Online
ISSN 1346-0447

ISSN 1346-0862

Scientiae Mathematicae Japonicae

(Scientiae Mathematicae / Mathematica Japonica New Series)

Vol. 82, No. 1 Whole Number 291

April 2019

International Society for Mathematical Sciences

INTERNATIONAL SOCIETY FOR MATHEMATICAL SCIENCES
Scientiae Mathematicae Japonicae, Notices from the ISMS

The International Society for Mathematical Sciences (ISMS) is an international society consisting of mathematical scientists throughout the world.

The main activities of the ISMS are to publish (1) the (**print and online**) journal **Scientiae Mathematicae Japonicae (SCMJ)** and (2) Notices from the ISMS and to hold **assembly meeetings** in Japan and **international internet meetings** (distance symposium) of mathematical sciences (IVMS) accessible from all over the world.

SCMJ is the 21st Century New Unified Series of **Mathematica Japonica (MJ)** and **Scientiae Mathematicae (SCM). MJ** was first published in 1948 and was one of the oldest mathematical journals in Japan. **SCM** was an online and print journal started in 1998 in celebration of the semi-centurial anniversary and received 26000 visits per month from 50 countries in the world. **SCMJ** contains original papers in mathematical sciences submitted from all over the world and receives **38000** visits per month now. Not only papers in pure and applied mathematics but those devoted to mathematical statistics, operations research, informatics, computer science, biomathematics, mathematical economics and other mathematical sciences are also welcome. The journal is published in January, March, May, July, September, and November in each calendar year.

The ISMS has enhanced the journal, begining from July 1995, by including excellent Research-Expository papers in the section "**International Plaza** for Mathematical Sciences " as well as original research papers. The section provides papers dealing with broad overviews of contemporary mathmatical sciences, written by experts mainly at our invitation. Papers shedding lights on open problems or new directions or new breakthroughs for future research are especially welcome.

As is shown in the **Editorial Board of SCMJ**, we have invited **many distinguished professors** of 20 countries as editors, who will receive and referee the papers of their special fields with their **high standard**.

Beginning from 2007, we make the online version of SCMJ more readable and convenient to the readers by adding the specialized contents. By this, the readers can access to the online version, in which the papers appear in the order of acceptance, from (i) the contents of the printed version, and (ii) the specialized contents of a volume. From 2007, the subscription fee of the printed version plus the online version of SCMJ becomes lower and the same of the printed version only. Therefore, the subscribers of the printed version can read the online version without no additional cost.

For benefit of the ISMS members, we publish "Notices from the ISMS" 6 times a year. We are enhancing it by adding interesting articles, including book reviewing, written by eminent professors.

The ISMS has set up a **videoconferencing system** (**IVMS**) which can connect up to twenty sites of a reserch group in the same or **different countries in the world**. Using this system, speakers of the session can write on a white board or an OHP sheet or use PowerPoint. On the other hand participants can ask questions or make comments from any connected site in the world. All these are performed similarly to the traditional meetings.

To connect with our system, you can use your own videoconferencing system only if it satisfies the International Telecommunication Union-Technical Committee Standards (ITU-T Standard).

Copyright Transfer Agreement

A copyright transfer agreement is required before a paper is published in this journal. By submitting a paper to this journal, authors are regarded to certify that the manuscript has not been submitted to nor is it under consideration for publication by another journal, conference proceedings or similar publication.

For more information, please visit http://www.jams.or.jp.

A GENERALIZATION OF LLL LATTICE BASIS REDUCTION OVER IMAGINARY QUADRATIC FIELDS

KOICHI ARIMOTO AND YASUYUKI HIRANO

Received July 27, 2017

Abstract. In this paper we generalize LLL lattice basis reduction defined by Lenstra, Lenstra, and Lovász. We consider $\mathcal{O}_F$-lattice, where $\mathcal{O}_F$ is the ring of integers in algebraic number field F. We can prove that basic properties of reduced basis can hold over imaginary quadratic fields. We can reveal existence of a least positive element over other algebraic number fields.

1 Introduction Among all the $\mathbb{Z}$ bases of a lattice, some are better than others. The ones whose elements are the shortest are called *reduced*. Since the bases all have the same discriminant, to be reduced implies also that a basis is not too far from being orthogonal.

In 1982 A.K.Lenstra, H.W.Lenstra, Jr., and L.Lovász presented the LLL reduction algorithm. It was originally meant to find "short" vectors in lattices, i.e. to determine a so called reduced basis for a given lattice. H.Napias generalized LLL reduction algorithm over euclidean rings or orders([3]).

In this paper we define LLL reduced basis over imaginary quadratic fields. We consider a lattice in the n-dimensional linear space $V = F^n$, so F is an imaginary quadratic field. F is included by the field of complex numbers. Lenstra, Lenstra, and Lovász showed some properties about reduced bases over real number fields. We proved these properties hold over imaginary quadratic fields.

In last section, we consider a general algebraic number field F. Let $\mathcal{O}_F$ be the ring of integers in F, we state that $\mathcal{O}_F$ has a least positive element or not. And we show a necessary and sufficient condition for algebraic number field F to lead structure of lattice.

2 Basis reduction on $\mathbb{Z}$-modules We consider a lattice in n-dimensional linear space $\mathbb{R}^n$, where $\mathbb{R}$ is the field of real numbers.

A subset Λ of the n-dimensional real vector space $\mathbb{R}^n$ is called a *lattice* if there exists a basis $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$ of $\mathbb{R}^n$ such that

$$\Lambda = \sum_{i=1}^{n} \mathbb{Z}\boldsymbol{b}_i = \left\{ \sum_{i=1}^{n} r_i \boldsymbol{b}_i \;\middle|\; r_i \in \mathbb{Z} \ (1 \leq i \leq n) \right\}.$$

In this situation we say that the set $\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ of vectors forms a basis for Λ, or that it spans Λ. We call n the *rank* of Λ.

For a $\mathbb{Z}$-basis $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$ of Λ the *discriminant* $d(\Lambda)$ of Λ is defined by $d(\Lambda) = |\det(\boldsymbol{b}_i, \boldsymbol{b}_j)|^{\frac{1}{2}} > 0$, where (,) denotes the ordinary inner product on $\mathbb{R}^n$. This does not depend on the choice of the basis. And by Hadamard's inequality, we have $d(\Lambda) \leq \prod_{i=1}^{n} \|\boldsymbol{b}_i\|$.

In the sequel we consider the construction of special bases of lattices Λ. For the applications and for geometrical reasons we are interested in bases consisting of vectors of small norm. *Minkowski reduced* is an example of reduced basis. The computation of a Minkowski

2010 *Mathematics Subject Classification.* 11Y40, 11R04.
Key words and phrases. basis reduction, LLL-algorithm, imaginary quadratic fields.

reduced basis of a lattice can be very time consuming. Hence, in many cases one is satisfied with constructing bases of lattices which are reduced in a much weaker sense. The most important reduction procedure now in use is LLL-reduction which was introduced in 1982 by Lenstra, Lenstra, and Lovász in [2].

Let $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n \in \mathbb{R}^n$ be linearly independent. We recall the Gram-Schmidt orthogonalization process. The vectors $\boldsymbol{b}_i^*(1 \le i \le n)$ and the real numbers $\mu_{ij}(1 \le j < i \le n)$ are inductively defined by

$$\boldsymbol{b}_i^* := \boldsymbol{b}_i - \sum_{j=1}^{i-1} \mu_{ij}\boldsymbol{b}_j^*, \tag{1}$$

$$\mu_{ij} := \frac{(\boldsymbol{b}_i, \boldsymbol{b}_j^*)}{(\boldsymbol{b}_j^*, \boldsymbol{b}_j^*)}, \tag{2}$$

where (,) denotes the ordinary inner product on $\mathbb{R}^n$. We call a basis $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$ for a lattice *LLL-reduced* if

$$|\mu_{ij}| \le \frac{1}{2} \quad \text{for } 1 \le j < i \le n \tag{3}$$

and

$$\|\boldsymbol{b}_i^* + \mu_{i,i-1}\boldsymbol{b}_{i-1}^*\|^2 \ge \frac{3}{4}\|\boldsymbol{b}_{i-1}^*\|^2 \quad \text{for } 1 < i \le n \tag{4}$$

where $\|\cdot\|$ denotes the ordinary Euclidean length. Notice that the vectors $\boldsymbol{b}_i^* + \mu_{i,i-1}\boldsymbol{b}_{i-1}^*$ and $\boldsymbol{b}_{i-1}^*$ appearing in (4) are projections of $\boldsymbol{b}_i$ and $\boldsymbol{b}_{i-1}$ on the orthogonal complement of $\sum_{j=1}^{i-2} \mathbb{R}\boldsymbol{b}_j$. The constant $\frac{3}{4}$ in (4) is arbitrarily chosen, and may be replaced by any fixed real number y with $\frac{1}{4} < y < 1$.

We state without proof several key properties of LLL-reduced bases. The proof is given in [2].

Proposition 2.1 [2, Proposition(1.6), (1.11), (1.12)] *If $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$ is some reduced basis for a lattice Λ in $\mathbb{R}^n$, then*
(i) $\|\boldsymbol{b}_j\|^2 \le 2^{i-1}\|\boldsymbol{b}_i^*\|^2$ *for* $1 \le j \le i \le n$,
(ii) $d(\Lambda) \le \prod_{i=1}^n \|\boldsymbol{b}_i\| \le 2^{n(n-1)/4} d(\Lambda)$,
(iii) $\|\boldsymbol{b}_1\| \le 2^{(n-1)/4} d(\Lambda)^{1/n}$,
(iv) $\|\boldsymbol{b}_1\|^2 \le 2^{n-1}\|\boldsymbol{x}\|^2$ *for every* $\boldsymbol{x} \in \Lambda, \boldsymbol{x} \ne \boldsymbol{0}$,
(v) For any linearly independent set of vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_t \in \Lambda$ *we have*
$\|\boldsymbol{b}_j\|^2 \le 2^{n-1} \max\{\|\boldsymbol{x}_1\|^2, \cdots, \|\boldsymbol{x}_t\|^2\}$ *for* $1 \le j \le t \le n$,
where $\|\cdot\|$ denotes the ordinary Euclidean length.

3 Basis reduction on $\mathcal{O}_F$-modules Let F be an imaginary quadratic field and $\mathcal{O}_F$ be the ring of integers in F, now we consider a lattice in the n-dimensional linear space $V = F^n$.

Let n be a positive integer. A subset Λ of the n-dimensional vector space V is called a *$\mathcal{O}_F$-lattice* if there exists an $\mathcal{O}_F$-basis $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$ of V such that

$$\Lambda = \sum_{i=1}^n \mathcal{O}_F\boldsymbol{b}_i = \left\{ \sum_{i=1}^n r_i\boldsymbol{b}_i \;\middle|\; r_i \in \mathcal{O}_F \ (1 \le i \le n) \right\}.$$

Suppose that $\boldsymbol{a} = (a_1, \cdots, a_n)^t, \boldsymbol{b} = (b_1, \cdots, b_n)^t$ are vectors in $\mathbb{C}^n$. The *Hermitian inner product* of $\boldsymbol{a}$ and $\boldsymbol{b}$ is defined by

$$(\boldsymbol{a}, \boldsymbol{b}) = a_1\bar{b}_1 + \cdots + a_n\bar{b}_n. \tag{5}$$

Suppose that $\boldsymbol{x} = (x_1, \cdots, x_n)^t$ is vector in $\mathbb{C}^n$. The *norm of* $\boldsymbol{x}$ is defined by

$$\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})} = \sqrt{|x_1|^2 + \cdots + |x_n|^2}, \tag{6}$$

where, $x_i(\in \mathbb{C})$ is the i-th coordinate of $\boldsymbol{x}$, and $\|\boldsymbol{x}\| \in \mathbb{R}$.

Let $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n \in F^n$ be linearly independent. Similarly the vectors $\boldsymbol{b}_i^*(1 \leq i \leq n)$ and the complex numbers $\mu_{ij}(1 \leq j < i \leq n)$ are inductively defined by $\boldsymbol{b}_i^* := \boldsymbol{b}_i - \sum_{j=1}^{i-1} \mu_{ij}\boldsymbol{b}_j^*$, $\mu_{ij} := (\boldsymbol{b}_i, \boldsymbol{b}_j^*)/(\boldsymbol{b}_j^*, \boldsymbol{b}_j^*)$, where (,) denotes the Hermitian inner product on $\mathbb{C}^n$. And LLL-reduced basis is similarly defined by (3), (4).

From now on, we consider the imaginary quadratic field $F = \mathbb{Q}(\sqrt{m})$, where m is a square free negative integer, $R = \mathcal{O}_F$, the ring of integers in F.

Given imaginary quadratic field $\mathbb{Q}(\sqrt{m}) := \{a+b\sqrt{m} \mid a, b \in \mathbb{Q}\}$, the ring $\mathcal{O}_F$ of integers in $\mathbb{Q}(\sqrt{m})$ is the following:

(i) If $m \not\equiv 1 \pmod 4$, then $\mathcal{O}_F := \{a + b\sqrt{m} \mid a, b \in \mathbb{Z}\}$.

(ii) If $m \equiv 1 \pmod 4$, then $\mathcal{O}_F := \left\{\frac{a+b\sqrt{m}}{2} \mid a, b \in \mathbb{Z},\ a \equiv b \ (\mathrm{mod} 2)\right\}$.

For above two cases about m, we can prove its non-zero absolute values are greater than 1. So, we show below it as a lemma.

Lemma 3.1 *If $F = \mathbb{Q}(\sqrt{m})$, where $m < 0$, we get for any non-zero $r \in \mathcal{O}_F, |r|^2 \geq 1$.*

Proof. (i) In case $m \not\equiv 1 \pmod 4$. Let $r =: a + b\sqrt{m}$, where $a, b \in \mathbb{Z}$. Then we can rewrite r as $r = a + b\sqrt{-m}i$, therefore $|r|^2 = a^2 - mb^2$.
We assume $a \neq 0$. Then $|r|^2 \geq 1$. If $b \neq 0$, then $|r|^2 \geq -m \geq 1$. Hence if either $a \neq 0$ or $b \neq 0$, then $|r|^2 \geq 1$.
(ii) In case $m \equiv 1 \pmod 4$. Let $r =: \frac{a+b\sqrt{m}}{2}$, where $a, b \in \mathbb{Z}$, with $a \equiv b \pmod 2$. Then $|r|^2 = \frac{a^2-mb^2}{4}$. We show that if either $a \neq 0$ or $b \neq 0$ then $|r|^2 \geq 1$. Since $m < 0$, $m \equiv 1 \pmod 4$, the minimum value of $-m(> 0)$ is 3. Hence $|r|^2 \geq \frac{a^2+3b^2}{4}$.
(a) In case $a \equiv b \equiv 0 \pmod 2$. The minimum value of $a^2 + 3b^2$ is 4 ($a = \pm 2, b = 0$).
(b) In case $a \equiv b \equiv 1 \pmod 2$. The minimum value of $a^2 + 3b^2$ is 4 ($a = \pm 1,\ b = \pm 1$).
In any case, we have $|r|^2 \geq \frac{a^2+3b^2}{4} \geq 1$. ■

This lemma implies the following proposition.

Proposition 3.2 *Let F denote the imaginally quadratic field $\mathbb{Q}(\sqrt{m})$ and $R = \mathcal{O}_F$ be the ring of integers in F. Let $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$ be a basis of Λ, and $\boldsymbol{b}_i^*$ $(i = 1, 2, \cdots, n)$ be as above. Then we have*

$$\|\boldsymbol{x}\|^2 \geq \|\boldsymbol{b}_i^*\|^2 \quad \text{for some } i \leq n. \tag{7}$$

for any non-zero $\boldsymbol{x} \in \Lambda$.

Proof. For every $\boldsymbol{x} \in \Lambda$, we can write $\boldsymbol{x} =: \sum_{j=1}^n r_j\boldsymbol{b}_j = \sum_{j=1}^n s_j\boldsymbol{b}_j^*$, where $r_j \in \mathcal{O}_F$ and $s_j \in \mathbb{Q}(\sqrt{m})$. Let i be the largest index with $r_i \neq 0$. We claim that $\boldsymbol{x} = \sum_{j=1}^i s_j\boldsymbol{b}_j^*$ and

$r_i = s_i$. By $\boldsymbol{b}_i = \boldsymbol{b}_i^* + \sum_{j=1}^{i-1} \mu_{ij}\boldsymbol{b}_j^*$, we have

$$\boldsymbol{x} = \sum_{j=1}^{i} r_j \boldsymbol{b}_j = \sum_{j=1}^{i} \left(r_j + \sum_{k=j+1}^{i} r_k \mu_{kj} \right) \boldsymbol{b}_j^*. \tag{8}$$

We suppose $j = i$, we have $r_i = s_i$.
Next,

$$\|\boldsymbol{x}\|^2 = \|s_1 \boldsymbol{b}_1^*\|^2 + \|s_2 \boldsymbol{b}_2^*\|^2 + \cdots + \|s_i \boldsymbol{b}_i^*\|^2 \geq |s_i|^2 \|\boldsymbol{b}_i^*\|^2 \tag{9}$$

Now since $s_i = r_i, |r_i|^2 \geq 1$(by Lemma 3.1). Therefore we have

$$\|\boldsymbol{x}\|^2 \geq |r_i|^2 \|\boldsymbol{b}_i^*\|^2 \geq \|\boldsymbol{b}_i^*\|^2, \tag{10}$$

for some $i \leq n$. ■

These arguments imply the following main theorem. The idea of following proof is due to [2].

Theorem 3.3 *Let $F = \mathbb{Q}(\sqrt{m})$, where m is a square free negative integer, If $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$ is some reduced basis for a lattice Λ in V, then*
(i) $\|\boldsymbol{b}_j\|^2 \leq 2^{i-1}\|\boldsymbol{b}_i^*\|^2$ *for* $1 \leq j \leq i \leq n$,
(ii) $d(\Lambda) \leq \prod_{i=1}^{n} \|\boldsymbol{b}_i\| \leq 2^{n(n-1)/4} d(\Lambda)$,
(iii) $\|\boldsymbol{b}_1\| \leq 2^{(n-1)/4} d(\Lambda)^{1/n}$,
(iv) $\|\boldsymbol{b}_1\|^2 \leq 2^{n-1}\|\boldsymbol{x}\|^2$ *for every* $\boldsymbol{x} \in \Lambda, \boldsymbol{x} \neq \boldsymbol{0}$,
(v) *For any linearly independent set of vectors* $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_t \in \Lambda$ *we have*
$\|\boldsymbol{b}_j\|^2 \leq 2^{n-1}\max\{\|\boldsymbol{x}_1\|^2, \cdots, \|\boldsymbol{x}_t\|^2\}$ *for* $1 \leq j \leq t \leq n$,
where $\|\cdot\|$ *denotes the norm defined by (6).*

Proof. (i) From (4) and (3) we see that

$$\|\boldsymbol{b}_i^*\|^2 \geq \left(\frac{3}{4} - |\mu_{i,i-1}|^2\right) \|\boldsymbol{b}_{i-1}^*\|^2 \geq \frac{1}{2}\|\boldsymbol{b}_{i-1}^*\|^2$$

for $1 < i \leq n$, so by induction

$$\|\boldsymbol{b}_j^*\|^2 \leq 2^{i-j}\|\boldsymbol{b}_i^*\|^2 \quad \text{for } 1 \leq j \leq i \leq n.$$

From (1) and (3) we now obtain

$$\begin{aligned} \|\boldsymbol{b}_i\|^2 &= \|\boldsymbol{b}_i^*\|^2 + \sum_{j=1}^{i-1} |\mu_{ij}|^2 \|\boldsymbol{b}_j^*\|^2 \\ &\leq \|\boldsymbol{b}_i^*\|^2 + \sum_{j=1}^{i-1} \frac{1}{4} \cdot 2^{i-j} \|\boldsymbol{b}_i^*\|^2 \\ &= \left(1 + \frac{1}{4}(2^i - 2)\right) \|\boldsymbol{b}_i^*\|^2 \\ &\leq 2^{i-1}\|\boldsymbol{b}_i^*\|^2. \end{aligned}$$

It follows that

$$\|\boldsymbol{b}_j^*\|^2 \leq 2^{j-1}\|\boldsymbol{b}_j^*\|^2 \leq 2^{i-1}\|\boldsymbol{b}_i^*\|^2$$

for $1 \leq j \leq i \leq n$. This proves (i).
(ii) From $d(\Lambda) = |\det(\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n)|$ and (1), it follows that

$$d(\Lambda) = |\det(\boldsymbol{b}_1^*, \cdots, \boldsymbol{b}_n^*)|$$

and therefore, since the $\boldsymbol{b}_i^*$ are pairwise orthogonal

$$d(\Lambda) = \prod_{i=1}^{n} \|\boldsymbol{b}_i^*\|.$$

From $\|\boldsymbol{b}_i^*\| \leq \|\boldsymbol{b}_i\|$ and $\|\boldsymbol{b}_i\| \leq 2^{(i-1)/2}\|\boldsymbol{b}_i^*\|$ we now obtain (ii).
(iii) Putting $j = 1$ in (i) and taking the product over $i = 1, \cdots, n$ we find (iii).
(iv) By Proposition 3.2, for every non-zero $\boldsymbol{x} \in \Lambda$, $\|\boldsymbol{x}\|^2 \geq \|\boldsymbol{b}_i^*\|^2$ for some $i \leq n$. Putting $j = 1$ in (i), we have $\|\boldsymbol{b}_1\|^2 \leq 2^{i-1}\|\boldsymbol{b}_i^*\|^2 \leq 2^{n-1}\|\boldsymbol{b}_i^*\|^2$. This proves (iv).
(v) Write $\boldsymbol{x}_j = \sum_{i=1}^{n} r_{ij}\boldsymbol{b}_i$ with $r_{ij} \in \mathcal{O}_F (1 \leq i \leq n)$ for $1 \leq j \leq t$. For fixed j, let $i(j)$ denote the largest i for which $r_{ij} \neq 0$. Then we have, by the proof of Proposition 3.2,

$$\|\boldsymbol{x}_j\|^2 \geq \|\boldsymbol{b}_{i(j)}^*\|^2 \tag{11}$$

for $1 \leq j \leq t$. Renumber the $\boldsymbol{x}_j$ such that $i(1) \leq \cdots \leq i(t)$. We claim that $j \leq i(j)$ for $1 \leq j \leq t$. If not, then $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_j$ would all belong to $\mathcal{O}_F\boldsymbol{b}_1 + \cdots + \mathcal{O}_F\boldsymbol{b}_{j-1}$, a contradiction with the linear independence of $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_t$. From $j \leq i(j)$ and (i) we obtain, using (11):

$$\|\boldsymbol{b}_j\|^2 \leq 2^{i(j)-1} \cdot \|\boldsymbol{b}_{i(j)}^*\|^2 \leq 2^{n-1} \cdot \|\boldsymbol{b}_{i(j)}^*\|^2 \leq 2^{n-1} \cdot \|\boldsymbol{x}_j\|^2$$

for $j = 1, \cdots, t$. This proves (iv). ■

4 Absolute values of elements in some the rings of integers $\mathcal{O}_F$ In case F is the rational or an imaginary quadratic field the absolute value of the non-zero elements of $\mathcal{O}_F$ is greater than one. The situation is different for general number fields, as we shall show in the sequel.

Let F be a number field of degree n and $\mathcal{O}_F$ denote its ring of integers. It is well-known that $\mathcal{O}_F$ is a free abelian group of rank n.

Using the Pigeonhole Principle, we can prove the following lemma. It is a special case of Dirichlet's simultaneous approximation theorem. The proof is given in [7].

Lemma 4.1 *Suppoose that α and β are real numbers and at least one of α, β is in $\mathbb{R} \setminus \mathbb{Q}$. Then there are infinitely many triads (x, y, z) of integers such that $|x - z\alpha| < 1/\sqrt{z}$ and $|y - z\beta| < 1/\sqrt{z}$.*

Proposition 4.2 *Let L be a free abelian group of rank $n \geq 3$ in $\mathbb{C}$. Then, for any positive real number ϵ, there is a non-zero $z \in L$ such that $|z| < \epsilon$.*

Proof. We may assume that $n = 3$ and $L = \mathbb{Z}\boldsymbol{a} + \mathbb{Z}\boldsymbol{b} + \mathbb{Z}\boldsymbol{e}$. Since $\mathbb{C}$ is a 2-dimensional vector space over $\mathbb{R}$, there exist real numbers α and β such that $\boldsymbol{e} = \alpha\boldsymbol{a} + \beta\boldsymbol{b}$ and at least one of α, β is in $\mathbb{R} \setminus \mathbb{Q}$. By the lemma above, there are integers p, q, r such that $|p\alpha + q| < \epsilon/(\|\boldsymbol{a}\| + \|\boldsymbol{b}\|)$ and $|p\beta + r| < \epsilon/(\|\boldsymbol{a}\| + \|\boldsymbol{b}\|)$. Then we have $\|p\boldsymbol{e} + q\boldsymbol{a} + r\boldsymbol{b}\| = \|(p\alpha + q)\boldsymbol{a} + (p\beta + r)\boldsymbol{b}\| \leq |(p\alpha + q)|\|\boldsymbol{a}\| + |(p\beta + r)|\|\boldsymbol{b}\| < \epsilon$. ■

By similar way, we can prove the following.

Proposition 4.3 *Let L be a free abelian group of rank $n \geq 2$ in $\mathbb{R}$. Then, for any positive real number ϵ, there is a non-zero $z \in L$ such that $|z| < \varepsilon$.*

By these propositions, the ring of integers $\mathcal{O}_F$ has a least positive element, if and only if F is the rational number field or an imaginary quadratic field. Therefore we conclude the following theorem.

Theorem 4.4 *Let F be a number field and $\mathcal{O}_F$ is the ring of integers in F. Then $\mathcal{O}_F$ has a least positive element if and only if F is either the rational number field or an imaginary quadratic field.*

Acknowledgment The authors would like to express their hearty thanks to the referee for his valuable advices.

References

[1] H.Cohen, *A Course in Computational Algebraic Number Theory*, GTM **138**, Springer Verlag, 1993.
[2] A.K.Lenstra, H.W.Lenstra,Jr., and L.Lovász, *Factoring Polynomials with Rational Coefficients*, Math. Ann., **261**, 515-534, 1982.
[3] H.Napias, *A generalization of the LLL-algorithm over euclidean rings or orders*, Journal de Theorie des Nombres de Bordeaux, tome 8, no 2,387-396, 1996.
[4] K.Peter, *The LLL-Algorithm and some Applications*, 2009, available at `http://user.math.uzh.ch/dehaye/thesis_students/Karin`
[5] M.E.Pohst *Computational Algebraic Number Theory*, DMV Seminar **21**, Birkhäuser Verlag, 1993.
[6] M.Pohst and H.Zassenhaus, *Algorithmic Algebraic Number Theory*, Cambridge University Press, 1989.
[7] W.M.Schmidt, *Diophantine Approximation*, LNM **785**, Springer Verlag, 1980.

Communicated by *Pál Dömösi*

Koichi Arimoto:
Joint Graduate School (Ph.D. Program) in Science of School Education,
Hyogo University of Teacher Education,
Kato-Shi, Hyogo 673-1494, Japan

Yasuyuki Hirano:
Department of Mathematics,
Naruto University of Education,
Naruto-Shi, Tokushima 772-8502, Japan

VELOCITY AND ACCELERATION ON THE PATHS $A \natural_t B$ AND $A \sharp_{t,r} B$

DEDICATED TO THE MEMORY OF PROFESSOR TAKAYUKI FURUTA

Hiroshi Isa(1), Masatoshi Ito(2)*, Eizaburo Kamei(3),
Hiroaki Tohyama(4) and Masayuki Watanabe(5)

Received February 1, 2017 ; revised April 24, 2017

Abstract. Let A and B be strictly positive linear operators on a Hilbert space. The derivative of the path $A \natural_t B$ ($t \in \mathbf{R}$) gives the relative operator entropy, that is, $\frac{d}{dt} A \natural_t B = S_t(A|B)$, which we can regard as the velocity function along $A \natural_t B$. The derivative of velocity function is the acceleration function, so we define the acceleration by $\mathcal{A}_t(A|B) = \frac{d}{dt} S_t(A|B)$. In this paper, we discuss properties of $S_t(A|B)$ and $\mathcal{A}_t(A|B)$. Firstly, we interpret some properties of $S_t(A|B)$ concerning interpolational property and the noncommutative ratio from the viewpoint of velocity. Secondly, we show the properties of $\mathcal{A}_t(A|B)$ similar to those of $S_t(A|B)$.

1 Introduction. Let A and B be strictly positive linear operators on a Hilbert space $\mathcal{H}$. An operator T on $\mathcal{H}$ is said to be positive (we denote it by $T \geq 0$) if $(T\xi, \xi) \geq 0$ for all $\xi \in \mathcal{H}$ and T is said to be strictly positive (we denote it by $T > 0$) if T is invertible and positive.

For $A, B > 0$, we define a path $A \natural_t B$ as follows ([2, 3, 6, 8, 14] etc.):

$$A \natural_t B \equiv A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^t A^{\frac{1}{2}} \quad (t \in \mathbf{R}),$$

which is passing through $A = A \natural_0 B$ and $B = A \natural_1 B$. If $t \in [0,1]$, the path $A \natural_t B$ coincides with the weighted geometric operator mean denoted by $A \sharp_t B$ (cf. [15]). We remark that $A \natural_t B = B \natural_{1-t} A$ holds for $t \in \mathbf{R}$ (cf. [8]).

Fujii and Kamei [1] defined the following relative operator entropy for $A, B > 0$:

$$S(A|B) \equiv A^{\frac{1}{2}} \log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}}.$$

Furuta [7] defined generalized relative operator entropy as follows (see also [9]):

$$\begin{aligned} S_\alpha(A|B) &\equiv A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^\alpha \log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}} \\ &= (A \natural_\alpha B)A^{-1}S(A|B) \quad (\alpha \in \mathbf{R}). \end{aligned}$$

We know immediately $S_0(A|B) = S(A|B)$. We remark that

$$S(A|B) = \frac{d}{dt} A \natural_t B\Big|_{t=0} \quad \text{and} \quad S_\alpha(A|B) = \frac{d}{dt} A \natural_t B\Big|_{t=\alpha}.$$

Yanagi, Kuriyama and Furuichi [16] introduced the Tsallis relative operator entropy as follows:

$$T_\alpha(A|B) \equiv \frac{A \sharp_\alpha B - A}{\alpha} \quad (\alpha \in (0,1]).$$

* *This work was supported by JSPS KAKENHI Grant Number JP16K05181.*
2010 *Mathematics Subject Classification.* 47A63, 47A64 and 94A17.
Key words and phrases. velocity, acceleration, path, relative operator entropy.

Since $\lim\limits_{x\to 0}\dfrac{a^x-1}{x}=\log a$ holds for $a>0$, we have $T_0(A|B)\equiv\lim\limits_{\alpha\to 0}T_\alpha(A|B)=S(A|B)$. The Tsallis relative operator entropy can be defined for any $\alpha\in\mathbf{R}$ by using $\natural_\alpha$ instead of $\sharp_\alpha$.

For $A,B>0$, $t\in[0,1]$ and $r\in[-1,1]$, operator power mean $A\ \sharp_{t,r}\ B$ is defined as follows:

$$A\ \sharp_{t,r}\ B\equiv A^{\frac{1}{2}}\left\{(1-t)I+t(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^r\right\}^{\frac{1}{r}}A^{\frac{1}{2}}\ =\ A\ \natural_{\frac{1}{r}}\ \{A\ \nabla_t\ (A\ \natural_r\ B)\}.$$

We remark that $A\ \sharp_{t,r}\ B=B\ \sharp_{1-t,r}\ A$ holds for $t\in[0,1]$ and $r\in[-1,1]$ (cf. [10, 12]). The operator power mean is a path combining $A=A\ \sharp_{0,r}\ B$ and $B=A\ \sharp_{1,r}\ B$, and interpolates the arithmetic operator mean, the geometric operator mean and the harmonic operator mean.

arithmetic operator mean
$A\ \nabla_t\ B=(1-t)A+tB$
$\uparrow_{r=1}$
$A\ \sharp_{t,r}\ B\ \xrightarrow[r\to 0]{}$ geometric operator mean $A\ \sharp_t\ B=A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^tA^{\frac{1}{2}}$
$\downarrow_{r=-1}$
harmonic operator mean
$A\ \Delta_t\ B=\left(A^{-1}\ \nabla_t\ B^{-1}\right)^{-1}$

For $A,B>0$, $\alpha\in[0,1]$ and $r\in[-1,1]$, expanded relative operator entropy $S_{\alpha,r}(A|B)$ is defined as follows (cf. [10]):

$$\begin{aligned}
S_{\alpha,r}(A|B)\ &\equiv\ \frac{d}{dt}\ A\ \sharp_{t,r}\ B\,\Big|_{t=\alpha}\\
&=\ A^{\frac{1}{2}}\left[\left\{(1-\alpha)I+\alpha\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^r\right\}^{\frac{1}{r}-1}\frac{(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^r-I}{r}\right]A^{\frac{1}{2}}\\
&=\ (A\ \sharp_{\alpha,r}\ B)(A\ \nabla_\alpha\ (A\ \natural_r\ B))^{-1}S_{0,r}(A|B)\ (r\neq 0),\\
S_{\alpha,0}(A|B)\ &\equiv\ \lim_{r\to 0}S_{\alpha,r}(A|B)\ =\ S_\alpha(A|B).
\end{aligned}$$

We remark that $S_{0,r}(A|B)=T_r(A|B)$, $S_{1,r}(A|B)=-T_r(B|A)$ hold for $r\in[-1,1]$.

$S(A|B)$ and $S_{0,r}(A|B)$ are given as follows:

$$S(A|B)\ =\ \frac{d}{dt}\ A\ \natural_t\ B\Big|_{t=0}\ \text{ and }\ S_{0,r}(A|B)\ =\ \frac{d}{dt}\ A\ \sharp_{t,r}\ B\Big|_{t=0}.$$

We illustrate an image for $S(A|B)$ and $S_{0,r}(A|B)$ in Figure 1.

In [6], $S(A|B)$ and $S_{0,r}(A|B)$ are regarded as the velocities on the paths $A\ \sharp_t\ B$ and $A\ \sharp_{t,r}\ B$ at $t=0$ respectively. According to this viewpoint, it is natural to call $S_\alpha(A|B)$ and $S_{\alpha,r}(A|B)$ the velocities on the paths $A\ \natural_t\ B$ and $A\ \sharp_{t,r}\ B$ respectively. These interpretations inspire us to introduce the accelerations $\mathcal{A}_\alpha(A|B)$ and $\mathcal{A}_{\alpha,r}(A|B)$ on the paths $A\ \natural_t\ B$ and $A\ \sharp_{t,r}\ B$ at $t=\alpha$.

In this paper, we can show that the properties concerning the accelerations $\mathcal{A}_\alpha(A|B)$ and $\mathcal{A}_{\alpha,r}(A|B)$, interpolational property, the behavior of noncommutative ratio and so on, are inherited from those of velocities $S_\alpha(A|B)$ and $S_{\alpha,r}(A|B)$. The contents of this paper are as follows: In section 2, we show properties of the velocity $S_\alpha(A|B)$. In section 3, we introduce the acceleration $\mathcal{A}_\alpha(A|B)$ and we show some properties of $\mathcal{A}_\alpha(A|B)$. In section 4, we introduce the acceleration on the path $A \sharp_{t,r} B$ and we show some results for velocity $S_{\alpha,r}(A|B)$ and acceleration $\mathcal{A}_{\alpha,r}(A|B)$ on the path $A \sharp_{t,r} B$.

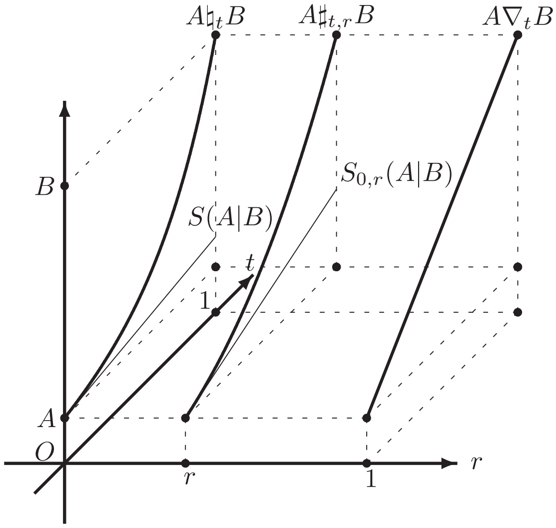

Figure 1: An image of $S(A|B)$ and $S_{0,r}(A|B)$.

2 Velocity on the path $A \natural_t B$. As mentioned in section 1, we regard $S_\alpha(A|B)$ as the velocity on the path $A \natural_t B$ at $t = \alpha$. In this section, we show some properties of the velocity $S_\alpha(A|B)$.

The next lemma shows interpolational property of the path $A \natural_t B$. This lemma is fundamental in our discussion.

Lemma 2.1. ([12]) *For $A, B > 0$ and $x, y, \alpha \in \mathbf{R}$,*

$$(A \natural_y B) \natural_\alpha (A \natural_x B) \;=\; A \natural_{(1-\alpha)y+\alpha x} B$$

holds.

Let $A \natural_x B$ and $A \natural_y B$ ($x, y \in \mathbf{R}$) be arbitrary points on the path $A \natural_t B$. Concerning the velocity $S_\alpha(A \natural_y B|A \natural_x B)$ at $t = \alpha$, we have the following theorem which was proved in [12].

Theorem 2.2. ([12]) *Let $A, B > 0$ and $\alpha, x, y \in \mathbf{R}$. Then*

$$S_\alpha(A \natural_y B|A \natural_x B) \;=\; (x-y)S_{(1-\alpha)y+\alpha x}(A|B).$$

In our discussion, for a given path $\gamma(t) = X \natural_t Y$ for $X, Y > 0$ and $t \in \mathbf{R}$, we imagine that an object moves through base points X ($t = 0$) and Y ($t = 1$) on the path $\gamma(t)$. Then $S_\alpha(A \natural_y B|A \natural_x B)$ means the velocity on the path $\gamma_1(t) = (A \natural_y B) \natural_t (A \natural_x B)$ at $t = \alpha$, and also $S_{(1-\alpha)y+\alpha x}(A|B)$ means the velocity on the path $\gamma_2(t) = A \natural_t B$ at $t = (1-\alpha)y + \alpha x$. Note that $\gamma_1(t)$ and $\gamma_2(t)$ represent the same path, and the point on $\gamma_1(t)$ at $t = \alpha$ and the point on $\gamma_2(t)$ at $t = (1-\alpha)y + \alpha x$ are the same point by Lemma 2.1. In this situation, we consider exchanging the path $\gamma_1(t)$ for $\gamma_2(t)$ to change unit length of the path. Then we can regard Theorem 2.2 as the result on the rate of change of velocities at the same point.

The next Corollary 2.3 is an immediate consequence of Theorem 2.2.

Corollary 2.3. *For $A, B > 0$ and $\alpha, x, y \in \mathbf{R}$, the following hold:*

(1) $S_\alpha(B|A) \;=\; -S_{1-\alpha}(A|B)$.

(2) $S_\alpha(A|A \natural_x B) \;=\; xS_{\alpha x}(A|B)$.

(3) $S_\alpha(A \natural_y B|A \natural_{y+1} B) \;=\; S_{\alpha+y}(A|B)$.

Proof. (1) is obtained by putting $x = 0$ and $y = 1$, (2) is obtained by putting $y = 0$ and (3) is obtained by putting $x = y + 1$. □

Next, from the above point of view, we discuss the noncommutative ratio $\mathcal{R}(v; A, B) \equiv (A \natural_v B)A^{-1}$ for $v \in \mathbf{R}$ which is defined in [11]. Note that it is independent of α in $S_\alpha(A|B)$.

Theorem 2.4. ([11]) *For $A, B > 0$ and $v \in \mathbf{R}$,*

$$\mathcal{R}(v; A, B)S_\alpha(A|B) \;=\; S_{\alpha+v}(A|B)$$

for all $\alpha \in \mathbf{R}$.

In particular, by putting $\alpha = 0$ in Theorem 2.4, we have following relation.

Corollary 2.5. ([11]) *For $A, B > 0$ and $v \in \mathbf{R}$, following hold:*

$$\mathcal{R}(v; A, B)S(A|B) \;=\; S_v(A|B).$$

By Theorem 2.4 and (3) in Corollary 2.3, we have

$$(\heartsuit) \qquad \mathcal{R}(v; A, B)S_\alpha(A|B) = S_\alpha(A \natural_v B|A \natural_{v+1} B) \quad \text{for } \alpha, v \in \mathbf{R}.$$

As an extension of this relation, we obtain the following Theorem 2.6. Here, we consider exchanging the path $\gamma_1(t) = (A \natural_y B) \natural_t (A \natural_x B)$ for $\gamma_2(t) = (A \natural_{y+v} B) \natural_t (A \natural_{x+v} B)$, that is, moving base points of the path preserving unit length. Then, Theorem 2.6 shows a relation between velocity on the path $\gamma_1(t)$ at $t = \alpha$ and velocity on the path $\gamma_2(t)$ at $t = \alpha$ by using the noncommutative ratio.

Theorem 2.6. *Let $A, B > 0$ and $\alpha, v, x, y \in \mathbf{R}$. Then*

$$\mathcal{R}(v; A, B)S_\alpha(A \natural_y B|A \natural_x B) \;=\; S_\alpha(A \natural_{y+v} B|A \natural_{x+v} B).$$

Proof. By Theorem 2.2 and Theorem 2.4, we have

$$\begin{aligned}
\mathcal{R}(v; A, B)S_\alpha(A \natural_y B|A \natural_x B) &= (x-y)\mathcal{R}(v; A, B)S_{(1-\alpha)y+\alpha x}(A|B) \\
&= (x-y)S_{(1-\alpha)y+\alpha x+v}(A|B) \\
&= \{(x+v)-(y+v)\}\, S_{(1-\alpha)(y+v)+\alpha(x+v)}(A|B) \\
&= S_\alpha(A \natural_{y+v} B|A \natural_{x+v} B).
\end{aligned}$$

□

Remark. We know that $\mathcal{R}(v(x-y); A, B) \;=\; \mathcal{R}(v; A \natural_y B, A \natural_x B)$ holds for $A, B > 0$ and $v, x, y \in \mathbf{R}$, since

$$\begin{aligned}
\mathcal{R}(v(x-y); A, B) &= (A \natural_{v(x-y)} B)A^{-1} \\
&= A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{v(x-y)+y}A^{\frac{1}{2}}A^{-\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{-y}A^{-\frac{1}{2}} \\
&= (A \natural_{(1-v)y+vx} B)(A \natural_y B)^{-1} \;=\; ((A \natural_y B) \natural_v (A \natural_x B))\,(A \natural_y B)^{-1} \\
&= \mathcal{R}(v; A \natural_y B, A \natural_x B).
\end{aligned}$$

From this relation and ($\heartsuit$), we can give an alternative proof of Theorem 2.6 as follows: By putting $u = v(x-y)$, we have $\mathcal{R}(u; A, B) = \mathcal{R}(\frac{u}{x-y}; A \natural_y B, A \natural_x B)$. Then

$$\begin{aligned}\mathcal{R}(u; A, B) S_\alpha(A \natural_y B | A \natural_x B) &= \mathcal{R}\left(\frac{u}{x-y}; A \natural_y B, A \natural_x B\right) S_\alpha(A \natural_y B | A \natural_x B)\\ &= S_\alpha((A \natural_y B) \natural_{\frac{u}{x-y}} (A \natural_x B) | (A \natural_y B) \natural_{\frac{u}{x-y}+1} (A \natural_x B))\\ &= S_\alpha(A \natural_{(1-\frac{u}{x-y})y + \frac{ux}{x-y}} B | A \natural_{-\frac{uy}{x-y} + (\frac{u}{x-y}+1)x} B)\\ &= S_\alpha(A \natural_{y+u} B | A \natural_{x+u} B).\end{aligned}$$

Corollary 2.5 means that $\mathcal{R}(v; A, B)$ is the ratio of $S_v(A|B)$ and $S(A|B)$. Related to it, the difference between $S_v(A|B)$ and $S(A|B)$ is as follows:

Proposition 2.7. *For $A, B > 0$ and $v \in \mathbf{R}$,*

$$S_v(A|B) - S(A|B) = v T_v(A|B) A^{-1} S(A|B)$$

holds.

