An Exponential Approximation to the Hockey Stick Function*

Ian Iscoe† Ken Jackson‡ Alex Kreinin§ Xiaofang Ma¶

March 19, 2010

Abstract

The hockey stick (HS) function plays an important role in pricing and risk management of many financial derivatives. This paper considers approximating the HS function by a sum of exponentials. This enables the efficient computation of an approximation to the expected value of the HS function applied to a sum of conditionally independent nonnegative random variables, a task that arises in pricing many financial derivatives, CDOs in particular. The algorithm proposed by Beylkin and Monzón is used to determine the parameters of the exponential approximation to the hockey stick function. Theoretical properties of the approximation are studied and numerical results are presented.

*This research was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

†Algorithmics Inc., 185 Spadina Ave., Toronto, ON, M5T 2C6, Canada; Ian.Iscoe@algorithmics.com

‡Computer Science Dept., University of Toronto, 10 King's College Rd., Toronto, ON, M5S 3G4, Canada; krj@cs.utoronto.ca

§Algorithmics Inc., 185 Spadina Ave., Toronto, ON, M5T 2C6, Canada; Alex.Kreinin@algorithmics.com

¶Market Risk, Bank of Montreal, 51st Floor, 1 First Canadian Place, 100 King St. W., Toronto, ON, M5X 1A1, Canada; Xiaofang.Ma@bmo.com

1 Introduction

In this paper, we describe how to approximate the *hockey stick* (HS) function

$$h(x) = \begin{cases} 1 - x & \text{if } 0 \le x \le 1; \\ 0 & \text{if } x > 1, \end{cases} \tag{1}$$

by a sum of exponentials

$$h_{\exp}(x) = \sum_{n=1}^{N} \omega_n \exp(\gamma_n x) \tag{2}$$

over $\mathbb{X} = [0, \infty)$, where ω_n and γ_n are complex numbers.

The function

$$h(x;t) = \begin{cases} t - x & \text{if } 0 \le x \le t; \\ 0 & \text{if } x > t, \end{cases}$$

where t is a positive number, is often also called the hockey stick function. The two functions, $h(x)$ and $h(x;t)$, are closely related. Obviously, $h(x) = h(x;1)$. On the other hand, for a fixed positive t, $h(x;t) = t \cdot h(x/t)$. Therefore, we can take $h(x)$ as the basic function. Consequently, we focus our attention on $h(x)$ throughout this paper.

The HS function is ubiquitous in financial engineering. For example, in the valuation of synthetic collateralized debt obligations [7], the central problem is how to compute efficiently the conditional expectation $\mathbb{E}[h(X)]$, where $\mathbb{E}$ denotes the expected value calculated under a given probability measure, $X = \sum_{k=1}^{K} X_k$ and X_k are conditionally independent nonnegative random variables. Due to the piecewise linearity of h, it is clear that once the distribution of X is known, the expected value $\mathbb{E}[h(X)]$ can be readily computed. All the methods described in [9] follow this approach. When the distribution function of X is computationally expensive to obtain from the distributions of the random variables X_k, this approach may be inefficient. However, if $h(x)$ can be approximated by a sum of exponentials as in (2), then $\mathbb{E}[h(X)] \approx \mathbb{E}[h_{\exp}(X)]$ and

$$\mathbb{E}[h_{\exp}(X)] = \mathbb{E}\left[\sum_{n=1}^{N} \omega_n \exp(\gamma_n X)\right] = \sum_{n=1}^{N} \omega_n \mathbb{E}\left[\exp\left(\gamma_n \sum_{k=1}^{K} X_k\right)\right] = \sum_{n=1}^{N} \omega_n \prod_{k=1}^{K} \mathbb{E}\left[\exp\left(\gamma_n X_k\right)\right],$$

where the last equality follows from the independence of the X_k's. This shows that to approximate $\mathbb{E}[h(X)]$, we need only to compute $\mathbb{E}[\exp(\gamma_n X_k)]$ for each of the K conditionally independent random variables, X_k. Thus, the second approach avoids the computation of the distribution of $X = \sum_{k=1}^{K} X_k$ and may be significantly more efficient than the first approach in some cases, particularly if $\mathbb{E}[\exp(\gamma_n X_k)]$ is known in closed form, as is often the case in practice. For a more complete discussion of the second approach, see [7].

The approximation problem considered here is in the sense of Chebyshev approximation. For such an approximation, the weights ω_n and the exponents γ_n should be chosen to solve the minimization problem

$$\min_{\omega_n, \gamma_n \in \mathbb{C}} \left\| h(x) - \sum_{n=1}^{N} \omega_n \exp(\gamma_n x) \right\|_{\infty}, \tag{3}$$

where $\mathbb{C}$ denotes the set of complex numbers, and $\|f\|_{\infty} = \sup_{x \in \mathbb{X}} |f(x)|$ is the Chebyshev norm (also known as the uniform norm or the sup-norm) of f over $\mathbb{X} = [0, \infty)$. Theoretically, the existence of such an optimal approximation is generally not guaranteed [3, Chapters VI and VII]. Classic numerical methods for linear Chebyshev approximations, such as the Remez exchange algorithm and its improvements, do not work well for solving nonlinear Chebyshev approximation problems such as (3) [10]. Most algorithms for nonlinear Chebyshev approximations resort to solving discrete Chebyshev approximation subproblems. For exponential approximation problems, such a discrete Chebyshev approximation subproblem is equivalent to an exponential fitting problem, which is often badly-conditioned [4]. Consequently, we must find a method that works well for (3). In this paper, we apply the method recently proposed by Beylkin and Monzón [2] to determine the coefficients ω_n and γ_n in (3).

The remainder of the paper is organized as follows. Beylkin and Monzón's method and its application to the HS function are discussed in Section 2. Properties of the exponential approximation are discussed in Section 3. The paper ends with numerical results in Section 4.

2 Beylkin and Monzón's method and its application to the HS function

2.1 Beylkin and Monzón's method

In a recent paper [2], Beylkin and Monzón proposed an effective numerical method to find a good exponential approximation to a function f. Instead of finding optimal ω_n and γ_n satisfying (3), their method finds such parameters so that the exponential approximation satisfies a given accuracy requirement. More specifically, for a given function f defined on $[0, 1]$ and a given ϵ, their method seeks to find the minimal (or nearly minimal) number of complex weights ω_n and nodes $\exp(\gamma_n)$ such that

$$\left| f(x) - \sum_{n=1}^{N} \omega_n \exp(\gamma_n x) \right| \le \epsilon, \qquad \forall x \in [0, 1]. \tag{4}$$

This continuous problem is in turn approximated by a discrete problem: Given a positive integer $\mathcal{M}$, find the minimal positive integer number $N \le \mathcal{M}$ of complex weights ω_n and complex nodes ζ_n such that

$$\left| f\left(\frac{m}{2\mathcal{M}}\right) - \sum_{n=1}^{N} \omega_n \zeta_n^m \right| \le \epsilon, \qquad \text{for all integers } m \in [0, 2\mathcal{M}]. \tag{5}$$

Then, for the continuous problem, the weights and the exponents are ω_n and

$$\gamma_n = 2\mathcal{M} \log \zeta_n, \tag{6}$$

respectively, where $\log z$ is the principal value of the logarithm.

To describe their method in more detail, we introduce some additional notation. For theoretical background and a more detailed description of the method, see [2].

For a real $(2\mathcal{M} + 1)$-vector $\mathbf{h} = (h_0, h_1, \ldots, h_{2\mathcal{M}})$, the $(\mathcal{M} + 1) \times (\mathcal{M} + 1)$ Hankel matrix

$\mathbf{H_h}$ defined in terms of $\mathbf{h}$ is

$$\mathbf{H_h} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{\mathcal{M}} \\ h_1 & \cdots & \cdots & h_{\mathcal{M}+1} \\ \vdots & & \iddots & \vdots \\ h_{\mathcal{M}-1} & h_{\mathcal{M}} & \cdots & h_{2\mathcal{M}-1} \\ h_{\mathcal{M}} & \cdots & h_{2\mathcal{M}-1} & h_{2\mathcal{M}} \end{bmatrix}.$$

That is, $\mathbf{H_h}(i,j) = h_{i+j}$ for $0 \leq i,j \leq \mathcal{M}$. It is clear that $\mathbf{H_h}$ is a real symmetric matrix. By the Corollary in §4.4.4 of [6, pp. 204], there exists a unitary matrix $\mathbf{U}$ and a *nonnegative* diagonal matrix Σ such that

$$\mathbf{H_h} = \mathbf{U}\Sigma\mathbf{U}^T,$$

where the superscript T denotes transposition. This decomposition is called the Takagi factorization [6, pp. 204].

