



Convergence analysis of approximation formulas for analytic functions via duality for potential energy minimization

Satoshi Hayakawa¹ · Ken'ichiro Tanaka² 

Received: 21 October 2022 / Revised: 2 March 2023 / Accepted: 30 March 2023 /

Published online: 4 May 2023

© The Author(s) 2023

Abstract

We investigate the approximation formulas that were proposed by Tanaka & Sugihara (IMA J. Numer. Anal. 39(4):1957–1984, 2019), in weighted Hardy spaces, which are analytic function spaces with certain asymptotic decay. Under the criterion of minimum worst error of n -point approximation formulas, we demonstrate that the formulas are nearly optimal. We also obtain the upper bounds of the approximation errors that coincide with the existing heuristic bounds in asymptotic order by a duality theorem for the minimization problem of potential energy.

Keywords Approximation · Weighted Hardy space · Convex optimization · Duality · Discrete energy minimization

Mathematics Subject Classification 65D05 · 65E05 · 41A25 · 49N15

1 Introduction

By taking over the arguments of [23], Tanaka & Sugihara [24] proposed an algorithm to design accurate approximation formulas in function spaces called weighted Hardy

This study was supported by the Japan Society for the Promotion of Science with KAKENHI (17K14241 to K.T.).

✉ Ken'ichiro Tanaka
kenichiro@mist.i.u-tokyo.ac.jp
Satoshi Hayakawa
hayakawa@maths.ox.ac.uk

¹ Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter (550), Woodstock Road, Oxford OX2 6GG, UK

² Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, Japan

spaces defined by

$$\mathbb{H}^\infty(\mathcal{D}_d, w) := \left\{ f : \mathcal{D}_d \rightarrow \mathbb{C} \mid f \text{ is analytic on } \mathcal{D}_d, \sup_{z \in \mathcal{D}_d} \left| \frac{f(z)}{w(z)} \right| < \infty \right\}, \quad (1)$$

where $d > 0$, $\mathcal{D}_d := \{z \in \mathbb{C} \mid |\operatorname{Im} z| < d\}$, and w is a weight function characterized later in Sect. 2.1. The spaces $\mathbb{H}^\infty(\mathcal{D}_d, w)$ are often considered as spaces of transformed functions for well-used sinc approximation formulas shown later in (2). The objective of [23] and [24] was to provide formulas outperforming the sinc formulas. However, their studies only provided heuristic analyses on the proposed formulas without any theoretical guarantees, although their methods have shown superiority to the sinc approximation formulas. In this study, we mathematically

- (1) prove near optimality of the formulas, and
- (2) provide a general upper bound of the errors of the proposed formulas and show that the bound coincides in asymptotic order with the heuristic bound derived by [23].

Below we describe the background of this study more precisely. The spaces $\mathbb{H}^\infty(\mathcal{D}_d, w)$ appear in literature as spaces of variable-transformed functions [18, 19, 21, 25]. For example, the double exponential (DE) transform, which is well-used in numerical analysis [22], has the form

$$f(x) = g\left(\tanh\left(\frac{\pi}{2} \sinh(x)\right)\right)$$

and shows a double-exponential decay. Also, TANH transform $g(\tanh(x/2))$ is commonly used [2, 15]. These variable transformations are employed for the accurate approximation of functions by yielding functions with rapid decay on \mathcal{D}_d , which enables us to neglect the values of the functions for large $|x|$. This motivates us to analyze the approximation possibility over weighted Hardy spaces with general weight functions w . After Sugihara [21] demonstrated near optimality of sinc approximation formulas

$$f(x) \approx \sum_{k=N_-}^{k=N_+} f(kh) \operatorname{sinc}\left(\frac{x}{h} - k\right) \quad (2)$$

for several weight functions w , attempts to construct an optimal formula for general weight functions was started in the literature.

For this purpose, Tanaka et al. [23] employed potential theoretical arguments to generate sampling points for the approximation of functions. Furthermore, Tanaka & Sugihara [24] simplified the arguments and proposed accurate formulas $L_n[a^*; f](x)$ given later by (6) with special sets a^* of sampling points. The formulas $L_n[a^*; f](x)$

outperform the sinc methods for functions $f \in \mathbb{H}^\infty(\mathcal{D}_d, w)$. The authors showed that

$$\sup_{\|f\| \leq 1, x \in \mathbb{R}} |f(x) - L_n[a^*; f](x)| \leq \exp\left(-\frac{F_{K,Q}^D(n)}{n-1}\right), \tag{3}$$

where $\|f\|$ is a norm of $f \in \mathbb{H}^\infty(\mathcal{D}_d, w)$ and $F_{K,Q}^D(n)$ is determined later in (13) by a ‘‘discrete’’ energy minimization problem. Furthermore, they considered the minimum worst error $E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w))$ in (5) of n -point approximation formulas in $\mathbb{H}^\infty(\mathcal{D}_d, w)$ and evaluated it as

$$\exp\left(-\frac{F_{K,Q}^C(n)}{n}\right) \leq E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w)), \tag{4}$$

where $F_{K,Q}^C(n)$ is determined later in (11) by a ‘‘continuous counterpart’’ of the above energy minimization problem. The following problems about the formula $L_n[a^*; f](x)$ were left unsolved in [24].

- (i) Since (the RHS of (4)) \leq (the LHS of (3)), the formula $L_n[a^*; f](x)$ is assured of ‘‘near optimality’’ if $F_{K,Q}^C(n)$ and $F_{K,Q}^D(n)$ are close. However, their difference was not estimated.
- (ii) To estimate the convergence rate of the error in the LHS of (3), we need to know how $F_{K,Q}^D(n)$ depends on n . However, it was not known.

In this paper, we provide solutions to these problems. Our contributions (1) and (2) mentioned in the first paragraph of this section correspond to the solutions to problems (i) and (ii), respectively. More precisely, we show the following statements.

- (1) We show an evaluation like

$$F_{K,Q}^D(n) \lesssim F_{K,Q}^C(n) \lesssim 2F_{K,Q}^D(n).$$

Its rigorous version is given by Theorem 23 in Sect. 2.3. The quantities $F_{K,Q}^D(n)$ and $F_{K,Q}^C(n)$ were obtained from the optimal solutions of the ‘‘discrete’’ energy minimization problem and its ‘‘continuous counterpart’’, respectively. Therefore we construct a feasible solution for the latter using the optimal solution of the former to show this theorem.

- (2) We show an inequality

$$\frac{F_{K,Q}^C(n)}{n} \geq \frac{Q(\alpha_n)}{2},$$

where $Q(x) = -\log w(x)$ and α_n is determined by a tractable inequality. Its details are given by Theorem 24 in Sect. 2.3. By combining this inequality, the above statement (1), and Inequality (3), we obtain explicit convergence rates of the proposed formulas. To show this theorem, we consider the dual problem of the ‘‘continuous’’ energy minimization problem and provide its feasible solution. For

preparation, we present a primal-dual theory of the energy minimization problem in Sect. 4.

As a result, we explicitly obtain lower bounds of $F_{K,Q}^C(n)$ and demonstrate that the rates of lower bounds coincide with those of heuristic bounds in [23].

The rest of this paper is organized as follows. In Sect. 2, we present a mathematical overview of the existing studies and describe our main results as mathematical statements. Section 3 describes the proof of the first result, i.e., Theorem 23. Section 4 contains general arguments, which introduce the concept of “positive semi-definite in measure”. Then, we show that the problem under our interest is a special case of that concept and derive the duality theorem. The evaluations for the second result, described by Theorem 24, are given in Sect. 5. We compare the bounds with those in [23] in Sect. 6. Finally, we describe the concluding remarks in Sect. 7.

