8-2024

# Probabilistic Frames and Concepts from Optimal Transport

Dongwei Chen
*Clemson University*, dongwec@clemson.edu

# Probabilistic Frames and Concepts from Optimal Transport

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Mathematical Science

by
Dongwei Chen
August 2024

Accepted by:
Dr. Martin Schmoll, Committee Chair
Dr. Mishko Mitkovski, Committee Co-Chair
Dr. Shitao Liu
Dr. Cody Stockdale

# Abstract

As the generalization of frames in the Euclidean space $\mathbb{R}^n$, a probabilistic frame is a probability measure on $\mathbb{R}^n$ that has a finite second moment and whose support spans $\mathbb{R}^n$. The p-Wasserstein distance with $p \geq 1$ from optimal transport is often used to compare probabilistic frames. It is particularly useful to compare frames of various cardinalities in the context of probabilistic frames. We show that the 2-Wasserstein distance appears naturally in the fundamental objects of frame theory and draws consequences leading to a geometric viewpoint of probabilistic frames.

We convert the classic lower bound estimates of 2-Wasserstein distance [Gelbrich, 1990, Cuesta-Albertos et al., 1996] from covariance operators to frame operators. As a consequence, we show that the sets of probabilistic frames with given frame operators are all homeomorphic, and the homeomorphism is an optimal transport map given by push-forward with a unique symmetric positive definite map that depends only on the two given frame operators. As an application, we generalize several recent results in probabilistic frames, such as finding the closest frame of a certain kind and showing the connectedness of the set of probabilistic frames with any given frame operator.

Furthermore, we consider the perturbation of probabilistic frames by generalizing the Paley–Wiener theorem to probabilistic frames. The Paley–Wiener Theorem is a classical result about the stability of a basis in a Banach space, claiming that if a sequence is close to a basis, then this sequence is also a basis. Similar results are extended to frames in Hilbert spaces. In this work, we generalize the Paley–Wiener theorem to probabilistic frames and claim that if a probability measure is close to a probabilistic frame in some sense, this probability measure is also a probabilistic frame.

In the end, we discuss some open problems about probabilistic frames. We first mention the minimization problem of probabilistic p-frames in the $p$-Wasserstein metric. Then we introduce probabilistic frames on the unit sphere $\mathbb{S}^{n-1}$. We show the existence of the closest probabilistic Parseval frame on $\mathbb{S}^{n-1}$ and give a lower bound estimate about the minimizing distance.

# Dedication

This dissertation is a gift to my family and friends.

# Acknowledgments

It has been such a long time since 2015 Spring when I decided to go to graduate school and pursue my PhD degree. After getting my Bachelor's degree in math from Fuzhou University, I went to the Chinese Academy of Sciences to study atmospheric turbulence, working with Dr. Fei Hu, who treated me as a 21-year-old young colleague. I visited École Polytechnique in the summer of 2017, which was my first trip abroad and motivated me to reconsider my academic career plan. Then I planned to pursue my PhD outside China, although it took me two years to get to Clemson.

During my PhD, I have met many people who encouraged me. Firstly, I must thank my advisor, Dr. Martin Schmoll. I learned optimal transport (which is my favorite) from Martin, and I am proud to be an optimal transporter. We have a weekly meeting to discuss research progress every Friday afternoon and it often takes me one hour to drive to Greenville for our meeting. Martin likes coffee and has very good taste in the cafe (and mathematics and skiing). We first met at Chestnut but it was always full of people. Then we switched to Cohesive. Sometimes, I am curious about other people's points of view in the coffee shop when they hear math words like Wasserstein distance from us.

I would also like to thank my co-advisor, Dr. Mishko Mitkovski, who played a role in the beginning of my research. In the fall of 2022, Mishko took me to Georgia Tech for a workshop on time-frequency analysis, where I found my current dissertation topic in probabilistic frames. One month later, I made progress and got the first lemma in my life. I also benefited greatly from taking courses with Mishko, who could lecture smoothly without lecture notes.

In addition, I would also like to thank Dr. Cody Stockdale and Dr. Shitao Liu. I used to attend Cody's reading group in the fall of 2022, where I learned a picture about harmonic analysis. Then in the spring of 2023, we continued the weekly morning meeting by using twenty minutes to share our research progress. In one of the early mornings when Trevor was absent, Cody moved our meeting to his office, and he suggested me to consider the perturbation of probabilistic frames, which now consists of the third chapter in my dissertation. I often meet Shitao in Schilletter and I appreciate every discussion we have. Shitao is well-known for his clear lecturing among graduate students, and I verified this by conducting a functional analysis with him in the fall of 2022. I also used Shitao's linear analysis lecture notes to prepare the analysis prelim. I would also like to thank Cody and Shitao for their recommendation letters during my postdoc hunting.

Furthermore, I want to thank Dr. Brian Fralix for his encouragement. I used Fralix's lecture notes for my stochastic process prelim, and his online lectures on probability theory during the COVID deepened my understanding of measure theory and optimal transport. I would like to thank Ms. Jennifer Brondell for her support in language improving, Ms. April Haynes for her help in class registration, and Ms. Jennifer Newton for her advising on teaching. I would also like to thank all my students in Business Calculus I and Business Calculus II.

I would like to thank my graduate fellows at Clemson, Trevor Camper and David Frost, for organizing the stochastic analysis reading group and mathematical smoking gathering, Haodong Li and Deborpita Biswas for our discussion on frames, and Yu Huang, Xuejian Li, and Tianyu Ni for our usual gathering.

I would like to thank all my old friends who always encouraged me. In particular, I would like to thank Qingsong Zhang, who tragically passed away in the 2014 Winter, and Lina Han, who encouraged and accompanied me during my tough time in 2018.

I would like to thank all my collaborators and those who encouraged me during my PhD. In particular, I would like to thank Dr. Kasso Okoudjou for visiting him at Tufts and Dr. Christopher Heil for his suggestion during my job hunting. I would also like to thank Dr. Emily King and Dr. Clayton Shonkwiler for offering me the postdoc position at Colorado State University.

Finally, I would like to thank my parents, Guobao Chen and Xianrong Wang, my sister Shumei Chen; my brother in law Leilei Liang; my nephews Yangyang Chen, and my nieces, Yifei Liang, Yilin Liang, and Yichen Liang. I would like to thank Dr. Tiantian Yang, who also graduated from Clemson in the summer of 2021. Lemma 3.3.3 is named Sweetie's Lemma since I proved it while talking to her on the phone at midnight in October 2023.

May the Fortune and Prosperity Be with All of You.

# Notation

$\mathscr{H}$: a separable Hilbert space.

$\mathbb{R}^n$: the $n$-dimensional Euclidean space.

$\mathbb{S}^{n-1}$: the unit sphere in the $n$-dimensional Euclidean space.

$\mathbf{Id}$: the $n \times n$ identity matrix in $\mathbb{R}^n$.

$\mathbf{f}^t$: the transpose of a column vector $\mathbf{f}$ in $\mathbb{R}^n$.

$\mathscr{P}(\mathbb{R}^n)$: the set of Borel probability measures on $\mathbb{R}^n$.

$M_p(\mu)$: the $p$-moments of probability measure $\mu$ where $p \in [1, \infty)$.

$\mathscr{P}_p(\mathbb{R}^n)$: the set of Borel probability measures on $\mathbb{R}^n$ with finite $p$-moments where $p \in [1, \infty)$.

$supp(\mu)$: the support of probability measure $\mu$.

$f_{\#}\mu$: the pushforward of $\mu$ by a measurable map $f$.

$\mathbf{A}_{\#}\mu$: the pushforward of $\mu$ by the linear map given by the matrix $\mathbf{A}$.

$\Gamma(\mu, \nu)$: the set of transport couplings with marginals $\mu$ and $\nu$.

$W_p(\mu, \nu)$: p-Wasserstein distance between $\mu \in \mathscr{P}_p(\mathbb{R}^n)$ and $\nu \in \mathscr{P}_p(\mathbb{R}^n)$ where $p \in [1, \infty)$.

$\mathbf{S}_\mu$: the frame operator of $\mu$.

$\|\mathbf{S}_\mu\|_2$: the 2-matrix norm of $\mathbf{S}_\mu$.