The main steps of Beylkin and Monzón's method are:

1. Sample the approximated function f at $2\mathcal{M}+1$ points uniformly distributed on $[0,1]$. That is, let $h_m = f\left(\frac{m}{2\mathcal{M}}\right)$, $0 \leq m \leq 2\mathcal{M}$.

2. Form $\mathbf{h} = (h_0, h_1, \ldots, h_{2\mathcal{M}})$ and the Hankel matrix $\mathbf{H_h}$.

3. Compute the Takagi factorization of $\mathbf{H_h} = \mathbf{U}\Sigma\mathbf{U}^T$, where $\Sigma = \mathrm{diag}(\sigma_0, \sigma_1, \ldots, \sigma_{\mathcal{M}})$ and $\sigma_0 \geq \sigma_1 \geq \ldots \geq \sigma_{\mathcal{M}} \geq 0$.

4. Find the largest σ_N satisfying $\sigma_N \leq \epsilon$.

5. Let $\mathbf{u} = (u_0, u_1, \ldots, u_{\mathcal{M}})^T$ be the $(N+1)$-st column of $\mathbf{U}$.

6. Find N roots of the polynomial $\sum_{m=0}^{\mathcal{M}} u_m z^m$ with the largest moduli and denote these roots by $\zeta_1, \zeta_2, \ldots, \zeta_N$.

7. Compute the N weights ω_n, $1 \leq n \leq N$, by solving the linear least squares problem for the overdetermined Vandermonde system

$$h_m = \sum_{n=1}^{N} \omega_n \zeta_n^m, \qquad \text{for } 0 \leq m \leq 2\mathcal{M}.$$

8. Compute parameters γ_n using formula (6).

Remark 1 *The algorithm outlined above is applicapable to functions defined on* $[0, 1]$*. To extend it to a function* f *defined on a finite interval* $[a, b]$*, we could apply the algorithm to the function* $\hat{f}(t) = f(t(b - a) + a)$ *for* $t \in [0, 1]$*. For a function* f *defined on an infinite interval, such as* $[0, \infty)$*, the interval could first be truncated to a finite interval, say* $[a, b] \subset [0, \infty)$*, then the approach outlined above could be used to compute an accurate exponential approximation to* f *on* $[a, b]$*. This exponential sum could also be viewed as an approximation to* f *on* $[0, \infty)$*, but one should check that the approximation is sufficiently accurate on* $[0, \infty)\backslash[a, b]$*, since the approach described above does not take the approximation error on* $[0, \infty)\backslash[a, b]$ *into account.*

Remark 2 *In practice it is not necessary to compute* $\mathbf{H_h}$*'s Takagi factorization explicitly. From the spectral theorem for Hermitian matrices [6, pp. 171], we know that there is a* **real** *orthogonal matrix* $\mathbf{V}$ *and a* **real** *diagonal matrix* $\Lambda = \mathrm{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{\mathcal{M}})$*, with* $|\lambda_i|$ *non-increasing, such that* $\mathbf{H_h} = \mathbf{V}\Lambda\mathbf{V}^T$*. Since* $\mathbf{H_h}$ *is not necessarily positive semidefinite,* Λ *may have negative elements. Thus* $\mathbf{V}\Lambda\mathbf{V}^T$ *is not necessarily the Takagi factorization of* $\mathbf{H_h}$*. However, we can easily construct a Takagi factorization based on its spectral decomposition in the following way. Let* $\Sigma = \mathrm{diag}(|\lambda_0|, |\lambda_1|, \ldots, |\lambda_{\mathcal{M}}|)$ *and* $\mathbf{U} = (\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{\mathcal{M}})$*, where* $\mathbf{u}_m = \mathbf{v}_m$ *if* $\lambda_m \geq 0$*; and* $\mathbf{u}_m = \sqrt{-1}\mathbf{v}_m$*, if* $\lambda_m < 0$*. It is easy to check that* $\mathbf{U}$ *is a unitary matrix and* $\mathbf{H_h} = \mathbf{U}\Sigma\mathbf{U}^T$ *is the Takagi factorization of* $\mathbf{H_h}$*.*

Remark 3 *To compute* ω_n *from the linear least squares problem in Step 7, the* N *roots determined in Step 6 must be distinct. If this condition is not met,* ω_n *should be computed by a different method, as discussed in [2, 8]. This condition may be difficult to verify in theory. For numerical solutions, we should check its validity, as suggested by Beylkin and Monzón.*

2.2 Application to the HS function

In this subsection, we apply Beylkin and Monzón's method to the HS function $h(x)$. Recall that $h(x)$ is defined on $[0, \infty)$. Following the approach outlined in Remark 1, the infinite interval is first truncated to a finite interval $[0, b]$ for a sufficiently large b. (In fact, $b = 2$ is sufficiently large, as explained below.) Then $h(b \cdot t)$, $t \in [0, 1]$, is sampled at $2\mathcal{M} + 1$ points:

$$h_m = h\left(bt_m\right) = \begin{cases} 1 - bt_m & \text{if } bt_m \leq 1 \\ 0 & \text{otherwise} \end{cases},$$

where $t_m = \frac{m}{2\mathcal{M}}$ for $m = 0, 1, \ldots, 2\mathcal{M}$. Numerical results suggest that more accurate approximations are obtained if the critical point $x = 1$ of $h(x)$ is included in the sample points $\{bt_m : m = 0, \ldots, 2M\}$. Therefore, in the remainder of this paper, we assume $bt_m = 1$ for some m. This implies that $\frac{2\mathcal{M}}{b}$ must be an integer.

The corresponding Hankel matrix $\mathbf{H_h}$ is

$$\mathbf{H_h} = \left[\begin{array}{ccccc:ccc} 1 & 1-\frac{b}{2\mathcal{M}} & 1-2\frac{b}{2\mathcal{M}} & \cdots & \frac{b}{2\mathcal{M}} & 0 & \cdots & 0 \\ 1-\frac{b}{2\mathcal{M}} & 1-2\frac{b}{2\mathcal{M}} & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 1-2\frac{b}{2\mathcal{M}} & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & & & & & & \\ \frac{b}{2\mathcal{M}} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hdashline \cdots & \cdots & & & & & & \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right].$$

To keep this neat form of $\mathbf{H_h}$, it is sufficient to choose $b \geq 2$. If $b < 2$, the last nonzero row of $\mathbf{H_h}$ may have more than one nonzero element. A direct consequence of this is that the properties of the approximation discussed in Section 3 may not hold. Thus, we assume $b \geq 2$ throughout the remainder of the paper.

Let $\mathcal{N} = \frac{2\mathcal{M}}{b}$ and

$$\mathbf{H}_{\mathcal{N}} = \begin{bmatrix} \mathcal{N} & \mathcal{N}-1 & \mathcal{N}-2 & \cdots & 1 \\ \mathcal{N}-1 & \mathcal{N}-2 & \cdots & \cdots & 0 \\ \mathcal{N}-2 & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & & & \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{7}$$

Then we have

$$\mathbf{H_h} = \frac{b}{2\mathcal{M}} \left[\begin{array}{c:c} \mathbf{H}_{\mathcal{N}} & \mathbf{0}_{12} \\ \hdashline \mathbf{0}_{12}^T & \mathbf{0}_{22} \end{array}\right] \tag{8}$$

where $\mathbf{0}_{12}$ and $\mathbf{0}_{22}$ are zero matrices of the proper dimensions.