2 Mathematical preliminaries and main results

2.1 General settings

We first give some definitions and formulate the problem mathematically. Let $d > 0$ and define the strip region $\mathcal{D}_d := \{z \in \mathbb{C} \mid |\text{Im } z| < d\}$. Throughout this paper, a *weight function* $w : \mathcal{D}_d \rightarrow \mathbb{C}$ is supposed to satisfy the following conditions:

1. w is analytic and does not vanish over the domain \mathcal{D}_d and takes values in $(0, 1]$ on \mathbb{R} ;
2. w satisfies $\lim_{x \rightarrow \pm\infty} \int_{-d}^d |w(x + iy)| dy = 0$ and $\lim_{y \nearrow d} \int_{-\infty}^{\infty} (|w(x + iy)| + |w(x - iy)|) dx < \infty$;
3. $\log w$ is strictly concave on \mathbb{R} .

For a weight function with the above conditions, we define the weighted Hardy space $\mathbb{H}^\infty(\mathcal{D}_d, w)$ on \mathcal{D}_d in (1). We define

$$\|f\| := \sup_{z \in \mathcal{D}_d} \left| \frac{f(z)}{w(z)} \right|$$

for $f \in \mathbb{H}^\infty(\mathcal{D}_d, w)$, and the expression $\|f\| < \infty$ shall also imply $f \in \mathbb{H}^\infty(\mathcal{D}_d, w)$ in the following.

For an approximation formula over $\mathbb{H}^\infty(\mathcal{D}_d, w)$, an evaluation criterion needs to be defined. Based on [21] and [24], we adopt the minimum worst-case error

$$E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w)) := \inf \left\{ \sup_{\|f\| \leq 1, x \in \mathbb{R}} \left| f(x) - \sum_{j=1}^l \sum_{k=0}^{n_j-1} f^{(k)}(a_j) \phi_{jk}(x) \right| \mid \begin{array}{l} 1 \leq l \leq n, n_1 + \dots + n_l = n, \\ a_j \in \mathcal{D}_d \text{ are distinct,} \\ \phi_{jk} : \mathcal{D}_d \rightarrow \mathbb{C} \text{ are analytic} \end{array} \right\} \tag{5}$$

as the optimal performance over all possible n -point interpolation formulas on \mathbb{R} , which is applicable to any $f \in \mathbb{H}^\infty(\mathcal{D}_d, w)$.

2.2 Properties of approximation formulas to be analyzed

Let us introduce some functions dependent on an n -sequence $a = \{a_j\}_{j=1}^n \subset \mathbb{R}$ as follows.

$$\begin{aligned}
 T_d(x) &:= \tanh\left(\frac{\pi}{4d}x\right), \\
 B_n(x; a, \mathcal{D}_d) &:= \prod_{j=1}^n \frac{T_d(x) - T_d(a_j)}{1 - T_d(a_j)T_d(x)}, \\
 B_{n;k}(x, a, \mathcal{D}_d) &:= \prod_{\substack{1 \leq j \leq n, \\ j \neq k}} \frac{T_d(x) - T_d(a_j)}{1 - T_d(a_j)T_d(x)}.
 \end{aligned}$$

Using these functions, we can give an n -point interpolation formula

$$L_n[a; f](x) := \sum_{k=1}^n f(a_k) \frac{B_{n;k}(x; a, \mathcal{D}_d)w(x)}{B_{n;k}(a_k; a, \mathcal{D}_d)w(a_k)} \frac{T'_d(x - a_k)}{T'_d(0)}, \tag{6}$$

which is known to characterize the value $E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w))$ as follows.

Proposition 21 [21, 24] *We have an upper bound of the error of (6) as*

$$\sup_{\|f\| \leq 1, x \in \mathbb{R}} |f(x) - L_n[a; f](x)| \leq \sup_{x \in \mathbb{R}} |B_n(x; a, \mathcal{D}_d)w(x)|$$

for any fixed sequence $a = \{a_j\}_{j=1}^n \subset \mathbb{R}$ (of distinct points). Moreover, by taking infimum of the above expression over all n -sequences, it holds that

$$\begin{aligned}
 E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w)) &= \inf_{a_j \in \mathbb{R}} \sup_{\|f\| \leq 1, x \in \mathbb{R}} |f(x) - L_n[a; f](x)| \\
 &= \inf_{a_j \in \mathbb{R}} \sup_{x \in \mathbb{R}} |B_n(x; a, \mathcal{D}_d)w(x)|.
 \end{aligned}$$

By this assertion, it is enough to consider interpolation formulas of the form (6). Additionally, this motivates us to analyze the value $\sup_{x \in \mathbb{R}} |B_n(x; a, \mathcal{D}_d)w(x)|$, which is simpler than the worst-case error of (6). In [23] and [24],

$$-\log\left(\inf_{a_j \in \mathbb{R}} \sup_{x \in \mathbb{R}} |B_n(x; a, \mathcal{D}_d)w(x)|\right)$$

is treated as an optimal value of an optimization problem (justifiable by the addition rule of tanh)

$$\begin{aligned}
 \text{(DC)} \quad &\text{maximize} \quad \inf_{x \in \mathbb{R}} \left(\sum_{i=1}^n K(x - a_i) + Q(x) \right) \\
 &\text{subject to} \quad a_1 < \dots < a_n,
 \end{aligned}$$

where K and Q are defined by

$$K(x) := -\log |T_d(x)| \left(= -\log \left| \tanh \left(\frac{\pi}{4d} x \right) \right| \right), \tag{7}$$

$$Q(x) := -\log w(x). \tag{8}$$

They considered a continuous relaxation of (DC) as

$$\begin{aligned} \text{(CT)} \quad & \text{maximize} \quad \inf_{x \in \mathbb{R}} \left(\int_{\mathbb{R}} K(x - y) \, d\mu(y) + Q(x) \right) \\ & \text{subject to} \quad \mu \in \mathcal{M}_c(\mathbb{R}, n), \end{aligned}$$

where, we define $\mathcal{M}(\mathbb{R}, n)$ as the set of all (positive) Borel measures μ over \mathbb{R} with $\mu(\mathbb{R}) = n$ and

$$\mathcal{M}_c(\mathbb{R}, n) := \{ \mu \in \mathcal{M}(\mathbb{R}, n) \mid \text{supp } \mu \text{ is compact} \}.$$

Because each feasible solution of (DC) can be interpreted as a combination of δ -measures being a feasible solution of (CT),

$$\text{(the optimal value of (DC))} \leq \text{(the optimal value of (CT))} \tag{9}$$

Potential theoretical arguments [5, 14, 24] lead to the following proposition.

Proposition 22 [24, Theorem 2.4, 2.5] *The energy of $\mu \in \mathcal{M}(\mathbb{R}, n)$ is defined as*

$$I_n^C(\mu) := \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) \, d\mu(x) \, d\mu(y) + 2 \int_{\mathbb{R}} Q(x) \, d\mu(x). \tag{10}$$

Then, there exists a unique minimizer μ_n^ over $\mathcal{M}(\mathbb{R}, n)$ of $I_n^C(\mu)$ with a compact support and μ_n^* is also an optimal solution of (CT). Furthermore, if we define*

$$\begin{aligned} F_{K,Q}^C(n) &:= I_n^C(\mu_n^*) - \int_{\mathbb{R}} Q(x) \, d\mu_n^*(x) \\ & \left(= \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) \, d\mu_n^*(x) \, d\mu_n^*(y) + \int_{\mathbb{R}} Q(x) \, d\mu_n^*(x) \right), \end{aligned} \tag{11}$$

the optimal value of (CT) coincides with $\frac{F_{K,Q}^C(n)}{n}$.