$\mu \otimes \nu$: the product measure of $\mu$ and $\nu$.

$\mathscr{T}_{\mathscr{P}}(\mathbb{R}^n)/\mathscr{T}(\mathbb{R}^n)$: the set of Parseval/tight probabilistic frames in $\mathbb{R}^n$.

$\pi_{\mathbf{x}^\perp}$: the projection on the subspace orthogonal to the (unit) vector $\mathbf{x}$.

$\mathbb{S}^n_+$: the set of real $n \times n$ positive semi-definite symmetric matrices.

$\mathbb{S}^n_{++}$: the set of real $n \times n$ positive definite symmetric matrices.

$\mathscr{P}_{++}$: the set of probabilistic frames on $\mathbb{R}^n$.

$\mathrm{tr}(\mathbf{A})$: the trace of matrix $\mathbf{A}$.

$C_c(\mathbb{R}^n)$: the set of continuous functions on $\mathbb{R}^n$ with compact support.

# Table of Contents

**Title Page** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **i**

**Abstract** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **ii**

**Dedication** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **iii**

**Acknowledgments** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **iv**

**Notation** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **vi**

**1 Introduction and Background** . . . . . . . . . . . . . . . . . . . . . . . . . . . **1**
1.1 Introduction to Frame Theory . . . . . . . . . . . . . . . . . . . . . . . . . . 1
1.2 Background on Optimal Transport . . . . . . . . . . . . . . . . . . . . . . . 3
1.3 Preliminaries to Probabilistic Frames . . . . . . . . . . . . . . . . . . . . . . 6
1.4 Summary of Main Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9

**2 Probabilistic Frames and Wasserstein Distances** . . . . . . . . . . . . . . . . . **11**
2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
2.2 Applications of the Main Results . . . . . . . . . . . . . . . . . . . . . . . . 14
2.3 Wasserstein Openness of the Set of Probabilistic p-Frames. . . . . . . . . . . 17
2.4 Wasserstein Distances for Frames: from Standard Estimates to Uniqueness . . . . . 21
2.5 Connection to Covariance and Centered Measures. . . . . . . . . . . . . . . . 31

**3 Probabilistic Frame Perturbations** . . . . . . . . . . . . . . . . . . . . . . . . . **34**
3.1 The Paley–Wiener Theorem for Frames . . . . . . . . . . . . . . . . . . . . . 34
3.2 Invertibility of Linear Operators on Banach Spaces . . . . . . . . . . . . . . 36
3.3 The Paley-Wiener Theorem for Probabilistic Frames . . . . . . . . . . . . . . 37
3.4 Perturbations Including Probabilistic Dual Frames . . . . . . . . . . . . . . . 41
3.5 Quadratic Closeness of Probabilistic Frames . . . . . . . . . . . . . . . . . . 47

**4 Discussion** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **50**
4.1 Minimization of Probabilistic p-Frames in $W_p$ . . . . . . . . . . . . . . . . . 51
4.2 Minimization of Probabilistic Frames on the Unit Sphere . . . . . . . . . . . . 51

**Bibliography** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **53**

# Chapter 1

# Introduction and Background

## 1.1 Introduction to Frame Theory

In the study of vector spaces, one of the most important concepts is that of a basis, which provides us with an expansion of all vectors in terms of "elementary building blocks" and hereby helps us by reducing many questions concerning general vectors to similar questions concerning only the basis elements [Christensen, 2016]. However, being a basis is conditional: linear dependence between the elements is not allowed, and sometimes, we even require orthogonality with respect to an inner product. This makes looking for bases satisfying other conditions challenging, and thus we need a more flexible tool. Frames are such tools, which were first introduced by Duffin and Schaeffer in the context of nonharmonic analysis [Duffin and Schaeffer, 1952].

**Definition 1.1.1** (Frame in Hilbert Space)**.** *A sequence* $\{\mathbf{f}_i\}_{i=1}^{\infty}$ *in a separable Hilbert space* $\mathscr{H}$ *is said to be a frame for* $\mathscr{H}$ *if there exist* $0 < A \leq B < \infty$ *such that for any* $\mathbf{f} \in \mathscr{H}$*,*

$$A\|\mathbf{f}\|^2 \leq \sum_{i=1}^{\infty} |\langle \mathbf{f}, \mathbf{f}_i \rangle|^2 \leq B\|\mathbf{f}\|^2,$$

*where* $\|\cdot\|$ *is the norm induced by the inner product in* $\mathscr{H}$*.* $A$ *and* $B$ *are called lower and upper frame bounds. Furthermore,* $\{\mathbf{f}_i\}_{i=1}^{\infty}$ *is called a tight frame if* $A = B$ *and Parseval if* $A = B = 1$*.*

In the finite $n$-dimensional Euclidean space $\mathbb{R}^n$, the infinite summation in the frame condition becomes finite summation. That is to say, a collection of points $\{\mathbf{y}_i\}_{i=1}^{N}$ where $N \geq n$ are said to be a (finite) frame in $\mathbb{R}^n$ if there exist $0 < A \leq B < \infty$ such that for any $\mathbf{x} \in \mathbb{R}^n$,

$$A\|\mathbf{x}\|^2 \leq \sum_{i=1}^{N} |\langle \mathbf{x}, \mathbf{y}_i \rangle|^2 \leq B\|\mathbf{x}\|^2.$$

Intuitively, one could think about a frame as a basis for adding more elements, but linear independence between elements is not required. A frame for an inner product space also allows each vector to be written as a linear combination of the frame elements, which provides a way to reconstruct signals. In signal processing terminology, a basis could only provide a unique signal representation, while a frame provides a redundant, stable, and robust way of representing a signal. If some data are missing in the signal transmission process, the signal cannot be reconstructed using a basis. However, the missing signal could be reconstructed using a frame due to the linear dependence of frame elements.

Frames have many applications in pure and applied mathematics, such as the Kadison-Singer problem [Casazza, 2013] and directional statistics [Ehler and Galanis, 2011]. Due to the redundancy and robustness, frames are effective tools in signal processing, and have been applied in time-frequency analysis [Gröchenig, 2001], wavelet analysis [Daubechies, 1992], coding theory [Strohmer and Heath Jr, 2003], sampling theory [Eldar, 2003], wireless communication [Strohmer, 2001], image process [Kutyniok and Labate, 2012], compressed sensing [Naidu and Murthy, 2020], the design and analysis of filter banks [Fickus et al., 2013], and, more generally, in applied mathematics, computer science, and engineering. For a complete treatment of frame theory, see [Casazza, 2000, Casazza, 2001, Han, 2007, Casazza and Kutyniok, 2012, Casazza and Lynch, 2016, Christensen, 2016] for more information.

Let $\{\mathbf{f}_i\}_{i=1}^{\infty}$ be a frame for $\mathscr{H}$ with bounds $0 < A \leq B < \infty$. One of the important tools for studying frames is operator theory. There are four operators related to the frame $\{\mathbf{f}_i\}_{i=1}^{\infty}$: analysis operator, synthesis operator, Gram operator, and frame operator $\mathbf{S} : \mathscr{H} \to \mathscr{H}$ which is given by

$$\mathbf{S} : \mathscr{H} \to \mathscr{H}, \quad \mathbf{S}(\mathbf{f}) = \sum_{i=1}^{\infty} \langle \mathbf{f}, \mathbf{f}_i \rangle \mathbf{f}_i, \ \forall \ \mathbf{f} \in \mathscr{H}.$$

The frame operator $\mathbf{S} : \mathscr{H} \to \mathscr{H}$ plays an essential role in this work. Indeed, the frame operator $\mathbf{S} : \mathscr{H} \to \mathscr{H}$ associated with a frame is bounded linear and invertible. In $\mathbb{R}^n$, the frame operator corresponds to a symmetric positive definite matrix, i.e.,

$$\mathbf{S} = \sum_{i=1}^{N} \mathbf{f}_i \mathbf{f}_i^t \in \mathbb{S}_{++}^n$$

where $\mathbf{f}_i^t$ is the transpose of $\mathbf{f}_i$. Let $\mathbf{Id}$ be the $n \times n$ identity matrix in $\mathbb{R}^n$. By using the frame operator, we have the following characterization about finite frames in $\mathbb{R}^n$.