Let $\mathbf{U}\Sigma\mathbf{U}^T$ be a Takagi factorization of $\mathbf{H}_{\mathcal{N}}$, where $\Sigma = \mathrm{diag}(\sigma_{\mathcal{N}1}, \sigma_{\mathcal{N}2}, \ldots, \sigma_{\mathcal{N}\mathcal{N}})$ and $\sigma_{\mathcal{N}1} \geq \sigma_{\mathcal{N}2} \geq \cdots \geq \sigma_{\mathcal{N}\mathcal{N}} \geq 0$. Then a Takagi factorization of $\mathbf{H}_{\mathbf{h}}$ can be obtained by

$$\mathbf{H}_{\mathbf{h}} = \frac{b}{2\mathcal{M}} \begin{bmatrix} \mathbf{U} & \mathbf{0}_{12} \\ \mathbf{0}_{12}^T & \mathbf{I}_{22} \end{bmatrix} \begin{bmatrix} \Sigma & \mathbf{0}_{12} \\ \mathbf{0}_{12}^T & \mathbf{0}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{U}^T & \mathbf{0}_{12}^T \\ \mathbf{0}_{12} & \mathbf{I}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{U} & \mathbf{0}_{12} \\ \mathbf{0}_{12}^T & \mathbf{I}_{22} \end{bmatrix} \begin{bmatrix} \widehat{\Sigma} & \mathbf{0}_{12} \\ \mathbf{0}_{12}^T & \mathbf{0}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{U}^T & \mathbf{0}_{12}^T \\ \mathbf{0}_{12} & \mathbf{I}_{22} \end{bmatrix}$$

where $\widehat{\Sigma} = \frac{b}{2\mathcal{M}}\Sigma$.

Remark 4 *Proposition 1 in Section 3, together with Theorems 2 and 3 of [2], implies that, for a given accuracy ϵ in (5), $\mathcal{M}$ must be large enough so that $\frac{1}{4}\frac{b}{2\mathcal{M}} \leq \epsilon$. From this relation and noting that $\mathcal{N} = \frac{2\mathcal{M}}{b}$, we can see that the only requirements are $b \geq 2$, $\frac{2\mathcal{M}}{b}$ is an integer and*

$$\mathcal{N} \geq \frac{1}{4\epsilon}. \tag{9}$$

Thus we choose $b = 2$ for simplicity and $\mathcal{N} = \mathcal{M} \geq \frac{1}{4\epsilon}$.

Thus, for $\mathcal{N} = N + 1$, Beylkin and Monzón's method described above for computing an exponential approximation to a general function $f(x)$ reduces to the following algorithm for determining the coefficients ω_n and γ_n for an accurate exponential approximation $h_{\exp}(x)$ of the form (2) to the HS function $h(x)$.

1. Input ϵ, the required accuracy.

2. Find the smallest integer $\mathcal{N}$ such that $\mathcal{N} \geq \frac{1}{4\epsilon}$.

3. Compute the spectral decomposition of the matrix $\mathbf{H}_{\mathcal{N}} = \mathbf{V}\Lambda\mathbf{V}^T$.

4. Let $\mathbf{u} = (u_0, u_1, \ldots, u_{\mathcal{N}-1})^T$ be the last column of $\mathbf{V}$.

5. Find all roots $\zeta_1, \zeta_2, \ldots, \zeta_{\mathcal{N}-1}$ of the polynomial $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$ and check whether they are distinct. If they are not distinct, then exit.

6. Solve $h_m = \sum_{n=1}^{\mathcal{N}-1} \omega_n \zeta_n^m$, $0 \leq m \leq 2\mathcal{N}$, in the least squares sense for ω_n.

7. Compute $\gamma_n = 2\mathcal{N} \log \zeta_n$.

Before ending this section, we want to say a little about ω_n and γ_n. As mentioned in Remark 2, $\mathbf{u}_\mathcal{N}$ is either a real vector or the product of a real vector and the imaginary unit, $\sqrt{-1}$. In either case, the roots of $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$ are either real or pairwise conjugate. Thus, the corresponding ω_n are also real or pairwise complex conjugate. That is, ζ_n is real if and only if ω_n is real; ζ_i and ζ_j are conjugate if and only if ω_i and ω_j are conjugate. Furthermore, since $\gamma_n = 2\mathcal{N} \log \zeta_n$, $\exp(\gamma_n) = \zeta_n^{2\mathcal{N}}$ possesses the same conjugacy property. Thus, $\omega_n \exp(\gamma_n x)$ are either real or pairwise conjugate for all real x. This result simplifies the calculation of $h_{\text{exp}}(x) = \sum_{n=1}^{\mathcal{N}-1} \omega_n \exp(\gamma_n x)$. For real ω_n, the term $\omega_n \exp(\gamma_n x)$ is evaluated as usual, whereas, for the complex conjugate pair indexed by i and j, only one term needs to be evaluated, say $\omega_i \exp(\gamma_i x)$, and then the contribution of the complex conjugate pair of terms is $2\Re(\omega_i \exp(\gamma_i x))$, where $\Re(z)$ denotes the real part of the complex number z.

Another point is that, although Remark 3 discusses the possibility of multiple roots in step 6 of Beylkin and Monzón's original algorithm outlined in subsection 2.1 and similarly in step 5 of our adapted algorithm described above in this subsection, we have not encountered this problem in any of the many numerical examples we have considered.

Finally, in our adapted algorithm described above in this subsection, the coefficients ω_n and γ_n are determined based on the sampled HS function values over $[0, 2]$. However, our intention is to use $h_{\text{exp}}(x)$ to approximate $h(x)$ over the infinite interval $[0, \infty)$. In the next section, we show that $\Re(\gamma_n) < 0$ for all $n = 1, \ldots, \mathcal{N} - 1$. Therefore, $h_{\text{exp}}(x) = \sum_{n=1}^{\mathcal{N}-1} \omega_n \exp(\gamma_n x) \to 0$ as $x \to \infty$. Since $h(x) = 0$ for $x \geq 1$, this ensures that $h_{\text{exp}}(x) \to h(x)$ as $x \to \infty$. Moreover, our adapted algorithm described above in this subsection ensures that

$$|h_{\text{exp}}(n/\mathcal{N})| = |h_{\text{exp}}(n/\mathcal{N}) - h(n/\mathcal{N})| \leq \epsilon \quad \text{for } n = \mathcal{N}, \mathcal{N}+1, \ldots, 2\mathcal{N},$$

where ϵ is the tolerance parameter in (5). This suggests that

$$|h_{\text{exp}}(x)| = |h_{\text{exp}}(x) - h(x)| \lessapprox \epsilon \quad \forall x \in [1, 2].$$

Furthermore, since $h_{\text{exp}}(x) = \sum_{n=1}^{\mathcal{N}-1} \omega_n \exp(\gamma_n x)$ and $\Re(\gamma_n) < 0$ for all $n = 1, \ldots, \mathcal{N} - 1$, it is not unreasonable to expect that

$$\max_{x\in(2,\infty)} |h_{\mathrm{exp}}(x) - h(x)| = \max_{x\in(2,\infty)} |h_{\mathrm{exp}}(x)| \lessapprox \max_{x\in[1,2]} |h_{\mathrm{exp}}(x)| = \max_{x\in[1,2]} |h_{\mathrm{exp}}(x) - h(x)| \lessapprox \epsilon.$$

This does not guarantee that

$$\max_{x\in[0,\infty)} |h_{\mathrm{exp}}(x) - h(x)| \le \epsilon$$

but it does suggest that $h_{\mathrm{exp}}(x)$ should be an accurate approximation to $h(x)$ for all $x \in [0,\infty)$. This is indeed the case for all of the many numerical examples we have considered, a sample of which is shown in Section 4. Nevertheless, as noted in Remark 1, it is advisable to check that $h_{\mathrm{exp}}(x)$ is a sufficiently accurate approximation to $h(x)$ over $(2,\infty)$, since no accuracy requirements are explicitly imposed there.

3 Properties of the approximation

In this section, we discuss some properties of the exponential approximation (2) to the HS function (1).

Since the diagonal matrix Σ appearing in the Takagi factorization of $\mathbf{H}_{\mathcal{N}}$ is the same as the diagonal matrix appearing in $\mathbf{H}_{\mathcal{N}}$'s singular value decomposition, we call $\sigma_{\mathcal{N}n}$, $n = 1, 2, \ldots, \mathcal{N}$, the singular values of $\mathbf{H}_{\mathcal{N}}$.