Following this proposition, Tanaka & Sugihara [24] considered a discrete counterpart of $I_n^C(\mu)$ and $F_{K,Q}^C$, which are defined for $a = \{a_i\}_{i=1}^n$ ($a_1 < \dots < a_n$)

as

$$I_{K,Q}^D(a) := \sum_{i \neq j} K(a_i - a_j) + \frac{2(n-1)}{n} \sum_{i=1}^n Q(a_i), \tag{12}$$

$$F_{K,Q}^D(n) := I_{K,Q}^D(a^*) - \frac{n-1}{n} \sum_{i=1}^n Q(a_i^*), \tag{13}$$

where $a^* = \{a_i^*\}_{i=1}^n$ is the unique minimizer of $I_{K,Q}^D(a)$, which certainly exists according to Theorem 3.3 in [24]. We can easily obtain a^* numerically as it is a solution of the convex programming and it is known to satisfy [24, Theorem 4.1]

$$\sup_{\|f\| \leq 1, x \in \mathbb{R}} |f(x) - L_n[a^*; f](x)| \leq \exp\left(-\frac{F_{K,Q}^D(n)}{n-1}\right). \tag{14}$$

Then $E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w))$ is evaluated as [24, Remark 4.2]

$$\exp\left(-\frac{F_{K,Q}^C(n)}{n}\right) \leq E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w)) \leq \exp\left(-\frac{F_{K,Q}^D(n)}{n-1}\right).$$

Indeed, the left inequality holds true by (9) and Proposition 22 and the right inequality follows from (14). By this evaluation, we can consider $L_n[a^*; f](x)$ as a nearly optimal approximation formula if $F_{K,Q}^C(n)/n$ and $F_{K,Q}^D(n)/(n-1)$ are sufficiently close.

2.3 Main results

In this paper, we demonstrate the following two theorems. The first and second theorems, respectively, correspond to (1) and (2) in Sect. 1.

Theorem 23 *For $n \geq 2$, the following holds true:*

$$\frac{F_{K,Q}^D(n)}{n-1} \leq \frac{F_{K,Q}^C(n)}{n} \leq \frac{n}{n-1} \left(\frac{2F_{K,Q}^D(n)}{n-1} + (3 + \log 2) \right).$$

Theorem 24 *Suppose w is even on \mathbb{R} . For $\alpha_n > 0$ that satisfies*

$$\frac{2\alpha_n}{\pi \tanh(d)} \frac{Q(\alpha_n)^2 + Q'(\alpha_n)^2}{Q(\alpha_n)} \leq n,$$

we have

$$\frac{F_{K,Q}^C(n)}{n} \geq \frac{Q(\alpha_n)}{2}.$$

Theorem 23 shows the near optimality of the approximation formula $L_n[a^*; f](x)$. In addition, Theorem 24 (combined with Theorem 23) gives an explicit upper bound of $E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w))$. We describe these results by the following theorem.

Theorem 25 *Let w be a weight function and let K and Q be given by (7) and (8), respectively. In addition, let $a^* = \{a_i^*\}_{i=1}^n$ be the unique minimizer of $I_{K, Q}^D(a)$ and let $L_n[a^*; f](x)$ be the formula given by (6) with $a = a^*$. Then, for arbitrary $\varepsilon > 0$, we have*

$$\sup_{\|f\| \leq 1, x \in \mathbb{R}} |f(x) - L_n[a^*; f](x)| \leq \sqrt{2e^3} E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w))^{\frac{1}{2+\varepsilon}}$$

for each sufficiently large n . In addition, we have

$$E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w)) \leq \sqrt{2e^3} \exp\left(-\frac{n-1}{4n} Q(\alpha_n)\right).$$

2.4 Basic ideas to show the main results

The left inequality of Theorem 23 is from Theorems 3.4 and 3.5 in [24]. To prove the right inequality of Theorem 23, we consider the optimization problem

$$\begin{aligned} \text{(P)} \quad & \text{minimize } \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) \, d\mu(x) \, d\mu(y) + 2 \int_{\mathbb{R}} Q(x) \, d\mu(x) \\ & \text{subject to } \mu \in \mathcal{M}(\mathbb{R}, n), \end{aligned}$$

whose solution provides $F_{K, Q}^C(n)$ as shown in Proposition 22. The quantity $F_{K, Q}^D(n)$ is obtained from the optimal solution of a discrete counterpart of (P) given by (12). Then, we construct a feasible solution of (P) given later by (16) from the optimal solution of the discrete counterpart. By using the feasible solution, we bound $F_{K, Q}^C(n)$ from above by using $F_{K, Q}^D(n)$.

To prove Theorem 24, we need a lower bound of the optimal value of (P). However, because (P) is a minimization problem, any concrete feasible solution does not help us. Therefore, we prove that (P) can be regarded as an infinite-dimensional convex quadratic programming, as K is *positive semi-definite in measure* (Definition 41), and take the dual problem [1, 6]. We also show that the dual problem

$$\begin{aligned} \text{(D)} \quad & \text{maximize } - \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) \, d\nu(x) \, d\nu(y) + 2ns \\ & \text{subject to } \nu \text{ is a signed Borel measure} \\ & s - \int_{\mathbb{R}} K(\cdot - y) \, d\nu(y) \leq Q \end{aligned} \tag{15}$$

satisfies the weak and strong duality (Theorem 43), i.e., the optimal value of (D) coincides with that of (P). By this, we can obtain a lower bound for the optimal value of (P), taking concrete ν and s . The practical advantage of taking (D) is that ν can be

a signed measure (though we indeed deal with a little wider class in Sect. 4), which means that we can define ν as some Fourier transform of the symmetric function, without confirming the non-negativity. This solves one of the improper points of the evaluation in [23].

Remark 1 Problem (D) in (15) needs to be more rigorously stated to realize a primal-dual theory for (P) and (D). In Sect. 4, we provide a rigorous form of (D) by introducing a set S_K for ν .

3 Proof of Theorem 23

To prove Theorem 23, we prepare the following lemmas.

Lemma 31 *For arbitrary $t > 0$, the following holds true.*

$$\int_0^1 K(tx) \, dx \leq K(t) + 1.$$

Proof Consider the function $g(x) := K(x) + \log\left(\frac{\pi}{4d}x\right)$ defined for $x > 0$. We first prove that $g(x)$ is strictly increasing and satisfies $\lim_{x \searrow 0} g(x) = 0$. Let $h(x) := \exp\left(g\left(\frac{2d}{\pi}x\right)\right)$. Then, we have

$$h(x) = \frac{x}{2 \tanh \frac{x}{2}} = \frac{x(e^x + 1)}{2(e^x - 1)}$$

and

$$h'(x) = \frac{(xe^x + e^x + 1)(e^x - 1) - x(e^x + 1)e^x}{2(e^x - 1)^2} = \frac{e^{2x} - 2xe^x - 1}{2(e^x - 1)^2}.$$

Because $(e^{2x} - 2xe^x - 1)' = 2(e^{2x} - e^x - xe^x) = 2e^x(e^x - 1 - x)$ is valid, we have $h'(x) > 0$ for $x > 0$. Evidently, we also have $\lim_{x \searrow 0} h(x) = 1$. Thus, g satisfies the above properties.

Because g is positive and increasing, $\int_0^1 g(tx) \, dx \leq g(t)$ is valid. Therefore, we have

$$\begin{aligned} \int_0^1 K(tx) \, dx &= \int_0^1 g(tx) \, dx - \int_0^1 \log\left(\frac{\pi}{4d}tx\right) \, dx \\ &\leq g(t) - \log\left(\frac{\pi}{4d}t\right) + 1 = K(t) + 1 \end{aligned}$$

as desired. □

Lemma 32 *For arbitrary $x > 0$, the following holds true.*

$$K\left(\frac{x}{2}\right) \leq K(x) + \log 2.$$

Proof By the definition of K , the assertion follows from the fact that $\tanh x \leq 2 \tanh \frac{x}{2}$. □

We can now prove the first theorem.

Proof of Theorem 23 The left inequality is from Theorem 3.4 and 3.5 in [24].