Proof. From Corollary 2.5, we have

$$\begin{aligned}S_v(A|B) - S(A|B) &= \mathcal{R}(v; A, B) S(A|B) - S(A|B)\\ &= (A \natural_v B - A) A^{-1} S(A|B) = v T_v(A|B) A^{-1} S(A|B).\end{aligned}$$

□

We remark that the above difference was also represented by using Petz-Bregman divergence (see [13]).

3 Acceleration on the path $A \natural_t B$. Since the relative operator entropy $S_\alpha(A|B)$ is regarded as the velocity on the path $A \natural_t B$ at $t = \alpha$, it is natural to call the derivative of $S_t(A|B)$ acceleration on $A \natural_t B$.

Definition 3.1. *For $A, B > 0$ and $\alpha \in \mathbf{R}$, we define the acceleration on the path $A \natural_t B$ at $t = \alpha$ as follows:*

$$\mathcal{A}_\alpha(A|B) \equiv \frac{d}{dt} S_t(A|B)\bigg|_{t=\alpha}.$$

The acceleration $\mathcal{A}_\alpha(A|B)$ is represented explicitly as follows:

Theorem 3.2. *Let $A, B > 0$ and $\alpha \in \mathbf{R}$. Then*

$$\mathcal{A}_\alpha(A|B) = S_\alpha(A|B) A^{-1} S(A|B) = S_\alpha(A|B)(A \natural_\alpha B)^{-1} S_\alpha(A|B).$$

In particular,

$$\mathcal{A}_0(A|B) = S(A|B) A^{-1} S(A|B).$$

Proof. For $a > 0$, we have

$$\frac{d}{dt} a^t \log a = a^t (\log a)^2.$$

Then

$$\begin{aligned}\mathcal{A}_\alpha(A|B) &= \frac{d}{dt}S_t(A|B)\Big|_{t=\alpha} \\ &= A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^\alpha(\log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}))^2A^{\frac{1}{2}} \\ &= A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^\alpha\log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}}A^{-1}A^{\frac{1}{2}}\log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}} \\ &= S_\alpha(A|B)A^{-1}S(A|B),\end{aligned}$$

which shows the first equality. On the other hand, we have

$$\begin{aligned}S_\alpha(A|B)A^{-1}S(A|B) &= S_\alpha(A|B)(A \natural_\alpha B)^{-1}(A \natural_\alpha B)A^{-1}S(A|B) \\ &= S_\alpha(A|B)(A \natural_\alpha B)^{-1}S_\alpha(A|B).\end{aligned}$$

□

Remark. Theorem 3.2 shows that if we put $\gamma(t) = A \natural_t B$, then it satisfies the geodesic equation $\ddot{\gamma}(t) - \dot{\gamma}(t)\big(\gamma(t)\big)^{-1}\dot{\gamma}(t) = 0$ since $\dot{\gamma}(t) = S_t(A|B)$ and $\ddot{\gamma}(t) = \mathcal{A}_t(A|B)$. Conversely, $A \natural_t B$ is given as the solution of the geodesic equation for initial conditions $\gamma(0) = A$ and $\gamma(1) = B$. We show it here according to [5] which treats matrices, but the same arguments are valid for operator valued functions, since, even for a operator valued function $\gamma(t)$, it holds that $(\gamma(t)^{-1})' = -(\gamma(t))^{-1}\gamma(t)'(\gamma(t))^{-1}$ and that $(\log\gamma(t))' = \gamma(t)'(\gamma(t))^{-1}$ if $\gamma(t)\gamma(t)' = \gamma(t)'\gamma(t)$.

By putting $f(t) = \gamma(0)^{-\frac{1}{2}}\gamma(t)\gamma(0)^{-\frac{1}{2}} = A^{-\frac{1}{2}}\gamma(t)A^{-\frac{1}{2}}$, we have

$$f''(t) - f'(t)\big(f(t)\big)^{-1}f'(t) = 0$$

and that $f(0) = A^{-\frac{1}{2}}\gamma(0)A^{-\frac{1}{2}} = I$ and $f(1) = A^{-\frac{1}{2}}\gamma(1)A^{-\frac{1}{2}} = A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$. Since

$$\begin{aligned}\big(f'(t)(f(t))^{-1}\big)' &= f''(t)(f(t))^{-1} - f'(t)(f(t))^{-1}f'(t)(f(t))^{-1} \\ &= f''(t)(f(t))^{-1} - f''(t)(f(t))^{-1} = 0,\end{aligned}$$

then we have $f'(t)(f(t))^{-1} = C$, that is, $f'(t) = Cf(t)$. It is known that $f(t)$ and $f'(t)$ are selfadjoint, so we have

$$C^* = f(0)C^* = (Cf(0))^* = (f'(0))^* = f'(0) = Cf(0) = C.$$

Hence

$$f'(t)(f(t))^{-1} = C = C^* = (f'(t)(f(t))^{-1})^* = (f(t))^{-1}f'(t),$$

and then $f'(t)f(t) = f(t)f'(t)$. So we have $(\log f(t))' = f'(t)(f(t))^{-1} = C$ and then $\log f(t) = Ct + D$. By $f(0) = I$ and $f(1) = A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$, we have $\exp C = A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$ and $D = 0$, that is, $f(t) = (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^t$. Therefore, we obtain

$$\gamma(t) = A^{\frac{1}{2}}f(t)A^{\frac{1}{2}} = A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^tA^{\frac{1}{2}} = A \natural_t B.$$

Through Theorem 3.2, we know that the acceleration $\mathcal{A}_\alpha(A|B)$ has the similar properties to the velocity $S_\alpha(A|B)$. First, we have Theorem 3.3 which corresponds to Theorem 2.2. As mentioned in section 2, the point on the path $\gamma_1(t) = (A \natural_y B) \natural_t (A \natural_x B)$ at $t = \alpha$ and the point on $\gamma_2(t) = A \natural_t B$ at $t = (1-\alpha)y + \alpha x$ are the same point. Then Theorem 3.3 shows the result on the rate of change of accelerations at the same point.

Theorem 3.3. *Let $A, B > 0$ and $\alpha, x, y \in \mathbf{R}$. Then*
$$\mathcal{A}_\alpha(A \natural_y B|A \natural_x B) = (x-y)^2\mathcal{A}_{(1-\alpha)y+\alpha x}(A|B).$$
Proof. By Theorem 3.2, Lemma 2.1 and Theorem 2.2, we have
$$\begin{aligned}
&\mathcal{A}_\alpha(A \natural_y B|A \natural_x B)\\
&\quad = S_\alpha(A \natural_y B|A \natural_x B)((A \natural_y B) \natural_\alpha (A \natural_x B))^{-1}S_\alpha(A \natural_y B|A \natural_x B)\\
&\quad = (x-y)^2 S_{(1-\alpha)y+\alpha x}(A|B)(A \natural_{(1-\alpha)y+\alpha x} B)^{-1}S_{(1-\alpha)y+\alpha x}(A|B)\\
&\quad = (x-y)^2\mathcal{A}_{(1-\alpha)y+\alpha x}(A|B).
\end{aligned}$$
□

Corollary 3.4. *For $A, B > 0$, and $\alpha, x, y \in \mathbf{R}$, the following hold:*

(1) $\mathcal{A}_\alpha(B|A) = \mathcal{A}_{1-\alpha}(A|B)$.

(2) $\mathcal{A}_\alpha(A|A \natural_x B) = x^2\mathcal{A}_{\alpha x}(A|B)$.

(3) $\mathcal{A}_\alpha(A \natural_y B|A \natural_{y+1} B) = \mathcal{A}_{\alpha+y}(A|B)$.

Secondly, related to the noncommutative ratio, we have Theorem 3.5 which corresponds to Theorem 2.4.

Theorem 3.5. *For $A, B > 0$ and $v \in \mathbf{R}$,*
$$\mathcal{R}(v; A, B)\mathcal{A}_\alpha(A|B) = \mathcal{A}_{\alpha+v}(A|B)$$
for all $\alpha \in \mathbf{R}$. In particular,
$$\mathcal{R}(v; A, B)\mathcal{A}_0(A|B) = \mathcal{A}_v(A|B).$$

Proof. By Theorem 3.2 and Theorem 2.4, we have
$$\begin{aligned}
\mathcal{R}(v; A, B)\mathcal{A}_\alpha(A|B) &= \mathcal{R}(v; A, B)S_\alpha(A|B)A^{-1}S(A|B)\\
&= S_{\alpha+v}(A|B)A^{-1}S(A|B) = \mathcal{A}_{\alpha+v}(A|B).
\end{aligned}$$
□

The following Theorem 3.6 is an extension of Theorem 3.5. Similarly to Theorem 2.6, Theorem 3.6 shows a relation between acceleration on the paths $\gamma_1(t) = (A \natural_y B) \natural_t (A \natural_x B)$ and $\gamma_2(t) = (A \natural_{y+v} B) \natural_t (A \natural_{x+v} B)$ at $t = \alpha$ by using the noncommutative ratio.

Theorem 3.6. *Let $A, B > 0$ and $\alpha, v, x, y \in \mathbf{R}$. Then*
$$\mathcal{R}(v; A, B)\mathcal{A}_\alpha(A \natural_y B|A \natural_x B) = \mathcal{A}_\alpha(A \natural_{y+v} B|A \natural_{x+v} B).$$

Proof. By Theorem 3.3 and Theorem 3.5, we have
$$\begin{aligned}
\mathcal{R}(v; A, B)\mathcal{A}_\alpha(A \natural_y B|A \natural_x B) &= (x-y)^2\mathcal{R}(v; A, B)\mathcal{A}_{(1-\alpha)y+\alpha x}(A|B)\\
&= (x-y)^2\mathcal{A}_{(1-\alpha)y+\alpha x+v}(A|B)\\
&= \{(x+v)-(y+v)\}^2\mathcal{A}_{(1-\alpha)(y+v)+\alpha(x+v)}(A|B)\\
&= \mathcal{A}_\alpha(A \natural_{y+v} B|A \natural_{x+v} B).
\end{aligned}$$
□

Lastly, the difference between $\mathcal{A}_v(A|B)$ and $\mathcal{A}_0(A|B)$ is gotten as follows.

Proposition 3.7. *For $A, B > 0$ and $v \in \mathbf{R}$,*

$$\mathcal{A}_v(A|B) - \mathcal{A}_0(A|B) \ = \ vT_v(A|B)A^{-1}\mathcal{A}_0(A|B).$$

holds.

Proof. By using Theorem 3.5, we have

$$\begin{aligned}\mathcal{A}_v(A|B) - \mathcal{A}_0(A|B) &= \mathcal{R}(v; A, B)\mathcal{A}_0(A|B) - \mathcal{A}_0(A|B) \\ &= (A \natural_v B - A)A^{-1}\mathcal{A}_0(A|B) \\ &= vT_v(A|B)A^{-1}\mathcal{A}_0(A|B).\end{aligned}$$

□

4 Velocity and acceleration on the path $A \sharp_{t,r} B$. In this section, we introduce the velocity and the acceleration on the path $A \sharp_{t,r} B$ and show their properties.

We know that the path $A \sharp_{t,r} B$ has interpolational property. The next lemma is the same property as Lemma 2.1.

Lemma 4.1. ([14]) *For $A, B > 0$, $\alpha, x, y \in [0, 1]$ and $r \in [-1, 1]$,*

$$(A \sharp_{y,r} B) \sharp_{\alpha,r} (A \sharp_{x,r} B) \ = \ A \sharp_{(1-\alpha)y+\alpha x,r} B$$

holds.

Although the noncommutative ratio discussed in section 2 can not be extended totally to the one concerning $A \sharp_{t,r} B$, the property Corollary 2.5 is extended as follows:

Theorem 4.2. ([12]) *For $A, B > 0$, $\alpha \in [0, 1]$ and $r \in [-1, 1]$,*

$$S_{\alpha,r}(A|B) \ = \ (A \sharp_{\alpha,r} B)(A \nabla_\alpha (A \natural_r B))^{-1}S_{0,r}(A|B).$$

holds.

We introduce the acceleration on the path $A \sharp_{t,r} B$ as follows:

Definition 4.3. *For $A, B > 0$, $\alpha \in [0, 1]$ and $r \in [-1, 1]$, we define $\mathcal{A}_{\alpha,r}(A|B)$ as*

$$\mathcal{A}_{\alpha,r}(A|B) \ \equiv \ \frac{d}{dt}S_{t,r}(A|B)\bigg|_{t=\alpha}.$$

We call it the acceleration on the path $A \sharp_{t,r} B$ at $t = \alpha$.

We remark that $\mathcal{A}_{\alpha,0}(A|B) \ = \ \mathcal{A}_\alpha(A|B)$ for $\alpha \in [0, 1]$ since $S_{t,0}(A|B) = S_t(A|B)$.

The acceleration $\mathcal{A}_{\alpha,r}(A|B)$ is represented explicitly as follows:

Theorem 4.4. *Let $A, B > 0$, $\alpha \in [0, 1]$ and $r \in [-1, 1]$. Then*

$$\begin{aligned}\mathcal{A}_{\alpha,r}(A|B) &= (1 - r)S_{\alpha,r}(A|B)(A \nabla_\alpha (A \natural_r B))^{-1}S_{0,r}(A|B) \\ &= (1 - r)S_{\alpha,r}(A|B)(A \sharp_{\alpha,r} B)^{-1}S_{\alpha,r}(A|B).\end{aligned}$$

In particular,

$$\mathcal{A}_{0,r}(A|B) \ = \ (1 - r)S_{0,r}(A|B)A^{-1}S_{0,r}(A|B).$$

Proof. We have shown the case $r = 0$ in Theorem 3.2. Hence, we have only to show the case $r \neq 0$. Since

$$\begin{aligned}
&\frac{d}{dt} S_{t,r}(A|B) \\
&= (1-r)A^{\frac{1}{2}} \left\{(1-t)I + t\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^r\right\}^{\frac{1}{r}-2} \left(\frac{(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^r - I}{r}\right)^2 A^{\frac{1}{2}} \\
&= (1-r)A^{\frac{1}{2}} \left[\left\{(1-t)I + t\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^r\right\}^{\frac{1}{r}-1} \frac{(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^r - I}{r}\right] A^{\frac{1}{2}} \\
&\quad \times A^{-\frac{1}{2}} \left\{(1-t)I + t\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^r\right\}^{-1} A^{-\frac{1}{2}} A^{\frac{1}{2}} \frac{(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^r - I}{r} A^{\frac{1}{2}} \\
&= (1-r)S_{t,r}(A|B)(A \nabla_t (A \natural_r B))^{-1} T_r(A|B) \\
&= (1-r)S_{t,r}(A|B)(A \nabla_t (A \natural_r B))^{-1} S_{0,r}(A|B),
\end{aligned}$$

we have

$$\mathcal{A}_{\alpha,r}(A|B) = (1-r)S_{\alpha,r}(A|B)(A \nabla_\alpha (A \natural_r B))^{-1} S_{0,r}(A|B).$$

On the other hand, by Theorem 4.2, we have

$$(1-r)S_{\alpha,r}(A|B)(A \nabla_\alpha (A \natural_r B))^{-1} S_{0,r}(A|B) = (1-r)S_{\alpha,r}(A|B)(A \sharp_{\alpha,r} B)^{-1} S_{\alpha,r}(A|B).$$

□

By Theorem 4.4 and Lemma 4.1, we give similar properties to those in sections 2 and 3. First, we have the next theorem and corollary.

Theorem 4.5. *Let $A, B > 0$, $\alpha, x, y \in [0,1]$ and $r \in [-1,1]$. Then*

(1) $S_{\alpha,r}(A \sharp_{y,r} B | A \sharp_{x,r} B) = (x-y) S_{(1-\alpha)y+\alpha x,r}(A|B)$.

(2) $\mathcal{A}_{\alpha,r}(A \sharp_{y,r} B | A \sharp_{x,r} B) = (x-y)^2 \mathcal{A}_{(1-\alpha)y+\alpha x,r}(A|B)$.

Proof. (1) By using Lemma 4.1, we have

$$\begin{aligned}
&S_{\alpha,r}(A \sharp_{y,r} B | A \sharp_{x,r} B) \\
&= \frac{d}{dt} (A \sharp_{y,r} B) \sharp_{t,r} (A \sharp_{x,r} B)\Big|_{t=\alpha} \\
&= \lim_{v\to 0} \frac{(A \sharp_{y,r} B) \sharp_{\alpha+v,r} (A \sharp_{x,r} B) - (A \sharp_{y,r} B) \sharp_{\alpha,r} (A \sharp_{x,r} B)}{v} \\
&= \lim_{v\to 0} \frac{A \sharp_{(1-(\alpha+v))y+(\alpha+v)x,r} B - A \sharp_{(1-\alpha)y+\alpha x,r} B}{v} \\
&= (x-y) \lim_{v\to 0} \frac{A \sharp_{(1-\alpha)y+\alpha x+v(x-y),r} B - A \sharp_{(1-\alpha)y+\alpha x,r} B}{(x-y)v} \\
&= (x-y) S_{(1-\alpha)y+\alpha x,r}(A|B).
\end{aligned}$$

(2) From Theorem 4.4, (1) in Theorem 4.5 and Lemma 4.1, we obtain

$$\begin{aligned}
&\mathcal{A}_{\alpha,r}(A \sharp_{y,r} B | A \sharp_{x,r} B) \\
&= (1-r)S_{\alpha,r}(A \sharp_{y,r} B | A \sharp_{x,r} B)\big((A \sharp_{y,r} B) \sharp_{\alpha,r} (A \sharp_{x,r} B)\big)^{-1} S_{\alpha,r}(A \sharp_{y,r} B | A \sharp_{x,r} B) \\
&= (x-y)^2 (1-r) S_{(1-\alpha)y+\alpha x,r}(A|B)(A \sharp_{(1-\alpha)y+\alpha x,r} B)^{-1} S_{(1-\alpha)y+\alpha x,r}(A|B) \\
&= (x-y)^2 \mathcal{A}_{(1-\alpha)y+\alpha x,r}(A|B).
\end{aligned}$$

□

Corollary 4.6. *For $A, B > 0$, $\alpha, x \in [0, 1]$ and $r \in [-1, 1]$, the following hold:*

(1) $S_{\alpha,r}(B|A) = -S_{1-\alpha,r}(A|B)$ *and* $S_{\alpha,r}(A|A \natural_{x,r} B) = xS_{\alpha x,r}(A|B)$.

(2) $\mathcal{A}_{\alpha,r}(B|A) = \mathcal{A}_{1-\alpha,r}(A|B)$ *and* $\mathcal{A}_{\alpha,r}(A|A \natural_{x,r} B) = x^2\mathcal{A}_{\alpha x,r}(A|B)$.

Acknowledgements. The authors would like to express their hearty thanks to Professor Jun Ichi Fujii for his kind advices.

REFERENCES

[1] J. I. Fujii and E. Kamei, Relative operator entropy in noncommutative information theory, Math. Japon., **34**(1989), 341–348.

[2] J. I. Fujii and E. Kamei, Interpolational paths and their derivatives, Math. Japon., **39**(1994), 557–560.

[3] J. I. Fujii and E. Kamei, Path of Bregman-Petz operator divergence, Sci. Math. Jpn., **70**(2009), 329–333.

[4] J. I. Fujii, Structure of Hiai-Petz parametrized geometry for positive definite matrices, Linear Algebra Appl., **432**(2010), 318–326.

[5] J. I. Fujii, Manifold of the Positive-definite Matrices and its Geodesic (in Japanese), Memoirs of Osaka Kyoiku University, Ser. III Natural Science and Applied Science, **59**(2010), 1–14.

[6] M. Fujii, J. Mićić, J. Pečarić and Y. Seo, Recent Developments of Mond-Pečarić Method in Operator Inequalities, Monographs in Inequalities 4, Element, Zagreb, 2012.

[7] T. Furuta, Parametric extensions of Shannon inequality and its reverse one in Hilbert space operators, Linear Algebra Appl., **381**(2004), 219–235.

[8] T. Furuta, J. M. Hot, J. Pečarić and Y. Seo, Mond-Pečarić Method in Operator Inequalities, Monographs in Inequalities 1, Element, Zagreb, 2005.

[9] H. Isa, M. Ito, E. Kamei, H. Tohyama and M. Watanabe, Relative operator entropy, operator divergence and Shannon inequality, Sci. Math. Jpn., **75**(2012), 289–298.

[10] H. Isa, M. Ito, E. Kamei, H. Tohyama and M. Watanabe, Extensions of Tsallis relative operator entropy and operator valued distance, Sci. Math. Jpn., **76**(2013), 427–435.

[11] H. Isa, M. Ito, E. Kamei, H. Tohyama and M. Watanabe, On relations between operator valued α-divergence and relative operator entropies, Sci. Math. Jpn., **78**(2015), 215–228. (online: e-2015 (2015), 215–228.)

[12] H. Isa, M. Ito, E. Kamei, H. Tohyama and M. Watanabe, Expanded relative operator entropies and operator valued α-divergence, J. Math. Syst. Sci., **5**(2015), 215–224.

[13] H. Isa, M. Ito, E. Kamei, H. Tohyama and M. Watanabe, Some operator divergences based on Petz-Bregman divergence, Sci., Math., Jpn. **80**(2017), 161-170.

[14] E. Kamei, Paths of operators parametrized by operator means, Math. Japon., **39**(1994), 395–400.

[15] F. Kubo and T. Ando, Means of positive linear operators, Math Ann., **248**(1980), 205–224.

[16] K. Yanagi, K. Kuriyama and S. Furuichi, Generalized Shannon inequalities based on Tsallis relative operator entropy, Linear Algebra Appl., **394**(2005), 109–118.

Communicated by *Junichi Fujii*

[1] MAEBASHI INSTITUTE OF TECHNOLOGY, 460-1, KAMISADORI-MACHI, MAEBASHI, GUNMA, JAPAN, 371-0816. isa@maebashi-it.ac.jp

[2] MAEBASHI INSTITUTE OF TECHNOLOGY, 460-1, KAMISADORI-MACHI, MAEBASHI, GUNMA, JAPAN, 371-0816. m-ito@maebashi-it.ac.jp

[3] 1-1-3, Sakuragaoka, Kanmakicho, Kitakaturagi-gun, Nara, Japan, 639-0202. ekamei1947@yahoo.co.jp

[4] Maebashi Institute of Technology, 460-1, Kamisadori-machi, Maebashi, Gunma, Japan, 371-0816. tohyama@maebashi-it.ac.jp

[5] Maebashi Institute of Technology, 460-1, Kamisadori-machi, Maebashi, Gunma, Japan, 371-0816. masayukiwatanabe@maebashi-it.ac.jp

ON A MISUSE OF SUFFICIENT STATISTICS IN THE EXPONENTIAL FAMILY

Etsuo Hamada

Received Febuary 24, 2017

Abstract. With respect to the sufficient statistics in the transformed exponential family based on a continuous probability distribution, we examine a misuse that the sufficient statistics ought to be distributed with a k-dimensional exponential family where k is the dimension of the sufficient statistics. Under the irreducibility of the sufficient statistics, we define two types of the transformed exponential family, i.e., regular and pseudo, so that the misuse is made explicit.

1 Introduction The exponential and curved exponential families cover a wide range of distributions ([4], [5], [14], [15]), and are widely used for generalized mixed linear models [7], and for methods in information geometry ([1], [2], [3], [6], [8], [13]).

Under the framework of information geometry, it is well assumed that the sufficient statistics are *linearly independent* in order to make a one-to-one correspondence between the parameter and the density [12] and that the dimension of parameters are equal to the dimension of the sufficient statistics which are linearly independent under the duality in the statistical manifold. [10] showed that linear independence of the score function is not a sufficient condition for a distribution to belong to the curved exponential family, showing that there exists a gap between the parameter space and the observations.

When we regard an original probability distribution as one of the exponential family, we should carefully examine the assumptions in the exponential family. In this article, we show that an appropriate transformation from an original probability distribution of random variable X to a natural exponential family is restricted by a structure corresponding to the expectation $\mu = E(X)$ with respect to the original distribution at most. Thus an extended structure corresponding up to the sufficient statistics with respect to the original distribution

2010 *Mathematics Subject Classification.* 62B10.

Key words and phrases. Exponential family, Sufficient statistics, Misuse, Linearly independent, Irreducible, Transformed regular/pseudo- exponential family.

implies that the transformed exponential family contradicts the assumption of the natural exponential family in the viewpoint of the probability measure of the sufficient statistics.

We define a transformed regular/pseudo- exponential family with respect to the transformation from a continuous probability distribution and examine whether the sufficient statistics in the transformed exponential family is distributed with the k-dimensional exponential family where k is the dimension of the sufficient statistics. A vague application to the sufficient statistics implies a misuse in the exponential family.

2 A misuse in the sufficient statistics Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and a random vector $\boldsymbol{X}$ on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ induces a probability space $(\mathbf{R}^m, \mathcal{B}^m, \mu)$ by $\mu(A) = \mathcal{P}(\boldsymbol{X}^{-1}(\boldsymbol{A})) = \mathcal{P}\{\boldsymbol{X} \in \boldsymbol{A}\}$ for any $\boldsymbol{A} \in \mathcal{B}^m$. Symbolically we may write it $\mu = \mathcal{P} \circ \boldsymbol{X}^{-1}$ and this μ is called the "probability distribution measure" of $\boldsymbol{X}$ and we denote the probability space $(\mathbf{R}^m, \mathcal{B}^m, \mu)$ as $(\mathcal{X}, \mathcal{A}, \mu)$ and the probability (density) function of $\boldsymbol{X}$ as $f(\boldsymbol{x}|\boldsymbol{\xi})$ where $\boldsymbol{\xi}$ is an s-dimensional parameter with respect to the probability distribution of $\boldsymbol{X}$, i.e.,

$$f(\boldsymbol{x}|\boldsymbol{\xi}), \quad \boldsymbol{x} \in \mathbf{R}^m,\ \boldsymbol{\xi} \in \mathbf{R}^s, \tag{1}$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)^T$, $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_s)^T$, and the notation ^{T} means the transpose.

Now we consider a transformed expression of $f(\boldsymbol{x}|\boldsymbol{\xi})$ as regarding an exponential family and we denote it as $f(\boldsymbol{y}|\boldsymbol{\eta})$, i.e.,

$$f(\boldsymbol{y}|\boldsymbol{\eta}) \;=\; \exp\left[C(\boldsymbol{x}) + \langle \boldsymbol{\eta}(\boldsymbol{\xi}),\, \boldsymbol{y}(\boldsymbol{x})\rangle - \phi(\boldsymbol{\xi})\right], \tag{2}$$

where the notation $\langle , \rangle$ means the inner product, $\boldsymbol{\eta}(\boldsymbol{\xi}) = (\eta^1(\boldsymbol{\xi}), \ldots, \eta^k(\boldsymbol{\xi}))^T$ is a transformed parameter, $\boldsymbol{y} = \boldsymbol{y}(\boldsymbol{x}) = (y_1(\boldsymbol{x}), \ldots, y_k(\boldsymbol{x}))^T$ is a transformed variable, $\phi(\boldsymbol{\xi})$ is the normalizing term, and $C(\boldsymbol{x})$ is a constant term. We call (2) a transformed exponential family for the original density (1) and we define the following three kind of types in (2):

$$\begin{cases} \text{(Case A)} \quad f(\boldsymbol{y}|\boldsymbol{\eta}) \;=\; \displaystyle\sum_{\boldsymbol{x}:\boldsymbol{y}(\boldsymbol{x})=\boldsymbol{y}} f(\boldsymbol{x}|\boldsymbol{\xi}), \\ \text{(Case B)} \quad f(\boldsymbol{y}|\boldsymbol{\eta})\, d\boldsymbol{y} \;=\; \begin{cases} \text{(B1)} \quad f(\boldsymbol{x}|\boldsymbol{\xi}) \left|\dfrac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}}\right| d\boldsymbol{y} \quad \text{under } \left|\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}}\right| \neq 0, \\ \text{(B2)} \quad f(\boldsymbol{y}(\boldsymbol{x})|\boldsymbol{\eta})\, d\boldsymbol{x} \quad \text{under } \left|\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}}\right| = 0. \end{cases} \end{cases} \tag{3}$$

Definition **2.1** *We define the above three types (3) of the transformed exponential family as follows:*

- *(Case A) a transformed discrete exponential family,*
- *(Case B1) a transformed regular exponential family,*
- *(Case B2) a transformed pseudo-exponential family.*

Remark that Case A is for a discrete random vector $\boldsymbol{X}$ and it does not need one-to-one correspondence between $\boldsymbol{X}$ and $\boldsymbol{Y}$ and Case B1 requires a one-to-one correspondence between $\boldsymbol{X}$ and $\boldsymbol{Y}$ in order to the non-zero Jacobian $\left|\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}}\right| \neq 0$, so that both dimensions are equal, i.e., $k = m$ and the sufficient statistics $\boldsymbol{Y}$ is a k-dimensional random vector distributed with a k-dimensional probability distribution.

Note that, because of $k \neq m$ in Case B2, the sufficient statistics $\boldsymbol{Y}(\boldsymbol{X})$ is a k-dimensional random vector distributed with $\left|\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}}\right| = 0$, so that we should recognize that the sufficient statistics $\boldsymbol{Y}$ in the transformed pseudo-exponential family does not correspond to $\boldsymbol{X}$ one-on-one.

We consider a k-dimensional canonical exponential family as follows:

$$g(\boldsymbol{z}|\boldsymbol{\theta}) = \exp\left[C(\boldsymbol{z}) + \langle \boldsymbol{\theta}, \boldsymbol{z}\rangle - \phi(\boldsymbol{\theta})\right], \quad \boldsymbol{z} \in \mathbf{R}^k, \ \boldsymbol{\theta} \in \mathbf{R}^k. \tag{4}$$

It is obvious that the sufficient statistics is a k-dimensional $\boldsymbol{z}$. Under the framework of information geometry for the density (4), it is well assumed that $k+1$ functions $\{1, z_1, \ldots, z_k\}$ are *linearly independent* in order to make a one-to-one correspondence between $\boldsymbol{\theta}$ and $g(\boldsymbol{z}|\boldsymbol{\theta})$ for arbitrary $\boldsymbol{\theta} \in \mathbf{R}^k$, where 1 is the identity mapping. For example, [12] defined the linear independence for the density (2) as follows:

Definition **2.2** *The functions $\{1, y_1(\boldsymbol{x}), \ldots, y_k(\boldsymbol{x})\}$ in (2) are said to be linearly independent if the following holds: $a_0 + \sum_{j=1}^{k} a_j y_j(\boldsymbol{x}) = 0$ for any $\boldsymbol{x}$ in an open set if and only if $a_0 = \cdots = a_k = 0$.*

Here we consider the following definition for Definition 2.2:

DEFINITION **2.3** *The functions* $\{y_1(\boldsymbol{x}), \ldots, y_k(\boldsymbol{x})\}$ *in (2) are said to be reducible if the following holds: for some* $y_i(\boldsymbol{x})$*, there exists constants* d_i, c_j $(j \neq i) \in \mathbf{R}$ *such that*

$$y_i(\boldsymbol{x}) \;=\; \sum_{j \neq i} c_j\, y_j(\boldsymbol{x}) + d_i\,,$$

where at least one of $\{c_j\}$ *is not zero. If the functions* $\{y_1(\boldsymbol{x}), \ldots, y_k(\boldsymbol{x})\}$ *are not reducible for arbitrary* $y_i(\boldsymbol{x})$*, then we call them irreducible.*

We have a relationship between Definition 2.2 and Definition 2.3 in the following lemma:

LEMMA **2.1** *If the functions* $\{y_1(\boldsymbol{x}), \ldots, y_k(\boldsymbol{x})\}$ *in (2) are irreducible for any* $\boldsymbol{x}$ *in an open set, then they satisfy the linear independence in Definition 2.2.*

Proof: If the functions $\{y_1(\boldsymbol{x}), \ldots, y_k(\boldsymbol{x})\}$ are irreducible, the condition $a_0 + \sum_{j=1}^{k} a_j y_j(\boldsymbol{x}) = 0$ in Definition 2.2 can be regarded as follows:

$$\left\langle \mathbf{1},\; a_0\, \boldsymbol{e}_0 + \sum_{j=1}^{k} \big(a_j y_j(\boldsymbol{x})\big)\, \boldsymbol{e}_j \right\rangle \;=\; 0,$$

where the vector $\mathbf{1} = (1, 1, \ldots, 1)$ and $\boldsymbol{e}_j$ is the $(j+1)$-th unit vector $(j = 0, 1, \ldots, k)$. Suppose that there exists some $a_i y_i(\boldsymbol{x}) \neq 0$. Then $a_i y_i(\boldsymbol{x}) = -a_0 - \sum_{j \neq i} a_j y_j(\boldsymbol{x})$. If $a_i \neq 0$, then

$$y_i(\boldsymbol{x}) \;=\; -\frac{a_0}{a_i} - \sum_{j \neq i} \frac{a_j}{a_i}\, y_j(\boldsymbol{x}),$$

i.e., $y_i(\boldsymbol{x})$ is reducible and this is a contradiction, so that $a_i = 0$ and this also contradicts the assumption $a_i y_i(\boldsymbol{x}) \neq 0$. Thus we have $a_0 = a_1 y_1(\boldsymbol{x}) = \cdots = a_k y_k(\boldsymbol{x}) = 0$ and $\forall y_i(\boldsymbol{x}) \neq 0$ for any $\boldsymbol{x}$ in an open set, so that $a_0 = a_1 = \cdots = a_k = 0$. Therefore the functions $\{1, y_1(\boldsymbol{x}), \ldots, y_k(\boldsymbol{x})\}$ satisfy the linear independence in Definition 2.2. □

Therefore, as a matter of principle, we suppose that $k+1$ functions $\{1, y_1(\boldsymbol{x}), \ldots, y_k(\boldsymbol{x})\}$ in (2) are irreducible. We show two typical examples as follows:

EXAMPLE **2.1** *In the multinomial distribution with* $k+1$ *cells, the probability function is*

$$f(\boldsymbol{x}|\boldsymbol{\xi}) \;=\; \binom{n}{x_1\, x_2\, \cdots\, x_{k+1}} \prod_{i=1}^{k+1} p_i^{x_i}$$

where $\boldsymbol{x} = (x_1, \ldots, x_k)$, $\boldsymbol{\xi} = (p_1, \ldots, p_k)$, $\sum_{i=1}^{k+1} p_i = 1$ $(p_i \geq 0 \ (i = 1, \ldots, k+1))$, *and where* $\forall x_i \geq 0$, $\sum_{i=1}^{k+1} x_i = n$, *and each integer* x_i $(i = 1, \ldots, k)$ *is the frequency in the i-th cell respectively. For* $f(\boldsymbol{x}|\boldsymbol{\xi})$, *the transformed exponential family is*

$$f(\boldsymbol{y}|\boldsymbol{\eta}) \ = \ \exp\left\{C(\boldsymbol{x}) + \sum_{i=1}^{k} \eta^i(\boldsymbol{\xi}) y_i(\boldsymbol{x}) - \phi(\boldsymbol{\xi})\right\}$$

with respect to $\boldsymbol{Y} = \boldsymbol{Y}(\boldsymbol{X})$, *where* $y_i(\boldsymbol{x}) = x_i$, $\eta^i(\boldsymbol{\xi}) = \log(\xi_i/(1 - \sum_{j=1}^{k} \xi_j)$, $(i = 1, \ldots, k)$,

$$\phi(\boldsymbol{\xi}) = -n \log\left(1 - \sum_{j=1}^{k} \xi_j\right), \quad \textit{and} \quad C(\boldsymbol{x}) = \log \binom{n}{x_1\, x_2\, \cdots\, x_{k+1}}.$$

Here the linear independence of $\{1, y_1(\boldsymbol{x}), \ldots, y_k(\boldsymbol{x})\}$ *holds.*

EXAMPLE **2.2** *In the normal distribution* $N(\mu, \sigma^2)$ *whose density is*

$$f(x|\boldsymbol{\xi}) \ = \ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} \quad (x \in \mathbf{R})$$

with $\boldsymbol{\xi} = (\mu, \sigma)$ *for* $-\infty < \mu < \infty$ *and* $0 < \sigma < \infty$, *the transformed exponential family is*

$$f(\boldsymbol{y}|\boldsymbol{\eta}) \ = \ \exp\left\{C(x) + \sum_{i=1}^{2} \eta^i(\xi)\, y_i(x) - \phi(\boldsymbol{\xi})\right\}$$

with respect to $\boldsymbol{Y} = \boldsymbol{Y}(X)$, *where* $\eta^1(\boldsymbol{\xi}) = \mu/\sigma^2$, $\eta^2(\boldsymbol{\xi}) = -1/(2\sigma^2)$, $y_1(x) = x$, $y_2(x) = x^2$,

$$\phi(\boldsymbol{\xi}) = \frac{\mu^2}{2\sigma^2} + \log(\sigma), \quad \textit{and} \quad C(x) = \log\frac{1}{\sqrt{2\pi}}.$$

Here the region $\{(y_1(x), y_2(x)) : x \in \mathbf{R}\}$ *is equivalent to a parabolic curve in the 2-dimensional space and the functions* $\{1, y_1(x), y_2(x)\}$ *are irreducible by Definition 2.3. This example is of (Case B2) in the relationship (3) unless the parameter* σ^2 *is supposed to be known.*

Since it is well known that the exponential family includes a lot of probability distributions, we are apt to confuse the sufficient statistics in the exponential family with the sufficient statistics distributed with the exponential family. The following theorem shows a solution to the above confusion.

THEOREM **2.1** *Let $\boldsymbol{X}$ be an m-dimensional random vector distributed with a probability density function (1) and, for a transformed exponential family (2) of $\boldsymbol{X}$, let the k-dimensional random vector $\boldsymbol{Y}(\boldsymbol{X}) = (\boldsymbol{Y}^{(1)}(\boldsymbol{X}), \boldsymbol{Y}^{(2)}(\boldsymbol{X}))$ which is the sufficient statistics in (2). Suppose that these $k+1$ functions $\{1, \boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)}\}$ are irreducible. There exists three cases as follows: (Case 1) $k < m$, (Case 2) $k = m$, (Case 3) $k > m$. In Case 1, since we have a loss of information based on $\boldsymbol{X}$, it contradicts that $\boldsymbol{Y}$ is the sufficient statistics. In Case 2, since $\boldsymbol{X}$ and $\boldsymbol{Y}$ have a one-to-one correspondence, this is a transformed regular exponential family. In Case 3, since $\boldsymbol{Y}^{(1)}(\boldsymbol{X})$ is regarded as corresponding to $\boldsymbol{X}$ under $k > m$ and the dimension of $\boldsymbol{Y}^{(2)}(\boldsymbol{X})$ is $k - m$, there exists a measurable and irreducible function u such that $\boldsymbol{y}^{(2)} = u\left(\boldsymbol{y}^{(1)}\right)$ and the conditional density of $\boldsymbol{y}^{(2)}$ given $\boldsymbol{y}^{(1)}$ is the indicator function, i.e., this is a transformed pseudo-exponential family.*

Proof: Both Case 1 and Case 2 are obvious, so we prove Case 3 only.

Since the joint density $h(\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)})$ of $\boldsymbol{Y}$ is equivalent to the transformed exponential family (2), i.e.,

$$h(\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)}) \;=\; f(\boldsymbol{y}|\boldsymbol{\eta}) \;=\; \exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(1)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(1)}\rangle + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle - \phi(\boldsymbol{\xi})\},$$

where $\boldsymbol{\eta}(\boldsymbol{\xi}) = (\boldsymbol{\eta}^{(1)}(\boldsymbol{\xi}),\, \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}))$ and the dimension of $\boldsymbol{\eta}^{(1)}(\boldsymbol{\xi})$ is m, we have the following representation:

$$f(\boldsymbol{y}|\boldsymbol{\eta}) \;=\; f(\boldsymbol{y}^{(2)} \,|\, \boldsymbol{y}^{(1)})\, f(\boldsymbol{y}^{(1)}),$$

where the marginal of $\boldsymbol{Y}^{(1)}$ and the conditional density of $\boldsymbol{Y}^{(2)}$ given $\boldsymbol{Y}^{(1)}$ are

$$\begin{aligned} f(\boldsymbol{y}^{(1)}) &= \exp\{\langle \boldsymbol{\eta}^{(1)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(1)}\rangle - \phi(\boldsymbol{\xi})\} \int \exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle\} d\boldsymbol{y}^{(2)}, \\ f(\boldsymbol{y}^{(2)} \,|\, \boldsymbol{y}^{(1)}) &= \frac{\exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle\}}{\int \exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle\}\, d\boldsymbol{y}^{(2)}} \end{aligned} \tag{5}$$

under the assumption that $0 < \int \exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle\} d\boldsymbol{y}^{(2)} < \infty$. Since these $k+1$ functions $\{1, \boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)}\}$ are irreducible, $\boldsymbol{Y}^{(1)} = \boldsymbol{Y}^{(1)}(\boldsymbol{X})$ corresponds $\boldsymbol{X}$ one-on-one, and $\boldsymbol{Y}^{(2)} = \boldsymbol{Y}^{(2)}(\boldsymbol{X})$ is a measurable function of $\boldsymbol{X}$, any element of $\boldsymbol{Y}^{(2)}$ is not of a linear combination of $\boldsymbol{Y}^{(1)}$, so that there exists a measurable function u such that $\boldsymbol{Y}^{(2)} = u(\boldsymbol{Y}^{(1)})$ and $\{1, \boldsymbol{y}^{(1)}, u(\boldsymbol{y}^{(1)})\}$ are irreducible.