Direct calculation shows that

$$\mathbf{H}_{\mathcal{N}}^{-1} = \begin{bmatrix} & & & & & 1 \\ & & & & 1 & -2 \\ & & & 1 & -2 & 1 \\ & & \iddots & \iddots & \iddots & \\ & \iddots & \iddots & \iddots & & \\ 1 & -2 & 1 & & & \\ 1 \;\; -2 & 1 & & & & \end{bmatrix} \tag{10}$$

Proposition 1 *As $\mathcal{N} \to \infty$, the smallest singular value $\sigma_{\mathcal{N}\mathcal{N}}$ of the matrix $\mathbf{H}_\mathcal{N}$ tends to $1/4$.*

Proof Since $\mathbf{H}_\mathcal{N}$ is nonsingular, its singular values are all positive. Proving that $\sigma_{\mathcal{N}\mathcal{N}} \to 1/4$ as $\mathcal{N} \to \infty$ is equivalent to proving that $\sigma_{\mathcal{N}\mathcal{N}}^{-1}$, the largest singular value of $\mathbf{H}_\mathcal{N}^{-1}$, tends to 4 as $\mathcal{N} \to \infty$.

From Gerschgorin's theorem [5, pp. 320] [6, pp. 344], we know that all eigenvalues of $\mathbf{H}_\mathcal{N}^{-1}$ lie in the disc

$$D = \{z \in \mathbb{C} : |z| \leq 4\}.$$

Since $\mathbf{H}_\mathcal{N}^{-1}$ is a real symmetric matrix, all its singular values are bounded by 4. To show that $\sigma_{\mathcal{N}\mathcal{N}}^{-1} \to 4$ as $\mathcal{N} \to \infty$, note that, if $\mathbf{A}$ is a real symmetric matrix, then $\max_{\mathbf{x}\neq 0} \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_{\max}$, where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{A}$. Hence, if for each $\mathcal{N}$, we can find a vector $\mathbf{x}_\mathcal{N}$ such that the Rayleigh quotient $\frac{\mathbf{x}_\mathcal{N}^T\mathbf{H}_\mathcal{N}^{-T}\mathbf{H}_\mathcal{N}^{-1}\mathbf{x}_\mathcal{N}}{\mathbf{x}_\mathcal{N}^T\mathbf{x}_\mathcal{N}} \to 16$ as $\mathcal{N} \to \infty$, then we can conclude that $\sigma_{\mathcal{N}\mathcal{N}}^{-1} \to 4$ as $\mathcal{N} \to \infty$.

For even $\mathcal{N} \geq 6$, let $\mathcal{N} = 2n$. Define a vector $\mathbf{x}_\mathcal{N} = (x_1, x_2, \ldots, x_\mathcal{N})^T$ by

$$\begin{aligned}
&x_1 = 1, \\
&x_i = x_{\mathcal{N}-i+2} = (-1)^{i-1}(i-1), \text{ for } i = 2, 3, \ldots, n, \\
&x_{n+1} = -x_n.
\end{aligned}$$

By direct calculation, we obtain

$$\mathbf{x}_\mathcal{N}^T\mathbf{x}_\mathcal{N} = 1 + 2\sum_{i=1}^{n-1} i^2 + (n-1)^2, \tag{11}$$

and

$$
\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{x}_{\mathcal{N}} = \left[\begin{array}{c} -1 \\ 4 \\ \hdashline -8 \\ \vdots \\ (-1)^{n-1}4(n-2) \\ \hdashline (-1)^{n}(4n-5) \\ (-1)^{n+1}4(n-1) \\ (-1)^{n}(4n-5) \\ \hdashline (-1)^{n-1}4(n-2) \\ \vdots \\ -8 \\ \hdashline 5 \end{array}\right],
$$

which implies

$$
\mathbf{x}_{\mathcal{N}}^T \mathbf{H}_{\mathcal{N}}^{-T} \mathbf{H}_{\mathcal{N}}^{-1} \mathbf{x}_{\mathcal{N}} = 10 + 2 \cdot 4^2 \sum_{i=1}^{n-2} i^2 + 4^2(n-1)^2 + 2(4n-5)^2. \tag{12}
$$

Equations (11) and (12) imply that, for $\mathcal{N}$ even,

$$
\frac{\mathbf{x}_{\mathcal{N}}^T \mathbf{H}_{\mathcal{N}}^{-T} \mathbf{H}_{\mathcal{N}}^{-1} \mathbf{x}_{\mathcal{N}}}{\mathbf{x}_{\mathcal{N}}^T \mathbf{x}_{\mathcal{N}}} \to \frac{2 \cdot 4^2 \sum_{i=1}^{n-1} i^2}{2 \sum_{i=1}^{n-1} i^2} = 16
$$

as $\mathcal{N} \to \infty$. Therefore, for $\mathcal{N}$ even, the largest singular value of $\mathbf{H}_{\mathcal{N}}^{-1} \to 4$ as $\mathcal{N} \to \infty$.

For odd $\mathcal{N} \geq 7$, let $\mathcal{N} = 2n + 1$. Define a vector $\mathbf{x}_{\mathcal{N}} = (x_1, x_2, \ldots, x_{\mathcal{N}})^T$ by

$$
\begin{aligned}
&x_i = x_{\mathcal{N}-i+1} = (-1)^i i, \text{ for } i = 1, 2, \ldots, n, \\
&x_{n+1} = -x_n.
\end{aligned}
$$

Similar to the case for even $\mathcal{N}$, we have

$$
\mathbf{x}_{\mathcal{N}}^T \mathbf{x}_{\mathcal{N}} = 2 \sum_{i=1}^{n} i^2 + n^2, \tag{13}
$$

$$
\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{x}_{\mathcal{N}} = \left[\begin{array}{c} -1 \\ 4 \\ \hdashline -8 \\ \vdots \\ (-1)^n 4(n-1) \\ \hdashline (-1)^{n+1}(4n-1) \\ (-1)^n 4n \\ (-1)^{n+1}(4n-1) \\ \hdashline (-1)^n 4(n-1) \\ \vdots \\ -8 \end{array}\right],
$$

$$
\mathbf{x}_{\mathcal{N}}^T \mathbf{H}_{\mathcal{N}}^{-T} \mathbf{H}_{\mathcal{N}}^{-1} \mathbf{x}_{\mathcal{N}} = 17 + 2 \cdot 4^2 \sum_{i=3}^{n} (i-1)^2 + 2(4n-1)^2 + 4^2 n^2. \tag{14}
$$

Thus, from (13) and (14), we have that, for $\mathcal{N}$ odd,

$$
\frac{\mathbf{x}_{\mathcal{N}}^T \mathbf{H}_{\mathcal{N}}^{-T} \mathbf{H}_{\mathcal{N}}^{-1} \mathbf{x}_{\mathcal{N}}}{\mathbf{x}_{\mathcal{N}}^T \mathbf{x}_{\mathcal{N}}} \to \frac{2 \cdot 4^2 \sum_{i=1}^{n} i^2}{2 \sum_{i=1}^{n} i^2} = 16
$$

as $\mathcal{N} \to \infty$. Therefore, for $\mathcal{N}$ odd, the largest singular value of $\mathbf{H}_{\mathcal{N}}^{-1} \to 4$ as $\mathcal{N} \to \infty$. This completes the proof.

As explained before, $\mathbf{u}_{\mathcal{N}}$ is either a real vector or the product of a real vector and the imaginary unit, $\sqrt{-1}$. In either case, the results in the remainder of this section hold. For simplicity, we assume in the proofs that $\mathbf{u}_{\mathcal{N}}$ is a real vector.

Proposition 2 *For all $\mathcal{N} \geq 2$, the smallest singular value $\sigma_{\mathcal{N}\mathcal{N}}$ and the associated eigenvector $\mathbf{u}_{\mathcal{N}} = (u_0, u_1, \ldots, u_{\mathcal{N}-1})^T$ of $\mathbf{H}_{\mathcal{N}}$ have the following properties:*

1. *$\sigma_{\mathcal{N}+1,\mathcal{N}+1} < \sigma_{\mathcal{N}\mathcal{N}}$ and*

2. *$u_0 \neq 0$ and $u_{\mathcal{N}-1} \neq 0$.*

Proof We prove these two results in three steps. First we prove a result that is weaker than result 1:

$$
\sigma_{\mathcal{N}+2,\mathcal{N}+2} < \sigma_{\mathcal{N}\mathcal{N}} \quad \text{for all } \mathcal{N} \geq 2. \tag{15}
$$

Then we prove result 2 using (15). Finally, we prove result 1 using result 2.