**Lemma 1.1.2** (Characterization of Finite Frames)**.** (1) $\{\mathbf{f}_i\}_{i=1}^N$ *is a frame for* $\mathbb{R}^n \Leftrightarrow$ *the frame operator* $\mathbf{S}$ *is positive definite* $\Leftrightarrow$ *span*$\{\mathbf{f}_1, \cdots, \mathbf{f}_N\} = \mathbb{R}^n$.

(2) $\{\mathbf{f}_i\}_{i=1}^N$ *is a tight frame with bound* $A > 0 \Leftrightarrow \mathbf{S} = A\ \mathbf{Id}$.

(3) $\{\mathbf{f}_i\}_{i=1}^N$ *is a Parseval frame* $\Leftrightarrow \mathbf{S} = \mathbf{Id}$.

There are two ways to reconstruct the signal using frames in signal processing. One way is to employ tight frames. If $\{\mathbf{f}_i\}_{i=1}^\infty$ is a tight frame for the separable Hilbert space $\mathscr{H}$ with the frame bound $0 < A < \infty$, then for any $\mathbf{f} \in \mathscr{H}$, we have the following reconstruction formula

$$\mathbf{f} = \frac{1}{A} \sum_{i=1}^{\infty} \langle \mathbf{f}, \mathbf{f}_i \rangle \, \mathbf{f}_i.$$

The other way to reconstruct the signal is through dual frames.

**Definition 1.1.3** (Dual Frames)**.** *Suppose* $\{\mathbf{f}_i\}_{i=1}^\infty$ *is a frame for the Hilbert space* $\mathscr{H}$ *with bounds* $0 < A \leq B < \infty$. *A frame* $\{\mathbf{g}_i\}_{i=1}^\infty$ *for* $\mathscr{H}$ *is said to be a dual frame of* $\{\mathbf{f}_i\}_{i=1}^\infty$ *if for any* $\mathbf{f} \in \mathscr{H}$,

$$\mathbf{f} = \sum_{i=1}^{\infty} \langle \mathbf{f}, \mathbf{g}_i \rangle \, \mathbf{f}_i = \sum_{i=1}^{\infty} \langle \mathbf{f}, \mathbf{f}_i \rangle \, \mathbf{g}_i.$$

An example of dual frames for $\{\mathbf{f}_i\}_{i=1}^\infty$ is $\{\mathbf{S}^{-1}\mathbf{f}_i\}_{i=1}^\infty$ with bounds $0 < \frac{1}{B} \leq \frac{1}{A} < \infty$, which is known as the canonical dual frame.

## 1.2 Background on Optimal Transport

This section introduces optimal transport and Wasserstein distance, which is used to quantify the distance and similarities between two probability measures [Figalli and Glaudo, 2021]. Furthermore, 2−Wasserstein distance is useful to quantify the distance between two finite frames with different cardinalities in the sense of probabilistic frame.

Before talking about optimal transport, we need the following notations and definitions. Let $\mathscr{P}(\mathbb{R}^n)$ be the set of Borel probability measures on $\mathbb{R}^n$ and for $p \in [1, \infty)$, let $\mathscr{P}_p(\mathbb{R}^n)$ be the set of Borel probability measures on $\mathbb{R}^n$ with finite $p$-moments $M_p(\mu)$ , i.e.,

$$\mathscr{P}_p(\mathbb{R}^n) = \{\mu \in \mathscr{P}(\mathbb{R}^n) : M_p(\mu) = \int_{\mathbb{R}^n} \|\mathbf{x}\|^p d\mu(\mathbf{x}) < +\infty\}.$$

**Definition 1.2.1** (Support of Probability Measure)**.** *Let $B_r(\mathbf{x})$ be the open ball centered at the vector $\mathbf{x}$ with radius $r > 0$. The support of $\mu \in \mathscr{P}(\mathbb{R}^n)$ is defined by*

$$supp(\mu) = \Big\{\mathbf{x} \in \mathbb{R}^n : \textit{for any } r > 0, \mu(B_r(\mathbf{x})) > 0\Big\}.$$

**Definition 1.2.2** (Pushforward)**.** *Let $M, N > 0$ and $\mu \in \mathscr{P}(\mathbb{R}^M)$. The pushforward of $\mu$ by a measurable map $f : \mathbb{R}^M \to \mathbb{R}^N$ is a probability measure in $\mathbb{R}^N$, which is denoted by $f_\# \mu$ and*

$$f_\#\mu(E) := (\mu \circ f^{-1})(E) = \mu\big(f^{-1}(E)\big) \textit{ for any Borel set } E \subset \mathbb{R}^N.$$

*If $f : \mathbb{R}^M \to \mathbb{R}^N$ is a linear map given by the matrix $\mathbf{A}$, then we use $\mathbf{A}_\#\mu$ to denote $f_\#\mu$. Furthermore, we have the change-of-variable formula:*

$$\int_{\mathbb{R}^N} g(\mathbf{y}) d(f_\#\mu)(\mathbf{y}) = \int_{\mathbb{R}^M} g(f(\mathbf{x})) d\mu(\mathbf{x}),$$

*where $g$ is measurable such that $g \in L^1(\mathbb{R}^N, f_\#\mu)$ and $g \circ f \in L^1(\mathbb{R}^M, \mu)$.*

Now, let us switch to optimal transport. In 1781, Gaspard Monge proposed the concept of optimal transport from one practical situation: if one transports soil to construct fortifications, what is the cheapest way to move the soil? Let $\mu$ and $\nu$ be two probability measures on $\mathbb{R}^n$. This scenario leads to the well-known Monge optimal transportation problem

$$\inf_{T_\#\mu = \nu} \int_{\mathbb{R}^n} c(\mathbf{x}, T(\mathbf{x}))\ d\mu(\mathbf{x}),$$

where $T : \mathbb{R}^n \to \mathbb{R}^n$ is a map and $c(\mathbf{x}, T(\mathbf{x}))$ is the cost of transporting unit mass from $\mathbf{x}$ to $T(\mathbf{x})$. Such map $T$ is called a transport map. Since $T$ is a map, the mass at $\mathbf{x}$ can only be transported to one destination, meaning Monge's formulation does not allow mass splitting.

In the 1940s, Leonid Kantorovich revisited Monge's problem by relaxing Monge's formulation to allow mass splitting.

$$\inf_{\gamma\in\Gamma(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} c(\mathbf{x},\mathbf{y})\ d\gamma(\mathbf{x},\mathbf{y}),$$

where $c(\mathbf{x},\mathbf{y})$ is the cost of transporting unit mass from $\mathbf{x}$ to $\mathbf{y}$, and $\Gamma(\mu,\nu)$ is the set of transport couplings with marginals $\mu$ and $\nu$, which is given by

$$\Gamma(\mu,\nu) = \Big\{\gamma \in \mathscr{P}(\mathbb{R}^n\times\mathbb{R}^n) : \pi_{x\#}\gamma = \mu,\ \pi_{y\#}\gamma = \nu\Big\},$$

and $\pi_x$ and $\pi_y$ are the projections on the corresponding coordinate, i.e., for any $(\mathbf{x},\mathbf{y}) \in \mathbb{R}^n\times\mathbb{R}^n$, $\pi_x(\mathbf{x},\mathbf{y}) = \mathbf{x}, \pi_y(\mathbf{x},\mathbf{y}) = \mathbf{y}$.

When the cost function $c(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^p$ where $p\in[1,\infty)$, the metric side of Kantorovich formulation makes it able to quantify the distance between probability measures $\mu\in\mathscr{P}_p(\mathbb{R}^n)$ and $\nu\in\mathscr{P}_p(\mathbb{R}^n)$ via p-Wasserstein distance, which is defined as below

$$W_p(\mu,\nu) = \left(\inf_{\gamma\in\Gamma(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} \|\mathbf{x}-\mathbf{y}\|^p\ \ d\gamma(\mathbf{x},\mathbf{y})\right)^{\frac{1}{p}} \tag{1.2.1}$$

As we mentioned in the later section, (probabilistic) frames for $\mathbb{R}^n$ could be seen as probability measures with finite second moments whose support span $\mathbb{R}^n$. Therefore, we could use 2-Wasserstein metric $W_2(\mu,\nu)$ to quantify the distance between two (probabilistic) frames $\mu$ and $\nu$.

