Higher Dimensional Hermite-Hadamard Inequality for Semiconvex Functions of Rate $(k_1, k_2, ..., k_n)$ on the Co-ordinates and Optimal Mass Transportation

Ping Chen and Wing-Sum Cheung

August 7, 2020

Abstract

In this paper, we give a new higher dimensional Hermite-Hadamard inequality for a function $f : \prod_{i=1}^{n} [a_i, b_i] \subset \mathbb{R}^n \to \mathbb{R}$ which is semiconvex of rate $(k_1, k_2, ..., k_n)$ on the co-ordinates. This generalizes some existing results on Hermite-Hadamard inequalities of S.S. Dragomir. In addition, we explain the Hermite-Hadamard inequality from the point of view of optimal mass transportation with cost function $c(x, y) := f(y - x) + \sum_{i=1}^{n} \frac{k_i}{2} |x_i - y_i|^2$, where $f(\cdot) : \prod_{i=1}^{n} [a_i, b_i] \to [0, \infty)$ is semiconvex of rate $(k_1, k_2, ..., k_n)$ on the co-ordinates and $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_n) \in \prod_{i=1}^{n} [a_i, b_i]$. Furthermore, by using the higher dimensional Hermite-Hadamard inequality, we compare the transport cost in different transport models on the sphere $\mathbb{S}^2$.

Keywords. Convex functions, Hermite-Hadamard inequality, Optimal mass transportation

1 Introduction

The classical Hermite-Hadamard inequality for convex functions f on $[a, b]$ is usually stated as

$$f(\frac{a+b}{2}) \leq \int_a^b f(x)dx \leq \frac{f(a) + f(b)}{2}. \tag{1.1}$$

which was first published in [4].

Ping Chen: School of Mathematics and Information Technology, Jiangsu Second Normal University, Nanjing 210013, P.R.China.; e-mail: chenping200507@126.com

Wing-Sum Cheung: Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong, P.R. China.e-mail: wscheung@hku.hk

Mathematics Subject Classification (2010): 26B25; 26D15; 49Q20

In the 2-dimensional situation, for any function $f : [a,b] \times [c,d] \to \mathbb{R}$ which is convex on the co-ordinates on $[a,b] \times [c,d]$, Dragomir proved in 2001 the following two-dimensional Hermite-Hadamard inequality (Theorem 1 in [2]):

$$\begin{aligned}
f(\frac{a+b}{2},\frac{c+d}{2}) \leq\ & \frac{1}{2}\left[\frac{1}{b-a}\int_a^b f(x_1,\frac{c+d}{2})dx_1 + \frac{1}{d-c}\int_c^d f(\frac{a+b}{2},x_2)dx_2\right] \\
\leq\ & \frac{1}{(b-a)(c-d)}\int_a^b\int_c^d f(x_1,x_2)dx_1dx_2 \\
\leq\ & \frac{1}{4}\left[\frac{1}{b-a}\int_a^b f(x_1,c)dx_1 + \frac{1}{b-a}\int_a^b f(x_1,d)dx_1\right. \\
& \left. + \frac{1}{d-c}\int_c^d f(a,x_2)dx_2 + \frac{1}{d-c}\int_c^d f(b,x_2)dx_2\right] \\
\leq\ & \frac{f(a,c)+f(a,d)+f(b,c)+f(b,d)}{4} \qquad (1.2)
\end{aligned}$$

Interested readers are also refered to [3] for more details. On the other hand, Hermite-Hadamard Inequality has also been extended to various other contexts, including for example Hermite-Hadamard's type inequalities involving two functions, Hermite-Hadamard inequality for log-convex functions, etc. Details can be found in [7] and [11], respectively.

In this paper, inspired by the Hermite-Hadamard inequality for semiconvex functions $f : [a,b] \to \mathbb{R}$ in [6], in Section 2, we establish a higher dimensional refinement of the Hermite-Hadamard inequality (1.2) for a function $f : \prod_{i=1}^{n}[a_i,b_i] \subset \mathbb{R}^n \to \mathbb{R}$ which is continuous and semiconvex of rate $(k_1,k_2,...,k_n)$ on the co-ordinates. Here $a_i < b_i (i = 1,...,n)$. Dragomir's result (1.2), and the Hermite-Hadamard inequality for 1- and higher dimensional semiconvex functions (see Theorem 2.2 and Remark 2.5, respectively) can all be seen as special cases. In Section 3, we interpret the meaning of the new higher dimensional Hermite-Hadamard inequality from the point of view of optimal mass transportation problems by studying and comparing the optimal transportation costs of five transport models in the hyper-rectangles $\prod_{i=1}^{n}[a_i,b_i]$. Five transport models on the sphere $\mathbb{S}^2$ are also studied. In addition, by making use of the the new higher dimensional Hermite-Hadamard inequality, comparison of the transport costs in such models on the sphere is given.

2 Higher dimensional Hermite-Hadamard inequality for semiconvex functions of rate $(k_1,k_2,...,k_n)$ on the co-ordinates

We first recall some preliminaries on semiconvexity and the one-dimensional Hermite-Hadamard inequality for semiconvex functions $f : [a,b] \to \mathbb{R}$ of rate $k \in \mathbb{R}$.

Definition 2.1. ([5], [6],[8]) A function f defined on a convex set in $\mathbb{R}^n$ is said to be *semiconvex of rate* k if the function

$$h(\cdot) := f(\cdot) + \frac{k}{2}||\cdot||^2$$

is convex for some real constant k. Here $||\cdot||$ is the usual Euclidean norm.

Theorem 2.2. *([6]) If μ is a Borel probability measure on an interval $[a,b]$ with barycenter*

$$b_\mu = \int_a^b x d\mu(x) , \tag{2.1}$$

then for every semiconvex function $f : [a,b] \to \mathbb{R}$ of rate k, we have

$$\begin{aligned} f(b_\mu) &\le \int_a^b f(x) d\mu(x) + \frac{k}{2}\int_a^b |x-b_\mu|^2 d\mu(x) & (2.2)\\ &\le \frac{b-b_\mu}{b-a} f(a) + \frac{b_\mu - a}{b-a} f(b) + \frac{k}{2}(b_\mu - a)(b-b_\mu) . & (2.3)\end{aligned}$$

Motivated by these results, we establish a higher dimensional Hermite-Hadamard inequality for semiconvex function of rate $(k_1, k_2, ..., k_n)$ on the co-ordinates. We start by recalling the following definition.

Definition 2.3. Let $k_1, k_2, ..., k_n \in \mathbb{R}$, a function $f : \prod_{i=1}^{n}[a_i, b_i] \to \mathbb{R}$ is said to be *semiconvex of rate* $(k_1, k_2, ..., k_n)$ *on the co-ordinates* if for all $x_i \in [a_i, b_i]$, the partial maps

$$f_{\hat{x_i}}(u) : [a_i, b_i] \to \mathbb{R}, \quad f_{\hat{x_i}}(u) := f(x_1, ..., x_{i-1}, u, x_{i+1}, ...x_n), \ i = 1, 2, ..., n$$

are semiconvex of rate $k_i, \ i = 1, 2, ..., n$, respectively.