Let $\lambda_{\mathcal{N}}$ be the eigenvalue of $\mathbf{H}_{\mathcal{N}}$ corresponding to $\sigma_{\mathcal{N}\mathcal{N}}$ and let $\mathbf{u}_{\mathcal{N}} = (u_0, u_1, \ldots, u_{\mathcal{N}-1})^T \in \mathbb{R}^{\mathcal{N}}$ be an associated eigenvector. Thus, $\mathbf{H}_{\mathcal{N}}\mathbf{u}_{\mathcal{N}} = \lambda_{\mathcal{N}}\mathbf{u}_{\mathcal{N}}$. Without loss of generality, assume $\mathbf{u}_{\mathcal{N}}$ is normalized so that $\|\mathbf{u}_{\mathcal{N}}\|_2 = 1$. Consequently, $\mathbf{u}_{\mathcal{N}}^T\mathbf{H}_{\mathcal{N}}^{-2}\mathbf{u}_{\mathcal{N}} = \sigma_{\mathcal{N}\mathcal{N}}^{-2}$.

To begin, we show that, for any $\mathcal{N} \geq 2$, u_0 and u_1 cannot both be zero. For $\mathcal{N} = 2$, this follows immediately from $\mathbf{u}_2 = (u_0, u_1)^T$ and $\|\mathbf{u}_2\|_2 = 1$. For $\mathcal{N} > 2$, we prove the result by contradiction. To this end, suppose $u_0 = u_1 = 0$ for some $\mathcal{N} > 2$. Note that $\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{u}_{\mathcal{N}} = \lambda_{\mathcal{N}}^{-1}\mathbf{u}_{\mathcal{N}}$. That is,

$$\begin{bmatrix} & & & & & & 1 \\ & & & & & 1 & -2 \\ & & & & 1 & -2 & 1 \\ & & & \iddots & \iddots & \iddots & \\ & & \iddots & \iddots & \iddots & & \\ & 1 & -2 & 1 & & & \\ 1 & -2 & 1 & & & & \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \\ u_{\mathcal{N}-1} \end{bmatrix} = \lambda_{\mathcal{N}}^{-1} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \\ u_{\mathcal{N}-1} \end{bmatrix}. \tag{16}$$

By comparing the two sides of the system of equations (16), we obtain

1. $u_{\mathcal{N}-1} = \lambda_{\mathcal{N}}^{-1}u_0 = 0$ from the first equation in the system (16) and the assumption $u_0 = 0$;
2. $u_{\mathcal{N}-2} = u_{\mathcal{N}-2} - 2u_{\mathcal{N}-1} = \lambda_{\mathcal{N}}^{-1}u_1 = 0$ from $u_{\mathcal{N}-1} = 0$ proved in item 1 above, the second equation in the system (16) and the assumption $u_1 = 0$;
3. $u_2 = u_0 - 2u_1 + u_2 = \lambda_{\mathcal{N}}^{-1}u_{\mathcal{N}-1} = 0$ from the assumption $u_0 = u_1 = 0$, the last equation in the system (16) and $u_{\mathcal{N}-1} = 0$ proved in item 1 above;
4. $u_3 = u_1 - 2u_2 + u_3 = \lambda_{\mathcal{N}}^{-1}u_{\mathcal{N}-2} = 0$ from the assumption $u_1 = 0$, the result $u_2 = 0$ proved in item 3 above, the second to last equation in the system (16) and $u_{\mathcal{N}-2} = 0$ proved in item 2 above.

Continuing this process, we obtain $\mathbf{u}_{\mathcal{N}} = 0$, which contradicts $\|\mathbf{u}_{\mathcal{N}}\|_2 = 1$. Therefore, we have shown that u_0 and u_1 cannot both be zero.

For any $\mathcal{N} \geq 2$, let $\bar{\mathbf{u}}_{\mathcal{N}+2} = (0, 0, \mathbf{u}_{\mathcal{N}}^T)^T = (0, 0, u_0, u_1, \ldots, u_{\mathcal{N}-1})^T$. Then $\|\bar{\mathbf{u}}_{\mathcal{N}+2}\|_2 = 1$ and

$$\mathbf{H}_{\mathcal{N}+2}^{-1}\bar{\mathbf{u}}_{\mathcal{N}+2} = \begin{bmatrix} \mathbf{H}_{\mathcal{N}}^{-1}\mathbf{u}_{\mathcal{N}} \\ -2u_0 + u_1 \\ u_0 \end{bmatrix}.$$

Furthermore, we have

$$\begin{aligned} \sigma_{\mathcal{N}+2,\mathcal{N}+2}^{-2} = \max_{\|\mathbf{x}\|=1} \mathbf{x}^T\mathbf{H}_{\mathcal{N}+2}^{-2}\mathbf{x} &\geq \bar{\mathbf{u}}_{\mathcal{N}+2}^T\mathbf{H}_{\mathcal{N}+2}^{-2}\bar{\mathbf{u}}_{\mathcal{N}+2} \\ &= \mathbf{u}_{\mathcal{N}}^T\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{u}_{\mathcal{N}} + (u_1 - 2u_0)^2 + u_0^2 \\ &= \sigma_{\mathcal{N}\mathcal{N}}^{-2} + (u_1 - 2u_0)^2 + u_0^2 \\ &> \sigma_{\mathcal{N}\mathcal{N}}^{-2}, \end{aligned}$$

where the last inequality follows from the observation above that u_0 and u_1 cannot both be zero. This proves (15).

Next we prove result 2. From the first equation in the system (16), we see that $u_0 = 0$ if and only if $u_{\mathcal{N}-1} = 0$. So we need only show that $u_0 \neq 0$. We prove this result by contradiction.

First consider the case $\mathcal{N} = 2$. Suppose $u_0 = 0$. From the discussion above, this implies $u_{\mathcal{N}-1} = 0$ too. However, $\mathcal{N} - 1 = 1$ in this case. So, $u_1 = u_{\mathcal{N}-1} = 0$. Consequently, $\mathbf{u}_2 = (u_0, u_1)^T = (0, 0)^T$, which contradicts $\|\mathbf{u}_2\|_2 = 1$. Therefore, $u_0 \neq 0$ in this case.

Next consider the case $\mathcal{N} = 3$. Suppose $u_0 = 0$. Again, from the discussion above, this implies $u_{\mathcal{N}-1} = 0$ too. However, $\mathcal{N} - 1 = 2$ in this case. So, $u_2 = u_{\mathcal{N}-1} = 0$. From $u_0 = u_2 = 0$ and the last equation of the system (16), we get $-2u_1 = u_0 - 2u_1 + u_2 = \lambda_3^{-1}u_2 = 0$, whence $u_1 = 0$ too. Consequently, $\mathbf{u}_3 = (u_0, u_1, u_2)^T = (0, 0, 0)^T$, which contradicts $\|\mathbf{u}_3\|_2 = 1$. Therefore, $u_0 \neq 0$ in this case either.

For the general case, suppose that, for some $\mathcal{N} \geq 4$, $u_0 = 0$. Again, from the discussion above, this implies $u_{\mathcal{N}-1} = 0$ too. Since $u_0 = u_{\mathcal{N}-1} = 0$, deleting the first and the last rows as well as the first and the last columns of $\mathbf{H}_{\mathcal{N}}^{-1}$ and correspondingly the first and the last elements of $\mathbf{u}_{\mathcal{N}}$ (i.e., u_0 and $u_{\mathcal{N}-1}$) results in the new system of equations

$$\begin{bmatrix} & & & & 1 \\ & & & 1 & -2 \\ & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \\ & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & & \\ 1 & -2 & 1 & & \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \end{bmatrix} = \lambda_{\mathcal{N}}^{-1} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \end{bmatrix}$$

which is equivalent to $\mathbf{H}_{\mathcal{N}-2}^{-1}\hat{\mathbf{u}}_{\mathcal{N}-2} = \lambda_{\mathcal{N}}^{-1}\hat{\mathbf{u}}_{\mathcal{N}-2}$, where $\hat{\mathbf{u}}_{\mathcal{N}-2} = (u_1, \ldots, u_{\mathcal{N}-2})^T$. Moreover, $\hat{\mathbf{u}}_{\mathcal{N}-2} \neq 0$, since $\mathbf{u}_{\mathcal{N}} = (0, \hat{\mathbf{u}}_{\mathcal{N}-2}^T, 0)^T$ and $\|\mathbf{u}_{\mathcal{N}}\|_2 = 1$. Consequently, $\lambda_{\mathcal{N}}$ is an eigenvalue of $\mathbf{H}_{\mathcal{N}-2}$. By definition, $\sigma_{\mathcal{N}-2,\mathcal{N}-2} \leq |\lambda_{\mathcal{N}}|$, since $\sigma_{\mathcal{N}-2,\mathcal{N}-2}$ is the smallest singular value of $\mathbf{H}_{\mathcal{N}-2}$, but we also have the $|\lambda_{\mathcal{N}}| = \sigma_{\mathcal{N}\mathcal{N}}$, whence $\sigma_{\mathcal{N}\mathcal{N}} \geq \sigma_{\mathcal{N}-2,\mathcal{N}-2}$. Since $\mathcal{N}-2 \geq 2$, this contradicts (15). Thus, we conclude that $u_0 \neq 0$, which completes the proof of result 2.