For the joint density function $h(\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)})$ with the relationship $\boldsymbol{y}^{(2)} = u(\boldsymbol{y}^{(1)})$, it holds that $h(\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)}) = \delta_{u(y^{(1)})}(\boldsymbol{y}^{(2)})\, h(\boldsymbol{y}^{(1)})$, where $h(\boldsymbol{y}^{(1)})$ is the marginal function of $\boldsymbol{Y}^{(1)}$, the function $\delta_{u(y^{(1)})}(\boldsymbol{y}^{(2)})$ is the Dirac's delta function at the point $u(\boldsymbol{y}^{(1)})$, that is, the conditional density function of $\boldsymbol{Y}^{(2)}$ given $\boldsymbol{y}^{(1)}$ is the delta function $\delta_{u(y^{(1)})}(\boldsymbol{Y}^{(2)})$:

$$h(\boldsymbol{y}^{(2)} \,|\, \boldsymbol{y}^{(1)}) \;=\; \frac{h(\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)})}{h(\boldsymbol{y}^{(1)})} \;=\; \frac{\delta_{u(y^{(1)})}(\boldsymbol{y}^{(2)})\, h(\boldsymbol{y}^{(1)})}{h(\boldsymbol{y}^{(1)})} \;=\; \delta_{u(y^{(1)})}(\boldsymbol{y}^{(2)}), \tag{6}$$

so that, since the conditional probability function (5) is equivalent to the conditional (6), we have the following relationship:

$$\begin{aligned} \delta_{u(y^{(1)})}(\boldsymbol{y}^{(2)}) \;&=\; \frac{\exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle\}}{\int \exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle\} d\boldsymbol{y}^{(2)}} \\ &=\; \begin{cases} 1, & \text{if } \boldsymbol{Y}^{(2)} = u(\boldsymbol{y}^{(1)}), \\ 0, & \text{if } \boldsymbol{Y}^{(2)} \neq u(\boldsymbol{y}^{(1)}). \end{cases} \end{aligned} \tag{7}$$

If $\boldsymbol{Y}^{(2)} = u(\boldsymbol{y}^{(1)})$, then the numerator is equivalent to the denominator in (7), i.e.,

$$\exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle\} \;=\; \int \exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle\} d\boldsymbol{y}^{(2)},$$

which implies that the probability distribution of $\boldsymbol{Y}^{(2)}$ should be one point distribution because the based random variable $\boldsymbol{X}$ is a continuous distribution. If $\boldsymbol{Y}^{(2)} \neq u(\boldsymbol{y}^{(1)})$, then the numerator in (7) should be zero, i.e.,

$$\exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(2)}\rangle\} \;=\; 0,$$

and this is impossible, but we need not directly consider the conditional density of $\boldsymbol{Y}^{(2)}$ given $\boldsymbol{Y}^{(1)}$ in this situation and the transformed density is zero, i.e., $f(\boldsymbol{y}|\boldsymbol{\eta}) = 0$.

Therefore, for the sufficient statistics $\boldsymbol{Y} = (\boldsymbol{Y}^{(1)}, \boldsymbol{Y}^{(2)})$, the transformed density (2) is represented by

$$f(\boldsymbol{y}|\boldsymbol{\eta}) \;=\; \begin{cases} f(\boldsymbol{y}^{(1)}), & \text{if } \boldsymbol{Y}^{(2)} = u(\boldsymbol{y}^{(1)}), \\ 0, & \text{if } \boldsymbol{Y}^{(2)} \neq u(\boldsymbol{y}^{(1)}), \end{cases} \tag{8}$$

where $f(\boldsymbol{y}^{(1)}) = \exp\{C(\boldsymbol{x}) + \langle \boldsymbol{\eta}^{(1)}(\boldsymbol{\xi}),\, \boldsymbol{y}^{(1)}\rangle + \langle \boldsymbol{\eta}^{(2)}(\boldsymbol{\xi}),\, u(\boldsymbol{y}^{(1)})\rangle - \phi(\boldsymbol{\xi})\}$, so that we can regard the second element $\boldsymbol{Y}^{(2)}$ as either a random variable with one point distribution or a non-random (deterministic) variable given $u(\boldsymbol{y}^{(1)})$. In the equation (8), the left-hand side is the density of k-dimensional random variable and the right-hand side is that of m-dimensional random variable ($k > m$), which implies that the k-dimensional sufficient statistics $\boldsymbol{Y}$ is distributed with a transformed pseudo-exponential family in (3). □

Although the conditional expectation and variance of $\boldsymbol{Y}^{(2)}$ given $\boldsymbol{Y}^{(1)}$ in Case 3 are

$$E[\boldsymbol{Y}^{(2)} \,|\, \boldsymbol{Y}^{(1)}] \;=\; u(\boldsymbol{Y}^{(1)}) \quad \text{and} \quad V[\boldsymbol{Y}^{(2)} \,|\, \boldsymbol{Y}^{(1)}] \;=\; \mathbf{0},$$

the expectation and variance of $\boldsymbol{Y}^{(2)}$ are

$$\begin{aligned} E[\boldsymbol{Y}^{(2)}] &= E\left[E[\boldsymbol{Y}^{(2)} \,|\, \boldsymbol{Y}^{(1)}]\right] \;=\; E[u(\boldsymbol{Y}^{(1)})], \\ V[\boldsymbol{Y}^{(2)}] &= V\left[E[\boldsymbol{Y}^{(2)} \,|\, \boldsymbol{Y}^{(1)}]\right] + E\left[V[\boldsymbol{Y}^{(2)} \,|\, \boldsymbol{Y}^{(1)}]\right] \;=\; V[u(\boldsymbol{Y}^{(1)})], \end{aligned}$$

and the covariance between $\boldsymbol{Y}^{(1)}$ and $\boldsymbol{Y}^{(2)}$ is $Cov[\boldsymbol{Y}^{(1)}, \boldsymbol{Y}^{(2)}] = Cov[\boldsymbol{Y}^{(1)}, u(\boldsymbol{Y}^{(1)})]$. On the other hand, based on the structure of exponential family, we have the following relationships with respect to the sufficient statistics $\boldsymbol{Y}$:

$$\begin{aligned} E[\boldsymbol{Y}] &= \frac{\partial \phi(\boldsymbol{\xi})}{\partial \boldsymbol{\eta}(\boldsymbol{\xi})} = \begin{pmatrix} E[\boldsymbol{Y}^{(1)}] \\ E[\boldsymbol{Y}^{(2)}] \end{pmatrix}, \\ V[\boldsymbol{Y}] &= \frac{\partial^2 \phi(\boldsymbol{\xi})}{\partial \boldsymbol{\eta}(\boldsymbol{\xi})\, \partial \boldsymbol{\eta}(\boldsymbol{\xi})^T} = \begin{pmatrix} V[\boldsymbol{Y}^{(1)}] & Cov[\boldsymbol{Y}^{(1)}, \boldsymbol{Y}^{(2)}] \\ Cov[\boldsymbol{Y}^{(1)}, \boldsymbol{Y}^{(2)}]^T & V[\boldsymbol{Y}^{(2)}] \end{pmatrix}, \end{aligned}$$

so that the transformed exponential family (2) with respect to k-dimensional $\boldsymbol{Y} = (\boldsymbol{Y}^{(1)}, \boldsymbol{Y}^{(2)})$ has the same properties with respect to the usual exponential family (4) with respect to $\boldsymbol{Z} = (Z_1, \ldots, Z_k)$ on the surface, but the transformed pseudo-exponential family in (3) is not like the k-dimensional exponential density (4) because the pseudo-density is reduced to the density (8) with respect to only $\boldsymbol{Y}^{(1)}$. Thus, even if the sufficient statistics $\boldsymbol{Y}(\boldsymbol{X})$ in the transformed exponential family (2) is irreducible, it might belong to the k-dimensional transformed pseudo-exponential family.

For the m-dimensional normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the density is represented as follows:

$$f(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \;=\; \exp\left\{ \langle \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}, \boldsymbol{x} \rangle - \frac{1}{2} \langle \boldsymbol{\Sigma}^{-1}\boldsymbol{x}, \boldsymbol{x} \rangle - \frac{\langle \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}, \boldsymbol{\mu} \rangle + \log(|\boldsymbol{\Sigma}|)}{2} \right\} \frac{1}{(2\pi)^{m/2}},$$

so that the sufficient statistics $\boldsymbol{Y}$ in the transformed exponential family under unknown parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ is

$$\boldsymbol{Y} \;=\; (X_1, \ldots, X_m,\, X_1^2, \ldots, X_m^2,\, X_1X_2, \ldots, X_{m-1}X_m)^T$$

whose dimension is $2m + (m^2 - m)/2$. Note that [9] studied the circular mechanism as a limitation to the transformed exponential family.

3 Conclusion In this article, we considered a misuse of the sufficient statistics in the transformed exponential family from a continuous probability distribution based on the linear independence of the sufficient statistics which are assumed in the information geometry. We defined new two terms, the transformed regular exponential family and the transformed pseudo-exponential family and we determined properties of the sufficient statistics under the irreducibility of the transformed exponential family. We recognized an importance of the Jacobian matrix with respect to the transformation of random variables.

We hope that it is decreasing to misuse that the sufficient statistics in a transformed pseudo-exponential family ought to be distributed with the k-dimensional regular exponential family where k is the dimension of the sufficient statistics.

References

[1] S.Amari (1985), *Differential-Geometrical Methods in Statistics*, Springer.

[2] S.Amari and H.Nagaoka (2000), *Methods of Information Geometry*, American Mathematical Society.

[3] O. Barndorff-Nielsen (1978), *Information and Exponential Families in Statistical Theory*, John Wiley & Sons

[4] D.D.Boos and L.A.Stefanski (2013), *Essential Statistical Inference Theory and Methods*, Springer.

[5] L.D.Brown (1986), *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*, Institute of Mathematical Statistics.

[6] B.Efron (1975), Defining the Curvature of a Statistical Problem (with Applications to Second Order Efficiency), *Ann. Stat.*, **3**, 1189–1242.

[7] J.Jiang (2007), *Linear and Generalized Linear Mixed Models and Their Applications*, Springer.

[8] R.E.Kass and P.W.Voss (1997), *Geometrical Foundations of Asymptotic Inference*, John Wiley & Sons, Inc.

[9] E.Kumagai and N.Inagaki (1999), The circular mechanism in the curved exponential family, *Mathematicae Japonicae*, **49**(1), 27–38.

[10] E.Kumagai (2015), A comment on a regularity condition in a curved exponential family, *Scientiae Mathematicae Japonicae*, **78**(1), 97–101.

[11] M.K.Murray and J.W.Rice (1993), *Differential Geometry and Statistics*, Chapman & Hall.

[12] T.Ohnishi and T.Yanagimoto (2010), Duality induced from conjugacy in the curved exponential family, *Journal of Japan Statistical Society*, **40**(1), 23–43.

[13] G.Pistone and M.P.Rogantin (1999), The exponential statistical manifold: mean parameters, orthogonality and space transformations, *Bernoulli*, **5**(4), 721–760.

[14] M.J.Schervish (1995), *Theory of Statistics*, Springer.

[15] J.Shao (2003), *Mathematical Statistics 2nd ed.*, Springer.

Communicated by *Hisao Nagao*

Division of Mathematical Science,

Graduate School of Engineering Science,

Osaka University,

1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan.

AN INVESTIGATION OF A GENERALIZED LEAST SQUARES ESTIMATOR FOR NON-LINEAR TIME SERIES MODEL

Xiaoling Dou

Received March 11, 2017 ; revised April 7, 2017

Abstract.

Ochi(1983) proposed an estimator for the autoregressive coefficient of the first-order autoregressive model (AR(1)) by using two constants for the end points of the process. Classical estimators for AR(1) , such as the least squares estimator, Burg's estimator, and Yule-Walker estimator are obtained as special cases by choice of the constants in Ochi's estimator. By writing the first-order autoregressive conditional heteroskedastic model, ARCH(1), in a form similar to that of AR(1), we extend Ochi ' s estimator to ARCH(1) models. This allows introducing analogues of the least squares estimator, Burg's estimator and Yule-Walker estimator, and we compare the relations of these with Ochi's estimator for ARCH(1) models. We then provide a simulation for AR(1) models and examine the performance of Ochi ' s estimator. Also, we simulate Ochi's estimator for ARCH(1) with different parameter values and sample sizes.

1 Introduction

Let $\{x_1, \cdots, x_T\}$ be generated from the first order autoregressive process, AR(1),

$$x_t = \alpha x_{t-1} + \epsilon_t, \quad |\alpha| < 1, \quad \epsilon_t \sim N(0, \sigma^2), \quad t \in [2, \cdots, T] \tag{1}$$

with an unknown coefficient α, and independent and identically distributed (iid) errors ϵ_t. Ochi (1983) proposed an estimator of the autoregressive coefficient

$$\text{Ochi}(c_1, c_2) = \hat{\alpha}_{c_1,c_2} = \frac{\sum_{t=2}^{T} x_t x_{t-1}}{\sum_{t=2}^{T-1} x_t^2 + c_1 x_1^2 + c_2 x_T^2}, \tag{2}$$

where c_1 and c_2 are nonnegative constants, and can be considered as weights of the end points. It is also known that Ochi$(1, 0)$, Ochi$(0.5, 0.5)$, and Ochi$(1, 1)$ are the least squares estimator(LSE), Burg's estimator, and Yule-Walker estimator, respectively.

Recently, non-linear time series models have been increasing in popularity. Autoregressive conditional heteroskedastic (ARCH) models were proposed by Engle (1982). Chan and Tong (1986) and Tong (1990) introduced some threshold models, such as threshold

2010 *Mathematics Subject Classification.* subject classifications .

Key words and phrases. Burg's method, Least squares estimator, Mean Square Error, Ochi's estimator, Yule-Walker estimator.

autoregressive (TAR) models, self-exciting threshold (SETAR) models and smooth threshold autoregressive(STAR) models. Markov switching autoregressive(MAR) models were developed by Hamilton (1989). Davis et al. (2008) introduced some segmented time series. For more non-linear models, see Turkman et al. (2014). The ARCH process of order 1, ARCH(1) is one of the most famous and can be modeled as

$$y_t = \sqrt{\theta_0 + \theta_1 y_{t-1}^2}\, u_t, \quad u_t \sim \text{iid } N(0,1), \tag{3}$$

with parameters $\theta_0 > 0$ and $|\theta_1| < 1$. Rewriting (3), we have

$$\begin{aligned} y_t^2 &= (\theta_0 + \theta_1 y_{t-1}^2) u_t^2 \\ &= \theta_0 + \theta_1 y_{t-1}^2 + (\theta_0 + \theta_1 y_{t-1}^2)(u_t^2 - 1) \\ &= \theta_0 + \theta_1 y_{t-1}^2 + \xi_t, \end{aligned} \tag{4}$$

which has a form similar to that of AR(1) in (1). Here, $\xi_t := (\theta_0 + \theta_1 y_{t-1}^2)(u_t^2 - 1)$ is an uncorrelated process with mean 0 and variance

$$\begin{aligned} \operatorname{Var}(\xi_t) &= E[(\theta_0 + \theta_1 y_{t-1}^2)^2 (u_t^2 - 1)^2] - \{E[(\theta_0 + \theta_1 y_{t-1}^2)(u_t^2 - 1)]\}^2 \\ &= E[(\theta_0 + \theta_1 y_{t-1}^2)^2] E[(u_t^2 - 1)^2] - \{\theta_0 + \theta_1 E[y_{t-1}^2]\}^2 \{E[u_t^2] - 1\}^2 \\ &= E[\theta_0^2 + \theta_1^2 y_{t-1}^4 + 2\theta_0\theta_1 y_{t-1}^2] E[u_t^4 + 1 - 2u_t^2] - 0 \\ &= 2\left(\theta_0^2 + \theta_1^2 E[y_{t-1}^4] + 2\theta_0\theta_1 E[y_{t-1}^2]\right). \end{aligned}$$

Since

$$E[y_t^2] = \theta_0 + \theta_1 E[y_{t-1}^2] + 0 = \frac{\theta_0}{1-\theta_1}, \quad E[y_t^2] = E[y_{t-1}^2],$$

$E[y_t^4] = E[y_{t-1}^4]$, $u_t^2 \sim \chi_1^2$, and $E[u_t^4] = 3$, we have

$$E[y_t^4] = E[(\theta_0 + \theta_1 y_{t-1}^2)^2]\, E[u_t^4] = 3(\theta_0^2 + \theta_1^2\, E[y_{t-1}^4] + 2\theta_0\theta_1 E[y_{t-1}^2]) = \frac{3\theta_0^2(1+\theta_1)}{(1-\theta_1)(1-3\theta_1^2)}.$$

Hence the expression for the variance of ξ_t can be simplified to

$$\operatorname{Var}(\xi_t) = \frac{2\theta_0^2(1+\theta_1)}{(1-3\theta_1^2)(1-\theta_1)}, \tag{5}$$

and the variance of y_t^2 can be easily found

$$\operatorname{Var}(y_t^2) = E[y_t^4] - (E[y_t^2])^2 = \frac{2\theta_0^2}{(1-3\theta_1^2)(1-\theta_1)^2}. \tag{6}$$

From variances (5) and (6), we see that $\theta_1 < \sqrt{1/3}$ is required. This is also discussed in Shumway and Stoffer (2011).

Suppose the process $\{y_1, \cdots, y_T\}$ is generated from (3). To estimate the parameters $\boldsymbol{\theta} = (\theta_0, \theta_1)'$ in the ARCH(1) process, we apply Ochi's estimator to the squared process (4). The main purpose of this paper is to investigate the performance of Ochi's estimator for (4)

$$\text{(7)} \qquad \text{Ochi}^*(c_1, c_2) = \hat{\boldsymbol{\theta}}_{c_1,c_2} = \left(\sum_{t=2}^{T-1} \boldsymbol{y}_t \boldsymbol{y}_t' + c_1 \boldsymbol{y}_1 \boldsymbol{y}_1' + c_2 \boldsymbol{y}_T \boldsymbol{y}_T' \right)^{-1} \sum_{t=1}^{T-1} \boldsymbol{y}_t y_{1+t}^2,$$

by simulation. In this, $c_1,\ c_2 \geq 0$, and

$$\boldsymbol{y}_t = \begin{pmatrix} 1 \\ y_t^2 \end{pmatrix}.$$

To compare Ochi's estimator (7) with the LSE, Burg's estimator, and Yule-Walker estimator, we give the derivations of the three estimators in the ARCH(1) case.

The paper is organized as follows. In Section 2, we extend the LSE, Burg's estimator and Yule-Walker estimator to the ARCH(1) model. In Section 3, we evaluate and compare Ochi's estimator and the three estimators in AR(1) and ARCH(1) models by simulation. Finally, in Section 4, we discuss these results and conclude.

2 LSE, Burg's estimator, and Yule-Walker estimator for ARCH(1) model

2.1 The least squares estimator By minimizing the sum of squared errors

$$\sum_{t=2}^{T} \{y_t^2 - (\theta_0 + \theta_1 y_{t-1}^2)\}^2,$$

we can obtain the LSE (e.g., Taniguchi et al., 2008)

$$\hat{\boldsymbol{\theta}}_{LSE} = \begin{pmatrix} \hat{\theta}_0 \\ \hat{\theta}_1 \end{pmatrix} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{Y},$$

where

$$\text{(8)} \qquad \boldsymbol{Y} = (y_2^2, \cdots, y_T^2)', \quad \boldsymbol{Z} = \begin{pmatrix} 1 & y_1^2 \\ \vdots & \vdots \\ 1 & y_{t-1}^2 \\ \vdots & \vdots \\ 1 & y_{T-1}^2 \end{pmatrix},$$

and

$$\boldsymbol{Z}' = \left(\begin{pmatrix} 1 \\ y_1^2 \end{pmatrix}, \begin{pmatrix} 1 \\ y_2^2 \end{pmatrix}, \cdots, \begin{pmatrix} 1 \\ y_{T-1}^2 \end{pmatrix} \right) = (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{T-1}).$$

Then the LSE can be rewritten as

$$\hat{\boldsymbol{\theta}}_{LSE} = \left(\sum_{t=1}^{T-1} \boldsymbol{y}_t \boldsymbol{y}_t' \right)^{-1} \sum_{t=1}^{T-1} \boldsymbol{y}_t y_{t+1}^2. \tag{9}$$

Recalling the form of Ochi's estimator (7) with constants c_1 and c_2, we see that when $c_1 = 1$ and $c_2 = 0$, Ochi's estimator becomes the LSE.

2.2 Burg's method Burg's idea (Burg, 1975) is simple. With a previous given value y_{t-1} and a next given value y_{t+1}, forward and backward linear prediction can be represented as

$$\hat{y}_t^2 = \hat{\theta}_0 + \hat{\theta}_1 y_{t-1}^2, \quad t \in \{2, 3, \cdots, T\} \tag{10}$$

and

$$\tilde{y}_t^2 = \hat{\theta}_0 + \hat{\theta}_1 y_{t+1}^2, \quad t \in \{1, 2, \cdots, T-1\}, \tag{11}$$

respectively. The sum of the squared errors for (10) is

$$S_f = \sum_{t=2}^{T} (y_t^2 - \theta_0 - \theta_1 y_{t-1}^2)^2 \tag{12}$$

and for (11) is

$$S_b = \sum_{t=1}^{T-1} (y_t^2 - \theta_0 - \theta_1 y_{t+1}^2)^2. \tag{13}$$

Minimizing the sum of (12) and (13),

$$\begin{aligned} S = S_f + S_b &= \sum_{t=2}^{T} (y_t^2 - \boldsymbol{\theta}' \boldsymbol{y}_{t-1})^2 + \sum_{t=1}^{T-1} (y_t^2 - \boldsymbol{\theta}' \boldsymbol{y}_{t+1})^2 \\ &= y_1^4 + 2 \sum_{t=2}^{T-1} y_t^4 + y_T^4 + (\boldsymbol{\theta}' \boldsymbol{y}_1)^2 + 2 \sum_{t=2}^{T-1} (\boldsymbol{\theta}' \boldsymbol{y}_t)^2 + (\boldsymbol{\theta}' \boldsymbol{y}_T)^2 - 2\boldsymbol{\theta}' \left(\sum_{t=2}^{T} y_t^2 \boldsymbol{y}_{t-1} + \sum_{t=1}^{T-1} y_t^2 \boldsymbol{y}_{t+1} \right), \end{aligned}$$

by setting the gradient with respect to $\boldsymbol{\theta}$ as $\mathbf{0}$

$$\frac{\partial S}{\partial \boldsymbol{\theta}} = 2\boldsymbol{y}_1 (\boldsymbol{\theta}' \boldsymbol{y}_1)' + 4 \sum_{t=2}^{T-1} \boldsymbol{y}_t (\boldsymbol{\theta}' \boldsymbol{y}_t)' + 2\boldsymbol{y}_T (\boldsymbol{\theta}' \boldsymbol{y}_T)' - 2 \left(\sum_{t=2}^{T} y_t^2 \boldsymbol{y}_{t-1} + \sum_{t=1}^{T-1} y_t^2 \boldsymbol{y}_{t+1} \right) = \mathbf{0},$$

we have

$$\hat{\boldsymbol{\theta}}_{Burg} = \left(\boldsymbol{y}_1 \boldsymbol{y}_1' + 2 \sum_{t=2}^{T-1} \boldsymbol{y}_t \boldsymbol{y}_t' + \boldsymbol{y}_T \boldsymbol{y}_T' \right)^{-1} \left(\sum_{t=2}^{T} y_t^2 \boldsymbol{y}_{t-1} + \sum_{t=1}^{T-1} y_t^2 \boldsymbol{y}_{t+1} \right)$$

$$= \left(\boldsymbol{y}_1 \boldsymbol{y}_1' + 2 \sum_{t=2}^{T-1} \boldsymbol{y}_t \boldsymbol{y}_t' + \boldsymbol{y}_T \boldsymbol{y}_T' \right)^{-1} \begin{pmatrix} y_1^2 + 2\sum_{t=2}^{T-1} y_t^2 + y_T^2 \\ 2\sum_{t=1}^{T-1} y_t^2 y_{t+1}^2 \end{pmatrix}.$$

We see that Burg's estimator gives $\hat{\theta}_0$ a different form from Ochi's estimator. However, θ_1 is estimated as a special case of Ochi's estimator (7) by setting $c_1 = c_2 = 0.5$.

2.3 Yule-Walker estimator The Yule-Walker method is derived by considering the following expectations

$$\begin{aligned} E[y_t^2 y_t^2] &= \theta_0 E[y_t^2] + \theta_1 E[y_t^2 y_{t-1}^2] + E[\xi_t y_t^2] \\ &= \theta_0 \mu + \theta_1 E[y_t^2 y_{t-1}^2] + V[\xi_t], \end{aligned} \tag{14}$$

$$\begin{aligned} E[y_t^2 y_{t-1}^2] &= \theta_0 E[y_{t-1}^2] + \theta_1 E[y_{t-1}^2 y_{t-1}^2] + E[\xi_t y_{t-1}^2] \\ &= \theta_0 \mu + \theta_1 E[y_{t-1}^2 y_{t-1}^2] \\ &= \theta_0 \mu + \theta_1 E[y_t^2 y_t^2], \end{aligned} \tag{15}$$

where

$$\mu = E[y_t^2] = \theta_0 + \theta_1 E[y_{t-1}^2] + E[\xi_t] = \theta_0 + \theta_1 \mu. \tag{16}$$

Hence,

$$\mu = \frac{\theta_0}{1 - \theta_1}, \quad \text{and} \quad \theta_0 = \mu(1 - \theta_1). \tag{17}$$

From (16) and(15), we can see that

$$\theta_1 = \frac{E[y_t^2 y_{t-1}^2] - \mu^2}{E[y_t^2 y_t^2] - \mu^2} = \frac{C_1}{C_0}.$$

That is, θ_1 can be estimated by the lag 1 autocorrelation function of the series. Then the Yule-Walker estimator of the parameters $\boldsymbol{\theta} = (\theta_0, \theta_1)'$ can be obtained by

$$\hat{\boldsymbol{\theta}}_{YW} = \begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T} y_t^2 \left(1 - \frac{\hat{C}_1}{\hat{C}_0}\right) \\ \hat{C}_1/\hat{C}_0 \end{pmatrix}, \tag{18}$$

where $\hat{C}_0$ and $\hat{C}_1$ are the sample autocovariance functions of y_t^2 for lags 0 and 1, respectively. Comparing this with Ochi's estimator (7), we can see that the Yule-Walker estimator of θ_1 is a centered version of Ochi*(1,1).

3 Simulation and results In this section we provide a simulation study for Ochi's estimator, the LSE, Burg's estimator and Yule-Walker estimator for AR(1) models in (1) and ARCH(1) models in the form (3). The models and parameters used in the simulations are given in Table 1.

Table 1: Simulation setting

AR(1)	ARCH(1)
$x_t = \alpha x_{t-1} + \epsilon_t, \qquad t \in \{2, \cdots, T\}$	$y_t^2 = (\theta_0 + \theta_1 y_{t-1}^2) u_t^2, \qquad t \in \{2, \cdots, T\}$
	$\theta_0 = 1$
$\alpha = (0, 0.05, 0.1, \cdots, 0.95)'$	$\theta_1 \in \{0, 0.05, 0.1, \cdots, 0.5, 0.55 < \sqrt{1/3}\}$
$\epsilon_t \quad \sim$ iid $N(0, 1)$	$u_t \quad \sim$ iid $N(0, 1)$
$\{x_1, x_2, \cdots, x_T\}$	$\{y_1^2,\ y_2^2, \ldots, y_T^2\}$
$T \in \{100,\ 200,\ 300,\ 500,\ 1000\}$	
Replications of time series sequences $N = 1000$	
$c_1,\ c_2 \in \{0, 0.2, 0.5, 0.7, 1\}$	

3.1 Simulation of Ochi's estimator for AR(1) For model (1) with $\sigma^2 = 1$ and α as each of $(0, 0.05, 0.1, \cdots, 0.95)'$, we generate $N = 1000$ sequences as time series with length 1000. That is, 20×1000 sequences of length 1000 are generated. For each sequence, we consider five different sample sizes T,

$$x_1,\ x_2, \cdots, x_T, \quad T \in \{100,\ 200,\ 300,\ 500,\ 1000\}$$

and estimate the parameter α by Ochi's estimator (2) for each T. In the calculation, we try all 25 pairs of $(c_1, c_2),\ c_1,\ c_2 \in \{0, 0.2, 0.5, 0.7, 1\}$ for the constants of Ochi's estimator. Then, we evaluate the performance of the different pairs (c_1, c_2) by comparing the resulting mean square errors (MSEs) for different sample sizes as α changes.

Figure 1 shows a part of the simulation results for $T = 100$ and $T = 200$. As expected, a bigger sample size T gives a smaller MSE, and hence provides a better estimate of α. For each sample size, MSE is calculated for all 25 pairs of (c_1, c_2). The MSE curves are plotted as dashed lines. MSE curves for four special pairs are plotted as solid lines, in red for $(0, 0)$, green for $(0.5, 0.5)$, blue for $(1, 0)$, and light blue for $(1, 1)$. We see that as α grows, the MSE curves decrease.This means that better estimation is obtained when α is bigger. At around $\alpha = 0.5$, the MSEs for different pairs of (c_1, c_2) do not show big differences. When $\alpha < 0.4$, the two extreme cases, $(c_1, c_2) = (0, 0)$ and $(c_1, c_2) = (1, 1)$, give the biggest and smallest MSEs, respectively. However, when $\alpha > 0.6$, these two curves exchange their positions. In contrast, Burg's method(Ochi(0.5, 0.5)) and the LSE (Ochi(1,0)) give intermediate MSEs. In particular, when α is close to 1, Burg's method is slightly better than the LSE. When the sample size is large, different pairs of (c_1, c_2) give only small differences in MSE.

The figures for the variance and squared bias show that as α increases the variance becomes smaller but the squared bias becomes bigger. In particular, Ochi$(1, 1)$ (the Yule-Walker estimator) shows the largest squared bias among these methods. We can also see from the last panel of the figure that the mean of the estimated α is slightly less than the real α (gray line).

3.2 Simulation of Ochi's estimator for ARCH(1) We set the true parameters in model (4) as $\theta_0 = 1$ and $\theta_1 \in \{0, 0.05, 0.1, \cdots, 0.5, 0.55 < \sqrt{1/3}\}$, and then we use Ochi's estimator (7) to estimate the parameters by simulation with different constants $c_1,\ c_2 \in \{0, 0.2, 0.5, 0.7, 1\}$. For every value of θ_1, $N = 1000$ sequences of length 1000 are generated. For each sequence, we consider different values of T;

$$y_1^2,\ y_2^2, \ldots, y_T^2, \qquad T \in \{100, 200, 300, 500, 1000\}.$$

We estimate θ_0 and θ_1 and then compute the MSEs, variances, and squared biases of the estimates for each case.

Figure 2 shows the MSE, variance, and squared biase for Ochi's estimator for estimating θ_0 and θ_1. Different colors indicate different lengths of the time series (or sample sizes), with these sizes $T \in \{100, 200, 300, 500, 1000\}$. In each panel, for each T, 25 curves obtained from different pairs of (c_1, c_2), $c_1,\ c_2 \in \{0, 0.2, 0.5, 0.7, 1\}$, are plotted with respect to θ_1.

The first panel shows MSE curves obtained in estimating θ_0. The MSE curve obtained from $(c_1 = 1, c_2 = 1)$ is plotted as a solid line, and the other 24 MSE curves are plotted as dashed lines. The graph shows that as the sample size T increases, the corresponding MSE becomes smaller. For big sample sizes, such as $T = 1000$, the choice of (c_1, c_2) makes almost no difference. In contrast, with smaller sample sizes, there are bigger differences

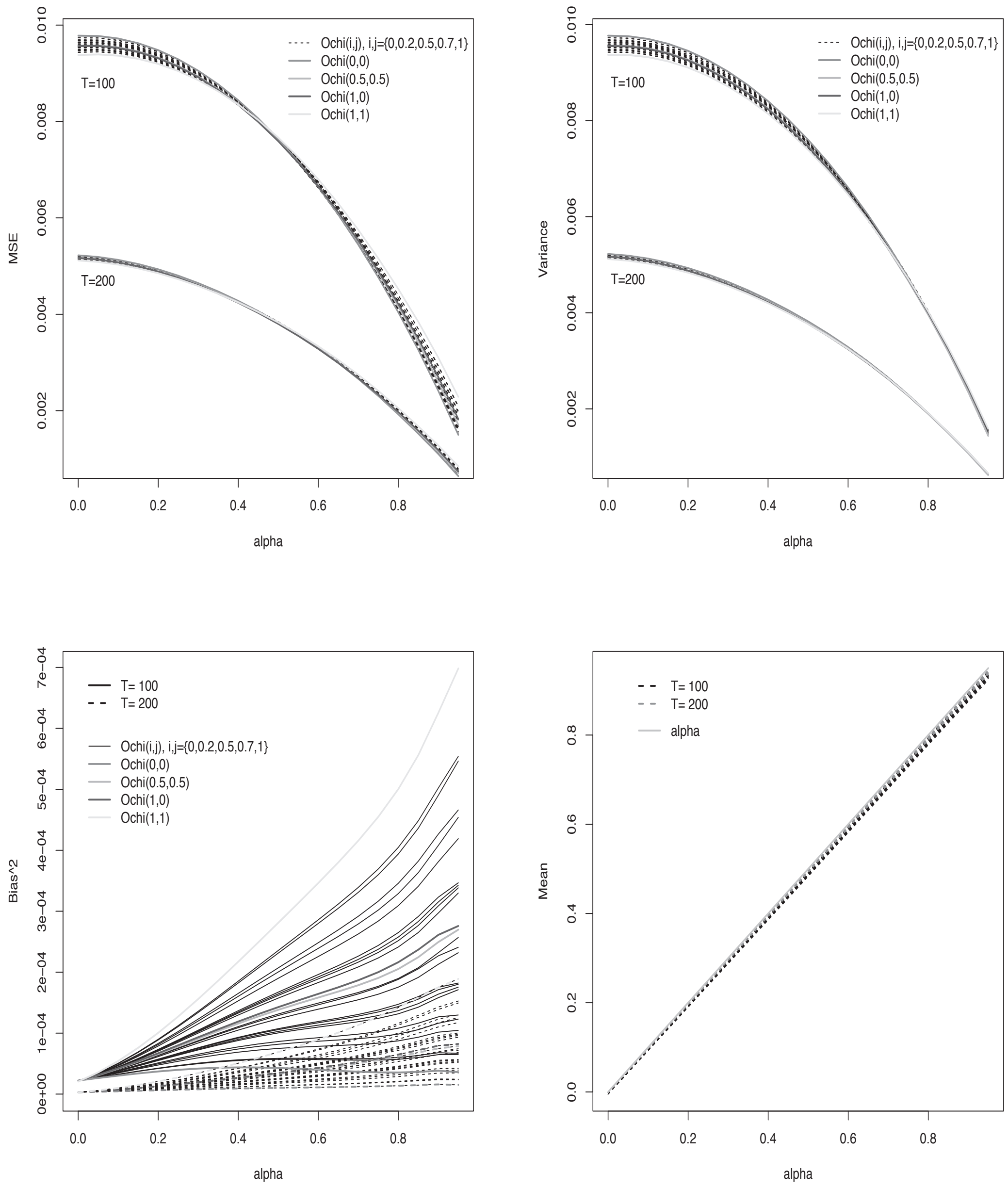

Figure 1: Comparison of Ochi(c_1, c_2) estimators for AR(1) model.

among different pairs of (c_1, c_2). We also see that $(c_1, c_2) = (1, 1)$ gives a better estimation for θ_0 than the other pairs do. However, increasing θ_1 enlarges MSE and gives worse estimation of θ_0.

The second panel shows MSE curves in estimating θ_1. Comparing this with the first panel, we see that a similar trend is obtained, except that Ochi*$(c_1 = 1, c_2 = 1)$ works better than the others when $\theta_1 \leq 0.3$; around a θ_1 of $[0.3, 0.4]$, the difference of attributable to (c_1, c_2) is very small; after that, Ochi*$(c_1 = 1, c_2 = 1)$ becomes worse than the others, and Ochi*$(c_1 = 0, c_2 = 0)$ works better, instead. As special cases of Ochi's estimator, the LSE (Ochi*$(1, 0)$) and Burg's estimator (Ochi*$(0.5, 0.5)$) perform similarly in between Ochi*$(1, 1)$ and Ochi*$(0, 0)$.

Plots of variance and squared biases for both θ_0 and θ_1 are given in the second and third rows, respectively in Figure 2. We can see that $(c_1 = 1, c_2 = 1)$ has better performance than the other pairs of (c_1, c_2) for estimating θ_0. For θ_1, Ochi*$(c_1 = 1, c_2 = 1)$ has smaller variance than with other constants, but its squared bias is bigger.

Since Ochi*$(c_1 = 1, c_2 = 1)$ works well in all the cases for estimating θ_0, we compare it with the methods of LSE, Burg and Yule-Walker in Figure 3 by evaluating their MSE, variance and squared bias curves. In each panel of Figures 3, the results for different sample sizes are indicated by different colors. For each sample size, four curves with respect to θ_1 are plotted, one for each of four different methods. Figure 3 shows that Ochi*$(c_1 = 1, c_2 = 1)$ works well in all the cases for θ_0.

The last two graphs in Figure 3 show the means of the estimates, $\hat{\theta}_0$ and $\hat{\theta}_1$, obtained with different methods and different sample sizes. In the simulation, the true value of θ_0 is fixed to 1, and the true θ_1 takes values from $\{0, 0.05, 0.1, \cdots, 0.55\}$. We see that when θ_1 becomes bigger, the means of the estimates spread from above (with $\hat{\theta}_0$) and below (with $\hat{\theta}_1$) the gray lines in the two graphs. That is, θ_0 is over estimated and θ_1 is under estimated in the ARCH(1) model. The simulation also shows that, in estimating θ_1, Ochi*$(1, 1)$ and the Yule-Walker estimator are not exactly the same but their results are particularly close.

3.3 Data with heavy-tailed distributions In time series analysis, data with heavy-tailed distributions are often of interest. Here, we also evaluate Ochi's estimator, the LSE, Burg's estimator, and the Yule-Walker estimator by simulation when errors ϵ_t and u_t have t distributions with $4, \cdots, 10$ degrees of freedom. Since the simulations give similar results, we show only here the cases of $\epsilon_t \sim$ iid $t(5)$ for the AR(1) model, and $u_t \sim$ iid $t(5)$ for the ARCH(1) model.

For the AR(1) model, Ochi(c_1, c_2) performs similar to its performance in the case of $\epsilon_t \sim N(0, 1)$. This can be seen by comparing Figures 1 and 4.

When $u_t \sim$ iid $t(5)$, among Ochi*(c_1, c_2), $c_1, c_2 \in \{0, 0.2, 0.5, 0.7, 1\}$, Ochi*(1, 1) is better at estimating θ_0. For θ_1, Ochi*(1, 1) maintains a small MSE and shows stability, as seen in the first row of Figure 5. Figure 5 also shows the variance and squared bias for $\hat{\theta}_0$ and $\hat{\theta}_1$ compared for instances of Ochi*(c_1, c_2).

Comparing Ochi*(1, 1) with the LSE, Burg's estimator, and the Yule-Walker estimator in Figure 6, we see that Ochi*(1, 1) works well in most cases for estimating θ_0. In estimating θ_1, from the last panel of Figure 6, we can see that Ochi*(1, 1) and the Yule-Walker estimator give similar performance, and this performance is close to that of Burg's estimator and is more stable than the LSE.

From the ranges of MSE and mean in Figures 5 and 6, we also see that, for the ARCH(1) model, estimation of θ_0 is difficult when u_t has heavy-tailed distributions.

4 Conclusions Ochi's estimator is examined for estimating the parameters in both AR(1) and ARCH(1) models. The simulation for AR(1) models shows that Ochi(1,1) (equivalently Yule-Walker estimator) works well when $\alpha < 0.4$, and Ochi(0,0) gives a smaller MSE when $\alpha > 0.6$. Around $\alpha = 0.5$, there is not much difference in MSEs of Ochi(c_1, c_2). Ochi(1,1) also has bigger squared biases than the other methods. Ochi(1,0) and Ochi(0.5, 0.5) are the LSE and Burg's estimator, respectively. They give intermediate MSE values.

Since ARCH(1) models can be written as a form of AR(1), we introduced Ochi's estimator to ARCH(1). With different pairs of Ochi parameters (c_1, c_2), we investigated its performance by simulation and found that Ochi*($c_1 = 1, c_2 = 1$) works well for estimating θ_0. Ochi*($c_1 = 1, c_2 = 1$) performs similarly to the Yule-Walker estimator, having relatively smaller MSEs than given by the LSE and Burg's estimator for $\theta_1 < 0.3$; when $\theta_1 > 0.4$, the LSE and Burg's estimator work better.

When the data are heavy-tailed, such as when $\epsilon_t \sim t(5)$, for AR(1), Ochi's estimator robustly estimates α. However, for ARCH(1), from the large MSEs, we see that Ochi's estimator gives poor estimation of θ_0. The MSEs for θ_1 are also big, but much smaller than those for θ_0. Moreover, Ochi*(1, 1) shows performance similar to the performance of the

Yule-Walker estimator and Burg's estimator for θ_1.

Acknowledgments The author thanks the editors and the referee for their constructive comments that led to an improved article. The author also expresses her sincere thanks and appreciation to Professor Masanobu Taniguchi for his great help and valuable advices on the paper.

References

[1] Burg, J. P. (1975). Maximum entropy spectral analysis, *PhD thesis, Stanford University* , Available from `http://sepwww.stanford.edu/theses/sep06`.

[2] Chan, K. S., Tong, H. (1986). On estimating thresholds in autoregressive models, *Journal of Time Series Analysis*, **7**, 179–190.

[3] Davis, R.A., Lee, T., and Rodriguez-Yam, G. (2008). Break Detection for a Class of Nonlinear Time Series Models, Journal of Time Series Analysis, **29**, 834–867.

[4] Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, **50** (4), 987–1007.

[5] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, **57** 357–384.

[6] Ochi, Y. (1983). Asymptotic expansions for the distribution of an estimator in the first-order autoregressive process, *Journal of time series analysis*, **4** (1), 57–67.

[7] Shumway, R. H. and Stoffer, D. S. (2011). *Time Series Analysis and Its Applications: With R Examples*, Third Edition. Springer, New York.

[8] Taniguchi, M., Hirukawa, J., Tamaki, K. (2008). *Optimal Statistical Inference in Financial Engineering*, Chapman & Hall/CRC, New York.

[9] Tong, H. (1990). *Non-linear time series*, Oxford Science Publications, Oxford.

[10] Turkman, K., Scotto, M. G., de Zea Bermudez, P. (2014). *Non-Linear Time Series: Extreme Events and Integer Value Problems*, Springer, New York.

[11] Vos, K. (2013). A fast implementation of Burg's method, Available from `http://creativecommons.org/licenses/by/3.0`.