We also fix our notations as follows. For any $i = 1, ..., n$, $b_{\mu_i} = \int_{a_i}^{b_i} x_i d\mu_i(x_i)$ is the barycenter of a Borel probability measure μ_i on an interval $[a_i, b_i]$, respectively,

$$\begin{aligned} &\hat{x_i} = (x_1, ..., x_{i-1}, x_{i+1}, ...x_n) \in \mathbb{R}^{n-1} , \\ &(\hat{x_i}, u) = (x_1, ..., x_{i-1}, u, x_{i+1}, ...x_n) \in \mathbb{R}^n , \\ &b_\mu = (b_{\mu_1}, ..., b_{\mu_n}) \in \mathbb{R}^n , \\ &(\hat{b_{\mu_i}}, u) = (b_{\mu_1}, ..., b_{\mu_{i-1}}, u, b_{\mu_{i+1}}, ...b_{\mu_n}) \in \mathbb{R}^n , \end{aligned}$$

$$\iint\limits_{\prod_{j\ne i}[a_j,b_j]} f(\hat{x_i}, u) d\widehat{\mu_i}(x_i) = \iint\limits_{\prod_{j\ne i}[a_j,b_j]} f(\hat{x_i}, u) d\mu_1(x_1) \otimes ... \otimes \mu_{i-1}(x_{i-1}) \otimes \mu_{i+1}(x_{i+1}) ... \otimes \mu_n(x_n).$$

Theorem 2.4. *If $\mu_i, i = 1, 2..., n$ are Borel probability measures on the intervals $[a_i, b_i], i = 1, 2..., n$, respectively, then for every continuous and semiconvex function $f : \prod_{i=1}^{n}[a_i, b_i] \to \mathbb{R}$ of rate $(k_1, k_2, ..., k_n)$ on the co-ordinates, we have*

$$\begin{aligned} & f(b_{\mu_1}, ..., b_{\mu_n}) & (2.4)\\ &\le \frac{1}{n}\sum_{i=1}^{n}\int_{a_i}^{b_i} f(\widehat{b_{\mu_i}}, x_i) d\mu_i(x_i) + \sum_{i=1}^{n}\frac{k_i}{2n}\int_{a_i}^{b_i}|x_i - b_{\mu_i}|^2 d\mu_i(x_i) & (2.5)\\ &\le \iint\limits_{\prod_{i=1,...,n}[a_i,b_i]} f(x_1, ..., x_n) d\mu_1(x_1) \otimes ... \otimes \mu_n(x_n) + \sum_{i=1}^{n}\frac{k_i}{2}\int_{a_i}^{b_i}|x_i - b_{\mu_i}|^2 d\mu_i(x_i) & (2.6)\end{aligned}$$

$$\leq \sum_{i=1}^{n} \iint\limits_{\prod_{j\neq i}[a_j,b_j]} \frac{b_i - b_{\mu_i}}{n(b_i - a_i)} f(\widehat{x_i}, a_i) d\widehat{\mu_i}(x_i) + \sum_{i=1}^{n} \iint\limits_{\prod_{j\neq i}[a_j,b_j]} \frac{b_{\mu_i} - a_i}{n(b_i - a_i)} f(\widehat{x_i}, b_i) d\widehat{\mu_i}(x_i)$$

$$+ \sum_{i=1}^{n} \frac{(n-1)k_i}{2n} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i) + \sum_{i=1}^{n} \frac{k_i}{2n}(b_{\mu_i} - a_i)(b_i - b_{\mu_i}) \tag{2.7}$$

$$\leq \sum_{\substack{x_i = a_i \ or \ b_i, \\ i=1,2,\ldots,n}} \frac{g(x_1)\ldots g(x_n)}{\prod_{i=1}^{n}(b_i - a_i)} f(x_1, \ldots, x_n) + \sum_{i=1}^{n} \frac{k_i}{2}(b_{\mu_i} - a_i)(b_i - b_{\mu_i}) \ , \tag{2.8}$$

where $g(x_i) = \begin{cases} b_i - b_{\mu_i} & \textit{if } x_i = a_i \\ b_{\mu_i} - a_i & \textit{if } x_i = b_i \end{cases}, i = 1, \ldots, n.$

Proof. It follows from Definition 2.3 that $f_{\hat{x_i}}(u) : [a_i, b_i] \to \mathbb{R},\ i = 1, 2, \ldots, n$ are semiconvex of rate $k_i,\ i = 1, 2, \ldots, n$, respectively.

Applying Theorem 2.2 to $f_{\hat{x_i}}(u),\ i = 1, 2, \ldots, n$, respectively, one has

$$f_{\hat{x_i}}(b_{\mu_i}) \leq \int_{a_i}^{b_i} f_{\hat{x_i}}(x_i) d\mu_i(x_i) + \frac{k_i}{2} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i) \ . \tag{2.9}$$

Taking $\hat{x_i} = \hat{b_{\mu_i}}, i = 1, 2, \ldots, n$ in (2.9), and adding the n resulting inequalities, one has

$$nf(b_{\mu_1}, b_{\mu_2}, \ldots, b_{\mu_n}) \leq \sum_{i=1}^{n} \left[\int_{a_i}^{b_i} f_{\hat{b_{\mu_i}}}(x_i) d\mu_i(x_i) + \frac{k_i}{2} \int_{a_i}^{b_i} |x - b_{\mu_i}|^2 d\mu_i(x_i) \right]$$

and so

$$f(b_{\mu_1}, b_{\mu_2}, \ldots, b_{\mu_n}) \leq \frac{1}{n} \sum_{i=1}^{n} \left[\int_{a_i}^{b_i} f_{\hat{b_{\mu_i}}}(x_i) d\mu_i(x_i) + \frac{k_i}{2} \int_{a_i}^{b_i} |x - b_{\mu_i}|^2 d\mu_i(x_i) \right] \ ,$$

which proves inequality (2.5).

For all $i = 1, 2, \ldots, n$, $f_{\hat{b_{\mu_i}}}(x_i)$ can be seen as a function of b_{μ_1}. By using (2.2), one gets

$$\int_{a_i}^{b_i} f_{\hat{b_{\mu_i}}}(x_i) d\mu_i(x_i) + \frac{k_i}{2} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i)$$

$$\leq \int_{a_i}^{b_i} \left[\int_{a_1}^{b_1} f(x_1, b_{\mu_2}, \ldots, b_{\mu_{i-1}}, x_i, b_{\mu_{i+1}}, \ldots, b_{\mu_n}) d\mu_1(x_1) + \frac{k_1}{2} \int_{a_1}^{b_1} |x_1 - b_{\mu_1}|^2 d\mu_1(x_1) \right] d\mu_i(x_i)$$

$$+ \frac{k_i}{2} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i)$$

$$= \int_{a_i}^{b_i} \left[\int_{a_1}^{b_1} f(x_1, b_{\mu_2}, \ldots, b_{\mu_{i-1}}, x_i, b_{\mu_{i+1}}, \ldots, b_{\mu_n}) d\mu_1(x_1) \right] d\mu_i(x_i)$$

$$+ \frac{k_1}{2} \int_{a_1}^{b_1} |x_1 - b_{\mu_1}|^2 d\mu_1(x_1) + \frac{k_i}{2} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i).$$