Let us prove the right inequality. Let $a = (a_1, \dots, a_n)$ (with $a_1 < \dots < a_n$) be the minimizer of the discrete energy, satisfying

$$F_{K,Q}^D(n) = \sum_{i \neq j} K(a_i - a_j) + \frac{n-1}{n} \sum_{i=1}^n Q(a_i).$$

Let μ be a measure with the density function p defined by

$$p(x) = \begin{cases} \frac{n}{(n-1)(a_{i+1}-a_i)} & (x \in [a_i, a_{i+1}), i = 1, \dots, n-1), \\ 0 & (\text{otherwise}). \end{cases} \tag{16}$$

Then, we have

$$F_{K,Q}^C(n) \leq I_n^C(\mu_n^*) \leq I_n^C(\mu). \tag{17}$$

In the following, we obtain an upper bound of $I_n^C(\mu)$. First, we evaluate $\int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) d\mu(x) d\mu(y)$. For $1 \leq k \leq n-1$ and $y \in [a_k, a_{k+1})$, we have

$$\begin{aligned} \int_{\mathbb{R}} K(x-y) d\mu(x) &= \int_{\mathbb{R}} K(x-y) p(x) dx \\ &= \sum_{i=1}^{n-1} \frac{n}{(n-1)(a_{i+1}-a_i)} \int_{a_i}^{a_{i+1}} K(x-y) dx \\ &= \frac{n}{n-1} \sum_{i=1}^{n-1} \int_0^1 K(a_i + (a_{i+1}-a_i)z - y) dz. \end{aligned}$$

Here, because $y \in [a_k, a_{k+1})$, for $i \notin \{k-1, k, k+1\}$, the convexity and monotonicity of K over $(-\infty, 0)$ or $(0, \infty)$ shows that

$$\int_0^1 K(a_i + (a_{i+1}-a_i)z - y) dz \leq \begin{cases} \frac{1}{2} (K(a_i - a_k) + K(a_{i+1} - a_k)) & (i \leq k-2), \\ \frac{1}{2} (K(a_i - a_{k+1}) + K(a_{i+1} - a_{k+1})) & (i \geq k+2). \end{cases}$$

Therefore, by considering that K is non-negative, we have

$$\begin{aligned} \sum_{i \neq k-1, k, k+1} \int_0^1 K(a_i + (a_{i+1}-a_i)z - y) dz &\leq \sum_{j \leq k-2} K(a_j - a_k) + \sum_{j \geq k+3} K(a_j - a_{k+1}) \\ &\quad + \frac{1}{2} (K(a_{k-1} - a_k) + K(a_{k+2} - a_{k+1})). \end{aligned} \tag{18}$$

Here, the terms that include an index of a outside the domain $\{1, \dots, n\}$ are void. Next, we consider the cases $i = k \pm 1$. If $k - 1 \geq 1$ is valid, we have

$$\begin{aligned} \int_0^1 K(a_{k-1} + (a_k - a_{k-1})z - y) dz &\leq \int_0^1 K(a_{k-1} + (a_k - a_{k-1})z - a_k) dz \\ &= \int_0^1 K((a_k - a_{k-1})w) dw \\ &\leq K(a_k - a_{k-1}) + 1 = K(a_{k-1} - a_k) + 1. \end{aligned} \tag{19}$$

Similarly, if $k + 2 \leq n$ is valid, we have, by Lemma 31,

$$\int_0^1 K(a_{k+1} + (a_{k+2} - a_{k+1})z - y) dz \leq K(a_{k+2} - a_{k+1}) + 1. \tag{20}$$

Finally, we deal with the case $i = k$. We show that the integral

$$L_k(y) := \int_0^1 K(a_k + (a_{k+1} - a_k)z - y) dz$$

is maximized at $y = \frac{a_k + a_{k+1}}{2}$ (over $y \in [a_k, a_{k+1})$). If we define $t := \frac{y - a_k}{a_{k+1} - a_k}$ ($t \in [0, 1)$), the following holds true.

$$L_k(y) = \int_0^t K((a_{k+1} - a_k)w) dw + \int_0^{1-t} K((a_{k+1} - a_k)w) dw.$$

For $t < \frac{1}{2}$, we have

$$\begin{aligned} &L_k\left(\frac{a_k + a_{k+1}}{2}\right) - L_k(y) \\ &= \int_t^{\frac{1}{2}} K((a_{k+1} - a_k)w) dw - \int_{\frac{1}{2}}^{1-t} K((a_{k+1} - a_k)w) dw \\ &= \int_0^{\frac{1}{2}-t} \left(K((a_{k+1} - a_k)(t + w)) - K\left((a_{k+1} - a_k)\left(\frac{1}{2} + w\right)\right) \right) dw > 0. \end{aligned}$$

By symmetry, $L_k(y) < L_k\left(\frac{a_k+a_{k+1}}{2}\right)$ is valid for $t > \frac{1}{2}$. Therefore, by Lemma 31 and 32,

$$\begin{aligned} & \int_0^1 K(a_k + (a_{k+1} - a_k)z - y) \, dz \\ & \leq L_k\left(\frac{1}{2}\right) = 2 \int_0^{\frac{1}{2}} K((a_{k+1} - a_k)w) \, dw = \int_0^1 K\left(\frac{a_{k+1} - a_k}{2}v\right) \, dv \\ & \leq K\left(\frac{a_{k+1} - a_k}{2}\right) + 1 \leq K(a_{k+1} - a_k) + 1 + \log 2 \end{aligned} \tag{21}$$

By (18–21), we have the bound

$$\begin{aligned} & \left(\frac{n-1}{n}\right)^2 \int_{a_k}^{a_{k+1}} \int_{\mathbb{R}} K(x-y) \, d\mu(x) \, d\mu(y) \leq \frac{n-1}{n} \sup_{y \in \{a_k, a_{k+1}\}} \int_{\mathbb{R}} K(x-y) \, d\mu(x) \\ & \leq \sum_{j \leq k-2} K(a_j - a_k) + \sum_{j \geq k+3} K(a_j - a_{k+1}) + 3 + \log 2 \\ & \quad + \frac{3}{2}K(a_{k-1} - a_k) + \frac{1}{2}K(a_k - a_{k+1}) + \frac{1}{2}K(a_{k+1} - a_k) + \frac{3}{2}K(a_{k+2} - a_{k+1}). \end{aligned}$$

Considering the sum of the right-hand side with respect to $k = 1, \dots, n-1$, the coefficient of each $K(a_i - a_j)$ with $|i - j| \geq 2$ is at most 1, and that of $K(a_i - a_j)$ with $|i - j| = 1$ is at most 2 ($= \frac{1}{2} + \frac{3}{2}$), where we have distinguished $K(a_i - a_j)$ from $K(a_j - a_i)$. Therefore, we have

$$\left(\frac{n-1}{n}\right)^2 \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) \, d\mu(x) \, d\mu(y) \leq 2 \sum_{i \neq j} K(a_i - a_j) + (n-1)(3 + \log 2). \tag{22}$$

Let us now evaluate the second term of $I_n^C(\mu)$, i.e., $\int_{\mathbb{R}} Q(x) \, d\mu(x)$. By the convexity of Q , we have

$$\begin{aligned} \int_{\mathbb{R}} Q(x) \, d\mu(x) &= \frac{n}{n-1} \sum_{i=1}^{n-1} \int_0^1 Q(a_i + (a_{i+1} - a_i)z) \, dz \\ &\leq \frac{n}{n-1} \sum_{i=1}^{n-1} \max\{Q(a_i), Q(a_{i+1})\}. \end{aligned}$$

To estimate the sum, we consider the following two cases:

1. Q is not monotone in $[a_1, a_n]$,
2. Q is monotone in $[a_1, a_n]$.

In the former case, the unique minimizer q^* of Q on \mathbb{R} exists in $[a_1, a_n]$ because of the strict convexity. Then, by the strict convexity of Q , we have

$$\max\{Q(a_i), Q(a_{i+1})\} = \begin{cases} Q(a_i) & (a_i, a_{i+1} < q^*), \\ Q(a_{i+1}) & (q^* < a_i, a_{i+1}), \\ Q(a_{j^*}) & (q^* \in [a_i, a_{i+1}]), \end{cases}$$

where

$$j^* \in \operatorname{argmax}_{j \in \{i, i+1\}} Q(a_j).$$

Therefore

$$\sum_{i=1}^{n-1} \max\{Q(a_i), Q(a_{i+1})\} = \sum_{i \in \{1, \dots, n\} \setminus \{k\}} Q(a_i) \tag{23}$$

holds for some $k \in \{1, \dots, n\}$. In the latter case (case 2 above), we have equality (23) for $k = 1$ or $k = n$. Therefore, in both cases, the following holds true:

$$\int_{\mathbb{R}} Q(x) \, d\mu(x) \leq \frac{n}{n-1} \sum_{i=1}^n Q(a_i). \tag{24}$$

Combining (22) and (24), we obtain

$$\begin{aligned} I_n^C(\mu) &= \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) \, d\mu(x) \, d\mu(y) + 2 \int_{\mathbb{R}} Q(x) \, d\mu(x) \\ &\leq 2 \left(\frac{n}{n-1}\right)^2 \sum_{i \neq j} K(a_i - a_j) + \frac{2n}{n-1} \sum_{i=1}^n Q(a_i) + \frac{n^2}{n-1} (3 + \log 2) \\ &= 2 \left(\frac{n}{n-1}\right)^2 \left(\sum_{i \neq j} K(a_i - a_j) + \frac{n-1}{n} \sum_{i=1}^n Q(a_i) \right) + \frac{n^2}{n-1} (3 + \log 2) \\ &= 2 \left(\frac{n}{n-1}\right)^2 F_{K,Q}^D(n) + \frac{n^2}{n-1} (3 + \log 2). \end{aligned}$$

Now, using (17), we reach the conclusion. □

4 Duality theorem for convex programming of measures

The following definition is a variant of the existing definitions of positive definite kernel [4, 17, 20].

Definition 41 Let X be a topological space. A non-negative measurable function $k : X \times X \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ is called *positive semi-definite in measure* if it satisfies

$$\begin{aligned} \int_X \int_X k(x, y) \, d\mu(x) \, d\mu(y) + \int_X \int_X k(x, y) \, d\nu(x) \, d\nu(y) \\ \geq \int_X \int_X k(x, y) \, d\mu(x) \, d\nu(y) \\ + \int_X \int_X k(x, y) \, d\nu(x) \, d\mu(y) \end{aligned} \tag{25}$$

for arbitrary (positive) σ -finite Borel measures μ, ν on X .

Remark 2 Let k be positive semi-definite in measure. Considering the Hahn-Jordan decomposition of a signed measure, we have

$$\int_X \int_X k(x, y) \, d|\mu|(x) \, d|\mu|(y) < \infty \implies \int_X \int_X k(x, y) \, d\mu(x) \, d\mu(y) \geq 0$$

for an arbitrary signed Borel measure μ on X with $|\mu|$ being σ -finite, where $|\mu|$ denotes the total variation of μ . This is the generalization of the ordinary positive semi-definiteness. Notice that this non-negativity holds for a wider class of “measure”. Indeed, if we define

$$\mathcal{S}_k := \left\{ (\mu_+, \mu_-) \mid \begin{array}{l} \mu_+ \text{ and } \mu_- \text{ are } \sigma\text{-finite Borel measures} \\ \int_X \int_X k(x, y) \, d\mu_+(x) \, d\mu_+(y), \int_X \int_X k(x, y) \, d\mu_-(x) \, d\mu_-(y) < \infty \end{array} \right\}$$

and for each $\nu = (\nu_+, \nu_-) \in \mathcal{S}_k$ define

$$\begin{aligned} \int_X \int_X k(x, y) \, d\nu(x) \, d\nu(y) := \int_X \int_X k(x, y) \, d\nu_+(x) \, d\nu_+(y) + \int_X \int_X k(x, y) \, d\nu_-(x) \, d\nu_-(y) \\ - \int_X \int_X k(x, y) \, d\nu_+(x) \, d\nu_-(y) - \int_X \int_X k(x, y) \, d\nu_-(x) \, d\nu_+(y), \end{aligned}$$

then this integral is well-defined and the generalization of quadratic forms for ordinary signed measures. We formally write $\nu = \nu_+ - \nu_-$ in such a situation, and call it also the Hahn-Jordan decomposition of ν .

Lemma 42 Let $K : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ be an even function. If $K \in L^1(\mathbb{R})$ and K is convex on $[0, \infty)$, and satisfies $\lim_{x \searrow 0} K(x) = K(0)$, then $K(x - y)$ is positive semi-definite in measure.

Proof Because K is integrable and convex, K is continuous over $(0, \infty)$ and $\lim_{x \rightarrow \infty} K(x) = 0$ holds true. If $K(0) < \infty$, K becomes continuous and this type of function is called *Pólya-type*. Pólya-type functions are known to be a characteristic function of a positive bounded Borel measure, i.e., there exists a positive bounded

measure α on \mathbb{R} such that

$$K(x) = \int_{\mathbb{R}} e^{-i\omega x} d\alpha(\omega) \tag{26}$$

is valid [4, 12]. Let μ be a signed Borel measure with $\int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) d\mu(x) d\mu(y)$ being finite and $|\mu|$ being σ -finite. Then, we can take a sequence of increasing Borel sets $A_1 \subset A_2 \subset \dots \rightarrow \mathbb{R}$ satisfying $|\mu|(A_k) < \infty$ for all k . Let $\mu = \mu_+ - \mu_-$ be the Hahn-Jordan decomposition and $\mu_+^k := \mu_+(A_k \cap \cdot)$, $\mu_-^k := \mu_-(A_k \cap \cdot)$. For each k , by Fubini's theorem and (26), we have

$$\int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) d(\mu_+^k - \mu_-^k)(x) d(\mu_+^k - \mu_-^k)(y) = \int_{\mathbb{R}} \left| \int_{\mathbb{R}} e^{-ikx} d(\mu_+^k - \mu_-^k) \right|^2 d\alpha(\omega) \geq 0.$$

This can be rewritten as

$$\begin{aligned} & \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) d\mu_+^k(x) d\mu_+^k(y) + \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) d\mu_-^k(x) d\mu_-^k(y) \\ & \geq 2 \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) d\mu_+^k(x) d\mu_-^k(y). \end{aligned} \tag{27}$$

The integrals in (27) are given by integrands that are monotone increasing with respect to k . Indeed, the first term of the left-hand side is written in the form

$$\int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) d\mu_+^k(x) d\mu_+^k(y) = \int_{\mathbb{R}} \int_{\mathbb{R}} 1_{A_k \times A_k}(x, y) K(x - y) d\mu_+(x) d\mu_+(y)$$

and its integrand $1_{A_k \times A_k}(x, y)K(x - y)$ is monotone increasing with respect to k because $A_1 \subset A_2 \subset \dots$. Similar arguments can be applied to the other terms. Therefore we get the desired inequality by letting $k \rightarrow \infty$ and using the monotone convergence theorem in (27).