Communicated by *Masanobu Taniguchi*

Xiaoling Dou

Center for Data Science

Waseda University

27 Waseda-cho, Shinjuku-ku, Tokyo 162-0044, Japan

E-mail: xiaoling@aoni.waseda.jp

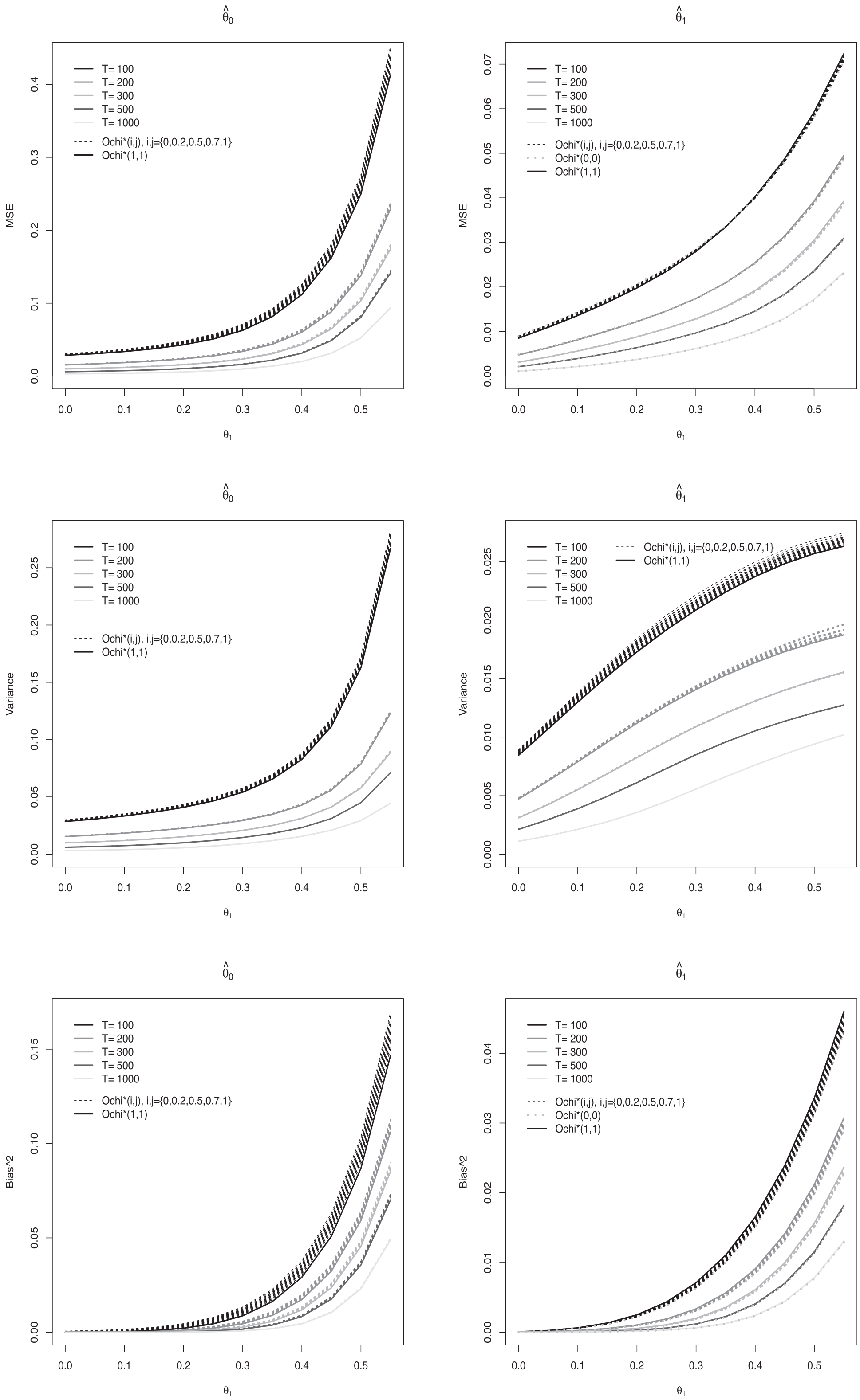

Figure 2: Comparison of Ochi's estimators with different pairs of (c_1, c_2). Left panels show θ_0, right panels show θ_1.

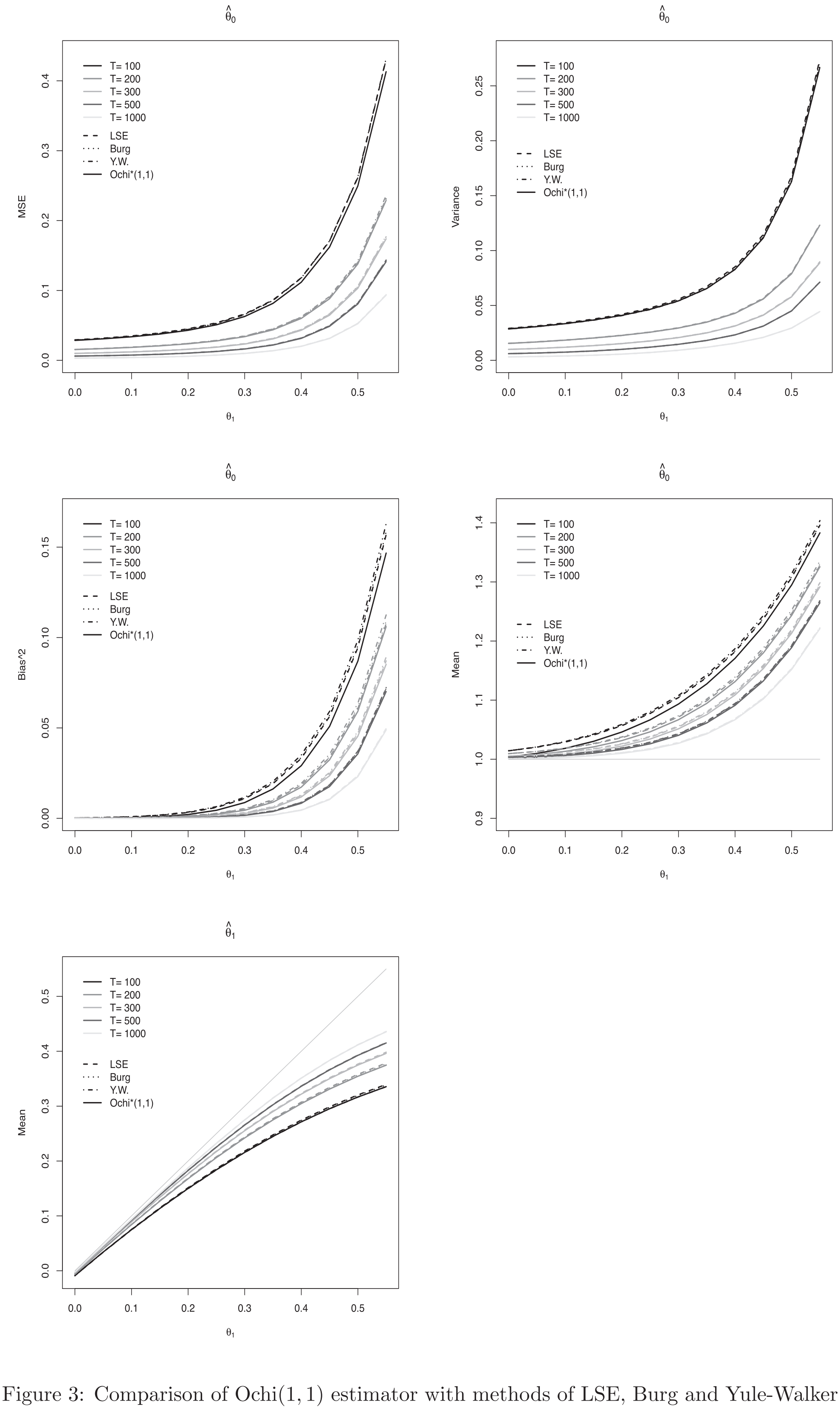

Figure 3: Comparison of Ochi(1, 1) estimator with methods of LSE, Burg and Yule-Walker for θ_0. The last two panels show means of the estimates of parameters θ_0 and θ_1.

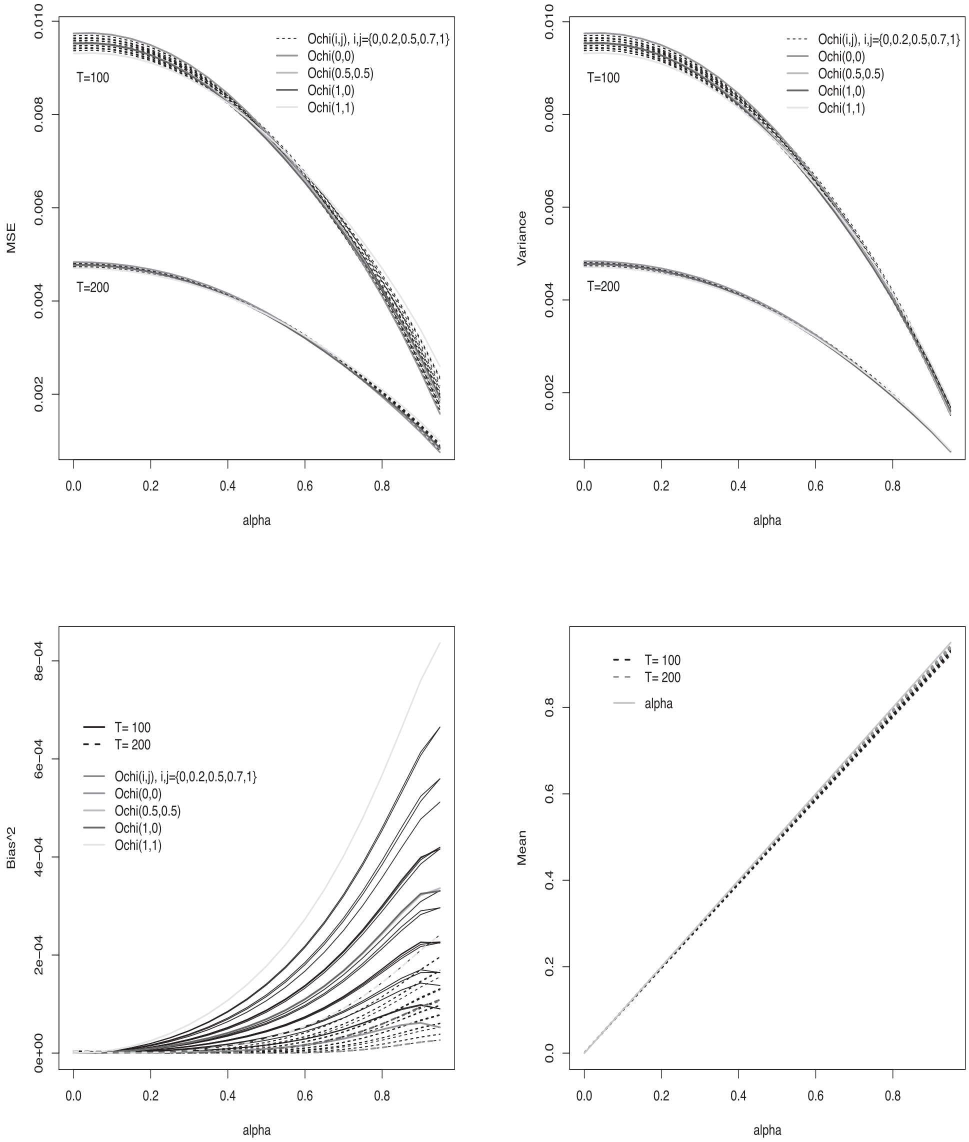

Figure 4: For AR(1), when $\epsilon_t \sim$ iid $t(5)$, Ochi's estimators perform similar to case with $\epsilon_t \sim$ iid $N(0,1)$.

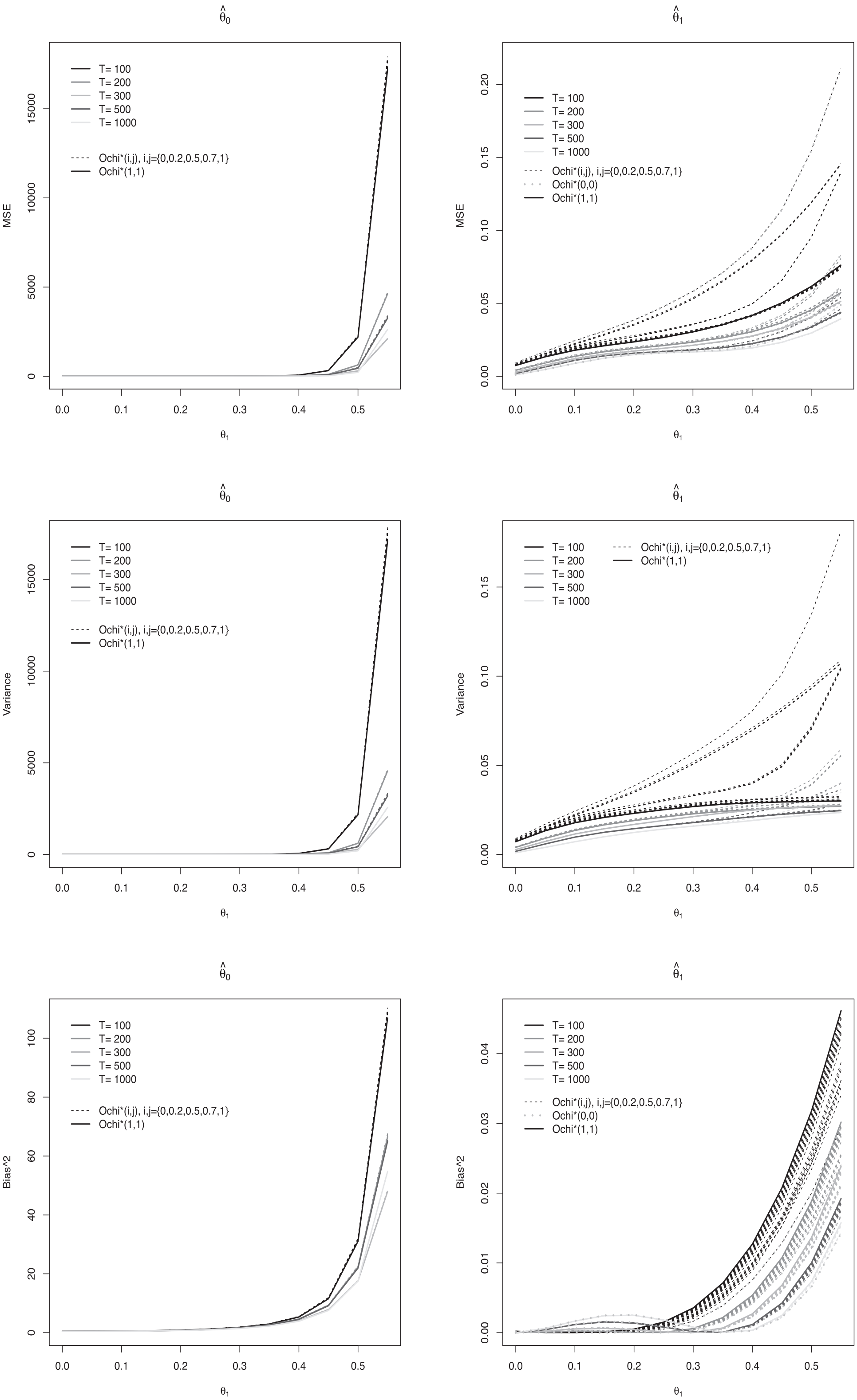

Figure 5: Comparison of Ochi's estimators with different pairs of (c_1, c_2) in case of $u_t \sim$iid $t(5)$. Left panels show θ_0, right panels show θ_1.

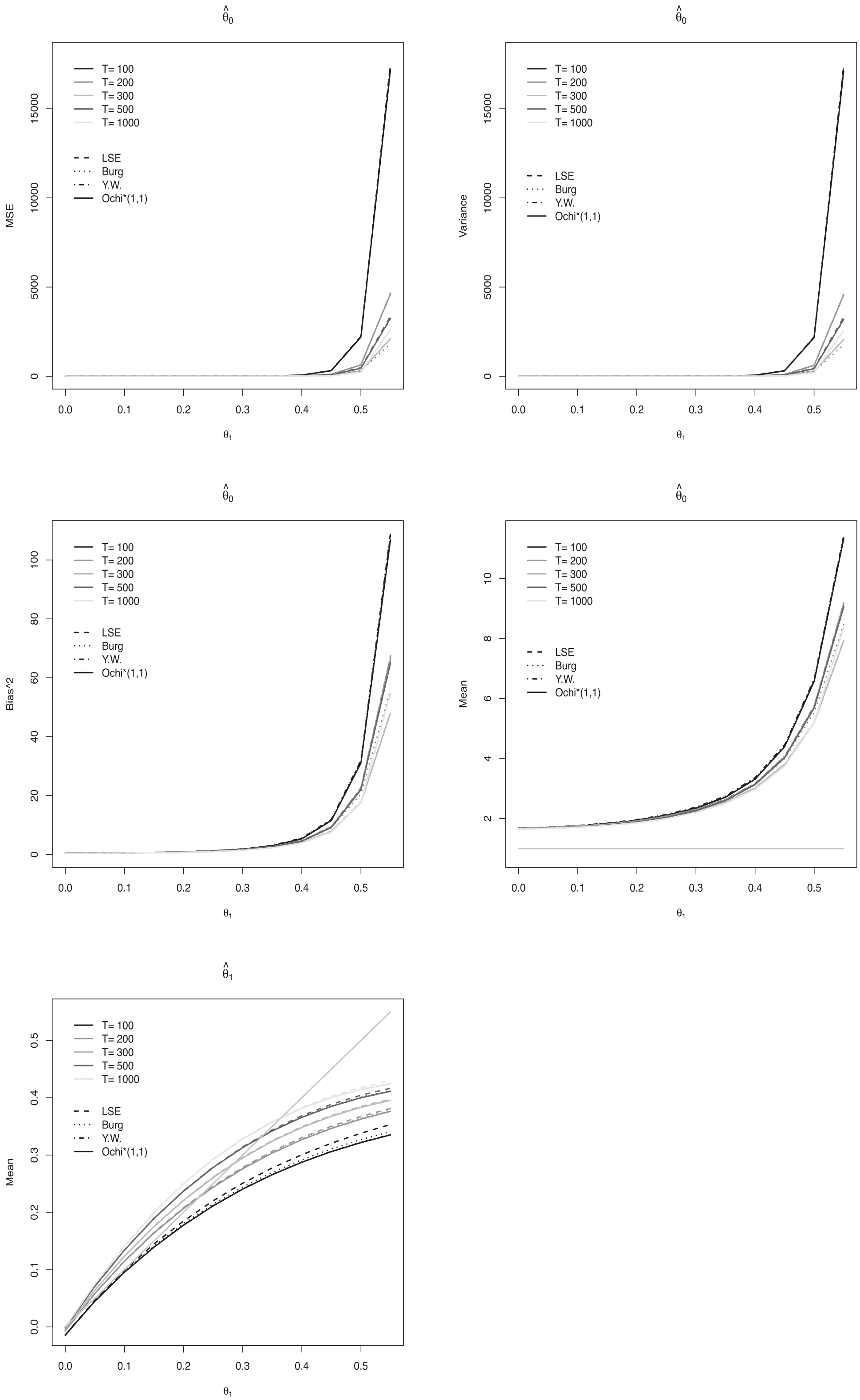

Figure 6: Comparison of Ochi$(1,1)$ estimator and methods of LSE, Burg and Yule-Walker for θ_0 in the case of u_t ∼iid $t(5)$. The last two panels show means of the estimates of parameters θ_0 and θ_1 in the case of u_t ∼iid $t(5)$.

CLASS p-$wA(s,t)$ OPERATORS AND RANGE KERNEL ORTHOGONALITY

T. PRASAD, M. CHŌ, M.H.M RASHID, K. TANAHASHI AND A. UCHIYAMA

Received March 27, 2017; revised June 19, 2017

Abstract. Let $T = U|T|$ be a polar decomposition of a bounded linear operator T on a complex Hilbert space with $\ker U = \ker |T|$. T is said to be class p-$wA(s,t)$ if $\left(|T^*|^t|T|^{2s}|T^*|^t\right)^{\frac{tp}{s+t}} \geq |T^*|^{2tp}$ and $|T|^{2sp} \geq \left(|T|^s|T^*|^{2t}|T|^s\right)^{\frac{sp}{s+t}}$ with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$. This is a generalization of p-hyponormal or class A operators. In this paper we prove following assertions. (i) If T is class p-$wA(s,t)$, then T is normaloid and isoloid. (ii) If T is class p-$wA(s,t)$ and $\sigma(T) = \{\lambda\}$, then $T = \lambda$. (iii) If T is class p-$wA(s,t)$, then T is finite and the range of generalized derivation $\delta_T : B(\mathcal{H}) \ni X \to TX - XT \in B(\mathcal{H})$ is orthogonal to its kernel. (iv) If S is class p-$wA(s,t), T^*$ is an invertible p-$wA(t,s)$ operator and X is a Hilbert-Schmidt operator such that $SX = XT$, then $S^*X = XT^*$.

Dedicated to the memory of Professor Takayuki Furuta with deep gratitude.

1. Introduction and Preliminaries

Let $B(\mathcal{H})$ denote the algebra of all bounded linear operators on a complex Hilbert space $\mathcal{H}$ and let $\ker(T)$, $\operatorname{ran}(T)$ and $\sigma(T)$ denote the kernel, the range and the spectrum of $T \in B(\mathcal{H})$, respectively. Recall that an operator T is said to be hyponormal if $T^*T \geq TT^*$. Aluthge [1] defined p-hyponormal operator as $(T^*T)^p \geq (TT^*)^p$ with $p \in (0,1]$, and he proved many interesting properties of p-hyponormal operators by using Furuta's inequality [9]. An invertible operator T is said to be log-hyponormal if $\log(T^*T) \geq \log(TT^*)$. It is known that invertible p-hyponormal operator is log-hyponormal, but the reverse does not hold by [16]. Moreover, by using Furuta's inequality, Furuta, Ito and Yamazaki [10] define class A operator as

$$|T^2| \geq |T|^2$$

and class $A(k)$ operator as

$$\left(T^*|T|^{2k}T\right)^{\frac{1}{k+1}} \geq |T|^2.$$

These classes are an extension of p-hyponormal, log-hyponormal operators, and moreover, class A and class $A(k)$ operator are extended to class $wA(s,t)$ operators with $0 < s, t$ as

$$\left(|T^*|^t|T|^{2s}|T^*|^t\right)^{\frac{t}{s+t}} \geq |T^*|^{2t} \tag{1.1}$$

and

$$|T|^{2s} \geq \left(|T|^s|T^*|^{2t}|T|^s\right)^{\frac{s}{s+t}}. \tag{1.2}$$

2010 *Mathematics Subject Classification.* 47B20, 47A10.
Key words and phrases. class p-$wA(s,t)$, normaloid, isoloid, finite, orthogonality.

In [8], an operator T is said to be class $A(s,t)$ if T satisfies (1.1). However Ito and Yamazaki [12] proved that (1.1) implies (1.2). This is a striking result. An operator T is said to be class $A(s,t)$ if T satisfies (1.1). Hence Ito and Yamazaki proved that class $wA(s,t)$ coincides with class $A(s,t)$. It is known that every invertible p-hyponormal operator is log-hyponormal, every p-hyponormal, log-hyponormal operator is class $A(s,t)$ for all $0 < s,t$ and if T is invertible and class $A(s,t)$ for all $0 < s,t$ then T is log-hyponormal ([8], [11], [12], [16]).

It is well known that class $A(s,t)$ operators enjoy many interesting properties as hyponormal operators, for example, Fuglede-Putnam type theorem, Weyl type theorem, subscalarity and Putnam's inequality. Although there are many outstanding problems which are still open for hyponormal operators, for example, the invariant subspace problem, investigating new generalizations of hyponormal operators is one of recent interest in operator theory.

For $T \in B(\mathcal{H})$, set $|T| = (T^*T)^{\frac{1}{2}}$ as usual. By taking $U|T|x = Tx$ for $x \in \mathcal{H}$ and $Ux = 0$ for $x \in \ker|T|$, T has a unique polar decomposition $T = U|T|$ with $\ker U = \ker|T|$. An operator T is said to be class p-$wA(s,t)$ [15] if

$$\left(|T^*|^t|T|^{2s}|T^*|^t\right)^{\frac{tp}{s+t}} \geq |T^*|^{2tp} \tag{1.3}$$

and

$$|T|^{2sp} \geq \left(|T|^s|T^*|^{2t}|T|^s\right)^{\frac{sp}{s+t}} \tag{1.4}$$

where $0 < p \leq 1$ and $0 < s,t,s+t \leq 1$. In [5], the authors proved that a set of class p-$wA(s,t)$ operators are increasing for $0 < s,t$ and decreasing for $0 < p \leq 1$.

Lemma 1.1. [5] *If $T \in B(\mathcal{H})$ is class p-$wA(s,t)$ and $0 < s \leq s_1, 0 < t \leq t_1, 0 < p_1 \leq p \leq 1$, then T is class p_1-$wA(s_1,t_1)$.*

Ito and Yamazaki [12] proved that (1.1) implies (1.2). However it is not known that whether (1.3) implies (1.4) or not. Class $A(1,1)$ is said to be class A and class $A(\frac{1}{2},\frac{1}{2})$ is said to be w-hyponormal (see [8, 11, 12, 20]). It is known that an operator T of class A is normaloid, i.e., its spectral radius $r(T)$ coincides with its norm $\|T\|$. Also, class A operator T are isoloid, i.e., its isolated point of spectrum $\sigma(T)$ is a point spectrum of T. The first aim of this paper is to prove that class p-$wA(s,t)$ operator is normaloid and isoloid.

Following [19], we say that an operator $T \in B(\mathcal{H})$ is finite if

$$\|I - (TX - XT)\| \geq 1$$

holds for all $X \in B(\mathcal{H})$. The above inequality is the starting point of the study of commutator approximations, a topic with roots in quantum theory [18]. Let $\mathcal{B}$ denote a Banach algebra. Recall that $b \in \mathcal{B}$ is said to be orthogonal to $a \in \mathcal{B}$, written $b \perp a$, if the inequality

$$\|a\| \leq \|a + \mu b\|$$

holds for all $\mu \in \mathbb{C}$. The above definition of orthogonality has natural geometric meaning, namely, $b \perp a$ if and only if the line $\{a + \mu b : \mu \in \mathbb{C}\}$ is tangent to the ball of center zero and radius $\|a\|$. If $\mathcal{B} = \mathcal{H}$, then the orthogonality means usual sense $\langle a, b\rangle = 0$.

The generalized derivation $\delta_{S,T} : B(\mathcal{H}) \to B(\mathcal{H})$ for $S,T \in B(\mathcal{H})$ is defined by $\delta_{S,T}(X) = SX - XT$ for $X \in B(\mathcal{H})$, and we note $\delta_{T,T} = \delta_T$. If the following inequality

$$\|S-(TX-XT)\| \geq \|S\|$$

holds for all $S \in \ker \delta_T$ and for all $X \in B(\mathcal{H})$, then we say that the range of δ_T is orthogonal to the kernel of δ_T .

Let $T \in B(\mathcal{H})$ and let $\{e_n\}$ be an orthonormal basis of a Hilbert space $\mathcal{H}$. The Hilbert-Schmidt norm is given by

$$\|T\|_2 = \left(\sum_{n=1}^{\infty} \|Te_n\|^2\right)^{\frac{1}{2}}.$$

An operator T is said to be a Hilbert-Schmidt operator if $\|T\|_2 < \infty$ (see [7] for details). $C_2(\mathcal{H})$ denotes a set of all Hilbert-Schmidt operators. For $S,T \in B(\mathcal{H})$, the operator $\Gamma_{S,T}$ defined as $\Gamma_{S,T} : C_2(\mathcal{H}) \ni X \to SXT \in C_2(\mathcal{H})$ has been studied in [3]. It is known that $|\Gamma| \leq \|S\|\|T\|$ and $(\Gamma_{S,T})^*X = S^*XT^* = \Gamma_{S^*,T^*}X$. For more information see [3].

In [19], J. P. Williams proved that normal operators and operators with a compact direct summand are finite. S. Mecheri ([13],[14]) extended Williams's results to more general classes of operators containing the classes of hyponormal operators and paranormal operators and studied range kernel orthogonality for these classes.

The second aim of this paper is to prove that (1) class p-$wA(s,t)$ operators with $0 < s+t \leq 1, 0 < p \leq 1$ are finite, and (2) if $T \in B(\mathcal{H})$ is class p-$wA(s,t)$, then the range of generalized derivation δ_T is orthogonal to its kernel, and (3) if $S \in B(\mathcal{H})$ is class p-$wA(s,t)$ and if $T^* \in B(\mathcal{H})$ is an invertible class p-$wA(t,s)$ operator and X is a Hilbert-Schmidt operator such that $SX = XT$, then $S^*X = XT^*$.

2. Main Results

We begin with the definition of generalized Aluthge transformation.

Definition 2.1. Let $T = U|T| \in B(\mathcal{H})$ be the polar decomposition of T with $\ker U = \ker |T|$. For $s,t > 0$, the generalized Aluthge transformation $T(s,t)$ of T is defined by

$$T(s,t) = |T|^s U |T|^t.$$

Hence, we have

$$T(s,t)^* = |T|^t U^* |T|^s.$$

In [15], the authors proved that if $T \in B(\mathcal{H})$ is class p-$wA(s,t)$, then $T(s,t)$ is $\frac{\rho p}{s+t}$-hyponormal for any $\rho \in (0, \min\{s,t\}]$.

Proposition 2.2. *Let $T \in B(\mathcal{H})$ be class p-$wA(s,t)$ with $0 < p \leq 1$ and $0 < s,t,s+t \leq 1$. Then*

$$|T(s,t)|^{\frac{2tp}{s+t}} \geq |T|^{2tp}$$

and

$$|T|^{2sp} \geq |T(s,t)^*|^{\frac{2sp}{s+t}}.$$

Hence

$$|T(s,t)|^{\frac{2\rho p}{s+t}} \geq |T|^{2\rho p} \geq |T(s,t)^*|^{\frac{2\rho p}{s+t}} \tag{2.1}$$

for any $\rho \in (0, \min\{s,t\}]$.

A complex number λ is said to be an approximate eigenvalue of T if there exists a sequence $\{x_n\}$ of unit vectors such that

$$(T-\lambda)x_n \to 0 \quad (n \to \infty).$$

Also λ is said to be a joint approximate eigenvalue of T if there exists a sequence $\{x_n\}$ of unit vectors such that

$$(T-\lambda)x_n \to 0 \ \text{ and } \ (T-\lambda)^* x_n \to 0 \quad (n \to \infty).$$

We denote the set of all approximate eigenvalues of T by $\sigma_a(T)$ and denote the set of all joint approximate eigenvalues of T by $\sigma_{ja}(T)$. We say that $\lambda \in \sigma(T)$ belongs to the (Xia's) residual spectrum $\sigma_r^X(T)$ of T if $(T-\lambda)\mathcal{H} \neq \mathcal{H}$ and there exists a positive number $c > 0$ such that

$$\|(T-\lambda)x\| \geq c\|x\| \quad \text{for} \quad x \in \mathcal{H}.$$

By the definition, $\sigma(T)$ is a disjoint union of $\sigma_a(T)$ and $\sigma_r^X(T)$.

Recently, the following result was proved by M. Chō, M.H.M. Rashid, K. Tanahashi and A. Uchiyama [5].

Proposition 2.3. [5] *Let $T \in B(\mathcal{H})$ be class p-$wA(s,t)$ with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$. Let $re^{i\theta} \in \mathbb{C}$ with $0 < r$ and $(T - re^{i\theta})x_n \to 0$. Then $(|T|-r)x_n, (U-e^{i\theta})x_n, (U-e^{i\theta})^* x_n, (T-re^{i\theta})^* x_n \to 0$.*

Lemma 2.4. *Let $T = U|T| \in B(\mathcal{H})$ be the polar decomposition of T with* $\ker U = \ker |T|$ *and let $T_\alpha = U|T|^\alpha$ with $0 < \alpha$. Then*

$$0 \in \sigma_a(T) \Longleftrightarrow 0 \in \sigma_a(T_\alpha),$$
$$0 \in \sigma_r^X(T) \Longleftrightarrow 0 \in \sigma_r^X(T_\alpha),$$
$$0 \in \sigma(T) \Longleftrightarrow 0 \in \sigma(T_\alpha).$$

Proof. Let $0 \in \sigma_a(T)$. Then there exist unit vectors x_n such that $Tx_n \to 0$. Then $|T|x_n = U^*U|T|x_n = U^*Tx_n \to 0$. Hence $T_\alpha x_n = U|T|^\alpha x_n \to 0$ and $0 \in \sigma_a(T_\alpha)$. The converse is similar. Let $0 \notin \sigma(T)$. Then $|T|$ is invertible and U is unitary. Hence $T_\alpha = U|T|^\alpha$ is invertible and $0 \notin \sigma(T_\alpha)$. The converse is similar. Since $\sigma(T)$ is a disjoint union of $\sigma_a(T)$ and $\sigma_r^X(T)$, the proof is completed. □

Theorem 2.5. *If $T = U|T| \in B(\mathcal{H})$ is class p-$wA(s,t)$ with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$ and if $T_\alpha = U|T|^\alpha$ with $s+t \leq \alpha$, then*

$$\sigma_a(T_\alpha) = \{r^\alpha e^{i\theta} \mid re^{i\theta} \in \sigma_a(T)\}, \tag{2.2}$$

$$\sigma_r^X(T_\alpha) = \{r^\alpha e^{i\theta} \mid re^{i\theta} \in \sigma_r^X(T)\}, \tag{2.3}$$

$$\sigma(T_\alpha) = \{r^\alpha e^{i\theta} \mid re^{i\theta} \in \sigma(T)\}. \tag{2.4}$$

Proof. Let $T = U|T|$ be class p-$wA(s,t)$ with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$. Let $\lambda = re^{i\theta} \in \sigma_a(T) \setminus \{0\}$ with $0 < r$. Then there exists a sequence $\{x_n\}$ of unit vectors such that $(T - re^{i\theta})x_n \to 0$. Hence $(T-re^{i\theta})^* x_n \to 0, (|T|-r)x_n \to 0, (U-e^{i\theta})x_n \to 0$ and $(U-e^{i\theta})^* x_n \to 0$ by Proposition 2.3. Hence $\lambda_\alpha := r^\alpha e^{i\theta} \in \sigma_{ja}(T_\alpha) \subset \sigma_a(T_\alpha)$. Conversely, let $\mu = r'e^{\phi} \in \sigma_a(T_\alpha) \setminus \{0\}$ with $0 < r'$. Then there exists a sequence unit vectors $\{x_n\}$ such that $(T_\alpha - r'e^{\phi})x_n \to 0$. Since T_α is p-$wA(s/\alpha, t/\alpha)$ and $0 < s/\alpha + t/\alpha \leq 1$, we have that $\mu = r'e^{\phi} \in \sigma_{ja}(T_\alpha)$ by Proposition 2.3. Hence $\mu_{1/\alpha} = (r')^{1/\alpha}e^{i\phi} \in \sigma_{ja}(T) \subset \sigma_a(T)$. Therefore

$$\sigma_a(T_\alpha) \setminus \{0\} = \{r^\alpha e^{i\theta} \mid re^{i\theta} \in \sigma_a(T)\} \setminus \{0\}. \tag{2.5}$$

Hence we have (2.2) by Lemma 2.4.

Next we show (2.3). Let $\lambda = re^{i\theta} \in \sigma_r^X(T) \setminus \{0\}$ with $0 < r$. We claim $\lambda_\alpha = r^\alpha e^{i\theta} \in \sigma(T_\alpha)$.

Assume that $\lambda_\alpha = r^\alpha e^{i\theta} \notin \sigma(T_\alpha)$. Let J be a closed interval $[1,\alpha]$ (or $[\alpha,1]$) and let f be an operator valued continuous function $f(x) := T_x - r^x e^{i\theta}$ $(x \in J)$. Then $f(1)$ is a semi-Fredholm operator with the Fredholm index

$$\operatorname{ind}(f(1)) = \dim(\ker(T - re^{i\theta})) - \dim(\ker(T - re^{i\theta})^*) \leq -1,$$

and $f(\alpha)$ is invertible (so, it is Fredholm with index 0).

We claim that there exists a real number $x_0 \in J$ such that $f(x_0)$ is not semi-Fredholm. Assume that there exists no such $x \in J$. Since $F(J) = \{f(x)|x \in J\}$ is connected in the set of all semi-Fredholm operators of $\mathcal{H}$ and every operator in $F(J)$ has the same Fredholm index, we have that $f(1)$ and $f(\alpha)$ have same Fredholm index. But this is a contradiction.

Since there exists $x_0 \in J$ such that $f(x_0)$ is not semi-Fredholm, we have

$$r^{x_0} e^{i\theta} \in \sigma(T_{x_0}) \setminus \sigma_r^X(T_{x_0}) = \sigma_a(T_{x_0}).$$

Since $s + t \leq x_0$ and $0 < r$, we have $\lambda = re^{i\theta} \in \sigma_a(T)$ by (2.2). But it is a contradiction. Hence $\lambda_\alpha = r^\alpha e^{i\theta} \in \sigma(T_\alpha)$.

We claim $\lambda_\alpha = r^\alpha e^{i\theta} \notin \sigma_a(T_\alpha)$. Assume $\lambda_\alpha = r^\alpha e^{i\theta} \in \sigma_a(T_\alpha)$. Then $\lambda = re^{i\theta} \in \sigma_a(T)$ by (2.2). But it is a contradiction. Hence

$$\{r^\alpha e^{i\theta} \mid re^{i\theta} \in \sigma_r^X(T) \setminus \{0\}\} \subset \sigma_r^X(T_\alpha) \setminus \{0\}.$$

Similarly we have

$$\{(r')^{1/\alpha} e^{i\theta} \mid r' e^{i\theta} \in \sigma_r^X(T_\alpha) \setminus \{0\}\} \subset \sigma_r^X(T) \setminus \{0\}.$$

Hence (2.3) holds by Lemma 2.4. Since $\sigma(T)$ is a disjoint union of $\sigma_a(T)$ and $\sigma_r^X(T)$, the proof of (2.4) is completed,

□

The following result was proved by [5] if $s + t = 1$ and $\rho \neq 0$.

Theorem 2.6. *Let $T \in B(\mathcal{H})$ be class p-$wA(s,t)$ with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$. Let $re^{i\theta} \in \mathbb{C}$ with $0 \leq r$. Then*

$$\ker(T - re^{i\theta}) = \ker(T(s,t) - r^{s+t} e^{i\theta}).$$

Proof. Assume $0 < r$. Let $x \in \ker(T - re^{i\theta})$. Then $|T|x = rx, Ux = e^{i\theta}x$ by Theorem 2.2 of [5]. Hence $T(s,t)x = |T|^s U |T|^t x = r^{s+t} e^{i\theta} x$ and $x \in \ker(T(s,t) - r^{s+t} e^{i\theta})$.

Conversely, let $x \in \ker(T(s,t) - r^{s+t} e^{i\theta})$. Since

$$|T(s,t)|^{\frac{2\rho p}{s+t}} \geq |T|^{2\rho p} \geq |T(s,t)^*|^{\frac{2\rho p}{s+t}} \tag{2.6}$$

and $T(s,t)$ is ρp-hyponormal for any $\rho \in (0, \min\{s,t\}]$ by Proposition 2.2, we have

$$T(s,t)^* x = r^{s+t} e^{-i\theta} x$$

and

$$|T(s,t)|x = |T(s,t)^*|x = r^{s+t} x$$

by Theorem 4 of [4]. Then

$$0 \leq |T(s,t)|^{\frac{2\rho p}{s+t}} - |T|^{2\rho p} \leq |T(s,t)|^{\frac{2\rho p}{s+t}} - |T(s,t)^*|^{\frac{2\rho p}{s+t}},$$

and we have

$$\begin{aligned}
&\| \left(|T(s,t)|^{\frac{2\rho p}{s+t}} - |T|^{2\rho p}\right)^{\frac{1}{2}} x\|^2 = \langle \left(|T(s,t)|^{\frac{2\rho p}{s+t}} - |T|^{2\rho p}\right) x, x\rangle \\
&\leq \langle \left(|T(s,t)|^{\frac{2\rho p}{s+t}} - |T(s,t)^*|^{\frac{2\rho p}{s+t}}\right) x, x\rangle = 0.
\end{aligned}$$

Hence $\left(|T(s,t)|^{\frac{2\rho p}{s+t}} - |T|^{2\rho p}\right)^{\frac{1}{2}} x = 0$ and

$$|T|^{2\rho p} x = |T(s,t)|^{\frac{2\rho p}{s+t}} x = r^{2\rho p} x.$$

This implies $|T|x = rx$. Since

$$r^{s+t} e^{-i\theta} x = T(s,t)^* x = |T|^t U^* |T|^s x = r^s |T|^t U^* x,$$

we have

$$T^* x = |T|^{1-t} |T|^t U^* x = |T|^{1-t} r^t e^{-i\theta} x = r e^{-i\theta} x.$$

Then

$$\begin{aligned}
\|(T - re^{i\theta})x\|^2 &= \|Tx\|^2 - re^{i\theta}\langle x, Tx\rangle - re^{-i\theta}\langle Tx, x\rangle + r^2\|x\|^2 \\
&= \||T|x\|^2 - re^{i\theta}\langle T^*x, x\rangle - re^{-i\theta}\langle x, T^*x\rangle + r^2\|x\|^2 \\
&= (r^2 - r^2 - r^2 + r^2)\|x\|^2 = 0.
\end{aligned}$$

Hence $x \in \ker(T - re^{i\theta})$.

Assume $r = 0$. Let $x \in \ker(T)$. Then $|T|x = 0$ and $T(s,t)x = |T|^s U |T|^t x = 0$.

Conversely, let $x \in \ker(T(s,t))$. Then $|T(s,t)|x = 0$ and $|T|x = 0$ by (2.6). Thus $x \in \ker(T)$.