Note that $f(x_1, b_{\mu_2}, ..., b_{\mu_{i-1}}, x_i, b_{\mu_{i+1}}, ..., b_{\mu_n})$ can also be seen as a function of b_{μ_2}. Using (2.2) again and repeating the above steps, we get

$$\begin{aligned}
&\int_{a_i}^{b_i} f_{\hat{b_{\mu_i}}}(x_i)d\mu_i(x_i) + \frac{k_i}{2}\int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i) \\
&\leq \iint\limits_{\prod\limits_{i=1,...,n}[a_i,b_i]} f(x_1, ..., x_n)d\mu_1(x_1) \otimes ... \otimes \mu_n(x_n) \\
&\qquad + \sum_{i=1}^{n} \frac{k_i}{2}\int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i).
\end{aligned}$$

Hence

$$\begin{aligned}
&\sum_{i=1}^{n}\left[\int_{a_i}^{b_i} f_{\hat{b_{\mu_i}}}(x_i)d\mu_i(x_i) + \frac{k_i}{2}\int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i)\right] \\
&\leq n \iint\limits_{\prod\limits_{i=1,...,n}[a_i,b_i]} f(x_1, ..., x_n)d\mu_1(x_1) \otimes ... \otimes \mu_n(x_n) \\
&\qquad + n\sum_{i=1}^{n} \frac{k_i}{2}\int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i)\ ,
\end{aligned}$$

which implies inequality (2.6)

We proceed to prove (2.7) by using (2.3). Note that for all $i = 1, 2, ..., n$,

$$\begin{aligned}
&\iint\limits_{\prod\limits_{i=1,...,n}[a_i,b_i]} f(x_1, ..., x_n)d\mu_1(x_1) \otimes ... \otimes \mu_n(x_n) + \frac{k_i}{2}\int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i) \\
&= \iint\limits_{\prod\limits_{i=1,...,n}[a_i,b_i]} [f(x_1, x_2, ..., x_n) + \frac{k_i}{2}|x_i - b_{\mu_i}|^2]d\mu_i(x_i) \otimes \mu_1(x_1) \otimes \mu_2(x_2) \otimes ... \otimes \mu_n(x_n) \\
&\leq \iint\limits_{\prod\limits_{j\neq i}[a_j,b_j]} \frac{b_i - b_{\mu_i}}{b_i - a_i} f(\widehat{x_i}, a_i)d\widehat{\mu_i}(\widehat{x_i}) + \iint\limits_{\prod\limits_{j\neq i}[a_j,b_j]} \frac{b_{\mu_i} - a_i}{b_i - a_i} f(\widehat{x_i}, b_i)d\widehat{\mu_i}(\widehat{x_i}) + \frac{k_i}{2}(b_{\mu_i} - a_i)(b_i - b_{\mu_i}).
\end{aligned}$$

Adding these n inequalities and dividing the sum by n, one gets

$$\begin{aligned}
&\iint\limits_{\prod\limits_{i=1,...,n}[a_i,b_i]} f(x_1, ..., x_n)d\mu_1(x_1) \otimes ... \otimes \mu_n(x_n) + \sum_{i=1}^{n} \frac{k_i}{2n}\int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i) \\
&\leq \sum_{i=1}^{n} \iint\limits_{\prod\limits_{j\neq i}[a_j,b_j]} \frac{b_i - b_{\mu_i}}{n(b_i - a_i)} f(\widehat{x_i}, a_i)d\widehat{\mu_i}(\widehat{x_i}) + \sum_{i=1}^{n} \iint\limits_{\prod\limits_{j\neq i}[a_j,b_j]} \frac{b_{\mu_i} - a_i}{n(b_i - a_i)} f(\widehat{x_i}, b_i)d\widehat{\mu_i}(\widehat{x_i}) \\
&\qquad + \sum_{i=1}^{n} \frac{k_i}{2n}(b_{\mu_i} - a_i)(b_i - b_{\mu_i})\ . \qquad (2.10)
\end{aligned}$$

Adding

$$\sum_{i=1}^{n} \frac{(n-1)k_i}{2n} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i)$$

to both sides of (2.10), one proves inequality (2.7).

To end the prove we show inequality (2.8) by induction on n. If $n = 1$, then (2.8) is actually an equality. Assume (2.8) hods for $n-1$, then

$$\begin{aligned}
&\sum_{i=1}^{n} \iint\limits_{\prod\limits_{j\neq i, j=1,..,n}[a_j,b_j]} \frac{b_i - b_{\mu_i}}{n(b_i - a_i)} f(\widehat{x_i}, a_i) d\widehat{\mu_i}(x_i) + \sum_{i=1}^{n} \iint\limits_{\prod\limits_{j\neq i, j=1,..,n}[a_j,b_j]} \frac{b_{\mu_i} - a_i}{n(b_i - a_i)} f(\widehat{x_i}, b_i) d\widehat{\mu_i}(x_i) \\
&\quad + \sum_{i=1}^{n} \frac{(n-1)k_i}{2n} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i) + \sum_{i=1}^{n} \frac{k_i}{2n} (b_{\mu_i} - a_i)(b_i - b_{\mu_i}) \\
=& \frac{n-1}{n} \int_{a_n}^{b_n} [\sum_{i=1}^{n-1} \iint\limits_{\prod\limits_{j\neq i, j=1,..,n-1}[a_j,b_j]} \frac{b_i - b_{\mu_i}}{(n-1)(b_i - a_i)} f(\widehat{x_i}, a_i) d\mu_1(x_1) \otimes ... \otimes \mu_{n-1}(x_{n-1})] d\mu_n(x_n) \\
&+ \frac{n-1}{n} \int_{a_n}^{b_n} [\sum_{i=1}^{n-1} \iint\limits_{\prod\limits_{j\neq i, j=1,..,n-1}[a_j,b_j]} \frac{b_{\mu_i} - a_i}{(n-1)(b_i - a_i)} f(\widehat{x_i}, b_i) d\mu_1(x_1) \otimes ... \otimes \mu_{n-1}(x_{n-1})] d\mu_n(x_n) \\
&\quad + \frac{n-1}{n} \int_{a_n}^{b_n} \sum_{i=1}^{n-1} \frac{(n-2)k_i}{2(n-1)} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i)] d\mu_n(x_n) \\
&\quad + \frac{n-1}{n} \int_{a_n}^{b_n} [\sum_{i=1}^{n-1} \frac{k_i}{2(n-1)} (b_{\mu_i} - a_i)(b_i - b_{\mu_i})] d\mu_n(x_n) \\
&\quad + \iint\limits_{\prod\limits_{j=1,..,n-1}[a_j,b_j]} \frac{b_n - b_{\mu_n}}{n(b_n - a_n)} f(\widehat{x_n}, a_n) + \frac{b_{\mu_n} - a_n}{n(b_n - a_n)} f(\widehat{x_n}, b_n) d\mu_1(x_1) \otimes .... \otimes \mu_{n-1}(x_{n-1}) \\
&\quad + \sum_{i=1}^{n-1} \frac{k_i}{2n} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i) + \frac{(n-1)k_n}{2n} \int_{a_n}^{b_n} |x_n - b_{\mu_n}|^2 d\mu_n(x_n) + \frac{k_n}{2n} (b_{\mu_n} - a_n)(b_n - b_{\mu_n}) \\
\leq& \frac{n-1}{n} \int_{a_n}^{b_n} \sum_{\substack{x_i = a_i \ or \ b_i, \\ i=1,2,...,n-1}} \frac{g(x_1)...g(x_{n-1})}{\prod\limits_{i=1}^{n-1}(b_i - a_i)} f(x_1, ..., x_n) + \sum_{i=1}^{n-1} \frac{k_i}{2} (b_{\mu_i} - a_i)(b_i - b_{\mu_i})] d\mu_n(x_n) \\
&\quad + \iint\limits_{\prod\limits_{j=1,..,n-1}[a_j,b_j]} \frac{b_n - b_{\mu_n}}{n(b_n - a_n)} f(\widehat{x_n}, a_n) d\widehat{\mu_n}(x_n) + \iint\limits_{\prod\limits_{j=1,..,n-1}[a_j,b_j]} \frac{b_{\mu_n} - a_n}{n(b_n - a_n)} f(\widehat{x_n}, b_n) d\widehat{\mu_n}(x_n) \\
&\quad + \sum_{i=1}^{n-1} \frac{k_i}{2n} \int_{a_i}^{b_i} |x_i - b_{\mu_i}|^2 d\mu_i(x_i) + \frac{(n-1)k_n}{2n} \int_{a_n}^{b_n} |x_n - b_{\mu_n}|^2 d\mu_n(x_n) + \frac{k_n}{2n} (b_{\mu_n} - a_n)(b_n - b_{\mu_n})
\end{aligned}$$