In this work, we also need the gluing lemma, a standard tool in optimal transport to "glue" two transport couplings together. Similarly, let $\pi_x, \pi_y, \pi_{xy}, \pi_{yz}$ be projections on the corresponding coordinates, i.e., for any $(\mathbf{x},\mathbf{y},\mathbf{z})\in\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n$,

$$\pi_x(\mathbf{x},\mathbf{y},\mathbf{z}) = \mathbf{x},\ \pi_y(\mathbf{x},\mathbf{y},\mathbf{z}) = \mathbf{y},\ \pi_{xy}(\mathbf{x},\mathbf{y},\mathbf{z}) = (\mathbf{x},\mathbf{y}),\ \pi_{yz}(\mathbf{x},\mathbf{y},\mathbf{z}) = (\mathbf{y},\mathbf{z}).$$

**Lemma 1.2.3** (Gluing Lemma, [Figalli and Glaudo, 2021, pp.59])**.** *Let* $\mu_1,\mu_2,\mu_3\in\mathscr{P}_2(\mathbb{R}^n)$. *Suppose* $\gamma^{12}\in\Gamma(\mu_1,\mu_2)$ *and* $\gamma^{23}\in\Gamma(\mu_2,\mu_3)$ *such that* $\pi_{y\#}\gamma^{12} = \pi_{x\#}\gamma^{23} = \mu_2$. *Then there exists* $\gamma^{123}\in\mathscr{P}(\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n)$ *such that* $\pi_{xy\#}\gamma^{123} = \gamma^{12}$ *and* $\pi_{yz\#}\gamma^{123} = \gamma^{23}$.

## 1.3 Preliminaries to Probabilistic Frames

Recall that if $\{\mathbf{y}_i\}_{i=1}^N$ is a frame in $\mathbb{R}^n$ with bounds $0 < A \le B < \infty$, then for any $\mathbf{x} \in \mathbb{R}^n$,

$$A\|\mathbf{x}\|^2 \le \sum_{i=1}^{N} |\langle \mathbf{x}, \mathbf{y}_i \rangle|^2 \le B\|\mathbf{x}\|^2.$$

By taking $\mu_f = \sum_{i=1}^{N} \frac{1}{N}\delta_{\mathbf{y}_i} \in \mathscr{P}(\mathbb{R}^n)$, the definition of frame is equivalent to

$$\frac{A}{N}\|\mathbf{x}\|^2 \le \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{y} \rangle|^2 d\mu_f(\mathbf{y}) \le \frac{B}{N}\|\mathbf{x}\|^2, \text{ for any } \mathbf{x} \in \mathbb{R}^n.$$

This allows us to consider finite frames as discrete probabilistic measures on $\mathbb{R}^n$. Inspired by this insight, M. Ehler extended the above to a general probabilistic measure [Ehler, 2012], and we have the following definitions for probabilistic frames.

**Definition 1.3.1** (Probabilistic Frame, [Ehler and Okoudjou, 2013]). *$\mu \in \mathscr{P}(\mathbb{R}^n)$ is said to be a probabilistic frame for $\mathbb{R}^n$ if there exist $0 < A \le B < \infty$ such that for any $\mathbf{x} \in \mathbb{R}^n$,*

$$A\|\mathbf{x}\|^2 \le \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{y} \rangle|^2 d\mu(\mathbf{y}) \le B\|\mathbf{x}\|^2.$$

*$\mu$ is said to be a tight probabilistic frame if $A = B$ and Parseval if $A = B = 1$. Moreover, $\mu$ is said to be a Bessel probability measure if only the upper bound holds.*

By Cauchy-Schwartz inequality, it is easy to show that if $\mu \in \mathscr{P}_2(\mathbb{R}^n)$, then $\mu$ is a Bessel probability measure with upper bound $M_2(\mu)$. We could also define the frame operator for probabilistic frames, which is symmetric and positive definite.

**Definition 1.3.2** (Frame Operator). *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$. The frame operator $\mathbf{S}_\mu$ for $\mu$ is defined by*

$$\mathbf{S}_\mu := \int_{\mathbb{R}^n} \mathbf{y}\mathbf{y}^t d\mu(\mathbf{y}).$$

Then we have the following characterization for probabilistic frames.

**Proposition 1.3.3** (Characterization of Probabilistic Frames, Theorem 12.1 in [Ehler and Okoudjou, 2013], Proposition 3.1 in [Maslouhi and Loukili, 2019]). *Let $\mu \in \mathscr{P}(\mathbb{R}^n)$. Then we have*

(1) *$\mu$ is a probabilistic frame $\Leftrightarrow$ $\mathbf{S}_\mu$ is positive definite $\Leftrightarrow$ $\mu \in \mathscr{P}_2(\mathbb{R}^n)$ and span$\{supp(\mu)\} = \mathbb{R}^n$.*

(2) *$\mu$ is a tight probabilistic frame with bound $A > 0 \Leftrightarrow \mathbf{S}_\mu = A\ \mathbf{Id}$.*

(3) *$\mu$ is a Parseval probabilistic frame $\Leftrightarrow \mathbf{S}_\mu = \mathbf{Id}$.*

Note that if $\mu$ is tight with bound $A > 0$, then for any $\mathbf{f} \in \mathbb{R}^n$, we have the following reconstruction formula

$$\mathbf{f} = \frac{1}{A}\int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{x}\rangle\, \mathbf{x}\ d\mu(\mathbf{x}).$$

And if $\mu \in \mathscr{P}(\mathbb{R}^n)$ is a probabilistic frame with bounds $0 < A \leq B < \infty$, then $\mathbf{S}_\mu \in \mathbf{S}^n_{++}$. Let $\|\mathbf{S}_\mu\|_2$ denote the 2-matrix norm of $\mathbf{S}_\mu$. Since $\mathbf{S}_\mu$ is symmetric, then $\|\mathbf{S}_\mu\|_2$ is the largest eigenvalue of $\mathbf{S}_\mu$. Since each eigenvalue of $\mathbf{S}_\mu$ is within $A$ and $B$, then

$$A \leq \|\mathbf{S}_\mu\|_2 \leq B,\ \frac{1}{B} \leq \|\mathbf{S}_\mu^{-1}\|_2 \leq \frac{1}{A},\ \text{and}\ \frac{1}{\sqrt{B}} \leq \|\mathbf{S}_\mu^{-1/2}\|_2 \leq \frac{1}{\sqrt{A}}.$$

Similarly, one could define the probabilistic dual frames for a given probabilistic frame $\mu$.

**Definition 1.3.4** (Transport Duals, Definition 3.1, Theorem 3.6 and 3.7 in [Wickman, 2014])**.** *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$. The set of transport duals for $\mu$ is defined as*

$$D_\mu := \Big\{\nu \in \mathscr{P}_2(\mathbb{R}^n) : \exists\ \gamma \in \Gamma(\mu,\nu)\ s.t.\ \int_{\mathbb{R}^n\times\mathbb{R}^n} \mathbf{x}\mathbf{y}^t d\gamma(x,y) = \mathbf{Id}\Big\}.$$

*Furthermore, $D_\mu$ is not empty and a closed subset of $\mathscr{P}_2(\mathbb{R}^n)$ in the weak topology.*

**Proposition 1.3.5** ( Proposition 3.2 in [Wickman, 2014])**.** *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$ and take $\nu \in D_\mu$. Then $\nu$ is also a probabilistic frame.*

Therefore, the following is well-defined.