□

Corollary 2.7. *If T is class p-$wA(s,t)$ with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$, then T is normaloid.*

Proof. Since $T(s,t)$ is $\frac{\rho p}{s+t}$-hyponormal and satisfies

$$|T(s,t)|^{\frac{2\rho p}{s+t}} \geq |T|^{2\rho p} \geq |T(s,t)^*|^{\frac{2\rho p}{s+t}} \tag{2.7}$$

for all $\rho \in (0, \min\{s,t\}]$ by Proposition 2.2, we have

$$\sigma(T(s,t)) = \sigma(|T|^s U |T|^t) = \sigma(U|T|^{s+t}) = \{r^{s+t} e^{i\theta} \mid re^{i\theta} \in \sigma(T)\}$$

by Lemma 6 of [17] and Theorem 2.5. Since $T(s,t)$ is normaloid, we have

$$\begin{aligned}
\||T(s,t)|^{\frac{2\rho p}{s+t}}\| &= \||T(s,t)|\|^{\frac{2\rho p}{s+t}} = \|T(s,t)\|^{\frac{2\rho p}{s+t}} \\
&= r\left(T(s,t)\right)^{\frac{2\rho p}{s+t}} = \left(r(T)^{s+t}\right)^{\frac{2\rho p}{s+t}} = r(T)^{2\rho p},
\end{aligned}$$

and

$$\|T\|^{2\rho p} = \||T|\|^{2\rho p} = \||T|^{2\rho p}\| \leq \||T(s,t)|^{\frac{2\rho p}{s+t}}\| = r(T)^{2\rho p}$$

by (2.7). Hence $\|T\| \leq r(T)$ and therefore $\|T\| = r(T)$. □

Corollary 2.8. *If T is class p-$wA(s,t)$ with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$, then T is isoloid.*

Proof. Let $re^{i\theta}$ be an isolated point of $\sigma(T)$ with $0 \leq r$. Since

$$\sigma(T(s,t)) = \sigma(|T|^s U|T|^t) = \sigma(U|T|^{s+t})$$

by Lemma 6 of [17] and

$$\sigma(U|T|^{s+t}) = \{r^{s+t}e^{i\theta} \mid re^{i\theta} \in \sigma(T)\}$$

by Theorem 2.5, we have $r^{s+t}e^{i\theta}$ is an isolated point of $\sigma(T(s,t))$. We remark $T(s,t)$ is $\frac{\rho p}{s+t}$-hyponormal for any $\rho \in (0, \min\{s,t\}]$ by Proposition 2.2.

Assume $re^{i\theta} = 0$. Since $T(s,t)$ is $\frac{\rho p}{s+t}$-hyponormal, we have $E_0(s,t) = \ker T(s,t)$ where $E_0(s,t)$ is the Riesz idempotent of $T(s,t)$ for $0 \in \operatorname{iso}\sigma(T(s,t))$ by Theorem 5 of [6]. Hence there exists non-zero vector $x \in \mathcal{H}$ such that $T(s,t)x = 0$. Hence $Tx = 0$ by (2.7).

Assume $re^{i\theta} \neq 0$. Then

$$E_{r^{s+t}e^{i\theta}}(s,t) = \ker(T(s,t) - r^{s+t}e^{i\theta}) = \ker((T(s,t) - r^{s+t}e^{i\theta})^*)$$

where $E_{r^{s+t}e^{i\theta}}(s,t)$ is the Riesz idempotent of $T(s,t)$ for $r^{s+t}e^{i\theta} \in \operatorname{iso}\sigma(T(s,t))$ by Theorem 5 of [6]. Hence there exists non-zero vector $x \in \ker(T(s,t) - r^{s+t}e^{i\theta})$ such that $T(s,t)^*x = r^{s+t}e^{-i\theta}x$ and $|T(s,t)|x = |T(s,t)^*|x = r^{s+t}x$ by Theorem 5 of [6]. Then we have

$$\begin{aligned} 0 &= \langle \left(|T(s,t)|^{\frac{2\rho p}{s+t}} - r^{2\rho p} \right) x, x\rangle \geq \langle \left(|T|^{2\rho p} - r^{2\rho p} \right) x, x\rangle \\ &\geq \langle \left(|T(s,t)^*|^{\frac{2\rho p}{s+t}} - r^{2\rho p} \right) x, x\rangle = 0 \end{aligned}$$

by (2.7). Hence $\langle \left(|T|^{2\rho p} - r^{2rp} \right) x, x\rangle = 0$. Since $0 < \rho \leq \min\{s,t\}$ is arbitrary, we have $\langle (|T|^{\rho p} - r^{\rho p}) x, x\rangle = 0$ by the same arguement. Then

$$\begin{aligned} \| (|T|^{\rho p} - r^{\rho p}) x\|^2 &= \langle (|T|^{\rho p} - r^{\rho p})^2 x, x\rangle \\ &= \langle \left(|T|^{2\rho p} - r^{2\rho p} \right) x, x\rangle - 2r^{\rho p}\langle (|T|^{\rho p} - r^{\rho p}) x, x\rangle = 0. \end{aligned}$$

Hence $(|T|^{\rho p} - r^{\rho p}) x = 0$ and this implies $|T|x = rx$. Then $U^*Ux = U^*U|T|r^{-1}x = |T|r^{-1}x = x$. Since $r^{s+t}e^{-i\theta}x = T(s,t)^*x = |T|^tU^*|T|^sx = |T|^tU^*r^sx$, we have $|T|^tU^*x = r^te^{-i\theta}x = |T|^te^{-i\theta}x$. Hence $\left(U^* - e^{-i\theta}\right) x \in \ker|T|^t = \ker|T| = \ker U$. Hence $U\left(U^* - e^{-i\theta}\right) x = 0$ and $UU^*x = e^{-i\theta}Ux$. Then

$$U^*x = U^*UU^*x = e^{-i\theta}U^*Ux = e^{-i\theta}x$$

because $U^*Ux = x$. Then

$$\begin{aligned} \| \left(U - e^{i\theta}\right) x\|^2 &= \langle \left(U - e^{i\theta}\right) x, \left(U - e^{i\theta}\right) x\rangle \\ &= \langle \left(U - e^{i\theta}\right)^* \left(U - e^{i\theta}\right) x, x\rangle \\ &= \langle U^*Ux - e^{-i\theta}\left(U - e^{i\theta}\right) x - e^{i\theta}\left(U^* - e^{-i\theta}\right) x - x, x\rangle \\ &= \langle -e^{-i\theta}x, \left(U - e^{i\theta}\right)^* x\rangle = 0. \end{aligned}$$

Hence $Ux = e^{i\theta}x$. Thus $Tx = U|T|x = re^{i\theta}x$ and the proof is completed. □

Theorem 2.9. *Let $T \in B(\mathcal{H})$ be a class p-$wA(s,t)$ operator with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$ and $\sigma(T) = \{\lambda\}$. Then $T = \lambda$.*

Proof. Let $\lambda = 0$. Since T is normaloid by Corollary 2.7, we have $\|T\| = r(T) = 0$. Hence $T = 0$. Let $\lambda \neq 0$. Then $S := T/\lambda$ is class p-$wA(s,t)$ and $\sigma(S) = \{1\}$. Hence $\|S\| = r(S) = 1$ by Corollary 2.7. Since S^{-1} is class p-$wA(t,s)$ by [17], we have $\|S^{-1}\| = r(S^{-1}) = 1$ by Corollary 2.7. This implies $S = 1$. Hence $T = \lambda$.

□

Theorem 2.10. *Let $T \in B(\mathcal{H})$ be a class p-$wA(s,t)$ operator with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$. Then T is finite.*

Proof. We may assume $T \neq 0$. If $\sigma(T) = \{0\}$, then $T = 0$ by Theorem 2.9. Hence $\sigma(T) \neq \{0\}$. Hence T has an approximate point spectrum $\mu \neq 0$. Hence there exists a sequence $\{x_n\}$ of unit vectors such that $(T-\mu)x_n \to 0$. Then $(T-\mu)^* x_n \to 0$ by Proposition 2.3. Hence $\sigma_{ja}(T) \neq \emptyset$ and $T \in \overline{\mathcal{R}_1}$ where $\mathcal{R}_1$ is a class of all operators with a one-dimensional reducing subspace. Thus T is finite by Theorem 6 of [19]. □

Remark. The referee pointed us a simple proof of Theorem 2.10, that is, since T is normaloid by Corollary 2.7, T is finite by Theorem 5 of [19].

Next we consider a generalization of Theorem 2.10; in other words, we show the range kernel orthogonality of class p-$wA(s,t)$ operator with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$ by the method of [14]. We begin with the following lemma.

Lemma 2.11. *If $T \in B(\mathcal{H})$ is a class p-$wA(s,t)$ operator with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$ and if S is a normal operator such that $TS = ST$, then we have*

$$\|S - (TX - XT)\| \geq |\mu|$$

for all $\mu \in \sigma_p(S)$ and for all $X \in B(\mathcal{H})$.

Proof. Let $\mathcal{M}_\mu$ be an eigen space of $\mu \in \sigma_p(S)$. Since S is normal, the Fuglede-Putnam theorem ensures $TS = ST$ implies $S^*T = ST^*$. Hence $\mathcal{M}_\mu$ reduces both T and S. Now we write matrix representations of T, S and X as

$$T = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}, S = \begin{pmatrix} \mu & 0 \\ 0 & S_2 \end{pmatrix} \text{ and } X = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}.$$

on $\mathcal{H} = \mathcal{M}_\mu \oplus \mathcal{M}_\mu^\perp$. Hence we have

$$S - (TX - XT) = \begin{pmatrix} \mu - (T_1X_1 - X_1T_1) & A \\ B & C \end{pmatrix}.$$

for some operators A, B and C and so

$$\|S - (TX - XT)\| \geq \|\mu - (T_1X_1 - X_1T_1)\|. \tag{2.8}$$

Since T is a class p-$wA(s,t)$ operator and $\mathcal{M}_\mu$ is a reducing subspace of T, the restriction $T_1 = T|_{\mathcal{M}_\mu}$ is a class p-$wA(s,t)$ operator. Since T_1 is finite by Theorem 2.10, we have

$$\|(T_1X_1 - X_1T_1) - \mu\| \geq \|T_1(\frac{X_1}{\mu}) + (\frac{X_1}{\mu})T_1 - 1\||\mu| \geq |\mu|. \tag{2.9}$$

From (2.8) and (2.9), we have

$$\|S - (TX - XT)\| \geq |\mu|$$

for all $X \in B(\mathcal{H})$. □

The following result due to S.K. Berberian [2] is well known.

Proposition 2.12. [2] [Berberian technique] *Let $\mathcal{H}$ be a complex Hilbert space. Then there exist a Hilbert space $\mathcal{K} \supset \mathcal{H}$ and $\psi : B(\mathcal{H}) \to B(\mathcal{K})$ such that ψ is an $*$-isometric isomorphism preserving the order satisfying*

(i) $\psi(T^*) = \psi(T)^*$, $\psi(I_{\mathcal{H}}) = I_{\mathcal{K}}$, $\psi(\alpha T + \beta S) = \alpha\psi(T) + \beta\psi(S)$, $\psi(TS) = \psi(T)\psi(S)$, $\|\psi(T)\| = \|T\|$, $\psi(T) \leq \psi(S)$ *if* $T \leq S$ *for all* $T, S \in B(\mathcal{H})$ *and for all* $\alpha, \beta \in \mathbb{C}$.

(ii) $\sigma(T) = \sigma(\psi(T)), \sigma_a(T) = \sigma_a(\psi(T)) = \sigma_p(\psi(T))$, *where* $\sigma_p(T)$ *is the point spectrum of* T.

Theorem 2.13. *Let $T \in B(\mathcal{H})$ be a class p-$wA(s,t)$ operator operator with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$, and let S be a normal operator such that $TS = ST$. Then*

$$\|S\| \leq \|S - (TX - XT)\|$$

for all $X \in B(\mathcal{H})$.

Proof. By Proposition 2.12, it follows that $\psi(S)$ is normal, $\psi(T)$ is p-$wA(s,t)$ and $\psi(T)\psi(S) = \psi(S)\psi(T)$. Since $\sigma_p(\psi(S)) = \sigma_a(\psi(S)) = \sigma_a(S) = \sigma(S)$, we have

$$|\mu| \leq \|\psi(S) - \psi(T)\psi(X) - \psi(X)\psi(T)\| = \|S - (TX - XT)\|$$

for all $\mu \in \sigma(S)$ and for all $X \in B(\mathcal{H})$ by Lemma 2.11. Hence

$$\sup_{\mu \in \sigma(S)} |\mu| = r(S) = \|S\| \leq \|S - (TX - XT)\|.$$

This completes the proof. $\square$

Now we prove if $S \in B(\mathcal{H})$ is a class p-$wA(s,t)$ operator, $T^* \in B(\mathcal{H})$ is an invertible class p-$wA(t,s)$ operator and $X \in B(\mathcal{H})$ is a Hilbert-Schmidt operator such that $SX = XT$, then $S^*X = XT^*$. The following key lemma is necessary for the proof of theorem 2.15.

Lemma 2.14. *Let $S, T^* \in B(\mathcal{H})$ be class p-$wA(s,t)$ operators with $0 < p \leq 1$ and $0 < s, t, s+t \leq 1$ and let $X \in B(\mathcal{H})$ be a Hilbert-Schmidt operator. Then the operator $\Gamma = \Gamma_{S,T} : C_2(\mathcal{H}) \ni X \to SXT \in C_2(\mathcal{H})$ is class p-$wA(s,t)$.*

Proof. Since $\Gamma^* X = S^*XT^*, |\Gamma|X = |S|X|T^*|, |\Gamma^*|X = |S^*|X|T|$, we have

$$\begin{aligned}
&\left(\left(|\Gamma^*|^t|\Gamma|^{2s}|\Gamma^*|^t\right)^{\frac{tp}{s+t}} - |\Gamma^*|^{2tp}\right)X \\
&= \left(|S^*|^t|S|^{2s}|S^*|^t\right)^{\frac{tp}{s+t}} X \left(|T|^t|T^*|^{2t}|T|^t\right)^{\frac{tp}{s+t}} - |S^*|^{2tp}X|T|^{2tp} \\
&= \left(\left(|S^*|^t|S|^{2s}|S^*|^t\right)^{\frac{tp}{s+t}} - |S^*|^{2tp}\right) X \left(|T|^t|T^*|^{2t}|T|^t\right)^{\frac{tp}{s+t}} \\
&\quad + |S^*|^{2tp}X\left(\left(|T|^t|T^*|^{2t}|T|^t\right)^{\frac{tp}{s+t}} - |T|^{2tp}\right)
\end{aligned}$$

and

$$\begin{aligned}
&\left(|\Gamma|^{2sp} - \left(|\Gamma|^s|\Gamma^*|^{2t}|\Gamma|^s\right)^{\frac{sp}{s+t}}\right)X \\
&= |S|^{2sp}X|T^*|^{2sp} - \left(|S|^s|S^*|^{2t}|S|^s\right)^{\frac{sp}{s+t}} X \left(|T^*|^s|T|^{2t}|T^*|^s\right)^{\frac{sp}{s+t}} \\
&= \left(|S|^{2sp} - \left(|S|^s|S^*|^{2t}|S|^s\right)^{\frac{sp}{s+t}}\right) X|T^*|^{2sp} \\
&\quad + \left(|S|^s|S^*|^{2t}|S|^s\right)^{\frac{sp}{s+t}} X \left(|T^*|^{2sp} - \left(|T^*|^s|T|^{2t}|T^*|^s\right)^{\frac{sp}{s+t}}\right).
\end{aligned}$$

Hence $|\Gamma|^{2sp} - (|\Gamma|^s|\Gamma^*|^{2t}|\Gamma|^s)^{\frac{sp}{s+t}} \geq 0$ and $(|\Gamma^*|^t|\Gamma|^{2t}|\Gamma^*|^t)^{\frac{tp}{s+t}} - |\Gamma^*|^{2tp} \geq 0$. Thus Γ is class p-$wA(s,t)$. $\square$

Theorem 2.15. *Let $S \in B(\mathcal{H})$ be a class p-$wA(s,t)$ operator, $T^* \in B(\mathcal{H})$ be an invertible class p-$wA(t,s)$ operator and $X \in B(\mathcal{H})$ be a Hilbert-Schmidt operator such that $SX = XT$. Then $S^*X = XT^*$.*

Proof. Let $\Gamma_{S,T^{-1}} : C_2(\mathcal{H}) \ni X \to SXT^{-1} \in C_2(\mathcal{H})$. Since S and $(T^*)^{-1}$ are class p-$wA(s,t)$ operators by Corollary 2.4 of [15], Lemma 2.14 ensures that $\Gamma_{S,T^{-1}}$ is class p-$wA(s,t)$. Since $SX = XT$, we have $\Gamma_{S,T^{-1}}X = SXT^{-1} = X$. Applying Proposition 2.3, it follows that $(\Gamma_{S,T^{-1}})^*X = X$. Hence $S^*X(T^{-1})^* = X$ and $S^*X = XT^*$.

$\square$

Acknowledgement. The authors would like to express their sincere thanks to the referee for kind advices and for pointing us a simple proof of Theorem 2.10. The first author is supported by Indian Institute of Science Education and Research-Thiruvananthapuram.

References

[1] A. Aluthge, *On p-hyponormal operators for* $0 < p < 1$, Integral Equations Operator Theory, **13** (1990), 307 - 315.
[2] S. K. Berberian, *Approximate proper vectors*, Proc. Amer. Math. Soc., **13** (1962), 111-114.
[3] S. K. Berberian, *Extensions of a theorem of Fuglede and Putnam*, Proc. Amer. Math. Soc., **71** (1978), 113-114.
[4] M. Chō and T. Huruya, *p-hyponormal operator for* $0 < p < \frac{1}{2}$, Commentations Mthematicae, **33** (1993), 23–29.
[5] M. Chō, M.H.M. Rashid, K. Tanahashi and A. Uchiyama, *Spectrum of class p-wA(s, t) operators*, Acta Sci. Math. (Szeged), **82** (2016), 641–649.
[6] M. Chō and K. Tanahashi, *Isolated point of spectrum of p-hyponormal, log-hyponormal operators*, Integral Equations Operator Theory., **43** (2002), 379–384.
[7] J. B. Conway, *A Course in Functional Analysis, 2nd ed*, Springer, New York, 1990.
[8] M. Fujii, D. Jung, S.H. Lee., M.Y. Lee., and R. Nakamoto, *Some classes of operators related to paranormal and log hyponormal operators*, Math. Japon., **51** (2000), 395-402.
[9] T. Furuta, *$A \geq B \geq O$ assures $(B^rA^pB^r)^{\frac{1}{q}} \geq B^{\frac{p+2r}{q}}$ for $r \geq 0$, $p \geq 0$, $q \geq 1$ with $(1+2r)q \geq (p+2r)$,* Proc. Amer. Math. Soc., **101** (1987), 85–88.
[10] T. Furuta, M. Ito and T. Yamazaki, *A subclass of paranormal operators including class of log-hyponormal and several related classes,* Scientiae Mathematicae, **1** (1998), 389–403.
[11] M. Ito, *Some classes of operators with generalised Aluthege transformations*, SUT J. Math., **35** (1999), 149-165.
[12] M. Ito and T. Yamazaki, *Relations between two inequalities $(B^{\frac{r}{2}}A^pB^{\frac{r}{2}})^{\frac{r}{P+r}} \geq B^r$ and $A^p \geq (A^{\frac{p}{2}}B^rA^{\frac{p}{2}})^{\frac{r}{P+r}}$ and their applications*, Integral Equations Operator Theory, **44** (2002), 442-450.
[13] S. Mecheri, *Finite operators*, Demonstrato. Math., **35** (2002), 355-366.
[14] S. Mecheri, *Finite operators and orthogonality*, Nihonkai Math., **19** (2008) 53-60.
[15] T. Prasad and K. Tanahashi, *On class p-wA(s, t) operators*, Functional Analysis, Approximation and Computation, **6** (2) (2014), 39-42.
[16] K. Tanahashi, *On* log-*hyponormal operators,* Integral Equations Operator Theory, **34** (1999), 364–372.
[17] A. Uchiyama, K. Tanahashi and J. I. Lee, *Spectrum of class A(s, t) operators*, Acta Sci. Math. (Szeged), **70** (2004), 279–287.
[18] H. Wielandt, *Helmut ber die Unbeschrnktheit der Operatoren der Quantenmechanik*, (German) Math. Ann., **121** (1949), 21.
[19] J.P. Williams, *Finite operators*, Proc. Amer. Math. Soc., **26** (1970), 129-135.

[20] M. Yanagida, *Powers of class $wA(s,t)$ operators with generalised Aluthge transformation*, J. Inequal. Appl., **7** (2002), 143-168.

Communicated by *Masatoshi Fujii*

T.Prasad
School of Mathematics
Indian Institute of Science Education and Research-Thiruvananthapuram
Thiruvananthapuram-695016
Kerala
India
E-mail address: prasadvalapil@gmail.com

Muneo Chō
Department of Mathematics, Faculty of Science
Kanagawa University
Hiratsuka 259-1293
Japan
E-mail address: chiyom01@kanagawa-u.ac.jp

M.H.M.Rashid
Department of Mathematics
Faculty of Science P.O.Box(7), Mu'tah university
Al-Karak
Jordan
E-mail address: malik okasha@yahoo.com

Kotaro Tanahashi
Department of Mathematics
Tohoku Medical and Medical and Pharmaceutical University
Sendai, 981-8558
Japan
E-mail address: tanahasi@tohoku-mpu.ac.jp

Atsushi Uchiyama
Department of Mathematical Science, Faculty of Science
Yamagata University
Yamagata, 990-8560
Japan
E-mail address: uchiyama@sci.kj.yamagata-u.ac.jp

MARTINGALE BESOV SPACES AND MARTINGALE TRIEBEL-LIZORKIN SPACES

Gaku Sadasue

Received November 8, 2017; revised December 24, 2017

Abstract. In this paper, we give a definition of martingale Besov spaces and martingale Triebel-Lizorkin spaces for general filtrations. We investigate several fundamental properties of these spaces.

1 Introduction The theory of Besov spaces and Triebel-Lizorkin spaces provides us a unified approach to various important function spaces such as L_p-spaces, Hardy spaces, BMO spaces, Lipschitz spaces and Sobolev spaces. From such diversity, Besov spaces and Triebel-Lizorkin spaces are useful in various mathematical branches.

In martingale theory, Chao and Peng [5] gave a definition of Besov spaces and Triebel-Lizorkin spaces for p-adic martingales and pointed out some fundamental properties of these spaces. They used martingale Besov spaces for characterization of Schatten-von Neumann properties of commutators. For general filtrations, Weisz [17] proved duality theorems among martingale Hardy spaces of q-variations, including the duality between martingale Hardy spaces and martingale BMO spaces of q-variations. We note that these spaces coincide with martingale Triebel-Lizorkin spaces when the smoothness parameter equals to 0, and that Weisz's duality theorem is an early general result on martingale Triebel-Lizorkin spaces.

In this paper, we give a definition of martingale Besov spaces and martingale Triebel-Lizorkin spaces for general filtrations. We give proofs for several fundamental properties of these spaces such as duality, complex interpolation and norm equivalence in a general framework. We also study some embeddings under additional assumptions on filtrations. It relates to recent progress of the theory of fractional integral of martingales ([4], [7], [8], [11], [14]). In fact, we apply our results to the boundedness of fractional integrals of martingales and obtain some improvement.

The organization of this paper is as follows. In the next section, we give the definition of martingale Besov-Triebel-Lizorkin spaces for general filtrations and describe our results. In Section 3, we prove some basic properties of martingale Besov-Triebel-Lizorkin spaces. In Section 4, we show a duality between martingale Besov-Triebel-Lizorkin spaces. In Section 5, we study complex interpolation of martingale Besov-Triebel-Lizorkin spaces. In Section 6, we show a norm equivalence in terms of mean oscillations. In Section 7, we prove some embedding theorem under additional assumptions on filtrations. Finally in Section 8, we give an application of our results to the boundedness of fractional integral of martingales.

2 Notations, definitions and results Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{\mathcal{F}_n\}_{n\geq 0}$ be a filtration, that is, nondecreasing sequence of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F} = \sigma(\bigcup_n \mathcal{F}_n)$. The expectation operator and the conditional expectation operators relative to $\mathcal{F}_n$ are denoted by E and E_n, respectively. For simplicity, we use the convention $E_{-1} = 0$.

2010 *Mathematics Subject Classification.* Primary 60G46; Secondary 46E30, 42B35, 60G42.
Key words and phrases. martingale, Besov space, Triebel-Lizorkin space, fractional integral.

We say a sequence of measurable functions $f=(f_n)_{n\geq 0}$ is adapted if f_n is $\mathcal{F}_n$-measurable for every $n\geq 0$.

We denote by $\mathcal{V}$ the set of all adapted sequence of functions $v=(v_n)_{n\geq 0}$ satisfying that $v_0=1$ and that there exist constants $\delta_2\geq\delta_1>1$ such that

$$\delta_1 v_{n-1}\leq v_n\leq \delta_2 v_{n-1} \quad \text{for all} \quad n\geq 1. \tag{2.1}$$

By $v_0=1$ and (2.1), if $(v_n)_{n\geq 0}\in\mathcal{V}$, then

$$\delta_1^n\leq v_n\leq \delta_2^n \quad \text{for all} \quad n\geq 0 \tag{2.2}$$

for some $\delta_2\geq\delta_1>1$. For $(v_n)_{n\geq 0}\in\mathcal{V}$, we use the convention $v_{-1}=v_0$.

Let $(f_n)_{n\geq 0}$ be a sequence of integrable functions. We say $(f_n)_{n\geq 0}$ is a martingale relative to $\{\mathcal{F}_n\}_{n\geq 0}$ if it is adapted and satisfies $E_n[f_m]=f_n$ for every $n\leq m$. For a martingale $f=(f_n)_{n\geq 0}$, let $d_nf=f_n-f_{n-1}$ with convention $f_{-1}=0$. We denote by $\mathcal{M}$ the set of all martingales.

For $p\in[1,\infty)$, let $\mathcal{M}_p$ be the set of all L_p-bounded martingales. It is known that, if $p\in(1,\infty)$, then any L_p-bounded martingale converges in L_p. Moreover, if $f\in L_p$, $p\in[1,\infty)$, then $(f_n)_{n\geq 0}$ with $f_n=E_nf$ is in $\mathcal{M}_p$ and converges to f in L_p (see for example [10]). For this reason a function $f\in L_1$ and the corresponding martingale $(f_n)_{n\geq 0}$ with $f_n=E_nf$ will be denoted by the same symbol f. Note also that $\|f\|_{L_p}=\sup_{n\geq 0}\|E_nf\|_{L_p}$.

We now introduce martingale Besov spaces and martingale Triebel-Lizorkin spaces. Our definition is a generalization of Chao and Peng's one in [5].

Definition 2.1. Let $p\in(0,\infty]$, $q\in(0,\infty]$, $s\in\mathbb{R}$ and $v=(v_n)_{n\geq 0}\in\mathcal{V}$. For $f=(f_n)_{n\geq 0}\in\mathcal{M}$, define $\|f\|_{B^s_{pq}}=\|f\|_{B^s_{pq}(v)}$ and $\|f\|_{F^s_{pq}}=\|f\|_{F^s_{pq}(v)}$ by

$$\|f\|_{B^s_{pq}}=\left(\sum_{n=0}^{\infty}\|v^s_{n-1}d_nf\|^q_{L_p}\right)^{1/q} \quad\text{and}\quad \|f\|_{F^s_{pq}}=\left\|\left(\sum_{n=0}^{\infty}|v^s_{n-1}d_nf|^q\right)^{1/q}\right\|_{L_p} \tag{2.3}$$

respectively if $p<\infty$ and $q<\infty$ with convention $v_{-1}=v_0$ and $f_{-1}=0$.

If $p<\infty$ and $q=\infty$, then define

$$\|f\|_{B^s_{p\infty}}=\sup_{n\geq 0}\|v^s_{n-1}d_nf\|_{L_p} \quad\text{and}\quad \|f\|_{F^s_{p\infty}}=\left\|\sup_{n\geq 0}|v^s_{n-1}d_nf|\right\|_{L_p},$$

and if $p=\infty$ and $q<\infty$, then define

$$\|f\|_{B^s_{\infty q}}=\left(\sum_{n=0}^{\infty}\|v^s_{n-1}d_nf\|^q_{L_\infty}\right)^{1/q} \quad\text{and}\quad \|f\|_{F^s_{\infty q}}=\sup_{n\geq 0}\left\|E_n\left[\sum_{k=n}^{\infty}|v^s_{k-1}d_kf|^q\right]^{1/q}\right\|_{L_\infty},$$

and if $p=q=\infty$, then define

$$\|f\|_{B^s_{\infty\infty}}=\sup_{n\geq 0}\|v^s_{n-1}d_nf\|_{L_\infty} \quad\text{and}\quad \|f\|_{F^s_{\infty\infty}}=\left\|\sup_{n\geq 0}|v^s_{n-1}d_nf|\right\|_{L_\infty}$$

respectively with the same convention as in (2.3).

Then, the spaces $B^s_{pq}=B^s_{pq}(v)$ and $F^s_{pq}=F^s_{pq}(v)$ are defined by

$$B^s_{pq}=\{f\in\mathcal{M}:\|f\|_{B^s_{pq}}<\infty\} \quad\text{and}\quad F^s_{pq}=\{f\in\mathcal{M}:\|f\|_{F^s_{pq}}<\infty\}$$

respectively.

$\|f\|_{B^s_{pq}}$ and $\|f\|_{F^s_{pq}}$ are quasi-norms on B^s_{pq} and F^s_{pq} respectively. We call $B^s_{pq} = B^s_{pq}(v)$ a martingale Besov space associated to v and call $F^s_{pq} = F^s_{pq}(v)$ a martingale Triebel-Lizorkin space associated to v.

Remark 2.1. For $f = (f_n)_{n\geq 0} \in \mathcal{M}$, the square functions $S_n(f)$, where $n \geq 0$, and $S(f)$ are defined by

$$S_n(f) = \left(\sum_{k=0}^{n} |d_k f|^2\right)^{1/2} \quad \text{and} \quad S(f) = \left(\sum_{n=0}^{\infty} |d_n f|^2\right)^{1/2}$$

with convention $f_{-1} = 0$. Then, for $p \in (0, \infty)$, the martingale Hardy spaces H^S_p is defined by

$$H^S_p = \{f \in \mathcal{M} : \|S(f)\|_p < \infty\}.$$

The space F^0_{p2} coincides with H^S_p for $p \in (0, \infty)$. Moreover, if $p > 1$, then $F^0_{p2} = H^S_p \sim L_p$. Furthermore, martingale space BMO^{S-}_2 is defined by

$$\mathrm{BMO}^{S-}_2 = \{f \in \mathcal{M} : \|f\|_{\mathrm{BMO}^{S-}_2} < \infty\},$$

where

$$\|f\|_{\mathrm{BMO}^{S-}_2} = \sup_{n\geq 0} \|E_n[S(f)^2 - S_{n-1}(f)^2]^{1/2}\|_\infty$$

with convention $S_{-1}(f) = 0$. The space $F^0_{\infty 2}$ coincides with BMO^{S-}_2. For the theory of martingale Hardy spaces and martingale BMO spaces, we refer to [6], [10] and [16].

For $v = (v_n)_{n\geq 0} \in \mathcal{V}$, define $u = (u_n)_{n\geq 0}$ by $u_n = v_n^{-1}$ for $n \geq 0$. For $\alpha \in \mathbb{R}$ and $f = (f_n)_{n\geq 0} \in \mathcal{M}$, define a martingale $I^u_\alpha f = ((I^u_\alpha f)_n)_{n\geq 0}$ by

$$(I^u_\alpha f)_n = \sum_{k=0}^{n} u^\alpha_{k-1} d_k f$$

with convention $u_{-1} = u_0$, $f_{-1} = 0$ and $(I^u_\alpha f)_{-1} = 0$.

Our first result is a lifting property of I^u_α. It is a direct consequence of the definition, but for its importance, we give a proof.

Theorem 2.1. *Let $v = (v_n)_{n\geq 0} \in \mathcal{V}$. Define $u = (u_n)_{n\geq 0}$ by $u_n = v_n^{-1}$ for $n \geq 0$. Let $\alpha \in \mathbb{R}$. Then, I^u_α is an isometric isomorphism from B^s_{pq} to $B^{s+\alpha}_{pq}$ and F^s_{pq} to $F^{s+\alpha}_{pq}$ respectively for $p \in (0, \infty]$, $q \in (0, \infty]$ and $s \in \mathbb{R}$.*

Proof. Since $d_n(I^u_\alpha f) = u^\alpha_{n-1} d_n f$ for $n \geq 0$, it is clear that I^u_α is a bijection from $\mathcal{M}$ to $\mathcal{M}$ with the inverse map $I^u_{-\alpha}$. Moreover, we have

$$v^{s+\alpha}_{n-1} d_n(I^u_\alpha f) = v^s_{n-1} d_n f \quad \text{for all} \quad n \geq 0. \tag{2.4}$$

By (2.4), we have

$$\|I^u_\alpha f\|_{B^{s+\alpha}_{pq}} = \|f\|_{B^s_{pq}} \quad \text{and} \quad \|I^u_\alpha f\|_{F^{s+\alpha}_{pq}} = \|f\|_{F^s_{pq}}.$$

This is the desired conclusion. □

Our next result is a duality between martingale Besov-Triebel-Lizorkin spaces. For $p \in [1, \infty]$, we denote by p' the conjugate exponent of p, that is,

$$p' = \begin{cases} p/(p-1) & \text{if} \quad 1 < p < \infty, \\ 1 & \text{if} \quad p = \infty, \\ \infty & \text{if} \quad p = 1. \end{cases}$$

We use the notation A^s_{pq} to denote either B^s_{pq} or F^s_{pq} for short.

Theorem 2.2. *Let $v = (v_n)_{n\geq 0} \in \mathcal{V}$, $s \in \mathbb{R}$, $p \in [1,\infty)$ and $q \in [1,\infty)$. Denote by p' and q' the conjugate exponents of p and q respectively. Let $(A^s_{pq})'$ denote the topological dual space of A^s_{pq}. Then, $(A^s_{pq})'$ is isomorphic to $A^{-s}_{p'q'}$ under the pairing $(g,f) \mapsto \sum_{n=0}^{\infty} E[d_n\overline{g}d_nf]$ with convention $g_{-1} = f_{-1} = 0$. More precisely, there exists a positive constant C depending only on p and q such that the following (1) and (2) hold:*

(1) If $g \in A^{-s}_{p'q'}$, then the infinite sum $\sum_{n=0}^{\infty} E[d_n\overline{g}d_nf]$ converges for every $f \in A^s_{pq}$. Moreover,

$$\left|\sum_{n=0}^{\infty} E[d_n\overline{g}d_nf]\right| \leq C\|g\|_{A^{-s}_{p'q'}}\|f\|_{A^s_{pq}} \quad (f \in A^s_{pq}).$$

(2) Conversely, for each $\Phi \in (A^s_{pq})'$, there exists $h \in A^{-s}_{p'q'}$ such that

$$\Phi(f) = \sum_{n=0}^{\infty} E[d_n\overline{h}d_nf] \quad (f \in A^s_{pq})$$

and that $\|h\|_{A^{-s}_{p'q'}} \leq C\|\Phi\|_{(A^s_{pq})'}$.

The proof of Theorem 2.2 is given in Section 4.

Remark 2.2. The duality of the case $s = 0$ and $A = F$ was proved in [17, Theorem 14 and 17].

Further, we investigate the complex interpolation between martingale Besov-Triebel-Lizorkin spaces. We recall the definition of the first Calderón's complex interpolation functor.

Let $S = \{z \in \mathbb{C} : 0 \leq \mathrm{Re}z \leq 1\}$ and $S_0 = \{z \in \mathbb{C} : 0 < \mathrm{Re}z < 1\}$. Let (A_0, A_1) be a compatible couple of Banach spaces. We denote by $\mathcal{F}(A_0, A_1)$ the set of all $(A_0 + A_1)$-valued bounded continuous functions F on S which is holomorphic in S_0 and moreover, $t \mapsto F(j + it)$ $(j = 0, 1)$ is a function from $\mathbb{R}$ into A_j satisfying $\|F(j + it)\|_{A_j} \to 0$ as $|t| \to \infty$. As is shown in [2, Lemma 4.1.1], the space $\mathcal{F}(A_0, A_1)$ equipped with the norm

$$\|F\|_{\mathcal{F}(A_0,A_1)} = \max\left(\sup_{t\in\mathbb{R}}\|F(it)\|_{A_0}, \sup_{t\in\mathbb{R}}\|F(1+it)\|_{A_1}\right)$$

is a Banach space.

Definition 2.2. Let (A_0, A_1) be a compatible couple of Banach spaces. For $\theta \in [0, 1]$, define $[A_0, A_1]_\theta$ by

$$[A_0, A_1]_\theta = \{f \in A_0 + A_1 : f = F(\theta) \quad \text{for some} \quad F \in \mathcal{F}(A_0, A_1)\}$$

equipped with the norm

$$\|f\|_{[A_0,A_1]_\theta} = \inf_{F(\theta)=f} \|F\|_{\mathcal{F}(A_0,A_1)}.$$

We now state our result on complex interpolation of martingale Besov-Triebel-Lizorkin spaces.

Theorem 2.3. *Let $v \in \mathcal{V}$, $\theta \in (0, 1)$, $s_0, s_1 \in \mathbb{R}$ and $p_0, p_1, q_0, q_1 \in [1,\infty]$ with $\min(q_0, q_1) < \infty$. Define s, p and q by*

$$s = (1-\theta)s_0 + \theta s_1, \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1} \tag{2.5}$$

with convention $1/\infty = 0$. Then, the following (i) and (ii) hold.

(i) $[B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1}]_\theta = B^s_{pq}$ *with equivalence of norms.*

(ii) Assume that $1 < p_0, p_1 < \infty$*. Then,* $[F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1}]_\theta = F^s_{pq}$ *with equivalence of norms.*

The proof of Theorem 2.3 is given in Section 5.

Remark 2.3. In the theory of Besov-Triebel-Lizorkin spaces on Euclidean spaces, the complex interpolation is investigated for $p_0, p_1, q_0, q_1 \in (0, \infty]$ by using the framework of distribution valued analytic functions ([15, Section 2.4.4]) and by using isomorphisms to sequence spaces ([9, Theorem 9.1]). Since these methods are not known for martingales of general filtrations, we restrict ourselves to the case where $p_0, p_1, q_0, q_1 \in [1, \infty]$.

In the next section, we will show that if $s \in (0, \infty)$, $p \in [1, \infty]$ and $q \in (0, \infty]$, then $B^s_{pq} \subset L_p$ and $F^s_{pq} \subset L_p$. Further, in Section 6, we prove the following norm equivalence in terms of mean oscillations.

Theorem 2.4. *Let* $v \in \mathcal{V}, s \in (0, \infty)$*,* $p \in [1, \infty]$ *and* $q \in (0, \infty]$*. Let* $f \in L_p$ *and identify* f *with the corresponding martingale* $(f_n)_{n\geq 0} = (E_n f)_{n\geq 0}$*. Then, the following norm equivalence holds:*

$$\|f\|_{B^s_{pq}} \sim \left\|(\|v^s_{n-1} E_n|f - f_{n-1}|\|_{L_p})_{n\geq 0}\right\|_{\ell_q}. \tag{2.6}$$

Moreover, if $1 < p < \infty$ *and* $q \geq 1$*, then*

$$\|f\|_{F^s_{pq}} \sim \left\|\|(v^s_{n-1} E_n|f - f_{n-1}|)_{n\geq 0}\|_{\ell_q}\right\|_{L_p}. \tag{2.7}$$

Note that we do not need any assumption on $\{\mathcal{F}_n\}_{n\geq 0}$ in Theorems 2.1, 2.2, 2.3 and 2.4.

To study embeddings, we need some assumptions on $\{\mathcal{F}_n\}_{n\geq 0}$. $B \in \mathcal{F}_n$ is called an atom (more precisely a $(\mathcal{F}_n, P)$-atom), if any $A \subset B$ with $A \in \mathcal{F}_n$ satisfies $P(A) = P(B)$ or $P(A) = 0$. Below, we assume that

$$\text{every } \sigma\text{-algebra } \mathcal{F}_n \text{ is generated by countable atoms.} \tag{2.8}$$

We denote by $A(\mathcal{F}_n)$ the set of all atoms in $\mathcal{F}_n$. We define $\mathcal{F}_n$-measurable functions b_n and v_n by

$$b_n = \sum_{B\in A(\mathcal{F}_n)} P(B)\chi_B, \quad v_n = b_n^{-1}. \tag{2.9}$$

We also assume that $\{\mathcal{F}_n\}_{n\geq 0}$ is regular, that is, there exists $R \geq 2$ such that

$$E_n f \leq R E_{n-1} f \text{ for all } n \geq 1 \text{ and non-negative integrable function } f. \tag{2.10}$$

Further, for the sake of simplicity, we assume that

$$\begin{aligned} &\text{If } B \in A(\mathcal{F}_{n-1}),\ B' \in A(\mathcal{F}_n) \text{ and } B' \subset B, \\ &\text{then } P(B') < P(B) \text{ for every } n \geq 1. \end{aligned} \tag{2.11}$$

$$\mathcal{F}_0 = \{\emptyset, \Omega\}. \tag{2.12}$$

If (2.8), (2.10), (2.11) and (2.12) hold, then, by [11, Lemma 3.3],

$$\left(1 + \frac{1}{R}\right) b_n \leq b_{n-1} \leq R b_n$$

for every $n \geq 1$. Hence, we obtain that the sequence $v = (v_n)_{n\geq 0}$ defined in (2.9) belongs to $\mathcal{V}$.

As for embeddings, we show the following two theorems. For quasi-normed space X and Y, we denote by $X \hookrightarrow Y$ if the identity map from X is a continuous map into Y.

Theorem 2.5. *Suppose that every σ-algebra $\mathcal{F}_n$ is generated by countable atoms. Furthermore, assume that $\{\mathcal{F}_n\}_{n\geq 0}$ is regular with* (2.11) *and* (2.12)*. Let $v = (v_n)_{n\geq 0}$ be the sequence of functions defined in* (2.9)*. Let $s \in \mathbb{R}$, $q \in (0,\infty)$ and $p_0, p_1 \in (0,\infty)$ with $p_0 < p_1$. Let $\alpha = 1/p_0 - 1/p_1$. Then,*

$$B^{s+\alpha}_{p_0 q} \hookrightarrow B^s_{p_1 q} \quad \textit{and} \quad F^{s+\alpha}_{p_0 \infty} \hookrightarrow F^s_{p_1 q}. \tag{2.13}$$

Theorem 2.6. *Suppose that every σ-algebra $\mathcal{F}_n$ is generated by countable atoms. Furthermore, assume that $\{\mathcal{F}_n\}_{n\geq 0}$ is regular with* (2.11) *and* (2.12)*. Let $v = (v_n)_{n\geq 0}$ be the sequence of functions defined in* (2.9)*. Let $s \in \mathbb{R}$, $p \in (0,\infty)$ and $q \in (0,\infty]$. Let A^s_{pq} denote either B^s_{pq} or F^s_{pq}. If $s > 1/p$, then*

$$A^s_{pq} \hookrightarrow B^{s-1/p}_{\infty\infty}. \tag{2.14}$$

The proofs of Theorems 2.5 and 2.6 are given in Section 7.

We apply our results to the boundedness of fractional integral for martingales. To explain this application, we recall the definition of fractional integrals for martingales.

Definition 2.3. Let $\alpha \in \mathbb{R}$. Suppose that every σ-algebra $\mathcal{F}_n$ is generated by countable atoms. Let b_n be the function defined in (2.9). For a martingale $(f_n)_{n\geq 0}$, define a martingale $I_\alpha f = ((I_\alpha f)_n)_{n\geq 0}$ by

$$(I_\alpha f)_n = \sum_{k=0}^{n} b^\alpha_{k-1} d_k f$$

with convention $b_{-1} = b_0$ and $f_{-1} = 0$. If $\alpha > 0$, then we call $I_\alpha f$ the fractional integral of f of order α.