Let us consider the case $K(0) = \infty$. In this case, K is continuous on $(0, \infty)$ and has a limit $\lim_{x \searrow 0} K(x)$. For any $\varepsilon > 0$, define

$$K_\varepsilon(x) := \frac{1}{\varepsilon} \int_0^\varepsilon K(|x| + z) dz, \quad x \in \mathbb{R}.$$

Then, by $K \in L^1(\mathbb{R})$, K is bounded everywhere by $\varepsilon^{-1} \|K\|_{L^1}$. Moreover, K_ε is still convex, such that $K_\varepsilon(x - y)$ is positive semi-definite in measure. Now, the continuity of K leads to

$$K_\varepsilon(x) = \int_0^1 K(|x| + \varepsilon z) dz \nearrow K(|x|) = K(x) \quad (\varepsilon \searrow 0)$$

by the monotone convergence theorem. Applying the monotone convergence theorem to both sides of (25) with $K = K_\varepsilon$, we obtain the conclusion. \square

The function $K = -\log \left| \tanh \left(\frac{\pi}{4d} \cdot \right) \right|$ satisfies the condition of Lemma 42. Thus, we can observe the optimization problem

$$(P) \quad \begin{array}{l} \text{minimize} \quad \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) \, d\mu(x) \, d\mu(y) + 2 \int_{\mathbb{R}} Q(x) \, d\mu(x) \\ \text{subject to} \quad \mu \in \mathcal{M}(\mathbb{R}, n) \end{array}$$

as convex quadratic programming. We can analogously make the dual problem to the finite-dimensional case in [1], as

$$(D) \quad \begin{array}{l} \text{maximize} \quad - \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) \, d\nu(x) \, d\nu(y) + 2ns \\ \text{subject to} \quad \nu \in \mathcal{S}_K, \quad s - \int_{\mathbb{R}} K(\cdot - y) \, d\nu(y) \leq Q. \end{array}$$

Note that this is a rigorous version of problem (D) in (15). It should be noted here that we have not justified (D) as a formal (topologically) dual problem. There are arguments limited to the optimization of Radon measure over compact space [10, 11, 27]. While they are on quadratic programming problems, there exist more general theories on duality, such as [3], von Neumann's minimax theorem [8, 16] and Fenchel-Rockafellar duality theorem [13, 26]. However, as it is essential that our duality can treat infinite measure ν with unbounded support (we indeed later use such a measure as a dual feasible solution), it is difficult to just apply existing studies and check all the conditions for (D) to be a topologically dual problem. Therefore, we here do not go deeper in this aspect, but just prove the assertion of Theorem 43. This assertion is sufficient to derive a lower bound of the optimal value of (P), which is our objective.

In the following, we demonstrate that the weak duality and strong duality are still valid in this infinite-dimensional primal-dual pair. It should be noted that $s = 0$, $\nu \equiv 0$ is a trivial feasible solution of (D) such that there exists an optimal value of (D).

Theorem 43 *The optimal value of (D) is equal to the optimal value of (P).*

Proof First, we present the weak duality. Let μ and (ν, s) be feasible solutions of (P) and (D), respectively, and $\nu = \nu_+ - \nu_-$ be the Hahn-Jordan decomposition. If we write $\langle \alpha, \beta \rangle_K := \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) \, d\alpha(x) \, d\beta(y)$ for measures α and β ,

$$\langle \nu, \nu \rangle_K = \langle \nu_+, \nu_+ \rangle_K + \langle \nu_-, \nu_- \rangle_K - 2\langle \nu_+, \nu_- \rangle_K$$

holds true. Because $\langle \mu, \mu \rangle_K, \langle v_+, v_+ \rangle_K, \langle v_-, v_- \rangle_K < \infty$, we have $\langle \mu, v_+ \rangle_K, \langle \mu, v_- \rangle_K, \langle v_+, v_- \rangle_K < \infty$ by K 's positive semi-definiteness in measure. Therefore, we have

$$\begin{aligned} & \left(\int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) \, d\mu(x) \, d\mu(y) + 2 \int_{\mathbb{R}} Q(x) \, d\mu(x) \right) \\ & \quad - \left(- \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) \, dv(x) \, dv(y) + 2ns \right) \\ & = \langle \mu, \mu \rangle_K + \langle v, v \rangle_K + 2 \int_{\mathbb{R}} (Q(x) - s) \, d\mu(x) \\ & \geq \langle \mu, \mu \rangle_K + (\langle v_+, v_+ \rangle_K + \langle v_-, v_- \rangle_K - 2\langle v_+, v_- \rangle_K) \\ & \quad + 2 \int_{\mathbb{R}} \left(- \int_{\mathbb{R}} K(x - y) \, dv(y) \right) \, d\mu(x) \\ & = \langle \mu, \mu \rangle_K + \langle v_+, v_+ \rangle_K + \langle v_-, v_- \rangle_K - 2\langle v_+, v_- \rangle_K - 2\langle \mu, v_+ \rangle_K + 2\langle \mu, v_- \rangle_K \\ & = \langle \mu + v_-, \mu + v_- \rangle_K + \langle v_+, v_+ \rangle_K - 2\langle \mu + v_-, v_+ \rangle_K \geq 0 \end{aligned}$$

by the positive semi-definiteness in measure. This indicates the weak duality. Note that we have the last inequality above by replacing μ and v in (25) in Definition 41 with $\mu + v_+$ and v_+ , respectively.

To prove the strong duality, we construct the optimal solution of (D) using that of (P). By Theorem 2.4 in [24], μ^* , the optimal solution of (P), satisfies

$$\int_{\mathbb{R}} K(x - y) \, d\mu^*(y) + Q(x) \geq \frac{F_{K,Q}^C(n)}{n} \tag{28}$$

for all $x \in \mathbb{R}$. Now, μ^* and $n^{-1}F_{K,Q}^C(n)$ is a feasible solution for (D). Moreover, the equation that we obtain by replacing the inequality of (28) with an equality is valid on the support of μ^* . Therefore we have

$$\begin{aligned} & - \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) \, d\mu^*(x) \, d\mu^*(y) + 2n \frac{F_{K,Q}^C(n)}{n} \\ & = - \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) \, d\mu^*(x) \, d\mu^*(y) + 2 \int_{\mathbb{R}} \left(Q(x) + \int_{\mathbb{R}} K(x - y) \, d\mu^*(y) \right) \, d\mu^*(x) \\ & = \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y) \, d\mu^*(x) \, d\mu^*(y) + 2 \int_{\mathbb{R}} Q(x) \, d\mu^*(x). \end{aligned}$$

This shows the strong duality. □

5 Proof of Theorem 24

We can now give a lower bound of $F_{K,Q}^C(n)$ by using the dual problem (D) and prove Theorem 24. Let $\alpha > 0$ be a constant and f be the inverse Fourier transform of

$$(\mathcal{F}[f](\omega) =) \frac{\omega}{\pi \tanh(d\omega)} \int_{-\alpha}^{\alpha} (Q(\alpha) - Q(x)) e^{-i\omega x} dx$$

Along with this, f is L^2 -integrable by Theorem 4.4 in [23]. Here, the Fourier transform of a function $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is defined by

$$\mathcal{F}[g](\omega) := \int_{\mathbb{R}} g(x) e^{-i\omega x} dx$$

and for the whole space $L^2(\mathbb{R})$, $\mathcal{F}[\cdot]$ is defined as the continuous extension of $\mathcal{F}[\cdot]|_{L^1 \cap L^2}$. Because $Q(x)$ is even by the assumption, f is an inverse Fourier transform of an even real function, so that f itself is an even real function. Then, the formula [9, p.43, 7.112]

$$\mathcal{F} \left[\log \left| \tanh \left(\frac{\pi}{4d} \cdot \right) \right| \right] (\omega) = -\frac{\pi}{\omega} \tanh(d\omega)$$

leads to the (almost everywhere) equation

$$\mathcal{F} \left[\int_{\mathbb{R}} K(x - y) f(y) dy \right] (\omega) = \mathcal{F}[K](\omega) \cdot \mathcal{F}[f](\omega) = \int_{-\alpha}^{\alpha} (Q(\alpha) - Q(x)) e^{-i\omega x} dx, \tag{29}$$

where $K \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and $f \in L^2(\mathbb{R})$ are used for the justification of the first equality. The former statement $K \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ follows from

$$\begin{aligned} \int_{-\infty}^{\infty} \left(-\log \tanh \left(\frac{\pi}{4d} x \right) \right) dx &< \infty \quad \text{and} \\ \int_{-\infty}^{\infty} \left(-\log \tanh \left(\frac{\pi}{4d} x \right) \right)^2 dx &< \infty. \end{aligned}$$