$$
\begin{aligned}
=&\frac{n-1}{n}\sum_{\substack{x_i=a_i\ or\ b_i,\\ i=1,2,\dots,n-1}}\frac{g(x_1)...g(x_{n-1})}{\prod\limits_{i=1}^{n-1}(b_i-a_i)}\int_{a_n}^{b_n}[f(x_1,...,x_n)+\frac{k_n}{2}\int_{a_n}^{b_n}|x_n-b_{\mu_n}|^2]d\mu_n(x_n)\\
&+\iint\limits_{\prod\limits_{j=1,..,n-2}[a_j,b_j]}\int\limits_{a_{n-1}}^{b_{n-1}}\frac{b_n-b_{\mu_n}}{n(b_n-a_n)}[f(\widehat{x_n},a_n)+\frac{k_{n-1}}{2}|x_{n-1}-b_{\mu_{n-1}}|^2]]\\
&+\frac{b_{\mu_n}-a_n}{n(b_n-a_n)}[f(\widehat{x_n},b_n)+\frac{k_{n-1}}{2}|x_{n-1}-b_{\mu_{n-1}}|^2]d\mu_{n-1}(x_{n-1})\otimes d\mu_1\otimes...\otimes d\mu_{n-2}\\
&+\sum_{i=1}^{n-2}\frac{k_i}{2n}\int_{a_i}^{b_i}|x_i-b_{\mu_i}|^2d\mu_i(x_i)+\frac{k_n}{2n}(b_{\mu_n}-a_n)(b_n-b_{\mu_n})+\sum_{i=1}^{n}\frac{(n-1)k_i}{2n}(b_{\mu_i}-a_i)(b_i-b_{\mu_i})\\
\le&\frac{n-1}{n}\sum_{\substack{x_i=a_i\ or\ b_i,\\ i=1,2,\dots,n-1}}\frac{g(x_1)...g(x_{n-1})}{\prod\limits_{i=1}^{n-1}(b_i-a_i)}[\frac{b_n-b_{\mu_n}}{b_n-a_n}f(x_1,x_2,...,a_n)\\
&+\frac{b_{\mu_n}-a_n}{b_n-a_n}f(x_1,x_2,...,b_n)+\frac{k_n}{2}(b_{\mu_n}-a_n)(b_n-b_{\mu_n})]\\
&+\iint\limits_{\prod\limits_{j=1,..,n-2}[a_j,b_j]}\frac{b_n-b_{\mu_n}}{n(b_n-a_n)}[\frac{b_{n-1}-b_{\mu_{n-1}}}{b_{n-1}-a_{n-1}}f(x_1,x_2,...,a_{n-1},a_n)\\
&+\frac{b_{\mu_{n-1}}-a_{n-1}}{b_{n-1}-a_{n-1}}f(x_1,x_2,...,b_{n-1},a_n)+\frac{k_{n-1}}{2}(b_{\mu_{n-1}}-a_{n-1})(b_{n-1}-b_{\mu_{n-1}})]\\
&+\iint\limits_{\prod\limits_{j=1,..,n-2}[a_j,b_j]}\frac{b_{\mu_n}-a_n}{n(b_n-a_n)}[\frac{b_{n-1}-b_{\mu_{n-1}}}{b_{n-1}-a_{n-1}}f(x_1,x_2,...,a_{n-1},b_n)\\
&+\frac{b_{\mu_{n-1}}-a_{n-1}}{b_{n-1}-a_{n-1}}f(x_1,x_2,...,b_{n-1},b_n)+\frac{k_{n-1}}{2}(b_{\mu_{n-1}}-a_{n-1})(b_{n-1}-b_{\mu_{n-1}})]\\
&+\sum_{i=1}^{n-2}\frac{k_i}{2n}\int_{a_i}^{b_i}|x_i-b_{\mu_i}|^2d\mu_i(x_i)+\frac{k_n}{2n}(b_{\mu_n}-a_n)(b_n-b_{\mu_n})+\sum_{i=1}^{n}\frac{(n-1)k_i}{2n}(b_{\mu_i}-a_i)(b_i-b_{\mu_i})\\
\le&\sum_{\substack{x_i=a_i\ or\ b_i,\\ i=1,2,\dots,n}}\frac{g(x_1)...g(x_n)}{\prod\limits_{i=1}^{n}(b_i-a_i)}f(x_1,...,x_n)+\sum_{i=1}^{n-1}\frac{k_i}{2n}(b_{\mu_i}-a_i)(b_i-b_{\mu_i})\\
&+\frac{k_n}{2n}(b_{\mu_n}-a_n)(b_n-b_{\mu_n})+\sum_{i=1}^{n}\frac{(n-1)k_i}{2n}(b_{\mu_i}-a_i)(b_i-b_{\mu_i})\\
=&\sum_{\substack{x_i=a_i\ or\ b_i,\\ i=1,2,\dots,n}}\frac{g(x_1)...g(x_{n-1})}{\prod\limits_{i=1}^{n-1}(b_i-a_i)}f(x_1,x_2,...,x_n)+\sum_{i=1}^{n}\frac{k_i}{2}(b_{\mu_i}-a_i)(b_i-b_{\mu_i})
\end{aligned}
$$

□

Remark 2.5. If $k_i=k, i=1,2,...,n$, then a function which is semiconvex of rate $(k_1,k_2,...,k_n)$ on the co-ordinates is actually semiconvex of rate k, and one has the Hermite-Hadamard inequality for n-dimensional semiconvex functions of rate k after taking $k_i = k, i = 1, 2, ..., n$ in Theorem 2.4.

Remark 2.6. In the two dimensional case, by taking $\mu_1 := \dfrac{1}{b-a}\, \nu\Big|_{[a,b]}$, $\mu_2 = \dfrac{1}{d-c}\, \nu\Big|_{[c,d]}$, where ν is the 1-dimensional Lebsegue measure, then $b_{\mu_1} = \dfrac{b+a}{2}, b_{\mu_2} = \dfrac{c+d}{2}$, and Theorem 2.4 reduces to Dragomir's result (1.2) for a function which is convex on the co-ordinates (see [2]). Such a function, in fact, is also semiconvex of rate $(0, 0)$ on the co-ordinates.