Further, we recall the definition of martingale Lipschitz spaces ([16, page 7]). For $s > 0$ and $f \in L_1$, let

$$\|f\|_{\Lambda_1^-(s)} = \sup_{n\geq 0} \sup_{B\in A(\mathcal{F}_n)} \frac{1}{P(B)^{1+s}} \int_B |f(\omega) - (E_{n-1}f)(\omega)|\, dP(\omega)$$

with convention $E_{-1}f = 0$. We do not assume $E_0 f = 0$, different from [16]. Then define

$$\Lambda_1^-(s) = \{f \in L_1 : \|f\|_{\Lambda_1^-(s)} < \infty\}. \tag{2.15}$$

We regard $\Lambda_1^-(s)$ as martingale spaces by the identification $f \in L_1$ with the corresponding martingale $(E_n f)_{n\geq 0}$.

We now state the application of our results. For two quasi-normed spaces X and Y, we denote by $B(X,Y)$ the set of all bounded linear maps from X to Y.

Theorem 2.7. *Suppose that every σ-algebra $\mathcal{F}_n$ is generated by countable atoms. Furthermore, assume that $\{\mathcal{F}_n\}_{n\geq 0}$ is regular with* (2.11) *and* (2.12)*. Let $v = (v_n)_{n\geq 0}$ be the sequence of functions defined in* (2.9)*. Let $s \in \mathbb{R}$, $p \in (0,\infty)$, $q \in (0,\infty)$ and $\alpha \in (0,\infty)$. If $\alpha < 1/p$, then define p_1 by $1/p_1 = 1/p - \alpha$. Then, the following boundedness holds for the fractional integral I_α:*

$$I_\alpha \in B(F^s_{p\infty}, F^s_{p_1 q}) \quad \textit{if} \quad \alpha < 1/p, \tag{2.16}$$

$$I_\alpha \in B(F^s_{pq}, F^s_{\infty q}) \quad \textit{if} \quad \alpha = 1/p \quad \textit{and} \quad q \geq 1, \tag{2.17}$$

$$I_\alpha \in B(F^s_{p\infty}, B^{s+\alpha-1/p}_{\infty\infty}) \quad \textit{if} \quad \alpha > 1/p. \tag{2.18}$$

Theorem 2.7 is an extension of the following known fact shown in [4], [11] and [14]. Indeed, we can obtain it as a corollary of Theorem 2.7.

Corollary 2.8. *Under the assumptions in Theorem 2.7, the following boundedness holds for the fractional integral I_α:*

$$I_\alpha \in B(H_p^S, H_{p_1}^S) \quad \text{if} \quad \alpha < 1/p, \tag{2.19}$$

$$I_\alpha \in B(H_p^S, \mathrm{BMO}_2^{S-}) \quad \text{if} \quad \alpha = 1/p, \tag{2.20}$$

$$I_\alpha \in B(H_p^S, \Lambda_1^-(\alpha - 1/p)) \quad \text{if} \quad \alpha > 1/p. \tag{2.21}$$

In Section 8, we give proofs of Theorem 2.7 and Corollary 2.8.

3 Some basic properties In this section, we show several basic properties of martingale Besov spaces and martingale Triebel-Lizorkin spaces.

Proposition 3.1. *Let $v \in \mathcal{V}, s \in \mathbb{R}$, $p \in [1, \infty]$ and $q \in (0, \infty]$. Then B_{pq}^s and F_{pq}^s are quasi-Banach spaces.*

Proof. Let A_{pq}^s denote either B_{pq}^s or F_{pq}^s. Let $(f^{(N)})_{N\geq 1}$ be a Cauchy sequence in A_{pq}^s. By (2.2), the sequence $(d_n f^{(N)})_{N\geq 1}$ is a Cauchy sequence in L_p for every $n \geq 0$. Let $g_n \in L_p$ be the limit function of the sequence $(d_n f^{(N)})_{N\geq 1}$. Noting that $p \geq 1$, we have $E_{n-1} g_n = 0$ for all $n \geq 1$. Therefore, the sequence $f = (f_n)_{n\geq 0}$ defined by $f_n = \sum_{k=0}^n g_k$ for $n \geq 0$ is a martingale. Hence, by a standard argument, we have that $(f^{(N)})_{N\geq 1}$ converges to f in A_{pq}^s. We obtain the desired conclusion. □

Proposition 3.2. *Let $v \in \mathcal{V}$, $s \in \mathbb{R}$, $p \in (0, \infty]$ and $q, q_1, q_2 \in (0, \infty]$.*

(1) If $p < \infty$ and $q_1 \leq q_2$, then

$$B_{pq_1}^s \hookrightarrow B_{pq_2}^s \quad \text{and} \quad F_{pq_1}^s \hookrightarrow F_{pq_2}^s. \tag{3.1}$$

(2) For each $s \in \mathbb{R}$, $p \in (0, \infty]$ and $q \in (0, \infty]$,

$$B_{p\min(p,q)}^s \hookrightarrow F_{pq}^s \hookrightarrow B_{p\max(p,q)}^s. \tag{3.2}$$

Proof. (3.1) is a consequence of the known fact $\|(a_n)_{n\geq 0}\|_{\ell_{q_2}} \leq \|(a_n)_{n\geq 0}\|_{\ell_{q_1}}$ for any sequence $(a_n)_{n\geq 0}$.

To show (3.2), we first note that

$$B_{pp}^s = F_{pp}^s, \tag{3.3}$$

which is derived from the definition. Furthermore, we recall the following fact for any sequence of measurable functions $(g_n)_{n\geq 0}$, which is proved by the use of Minkowski's inequality:

$$\big\|(\|g_n\|_{L_p})_{n\geq 0}\big\|_{\ell_q} \leq \big\|\|(g_n)_{n\geq 0}\|_{\ell_q}\big\|_{L_p} \quad \text{if} \quad p \leq q, \tag{3.4}$$

$$\big\|(\|g_n\|_{L_p})_{n\geq 0}\big\|_{\ell_q} \geq \big\|\|(g_n)_{n\geq 0}\|_{\ell_q}\big\|_{L_p} \quad \text{if} \quad p \geq q. \tag{3.5}$$

We now show (3.2) in case $p < \infty$. If $p \leq q$, then, using (3.3), (3.1) and (3.4), we have (3.2) as follows:

$$B_{p\min(p,q)}^s = B_{pp}^s = F_{pp}^s \hookrightarrow F_{pq}^s \hookrightarrow B_{pq}^s = B_{p\max(p,q)}^s.$$

Similarly, if $p \geq q$, then we have (3.2) as follows:

$$B^s_{p\min(p,q)} = B^s_{pq} \hookrightarrow F^s_{pq} \hookrightarrow F^s_{pp} = B^s_{pp} = B^s_{p\max(p,q)}.$$

Thus, we obtain (3.2) in case $p < \infty$.

If $p = \infty$ and $q < \infty$, then we have $\|f\|_{F^s_{\infty q}} \leq \|f\|_{B^s_{\infty q}}$ by the following inequality:

$$\sum_{k=n}^{\infty} |v^s_{k-1} d_k f|^q \leq \sum_{n=0}^{\infty} \|v^s_{n-1} d_n f\|^q_{L_\infty} = \|f\|^q_{B^s_{\infty q}}.$$

We also have $\|f\|_{B^s_{\infty\infty}} \leq \|f\|_{F^s_{\infty q}}$ for $q < \infty$ by the following inequality:

$$|v^s_{n-1} d_n f|^q = E_n[|v^s_{n-1} d_n f|^q] \leq E_n\left[\sum_{k=n}^{\infty} |v^s_{k-1} d_k f|^q\right] \leq \|f\|^q_{F^s_{\infty q}}.$$

The proof is completed. □

Concerning Theorem 2.4, we show the following proposition.

Proposition 3.3. *Let $v \in \mathcal{V}$, $s > 0$, $p \in [1, \infty]$, and $q \in (0, \infty]$. Then,*

$$B^s_{pq} \hookrightarrow L_p \quad \textit{and} \quad F^s_{pq} \hookrightarrow L_p \tag{3.6}$$

under the identification of $(f_n)_{n\geq 0} \in A^s_{pq}$ with its limit function, where A^s_{pq} denote either B^s_{pq} or F^s_{pq}.

Proof. By Proposition 3.2, we only have to show that

$$B^s_{p\infty} \hookrightarrow L_p.$$

Let $f = (f_n)_{n\geq 0} \in B^s_{p\infty}$. By (2.2), we have

$$\sum_{n=1}^{\infty} \|d_n f\|_{L_p} \leq \sum_{n=1}^{\infty} \delta_1^{-s(n-1)} \|v^s_{n-1} d_n f\|_{L_p} \leq \sum_{n=1}^{\infty} \delta_1^{-s(n-1)} \|f\|_{B^s_{p\infty}} < \infty.$$

Thus, $(f_n)_{n\geq 0} = (\sum_{k=0}^{n} d_k f)_{n\geq 0}$ converges in L_p. Denote the limit function by the same symbol f. Then we have $E_n f = f_n$ and $\|f\|_{L_p} \leq 2(1 - \delta_1^{-s})^{-1} \|f\|_{B^s_{p\infty}}$. The proof is completed. □

4 Proof of Theorem 2.2.

In this section, we prove Theorem 2.2. To do this, we need two lemmas.

Lemma 4.1. *Let $p \in [1, \infty]$ and $q \in [1, \infty]$. Let $(f_n)_{n\geq 0}$ be a sequence of integrable functions. If $1 \leq q \leq p < \infty$ or $1 < p \leq q \leq \infty$, then, there exists a constant $C_{p,q}$ depending only on p and q such that*

$$\left\| \|(E_n f_n)_{n\geq 0}\|_{\ell_q} \right\|_{L_p} \leq C_{p,q} \left\| \|(f_n)_{n\geq 0}\|_{\ell_q} \right\|_{L_p}.$$

For the proof of Lemma 4.1, we refer to [1, Theorem 3.1].

Remark 4.1. Since $\|E_n f_n\|_{L_p} \leq \|f_n\|_{L_p}$ by Jensen's inequality for E_n, it is clear that

$$\left\| (\|E_n f_n\|_{L_p})_{n\geq 0} \right\|_{\ell_q} \leq \left\| (\|f_n\|_{L_p})_{n\geq 0} \right\|_{\ell_q}$$

for $p \in [1, \infty]$ and $q \in (0, \infty]$.

Lemma 4.2. *Let $q \in [1,\infty)$. Denote by q' the conjugate exponent of q. Then, there exists a positive constant C depending only on q such that the following (1) and (2) hold:*

(1) If $g \in F^0_{\infty q'}$, then the infinite sum $\sum_{n=0}^\infty E[d_n\overline{g}d_nf]$ converges for every $f \in F^0_{1q}$. Moreover,

$$\left|\sum_{n=0}^\infty E[d_n\overline{g}d_nf]\right| \le C\|g\|_{F^0_{\infty q'}}\|f\|_{F^0_{1q}} \quad (f \in F^0_{1q}).$$

(2) Conversely, for each $\Phi \in (F^0_{1q})'$, there exists $g \in F^0_{\infty q'}$ such that

$$\Phi(f) = \sum_{n=0}^\infty E[d_n\overline{g}d_nf] \quad (f \in F^0_{1q})$$

and that $\|g\|_{F^0_{\infty q'}} \le C\|\Phi\|_{(F^0_{1q})'}$.

For the proof of Lemma 4.2, we refer to [17, Theorem 17 and Corollary 10].

Remark 4.2. We remark on the difference between our convention and the one in [17]. In [17, Corollary 10], it was shown the duality between $H_1^{S_q} = \{(f_n)_{n\ge0} \in F^0_{1q} : f_0 = 0\}$ and $\mathcal{BMO}^-_{q'} = \{(f_n)_{n\ge0} \in F^0_{\infty q'} : f_0 = 0\}$. For this difference, we note that

$$\begin{aligned} f \in F^0_{1q} \quad &\text{if and only if} \quad f - f_0 \in H_1^{S_q} \quad \text{and} \quad f_0 \in L_1, \\ f \in F^0_{\infty q'} \quad &\text{if and only if} \quad f - f_0 \in \mathcal{BMO}^-_{q'} \quad \text{and} \quad f_0 \in L_\infty \end{aligned}$$

with

$$\|f\|_{F^0_{1q}} \sim \|f - f_0\|_{F^0_{1q}} + \|f_0\|_{L_1}, \quad \|f\|_{F^0_{\infty q'}} \sim \|f - f_0\|_{F^0_{\infty q'}} + \|f_0\|_{L_\infty},$$

where $f = (f_n)_{n\ge0}$ and $f - f_0 = (f_n - f_0)_{n\ge0}$.

Proof of Theorem 2.2. The proof below is a modification of the one given in [17, Theorems 14-17], but, to include the Besov space case, we give a proof.

We first prove the case where $p \in (1,\infty)$. Let $g \in A^{-s}_{p'q'}$ and $f \in A^s_{pq}$. If $A^s_{pq} = F^s_{pq}$, then using Hölder's inequality, we have

$$\begin{aligned} \sum_{n=0}^\infty E\left[|d_ngd_nf|\right] &= E\left[\sum_{n=0}^\infty |v^{-s}_{n-1}d_ngv^s_{n-1}d_nf|\right] \\ &\le E\left[\|(v^{-s}_{n-1}d_ng)_{n\ge0}\|_{\ell_{q'}}\|(v^s_{n-1}d_nf)_{n\ge0}\|_{\ell_q}\right] \\ &\le \|g\|_{F^{-s}_{p'q'}}\|f\|_{F^s_{pq}}. \end{aligned} \tag{4.1}$$

If $A^s_{pq} = B^s_{pq}$, then similarly we have

$$\begin{aligned} \sum_{n=0}^\infty E\left[|d_ngd_nf|\right] &= \sum_{n=0}^\infty E\left[|v^{-s}_{n-1}d_ngv^s_{n-1}d_nf|\right] \\ &\le \sum_{n=0}^\infty \|v^{-s}_{n-1}d_ng\|_{L_{p'}}\|v^s_{n-1}d_nf\|_{L_p} \\ &\le \|g\|_{B^{-s}_{p'q'}}\|f\|_{B^s_{pq}}. \end{aligned}$$

Therefore, we have obtained (1) in case $p \in (1,\infty)$.

We next show (2) in case $p \in (1,\infty)$. Define A_{pq} by

$$A_{pq} = \begin{cases} \ell_q(L_p) & \text{if} \quad A^s_{pq} = B^s_{pq}, \\ L_p(\ell_q) & \text{if} \quad A^s_{pq} = F^s_{pq}. \end{cases}$$

Let $\Phi \in (A^s_{pq})'$ and let $u_n = v_n^{-1}$ for $n \geq 0$. By Theorem 2.1, the functional $f \mapsto \Phi \circ I^u_s(f)$ on A^0_{pq} is bounded. We denote by $i : A^0_{pq} \to A_{pq}$ the isometric embedding defined by $i(f) = (d_n f)_{n\geq 0}$ $(f \in A^0_{pq})$. Using Hahn-Banach's theorem, we take $\Psi \in (A_{pq})'$ such that $\|\Psi\|_{(A_{pq})'} = \|\Phi\|_{(A^s_{pq})'}$ and that $\Psi \circ i = \Phi \circ I^u_s$ on A^0_{pq}. Furthermore, using the fact that $(A_{pq})'$ is isometric to $A_{p'q'}$, we take $g = (g_n)_{n\geq 0} \in A_{p'q'}$ such that

$$\|g\|_{A_{p'q'}} = \|\Phi\|_{(A^s_{pq})'} \quad \text{and that} \quad \Phi(f) = \sum_{n=0}^{\infty} E[\overline{g_n} v^s_{n-1} d_n f] \quad \text{for} \quad f \in A^s_{pq}.$$

Then define $h = (h_n)_{n\geq 0}$ by

$$h_n = \sum_{k=0}^{n} v^s_{k-1}(E_k g_k - E_{k-1} g_k)$$

with convention $v_{-1} = v_0$ and $E_{-1} g_0 = 0$. It is clear that $h = (h_n)_{n\geq 0}$ is a martingale. If $A^s_{pq} = B^s_{pq}$, then by Remark 4.1,

$$\begin{aligned} \left\| (\|v^{-s}_{n-1} d_n h\|_{L_{p'}})_{n\geq 0} \right\|_{\ell_{q'}} &= \left\| (\|E_n g_n - E_{n-1} g_n\|_{L_{p'}})_{n\geq 0} \right\|_{\ell_{q'}} \\ &\leq \left\| (\|E_n g_n\|_{L_{p'}})_{n\geq 0} \right\|_{\ell_{q'}} + \left\| (\|E_{n-1} g_n\|_{L_{p'}})_{n\geq 0} \right\|_{\ell_{q'}} \\ &\leq 2 \left\| (\|g_n\|_{L_{p'}})_{n\geq 0} \right\|_{\ell_{q'}} = 2\|\Phi\|_{(B^s_{pq})'}, \end{aligned}$$

that is, we have $h \in B^{-s}_{p'q'}$ with $\|h\|_{B^{-s}_{p'q'}} \leq 2\|\Phi\|_{(B^s_{pq})'}$.

Similarly, if $A^s_{pq} = F^s_{pq}$, then by Lemma 4.1,

$$\begin{aligned} \left\| \|(v^{-s}_{n-1} d_n h)_{n\geq 0}\|_{\ell_{q'}} \right\|_{L_{p'}} &= \left\| \|(E_n g_n - E_{n-1} g_n)_{n\geq 0}\|_{\ell_{q'}} \right\|_{L_{p'}} \\ &\leq \left\| \|(E_n g_n)_{n\geq 0}\|_{\ell_{q'}} \right\|_{L_{p'}} + \left\| \|(E_{n-1} g_n)_{n\geq 0}\|_{\ell_{q'}} \right\|_{L_{p'}} \\ &\leq 2C_{p',q'} \left\| \|(g_n)_{n\geq 0}\|_{\ell_{q'}} \right\|_{L_{p'}} = 2C_{p',q'} \|\Phi\|_{(F^s_{pq})'}, \end{aligned}$$

that is, we have $h \in F^{-s}_{p'q'}$ with $\|h\|_{F^{-s}_{p'q'}} \leq 2C_{p',q'} \|\Phi\|_{(F^s_{pq})'}$.

Let $f \in A^s_{pq}$. Then, by the formal self-adjointness of E_n, we have

$$\begin{aligned} \sum_{n=0}^{\infty} E[d_n \overline{h} d_n f] &= \sum_{n=0}^{\infty} (E[E_n(v^s_{n-1}\overline{g_n}) d_n f] - E[E_{n-1}(v^s_{n-1}\overline{g_n}) d_n f]) \\ &= \sum_{n=0}^{\infty} (E[\overline{g_n} v^s_{n-1} E_n(d_n f)] - E[\overline{g_n} v^s_{n-1} E_{n-1}(d_n f)]) \\ &= \sum_{n=0}^{\infty} E[\overline{g_n} v^s_{n-1} d_n f] = \Phi(f). \end{aligned}$$

Hence, we have the desired conclusion for the case where $p \in (1, \infty)$.

For the case where $A^s_{pq} = B^s_{pq}$ with $p = 1$, we can obtain the desired conclusion by the same way as in the case where $p \in (1, \infty)$.

We now give a proof for the case where $A^s_{pq} = F^s_{pq}$ with $p = 1$. Let $g \in F^{-s}_{\infty q'}$ and let $u = (v_n^{-1})_{n \geq 0}$. Then, by (1) of Lemma 4.2 and Theorem 2.1, we obtain that the infinite sum

$$\sum_{n=0}^{\infty} E[d_n\overline{g}d_n f] = \sum_{n=0}^{\infty} E[d_n(I^u_s\overline{g})d_n(I^u_{-s}f)]$$

converges and that

$$\left|\sum_{n=0}^{\infty} E[d_n\overline{g}d_n f]\right| \leq C\|I^u_s g\|_{F^0_{\infty q'}} \|I^u_{-s}f\|_{F^0_{1q}} = C\|g\|_{F^{-s}_{\infty q'}} \|f\|_{F^s_{1q}}$$

for $f \in F^s_{1q}$.

We next show (2). Let $\Phi \in (F^s_{1q})'$. By Theorem 2.1, $\Phi \circ I^u_s$ belongs to $(F^0_{1q})'$. Using (2) of Lemma 4.2, we take $g \in F^0_{\infty q'}$ such that

$$\Phi \circ I^u_s(\tilde{f}) = \sum_{n=0}^{\infty} E[d_n\overline{g}d_n\tilde{f}] \quad (\tilde{f} \in F^0_{1q}) \tag{4.2}$$

and that $\|g\|_{F^0_{\infty q'}} \leq C\|\Phi \circ I^u_s\|_{(F^0_{1q})'} = C\|\Phi\|_{(F^s_{1q})'}$. Let $f \in F^s_{1q}$. We put $h = I^u_{-s}g$ and $\tilde{f} = I^u_{-s}f$ in (4.2). Then, we have $\|h\|_{F^{-s}_{\infty q'}} \leq C\|\Phi\|_{(F^s_{1q})'}$ and

$$\Phi(f) = \sum_{n=0}^{\infty} E[d_n(I^u_s\overline{h})d_n(I^u_{-s}f)] = \sum_{n=0}^{\infty} E[d_n\overline{h}d_n f].$$

Therefore, we have the desired conclusion for the case where $A^s_{pq} = F^s_{pq}$ with $p = 1$. The proof is completed.

□

5 Proof of Theorem 2.3. In this section, we give a proof of Theorem 2.3. For the proof, we need some lemmas.

For $0 \leq x \leq 1$ and $t \in \mathbb{R}$, let $\mu_0(z, t)$, $\mu_1(z, t)$ be the Poisson kernel on $S = \{0 \leq \mathrm{Re}z \leq 1\}$, that is,

$$\mu_j(x + iy, t) = \frac{e^{-\pi(t-y)} \sin \pi x}{\sin^2 \pi x + (\cos \pi x - e^{ij\pi - \pi(t-y)})^2}, \quad j = 0, 1.$$

Lemma 5.1. *Let (A_0, A_1) be a compatible couple of Banach spaces. Let $f \in \mathcal{F}(A_0, A_1)$. Then, for $0 < \theta < 1$,*

$$\begin{aligned}&\|f(\theta)\|_{[A_0,A_1]_\theta}\\ &\leq \left(\frac{1}{1-\theta}\int_{-\infty}^{\infty} \|f(it)\|_{A_0}\mu_0(\theta, t)dt\right)^{1-\theta} \left(\frac{1}{\theta}\int_{-\infty}^{\infty} \|f(1 + it)\|_{A_1}\mu_1(\theta, t)dt\right)^{\theta}.\end{aligned}$$

For the proof of Lemma 5.1, see [2, Lemma 4.3.2].

Lemma 5.2. *Let f be a non-negative bounded measurable function on Ω. Let $a, b \in \mathbb{R}$ and let $\rho(z) = az + b$, $z \in \mathbb{C}$. Suppose that either $\operatorname{ess\,inf} f > 0$ or both a and b are positive. Then, the map $F : S \to L_\infty$ defined by $F(z) = f^{\rho(z)}$ is holomorphic on S_0.*

Proof. We first give the proof for the case where $\operatorname{ess\,inf} f > 0$. Since $f^{\rho(z)} = f^b(f^a)^z$, we only have to prove in case where $\rho(z) = z$. Let $z \in S_0$ and let $h \in \mathbb{C} \setminus \{0\}$ such that $z + h \in S_0$. By the fundamental theorem of calculus, we have

$$\left\| \frac{f^{z+h} - f^z}{h} - f^z \log f \right\|_{L_\infty} = \left\| f^z (\log f)^2 \frac{1}{h} \int_0^h \left(\int_0^t f^s \, ds \right) dt \right\|_{L_\infty} \leq C|h|,$$

where $C = (1 + \|f\|_{L_\infty}^2)\{\log(\|f^{-1}\|_{L_\infty} + \|f\|_{L_\infty})\}^2$. Therefore, we have $F'(z) = f^z \log f$ in L_∞.

We next give the proof for the case where both a and b are positive. Since $f^{\rho(z)} = f^{a(z+b/a)}$, we only have to show in case where $\rho(z) = z + c$ with $c > 0$. Then, as above, we have

$$\begin{aligned} \left\| \frac{f^{z+c+h} - f^{z+c}}{h} - f^{z+c} \log f \right\|_{L_\infty} &= \left\| (f^{c/2} \log f)^2 \frac{f^z}{h} \int_0^h \left(\int_0^t f^s \, ds \right) dt \right\|_{L_\infty} \\ &\leq \sup_{0 \leq x \leq \|f\|_{L_\infty}} (x^{c/2} \log x)^2 (1 + \|f\|_{L_\infty}^2)|h|. \end{aligned}$$

We have the desired conclusion. □

Lemma 5.3. *Let $(c_n)_{n=1}^\infty$ be a sequence of positive numbers and $\alpha > 0$. Then,*

$$\sum_{n=1}^\infty c_n \left(\sum_{k=1}^n c_k \right)^{\alpha - 1} \leq \frac{1}{\min(\alpha, 1)} \left(\sum_{n=1}^\infty c_n \right)^\alpha.$$

For the proof of Lemma 5.3, see [15, Section 2.4.6] and [13, Lemma 2.17].

In the next lemma, we give a dense subspace of B_{pq}^s and F_{pq}^s. Let $\mathcal{M}_b$ be the set of all martingales $(f_n)_{n \geq 0}$ which satisfies $\sup_{n \geq 0} \|f_n\|_{L_\infty} < \infty$.

Then define

$$\mathcal{T} = \{(f_n)_{n \geq 0} \in \mathcal{M}_b : \text{there exists } N \geq 0 \text{ such that } f_n = f_N \text{ for all } n \geq N\}.$$

Lemma 5.4. *Let $v \in \mathcal{V}$. Let $p \in [1, \infty]$, $q \in [1, \infty)$ and $s \in \mathbb{R}$. Then, $\mathcal{T}$ is dense in B_{pq}^s. Moreover, if $p < \infty$, then $\mathcal{T}$ is also dense in F_{pq}^s.*

Proof. We first show that $\mathcal{T}$ is dense in F_{pq}^s if $p < \infty$. Let $f = (f_n)_{n \geq 0} \in F_{pq}^s$. For $N \geq 0$, let $f^N = (f_{n \wedge N})_{n \geq 0}$ where $n \wedge N = \min(n, N)$. Then,

$$\left(\sum_{n=0}^\infty |v_{n-1}^s d_n(f - f^N)|^q \right)^{1/q} = \left(\sum_{n=N+1}^\infty |v_{n-1}^s d_n f|^q \right)^{1/q}. \tag{5.1}$$

By Lebesgue's convergence theorem, we have

$$\lim_{N \to \infty} \|f - f^N\|_{F_{pq}^s} = 0.$$

Therefore, to obtain the conclusion, we only have to show that each f^N is approximated by some sequences in $\mathcal{T}$.

For $R > 0$, let $g(N,R) = (E_n[f_N \chi_{\{|f_N| \le R\}}])_{n \ge 0}$. It is clear that $g(N,R) \in \mathcal{T}$. Noting that $v_n \le \delta_2^n$ for some $\delta_2 > 1$, we have

$$\begin{aligned}
(5.2) \qquad \left(\sum_{n=0}^{\infty} |v_{n-1}^s d_n(f^N - g(N,R))|^q \right)^{1/q} &= \left(\sum_{n=0}^{N} |v_{n-1}^s d_n(f^N - g(N,R))|^q \right)^{1/q} \\
&\le \delta_2^{Ns^+} \left(\sum_{n=0}^{N} |d_n(f^N - g(N,R))|^q \right)^{1/q} \\
&\le \delta_2^{Ns^+} \sum_{n=0}^{N} |d_n(f^N - g(N,R))| \\
&\le 2\delta_2^{Ns^+} \sum_{n=0}^{N} E_n |f_N \chi_{\{|f_N| > R\}}|
\end{aligned}$$

where $s^+ = \max(s,0)$. By (5.2), we have

$$\begin{aligned}
\|f^N - g(N,R)\|_{F_{pq}^s} &\le 2\delta_2^{Ns^+} \sum_{n=0}^{N} \big\| E_n |f_N \chi_{\{|f_N| > R\}}| \big\|_{L_p} \\
&\le 2\delta_2^{Ns^+} \sum_{n=0}^{N} \|f_N \chi_{\{|f_N| > R\}}\|_{L_p}.
\end{aligned}$$

Since $f_N \in L_p$, we have

$$\lim_{R \to \infty} \|f^N - g(N,R)\|_{F_{pq}^s} = 0.$$

Therefore, we have that $\mathcal{T}$ is dense in F_{pq}^s.

We next show that $\mathcal{T}$ is dense in B_{pq}^s. As in (5.1), we have

$$(5.3) \qquad \left(\sum_{n=0}^{\infty} \|v_{n-1}^s d_n(f - f^N)\|_{L_p}^q \right)^{1/q} = \left(\sum_{n=N+1}^{\infty} \|v_{n-1}^s d_n f\|_{L_p}^q \right)^{1/q}.$$

Note that (5.3) holds even if $p = \infty$. Then we have

$$\lim_{N \to \infty} \|f - f^N\|_{B_{pq}^s} = 0.$$

Similarly, as in (5.2), we have

$$\left(\sum_{n=0}^{\infty} \|v_{n-1}^s d_n(f^N - g(N,R))\|_{L_p}^q \right)^{1/q} \le 2\delta_2^{Ns^+} \sum_{n=0}^{N} \|f_N \chi_{\{|f_N| > R\}}\|_{L_p}.$$

Hence, we obtain

$$\lim_{R \to \infty} \|f^N - g(N,R)\|_{B_{pq}^s} = 0.$$

Therefore, we have the desired conclusion. □

The following is the key lemma for the proof of Theorem 2.3.

Lemma 5.5. *Let $v \in \mathcal{V}$, $\theta \in (0,1)$, $s_0, s_1 \in \mathbb{R}$ and $p_0, p_1, q_0, q_1 \in [1,\infty]$. Define s, p and q by (2.5) with convention $1/\infty = 0$. Then, there exists a positive constant C_1 depending only on p_0, p_1, q_0, q_1 and θ such that the following (i) and (ii) hold.*

(i) For each $f \in \mathcal{T}$, there exists $H \in \mathcal{F}(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})$ such that

$$\|H\|_{\mathcal{F}(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})} \le C_1 \|f\|_{B^s_{pq}}, \quad H(\theta) = f \tag{5.4}$$

and that

$$H(z) \in \mathcal{T} \quad \text{for all} \quad z \in S \quad \text{and} \quad \sup_{n\ge 0}\sup_{z\in S} \|d_n H(z)\|_{L_\infty} < \infty. \tag{5.5}$$

(ii) Assume that $1 < p_0, p_1 < \infty$. Then, for each $f \in \mathcal{T}$, there exists $H \in \mathcal{F}(F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})$ such that

$$\|H\|_{\mathcal{F}(F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})} \le C_1 \|f\|_{F^s_{pq}}, \quad H(\theta) = f \tag{5.6}$$

and that (5.5).

Proof. We first show (i). To do this, we introduce functions ρ_1, ρ_2 and ρ_3 defined on $\mathbb{C}$ by

$$\begin{aligned}
\rho_1(z) &= \left(s\frac{p}{p_0} - s_0\right)(1-z) + \left(s\frac{p}{p_1} - s_1\right)z,\\
\rho_2(z) &= \frac{p}{p_0}(1-z) + \frac{p}{p_1}z,\\
\rho_3(z) &= \left(\frac{q}{q_0} - \frac{p}{p_0}\right)(1-z) + \left(\frac{q}{q_1} - \frac{p}{p_1}\right)z,
\end{aligned}$$

with convention $1/\infty = 0$ and $\infty/\infty = 1$.

Furthermore, define $\mathrm{sgn} : \mathbb{C} \to \mathbb{C}$ by

$$\mathrm{sgn}(z) = \begin{cases} z/|z| & \text{if} \quad z \neq 0,\\ 0 & \text{if} \quad z = 0. \end{cases}$$

Let $f \in \mathcal{T}$ such that $\|f\|_{B^s_{pq}} = 1$. For $n \ge 0$, $z \in S$ and $\omega \in \Omega$, define $g_n(z, \omega)$ by

$$g_n(z,\omega) = v_{n-1}(\omega)^{\rho_1(z)} |d_n f(\omega)|^{\rho_2(z)} \mathrm{sgn}(d_n f(\omega)) \|v^s_{n-1} d_n f\|^{\rho_3(z)}_{L_p}.$$

Then, define $H(z) = (H_n(z))_{n\ge 0}$ by

$$h_n(z) = g_n(z) - E_{n-1}[g_n(z)], \quad H_n(z) = \sum_{k=0}^{n} h_k(z) \tag{5.7}$$

with convention $E_{-1}[g_0(z)] = 0$. $H(z)$ is a martingale for every $z \in S$. Noting that $\rho_1(\theta) = \rho_3(\theta) = 0$ and $\rho_2(\theta) = 1$, we have $g_n(\theta, \omega) = d_n f(\omega)$ and then have

$$H(\theta) = f. \tag{5.8}$$

By Lemma 5.2, we obtain that g_n is an L_∞-valued holomorphic function on S_0. Moreover, since $f \in \mathcal{T}$ and $\mathrm{Re}\rho_j$ $(j = 1, 2, 3)$ is bounded on S, we have

$$H(z) \in \mathcal{T} \quad \text{for all} \quad z \in S. \tag{5.9}$$

Thus, H is a $(B^{s_0}_{p_0q_0} + B^{s_1}_{p_1q_1})$-valued holomorphic function on S_0 with

$$\sup_{n\ge 0}\sup_{z\in S} \|d_n H(z)\|_{L_\infty} \le 2 \sup_{n\ge 0}\sup_{z\in S} \|g_n(z)\|_{L_\infty} < \infty. \tag{5.10}$$

For $\delta > 0$, let $H_\delta(z) = e^{\delta(z-\theta)^2} H(z)$. Then, H_δ also satisfies $H_\delta(\theta) = f$ and (5.5). We now show that H_δ belongs to $\mathcal{F}(B^{s_0}_{p_0 q_0}, B^{s_1}_{p_1 q_1})$. For $j \in \{0, 1\}$, noting that

$$\mathrm{Re}\rho_1(j+it) = s\frac{p}{p_j} - s_j, \quad \mathrm{Re}\rho_2(j+it) = \frac{p}{p_j}, \quad \mathrm{Re}\rho_3(j+it) = \frac{q}{q_j} - \frac{p}{p_j},$$

we have

$$|v^{s_j}_{n-1} g_n(j+it)|^{p_j} = |v^s_{n-1} d_n f|^p (\|v^s_{n-1} d_n f\|^p_{L_p})^{(qp_j/pq_j)-1}.$$

Hence, we have

$$\|v^{s_j}_{n-1} g_n(j+it)\|_{L_{p_j}} = (\|v^s_{n-1} d_n f\|_{L_p})^{q/q_j}.$$

Therefore,

$$\begin{aligned} &\|H(j+it)\|_{B^{s_j}_{p_j q_j}} \\ &\le \left\|(\|v^{s_j}_{n-1} g_n(j+it)\|_{L_{p_j}})_{n\ge 0}\right\|_{\ell_{q_j}} + \left\|(\|v^{s_j}_{n-1} E_{n-1}[g_n(j+it)]\|_{L_{p_j}})_{n\ge 0}\right\|_{\ell_{q_j}} \\ &\le 2 \left\|(\|v^s_{n-1} d_n f\|_{L_p})_{n\ge 0}\right\|^{1/q_j}_{\ell_q} = 2. \end{aligned} \tag{5.11}$$

By (5.8) and (5.11), we obtain $H_\delta \in \mathcal{F}(B^{s_0}_{p_0 q_0}, B^{s_1}_{p_1 q_1})$ with

$$\|H_\delta\|_{\mathcal{F}(B^{s_0}_{p_0 q_0}, B^{s_1}_{p_1 q_1})} \le 2\max(e^{\delta\theta^2}, e^{\delta(1-\theta)^2}), \quad H_\delta(\theta) = f. \tag{5.12}$$

Thus, by (5.12), (5.9) and (5.10), we obtain (i).

We now show (ii). In this case, we define ρ_1, ρ_2 and ρ_3 by

$$\begin{aligned} \rho_1(z) &= \left(s\frac{q}{q_0} - s_0\right)(1-z) + \left(s\frac{q}{q_1} - s_1\right) z, \\ \rho_2(z) &= \frac{q}{q_0}(1-z) + \frac{q}{q_1} z, \\ \rho_3(z) &= \left(\frac{p}{p_0} - \frac{q}{q_0}\right)(1-z) + \left(\frac{p}{p_1} - \frac{q}{q_1}\right) z. \end{aligned}$$

Let $f \in \mathcal{T}$ such that $\|f\|_{F^s_{pq}} = 1$. For $n \ge 0$, $z \in S$ and $\omega \in \Omega$, define $g_n(z,\omega)$ by

$$g_n(z,\omega) = v_{n-1}(\omega)^{\rho_1(z)} |d_n f(\omega)|^{\rho_2(z)} \mathrm{sgn}(d_n f(\omega)) \, (1 + G_n(\omega))^{\rho_3(z)}$$

where $G_n(\omega)$ denotes

$$G_n(\omega) = \|(v_{k-1}(\omega)^s d_k f(\omega))_{0\le k\le n}\|_{\ell_q}.$$

Then, by the same way as in (5.7), we obtain martingales $H(z) = (H_n(z))_{n\ge 0}$ such that $H(\theta) = f$, $H(z) \in \mathcal{T}$ for all $z \in S$ and that $z \mapsto H(z)$ is holomorphic from S_0 into $F^{s_0}_{p_0 q_0} + F^{s_1}_{p_1 q_1}$. Furthermore, $H_\delta(z) = e^{\delta(z-\theta)^2} H(z)$ satisfies $H_\delta(\theta) = f$ and (5.5) for every $\delta > 0$.

We now show that H_δ belongs to $\mathcal{F}(F^{s_0}_{p_0 q_0}, F^{s_1}_{p_1 q_1})$ for every $\delta > 0$. We first show it in case where $q_0, q_1 < \infty$. Since $\rho_3(\theta) = 0$, we have $\rho_3(0)\rho_3(1) < 0$. We may assume that $\rho_3(0) < 0$ because the other case is proved by the same way. Note that

$$\mathrm{Re}\rho_1(j+it) = s\frac{q}{q_j} - s_j, \quad \mathrm{Re}\rho_2(j+it) = \frac{q}{q_j}, \quad \mathrm{Re}\rho_3(j+it) = \frac{p}{p_j} - \frac{q}{q_j}$$

for $j = 0, 1$. Then, by the assumption $\rho_3(0) < 0$, we have

$$\begin{aligned}(v_{n-1}^{s_0}|g_n(it)|)^{q_0} &= (v_{n-1}^{s}|d_n f|)^{q}(1+G_n)^{\rho_3(0)q_0}\\ &\le (v_{n-1}^{s}|d_n f|)^{q} G_n^{\rho_3(0)q_0}\\ &= (v_{n-1}^{s}|d_n f|)^{q}\left(\sum_{k=0}^{n}(v_{k-1}^{s}|d_k f|)^{q}\right)^{(pq_0/qp_0)-1}\end{aligned}$$

and

$$\begin{aligned}(v_{n-1}^{s_1}|g_n(1+it)|)^{q_1} &= (v_{n-1}^{s}|d_n f|)^{q}(1+G_n)^{\rho_3(1)q_1}\\ &\le C(v_{n-1}^{s}|d_n f|)^{q}\{1+G_n^{\rho_3(1)q_1}\}\\ &= C(v_{n-1}^{s}|d_n f|)^{q}\left\{1+\left(\sum_{k=0}^{n}(v_{k-1}^{s}|d_k f|)^{q}\right)^{(pq_1/qp_1)-1}\right\},\end{aligned}$$

where C is a positive constant depending only on $\rho_3(1)q_1$. Using Lemma 5.3 and the assumption $\rho_3(0) < 0$, which is equivalent to $pq_0 < qp_0$, we have

$$\begin{aligned}\left(\sum_{n=0}^{\infty}(v_{n-1}^{s_0}|g_n(it)|)^{q_0}\right)^{1/q_0} &\le \left(\sum_{n=0}^{\infty}(v_{n-1}^{s}|d_n f|)^{q}\left(\sum_{k=0}^{n}(v_{k-1}^{s}|d_k f|)^{q}\right)^{(pq_0/qp_0)-1}\right)^{1/q_0}\\ &\le \frac{1}{\min((pq_0/qp_0),1)}\left\{\left(\sum_{n=0}^{\infty}(v_{n-1}^{s}|d_n f|)^{q}\right)^{1/q}\right\}^{p/p_0}\\ &= \frac{qp_0}{pq_0}\left\{\left(\sum_{n=0}^{\infty}(v_{n-1}^{s}|d_n f|)^{q}\right)^{1/q}\right\}^{p/p_0}.\end{aligned}$$

Similarly, we have

$$\begin{aligned}&\left(\sum_{n=0}^{\infty}(v_{n-1}^{s_1}|g_n(1+it)|)^{q_1}\right)^{1/q_1}\\ &\le C^{1/q_1}\left(\sum_{n=0}^{\infty}(v_{n-1}^{s}|d_n f|)^{q}\right)^{1/q_1} + C^{1/q_1}\left\{\left(\sum_{n=0}^{\infty}(v_{n-1}^{s}|d_n f|)^{q}\right)^{1/q}\right\}^{p/p_1}.\end{aligned} \tag{5.13}$$

Since $\|f\|_{F_{pq}^s} = 1$, we have

$$\left\|\left\{\left(\sum_{n=0}^{\infty}(v_{n-1}^{s}|d_n f|)^{q}\right)^{1/q}\right\}^{p/p_j}\right\|_{L_{p_j}} = \left\|\left(\sum_{n=0}^{\infty}(v_{n-1}^{s}|d_n f|)^{q}\right)^{1/q}\right\|_{L_p}^{p/p_j} = 1$$

for $j = 0, 1$. Furthermore, since the assumption $\rho_3(0) < 0$ is equivalent to $p/p_1 > q/q_1$, we have

$$
\begin{aligned}
\left\| \left(\sum_{n=0}^{\infty} (v_{n-1}^{s} |d_n f|)^{q} \right)^{1/q_1} \right\|_{L_{p_1}} &= \left\| \left\{ \left(\sum_{n=0}^{\infty} (v_{n-1}^{s} |d_n f|)^{q} \right)^{1/q} \right\}^{q/q_1} \right\|_{L_{p_1}} \\
&\le \left\| \left\{ \left(\sum_{n=0}^{\infty} (v_{n-1}^{s} |d_n f|)^{q} \right)^{1/q} \right\}^{p/p_1} \right\|_{L_{p_1}}^{qp_1/pq_1} \\
&= \left\| \left(\sum_{n=0}^{\infty} (v_{n-1}^{s} |d_n f|)^{q} \right)^{1/q} \right\|_{L_p}^{q/q_1} = 1.
\end{aligned}
$$

Hence, by Lemma 4.1, we have

(5.14)

$$
\begin{aligned}
&\|H(j+it)\|_{F_{p_j q_j}^{s_j}} \\
&= \left\| \left(\sum_{n=0}^{\infty} (v_{n-1}^{s_j} |h_n(j+it)|)^{q_j} \right)^{1/q_j} \right\|_{L_{p_j}} \\
&\le \left\| \left(\sum_{n=0}^{\infty} (v_{n-1}^{s_j} |g_n(j+it)|)^{q_j} \right)^{1/q_j} \right\|_{L_{p_j}} + \left\| \left(\sum_{n=0}^{\infty} (v_{n-1}^{s_j} |E_{n-1}[g_n(j+it)]|)^{q_j} \right)^{1/q_j} \right\|_{L_{p_j}} \\
&\le (1+C_{p_j,q_j}) \left\| \left(\sum_{n=0}^{\infty} (v_{n-1}^{s_j} |g_n(j+it)|)^{q_j} \right)^{1/q_j} \right\|_{L_{p_j}} \le C',
\end{aligned}
$$

where C' is a positive constant depending only on p_0, p_1, q_0, q_1 and θ. Therefore, we obtain $H_\delta \in \mathcal{F}(F_{p_0 q_0}^{s_0}, F_{p_1 q_1}^{s_1})$ with

$$
\|H_\delta\|_{\mathcal{F}(F_{p_0 q_0}^{s_0}, F_{p_1 q_1}^{s_1})} \le C' \max(e^{\delta\theta^2}, e^{\delta(1-\theta)^2}).
$$

Hence, we have the desired conclusion for the case where $q_0, q_1 < \infty$.