**Definition 1.3.6** (Probabilistic Dual Frame)**.** *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$. $\nu \in \mathscr{P}_2(\mathbb{R}^n)$ is called a probabilistic dual frame of $\mu$ with respect to $\gamma \in \Gamma(\mu,\nu)$ if*

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \mathbf{x}\mathbf{y}^t d\gamma(x,y) = \mathbf{Id}$$

**Remark 1.3.1.** *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$ with bounds $0 < A \leq B < \infty$. ${\mathbf{S}_\mu^{-1}}_{\#}\mu$ is said to be the canonical probabilistic dual frame of $\mu$ with bounds $0 < \frac{1}{B} \leq \frac{1}{A} < \infty$, since $\nu := {\mathbf{S}_\mu^{-1}}_{\#}\mu$ is the transport dual of $\mu$ with respect to $\gamma := (\mathbf{Id}\times\mathbf{S}_\mu^{-1})_{\#}\mu \in \Gamma(\mu,\nu)$.*

Suppose $\mu$ is a probabilistic frame for $\mathbb{R}^n$. There are two canonical probabilistic frames related to $\mu$: the canonical probabilistic Parseval frame ${\mathbf{S}_\mu^{-1/2}}_{\#}\mu$ and the canonical probabilistic dual frame ${\mathbf{S}_\mu^{-1}}_{\#}\mu$. Therefore, we have the following two reconstruction formulas. For any $\mathbf{f} \in \mathbb{R}^n$,

$$\mathbf{f} = \int_{\mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{S}_\mu^{-1/2}\mathbf{x} \right\rangle \mathbf{S}_\mu^{-1/2}\mathbf{x}\, d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1/2}\mathbf{f}, \mathbf{x} \right\rangle \mathbf{S}_\mu^{-1/2}\mathbf{x}\, d\mu(\mathbf{x}), \tag{1.3.1}$$

$$\mathbf{f} = \int_{\mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{S}_\mu^{-1}\mathbf{x} \right\rangle \mathbf{x}\, d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}\mathbf{f}, \mathbf{x} \right\rangle \mathbf{x}\, d\mu(\mathbf{x}). \tag{1.3.2}$$

After Ehler's starting work on probabilistic frames, M. Ehler and K.A. Okoudjou further expanded their work [Ehler and Okoudjou, 2012] and gave an overview of probabilistic frames [Ehler and Okoudjou, 2013]. They listed several open problems in probabilistic frames and rephrased the hyperplane conjecture in convex bodies proposed by Bourgain [Bourgain, 1986]. Under the guidance of K.A. Okoudjou, C.G. Wickman further studied some problems in frame theory by an optimal transport approach in her dissertation [Wickman, 2014]. Wickman uses Otto's calculus to construct gradient flows for the probabilistic p-frame potential and gives a more general definition of the concepts of duality, analysis, and synthesis in frame theory [Wickman, 2014, Lau and Okoudjou, 2015, Wickman and Okoudjou, 2017, Wickman and Okoudjou, 2023]. Some equalities and inequalities about probabilistic frames were also shown [Li et al., 2016].

Furthermore, M. Maslouhi and S. Loukili established a 1-1 correspondence between tight probabilistic frames and positive operator-valued measures [Maslouhi and Loukili, 2019] and answered an open problem in [Ehler and Okoudjou, 2013]. Distance minimizing problem over a given class of probabilistic frames in the Wasserstein metric is an open and active topic. In [Cheng, 2018, Cheng and Okoudjou, 2019], the authors showed that the canonical probabilistic Parseval frame is the unique closest probabilistic Parseval frame to a given probabilistic frame: let $\mu$ be a probabilistic frame in $\mathbb{R}^n$ and $\mathcal{T}_{\mathcal{P}}(\mathbb{R}^n)$ the set of Parseval probabilistic frames in $\mathbb{R}^n$, then the the following minimizing problem

$$I_{\mathcal{P}}(\mu, \mathbb{R}^n) := \inf_{\nu \in \mathcal{T}_{\mathcal{P}}(\mathbb{R}^n)} W_2(\mu, \nu)$$

admits an unique optimizer given by $\mu^* = {\mathbf{S}_\mu^{-\frac{1}{2}}}_{\#}\mu$. By using different techniques, M. Maslouhi and S. Loukili not only showed the same result about the closest Parseval frame but also showed that there exists a unique closest probabilistic tight frame [Maslouhi and Loukili, 2019, Loukili and Maslouhi, 2020]: let $\mathcal{T}(\mathbb{R}^n)$ be the set of tight probabilistic frames in $\mathbb{R}^n$, then the minimizing problem

$$I(\mu, \mathbb{R}^n) := \inf_{\nu \in \mathcal{T}(\mathbb{R}^n)} W_2(\mu, \nu)$$

admits a unique solution given by $\mu^* = (\lambda \mathbf{S}_\mu)^{-\frac{1}{2}}{}_{\#}\mu$ with $\lambda^{-1/2} = \frac{\operatorname{tr}(\mathbf{S}_\mu^{1/2})}{n}$.

For a complete introduction to probabilistic frames, we refer to [Ehler and Okoudjou, 2013, Wickman, 2014] for more details.

## 1.4 Summary of Main Results

It is well-known that the canonical Parseval frame is the closest Parseval frame to a given frame [Bodmann and Casazza, 2010, Casazza and Kutyniok, 2007] and the canonical Probabilistic Parseval frame is the unique closest Parseval probabilistic frame to a given probabilistic frame under the 2-Wassertein metric [Cheng and Okoudjou, 2019,Loukili and Maslouhi, 2020]. This work extends the above result into a more general setting.

Recall that $\mathbb{S}^n_{++}$ is the set of $n \times n$ real positive definite symmetric matrices. For $\mathbf{T} \in \mathbb{S}^n_{++}$, let $\mathcal{P}_{\mathbf{T}}$ be the set of probabilistic frames with frame operator $\mathbf{T}$. We first show that $\mathcal{P}_{++}$, the space of probabilistic frames, and $\mathcal{P}_{\mathbf{T}}$ are homotopy equivalent and pathwise connected. Then for $\mu \in \mathcal{P}_{\mathbf{S}}$, we show that the following problem

$$W_2(\mu, \mathcal{P}_{\mathbf{T}}) := \inf_{\nu \in \mathcal{P}_{\mathbf{T}}} W_2(\mu, \nu).$$

admits a unique optimizer $\nu := \mathbf{A}(\mathbf{S}, \mathbf{T})_{\#}\mu$ where $\mathbf{A}(\mathbf{S}, \mathbf{T}) := \mathbf{S}^{-1/2}(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}\mathbf{S}^{-1/2}$.

As an application, the above enables us to study unit-norm probabilistic tight frames. A probabilistic frame $\mu$ is unit-norm if $M_2(\mu) := \int_{\mathbb{R}^n} \|\mathbf{x}\|^2 d\mu(\mathbf{x}) = 1$. It is clear that $\mu$ is a unit-norm probabilistic tight frame if and only if $\mathbf{S}_\mu = \frac{1}{n}\mathbf{Id}$. Therefore, given a probabilistic frame $\mu$ with frame operator $\mathbf{S}_\mu$, the closest unit-norm probabilistic tight frame to $\mu$ in the $W_2$ topology is $\mathbf{A}(\mathbf{S}_\mu, \frac{1}{n}\mathbf{Id})_{\#}\mu = (n\mathbf{S}_\mu)^{-\frac{1}{2}}{}_{\#}\mu$, which answers an open problem (Problem 12.1 (b)) in [Ehler and Okoudjou, 2013]. Furthermore, we claim that $\mathcal{P}_{\frac{1}{n}\mathbf{Id}}$, the set of unit-norm probabilistic tight frames, is pathwise connected, which is similar to the Larson's frame homotopy problem proposed in an REU program at 2002 and later proved by [Cahill et al., 2017]: the set of finite unit-norm tight frames is pathwise connected.