Now we prove result 1. First consider the case $\mathcal{N} = 2$. Direct calculation shows that the eigenvalues of $\mathbf{H}_2$ are $1 \pm \sqrt{2}$. Therefore, the smallest singular value of $\mathbf{H}_2$ is $\sigma_{22} = \sqrt{2} - 1 > 0.4$. On the other hand, the characteristic polynomial of $\mathbf{H}_3$ is $p(x) = x^3 - 4x^2 - 2x + 1$ and $p(0) = 1$ and $p(0.4) = -0.376$. Thus, $p(x)$ has a root in the interval $(0, 0.4)$, whence $0 < \sigma_{33} < 0.4 < \sigma_{22}$. This proves result 1 for the case $\mathcal{N} = 2$.

For the general case $\mathcal{N} \geq 3$, let $\bar{\mathbf{u}}_{\mathcal{N}+1} = (0, \mathbf{u}_{\mathcal{N}}^T)^T = (0, u_0, u_1, \ldots, u_{\mathcal{N}-1})^T$. Note that $\|\bar{\mathbf{u}}_{\mathcal{N}+1}\|_2 = 1$, since $\|\mathbf{u}_{\mathcal{N}}\|_2 = 1$, and that

$$\mathbf{H}_{\mathcal{N}+1}^{-1}\bar{\mathbf{u}}_{\mathcal{N}+1} = \begin{bmatrix} \mathbf{H}_{\mathcal{N}}^{-1}\mathbf{u}_{\mathcal{N}} \\ u_1 - 2u_0 \end{bmatrix}.$$

Therefore,

$$\begin{aligned} \sigma_{\mathcal{N}+1,\mathcal{N}+1}^{-2} = \max_{\|\mathbf{x}\|=1} \mathbf{x}^T \mathbf{H}_{\mathcal{N}+1}^{-2}\mathbf{x} \ &\geq \ \bar{\mathbf{u}}_{\mathcal{N}+1}^T \mathbf{H}_{\mathcal{N}+1}^{-2} \bar{\mathbf{u}}_{\mathcal{N}+1} \\ &= \ \mathbf{u}_{\mathcal{N}}^T \mathbf{H}_{\mathcal{N}}^{-1}\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{u}_{\mathcal{N}} + (u_1 - 2u_0)^2 \\ &= \ \sigma_{\mathcal{N}\mathcal{N}}^{-2} + (u_1 - 2u_0)^2. \end{aligned} \tag{17}$$

Thus, to prove that $\sigma_{\mathcal{N}+1,\mathcal{N}+1}^{-2} > \sigma_{\mathcal{N}\mathcal{N}}^{-2}$, it suffices to prove that $(u_1 - 2u_0)^2 > 0$ (i.e., $u_1 \neq 2u_0$). We prove this by contradiction.

Suppose that $u_1 = 2u_0$. Since $u_0 \neq 0$, we can define $\mathbf{v}_{\mathcal{N}} = \frac{1}{u_0}\mathbf{u}_{\mathcal{N}} = (v_0, v_1, \ldots, v_{\mathcal{N}-1})^T$. Then $v_0 = 1$, $v_1 = 2$ and $\mathbf{v}_{\mathcal{N}}$ is an eigenvector of $\mathbf{H}_{\mathcal{N}}^{-2}$ with corresponding eigenvalue $\sigma_{\mathcal{N}\mathcal{N}}^{-2}$.

For $\mathcal{N} \geq 3$, direct calculation shows that

$$\mathbf{H}_{\mathcal{N}}^{-2} = \begin{bmatrix} 1 & -2 & 1 & & & & & & \\ -2 & 5 & -4 & 1 & & & & & \\ 1 & -4 & 6 & -4 & 1 & & & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & & 1 & -4 & 6 & -4 & 1 & \\ & & & & 1 & -4 & 6 & -4 \\ & & & & & 1 & -4 & 6 \end{bmatrix}.$$

We begin by showing that $\sigma_{\mathcal{N}\mathcal{N}}^{-2} > 10$ for all $\mathcal{N} \geq 3$. To this end, note that, from the theorem [6, Theorem 4.3.8] about the interlacing of eigenvalues for bordered matrices, we know that $\sigma_{\mathcal{N}\mathcal{N}}^{-2} \leq \sigma_{\mathcal{N}+1,\mathcal{N}+1}^{-2}$. Therefore, it suffices to show that $\sigma_{3,3}^{-2} > 10$. To show this, note that the characteristic polynomial of $\mathbf{H}_3^{-2}$ is $p(x) = x^3 - 12x^2 + 20x - 1$ and that $p(10) = -1$ and $p(11) = 98$. Thus, $\mathbf{H}_3^{-2}$ has an eigenvalue in the interval $(10, 11)$, whence $\sigma_{3,3}^{-2} > 10$.

Next, for any $\mathcal{N} \geq 3$, we show by induction on k that

$$v_{k+1} > 4v_k > 0, \text{ for all } k = 1, 2, \ldots, \mathcal{N} - 2. \tag{18}$$

For $k = 1$, we have from the first equation of the system

$$\mathbf{H}_{\mathcal{N}}^{-2}\mathbf{v}_{\mathcal{N}} = \sigma_{\mathcal{N}\mathcal{N}}^{-2}\mathbf{v}_{\mathcal{N}} \tag{19}$$

that

$$v_0 - 2v_1 + v_2 = \sigma_{\mathcal{N}\mathcal{N}}^{-2} v_0,$$

whence

$$v_2 - 4v_1 = \sigma_{\mathcal{N}\mathcal{N}}^{-2} v_0 - 2v_1 - v_0 = \sigma_{\mathcal{N}\mathcal{N}}^{-2} - 5 > 0.$$

Thus, (18) holds for $k = 1$.

For $k = 2$, we have from the second equation of the system (19) that

$$v_3 - 4v_2 = \sigma_{\mathcal{N}\mathcal{N}}^{-2} v_1 - 5v_1 + 2v_0 > 20 - 10 + 2 > 0.$$

Thus, (18) holds for $k = 2$.

For the induction step, assume that $2 \le K < \mathcal{N} - 2$ and that (18) holds for all $k \le K$. Then, from the $(K+1)$-st equation of the system (19), we obtain that

$$v_{K+2} - 4v_{K+1} = (\sigma_{\mathcal{N}\mathcal{N}}^{-2} - 6)v_K + 4v_{K-1} - v_{K-2},$$

which together with (18) and $\sigma_{\mathcal{N}\mathcal{N}}^{-2} > 10$ implies that

$$v_{K+2} > 4v_{K+1} > 0.$$

Thus, (18) holds for $k = K + 1$ also, completing the induction proof of (18).

For $k = \mathcal{N} - 2$, (18) becomes

$$v_{\mathcal{N}-1} > 4v_{\mathcal{N}-2} > 0. \tag{20}$$

On the other hand, the first two equations of (16) together with $u_1 = 2u_0$ imply that $u_{\mathcal{N}-2} = 4u_{\mathcal{N}-1}$, whence, by the definition of $\mathbf{v}_{\mathcal{N}}$, we have that $v_{\mathcal{N}-2} = 4v_{\mathcal{N}-1}$, which contradicts (20).