For the case where $q_0 < \infty$ and $q_1 = \infty$, we replace (5.13) by

$$
\sup_{n\ge 0} |v_{n-1}^{s_1} g_n(1+it)| = \sup_{n\ge 0} (1+G_n)^{p/p_1} \le C(1 + \|(v_{n-1}^{s} d_n f)_{n\ge 0}\|_{\ell_q}^{p/p_1})
$$

where C is a positive constant depending only on p/p_1. Furthermore, we replace (5.14) for $j=1$ by

$$
\begin{aligned}
\|H(1+it)\|_{F_{p_1 \infty}^{s_1}} &= \left\| \sup_{n\ge 0} v_{n-1}^{s_1} |h_n(1+it)| \right\|_{L_{p_1}} \\
&\le (1+C_{p_1,\infty}) \left\| \sup_{n\ge 0} v_{n-1}^{s_1} |g_n(1+it)| \right\|_{L_{p_1}} \\
&\le C(1+C_{p_1,\infty}) \left(1 + \left\| \|(v_{n-1}^{s} d_n f)_{n\ge 0}\|_{\ell_q}^{p/p_1} \right\|_{L_{p_1}} \right) \le C'.
\end{aligned}
$$

The rest of the proof is the same as in the case of $q_0, q_1 < \infty$. We have the desired conclusion for the case where $q_0 < \infty$ and $q_1 = \infty$. Similarly, we can prove the case where $q_0 = \infty$ and $q_1 < \infty$.

We now prove the case where $q_0 = q_1 = \infty$. We replace (5.13) by

$$\sup_{n\geq 0} |v_{n-1}^{s_j} g_n(j+it)| = \sup_{n\geq 0}(1+G_n)^{p/p_j} \leq C(1+\|(v_{n-1}^s d_n f)_{n\geq 0}\|_{\ell_q}^{p/p_j})$$

where C is a positive constant depending only on p/p_0 and p/p_1.

Then, we replace (5.14) by

$$\begin{aligned}\|H(j+it)\|_{F^{s_j}_{p_j\infty}} &= \left\|\sup_{n\geq 0} v_{n-1}^{s_j}|h_n(j+it)|\right\|_{L_{p_j}} \\ &\leq (1+C_{p_j,\infty})\left\|\sup_{n\geq 0} v_{n-1}^{s_j}|g_n(j+it)|\right\|_{L_{p_j}} \\ &\leq C(1+C_{p_j,\infty})\left(1+\left\|\|(v_{n-1}^s d_n f)_{n\geq 0}\|_{\ell_q}^{p/p_j}\right\|_{L_{p_j}}\right) \leq C'.\end{aligned}$$

The rest of the proof is the same as in the case of $q_0, q_1 < \infty$. The proof is completed. □

We now give the proof of Theorem 2.3.

Proof of Theorem 2.3. Combining $\theta \in (0,1)$ and $\min(q_0, q_1) < \infty$, we have $q < \infty$. Hence, combining Lemma 5.4 and Lemma 5.5, we obtain that $\|f\|_{[B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1}]_\theta} \leq C_1\|f\|_{B^s_{pq}}$ for all $f \in B^s_{pq}$, where C_1 is the constant in Lemma 5.5. Similarly, if $1 < p_0, p_1 < \infty$, then we have $\|f\|_{[F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1}]_\theta} \leq C_1\|f\|_{F^s_{pq}}$ for all $f \in F^s_{pq}$. Therefore, we only have to show the converse inequality.

We first give a proof for the case of martingale Besov spaces. Let $f \in [B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1}]_\theta$. Let $F \in \mathcal{F}(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})$ such that $F(\theta) = f$. From the fact $\|d_n g\|_{L_{p_j}} \leq C\|g\|_{B^{s_j}_{p_jq_j}}$, where C is a positive constant independent of $g = (g_n)_{n\geq 0} \in \mathcal{M}$, we have that $v_{n-1}^{s_0(1-z)+s_1 z} d_n F(z)$ belongs to $\mathcal{F}(L_{p_0}, L_{p_1})$ by a standard argument. Hence, by Lemma 5.1 with the fact $[L_{p_0}, L_{p_1}]_\theta = L_p$ ([2, Theorem 5.1.1]), we obtain that

$$\|v_{n-1}^s d_n F(\theta)\|_{L_p} \leq a_n^{1-\theta} b_n^\theta \tag{5.15}$$

where

$$\begin{aligned}a_n &= \frac{1}{1-\theta}\int_{-\infty}^{\infty} \|v_{n-1}^{s_0} d_n F(it)\|_{L_{p_0}} \mu_0(\theta,t)\,dt, \\ b_n &= \frac{1}{\theta}\int_{-\infty}^{\infty} \|v_{n-1}^{s_1} d_n F(1+it)\|_{L_{p_1}} \mu_1(\theta,t)\,dt.\end{aligned}$$

Using Minkowski's inequality and the fact that

$$\frac{1}{1-\theta}\int_{-\infty}^{\infty} \mu_0(\theta,t)\,dt = \frac{1}{\theta}\int_{-\infty}^{\infty} \mu_1(\theta,t)\,dt = 1, \tag{5.16}$$

we have

$$\begin{aligned}\|(a_n)_{n\geq 0}\|_{\ell_{q_0}} &\leq \frac{1}{1-\theta}\int_{-\infty}^{\infty} \|F(it)\|_{B^{s_0}_{p_0q_0}} \mu_0(\theta,t)\,dt \leq \|F\|_{\mathcal{F}(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})}, \\ \|(b_n)_{n\geq 0}\|_{\ell_{q_1}} &\leq \frac{1}{\theta}\int_{-\infty}^{\infty} \|F(1+it)\|_{B^{s_1}_{p_1q_1}} \mu_1(\theta,t)\,dt \leq \|F\|_{\mathcal{F}(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})}.\end{aligned} \tag{5.17}$$

Therefore, using (5.15), Hölder's inequality and (5.17), we obtain

$$\begin{aligned}
\|f\|_{B^s_{pq}} &= \|F(\theta)\|_{B^s_{pq}} \\
&= \left\|(\|v^s_{n-1} d_n F(\theta)\|_{L_p})_{n\geq 0}\right\|_{\ell_q} \\
&\leq \left\|(a_n^{1-\theta} b_n^\theta)_{n\geq 0}\right\|_{\ell_q} \\
&\leq \|(a_n)_{n\geq 0}\|_{\ell_{q_0}}^{1-\theta} \|(b_n)_{n\geq 0}\|_{\ell_{q_1}}^{\theta} \\
&\leq (\|F\|_{\mathcal{F}(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})})^{1-\theta} (\|F\|_{\mathcal{F}(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})})^{\theta} = \|F\|_{\mathcal{F}(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})}.
\end{aligned}$$

Thus, we obtain the desired conclusion for the case of martingale Besov spaces.

We next give the proof for the case of martingale Triebel-Lizorkin spaces.

Let $f \in [F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1}]_\theta$. Let $G \in \mathcal{F}(F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})$ such that $G(\theta) = f$. Let $h \in \mathcal{T}$ such that $\|h\|_{F^{-s}_{p'q'}} = 1$. Noting that $1 < p_0, p_1 < \infty$, we use Lemma 5.5 to take $H \in \mathcal{F}(F^{-s_0}_{p_0'q_0'}, F^{-s_1}_{p_1'q_1'})$ such that $H(\theta) = \overline{h}$, $\|H\|_{\mathcal{F}(F^{-s_0}_{p_0'q_0'}, F^{-s_1}_{p_1'q_1'})} \leq C_1$ and that H satisfies (5.5). Then define $D_h(z) = \sum_{n=0}^\infty E\left[d_n G(z) d_n H(z)\right]$. Since H satisfies these conditions mentioned above, we have that $D_h \in \mathcal{F}(\mathbb{C}, \mathbb{C})$. Moreover, as in (4.1), we have

$$\begin{aligned}
|D_h(j+it)| &\leq \|G(j+it)\|_{F^{s_j}_{p_jq_j}} \ \|H(j+it)\|_{F^{-s_j}_{p_j'q_j'}} \\
&\leq \|G\|_{\mathcal{F}(F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})} \ \|H\|_{\mathcal{F}(F^{-s_0}_{p_0'q_0'}, F^{-s_1}_{p_1'q_1'})} \\
&\leq C_1 \|G\|_{\mathcal{F}(F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})}
\end{aligned} \tag{5.18}$$

where $j = 0, 1$. Using Lemma 5.1, (5.18) and (5.16), we obtain that

$$\begin{aligned}
&|D_h(\theta)| \\
&\leq \left(\frac{1}{1-\theta}\int_{-\infty}^\infty |D_h(it)|\mu_0(\theta,t)\,dt\right)^{1-\theta} \left(\frac{1}{\theta}\int_{-\infty}^\infty |D_h(1+it)|\mu_0(\theta,t)\,dt\right)^{\theta} \\
&\leq (C_1\|G\|_{\mathcal{F}(F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})})^{1-\theta} (C_1\|G\|_{\mathcal{F}(F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})})^{\theta} = C_1\|G\|_{\mathcal{F}(F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})}
\end{aligned}$$

for all $h \in \mathcal{T}$ such that $\|h\|_{F^{-s}_{p'q'}} = 1$ and for all $G \in \mathcal{F}(F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})$ such that $G(\theta) = f$. Therefore, we have

$$\sup_{h\in\mathcal{T}:\|h\|_{F^{-s}_{p'q'}}=1} |D_h(\theta)| \leq C_1 \|f\|_{[F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1}]_\theta}. \tag{5.19}$$

For $g = (g_n)_{n\geq 0} \in \mathcal{M}$ and $N \geq 0$, let $g^N = (g_{n\wedge N})_{n\geq 0}$, where $n \wedge N = \min(n, N)$. Define

$$(F^s_{pq})^N = \{g^N : g \in F^s_{pq}\}, \quad \mathcal{T}^N = \{g^N : g \in \mathcal{T}\}.$$

Then, it is clear that $(F^s_{pq})^N$ is a closed subspace of F^s_{pq}. Moreover, by the same way as in Lemma 5.4, we have that $\mathcal{T}^N$ is dense in $(F^{-s}_{p'q'})^N$, even if $q' = \infty$. Hence, by Theorem 2.2 and (5.19), we have

$$\|f^N\|_{F^s_{pq}} \leq C \sup_{h\in\mathcal{T}^N:\|h\|_{F^{-s}_{p'q'}}=1} |D_h(\theta)| \leq CC_1 \|f\|_{[F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1}]_\theta}$$

where C and C_1 are positive constants in Theorem 2.2 and in Lemma 5.5 respectively. Using monotone convergence theorem, we obtain

$$\|f\|_{F^s_{pq}} = \sup_{N\geq 0} \|f^N\|_{F^s_{pq}} \leq CC_1 \|f\|_{[F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1}]_\theta}$$

for all $f \in [F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1}]_\theta$. Therefore, we obtain the desired conclusion for the case of martingale Triebel-Lizorkin spaces. The proof is completed. $\square$

6 Proof of Theorem 2.4.

In this section, we give a proof of Theorem 2.4.

Proof of Theorem 2.4. Since $|d_nf| = |E_n(f - f_{n-1})| \le E_n|f - f_{n-1}|$, we have

$$\|f\|_{B^s_{pq}} \le \left\|(\|v^s_{n-1}E_n|f - f_{n-1}|\|_{L_p})_{n\ge0}\right\|_{\ell_q},$$
$$\|f\|_{F^s_{pq}} \le \left\|\|(v^s_{n-1}E_n|f - f_{n-1}|)_{n\ge0}\|_{\ell_q}\right\|_{L_p}.$$

We now show the converse inequalities. Let δ_1 be the constant in (2.1). We first show (2.6) for $q = \infty$. By (2.1) and the assumption $s > 0$, we have

$$v^s_{n-1}|f - f_{n-1}| \le v^s_{n-1}\sum_{k=n}^{\infty}|d_kf| \le \sum_{k=n}^{\infty}\delta_1^{s(n-k)}|v^s_{k-1}d_kf|. \tag{6.1}$$

From Jensen's inequality for E_n and (6.1), we have

$$\left\|v^s_{n-1}E_n|f - f_{n-1}|\right\|_{L_p} \le \left\|v^s_{n-1}|f - f_{n-1}|\right\|_{L_p} \le \sum_{k=n}^{\infty}\delta_1^{s(n-k)}\|v^s_{k-1}d_kf\|_{L_p} \le \frac{\|f\|_{B^s_{p\infty}}}{1 - \delta_1^{-s}}.$$

Therefore, we have (2.6) for $q = \infty$.

To show (2.7) for $q = \infty$, let $G = \sup_{n\ge0} v^s_{n-1}|d_nf|$. Then, by (6.1), we have

$$v^s_{n-1}E_n|f - f_{n-1}| \le \sum_{k=n}^{\infty}\delta_1^{s(n-k)}E_n|v^s_{k-1}d_kf| \le \frac{E_nG}{1 - \delta_1^{-s}}.$$

Therefore, using Doob's inequality, we have

$$\begin{aligned}
\left\|\sup_{n\ge0} v^s_{n-1}E_n|f - f_{n-1}|\right\|_{L_p} &\le (1 - \delta_1^{-s})^{-1}\left\|\sup_{n\ge0} E_nG\right\|_{L_p} \\
&\le \frac{p}{(p-1)(1 - \delta_1^{-s})}\|G\|_{L_p} \\
&= \frac{p}{(p-1)(1 - \delta_1^{-s})}\|f\|_{F^s_{p\infty}}.
\end{aligned}$$

Thus, we have (2.7) for $q = \infty$.

We next show (2.6) for $0 < q < \infty$. If $q \le 1$, then we have

$$\begin{aligned}
\sum_{n=0}^{\infty}\left(\sum_{k=n}^{\infty}\delta_1^{s(n-k)}\|v^s_{k-1}d_kf\|_{L_p}\right)^q &\le \sum_{n=0}^{\infty}\sum_{k=n}^{\infty}\delta_1^{sq(n-k)}\|v^s_{k-1}d_kf\|^q_{L_p} \\
&= \sum_{k=0}^{\infty}\left(\sum_{n=0}^{k}\delta_1^{sq(n-k)}\right)\|v^s_{k-1}d_kf\|^q_{L_p} \\
&\le (1 - \delta_1^{-sq})^{-1}\sum_{k=0}^{\infty}\|v^s_{k-1}d_kf\|^q_{L_p}.
\end{aligned}$$

If $1 < q < \infty$, then, denoting by q' the conjugate exponent of q, we have

$$\begin{aligned}
(6.2)\quad \sum_{n=0}^{\infty}\left(\sum_{k=n}^{\infty}\delta_1^{s(n-k)}\|v_{k-1}^s d_k f\|_{L_p}\right)^q &= \sum_{n=0}^{\infty}\left(\sum_{k=n}^{\infty}\delta_1^{s(n-k)/q'}\delta_1^{s(n-k)/q}\|v_{k-1}^s d_k f\|_{L_p}\right)^q \\
&\le \sum_{n=0}^{\infty}\left(\sum_{k=n}^{\infty}\delta_1^{s(n-k)}\right)^{q/q'}\sum_{k=n}^{\infty}\delta_1^{s(n-k)}\|v_{k-1}^s d_k f\|_{L_p}^q \\
&= (1-\delta_1^{-s})^{-q/q'}\sum_{k=0}^{\infty}\left(\sum_{n=0}^{k}\delta_1^{s(n-k)}\right)\|v_{k-1}^s d_k f\|_{L_p}^q \\
&\le (1-\delta_1^{-s})^{-q}\sum_{k=0}^{\infty}\|v_{k-1}^s d_k f\|_{L_p}^q.
\end{aligned}$$

Therefore, we have

$$\sum_{n=0}^{\infty}\left(\sum_{k=n}^{\infty}\delta_1^{s(n-k)}\|v_{k-1}^s d_k f\|_{L_p}\right)^q \le C^q\sum_{k=0}^{\infty}\|v_{k-1}^s d_k f\|_{L_p}^q \tag{6.3}$$

where C is a positive constant depending only on s, q and δ_1. Combining Remark 4.1, (6.3) and (6.1), we have

$$\left(\sum_{n=0}^{\infty}\left\|v_{n-1}^s E_n|f-f_{n-1}|\right\|_{L_p}^q\right)^{1/q} \le \left(\sum_{n=0}^{\infty}\left\|v_{n-1}^s|f-f_{n-1}|\right\|_{L_p}^q\right)^{1/q} \le C\|f\|_{B_{pq}^s}.$$

Thus, we obtain (2.6).

We now show (2.7) for $1 \le q < \infty$. Similarly as in (6.2), we have

$$\sum_{n=0}^{\infty}\left(\sum_{k=n}^{\infty}\delta_1^{s(n-k)}v_{k-1}^s|d_k f|\right)^q \le (1-\delta_1^{-s})^{-q}\sum_{k=0}^{\infty}(v_{k-1}^s|d_k f|)^q. \tag{6.4}$$

Combining (6.1) and (6.4), we have

$$\left(\sum_{n=0}^{\infty}(v_{n-1}^s|f-f_{n-1}|)^q\right)^{1/q} \le (1-\delta_1^{-s})^{-1}\left(\sum_{k=0}^{\infty}(v_{k-1}^s|d_k f|)^q\right)^{1/q}.$$

Therefore, using Lemma 4.1, we have

$$\begin{aligned}
\left\|\left(\sum_{n=0}^{\infty}(v_{n-1}^s E_n|f-f_{n-1}|)^q\right)^{1/q}\right\|_{L_p} &\le C_{p,q}\left\|\left(\sum_{n=0}^{\infty}(v_{n-1}^s|f-f_{n-1}|)^q\right)^{1/q}\right\|_{L_p} \\
&\le \frac{C_{p,q}}{1-\delta_1^{-s}}\left\|\left(\sum_{k=0}^{\infty}(v_{k-1}^s|d_k f|)^q\right)^{1/q}\right\|_{L_p}.
\end{aligned}$$

We have the desired conclusion.

□

7 Proofs of Theorems 2.5 and 2.6. In this section, we give proofs of Theorems 2.5 and 2.6. As is described in Section 2, we postulate following conditions:

(7.1) Every σ-algebra $\mathcal{F}_n$ is generated by countable atoms.

(7.2) $\{\mathcal{F}_n\}_{n\geq 0}$ is regular.

(7.3) If $B \in A(\mathcal{F}_{n-1})$, $B' \in A(\mathcal{F}_n)$ and $B' \subset B$, then $P(B') < P(B)$ for every $n \geq 1$.

(7.4) $\mathcal{F}_0 = \{\emptyset, \Omega\}$,

where $A(\mathcal{F}_n)$ stands for the set of all atoms in $\mathcal{F}_n$. Define $\mathcal{F}_n$-measurable functions b_n and v_n by

$$b_n = \sum_{B\in A(\mathcal{F}_n)} P(B)\chi_B, \quad v_n = b_n^{-1}. \tag{7.5}$$

By [11, Lemma 3.3], b_n satisfy

$$\left(1+\frac{1}{R}\right) b_n \leq b_{n-1} \leq R b_n \tag{7.6}$$

where R is the constant in (2.10). Hence, $v = (v_n)_{n\geq 0}$ in (7.5) belongs to $\mathcal{V}$.

We start with the following lemma.

Lemma 7.1. *Let $p_0, p_1 \in (0,\infty)$ with $p_0 < p_1$. Let n be a non-negative integer. Let $\alpha = 1/p_0 - 1/p_1$. Let $f = (f_n)_{n\geq 0} \in \mathcal{M}$. If $d_n f \in L_{p_1}$, then*

$$\|d_n f\|_{L_{p_1}} \leq R^\alpha \|v_{n-1}^\alpha d_n f\|_{L_{p_0}} \tag{7.7}$$

with convention $v_{-1} = v_0$ and $f_{-1} = 0$, where R is the constant in (2.10)*.*

Proof. If $n = 0$, then $\|d_n f\|_{L_{p_1}} = \|v_{n-1}^\alpha d_n f\|_{L_{p_0}}$ because $d_0 f$ is constant and $v_{-1} = v_0 = 1$ by $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Since $R \geq 2$, we have (7.7) for $n = 0$.

For $n \geq 1$, let $B \in A(\mathcal{F}_n)$. Then, since $d_n f$ is constant on B, we have

$$\|\chi_B d_n f\|_{L_{p_1}} = P(B)^{1/p_1} \|\chi_B d_n f\|_{L_\infty} = P(B)^{1/p_1 - 1/p_0} \|\chi_B d_n f\|_{L_{p_0}}. \tag{7.8}$$

Using (7.8), $\alpha = 1/p_0 - 1/p_1$ and $v_n \leq R v_{n-1}$, we have

$$\|\chi_B d_n f\|_{L_{p_1}} = \|\chi_B v_n^\alpha d_n f\|_{L_{p_0}} \leq R^\alpha \|\chi_B v_{n-1}^\alpha d_n f\|_{L_{p_0}}. \tag{7.9}$$

Using (7.9), we have

$$\begin{aligned}
\|d_n f\|_{L_{p_1}}^{p_1} &= \sum_{B\in A(\mathcal{F}_n)} \|\chi_B d_n f\|_{L_{p_1}}^{p_1} \\
&\leq R^{\alpha p_1} \sum_{B\in A(\mathcal{F}_n)} \|\chi_B v_{n-1}^\alpha d_n f\|_{L_{p_0}}^{p_1} \\
&\leq R^{\alpha p_1} \left(\sum_{B\in A(\mathcal{F}_n)} \|\chi_B v_{n-1}^\alpha d_n f\|_{L_{p_0}}^{p_0} \right)^{p_1/p_0} = R^{\alpha p_1} \|v_{n-1}^\alpha d_n f\|_{L_{p_0}}^{p_1}.
\end{aligned}$$

We have the desired conclusion. □

We next show the following lemma.

Lemma 7.2. *Let $p \in (0,\infty)$ and $\alpha \in (0,\infty)$. Let n be a non-negative integer. Let $f \in F^{\alpha}_{p\infty}$ with $\|f\|_{F^{\alpha}_{p\infty}} = 1$. Then,*

$$|v^{\alpha}_{n-1} d_n f| \le R^{1/p} v^{1/p}_{n-1}$$

with convention $v_{-1} = v_0$ and $f_{-1} = 0$, where R is the constant in (2.10).

Proof. Let $B \in A(\mathcal{F}_n)$. Since $v^{\alpha}_{n-1} d_n f$ is constant on B, we have

$$\begin{aligned}
\chi_B |v^{\alpha}_{n-1} d_n f| &= \chi_B \left(\frac{1}{P(B)} \int_B |v^{\alpha}_{n-1}(\omega) d_n f(\omega)|^p \, dP(\omega) \right)^{1/p} \\
&\le \frac{\chi_B}{P(B)^{1/p}} \left(\int_\Omega |v^{\alpha}_{n-1}(\omega) d_n f(\omega)|^p \, dP(\omega) \right)^{1/p} \\
&\le \chi_B v_n^{1/p} \left(\int_\Omega \sup_{n\ge 0} |v^{\alpha}_{n-1}(\omega) d_n f(\omega)|^p \, dP(\omega) \right)^{1/p} \\
&\le \chi_B R^{1/p} v^{1/p}_{n-1} \|f\|_{F^{\alpha}_{p\infty}} = \chi_B R^{1/p} v^{1/p}_{n-1}.
\end{aligned}$$

The proof is completed. □

We now show Theorems 2.5 and 2.6.

Proof of Theorem 2.5. By Theorem 2.1, we only have to give a proof for the case where $s = 0$. We first show that

$$\|f\|_{B^0_{p_1 q}} \le R^{\alpha} \|f\|_{B^{\alpha}_{p_0 q}}. \tag{7.10}$$

Indeed, using Lemma 7.1, we have (7.10) as follows:

$$\|f\|_{B^0_{p_1 q}} = \left(\sum_{n=0}^{\infty} \|d_n f\|^q_{L_{p_1}} \right)^{1/q} \le R^{\alpha} \left(\sum_{n=0}^{\infty} \|v^{\alpha}_{n-1} d_n f\|^q_{L_{p_0}} \right)^{1/q} = R^{\alpha} \|f\|_{B^{\alpha}_{p_0 q}}.$$

We next show

$$\|f\|_{F^0_{p_1 q}} \le R^{\alpha} \|f\|_{F^{\alpha}_{p_0 \infty}}. \tag{7.11}$$

Let $f \in F^{\alpha}_{p_0\infty}$ with $\|f\|_{F^{\alpha}_{p_0\infty}} = 1$. Let

$$F(\omega) = \sup_{n\ge 0} |v^{\alpha}_{n-1}(\omega) d_n f(\omega)|, \quad G(\omega) = \left(\sum_{n=0}^{\infty} |d_n f(\omega)|^q \right)^{1/q}.$$

We show

$$G \le C F^{p_0/p_1} \tag{7.12}$$

where C is a positive constant depending only on p_0, p_1, q and R. Indeed, using Lemma 7.2 with $p = p_0$ and $\alpha = 1/p_0 - 1/p_1$, we have

$$\begin{aligned}
|d_n f| &= v^{-\alpha}_{n-1} |v^{\alpha}_{n-1} d_n f| \\
&\le \min(v^{-\alpha}_{n-1} F, v^{-\alpha}_{n-1} R^{1/p_0} v^{1/p_0}_{n-1}) \\
&= \min(v^{-\alpha}_{n-1} F, R^{1/p_0} v^{1/p_1}_{n-1}).
\end{aligned}$$

Therefore, noting the convention $v_{-1} = v_0$ and the fact $(1+1/R)v_{n-1} \le v_n \le Rv_{n-1}$, we have (7.12) as follows:

$$\begin{aligned}
G^q &\le \sum_{n=0}^{\infty} \min(v_{n-1}^{-\alpha}F, R^{1/p_0}v_{n-1}^{1/p_1})^q \\
&\le 2\sum_{n=1}^{\infty} \min(v_{n-1}^{-\alpha}F, R^{1/p_0}v_{n-1}^{1/p_1})^q \\
&\le 2\sum_{n=1}^{\infty} \frac{1}{\log(1+1/R)} \int_{v_{n-1}}^{v_n} \min(v_{n-1}^{-\alpha}F, R^{1/p_0}v_{n-1}^{1/p_1})^q \frac{dt}{t} \\
&\le \frac{2}{\log(1+1/R)} \sum_{n=1}^{\infty} \int_{v_{n-1}}^{v_n} \min(R^{\alpha}t^{-\alpha}F, R^{1/p_0}t^{1/p_1})^q \frac{dt}{t} \\
&\le \frac{2R^{q/p_0}}{\log(1+1/R)} \int_1^{\infty} \min(t^{-\alpha}F, t^{1/p_1})^q \frac{dt}{t} \\
&\le \frac{2R^{q/p_0}}{q\log(1+1/R)} \left(\frac{1}{\alpha} + p_1\right) F^{qp_0/p_1}.
\end{aligned}$$

By (7.12), we have

$$\|f\|_{F^0_{p_1q}} = \|G\|_{L_{p_1}} \le C\|F\|_{L_{p_0}}^{p_0/p_1} = C\|f\|_{F^{\alpha}_{p_0\infty}}^{p_0/p_1} = C.$$

We have the desired conclusion.

□

Proof of Theorem 2.6. By Theorem 2.1 and Proposition 3.2, we only have to show $B^0_{p\infty} \hookrightarrow B^{-1/p}_{\infty\infty}$. As in (7.8), we have

$$\|v_{n-1}^{-1/p}d_nf\|_{L_\infty} \le R^{1/p}\|v_n^{-1/p}d_nf\|_{L_\infty} \le R^{1/p} \sup_{B\in A(\mathcal{F}_n)} \|\chi_B d_nf\|_{L_p} \le R^{1/p}\|f\|_{B^0_{p\infty}}.$$

We obtain the desired conclusion. □

8 Proof of Theorem 2.7 and Corollary 2.8. In this section, we prove Theorem 2.7 and Corollary 2.8. To do this, we need the following John-Nirenberg type lemma.

Lemma 8.1. *Let $p \in (0,\infty)$ and $q \in [1,\infty)$. Then, the following equivalence holds:*

$$\|f\|_{F^0_{\infty q}} \sim \sup_{n\ge 0} \left\| E_n\left[\left(\sum_{k=n}^{\infty} |d_kf|^q\right)^{p/q}\right]^{1/p} \right\|_{L_\infty} \qquad (f \in F^0_{\infty q}).$$

For the proof of Lemma 8.1, we refer to [17, Theorem 2].

Lemma 8.2. *Suppose that every σ-algebra $\mathcal{F}_n$ is generated by countable atoms. Furthermore, assume that $\{\mathcal{F}_n\}_{n\ge0}$ is regular with (2.11) and (2.12). Let $v = (v_n)_{n\ge0}$ be the sequence of functions defined in (2.9). Let $s > 0$. Then, $B^s_{\infty\infty} = \Lambda_1^-(s)$ with equivalent norms.*

Proof. By Theorem 2.4 and the regularity of $\{\mathcal{F}_n\}_{n\geq 0}$, we have

$$\begin{aligned}
\|f\|_{B^s_{\infty\infty}} &\sim \sup_{n\geq 0} \left\| v^s_{n-1} E_n|f-f_{n-1}|\right\|_{L_\infty} \\
&\sim \sup_{n\geq 0} \|v^s_n E_n|f-f_{n-1}|\|_{L_\infty} \\
&= \sup_{n\geq 0} \sup_{B\in A(\mathcal{F}_n)} \frac{1}{P(B)^{1+s}} \int_B |f(\omega)-(E_{n-1}f)(\omega)|\, dP(\omega) \\
&= \|f\|_{\Lambda_1^-(s)}.
\end{aligned}$$

We have the desired conclusion. □

Proof of Theorem 2.7. We obtain (2.16) and (2.18) from (2.13) and (2.14) respectively. To show (2.17), let $f \in F^0_{pq}$ and $\alpha = 1/p$. Noting that $|d_k I_\alpha f| = |b^\alpha_{k-1} d_k f| \leq b^\alpha_{n-1}|d_k f|$ for $k \geq n$ and using the regularity of $\{\mathcal{F}_n\}_{n\geq 0}$, we have

$$\left(\sum_{k=n}^\infty |d_k I_\alpha f|^q\right)^{1/q} \leq b^\alpha_{n-1}\left(\sum_{k=n}^\infty |d_k f|^q\right)^{1/q} \leq R^\alpha b^\alpha_n \left(\sum_{k=0}^\infty |d_k f|^q\right)^{1/q}.$$

Then, for $B \in A(\mathcal{F}_n)$, we have

$$\begin{aligned}
E\left[\chi_B\left(\sum_{k=n}^\infty |d_k I_\alpha f|^q\right)^{p/q}\right] &\leq R^{\alpha p} P(B)^{\alpha p} E\left[\chi_B\left(\sum_{k=0}^\infty |d_k f|^q\right)^{p/q}\right] \\
&\leq R P(B)\|f\|^p_{F^0_{pq}}
\end{aligned}$$

by $\alpha = 1/p$. Since $\mathcal{F}_n$ is generated by $A(\mathcal{F}_n)$, this means

$$E_n\left[\left(\sum_{k=n}^\infty |d_k I_\alpha f|^q\right)^{p/q}\right]^{1/p} \leq R^\alpha \|f\|_{F^0_{pq}}. \tag{8.1}$$

Combining (8.1), Lemma 8.1 and Theorem 2.1, we obtain the desired conclusion. □

Proof of Corollary 2.8. Taking $s = 0$, $q = 2$ in (2.16) and combining with the fact $L_p(\ell_2) \hookrightarrow L_p(\ell_\infty)$, we obtain (2.19). Similarly, taking $s = 0$, $q = 2$ in (2.17), we obtain (2.20). Taking $s = 0$ in (2.18), we obtain (2.21) by Lemma 8.2 and by the fact $L_p(\ell_2) \hookrightarrow L_p(\ell_\infty)$. The proof is completed. □

Acknowledgement The author was supported by Grant-in-Aid for Scientific Research (C), No. 16K05203, Japan Society for the Promotion of Science.

References

[1] N. Asmar and S. Montgomery-Smith, *Littlewood-Paley theory on solenoids*, Colloq. Math. 65 (1993), 69-82.

[2] J. Bergh and J. Löfström, *Interpolation spaces: an introduction*, Grundlehren Math. Wiss. 223, Berlin, Springer-Verlag, 1976.

[3] A. Calderón, *Intermediate spaces and interpolation, the complex method*, Studia Math. 24 (1964), 113–190.

[4] J. A. Chao and H. Ombe, *Commutators on dyadic martingales*, Proc. Japan Acad. 61, Ser. A (1985), 35–38.

[5] J. A. Chao and L. Peng, *Schatten classes and commutators on simple martingales*, Colloq. Math 71 (1996), 7–21.

[6] A. W. Garsia, *Martingale inequalities: Seminar notes on recent progress*, Mathematics Lecture Notes Series. W. A. Benjamin, Mass.-London-Amsterdam, 1973.

[7] Z. Hao and Y. Jiao, *Fractional integral on martingale Hardy spaces with variable exponents*, Fract. Calc. Appl. Anal. 18 (2015), 1128–1145.

[8] Y. Jiao, L. Wu, A. Yang and R. Yi, *The predual and John-Nirenberg inequalities on generalized BMO martingale spaces*, Trans. Amer. Math. Soc. 369 (2017), 537–553.

[9] N. Kalton, S. Mayboroda and M. Mitrea, *Interpolation of Hardy-Sobolev-Besov-Triebel-Lizorkin spaces and applications to problems in partial differential equations*, Comtemp. Math. 445 (2007), 121-178.

[10] R. L. Long, *Martingale spaces and inequalities*, Peking University Press, Beijing, 1993.

[11] E. Nakai and G. Sadasue, *Martingale Morrey-Campanato spaces and fractional integrals*, J. Funct. Spaces Appl. 2012 (2012), Article ID 673929, 29 pages. DOI:10.1155/2012/673929.

[12] E. Nakai and G. Sadasue, *Pointwise multipliers on martingale Campanato spaces*, Studia Math. 220 (2014), 87–100.

[13] T. Noi and Y. Sawano, *Complex interpolation of Besov spaces and Triebel-Lizorkin spaces with variable exponents*, Jour. Math. Anal. Appl. 387 (2012), no. 2, 676–690.

[14] G. Sadasue, *Fractional integrals on martingale Hardy spaces for* $0 < p \leq 1$, Mem. Osaka Kyoiku Univ. Ser. III Nat. Sci. Appl. Sci. 60 (2011), 1–7.

[15] H. Triebel, *Theory of function spaces*, Monographs in mathematics 78, Basel, Boston, Birkhäuser, 1983.

[16] F. Weisz, *Martingale Hardy spaces and their applications in Fourier analysis*, Lecture Notes in Math. 1568, Springer-Verlag, Berlin, 1994.

[17] F. Weisz, *Martingale operators and Hardy spaces generated by them*, Studia Math. 114 (1995), 39–70.

Communicated by *Eiichi Nakai*

Gaku Sadasue
Department of Mathematics
Osaka Kyoiku University
Kashiwara, Osaka 582-8582, Japan
sadasue@cc.osaka-kyoiku.ac.jp

A NEW APPROACH TO REDUCE THE BURDEN OF PAIRWISE-COMPARISON ON THE AHP

SangHeon Han * Tsutomu Hiroshima † Toru Yarita ‡

Received April 18, 2017 ; revised May 15, 2017, July 4, 2017

Abstract. The pairwise comparisons in AHP (Analytic Hierarchy Process) are made using a scale list that indicates the importance of one entity over another entity with respect to a given criteria. Moreover, the pairwise comparison matrix represents the intensities of the decision maker's preference between individual pairs of alternatives. The matrix is usually determined from the 1-point to 9-point scale. Various methods for paired comparison method have been proposed, making more intuitive and highly accurate decision making possible. However, the number of pairwise comparisons increases as the number of criteria increases. Therefore, the burden of decision makers would become heavier.

In this paper, we propose an algorithm for the allocation problem of the burden and verify the algorithm by using a programming language called Haskell[1], which is specialized in the functional programming. This research contributes not only to allocation algorithm, but also aids researchers and decision makers in applying the AHPs effectively.

1 Introduction Pairwise Comparison Method is always used in comparing entities in pairs, whereby there are more than two entities to evaluate. Because the process of evaluating is very simple, the pairwise comparison has been broadly used in studies of preference, precedence and social choice etc.. Methods which utilize pairwise comparison focus on superiority/inferiority between entities or its differences. Pairwise Comparison Method is proposed as a typical method in AHP. In 1971, Thomas L. Saaty proposed AHP as a supporting tool for decision making to connect human's subjective evaluation with reasonable decision. In AHP, decision maker regards a problem as hierarchy relation of criteria /alternatives levels for systematic approaches to problem. As a result, AHP is useful in multiple criteria problem and in subjective decision-making problem. In real, AHP has been broadly applied to decision making in business or to consensus building in public works, etc.. However, AHP uses a great number of pairwise comparisons to calculate numerical weights of criteria and alternatives in general. While pairwise comparison methods have merits of simple process and accurate evaluation, many studies have indicated demerits that decision maker's comparing burden is heavy if comparing entities increase in Cowan (2001). There exists similar demerit in AHP because AHP needs to compare with many pairs of criteria if the number of criteria increases for more exact evaluation. Furthermore, if the number

2010 *Mathematics Subject Classification.* 91A10, 91A40, 91A50, 91A65.

Key words and phrases. Analytic Hierarchy Process (AHP), Combinatorial design method, Pairwise comparison .

*Department of Management Administration, Nagoya University of Commerce & Business, 4-4 Sagamine, Komenoki-cho, Nisshin-shi, Aichi-ken, Japan 470-0193, Phone: +81-(0)561-73-2111, E-mail: han-sh@nucba.ac.jp

†Department of Management Administration, Nagoya University of Commerce & Business, 4-4 Sagamine, Komenoki-cho, Nisshin-shi, Aichi-ken, Japan 470-0193, Phone: +81-(0)561-73-2111, E-mail: hiroshima@nucba.ac.jp

‡Department of Economics, Nagoya University of Commerce & Business, 4-4 Sagamine, Komenoki-cho, Nisshin-shi, Aichi-ken, Japan 470-0193, Phone: +81-(0)561-73-2111, E-mail: yarita@nucba.ac.jp

[1]Haskell is a relatively new language that was born in 1990 (Haskell 1.0) as a "standard word" in purely functional programming language based on lazy evaluation. In 1999, as a stable language definition Haskell 98 was enacted, and gradually became popular.

of alternatives increases then it is necessary to compare with many pairs of alternatives for each criterion at a higher level. Han (2014) pointed out the heavy burden of pairwise comparison in AHP and proposed a method to decrease the value of C.I..

Therefore, in this study we concentrate on the problem of sharing the heavy burden of pairwise comparison work with multiple decision makers under the assumption of making consensus building. The problem is transformed into group AHP problem, which is a necessary consensus building process to solve the problem.

Next, we outline the configuration of the study. Section 2 describes the background and the objective of the study and treat of past studies. In section 3, we describe basic theory related to burden sharing of pairwise comparison works. Section 4 considers fair sharing burden problem and propose efficient solution procedure. Section 5 discusses the conclusion of the study and subjects in future.

2 Review of Related study During last three decades, many attempts were made in directing investors so that AHP and Fuzzy are proposed as tools for portfolio optimization for financial decision making.

Saaty (1980) and Durer et al. (1994) analyzed a complex portfolio system. Furthermore, to select securities using data envelopment analysis (DEA), Tiryaki (2001) evaluate the financial performance of companies. Bolster et al. (1995) used AHP to determine investor suitability, based on age, to select among seven investment securities. The results of their study showed varying patterns of investment for the different age groups.

Thalheimer and Ali (1979) applied AHP to time series analysis and to portfolio selection on mutual savings banks to determine the optimal portfolio choice. They opined that an investor should invest in short-term stocks, risk-less assets, and well diversified investment to achieve the highest utility from investable fund. Oyatoye et al. (2010) applied AHP to investment portfolio selection in the banking sector. They underlines the importance of different criteria, factors and alternatives that are essential to successful investment decisions in the financial crisis in 2008. We find that each criterion plays an important role in the portfolio. Nevertheless, the effect of the AHP-portfolio on the total returns is trivial. Therefore, we determine that the decision maker prefers the government bond to the others. We also determine that the portfolio generated by AHP accords with the decision maker's preference.

3 Background and objective of study For proper number of criteria and alternatives in case single decision maker, concept of magical number is introduced to be 7 ± 2, as suggested by Miller (1956).

After that, magical number of 5 to 9 has been changed to number of less than equal to 4 in order to satisfy transitive law as possible in Cowan (2001). In short, proper number of criteria and alternatives is 3 or 4 in AHP for one decision maker. The number is regarded as physical limitation of decision maker in this study. Let's consider an interview test in entrance examination as an example. If the number of candidates for ranking increases then candidates will be split into multiple interviewers due to the physical limitation of interviewer. The ranking is decided by consensus building in meeting. In the above example, interview time is limited by physical situation of interviewer for one group of candidates. If proper interview time is given then the necessary numbers of interviewers and interview rooms are fixed.

In AHP pairwise comparison, decision maker continues to evaluate by referring to scale list. For example, when he/she considers comparison between alternative A and B, if A is more important than B then he/she evaluates as corresponding integer from 1 to 9 referred to scale list. In the contrary case, he/she evaluate as corresponding fraction from 1/1 to 1/9.

This procedure has different complexity of evaluating from questionaries' format. In AHP pairwise comparison, evaluating value is easy to lead into confusion, which results from violation of transitive law. It is necessary to investigate efficient method such as hybrid type of using interval and precedence criteria, simultaneously.