Remark 2.7. Theorem 2.2 which is a one-dimensional Hermite-Hadamard inequality can also be seen as a special case of Theroem 2.4. In fact, observe that when the intervals $[a_i, b_i], i = 2, ...n$ degenerate to points, the function $f(x_1, x_2, ..., x_n)$ in Theroem 2.4 reduces to a semi-convex function $f(x) : [a, b] \to \mathbb{R}$. With suitable modifications, (2.5), (2.6), (2.7) and (2.8) in Theroem 2.4 reduce to (2.2) and (2.3) in Theorem 2.2.

Remark 2.8. For higher dimensional convex functions f on a hyper-rectangle $\prod\limits_{i=1}^{n} [a_i, b_i]$, we also refer to [1] in which the Hermite-Hadamard inequality is expressed in probabilistic terms, that is, $f(E\xi) \le Ef(\xi) \le Ef(\xi^*)$, where E denotes mathematical expectation, and ξ (respectively ξ^*) is a random vector. By taking $\xi = (\xi_1, ..., \xi_i, ...\xi_n)$ a random vector with ξ_i having uniform distribution on $[a_i, b_i], i = 2, ...n$, then $f(E\xi) \le Ef(\xi) \le Ef(\xi^*)$ implies

$$f(\frac{a_1+b_1}{2}, ..., \frac{a_n+b_n}{2}) \le \iint\limits_{\prod\limits_{i=1,...,n} [a_i,b_i]} f(x_1, ..., x_n) dx_1...dx_n \le \sum_{\substack{x_i=a_i \ or \ b_i, \\ i=1,2,...,n}} \frac{1}{2^n} f(x_1, ..., x_n).$$

It can also be reduced by using Theroem 2.4 in which f is taken as a convex function and $\mu_i := \dfrac{1}{b_i - a_i}\, \nu\Big|_{[a_i,b_i]}$, $i = 1, ..., n$, where ν is the 1-dimensional Lebsegue measure, then $b_{\mu_i} = \dfrac{b_i + a_i}{2}, i = 1, ..., n.$

3 Mass transportation and Hermite-Hadamard inequality

In this section, we interpret the meaning of the new Hermite-Hadamard Inequality obtained in the previous section from the point of view of optimal mass transportation problems.

A typical optimal mass transport problem is the Kantorovich problem, which is formulated as:

$$\min_{\gamma \in \Pi(\nu_2, \nu_2)} \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y)\, d\gamma(x, y) \ , \tag{3.1}$$

where ν_1, $\nu_2 \in \mathcal{P}(\mathbb{R}^n)$ with $\mathcal{P}(\mathbb{R}^n)$ meaning the space of Borel probability measures on $\mathbb{R}^n$, $c(x, y) : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty)$ is a *cost function*, and

$$\Pi(\nu_1, \nu_2) := \{\gamma \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n) : (\pi_1)_\sharp \gamma = \nu_1,\ (\pi_2)_\sharp \gamma = \nu_2\}$$

is the set of transport plans between ν_1 and ν_2. Here $\pi_1, \pi_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ are the canonical projections on the first and second factors, respectively. We refer to [9, 10] for more information and references on optimal mass transportation theory.

Before proceeding, we first recall some standard notations. Let $\mu_i \in \mathcal{P}([a_i, b_i]), i = 1, 2, ..., n$ and δ_x denotes the Dirac measure at the point $x \in \mathbb{R}$. The product measure $\delta_{b_{\mu_1}} \otimes ... \otimes \delta_{b_{\mu_n}} \in \mathcal{P}(\mathbb{R}^n)$ of $\delta_{b_{\mu_i}}, i = 1, 2, ..., n$ is given by

$$\delta_{b_{\mu_1}} \otimes ... \otimes \delta_{b_{\mu_n}}(A_1 \times ... \times A_n) = \begin{cases} 1 & \text{if } b_{\mu_i} \in A_i, i = 1, 2, ..., n \\ 0 & \text{otherwise} \end{cases}$$

for any Borel measurable $A_i \subset [a_i, b_i], i = 1, 2, ..., n$.

3.1 Mass transportation meaning of the Hermite-Hadamard Inequality

After adding the constant $\sum_{i=1}^n \frac{k_i}{2} b_{\mu_i}^2$ to each term in the Hermite-Hadamard inequality in Theorem 2.4, we can prove that each new term equals to the mass transport cost in the following series of transportation models with initial measure $\nu_1 = \delta_0 \otimes ... \otimes \delta_0 \in \mathcal{P}(\mathbb{R}^n)$ and cost function $c(x, y) : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty)$ given by

$$c(x, y) = c(x_1.x_2, ..., x_n, y_1, y_2, ..., y_n) := f(y - x) + \sum_{i=1}^{n} \frac{k_i}{2} \left(x_i - y_i\right)^2, \tag{3.2}$$

where

$$f : \prod_{i=1}^{n} [a_i, b_i] \to [0, +\infty)$$

is continuous and semiconvex of rate $(k_1, k_2, ..., k_n)$ on the co-ordinates.

Example 3.1. Take $\nu_1 = \delta_0 \otimes ... \otimes \delta_0,\ \nu_2 = \delta_{b_{\mu_1}} \otimes ... \otimes \delta_{b_{\mu_n}} \in \mathcal{P}(\mathbb{R}^n)$, then $\Pi(\nu_1, \nu_2) = \{\nu_1 \otimes \nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 is the sum of expression (2.4) and $\sum_{i=1}^n \frac{k_i}{2} b_{\mu_i}^2$.

Example 3.2. Take $\nu_1 = \delta_0 \otimes ... \otimes \delta_0$, $\nu_2 = \frac{1}{n} \sum\limits_{i=1}^{n} \delta_{b_{\mu_1}} \otimes ... \delta_{b_{\mu_{i-1}}} \otimes \mu_i \otimes \delta_{b_{\mu_{i+1}}} \otimes \delta_{b_{\mu_n}} \in \mathcal{P}(\mathbb{R}^n)$, then $\Pi(\nu_1, \nu_2) = \{\nu_1 \otimes \nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 is the expression (2.5) in Theorem 2.4 plus $\sum_{i=1}^n \frac{k_i}{2} b_{\mu_i}^2$.

Example 3.3. Take $\nu_1 = \delta_0 \otimes ... \otimes \delta_0, \nu_2 = \mu_1 \otimes ... \otimes \mu_n \in \mathcal{P}(\mathbb{R}^n)$, then $\Pi(\nu_1, \nu_2) = \{\nu_1 \otimes \nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 is the expression (2.6) in Theorem 2.4 plus $\sum_{i=1}^n \frac{k_i}{2} b_{\mu_i}^2$.