Thus, $u_1 \neq 2u_0$, which implies that $(u_1 - 2u_0)^2 > 0$. Therefore, from (17), we know that $\sigma_{\mathcal{N}+1,\mathcal{N}+1}^{-2} > \sigma_{\mathcal{N},\mathcal{N}}^{-2}$, which implies that $\sigma_{\mathcal{N}+1,\mathcal{N}+1} < \sigma_{\mathcal{N}\mathcal{N}}$. This completes the proof.

Since $u_0 \neq 0$, zero is not a root of $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$. That is, $\zeta_n \neq 0$ for any n. Furthermore, we can prove that

Proposition 3 *All the roots $\zeta_1, \zeta_2, \ldots, \zeta_{\mathcal{N}-1}$ of the polynomial $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$ are within the unit circle and they are either real or pairwise complex conjugate.*

Proof As explained in Section 2.2, all roots $\zeta_1, \zeta_2, \ldots, \zeta_{\mathcal{N}-1}$ of $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$ are either real or pairwise complex conjugate. To prove that all roots are within the unit circle, it suffices to prove, according to the Schur-Cohn criterion [1], that the so-called Bezoutian matrix,

$$\mathbf{C}^T\mathbf{C} - \mathbf{D}^T\mathbf{D},$$

is positive definite, where

$$\mathbf{C} = \begin{bmatrix} u_{\mathcal{N}-1} & u_{\mathcal{N}-2} & \cdots & u_1 \\ & u_{\mathcal{N}-1} & \cdots & u_2 \\ & & \cdots & \cdots \\ & & & u_{\mathcal{N}-1} \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} u_0 & u_1 & \cdots & u_{\mathcal{N}-2} \\ & u_0 & \cdots & u_{\mathcal{N}-3} \\ & & \cdots & \cdots \\ & & & u_0 \end{bmatrix}. \tag{21}$$

From Proposition 2, we know that $u_{\mathcal{N}-1} \neq 0$, so $\mathbf{C}^{-1}$ exists. Therefore

$$\mathbf{C}^T\mathbf{C} - \mathbf{D}^T\mathbf{D} = \mathbf{C}^T\left(\mathbf{I} - \mathbf{C}^{-T}\mathbf{D}^T\mathbf{D}\mathbf{C}^{-1}\right)\mathbf{C}.$$

Let $\mathbf{Y} = \mathbf{D}\mathbf{C}^{-1}$. A straightforward, but tedious, calculation shows that

$$\mathbf{Y} = \lambda_{\mathcal{N}} \begin{bmatrix} 1 & -2 & 1 & & & & \\ & \cdots & \cdots & & & & \\ & & \cdots & \cdots & & & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \\ & & & & & 1 \end{bmatrix}_{(\mathcal{N}-1)\times(\mathcal{N}-1)} = \lambda_{\mathcal{N}}\mathbf{P}\mathbf{H}_{\mathcal{N}-1}^{-1},$$

where $\lambda_{\mathcal{N}} \neq 0$ is the eigenvalue of $\mathbf{H}_{\mathcal{N}}$ that corresponds to the eigenvector $\mathbf{u}_{\mathcal{N}}$ (i.e., $\mathbf{H}_{\mathcal{N}}\mathbf{u}_{\mathcal{N}} = \lambda_{\mathcal{N}}\mathbf{u}_{\mathcal{N}}$) and

$$\mathbf{P} = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}_{(\mathcal{N}-1)\times(\mathcal{N}-1)}$$

is a permutation matrix. Consequently,

$$\mathbf{Y}^T\mathbf{Y} = \lambda_{\mathcal{N}}^2\mathbf{H}_{\mathcal{N}-1}^{-T}\mathbf{P}^T \cdot \mathbf{P}\mathbf{H}_{\mathcal{N}-1}^{-1} = \sigma_{\mathcal{N}\mathcal{N}}^2\mathbf{H}_{\mathcal{N}-1}^{-T}\mathbf{H}_{\mathcal{N}-1}^{-1} = \sigma_{\mathcal{N}\mathcal{N}}^2\mathbf{H}_{\mathcal{N}-1}^{-2}$$

since $|\lambda_{\mathcal{N}}| = \sigma_{\mathcal{N}\mathcal{N}}$ and $\mathbf{P}^T\mathbf{P} = \mathbf{I}$. Let $\mathbf{H}_{\mathcal{N}-1} = \mathbf{V}_{\mathcal{N}-1}\Lambda_{\mathcal{N}-1}\mathbf{V}_{\mathcal{N}-1}^T$ be the spectral decomposition of $\mathbf{H}_{\mathcal{N}-1}$, where $\Lambda_{\mathcal{N}-1} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{\mathcal{N}-1})$, $|\lambda_n| = \sigma_{\mathcal{N}-1,n}$, for $1 \leq n \leq \mathcal{N}-1$, and $\mathbf{V}_{\mathcal{N}-1}^T\mathbf{V}_{\mathcal{N}-1} = \mathbf{I}$. Thus,

$$\mathbf{I}-\mathbf{Y}^T\mathbf{Y} = \mathbf{V}_{\mathcal{N}-1}\left[\mathbf{I} - \sigma_{\mathcal{N}\mathcal{N}}^2\Lambda_{\mathcal{N}-1}^{-2}\right]\mathbf{V}_{\mathcal{N}-1}^T = \mathbf{V}_{\mathcal{N}-1}\left[\mathbf{I} - \sigma_{\mathcal{N}\mathcal{N}}^2\left[\mathrm{diag}(\sigma_{\mathcal{N}-1,1}, \ldots, \sigma_{\mathcal{N}-1,\mathcal{N}-1})\right]^{-2}\right]\mathbf{V}_{\mathcal{N}-1}^T.$$

Since $\sigma_{\mathcal{N}-1,1} \geq \ldots \geq \sigma_{\mathcal{N}-1,\mathcal{N}-1} > \sigma_{\mathcal{N}\mathcal{N}} > 0$, $\mathbf{I} - \sigma_{\mathcal{N}\mathcal{N}}^2\left[\mathrm{diag}(\sigma_{\mathcal{N}-1,1}\ldots, \sigma_{\mathcal{N}-1,\mathcal{N}-1})\right]^{-2}$ is a diagonal matrix with all diagonal elements positive. Since $\mathbf{C}$ and $\mathbf{V}_{\mathcal{N}-1}$ are nonsingular, $\mathbf{C}^T\mathbf{C} - \mathbf{D}^T\mathbf{D}$ is positive definite. This completes the proof.

Since $\gamma_n = 2\mathcal{N}\log\zeta_n$, we have from Proposition 3 that

Corollary 1 *The real parts of all γ_n are negative (i.e., $\Re(\gamma_n) < 0$ for $n = 1, \ldots, \mathcal{N}-1$).*

4 Numerical results

We present some numerical results in the table and two figures below. Table 1 lists the values of ω_n and γ_n for the 25-term exponential approximation. Conjugacy of ω_n and also γ_n is clearly shown in the table. Figure 1 contains a plot of the singular values of the Hankel matrix $\frac{1}{\mathcal{N}}\mathbf{H}_{\mathcal{N}}$ associated with this 25-term exponential approximation. Figure 2 contains a plot of the error $|h(x) - h_{\exp}(x)|$ for each of the 25-, 50-, 100-, 200-, and 400-term exponential approximations to the HS function over the interval $[0, 30]$. In each of these five plots, we see that the error $|h(x) - h_{\exp}(x)|$ converges to zero as x increases. Moreover, as mentioned in Remark 4 in Section 2, the errors decrease linearly with $\mathcal{N}$ as $\mathcal{N}$ increases (i.e., $\epsilon = \frac{1}{4\mathcal{N}}$).

