michael, & Simon, Barry. 1975. *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-adjointness*. New York: Academic Press.

Reed, Michael, & Simon, Barry. 1980. *Methods of Modern Mathematical Physics I: Functional Analysis*. Revised and enlarged edn. New York: Academic Press.

DLMF. *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.8 of 2014-04-25. Online companion to Olver *et al.* (2010).

Riesz, Frigyes, & Szökefalvi-Nagy, Béla. 1955. *Functional Analysis.* New York: Frederick Ungar Publishing Co. Dover edition first published in 1990, an unabridged republication of the 1955 edition with the Appendix, Extensions of Linear Transformations in Hilbert Space Which Extend Beyond This Space, which was separately published by Ungar in 1960, added.

Rivlin, Theodore J. 1969. *An Introduction to the Approximation of Functions.* Waltham, Massachusetts: Blaisdell Publishing Company. Dover edition first published in 1981, an unabridged and corrected republication of the 1969 edition.

Rivlin, Theodore J. 1990. *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory.* Second edn. Pure and Applied Mathematics. New York: John Wiley & Sons.

Rokhlin, Vladimir. 1988. A Fast Algorithm for the Discrete Laplace Transform. *Journal of Complexity*, **4**, 12–32.

Rokhlin, Vladimir. 1995. Analysis-Based Fast Numerical Algorithms of Applied Mathematics. *Pages 1460–1467 of:* Chatterji, S. D. (ed), *Proceedings of the International Congress of Mathematicians, Zürich, Switzerland 1994, Volume II.* Basel: Birkhäuser Verlag.

Ross, Steve. 2014. The Recovery Theorem. *Journal of Finance*, **forthcoming**.

Rump, Siegfried M. 1999. INTLAB–INTerval LABoratory. *Pages 77–104 of:* Csendes, Tibor (ed), *Developments in Reliable Computing: Papers presented at the International Symposium on Scientific Computing, Computer Arithmetic, and Validated Numerics, SCAN-98, in Szeged, Hungary, Reliable Computing 5(3).* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Samuelson, Paul A. 1945. The Effect of Interest Rate Increases on the Banking System. *The American Economic Review*, **35**, 16–27.

Schoenberg, Isac. 1930. Über variationsvermindernde lineare Transformationen. *Mathematische Zeitschrift*, **32**, 321–328.

Singer, Ivan. 1970. *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces.* Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Band 171. Berlin: Springer-Verlag.

Slepian, D., & Pollak, H. O. 1961. Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty – I. *The Bell System Technical Journal*, **40**, 43–63.

Slepian, David. 1983. Some Comments on Fourier Analysis, Uncertainty and Modeling. *SIAM Review*, **25**, 379–393.

Steffens, Karl-Georg. 2006. *The History of Approximation Theory: From Euler to Bernstein.* Boston: Birkhäuser.

Strain, John. 1992. A Fast Laplace Transform Based on Laguerre Functions. *Mathematics of Computation*, **58**, 275–283.

Stratton, J. A., Morse, P. M., Chu, L. J., Little, J. D. C., & Corbat'o, F. J. 1956. *Spheroidal Wave Functions*. New York: The Technology Press of M. I. T. and John Wiley & Sons, Inc.

Stroock, Daniel W. 1999. *A Concise Introduction to the Theory of Integration*. Third edn. Boston: Birkhäuser.

Szegö, Gabor. 1975. *Orthogonal Polynomials*. Fourth edn. American Mathematical Society Colloquium Publications, vol. 23. Providence: American Mathematical Society.

Tikhomirov, V. M. 1960. Diameters of Sets in Function Spaces and the Theory of Best Approximations. *Russian Mathematical Surveys*, **15**, 75–111.

Tikhomirov, V. M. 1969. Best Methods of Approximation and Interpolation of Differentiable Functions in the Space $C[-1, 1]$. *Mathematics of the USSR Sbornik*, **9**, 275–289.

Tikhomirov, V. M. 1990. Approximation Theory. Encyclopaedia of Mathematical Sciences, vol. 14. Berlin: Springer-Verlag.

Traub, J. F., & Woźniakowski, H. 1980. *A General Theory of Optimal Algorithms*. ACM Monograph Series. New York: Academic Press.

Trefethen, Lloyd N. 2013. *Approximation Theory and Approximation Practice*. Other Titles in Applied Mathematics, vol. 128. Philadelphia: Society for Industrial and Applied Mathematics.

Tyrtyshnikov, Evgenij E. 1994. How Bad Are Hankel Matrices? *Numerische Mathematik*, **67**, 261–269.

Vasicek, Oldrich A. 1977. An Equilibrium Characterization of the Term Structure. *The Journal of Financial Economics*, **5**, 177–188.

Veidinger, L. 1960. On the Numerical Determination of the Best Approximations in the Chebyshev Sense. *Numerische Mathematik*, **2**, 99–105.

Weil, Roman L. 1973. Macaulay's Duration: An Appreciation. *The Journal of Business*, **46**, 589–592.

Zayed, Ahmed I. 2007. A Generalization of the Prolate Spheroidal Wave Functions. *Proceedings of the American Mathematical Society*, **135**, 2193–2203.

Zettl, Anton. 2005. *Sturm-Liouville Theory*. Mathematical Surveys and Monographs, vol. 121. Providence: American Mathematical Society. Reprinted in softcover by the American Mathematical Society, 2010.

Zhang, Shanjie, & Jin, Jianming. 1996. *Computation of Special Functions*. New York: John Wiley & Sons, Inc.

Zielke, Roland. 1979. *Discontinuous Čebyšev Systems*. Lecture Notes in Mathematics, vol. 707. Berlin: Springer-Verlag.