Example 3.4. Take $\nu_1 = \delta_0 \otimes ... \otimes \delta_0$,

$$\begin{aligned} \nu_2 = &\sum_{i=1}^{n} \frac{b_i - b_{\mu_i}}{n(b_i - a_i)} \mu_1 \otimes ... \otimes \mu_{i-1} \otimes \delta_{a_i} \otimes \mu_{i+1} \otimes ... \otimes \mu_n \\ &+ \sum_{i=1}^{n} \frac{b_{\mu_i} - a_i}{n(b_i - a_i)} \mu_1 \otimes ... \otimes \mu_{i-1} \otimes \delta_{b_i} \otimes \mu_{i+1} \otimes ... \otimes \mu_n \end{aligned}$$

in $\mathcal{P}(\mathbb{R}^n)$, then $\Pi(\nu_1, \nu_2) = \{\nu_1 \otimes \nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 is the expression (2.7) in Theorem 2.4 plus $\sum_{i=1}^n \frac{k_i}{2} b_{\mu_i}^2$.

Example 3.5. Take $\nu_1 = \delta_0 \otimes ... \otimes \delta_0$,

$$\nu_2 = \sum_{\substack{x_i = a_i \ or \ b_i, \\ i=1,2,...,n}} \frac{g(x_1)g(x_2)...g(x_n)}{\prod_{i=1}^n (b_i - a_i)} \delta_{x_1} \otimes \delta_{x_2} \otimes ... \otimes \delta_{x_n}$$

in $\mathcal{P}(\mathbb{R}^n)$. Here $g(\cdot)$ is defined as that in Theorem 2.4. Then $\Pi(\nu_1, \nu_2) = \{\nu_1 \otimes \nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 is the expression (2.8) in Theorem 2.4 plus $\sum_{i=1}^n \frac{k_i}{2} b_{\mu_i}^2$.

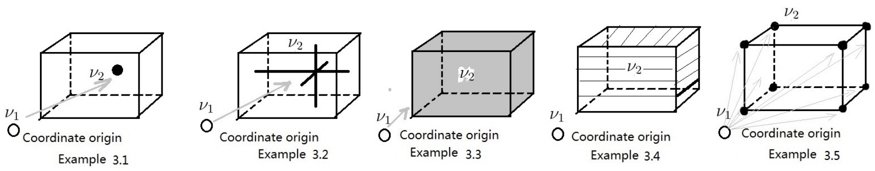

Figure 1: Mass transportation meaning of Theorem 2.4

As the transport costs from ν_1 to ν_2 in Examples 3.1, 3.2, 3.3, 3.4 and 3.5 equal to each term (adding $\sum_{i=1}^n \frac{k_i}{2} b_{\mu_i}^2$) in the Hermite-Hadamard inequality in Theorem 2.4, respectively, it follows that the transfer costs in these examples become more and more expensive (see Figure 1: For simplicity, the figures are drawn in three-dimension).

3.2 Mass transportation models in the unit ball

We consider mass transportation models in the unit ball. For simplicity, we give examples in 3-dimension. Here we take $c^*(x, y) = ||x - y||^2 : \mathbb{R}^3 \times \mathbb{R}^3 \to [0, +\infty)$ and the initial measure being $\delta_{(0,0,0)}$, the Dirac measure at the origin in $\mathbb{R}^3$. These transportation models in the unit ball can be seen as models in a hyper-rectangle $I_3 := [0, 1] \times [0, 2\pi] \times [0, \pi]$, and hence one can use the Hermite-Hadamard inequality in Theorem 2.4 and the interpretation in Section 3.1 to compare the transport costs in different models in the unit ball via the comparison of the transport costs in the corresponding models in hyper-rectangles.

Set $c(z, w) = (r_z - r_w)^2 : I_3 \times I_3 \to [0, +\infty)$, where $z = (r_z, \theta_z, \varphi_z), w = (r_w, \theta_w, \varphi_w)$. Set $\mu_r = 3r^2 dr \in \mathcal{P}([0, 1])$, $\mu_\theta = \frac{1}{2\pi} d\theta \in \mathcal{P}([0, 2\pi])$, and $\mu_\varphi = \frac{sin\varphi}{2} d\varphi \in \mathcal{P}([0, \pi])$. Then $b_{\mu_r} = \frac{3}{4}$, $b_{\mu_\theta} = \pi$, $b_{\mu_\varphi} = \frac{\pi}{2}$, and $\mu_r \otimes \mu_\theta \otimes \mu_\varphi \in \mathcal{P}(I_3)$. By performing the following change of coordinates

$$T: \begin{array}{ll} S^2 & \to I_3 \\ x = (x_1, x_2, x_3) & \to z = (r, \theta, \varphi) \end{array}, \quad \left\{ \begin{array}{ll} x_1 = & rsin\varphi cos\theta \\ x_2 = & rsin\varphi sin\theta \\ x_3 = & rcos\varphi \end{array} \right.$$

one has $c^*(0, x) = ||x||^2 = \sum_{i=1}^3 x_i^2 = r^2 = c(0, Tx)$.

Example 3.6. Take $\nu_1^* = \delta_{(0,0,0)}$, $\nu_2^* = \delta_{\frac{3}{4}} \otimes \delta_0 \otimes \delta_0 \in \mathcal{P}(\mathbb{S}^2)$. Then $\Pi(\nu_1^*, \nu_2^*) = \{\nu_1^* \otimes \nu_2^*\}$ is a singleton, and the optimal transportation cost from ν_1^* to ν_2^* for the cost function $c^*(x, y) = ||x - y||^2$ is

$$\int_{S^2 \times S^2} c^*(x, y) d\nu_1^* \otimes \nu_2^*(x, y) = (\frac{3}{4})^2 .$$

Correspondingly, take $\nu_1 = \delta_{(0,0,0)}$, $\nu_2 = \delta_{b_{\mu_r}} \otimes \delta_{b_{\mu_\theta}} \otimes \delta_{b_{\mu_\varphi}} = \delta_{(\frac{3}{4}, \pi, \frac{\pi}{2})} \in \mathcal{P}(I_3)$. Then $\Pi(\nu_1, \nu_2) = \{\nu_1 \otimes \nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 for the cost function $c(z, w) = (r_z - r_w)^2$ is

$$\int_{I_3 \times I_3} c(z, w) \nu_1 \otimes \nu_2(z, w) = (b_{\mu_r})^2 = (\frac{3}{4})^2 .$$

Hence the optimal transport cost in the two models coincide (see Figure 2).