n	$\Re(\omega_n)$	$\Im(\omega_n)$	$\Re(\gamma_n)$	$\Im(\gamma_n)$
1	1.68E-04	-3.16E-05	-5.68E-02	1.45E+02
2	1.68E-04	3.16E-05	-5.68E-02	-1.45E+02
3	2.04E-04	-5.98E-05	-1.72E-01	1.32E+02
4	2.04E-04	5.98E-05	-1.72E-01	-1.32E+02
5	2.69E-04	-1.02E-04	-3.51E-01	1.20E+02
6	2.69E-04	1.02E-04	-3.51E-01	-1.20E+02
7	3.87E-04	-1.70E-04	-5.98E-01	1.07E+02
8	3.87E-04	1.70E-04	-5.98E-01	-1.07E+02
9	6.02E-04	-2.94E-04	-9.25E-01	9.49E+01
10	6.02E-04	2.94E-04	-9.25E-01	-9.49E+01
11	1.01E-03	-5.39E-04	-1.35E+00	8.24E+01
12	1.01E-03	5.39E-04	-1.35E+00	-8.24E+01
13	1.87E-03	-1.09E-03	-1.89E+00	6.99E+01
14	1.87E-03	1.09E-03	-1.89E+00	-6.99E+01
15	3.86E-03	-2.53E-03	-2.58E+00	5.74E+01
16	3.86E-03	2.53E-03	-2.58E+00	-5.74E+01
17	9.17E-03	-7.44E-03	-3.50E+00	4.49E+01
18	9.17E-03	7.44E-03	-3.50E+00	-4.49E+01
19	2.45E-02	-3.19E-02	-4.73E+00	3.24E+01
20	2.45E-02	3.19E-02	-4.73E+00	-3.24E+01
21	7.57E-03	-2.10E-01	-6.44E+00	2.04E+01
22	7.57E-03	2.10E-01	-6.44E+00	-2.04E+01
23	3.81E+00	0.00E+00	-9.65E+00	0.00E+00
24	-1.46E+00	1.03E-01	-8.54E+00	9.38E+00
25	-1.46E+00	-1.03E-01	-8.54E+00	-9.38E+00

Table 1: Real and imaginary parts of ω_n and γ_n for the 25-term exponential approximation to the HS function.

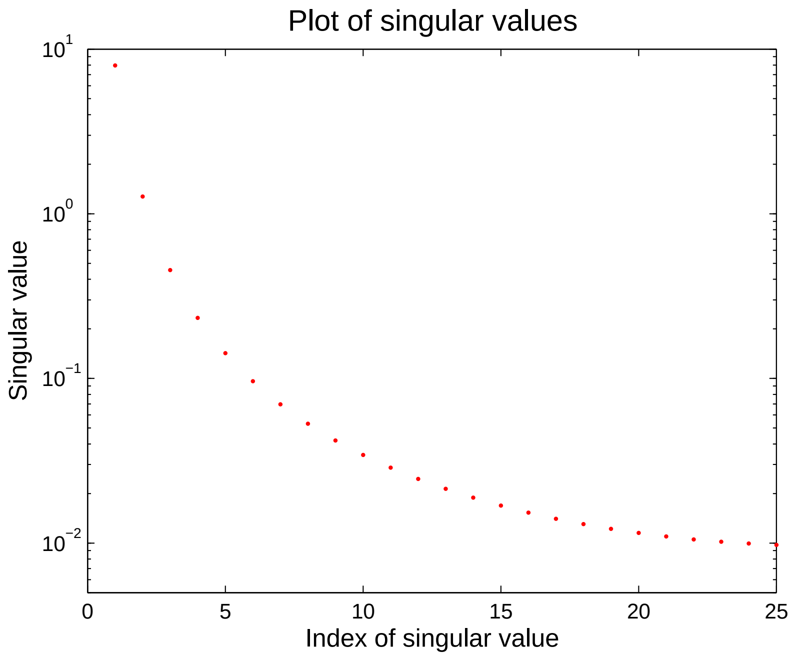

Figure 1: The singular values associated with the 25-term exponential approximation to the HS function.

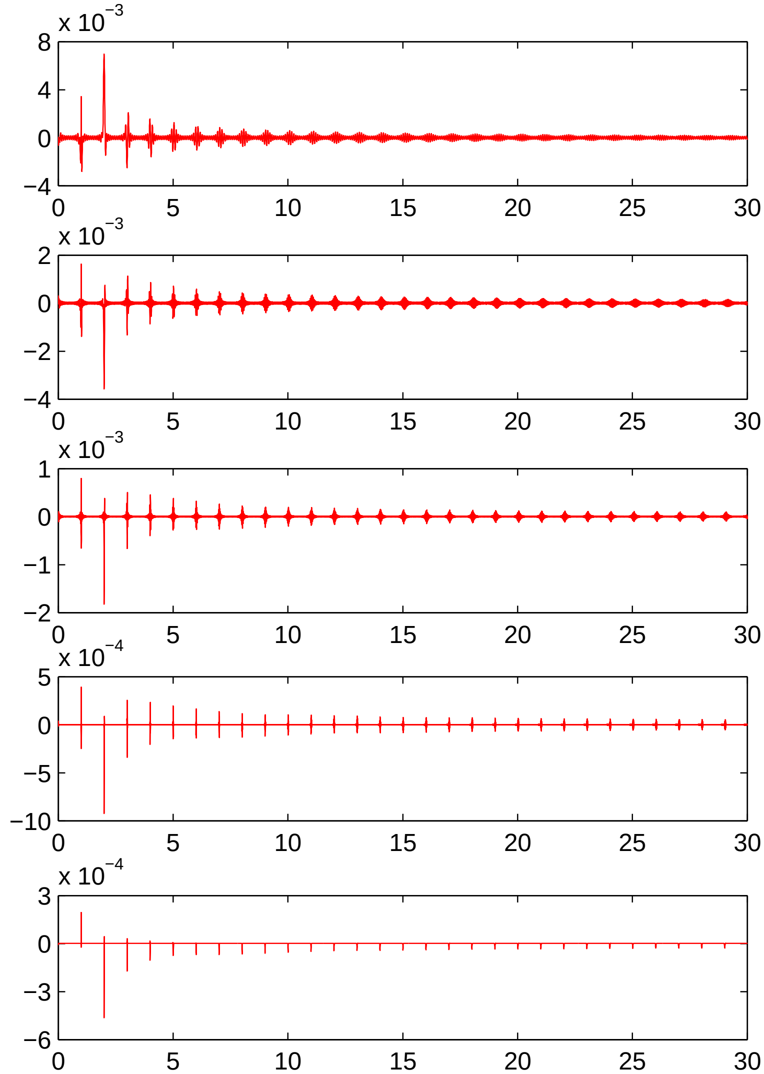

Figure 2: Plots from top to bottom of the errors $|h(x) - h_{\exp}(x)|$ for the 25-, 50-, 100-, 200-, and 400-term exponential approximations to the HS function over $[0, 30]$.

Acknowledgments

We thank Professor Christina Christara for some very helpful discussions on the theoretical properties of the exponential approximation. We also thank one anonymous referee for pointing out an error in the previous version of this paper and giving some useful suggestions on how to prove Proposition 3.

References

[1] Reuven Ackner and Thomas Kailath. On the Ptak-Young generalization of the Schur-Cohn theorem. *IEEE Transactions on Automatic Control*, 37(10):1601–1604, 1992.

[2] Gregory Beylkin and Lucas Monzón. On approximation of functions by exponential sums. *Applied and Computational Harmonic Analysis*, 19(1):17–48, 2005.

[3] Dietrich Braess. *Nonlinear Approximation Theory*, volume 7 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin Heidelberg, 1986.

[4] Gene Golub and Victor Pereyra. Separable nonlinear least squares: The variable projection method and its applications. *Inverse Problems*, 19(2):R1–R26, 2003.

[5] Gene H Golub and Charles F Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 2nd edition, 1991.

[6] Roger A Horn and Charles R Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[7] Ian Iscoe, Ken Jackson, Alex Kreinin, and Xiaofang Ma. Pricing correlation-dependent derivatives based on exponential approximations to the hockey stick function. Available from http://www.cs.toronto.edu/NA/reports.html, January 2007.

[8] Liviu Gr Ixaru and Guido Vanden Berghe. *Exponential Fitting*. Kluwer Academic Publishers, 2004.

[9] Ken Jackson, Alex Kreinin, and Xiaofang Ma. Loss distribution evaluation for synthetic CDOs. Available from http://www.cs.toronto.edu/NA/reports.html, February 2007.

[10] Zhongqi Jing and Adly T Fam. An algorithm for computing continuous Chebyshev approximations. *Mathematics of Computation*, 48(178):691–710, 1987.