Perturbation analysis is a classic topic in the frame theory. In this work, we show that the set of probabilistic $p$–frames is open with respect to the $W_p$ topology. Furthermore, we show that for a given probabilistic frame $\mu$ with the lower frame bound $A$ and $\nu \in \mathscr{P}_2(\mathbb{R}^d)$, if

$$\lambda := \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x} - \mathbf{y}\|^2 d\gamma(\mathbf{x}, \mathbf{y}) < A,$$

where $\gamma \in \Gamma(\mu, \nu)$. Then $\nu$ is a probabilistic frame with bounds

$$(\sqrt{A} - \sqrt{\lambda})^2 \text{ and } M_2(\nu).$$

In particular, if $W_2^2(\mu, \nu) < A$, then $\nu$ is a probabilistic frame with bounds

$$(\sqrt{A} - W_2(\mu, \nu))^2 \text{ and } \ M_2(\nu).$$

Finally, we generalize the Paley–Wiener Theorem for probabilistic frames, which claims that if a probability measure is close to a probabilistic frame in some sense, then this probability measure is also a probabilistic frame. That is to say, if $\mu$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $0 < A \leq B < \infty$ and $\nu \in \mathscr{P}_2(\mathbb{R}^n)$. If there exist constants $\lambda_1, \lambda_2, \delta \geq 0$ such that max $(\lambda_1 + \frac{\delta}{\sqrt{A}}, \lambda_2) < 1$ and for all $w \in C_c(\mathbb{R}^n)$,

$$\begin{aligned} &\Big\| \int_{\mathbb{R}^n} w(\mathbf{x})\mathbf{x} d\mu(\mathbf{x}) - \int_{\mathbb{R}^n} w(\mathbf{y})\mathbf{y} d\nu(\mathbf{y}) \Big\| \\ &\qquad \leq \lambda_1 \Big\| \int_{\mathbb{R}^n} w(\mathbf{x})\mathbf{x} d\mu(\mathbf{x}) \Big\| + \lambda_2 \Big\| \int_{\mathbb{R}^n} w(\mathbf{y})\mathbf{y} d\nu(\mathbf{y}) \Big\| + \delta \|w\|_{L^2(\mu)} \end{aligned}$$

where $C_c(\mathbb{R}^n)$ be the set of continuous functions on $\mathbb{R}^n$ with compact support. Then $\nu$ is a probabilistic frame for $\mathbb{R}^n$ with bounds

$$\frac{A^2(1 - (\lambda_1 + \frac{\delta}{\sqrt{A}}))^2}{(1 + \lambda_2)^2 M_2(\nu)} \text{ and } M_2(\nu).$$

# Chapter 2

# Probabilistic Frames and Wasserstein Distances

## 2.1 Introduction

In recent work [Ehler and Okoudjou, 2013,Loukili and Maslouhi, 2020,Maslouhi and Loukili, 2019,Wickman and Okoudjou, 2023], probabilistic frames, a subset of Borel probability measures on $\mathbb{R}^n$ that generalize, but also contain actual frames, have been considered. Probabilistic frames are Borel probability measures in $\mathbb{R}^n$ with finite second moments whose support, interpreted as a set of vectors, contains a basis of $\mathbb{R}^n$. Recall that we denote the set of Borel probability measures on $\mathbb{R}^n$ by $\mathscr{P}(\mathbb{R}^n)$ and by $\mathscr{P}_p(\mathbb{R}^n)$ those with finite p-th moments where $p \geq 1$, i.e. $\mu \in \mathscr{P}_p(\mathbb{R}^n)$ if $\mu \in \mathscr{P}(\mathbb{R}^n)$ and $M_p(\mu) := \int_{\mathbb{R}^n} \|\mathbf{x}\|^p \; d\mu(\mathbf{x}) < \infty$.

**Definition 2.1.1.** *$\mu \in \mathscr{P}_p(\mathbb{R}^n)$ is called probabilistic p-frame if there exist $0 < A \leq B$ such that for any $\mathbf{x} \in \mathbb{R}^n$,*

$$A\|\mathbf{x}\|^p \leq \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{y} \rangle|^p \; d\mu(\mathbf{y}) \leq B\|\mathbf{x}\|^p.$$

*Furthermore, $\mu$ is called tight (frame) if $A = B$, and $\mu$ is called Parseval (frame) if $A = B = 1$.*

The set $\mathscr{P}_p(\mathbb{R}^n)$ is equipped with the p-Wasserstein metric $W_p(\cdot, \cdot)$ and becomes a metric space. Convergence $\mu_n \to \mu$ in the p-Wasserstein metric is equivalent to weak-$*$ convergence together with convergence of p-th moments $\int_{\mathbb{R}^n} \|\mathbf{x}\|^p d\mu_n \to \int_{\mathbb{R}^n} \|\mathbf{x}\|^p d\mu$. For more background see [Figalli and Glaudo, 2021, Villani, 2003, Villani et al., 2009]. Starting with a simple observation:

**Proposition 2.1.2.** *Suppose $\mu \in \mathscr{P}_p(\mathbb{R}^n)$. For any unit-vector $\mathbf{x}$ and any $p \geq 1$*

$$W_p^p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{v} \rangle|^p d\mu(\mathbf{v}),$$

*where $\pi_{\mathbf{x}^\perp}$ denotes the orthogonal projection to the plane $\mathbf{x}^\perp$ of vectors perpendicular to $\mathbf{x}$.*

Since the "basis containment" condition for frames means that a probabilistic p-frame has positive p-Wasserstein distance to all linear subspaces we see:

**Proposition 2.1.3.** *$\mu \in \mathscr{P}_p(\mathbb{R}^n)$ is a probabilistic p-frame, if and only if $W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) > 0$ for all unit vectors $\mathbf{x}$.*

We call a probabilistic 2-frame simply probabilistic frame. If $\mu \in \mathscr{P}_2(\mathbb{R}^n)$ the linear operator $\langle \mathbf{x}, \mathbf{S}_\mu \mathbf{y} \rangle := \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{v} \rangle \langle \mathbf{y}, \mathbf{v} \rangle \; d\mu(\mathbf{v})$ declared for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. After the choice of a basis we see that $\mathbf{S}_\mu = \int_{\mathbb{R}^n} \mathbf{v}\mathbf{v}^t \; d\mu$ is a symmetric positive semi-definite matrix and for any unit vector $\mathbf{x} \in \mathbb{S}^{n-1}$,

$$W_2^2(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = \mathbf{x}^t \int_{\mathbb{R}^n} \mathbf{v}\mathbf{v}^t \; d\mu(\mathbf{v}) \; \mathbf{x} = \mathbf{x}^t \; \mathbf{S}_\mu \; \mathbf{x}. \tag{2.1.1}$$

Hence, $\mathbf{S}_\mu$ is positive definite if and only if $\mu$ is a (probabilistic) frame and, in this case, is known as the *frame operator* of $\mu$. Identity 2.1.1 gives a *geometric interpretation of the frame property.* More precisely, since any positive semi-definite $\mathbf{S}_\mu$ has a root $\mathbf{S}_\mu^{1/2}$ the ellipsoid $\mathscr{E}_\mu := \{\mathbf{S}_\mu^{1/2} \; \mathbf{x} : \; \|\mathbf{x}\| = 1\} \subset \mathbb{R}^n$ is a hyperellipsoid exactly if $\mathbf{S}_\mu$ is definite, that is, if $\mu$ is a probabilistic frame. We call the ellipsoid $\mathscr{E}_\mu$ the *Wasserstein ellipsoid* of $\mu$. It provides the 2-Wasserstein distance of a given (probabilistic) frame to the nearest non-frame in any given spacial direction. In an (orthogonal) eigenbasis $\{\mathbf{e}_i\}$ for $\mathbf{S}_\mu$ (not necessarily definite) points on $\mathscr{E}_\mu$ solve the equation

$$\sum_{\{i:\; W_2(\mu, (\pi_{\mathbf{e}_i^\perp})_\# \mu) \neq 0\}} \frac{x_i^2}{W_2^2(\mu, (\pi_{\mathbf{e}_i^\perp})_\# \mu)} = 1.$$

With respect to that eigenbasis the values $\{W_2(\mu, (\pi_{\mathbf{e}_i^\perp})_\# \mu)\}_{i=1}^n$ are the eigenvalues of $\mathbf{S}_\mu^{1/2}$. Because the Wasserstein ellipsoid is defined through Wasserstein distances, it is neither the Legendre- nor the Binet-ellipsoid; see [V.D. and Pajor, 1989] and the literature therein. Let $\mathbb{S}_+^n$ be the set of non-negative definite symmetric $n \times n$ matrices and $\mathbb{S}_{++}^n \subset \mathbb{S}_+^n$ those that are positive definite. For given $\mathbf{T} \in \mathbb{S}^n_{++}$ let $W_2(\mu, \mathscr{P}_{\mathbf{T}}) := \inf_{\nu \in \mathscr{P}_{\mathbf{T}}} W_2(\mu, \nu)$. Based on theorems of [Cuesta-Albertos et al., 1996] and [Gelbrich, 1990, Olkin and Pukelsheim, 1982] adapted to frames we show