The integrability of $K(x - \cdot) f(\cdot)$ comes from $K, f \in L^2(\mathbb{R})$. Indeed, we have

$$\begin{aligned} \left| \int_{\mathbb{R}} K(x - y) f(y) dy \right|^2 &\leq \left(\int_{\mathbb{R}} |K(x - y)|^{1/2} |K(x - y)|^{1/2} |f(y)| dy \right)^2 \\ &\leq \int_{\mathbb{R}} |K(x - y)| dy \int_{\mathbb{R}} |K(x - y)| |f(y)|^2 dy \\ &= \|K\|_{L^1} \int_{\mathbb{R}} |K(x - y)| |f(y)|^2 dy, \end{aligned}$$

where the Cauchy-Schwarz inequality is used on the second inequality. Therefore the integrability is shown as follows:

$$\begin{aligned} \left\| \int_{\mathbb{R}} K(\cdot - y)f(y) \, dy \right\|_{L^2}^2 &\leq \|K\|_{L^1} \int_{\mathbb{R}} \int_{\mathbb{R}} |K(x - y)||f(y)|^2 \, dy \, dx \\ &= \|K\|_{L^1} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |K(x - y)| \, dx \right) |f(y)|^2 \, dy \\ &= \|K\|_{L^1}^2 \|f\|_{L^2}^2 < \infty, \end{aligned}$$

where the Fubini theorem is used on the first equality. Considering the inverse Fourier transform of (29), we also have

$$\int_{\mathbb{R}} K(x - y)f(y) \, dy = 1_{[-\alpha, \alpha]}(x)(Q(\alpha) - Q(x)).$$

It should be noted that $f(x) \, dx \in \mathcal{S}_K$ follows from the inequality

$$\int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y)|f(x)f(y)| \, dx \, dy \leq \|K * f\|_{L^2} \|f\|_{L^2} \leq \|K\|_{L^1} \|f\|_{L^2}^2 < \infty.$$

These two relations imply that $(f(x) \, dx, Q(\alpha))$ is a feasible solution of (D). We can now evaluate the value of the objective function of (D). Let us define

$$F(\alpha) := - \int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y)f(x)f(y) \, dx \, dy + 2nQ(\alpha). \tag{30}$$

Because the first term can be considered as the inner product of $K * f$ and f in $L^2(\mathbb{R})$, it can be computed through the Fourier transform as

$$\begin{aligned} &\int_{\mathbb{R}} \int_{\mathbb{R}} K(x - y)f(x)f(y) \, dx \, dy \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\frac{\omega}{\pi \tanh(d\omega)} \int_{-\alpha}^{\alpha} (Q(\alpha) - Q(x)) e^{-i\omega x} \, dx \int_{-\alpha}^{\alpha} \overline{(Q(\alpha) - Q(x)) e^{-i\omega x}} \, dx \right) \, d\omega \\ &= \frac{1}{2\pi^2} \int_{\mathbb{R}} \frac{\omega}{\tanh(d\omega)} \left| \int_{-\alpha}^{\alpha} (Q(\alpha) - Q(x)) e^{-i\omega x} \, dx \right|^2 \, d\omega. \end{aligned} \tag{31}$$

Let $G(\alpha)$ be the value of the right-hand side. $G(\alpha)$ can be decomposed into two parts, which are defined as

$$G_1(\alpha) := \frac{1}{2\pi^2} \int_{-1}^1 \frac{\omega}{\tanh(d\omega)} \left| \int_{-\alpha}^{\alpha} (Q(\alpha) - Q(x)) e^{-i\omega x} \, dx \right|^2 \, d\omega$$

and

$$G_2(\alpha) := \frac{1}{2\pi^2} \int_{[-1, 1]^c} \frac{\omega}{\tanh(d\omega)} \left| \int_{-\alpha}^{\alpha} (Q(\alpha) - Q(x)) e^{-i\omega x} \, dx \right|^2 \, d\omega.$$

We first evaluate G_1 . Because the function $\omega / \tanh(d\omega)$ is monotonically increasing in $[0, \infty)$ (see the proof of Lemma 31), we have

$$\begin{aligned} G_1(\alpha) &\leq \frac{1}{\pi \tanh(d)} \cdot \frac{1}{2\pi} \int_{\mathbb{R}} \left| \int_{-\alpha}^{\alpha} (Q(\alpha) - Q(x)) e^{-i\omega x} dx \right|^2 d\omega \\ &= \frac{1}{\pi \tanh(d)} \|1_{[-\alpha, \alpha]}(x)(Q(\alpha) - Q(x))\|_{L^2}^2 \\ &\leq \frac{2}{\pi \tanh(d)} \alpha Q(\alpha)^2. \end{aligned} \quad (32)$$

Next, we similarly evaluate G_2 . By integration by parts, we get

$$\omega \int_{-\alpha}^{\alpha} (Q(\alpha) - Q(x)) e^{-i\omega x} dx = -\frac{1}{i} \int_{-\alpha}^{\alpha} Q'(x) e^{-i\omega x} dx.$$

Thus, we have

$$\begin{aligned} G_2(\alpha) &= \frac{1}{2\pi^2} \int_{[-1, 1]^c} \frac{1}{\omega \tanh(d\omega)} \left| \int_{-\alpha}^{\alpha} Q'(x) e^{-i\omega x} dx \right|^2 d\omega \\ &\leq \frac{1}{\pi \tanh(d)} \|1_{[-\alpha, \alpha]}(x) Q'(x)\|_{L^2}^2 \\ &\leq \frac{2}{\pi \tanh(d)} \alpha Q'(\alpha)^2. \end{aligned} \quad (33)$$

Finally, we reach the evaluation

$$\begin{aligned} G(\alpha) &\leq \frac{2\alpha}{\pi \tanh(d)} \left(Q(\alpha)^2 + Q'(\alpha)^2 \right), \\ F(\alpha) &\geq 2n Q(\alpha) - \frac{2\alpha}{\pi \tanh(d)} \left(Q(\alpha)^2 + Q'(\alpha)^2 \right). \end{aligned}$$

By letting α_n satisfy

$$\frac{2\alpha_n}{\pi \tanh(d)} \frac{Q(\alpha_n)^2 + Q'(\alpha_n)^2}{Q(\alpha_n)} \leq n,$$

we get $nQ(\alpha_n)$ as a lower bound for the optimal value of (P). For such α_n , we finally have

$$nQ(\alpha_n) \leq I_{K, Q}^C(\mu^*) \leq 2F_{K, Q}^C(n)$$

and this is equivalent to the assertion of Theorem 24.

6 Examples of convergence rates for several $Q(x)$'s

Although the asymptotic rates given in [23, Section 4.3] are derived through mathematically informal arguments, we here demonstrate that those rates roughly coincide with the bound in Theorem 24.

Example 61 (The case w is a single exponential) Consider the case

$$w(x) = \exp(-(\beta|x|)^\rho), \quad Q(x) = (\beta|x|)^\rho,$$

for $\beta > 0$ and $\rho \geq 1$. In this case, for a sufficiently large α (satisfying $\alpha \geq \rho$), we have

$$\frac{2\alpha}{\pi \tanh(d)} \frac{Q(\alpha)^2 + Q'(\alpha)^2}{Q(\alpha)} = \frac{2\alpha}{\pi \tanh(d)} \frac{(\beta\alpha)^{2\rho} + (\beta\rho)^2(\beta\alpha)^{2(\rho-1)}}{(\beta\alpha)^\rho} \leq \frac{4\beta^\rho \alpha^{\rho+1}}{\pi \tanh(d)}$$

and α_n can be taken as

$$\alpha_n = \left(\frac{\pi \tanh(d)}{4\beta^\rho} n \right)^{\frac{1}{\rho+1}},$$

$$\frac{Q(\alpha_n)}{2} = \frac{1}{2} \beta^\rho \left(\frac{\pi \tanh(d)}{4\beta^\rho} n \right)^{\frac{\rho}{\rho+1}} \quad \left(= \Theta \left(\beta^{\frac{\rho}{\rho+1}} n^{\frac{\rho}{\rho+1}} \right) \right), \tag{34}$$

for sufficiently large n . This rate roughly coincides with (4.37) in [23].