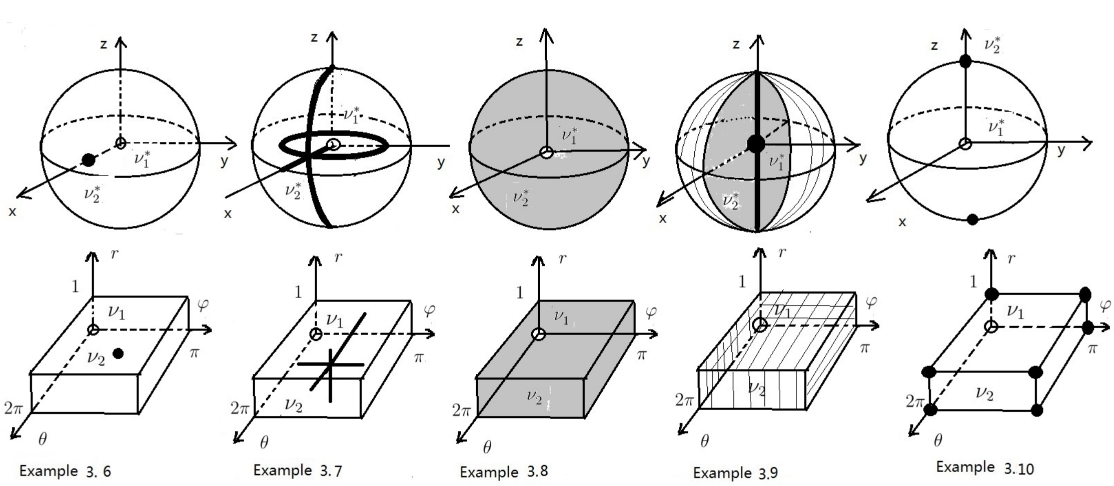

Figure 2: Comparison of optimal transport costs

Example 3.7. Take $\nu_1^* = \delta_{(0,0,0)}$, $\nu_2^* = \frac{1}{3}\mu_r \otimes \delta_0 \otimes \delta_0 + \frac{1}{3}\mu_{L_1} \otimes \delta_0 + \frac{1}{3}\mu_{L_2} \otimes \delta_0 \in \mathcal{P}(\mathbb{S}^2)$. Here $L_1 : \begin{cases} x^2 + y^2 = (\frac{3}{4})^2 \\ z = 0 \end{cases}$, $L_2 : \begin{cases} x = -\sqrt{(\frac{3}{4})^2 - z^2} \\ y = 0 \end{cases}$ are curves in $\mathbb{S}^2$ and $\mu_{L_i} \in \mathcal{P}(L_i)$ such that $\mu_{L_i}(A) = \int_{A \cap L_i} ds$, $i = 1, 2$, where ds is the arc length element. Then $\Pi(\nu_1^*, \nu_2^*) = \{\nu_1^* \otimes \nu_2^*\}$ is a singleton, and the optimal transportation cost from ν_1^* to ν_2^* for the cost function $c^*(x, y) = ||x - y||^2$ is

$$\begin{aligned} &\int_{S^2 \times S^2} c^*(x, y) d\nu_1^* \otimes \nu_2^*(x, y) \\ &= \frac{1}{3} \int_0^1 y_1^2 \cdot 3y_1^2 dy_1 + \frac{1}{3} \int_{L_1} (\frac{3}{4})^2 d\mu_{L_1} + \frac{1}{3} \int_{L_2} (\frac{3}{4})^2 d\mu_{L_2} = \frac{13}{40}. \end{aligned}$$

Correspondingly, take $\nu_1 = \delta_{(0,0,0)}$, $\nu_2 = \frac{1}{3}\delta_{b_{\mu_r}} \otimes \delta_{b_{\mu_\theta}} \otimes \mu_\varphi + \frac{1}{3}\delta_{b_{\mu_r}} \otimes \mu_\theta \otimes \delta_{b_{\mu_\varphi}} + \frac{1}{3}\mu_r \otimes \delta_{b_{\mu_\theta}} \otimes \delta_{b_{\mu_\varphi}} \in \mathcal{P}(I_3)$. Then $\Pi(\nu_1, \nu_2) = \{\nu_1 \otimes \nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 for the cost function $c(z, w) = (r_z - r_w)^2$ is

$$\begin{aligned}
&\int_{I_3\times I_3} c(z,w)\nu_1 \otimes \nu_2(z,w) \\
&= \frac{1}{3}\int_0^{2\pi}(\frac{3}{4})^2 d\mu_\theta + \frac{1}{3}\int_0^{\pi}(\frac{3}{4})^2 d\mu_\varphi + \frac{1}{3}\int_0^{1} r^2 d\mu_r \\
&= \frac{1}{3}\int_0^{1} 3r^4 d\mu_r + \frac{2}{3}\cdot(\frac{3}{4})^2 = \frac{13}{40}.
\end{aligned}$$

Hence the optimal transport cost in the two models coincide.

Example 3.8. Take $\nu_1^* = \delta_{(0,0,0)}$ and $\nu_2^* \in \mathcal{P}(\mathbb{S}^2))$ such that $d\nu_2^* = \frac{3}{4\pi}dx_1x_2x_3$, then $\Pi(\nu_1^*, \nu_2^*) = \{\nu_1^* \otimes \nu_2^*\}$ is a singleton, and the optimal transportation cost from ν_1^* to ν_2^* for the cost function $c^*(x, y) = ||x - y||^2$ is

$$\begin{aligned}
\int_{S^2\times S^2} c^*(x,y)d\nu_1^* \otimes \nu_2^*(x,y) &= \frac{3}{4\pi}\iiint_{\mathbb{S}^2}\sum_{i=1}^{3} y_i^2 dy_1y_2y_3 \\
&= \frac{3}{4\pi}\int_0^1\int_0^{2\pi}\int_0^{\pi} r^2r^2 sin\varphi dr d\theta d\varphi = \frac{3}{5}\ .
\end{aligned}$$

Correspondingly, take $\nu_1 = \delta_{(0,0,0)}$, $\nu_2 = \mu_r \otimes \mu_\theta \otimes \mu_\varphi \in \mathcal{P}(I_3)$. Then $\Pi(\nu_1, \nu_2) = \{\nu_1 \otimes \nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 for the cost function $c(z, w) = (r_z - r_w)^2$ is

$$\begin{aligned}
\int_{I_3\times I_3} c(z,w)\nu_1 \otimes \nu_2(z,w) &= \int_{I_3} r^2 d\mu_r \otimes \mu_\theta \otimes \mu_\varphi(r,\theta,\varphi) \\
&= \int_0^1\int_0^{2\pi}\int_0^{\pi} r^2 3r^2 dr \cdot \frac{1}{2\pi}d\theta \cdot \frac{sin\varphi}{2}d\varphi = \frac{3}{5}.
\end{aligned}$$

Once again, the optimal transport cost in the two models coincide.

Example 3.9. Take $\nu_1^* = \delta_{(0,0,0)}$, $\nu_2^* = \frac{1}{12}\delta_{(0,0,0)} + \frac{3}{12}\mu_{\partial\mathbb{S}^2} + \frac{1}{3}\mu_s + \frac{1}{3}\mu_L \in \mathcal{P}(\mathbb{S}^2)$, where $\mu_{\partial\mathbb{S}^2} \in \mathcal{P}(\partial\mathbb{S}^2))$, μ_s is the probability measure on the plane $s : \begin{cases} x_1^2 + x_3^2 \leq 1 \\ x_2 = 0 \end{cases}$ such that $d\mu_s = \frac{3}{2\pi}\sqrt{x_1^2 + x_3^2}dx_1dx_3$, and μ_L is the probability measure on the line $L : \begin{cases} 0 \leq x_3 \leq 1 \\ x_1 = x_2 = 0 \end{cases}$ such that $d\mu_L = 3x_3^2dx_3$. Then $\Pi(\nu_1^*, \nu_2^*) = \{\nu_1^* \otimes \nu_2^*\}$ is a singleton, and the optimal transportation cost from ν_1^* to ν_2^* for the cost function $c^*(x,y) = ||x-y||^2$ is

$$\begin{aligned}
&\int_{S^2\times S^2} c^*(x,y)d\nu_1^*\otimes\nu_2^*(x,y) \\
&= \frac{1}{12}\cdot 0 + \frac{3}{12}\int_{\partial\mathbb{S}^2} d\mu_{\partial\mathbb{S}^2} + \frac{1}{3}\cdot\frac{3}{\pi}\iint_s (x_1^2+x_3^2)\sqrt{x_1^2+x_3^2}dx_1dx_3 + \frac{1}{3}\int_0^1 x_3^2 d\mu_L \\
&= \frac{3}{12} + \frac{1}{3}\cdot\frac{3}{2\pi}\int_0^1\int_0^{2\pi} r^2\cdot r\cdot rdrd\theta + \frac{1}{3}\int_0^1 x_3^2\cdot 3x_3^2dx_3 = \frac{13}{20}\,.
\end{aligned}$$