**Theorem 2.1.4.** *For any given $\mathbf{S}, \mathbf{A} \in \mathbb{S}^n_{++}$ push-forward with $\mathbf{A}$ defines a homeomorphism $\mathbf{A}_\# : \mathscr{P}_{\mathbf{S}} \to \mathscr{P}_{\mathbf{ASA}}$ so that*

$$W_2^2(\mu, \mathbf{A}_\# \mu) = W_2^2(\mu, \mathscr{P}_{\mathbf{ASA}}) = \mathrm{tr}\ \mathbf{S}(\mathbf{Id} - \mathbf{A})^2 \tag{2.1.2}$$

*for all $\mu \in \mathscr{P}_{\mathbf{S}}$ and $W_2(\mu, \nu) > W_2(\mu, \mathbf{A}_\# \mu)$ for any $\nu \in \mathscr{P}_{\mathbf{ASA}}$ so that $\nu \neq \mathbf{A}_\# \mu$.*

One may be interested in the Wasserstein distance between frames with given frame operators, say $\mathbf{S}, \mathbf{T} \in \mathbb{S}^n_{++}$. In this case one applies the Theorem to the unique $\mathbf{A} \in \mathbb{S}^n_{++}$ that solves $\mathbf{T} = \mathbf{ASA}$ (see Proposition 2.4.1) given by

$$\mathbf{A} = \mathbf{A}(\mathbf{S}, \mathbf{T}) := \mathbf{S}^{-1/2}(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}\mathbf{S}^{-1/2}.$$

Since $W_2(\mu, \mathscr{P}_{\mathbf{T}})$ is the same for all $\mu \in \mathscr{P}_{\mathbf{S}}$ we may define $d_W(\mathbf{S}, \mathbf{T}) := W_2(\mu, \mathscr{P}_{\mathbf{T}})$ for one and any $\mu \in \mathscr{P}_{\mathbf{S}}$.

**Proposition 2.1.5.** *Given $\mathbf{S}, \mathbf{T} \in \mathbb{S}^n_{+}$. Then $d_W(\mathbf{S}, \mathbf{T}) := W_2(\mathscr{P}_{\mathbf{S}}, \mathscr{P}_{\mathbf{T}})$ is a metric on $\mathbb{S}^n_{+}$. More precisely, we have*

$$\begin{aligned} d_W(\mathbf{S}, \mathbf{T}) &= \mathrm{tr}(\mathbf{S} + \mathbf{T} - (\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2} - (\mathbf{T}^{1/2}\mathbf{S}\mathbf{T}^{1/2})^{1/2}) \\ &= \mathrm{tr}\ (\mathbf{S}(\mathbf{Id} - \mathbf{A}(\mathbf{S}, \mathbf{T})) + \mathbf{T}(\mathbf{Id} - \mathbf{A}(\mathbf{T}, \mathbf{S}))). \end{aligned} \tag{2.1.3}$$

*Up to roots of entries, this metric compares to the metrics induced by the operator norm $\|\cdot\|_{op}$ and the Frobenius norm $\|\cdot\|_F$ as follows.*

$$\|\mathbf{S}^{1/2} - \mathbf{T}^{1/2}\|_{op} \leq W_2(\mathscr{P}_{\mathbf{S}}, \mathscr{P}_{\mathbf{T}}) = d_W(\mathbf{S}, \mathbf{T}) \leq \|\mathbf{S}^{1/2} - \mathbf{T}^{1/2}\|_F. \tag{2.1.4}$$

*In particular, the topology generated by $d_W$ is equivalent to the standard norm topologies on $\mathbb{S}^n_{+}$.*

The importance of this proposition is that it shows the continuity of the "frame map" $S : \mathscr{P}_2(\mathbb{R}^n) \to \mathbb{S}^n_{+}$ given by $S(\mu) = \mathbf{S}_\mu$, for an alternative proof see [Wickman and Okoudjou, 2023]. The closely related metric $d(\mathbf{S}, \mathbf{T}) := \sqrt{d_W(\mathbf{S}^2, \mathbf{T}^2)}$ defined for symmetric matrices $\mathbf{S}, \mathbf{T} \in \mathbb{S}^n$ is by estimate 2.1.4 equivalent to norm induced metrics on $\mathbb{R}^n$, however it is not induced by a norm as shown in [Cuesta-Albertos et al., 1996].

**Corollary 2.1.6.** *The set of probabilistic $p$-frames is open in the $p$-Wasserstein topology on $\mathscr{P}_p(\mathbb{R}^n)$.*

We give the proof for general $p$ later. However, it is easy to see for $p = 2$. Just compose the (continuous) frame map $S$ with the determinant map $\det : \mathbb{S}^n_+ \to \mathbb{R}_{\geq 0}$, to get a continuous map $\det \circ S : \mathscr{P}_2 \to \mathbb{R}_{\geq 0}$. Then the set of frames $\{\mu \in \mathscr{P}_2(\mathbb{R}^n) : \det \circ S(\mu) > 0\}$ is open.

That means that the frame map $S : \mathscr{P}_2(\mathbb{R}^n) \to \mathbb{S}^n_+$ defines a foliation over the set $\mathbb{S}^n_+$ of $n \times n$ positive semi-definite symmetric matrices with real entries. Restricted to frames, this gives a foliation over $\mathbb{S}^n_{++}$, the set of positive definite symmetric matrices. Theorem 2.1.4 implies that the fibers $\mathscr{P}_{\mathbf{S}} \subset \mathscr{P}_2(\mathbb{R}^n)$, i.e. the measures in $\mathscr{P}_2(\mathbb{R}^n)$ with frame operator $\mathbf{S}$, can be identified.

## 2.2 Applications of the Main Results

Here is an application of Theorem 2.1.4. Given a frame operator, say $\mathbf{T}$, consider the ray $\mathbb{R}_+\mathbf{T} := \{\lambda\mathbf{T} : \lambda \in \mathbb{R}_+\}$ through $\mathbf{T}$. Let $W_2^2(\mu, \mathbb{R}_+\mathbf{T}) := \inf_{\lambda \in \mathbb{R}_+} W_2^2(\mu, \mathscr{P}_{\lambda\mathbf{T}})$.

**Corollary 2.2.1.** *Let $\mu$ be a probabilistic frame, then*
$W_2^2(\mu, \mathbb{R}_+\mathbf{T}) = W_2^2(\mu, (\lambda_{\min}^{1/2}\mathbf{A}(\mathbf{S}_\mu, \mathbf{T}))_{\#}\mu)$ *where* $\lambda_{\min}^{1/2} = \frac{\text{tr } (\mathbf{S}_\mu^{1/2}\mathbf{T}\mathbf{S}_\mu^{1/2})^{1/2}}{\text{tr } \mathbf{T}}$.

In particular, the closest tight frame to a given frame is obtained by putting $\mathbf{T} = \mathbf{Id}$, see also [Loukili and Maslouhi, 2020].

*Proof.* From theorem 2.1.4 we know that the probabilistic frame with given frame operator $\lambda\mathbf{T}$ closest to $\mu$ is given by $(\lambda^{1/2}\mathbf{A}(\mathbf{S}_\mu, \mathbf{T}))_{\#}\mu$. To determine the optimal $\lambda$, using Identity 2.1.2, so that:

$$W_2^2(\mu, (\lambda^{1/2}\mathbf{A}(\mathbf{S}_\mu, \mathbf{T}))_{\#}\mu) = \text{tr } \mathbf{S}_\mu + \lambda \cdot \text{tr } \mathbf{T} - 2\sqrt{\lambda} \cdot \text{tr } (\mathbf{S}_\mu^{1/2}\mathbf{T}\mathbf{S}_\mu^{1/2})^{1/2}$$

is differentiable in $\lambda$, with minimum as stated. □

Let us denote the set of probabilistic frames in $\mathscr{P}_2(\mathbb{R}^n)$ by $\mathscr{P}_{++}$, so that

$$\mathscr{P}_{++} = (\det \circ S)^{-1}(0, \infty) = S^{-1}\mathbb{S}^n_{++}.$$

We can reformulate Theorem 2.1.4 as follows.