In this study, we suppose a decision maker's physical constraints and focus on sharing heavy burden of AHP pairwise comparison work with multiple decision makers for improving process of pairwise comparison. The objective of this study is to propose an efficient solution procedure to support AHP pairwise comparison with decision maker's physical limitation and to investigate what pairwise comparisons should be.

4 Pairwise comparison and combinatorial design theory In this section, we consider a wine-tasting problem as a series of allocation problem. There are 7 wines of different brands, such as A, B, C, D, E, F, G which are evaluated based on some criteria by an expert sommelier.

In addition, we suppose a physical constraint of sommelier that they can test only 3 brands. That is, one sommelier can compare in three combinatorial pairs from three types of wines. The number of all pairs is from seven types of wine. Because one sommelier can test three pairs of wines, it is necessary to split all pairwise comparisons into seven sommeliers. All wines are named by A, B, ..., G and only three types of wines must be assigned to each sommelier. How to assign all types of wines to seven sommeliers fairly? The above problem has been studied in the design theory and some methods are proposed such as cyclic design method. As a result of utilizing cyclic design method, optimal design is obtained as follows;

$$\{A, B, C\}, \{A, D, E\}, \{A, F, G\}, \{B, D, F\}, \{B, E, G\}, \{C, D, G\}, \{C, E, F\}.$$

In Table 1, we illustrate the above optimal design as matrix, which called incidence matrix. Titles of row and column are wine and sommelier in matrix, respectively. If there is a design then 1 is assigned, otherwise 0. As a result, we obtain 7×7 matrix in Table 1.

Table 1: Incidence matrix of wine and sommelier

wine	1	2	3	4	5	6	7
wine A	1	1	1	0	0	0	0
wine B	1	0	0	1	1	0	0
wine C	1	0	0	0	0	1	1
wine D	0	1	0	1	0	1	0
wine E	0	1	0	0	1	0	1
wine F	0	0	1	1	0	0	1
wine G	0	0	1	0	1	1	0

In incidence matrix, there are three 1 in any column and it means that any sommelier test three brands of wines. And there are three 1 in any row and it means that any brand of wines tested by three sommeliers. Furthermore, we know that the matrix reflects all pairs of wine and sommelier. Partially, the combinatorial design theory traces its origin to recreational mathematics in middle of 19th century. However, the theory has been activated and developed by the design of experiments (DOE). It is the start of DOE that R. A. Fisher studied an agricultural test in Rothamsted in 1919. He introduced terms of plot, treatment, block and variety to his experiment. These terms and related symbols due to R. A. Fisher are used as they stand in design theory. For example, symbol of treatment for each plot is used to denote variety in DOE, which was suggested by Fisher (1937) and Ishii (1972).

Now we consider representation of design by symbol. The number of treatments is denoted by v. In the example of wine test, the number of wine brands is denoted by $v = 7$ and number of sommeliers is $b = 7$. Next, r denotes the number of blocks for each treatment and $r = 3$ in the wine test problem. Symbol k denotes the number of plots for each block and $k = 3$ in the wine test problem. Finally, λ denotes the number of blocks in the case there are any treatment pairs and $\lambda = 1$ in the wine test problem. (Table 2)

Table 2: Symbols in wine test problem

	plot (k)	treatment (v)	block (b)	r	λ
wine test	Sommelier's Constraints	Number of wines	Number of sommeliers	Number of Sommeliers on each wine	One testing
	3	7	7	3	1

Under the above setting of notation, (v, b, r, k, λ) is used to describe combinatorial design problems suggested by Mazur (2010). Note that v, b, r, k must satisfy the following equation:

$$ {}_v\mathrm{C}_2 = bk = vr. \tag{1} $$

For instances, wine test problem is expressed by $(7, 7, 3, 3, 1)$ design. The above design problem is efficiently utilized to sharing heavy burden of comparing work to multiple sommeliers for the AHP problem constrained by physical limitation of decision maker.

5 Fair sharing problem of burden Suppose the physical limitation k and the number of wines v are given, the required number of sommeliers b and the number of blocks of each wine r are determined by (1). Therefore, there exists a fair sharing if and only if

$$ \frac{{}_v\mathrm{C}_2}{k} \text{ is integer and } v \text{ is odd.} \tag{2} $$

In this study, we suppose $k = 3$, that is, each decision maker can cover with only 3 entities because of physical limitation. The assumption is based on a minimum number of entities for testing exactly and there are 3 combinatorial pairwise comparisons. For example, if wines of $\{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$ are assigned then decision maker test three pairwise comparison of (A, B), (B, C) and (A, C). However, there is no meaning to share comparison work when decision maker test only two wines of $\{\mathrm{A}, \mathrm{B}\}$ because there is only one pairwise comparison.

In general, accuracy of AHP pairwise comparison is calculated based on CI (Consistency Index) and Han (2014) pointed out that the greater the number of options, the greater the burden of decision-makers by pairwise comparisons.

Table 3 considers wine test problem with 7 wines ($v = 7$) and 3 physical limitations ($k = 3$). There are ${}_v\mathrm{C}_2 = 21$ pairwise comparisons for each wine and design of assigning 3 wines to 7 sommeliers ($b = 7$). In short, sommelier 1 tests 3 wines of $\{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$ and compares with 3 pairs of (A, B), (B, C) and (A, C). Figure 4 describes wine test problem with 9 wines ($v = 9$) and 3 physical limitations ($k = 3$). In case 8 wines ($v = 8$) there are all pairwise comparisons of ${}_v\mathrm{C}_2 = \frac{8\times 7}{1\times 2} = 28$ and necessary number of sommeliers is obtained $b = 28/3 = 9.33\ldots$ from the assumption of physical limitation. Therefore, there is no optimal design because the number of sommeliers is not positive integer.

Now Table 4 describes necessary number of sommeliers and the existence of optimal design based on situation with 9 wines ($v = 9$) and 3 physical limitations ($k = 3$).

Table 5 describes the judgements of existence of design. Column 6, 7 and 8 describe whether the condition (2) is satified.

Table 3: All pairwise comparisons and optimal design for fair sharing — Example of 7 wines constrained to 3 physical limitations

all pairwise comparisons					
AB	BC	CD	DE	EF	FG
AC	BD	CE	DF	EG	
AD	BE	CF	DG		
AE	BF	CG			
AG	BG				
AF					

assignment
sommelier-1:$\{A, B, C\}$
sommelier-2:$\{A, D, E\}$
sommelier-3:$\{A, F, G\}$
sommelier-4:$\{B, D, F\}$
sommelier-5:$\{B, E, G\}$
sommelier-6:$\{C, D, G\}$
sommelier-7:$\{C, E, F\}$

Table 4: All pairwise comparisons and optimal design for fair sharing — Example of 9 wines constrained to 3 physical limitations

all pairwise comparisons							
AB	BC	CD	DE	EF	FG	GH	HI
AC	BD	CE	DF	EG	FH	GI	
AD	BE	CF	DG	EH	FI		
AE	BF	CG	DH	EI			
AF	BG	CH	DI				
AG	BH	CI					
AH	BI						
AI							

assignment
Sommelier-1 :$\{A, B, C\}$
Sommelier-2 :$\{A, D, E\}$
Sommelier-3 :$\{A, F, G\}$
Sommelier-4 :$\{A, H, I\}$
Sommelier-5 :$\{B, D, F\}$
Sommelier-6 :$\{B, E, H\}$
Sommelier-7 :$\{B, G, I\}$
Sommelier-8 :$\{C, D, I\}$
Sommelier-9 :$\{C, E, G\}$
Sommelier-10:$\{C, F, H\}$
Sommelier-11:$\{D, G, H\}$
Sommelier-12:$\{E, F, I\}$

In Figure 1, we list up the optimal design by cyclic method after judging the existence of design. For detail cyclic method, refer to note of Rosa (1991).

In this study we consider an AHP problem to rank wines. The AHP problem is illustrated by AHP chart in Figure 1. For simplicity, we suppose that there are wines of less than 10 and one decision maker (sommelier, appraiser) has physical limitation of 3 wines for testing. For the problem, we propose an efficient algorithm to share heavy testing burden fairly with multiple tester (sub-sommeliers) based on the combinatorial design theory. It is necessary to investigate the existence of optimal design and construction of the design for fair burden sharing. The above AHP problem is revised by add sub-sommeliers layer such as Figure 1.

Solution procedure of fair sharing problem

- Step1: Set the wine test problem to classic AHP chart.
- Step2: Define a physical limitation of decision maker
- Step3: Find necessary number of sub decision makers
- Step4: Apply our algorithm to find optimal design assigning each wine to each sub decision maker, fairly.

Table 5: Judgement of existence of design

no. of pairwise comparison (t)	required no. of sommeliers (b)	Physical limitation (k)	no. of wines (v)	no. of blocks of each wine (r)	Final Judgement (Fairness Condition)
3	1	3	3	1	Passed
6	2	3	4	1.5	Failure
10	3.33	3	5	2	Failure
15	5	3	6	2.5	Failure
21	7	3	7	3	Passed
28	9.33	3	8	3.5	Failure
36	12	3	9	4	Passed
45	15	3	10	4.5	Failure

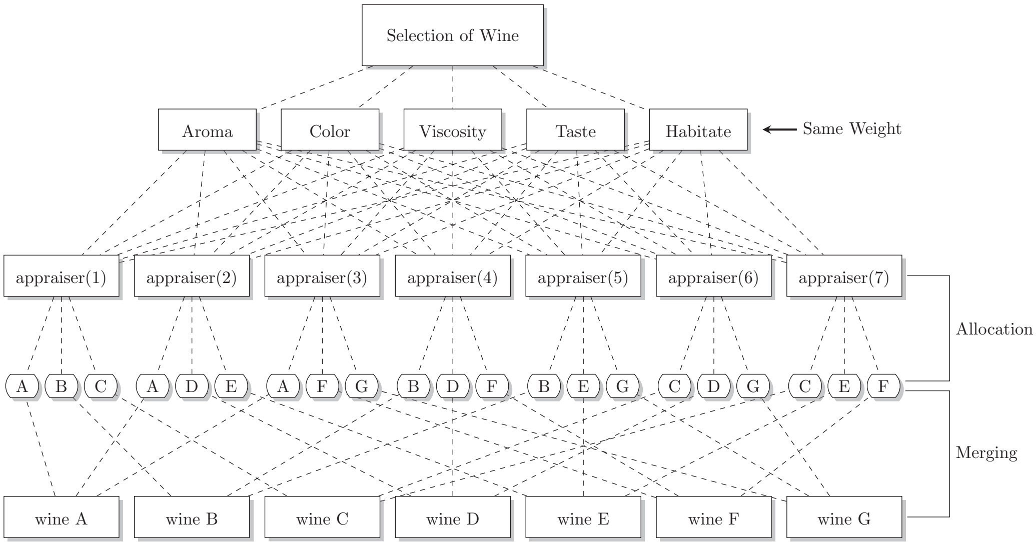

Figure 1: Result of assignment in AHP

- Step5: Revise the classic AHP chart by adding layer of sub decision makers
- Step6: Solve the problem by Group AHP

For the above solution procedure, computational complexity and validity are clear based on theory of combinatorial design and AHP.

6 Theoretical Approach and Methodology Let us consider optimal designs for $k = 3$ and $v = 7$. Identifying matrixes which become the same by permutation of their rows, we have 30 incidence matrixes as in Table 6.

Assume that there are 15 brands of wine ($v = 15$)

$$\mathrm{A, B, C, D, E, F, G, H, I, J, K, L, M, N, O.}$$

Table 6: All incidence matrixes for the case $k = 3$ and $v = 7$

$$\begin{pmatrix} 1&1&1&0&0&0&0\\ 1&0&0&1&1&0&0\\ 1&0&0&0&0&1&1\\ 0&1&0&1&0&1&0\\ 0&1&0&0&1&0&1\\ 0&0&1&1&0&0&1\\ 0&0&1&0&1&1&0 \end{pmatrix} \begin{pmatrix} 1&1&1&0&0&0&0\\ 1&0&0&1&1&0&0\\ 1&0&0&0&0&1&1\\ 0&1&0&1&0&0&1\\ 0&1&0&0&1&1&0\\ 0&0&1&1&0&1&0\\ 0&0&1&0&1&0&1 \end{pmatrix} \begin{pmatrix} 1&1&1&0&0&0&0\\ 1&0&0&1&0&1&0\\ 1&0&0&0&1&0&1\\ 0&1&0&1&1&0&0\\ 0&1&0&0&0&1&1\\ 0&0&1&1&0&0&1\\ 0&0&1&0&1&1&0 \end{pmatrix}$$

$$\begin{pmatrix} 1&1&1&0&0&0&0\\ 1&0&0&1&0&1&0\\ 1&0&0&0&1&0&1\\ 0&1&0&1&0&0&1\\ 0&1&0&0&1&1&0\\ 0&0&1&1&1&0&0\\ 0&0&1&0&0&1&1 \end{pmatrix} \begin{pmatrix} 1&1&1&0&0&0&0\\ 1&0&0&1&0&0&1\\ 1&0&0&0&1&1&0\\ 0&1&0&1&1&0&0\\ 0&1&0&0&0&1&1\\ 0&0&1&1&0&1&0\\ 0&0&1&0&1&0&1 \end{pmatrix} \begin{pmatrix} 1&1&1&0&0&0&0\\ 1&0&0&1&0&0&1\\ 1&0&0&0&1&1&0\\ 0&1&0&1&0&1&0\\ 0&1&0&0&1&0&1\\ 0&0&1&1&1&0&0\\ 0&0&1&0&0&1&1 \end{pmatrix}$$

$$\begin{pmatrix} 1&1&0&1&0&0&0\\ 1&0&1&0&1&0&0\\ 1&0&0&0&0&1&1\\ 0&1&1&0&0&1&0\\ 0&1&0&0&1&0&1\\ 0&0&1&1&0&0&1\\ 0&0&0&1&1&1&0 \end{pmatrix} \begin{pmatrix} 1&1&0&1&0&0&0\\ 1&0&1&0&1&0&0\\ 1&0&0&0&0&1&1\\ 0&1&1&0&0&0&1\\ 0&1&0&0&1&1&0\\ 0&0&1&1&0&1&0\\ 0&0&0&1&1&0&1 \end{pmatrix} \begin{pmatrix} 1&1&0&1&0&0&0\\ 1&0&1&0&0&1&0\\ 1&0&0&0&1&0&1\\ 0&1&1&0&1&0&0\\ 0&1&0&0&0&1&1\\ 0&0&1&1&0&0&1\\ 0&0&0&1&1&1&0 \end{pmatrix}$$

$$\begin{pmatrix} 1&1&0&1&0&0&0\\ 1&0&1&0&0&1&0\\ 1&0&0&0&1&0&1\\ 0&1&1&0&0&0&1\\ 0&1&0&0&1&1&0\\ 0&0&1&1&1&0&0\\ 0&0&0&1&0&1&1 \end{pmatrix} \begin{pmatrix} 1&1&0&1&0&0&0\\ 1&0&1&0&0&0&1\\ 1&0&0&0&1&1&0\\ 0&1&1&0&1&0&0\\ 0&1&0&0&0&1&1\\ 0&0&1&1&0&1&0\\ 0&0&0&1&1&0&1 \end{pmatrix} \begin{pmatrix} 1&1&0&1&0&0&0\\ 1&0&1&0&0&0&1\\ 1&0&0&0&1&1&0\\ 0&1&1&0&0&1&0\\ 0&1&0&0&1&0&1\\ 0&0&1&1&1&0&0\\ 0&0&0&1&0&1&1 \end{pmatrix}$$

$$\begin{pmatrix} 1&1&0&0&1&0&0\\ 1&0&1&1&0&0&0\\ 1&0&0&0&0&1&1\\ 0&1&1&0&0&1&0\\ 0&1&0&1&0&0&1\\ 0&0&1&0&1&0&1\\ 0&0&0&1&1&1&0 \end{pmatrix} \begin{pmatrix} 1&1&0&0&1&0&0\\ 1&0&1&1&0&0&0\\ 1&0&0&0&0&1&1\\ 0&1&1&0&0&0&1\\ 0&1&0&1&0&1&0\\ 0&0&1&0&1&1&0\\ 0&0&0&1&1&0&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&1&0&0\\ 1&0&1&0&0&1&0\\ 1&0&0&1&0&0&1\\ 0&1&1&1&0&0&0\\ 0&1&0&0&0&1&1\\ 0&0&1&0&1&0&1\\ 0&0&0&1&1&1&0 \end{pmatrix}$$

$$\begin{pmatrix} 1&1&0&0&1&0&0\\ 1&0&1&0&0&1&0\\ 1&0&0&1&0&0&1\\ 0&1&1&0&0&0&1\\ 0&1&0&1&0&1&0\\ 0&0&1&1&1&0&0\\ 0&0&0&0&1&1&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&1&0&0\\ 1&0&1&0&0&0&1\\ 1&0&0&1&0&1&0\\ 0&1&1&1&0&0&0\\ 0&1&0&0&0&1&1\\ 0&0&1&0&1&1&0\\ 0&0&0&1&1&0&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&1&0&0\\ 1&0&1&0&0&0&1\\ 1&0&0&1&0&1&0\\ 0&1&1&0&0&1&0\\ 0&1&0&1&0&0&1\\ 0&0&1&1&1&0&0\\ 0&0&0&0&1&1&1 \end{pmatrix}$$

$$\begin{pmatrix} 1&1&0&0&0&1&0\\ 1&0&1&1&0&0&0\\ 1&0&0&0&1&0&1\\ 0&1&1&0&1&0&0\\ 0&1&0&1&0&0&1\\ 0&0&1&0&0&1&1\\ 0&0&0&1&1&1&0 \end{pmatrix} \begin{pmatrix} 1&1&0&0&0&1&0\\ 1&0&1&1&0&0&0\\ 1&0&0&0&1&0&1\\ 0&1&1&0&0&0&1\\ 0&1&0&1&1&0&0\\ 0&0&1&0&1&1&0\\ 0&0&0&1&0&1&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&0&1&0\\ 1&0&1&0&1&0&0\\ 1&0&0&1&0&0&1\\ 0&1&1&1&0&0&0\\ 0&1&0&0&1&0&1\\ 0&0&1&0&0&1&1\\ 0&0&0&1&1&1&0 \end{pmatrix}$$

$$\begin{pmatrix} 1&1&0&0&0&1&0\\ 1&0&1&0&1&0&0\\ 1&0&0&1&0&0&1\\ 0&1&1&0&0&0&1\\ 0&1&0&1&1&0&0\\ 0&0&1&1&0&1&0\\ 0&0&0&0&1&1&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&0&1&0\\ 1&0&1&0&0&0&1\\ 1&0&0&1&1&0&0\\ 0&1&1&1&0&0&0\\ 0&1&0&0&1&0&1\\ 0&0&1&0&1&1&0\\ 0&0&0&1&0&1&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&0&1&0\\ 1&0&1&0&0&0&1\\ 1&0&0&1&1&0&0\\ 0&1&1&0&1&0&0\\ 0&1&0&1&0&0&1\\ 0&0&1&1&0&1&0\\ 0&0&0&0&1&1&1 \end{pmatrix}$$

$$\begin{pmatrix} 1&1&0&0&0&0&1\\ 1&0&1&1&0&0&0\\ 1&0&0&0&1&1&0\\ 0&1&1&0&1&0&0\\ 0&1&0&1&0&1&0\\ 0&0&1&0&0&1&1\\ 0&0&0&1&1&0&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&0&0&1\\ 1&0&1&1&0&0&0\\ 1&0&0&0&1&1&0\\ 0&1&1&0&0&1&0\\ 0&1&0&1&1&0&0\\ 0&0&1&0&1&0&1\\ 0&0&0&1&0&1&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&0&0&1\\ 1&0&1&0&1&0&0\\ 1&0&0&1&0&1&0\\ 0&1&1&1&0&0&0\\ 0&1&0&0&1&1&0\\ 0&0&1&0&0&1&1\\ 0&0&0&1&1&0&1 \end{pmatrix}$$

$$\begin{pmatrix} 1&1&0&0&0&0&1\\ 1&0&1&0&1&0&0\\ 1&0&0&1&0&1&0\\ 0&1&1&0&0&1&0\\ 0&1&0&1&1&0&0\\ 0&0&1&1&0&0&1\\ 0&0&0&0&1&1&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&0&0&1\\ 1&0&1&0&0&1&0\\ 1&0&0&1&1&0&0\\ 0&1&1&1&0&0&0\\ 0&1&0&0&1&1&0\\ 0&0&1&0&1&0&1\\ 0&0&0&1&0&1&1 \end{pmatrix} \begin{pmatrix} 1&1&0&0&0&0&1\\ 1&0&1&0&0&1&0\\ 1&0&0&1&1&0&0\\ 0&1&1&0&1&0&0\\ 0&1&0&1&0&1&0\\ 0&0&1&1&0&0&1\\ 0&0&0&0&1&1&1 \end{pmatrix}$$

Since $bk = {}_v\mathrm{C}_2 = 105$, we needs 35 sommeliers. An example of the optimal designs is

$$\{\mathrm{A,B,C}\},\{\mathrm{A,D,E}\},\{\mathrm{A,F,G}\},\{\mathrm{A,H,I}\},\{\mathrm{A,J,K}\},\{\mathrm{A,L,M}\},\{\mathrm{A,N,O}\},$$
$$\{\mathrm{B,D,F}\},\{\mathrm{B,E,G}\},\{\mathrm{B,H,J}\},\{\mathrm{B,I,K}\},\{\mathrm{B,L,N}\},\{\mathrm{B,M,O}\},\{\mathrm{C,D,G}\},$$
$$\{\mathrm{C,E,F}\},\{\mathrm{C,H,K}\},\{\mathrm{C,I,J}\},\{\mathrm{C,L,O}\},\{\mathrm{C,M,N}\},\{\mathrm{D,H,L}\},\{\mathrm{D,I,M}\},$$
$$\{\mathrm{D,J,N}\},\{\mathrm{D,K,O}\},\{\mathrm{E,H,M}\},\{\mathrm{E,I,L}\},\{\mathrm{E,J,O}\},\{\mathrm{E,K,N}\},\{\mathrm{F,H,N}\},$$
$$\{\mathrm{F,I,O}\},\{\mathrm{F,J,L}\},\{\mathrm{F,K,M}\},\{\mathrm{G,H,O}\},\{\mathrm{G,I,N}\},\{\mathrm{G,J,M}\},\{\mathrm{G,K,L}\}$$

and the corresponding incidence matrix is shown in Table 7.

It is very troublesome to do this procedure manually. We intend to apply these matrixes to results of questionnaires and verify the effectiveness of our methods by simulating allocation of the decision-makers. Threfore, we develop a computer program for this purpose.

In this section, we only consider the cases where v and k satisfy the fair sharing condition (2).

Let W be a finite set with $\#(W) = v$ and $b = \frac{{}_v\mathrm{C}_2}{k}$. Optimal design $\mathcal{D}$ of W is a set of subsets of W, such that

1. $\#(\mathcal{D}) = b$,
2. $\#(X) = k$ for all $X \in \mathcal{D}$,
3. $\#(X \cap Y) \leq 1$ for all X, $Y \in \mathcal{D}$, $X \neq Y$.

Let $\mathcal{W} = \{X | X \subset W, \#(X) = k\}$, then, the set of all optimal designs is defined by

$$\Delta = \{\mathcal{D} | \mathcal{D} \subset \mathcal{W}, \#(\mathcal{D}) = b, \#(X \cap Y) \leq 1 \text{ for all } X, Y \in \mathcal{D}, X \neq Y\},$$

which we would like to enumerate as a sequence of incidence matrixes. However, this definition is not suitable for computer algorithms. So we have to rebuild Δ recursively.

We proceed with our argument on more general assumptions. Suppose a finite set $U = \{\alpha, \beta, \ldots\}$ and a symmetric relation $\sim$ between two elements of U are given.

For a nonnegative integer a and a subset $S \subset U$, we define $\Delta(S, a)$ by

$$\Delta(S, a) = \{A | A \subset S, \#(A) = a, \alpha \sim \beta \text{ for all } \alpha, \beta \in A, \alpha \neq \beta\}.$$

Notice that

$$\Delta(S, 0) = \{\emptyset\}$$

and if $\#(S) < a$

$$\Delta(S, a) = \emptyset.$$

Now we assume $S \neq \emptyset$ and $a > 0$. Choosing $\alpha \in S$ arbitrarily, $\Delta(S, a)$ is expressed as a disjoint union as follows:

$$(3) \qquad \Delta(S, a) = \{A | A \in \Delta(S, a), A \ni \alpha\} \cup \{A | A \in \Delta(S, a), A \not\ni \alpha\}.$$

For a subset $T \subset U$, we define by $\alpha \sim T$ the subset

$$\alpha \sim T = \{\beta | \beta \in T, \beta \sim \alpha\}.$$

Table 7: An example of incidence matrix for the case $k = 3$ and $v = 15$

$$
\begin{pmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0
\end{pmatrix}
$$

Then, the first term of the right hand side of (3) is given by

$$\{A|A \in \Delta(S,a), A \ni \alpha\} = \{\{\alpha\} \cup B | B \in \Delta(\alpha \sim (S \setminus \{\alpha\}), a-1)\},$$

since

$$\{A \setminus \{\alpha\} | A \in \Delta(S,a), A \ni \alpha\} = \Delta(\alpha \sim (S \setminus \{\alpha\}), a-1).$$

For the second term of the right hand side of (3), we have

$$\{A|A \in \Delta(S,a), A \not\ni \alpha\} = \Delta(S \setminus \{\alpha\}, a),$$

and accordingly

$$\Delta(S,a) = \{\{\alpha\} \cup B | B \in \Delta(\alpha \sim (S \setminus \{\alpha\}), a-1)\} \cup \Delta(S \setminus \{\alpha\}, a).$$

Now we obtain

$$\text{(4)} \quad \Delta(S,a) = \begin{cases} \{\emptyset\} & \text{if } a = 0, \\ \emptyset & \text{if } \#(S) < a, \\ \{\{\alpha\} \cup B | B \in \Delta(\alpha \sim (S \setminus \{\alpha\}), a-1)\} \cup \Delta(S \setminus \{\alpha\}, a) & \text{otherwise.} \end{cases}$$

By using the programming language Haskell, the above construction (4) can be easily translated to a program code as below. That is the reason why we adopt Haskell.

```
delta :: [Int] -> Int -> [[Int]]
delta _ 0 = [[]]
delta ss@(alpha:sa) a
  | length ss < a = []
  | otherwise = (map (alpha:) $ delta (alpha \~ sa) (a - 1))
                ++
                (delta sa a)
```

We present the whole code in Appendix.

7 Concluding Remarks In this study we considered the case where one decision maker cannot compare all entities due to the heavy burden of pairwise comparison work. The necessity to allocate the burden to multiple decision makers leads us to the group-AHP problem. To solve this problem, we proposed an algorithm to generate an incidence matrix based on combinatorial design theory. For simplicity, our algorithm covers limited cases in which the fair sharing condition is satisfied. We will discuss more general cases with real data and verify the efficiency of our method in our further research.

A Source Code Now we present the whole source code.

```
module Main where

import System.Environment (getArgs)
import Data.Array ((!), array, listArray)

generateVectors :: Int -> Int -> [[Int]]
generateVectors v k
    | k == 0 = [replicate v 0]
    | v == k = [replicate v 1]
    | v > k = (map (1:) $ generateVectors (v - 1) (k - 1))
              ++
```

```
              (map (0:) $ generateVectors (v - 1) k)
    | otherwise = error "generateVectors"

relationR :: [Int] -> [Int] -> Bool
relationR w w' = (sum $ zipWith (*) w w') <= 1

takeTheFirstFromDelta :: Int -> Int -> Int -> [[Int]]
takeTheFirstFromDelta b v k
    = indexesToVectors $ head $ delta indexSet b
    where
      vectors = generateVectors v k
      l = length vectors
      indexSet = [1 .. l]
      indexedVectors = listArray (1, l) vectors
      indexesToVectors = map (indexedVectors !)

      matrixOfR = array (1, l)
                  [ (n, q n) | n <- [1 .. l] ]
          where
            q n = array (n + 1, l)
                  [ (m, r n m) | m <- [n + 1 .. l] ]
            r n m = relationR
                    (indexedVectors ! n)
                    (indexedVectors ! m)

      a \~ s = filter (matrixOfR ! a !) s

      delta :: [Int] -> Int -> [[Int]]
      delta _ 0 = [[]]
      delta ss@(alpha:sa) a
        | length ss < a = []
        | otherwise = (map (alpha:) $ delta (alpha \~ sa) (a - 1))
                      ++
                      (delta sa a)
      delta [] _ = []

readArgs :: [String] -> (Int, Int, Int)
readArgs = p . map read
    where
      p [] = (1, 3, 3)
      p [v] = p [v, 3]
      p [v, k]
          | m == 0 && n == 0 = (b, v, k)
          | otherwise = error "readArgs"
          where
            c = v*(v-1) `quot` 2
            (b,m) = c `quotRem` k
            n = c `rem` v
      p _ = error "readArgs"

main :: IO ()
main = do
  (b, v, k) <- fmap readArgs getArgs
  let s = takeTheFirstFromDelta b v k
```

```
    t = foldr1 (zipWith (+)) s
putStrLn $ unlines $ map show s
putStrLn $ show t
```

```
[1,1,1,0,0,0,0,0,0,0,0,0,0]
[1,0,0,1,1,0,0,0,0,0,0,0,0]
[1,0,0,0,0,1,1,0,0,0,0,0,0]
[1,0,0,0,0,0,0,1,1,0,0,0,0]
[1,0,0,0,0,0,0,0,0,1,1,0,0]
[1,0,0,0,0,0,0,0,0,0,0,1,1]
[0,1,0,1,0,1,0,0,0,0,0,0,0]
[0,1,0,0,1,0,1,0,0,0,0,0,0]
[0,1,0,0,0,0,0,1,0,1,0,0,0]
[0,1,0,0,0,0,0,0,1,0,0,1,0]
[0,1,0,0,0,0,0,0,0,0,1,0,1]
[0,0,1,1,0,0,0,1,0,0,0,0,0]
[0,0,1,0,1,0,0,0,1,0,0,0,0]
[0,0,1,0,0,1,0,0,0,1,0,0,0]
[0,0,1,0,0,0,1,0,0,0,0,0,1]
[0,0,1,0,0,0,0,0,0,0,1,1,0]
[0,0,0,1,0,0,1,0,0,0,1,0,0]
[0,0,0,1,0,0,0,0,1,0,0,0,1]
[0,0,0,1,0,0,0,0,0,1,0,1,0]
[0,0,0,0,1,1,0,0,0,0,0,1,0]
[0,0,0,0,1,0,0,1,0,0,1,0,0]
[0,0,0,0,1,0,0,0,0,1,0,0,1]
[0,0,0,0,0,1,0,1,0,0,0,0,1]
[0,0,0,0,0,1,0,0,1,0,1,0,0]
[0,0,0,0,0,0,1,1,0,0,0,1,0]
[0,0,0,0,0,0,1,0,1,1,0,0,0]

[6,6,6,6,6,6,6,6,6,6,6,6,6]
```

Figure 2: Computed result of our algorithm by haskell

References

Bolster, Paul J., Vahan Janjigian, and Emery A. Trahan (1995) "Determining investor suitability using the analytic hierarchy process," *Financial Analysts Journal*, pp. 63–75.

Cook, Wade D., Boaz Golany, Moshe Kress, Michal Penn, and Tal Raviv (2005) "Optimal allocation of proposals to reviewers to facilitate effective ranking," *Management Science*, Vol. 51, No. 4, pp. 655–661.

Cowan, Nelson (2001) "The magical number 4 in short-term memory: a reconsideration of mental storage capacity," *Behavioral and Brain Sciences*, Vol. 24, No. 1, pp. 87–114.

Durer, S., M. Ahlatcioglu, and F. Tiryaki (1994) "Ideal menkul kiymet portföyü olusturmada analitik hiyerarsi yaklasimi," in *Arastırma Sempozyumu'94, asım 1994, Devlet Istatistik Enstitüsü Konferans Salonu*, pp. 21–23, Ankara, Turkey.

Eason, George, Benjamin Noble, and I. N. Sneddon (1955) "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 247, No. 935, pp. 529–551.

Fisher, Ronald Aylmer (1937) *The design of experiments*, Edinburgh; London: Oliver And Boyd.

Han, Sangheon (2014) "A practical approaches to decrease the consistency index in AHP," in *proceeding paper in SCIS & ISIS 2014*, Kitakyushu, Japan.

Ishii, Goro (1972) *Theory of Design of Experiments*, Tokyo: Baihukan.

Maxwell, James Clerk (1892) *A treatise on electricity and magnetism*, Vol. 2: Clarendon Press, 3rd edition.

Mazur, David R. (2010) *Combinatorics: a guided tour*: The Mathematical Association of America.

Miller, George A. (1956) "The magical number seven, plus or minus two: some limits on our capacity for processing information.," *Psychological Review*, Vol. 63, No. 2, p. 81.

Mirkin, Boris G. (1979) *Group Choice*, New York: John Wiley & Sons.

Oyatoye, E. O., G. U. Okpokpo, and G. A. Adekoya (2010) "An application of analytic hierarchy process (AHP) to investment portfolio selection in the banking sector of the Nigerian capital market," *Journal of Economics and International Finance*, Vol. 2, No. 12, pp. 321–335.

Rosa, Alexander (1991) "Combinational design theory — notes," `http://kam.mff.cuni.cz/~atcagc14/materialy/DesignTheoryNotes.pdf`.

Saaty, Thomas L. (1980) *The analytic hierarchy process*, New York: McGraw-Hill, Inc.

——— (1986) "Axiomatic foundation of the analytic hierarchy process," *Management Science*, Vol. 32, No. 7, pp. 841–855.

——— (1994) "How to make a decision: the analytic hierarchy process," *Interfaces*, Vol. 24, No. 6, pp. 19–43.

——— (2000) *Fundamentals of decision making and priority theory with the analytic hierarchy process*, Vol. 6: Rws Publications.

Saaty, Thomas L. and Luis G. Vargas (2012) *Models, methods, concepts & applications of the analytic hierarchy process*, Vol. 175: Springer Science & Business Media.

Thalheimer, Richard and Mukhtar M. Ali (1979) "Time series analysis and portfolio selection: an application to mutual savings bank," *Southern Economic Journal*, Vol. 45, No. 3, pp. 821–837.

Tiryaki, Fatma (2001) "The use of data envelopment analysis for stocks selection on Istanbul stock exchange," in *Management of Engineering and Technology, 2001. PICMET'01. Portland International Conference on*, Vol. 1, pp. 373–vol, IEEE.

Communicated by *Yoshio Tabata*

Submission to the SCMJ

In September 2012, the way of submission to Scientiae Mathematicae Japonicae (SCMJ) was changed. Submissions should be sent electronically (in PDF file) to the editorial office of International Society for Mathematical Sciences (ISMS).

(1) Preparation of files and Submission

a. Authors who would like to submit their papers to the SCMJ should make source files of their papers in LaTeX2e using the ISMS style file (scmjlt2e.sty) Submissions should be in PDF file compiled from the source files. Send the PDF file to s1bmt@jams.jp .

b. Prepare a Submission Form and send it to the ISMS. The required items to be contained in the form are:

1. Editor's name whom the author chooses from the Editorial Board (http://www.jams.or.jp/hp/submission_f.html)and would like to take in charge of the paper for refereeing.
2. Title of the paper.
3. Authors' names.
4. Corresponding author's name, e-mail address and postal address (affiliation).
5. Membership number in case the author is an ISMS member.

Japanese authors should write 3 and 4 both in English and in Japanese.

At http://www.jams.or.jp/hp/submission_f.html, the author can find the Submission Form. Fulfill the Form and sent it to the editorial office by pushing the button "transmission". Or, without using the Form, the author may send an e-mail containing the items 1-5 to s1bmt@jams.jp

(2) Registration of Papers

When the editorial office receives both a PDF file of a submitted paper and a Submission Form, we register the paper. We inform the author of the registration number and the received date. At the same time, we send the PDF file to the editor whom the author chooses in the Submission Form and request him/her to begin the process of refereeing. (Authors need not send their papers to the editor they choose.)

(3) Reviewing Process

a. The editor who receives, from the editorial office, the PDF file and the request of starting the reviewing process, he/she will find an appropriate referee for the paper.

b. The referee sends a report to the editor. When revision of the paper is necessary, the editor informs the author of the referee's opinion.

c. Based on the referee report, the editor sends his/her decision (acceptance of rejection) to the editorial office.

(4) a. Managing Editor of the SCMJ makes the final decision to the paper valuing the editor's decision, and informs it to the author.

b. When the paper is accepted, we ask the author to send us a source file and a PDF file of the final manuscript.

c. The publication charges for the ISMS members are free if the membership dues have been paid without delay. If the authors of the accepted papers are not the ISMS members, they should become ISMS members and pay ¥6,000 (US$75, Euro55) as the membership dues for a year, or should just pay the same amount without becoming the members.

Items required in Submission Form

1. Editor's name who the authors wish will take in charge of the paper
2. Title of the paper
3. Authors' names

3'. 3. in Japanese for Japanese authors

4. Corresponding author's name and postal address (affiliation)

4'. 4. in Japanese for Japanese authors

5. ISMS membership number
6. E-mail address

Call for ISMS Members

Call for Academic and Institutional Members

Discounted subscription price: When organizations become the Academic and Institutional Members of the ISMS, they can subscribe our journal Scientiae Mathematicae Japonicae at the yearly price of US$225. At this price, they can add the subscription of the online version upon their request.

Invitation of two associate members: We would like to invite two persons from the organizations to the associate members with no membership fees. The two persons will enjoy almost the same privileges as the individual members. Although the associate members cannot have their own ID Name and Password to read the online version of SCMJ, they can read the online version of SCMJ at their organization.

To apply for the Academic and Institutional Member of the ISMS, please use the following application form.

Application for Academic and Institutional Member of ISMS

Subscription of SCMJ Check one of the two.	☐Print (US$225) ☐Print ＋ Online (US$225)
University (Institution)	
Department	
Postal Address where SCMJ should be sent	
E-mail address	
Person in charge	Name: Signature:
Payment Check one of the two.	☐Bank transfer ☐Credit Card (Visa, Master)
Name of Associate Membership	1.
	2.

Call for Individual Members

We call for individual members. The privileges to them and the membership dues are shown in "**Join ISMS !**" on the inside of the back cover.

Items required in Membership Application Form

1. Name
2. Birth date
3. Academic background
4. Affiliation
5. 4's address
6. Doctorate
7. Contact address
8. E-mail address
9. Special fields
10. Membership category (See **Table 1** in "**Join ISMS !**")

Individual Membership Application Form

1. **Name**	
2. **Birth date**	
3. **Academic background**	
4. **Affiliation**	
5. **4's address**	
6. **Doctorate**	
7. **Contact address**	
8. **E-mail address**	
9. **Special fields**	
10. **Membership category**	

Contributions (Gift to the ISMS)

We deeply appreciate your generous contributions to support the activities of our society.

The donation are used (1) to make medals for the new prizes (Kitagawa Prize, Kunugi Prize, and ISMS Prize), (2) to support the IVMS at Osaka University Nakanoshima Center, and (3) for a special fund designated by the contributors.

Your remittance to the following accounts of ours will be very much appreciated.

(1) Through a post office, remit to our giro account (in Yen only):
No. 00930-1-11872, Japanese Association of Mathematical Sciences (JAMS)
or send International Postal Money Order (in US Dollar or in Yen) to our address:
International Society for Mathematical Sciences
2-1-18 Minami Hanadaguchi, Sakai-ku, Sakai, Osaka 590-0075, Japan

(2) A/C 94103518, ISMS
CITIBANK, Japan Ltd., Shinsaibashi Branch
Midosuji Diamond Building
2-1-2 Nishi Shinsaibashi, Chuo-ku, Osaka 542-0086, Japan

****

Payment Instructions:

Payment can be made through a post office or a bank, or by credit card. Members may choose the most convenient way of remittance. Please note that we do not accept payment by bank drafts (checks). For more information, please refer to an invoice.

Methods of Overseas Payment:

Payment can be made through (1) a post office, (2) a bank, (3) by credit card, or (4) UNESCO Coupons.

Authors or members may choose the most convenient way of remittance as are shown below. Please note that **we do not accept payment by bank drafts (checks)**.

(1) Remittance through a post office to our giro account No. 00930-1-11872 or send International Postal Money Order to our postal address (2) Remittance through a bank to our account No. 94103518 at Shinsaibashi Branch of CITIBANK (3) **Payment by credit cards** (AMEX, VISA, MASTER or NICOS), or (4) Payment by UNESCO Coupons.

Methods of Domestic Payment:

Make remittance to:

(1) Post Office Transfer Account - 00930-3-73982 or

(2) Account No.7726251 at Sakai Branch, SUMITOMO MITSUI BANKING CORPORATION, Sakai, Osaka, Japan.

All of the correspondences concerning subscriptions, back numbers, individual and institutional memberships, should be addressed to the Publications Department, International Society for Mathematical Sciences.

Join ISMS !

ISMS Publications: We published **Mathematica Japonica (M.J.)** in print, which was first published in 1948 and has gained an international reputation in about sixty years, and its offshoot **Scientiae Mathematicae (SCM)** both online and in print. In January 2001, the two publications were unified and changed to **Scientiae Mathematicae Japonicae (SCMJ)**, which is the "21st Century New Unified Series of Mathematica Japonica and Scientiae Mathematicae" and published both online and in print. Ahead of this, the online version of SCMJ was first published in September 2000. The whole number of SCMJ exceeds 270, which is the largest amount in the publications of mathematical sciences in Japan. The features of SCMJ are:

1) About 80 eminent professors and researchers of not only Japan but also 20 foreign countries join the Editorial Board. The accepted papers are published both online and in print. SCMJ is reviewed by Mathematical Review and Zentralblatt from cover to cover.
2) SCMJ is distributed to many libraries of the world. The papers in SCMJ are introduced to the relevant research groups for the positive exchanges between researchers.
3) **ISMS Annual Meeting**: Many researchers of ISMS members and non-members gather and take time to make presentations and discussions in their research groups every year.

The privileges to the individual ISMS Members:

(1) No publication charges
(2) Free access (**including printing out**) to the online version of SCMJ
(3) Free copy of each printed issue

The privileges to the Institutional Members:

Two associate members can be registered, free of charge, from an institution.

Table 1: Membership Dues for 2019

Categories	Domestic	Overseas	Developing countries
1-year Regular member	￥8,000	US$80 , Euro75	US$50, Euro47
1-year Students member	￥4,000	US$50 , Euro47	US$30 , Euro28
Life member*	Calculated as below*	US$750 , Euro710	US$440, Euro416
Honorary member	Free	Free	Free

(Regarding submitted papers,we apply above presented new fee after April 15 in 2015 on registoration date.) * Regular member between 63 - 73 years old can apply the category.

(73 − age) × ￥3,000

Regular member over 73 years old can maintain the qualification and the privileges of the ISMS members, if they wish.

Categories of 3-year members were abolished.

CONTENTS

Notices from the ISMS