Correspondingly, take $\nu_1 = \delta_{(0,0,0)}$, $\nu_2 = \frac{1}{12}\delta_0\otimes\mu_\theta\otimes\mu_\varphi + \frac{3}{12}\delta_1\otimes\mu_\theta\otimes\mu_\varphi + \frac{1}{6}\mu_r\otimes\delta_0\otimes\mu_\varphi + \frac{1}{6}\mu_r\otimes\delta_{2\pi}\otimes\mu_\varphi + \frac{1}{6}\mu_r\otimes\mu_\theta\otimes\delta_0 + \frac{1}{6}\mu_r\otimes\mu_\theta\otimes\delta_\pi$ in $\mathcal{P}(I_3)$. Then $\Pi(\nu_1,\nu_2) = \{\nu_1\otimes\nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 for the cost function $c(z,w) = (r_z - r_w)^2$ is

$$\int_{I_3\times I_3} c(z,w)\nu_1\otimes\nu_2(z,w) = \frac{1}{12}\cdot 0 + \frac{3}{12}\cdot 1 + \frac{4}{6}\int_0^1 r^2\cdot 3r^2dr = \frac{13}{20}\,.$$

Again, the optimal transport cost in the two models coincide.

Example 3.10. Take $\nu_1^* = \delta_{(0,0,0)}$, $\nu_2^* = \frac{1}{4}\delta_{(0,0,0)} + \frac{3}{8}\delta_0\otimes\delta_0\otimes\delta_1 + \frac{3}{8}\delta_0\otimes\delta_0\otimes\delta_{-1} \in \mathcal{P}(\mathbb{S}^2)$. Then $\Pi(\nu_1^*,\nu_2^*) = \{\nu_1^*\otimes\nu_2^*\}$ is a singleton, and the optimal transportation cost from ν_1^* to ν_2^* for the cost function $c^*(x,y) = ||x-y||^2$ is

$$\int_{S^2\times S^2} c^*(x,y)d\nu_1^*\otimes\nu_2^*(x,y) = \frac{1}{4}\cdot 0 + \frac{3}{8}\cdot 1 + \frac{3}{8}\cdot 1 = \frac{3}{4}\,.$$

Correspondingly, take $\nu_1 = \delta_{(0,0,0)}$,

$$\begin{aligned}
\nu_2 \;=\; & \frac{1}{16}\delta_0\otimes\delta_{2\pi}\otimes\delta_0 + \frac{1}{16}\delta_0\otimes\delta_{2\pi}\otimes\delta_\pi + \frac{1}{16}\delta_0\otimes\delta_0\otimes\delta_0 \\
& + \frac{1}{16}\delta_0\otimes\delta_0\otimes\delta_\pi + \frac{3}{16}\delta_1\otimes\delta_{2\pi}\otimes\delta_0 + \frac{3}{16}\delta_1\otimes\delta_{2\pi}\otimes\delta_\pi \\
& + \frac{3}{16}\delta_1\otimes\delta_0\otimes\delta_0 + \frac{3}{16}\delta_1\otimes\delta_0\otimes\delta_\pi
\end{aligned}$$

in $\mathcal{P}(\mathbb{R}^n)$, then $\Pi(\nu_1,\nu_2) = \{\nu_1\otimes\nu_2\}$ is a singleton, and the optimal transportation cost from ν_1 to ν_2 for the cost function $c(z,w) = (r_z - r_w)^2$ is

$$\int_{I_3\times I_3} c(z,w)\nu_1\otimes\nu_2(z,w) = \frac{3}{16}\cdot 4 = \frac{3}{4}\,.$$

Once again, the optimal transport cost in the two models coincide.

Applying Theorem 2.4 to the convex continuous function $f(x_1,x_2,x_3) = x_1^2 : I_3 \to [0,1]$ and Borel probability measures $\mu_1 = \mu_r, \mu_2 = \mu_\theta, \mu_3 = \mu_\varphi$ on the intervals $[0,1]$, $[0,2\pi]$, $[0,\pi]$, respectively, one gets that each term in the Hermite-Hadamard inequality for $f(x_1, x_2, x_3) = x_1^2$ equals the optimal transfer cost $\int_{I_3 \times I_3} c(z, w)\nu_1 \otimes \nu_2(z, w)$ in the cube I_3 in Examples 3.6, 3.7, 3.8, 3.9 and 3.10, respectively. Hence the transfer costs in these examples become more and more expensive (see Figure 2). Furthermore, as the transport costs from ν_1^* to ν_2^* on the sphere $\mathbb{S}^2$ equals to the transport costs from ν_1 to ν_2 in the cube I_3, it follows that the transfer costs on the sphere in these examples become more and more expensive (see Figure 2).

Acknowledgments. The research of the first author was supported by the National Natural Science Foundation of China (No. 11601193), the Qing Lan Project of Jiangsu Province, and Jiangsu Overseas Visiting Scholar program for University Prominent Young & Middle-aged Teachers and Presidents.

References

[1] Cal, J. de la., Cárcamo, J.: Multidimensional Hermite-Hadamard inequalities and the convex order. J. Math. Anal. Appl., 324 (2006), 248-261.

[2] Dragomir, S. S.: On the Hadamard's inequlality for convex functions on the co-ordinates in a rectangle from the plane. Taiwanese J. of Math., 4 (2001), 775-788.

[3] Dragomir, S. S., Charles P.: Selected topics on Hermite-Hadamard inequalities and applications. RGMIA Monographs, Victoria University, 2000.

[4] Hadamard J.: Étude sur les propriétés des fonctions entières et en particulier d'une fonction considérée par Riemann, J. Math. Pures Appl., 58(1893), 171-215.

[5] Fu J.H.G.: Monge-Ampère Functions, I. Indiana Univ Math. J. 38 (1989), 745-771.

[6] Niculescu C. P., Pečarić J.: The equivalence of Chebyshevs inequality to the Hermite-Hadamard inequality. Math. Reports, 12 (2010), 145-156.

[7] Niculescu C. P.: The Hermite-Hadamard inequality for log-convex functions. Nonlinear Anal. TMA 75, (2012), 662-669.

[8] Niculescu C. P., Persson L. E.: Convex functions and their applications: a contemporary approach. Springer, Berlin (2004)

[9] Santambrogio F.: Optimal Transportation for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling, **87**, Birkäuser (2015).

[10] Villani, C.: Optimal transport, old and new. **338**, Grundlehren. Springer, Berlin (2009).

[11] Zhao C. J., Cheung W. S., Li X. Y.: On the Hermite-Hadamard type inequalities. J. Inequ. Appl. 2013;2013(228).