Example 62 (The case w is a double exponential) Consider the case

$$w(x) = \exp(-\beta \exp(\gamma|x|)), \quad Q(x) = \beta \exp(\gamma|x|),$$

for $\beta, \gamma > 0$. In this case,

$$\frac{2\alpha}{\pi \tanh(d)} \frac{Q(\alpha)^2 + Q'(\alpha)^2}{Q(\alpha)} = \frac{2\alpha\beta(1 + \gamma^2) \exp(\gamma\alpha)}{\pi \tanh(d)}$$

is valid. Let $\alpha_n > 0$ satisfy that the right-hand side is equal to n . Then, we have

$$\gamma\alpha_n = W \left(\frac{\pi \tanh(d)\gamma}{2\beta(1 + \gamma^2)} n \right) \quad \left(\sim \log \left(\frac{\gamma}{\beta(1 + \gamma^2)} n \right) \right),$$

where W is Lambert's W function, i.e., the inverse of $x \mapsto xe^x$. Using this, we get

$$\begin{aligned} \frac{Q(\alpha_n)}{2} &= \frac{\beta}{2\gamma\alpha_n} \cdot \gamma\alpha_n \exp(\gamma\alpha_n) = \frac{\beta}{2\gamma\alpha_n} \frac{\pi \tanh(d)\gamma}{2\beta(1 + \gamma^2)} n = \frac{\pi \tanh(d)}{4(1 + \gamma^2)} \frac{n}{\alpha_n} \\ &= \frac{\pi \tanh(d)\gamma}{4(1 + \gamma^2)} \frac{n}{W \left(\frac{\pi \tanh(d)\gamma}{2\beta(1 + \gamma^2)} n \right)} \quad \left(\sim \frac{\pi \tanh(d)\gamma}{4(1 + \gamma^2)} \frac{n}{\log \left(\frac{\gamma}{\beta(1 + \gamma^2)} n \right)} \right). \end{aligned} \tag{35}$$

This rate roughly coincides with the asymptotic order (4.44) in [23] for each fixed constant γ .

Remark 3 We choose the weight functions in Examples 61 and 62 for simplicity although they are not (necessarily) analytic in the strip region \mathcal{D}_d for any $d > 0$. This is because we just need their asymptotic properties for finding α_n .

7 Conclusion

In this study, we analyzed the approximation method proposed by [24] over weighted Hardy spaces $\mathbb{H}^\infty(\mathcal{D}_d, w)$. We provided (1) proof of the fact that the approximation formulas are nearly optimal from the viewpoint of minimum worst-case error $E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d, w))$; and (2) upper bounds of $E_n^{\min}(\mathbb{H}^\infty(\mathcal{D}_d))$ to evaluate the convergence rates of approximation errors with $n \rightarrow \infty$. To obtain (2), we introduced the concept “positive semi-definite in measure” and by using this, provided a lower bound for $F_{K, Q}^C(n)$. We also compared the given bounds with those mentioned in the study by [23], and demonstrated that they have the same convergence rate with $n \rightarrow \infty$.

The new bounds do not indicate that the approximation formulas in [24] are optimal. Another method to bound the error is recently considered by [7], although their bound do not show the optimality, either. We need tighter bounds to show the optimality, which may require more sophisticated analysis. We leave such analysis to future work.

Acknowledgements The authors are grateful to Ryunosuke Oshiro for his comment on signed measures. They also thank the anonymous reviewers for their valuable comments about this paper.

Funding Open access funding provided by The University of Tokyo.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Dorn, W.S.: Duality in quadratic programming. *Q. Appl. Math.* **18**(2), 155–162 (1960)
2. Haber, S.: The tanh rule for numerical integration. *SIAM J. Numer. Anal.* **14**(4), 668–685 (1977)
3. Isii, K.: Inequalities of the types of chebyshev and cramér-rao and mathematical programming. *Ann. Inst. Stat. Math.* **16**(1), 277–293 (1964)
4. Jaming, P., Matolcsi, M., Révész, S.G.: On the extremal rays of the cone of positive, positive definite functions. *J. Fourier Anal. Appl.* **15**(4), 561–582 (2009)
5. Levin, A., Lubinsky, D.: Green equilibrium measures and representations of an external field. *J. Approx. Theory* **113**(2), 298–323 (2001)
6. Luenberger, D.G.: *Optimization by vector space methods*. Wiley, New York (1997)
7. van Meurs, P., Tanaka, K.: Convergence rates for energies of interacting particles whose distribution spreads out as their number increases. *ESAIM: COCV* **29**, 4 (2023). <https://doi.org/10.1051/cocv/2022083>

8. von Neumann, J.: Zur theorie der gesellschaftsspiele. *Math. Ann.* **100**(1), 295–320 (1928)
9. Oberhettinger, F.: *Tables of Fourier transforms and Fourier transforms of distributions*. Springer, New York (1990)
10. Ohtsuka, M.: A generalization of duality theorem in the theory of linear programming. *J. Sci. Hiroshima Univ. Ser. A-I Math.* **30**(1), 31–39 (1966)
11. Ohtsuka, M.: Generalized capacity and duality theorem in linear programming. *J. Sci. Hiroshima Univ. Ser. A-I Math.* **30**(1), 45–56 (1966)
12. Pólya, G.: Remarks on characteristic functions. In: *Proceedings of First Berkeley Conference on Mathematical Statistics and Probability*, pp. 115–123. University of California Press, Berkeley (1949)
13. Rockafellar, R.T.: Extension of fenchel's duality theorem for convex functions. *Duke Math. J.* **33**(1), 81–89 (1966)
14. Saff, E.B., Totik, V.: *Logarithmic potentials with external fields*, vol. 316. Springer, New York (1997)
15. Schwartz, C.: Numerical integration of analytic functions. *J. Comput. Phys.* **4**(1), 19–29 (1969)
16. Sion, M.: On general minimax theorems. *Pac. J. Math.* **8**(1), 171–176 (1958)
17. Sriperumbudur, B.K., Gretton, A., Fukumizu, K., Schölkopf, B., Lanckriet, G.R.: Hilbert space embeddings and metrics on probability measures. *J. Mach. Learn. Res.* **11**, 1517–1561 (2010)
18. Stenger, F.: *Numerical methods based on sinc and analytic functions*, vol. 20. Springer, New York (1993)
19. Stenger, F.: *Handbook of Sinc numerical methods*. CRC Press, Boca Raton (2011)
20. Stewart, J.: Positive definite functions and generalizations, an historical survey. *Rocky Mountain J. Math.* **6**(3), 409–434 (1976). <https://doi.org/10.1216/RMJ-1976-6-3-409>
21. Sugihara, M.: Near optimality of the sinc approximation. *Math. Comput.* **72**(242), 767–786 (2003). <https://doi.org/10.1090/S0025-5718-02-01451-5>
22. Takahasi, H., Mori, M.: Double exponential formulas for numerical integration. *Publ. Res. Inst. Math. Sci.* **9**(3), 721–741 (1974)
23. Tanaka, K., Okayama, T., Sugihara, M.: Potential theoretic approach to design of accurate formulas for function approximation in symmetric weighted hardy spaces. *IMA J. Numer. Anal.* **37**(2), 861–904 (2017)
24. Tanaka, K., Sugihara, M.: Design of accurate formulas for approximating functions in weighted hardy spaces by discrete energy minimization. *IMA J. Numer. Anal.* **39**(4), 1957–1984 (2019)
25. Tanaka, K., Sugihara, M., Murota, K.: Function classes for successful de-sinc approximations. *Math. Comput.* **78**(267), 1553–1571 (2009)
26. Villani, C.: *Topics in optimal transportation*. 58. American Mathematical Society, Providence, Rhode Island (2003)
27. Wu, S.: A cutting plane approach to solving quadratic infinite programs on measure spaces. *J. Global Optim.* **21**(1), 67–87 (2001)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.