**Theorem 2.2.2.** *Push forward with $\mathbf{A} \in \mathbb{S}^n_{++}$ lifts the congruence action $C_{\mathbf{S}}(\mathbf{A}) := \mathbf{A}\mathbf{S}\mathbf{A}^t$ of the multiplicative group $(\mathbb{S}^n_{++}, \cdot)$ on $\mathbb{S}^n_{++}$ to the foliation $S : \mathscr{P}_{++} \to \mathbb{S}^n_{++}$. More precisely, push forward* *with* $\mathbf{A}$ *is a group action on* $\mathscr{P}_{++}$ *so that* $C_{\mathbf{A}} \circ S = S \circ \mathbf{A}_{\#}$. *The lifted action is faithful, continuous, and distance minimizing with respect to* $W_2$. *More precisely, if* $\mathbf{A} \in \mathbb{S}^n_{++}$ *then for every* $\mu \in \mathscr{P}_{++}$

$$W_2^2(\mu, \mathbf{A}_{\#}\mu) = W_2^2(\mathscr{P}_{\mathbf{S}_\mu}, \mathscr{P}_{\mathbf{A}\mathbf{S}_\mu\mathbf{A}}) = \text{tr } \mathbf{S}_\mu(\mathbf{Id} - \mathbf{A})^2 \tag{2.2.1}$$

*Moreover push-forward with the interpolation maps* $I_{\mathbf{A}}(t) := (1-t)\mathbf{Id} + t\mathbf{A}$ *defines 2-Wasserstein constant speed geodesic curves* $((I_{\mathbf{A}}(t))_{\#}\mu)_{t\in[0,1]}$ *in* $(\mathscr{P}_{++}, W_2)$.

The proofs are formal consequences of Theorem 2.1.4, instead of presenting those we show the $\mathbb{S}^n_{++}$ action in a commutative diagram:

$$\begin{array}{ccc} \mathbb{S}^n_{++} \times \mathscr{P}_{++} & \longrightarrow & \mathscr{P}_{++} \\ {\scriptstyle \mathbf{Id}\times S}\downarrow & & \downarrow{\scriptstyle S} \\ \mathbb{S}^n_{++} \times \mathbb{S}^n_{++} & \longrightarrow & \mathbb{S}^n_{++} \end{array} \qquad \begin{array}{ccc} (\mathbf{A}, \mu) & \longmapsto & \mathbf{A}_{\#}\mu \\ \downarrow & & \downarrow \\ (\mathbf{A}, \mathbf{S}_\mu) & \longmapsto & \mathbf{A}\mathbf{S}_\mu\mathbf{A}^t \end{array}$$

The last statement about interpolation geodesics is standard; see, for example, [Figalli and Glaudo, 2021] section 3.1.1.

Noticing that push-forward with $t \mapsto I_{\mathbf{A}(\mathbf{S}_\mu,\mathbf{T})}(t)$ defines a homotopy between $\mathscr{P}_{++}$ and the fiber $\mathscr{P}_{\mathbf{T}}$ that is the identity on $\mathscr{P}_{\mathbf{T}}$ now gives:

**Proposition 2.2.3.** *For any* $\mathbf{S} \in \mathbb{S}^n_{++}$ *the set theoretical inclusion* $i : \mathscr{P}_{\mathbf{S}} \hookrightarrow \mathscr{P}_{++}$ *is a deformation retraction with respect to the retraction map* $r : \mathscr{P}_{++} \to \mathscr{P}_{\mathbf{S}}$ *given by* $r(\mu) = \mathbf{A}(\mathbf{S}_\mu, \mathbf{S})_{\#}\mu$. *In particular the spaces* $\mathscr{P}_{++}$ *and* $\mathscr{P}_{\mathbf{S}}$ *are homotopy equivalent.*

*Proof.* We note that all maps stated in the proposition are well-defined and continuous on Wasserstein space $(\mathscr{P}_{++}, W_2)$. This is because push-forward with a continuous function is continuous. Now if $\mu \in \mathscr{P}_{\mathbf{S}}$, then $r(\mu) = \mathbf{A}(\mathbf{S}, \mathbf{S})_{\#}\mu = (\mathbf{Id})_{\#}\mu = \mu$, so that $r \circ i = \mathbf{Id}_{\mathscr{P}_{\mathbf{S}}}$. All we need to show is that $i \circ r = r$ is homotopic to the identity map on $\mathscr{P}_{++}$. Such a homotopy is given by

$$H(t, \mu) = (I_{\mathbf{A}(\mathbf{S}_\mu,\mathbf{S})}(t))_{\#}\mu$$

for $(t, \mu) \in [0, 1] \times \mathscr{P}_{++}$. □

**Theorem 2.2.4.** *The space* $\mathscr{P}_{\mathbf{S}}$ *is pathwise connected for any* $\mathbf{S} \in \mathbb{S}^n_{++}$.

*Proof.* By the previous statement, it suffices to show $\mathscr{P}_{++}$ is path-connected.

To do this, we first show that a 2-Wasserstein open ball of a given probabilistic frame $\nu \in \mathscr{P}_{++}$ is connected if it is small enough. Indeed since $\nu \in \mathscr{P}_{++}$ and $\mathscr{P}_{++}$ is open, there is a $\delta > 0$ such that the open ball $B_\delta(\nu) := \{\eta \in \mathscr{P}_2(\mathbb{R}^n) : \; W_2(\eta, \nu) < \delta\}$ is contained in $\mathscr{P}_{++}$. If $\mu \in B_\delta(\nu)$ given an optimal coupling $\gamma \in \Gamma(\nu, \mu)$, there is a unit speed geodesic $(\mu_t)_{t\in[0,1]}$ in $\mathscr{P}_2(\mathbb{R}^n)$ that never leaves $B_\delta(\nu)$ because it decreases distance. More precisely, the optimal coupling $\gamma \in \Gamma(\mu, \nu)$ induces a geodesic curve $\mu(t)$ connecting $\mu$ and $\nu$ as follows. Let $\pi_t(x, y) := (1-t)x + ty$, so that $\pi_0(x, y) = x$ and $\pi_1(x, y) = y$, then put $\mu_t := (\pi_t)_\# \gamma$ (for $t \in [0, 1]$), so that $\mu_0 = (\pi_0)_\# \gamma = \mu$ and $\mu_1 := (\pi_1)_\# \gamma_B = \nu$. An optimal coupling between any two points of the geodesic curve is given by $\gamma_B(s, t) := (\pi_s, \pi_t)_\# \gamma_B$. Use this coupling to show that the curve $(\mu_t)$ is a unit speed geodesic that linearly decreases distance to $\nu$ in $t$, in fact $W_2(\mu_t, \nu) = (1-t)W_2(\mu, \nu)$ for $t \in [0, 1]$. This shows $B_\delta(\nu)$ itself is connected.

Now we show that there is a curve within the set of probabilistic frames that starts at a specific measure and ends in $B_\delta(\nu)$. First, the specific measure, say $\mu_r$ corresponds to the equally distributed mass in an open ball $D_r$ of a radius $r$, so that $\mu_r(D_r) = 1$. Note that this measure is absolutely continuous with respect to the Lebesgue measure. Denote the absolutely continuous measures in $\mathscr{P}_2(\mathbb{R}^n)$ by $\mathscr{P}_{2,ac}(\mathbb{R}^n)$. Now every probability measure can be approximated by a probability that is a finite combination of delta measures in $W_2$. Those in turn can be approximated in $W_2$ by an absolutely continuous measure that is the union of "thickenings" of the delta measures by masses equally supported on small open balls centered at the support of the given delta distribution. Taking the supporting sets small enough, we can make sure such measure, say $\mu_\delta$, lies in $B_\delta(\nu)$. Since $\mu_\delta$ and $\mu_r$ are both in $\mathscr{P}_{2,ac}$ the minimal coupling $\gamma$ between the two is a given by a transport map. Moreover, see Villani [Villani et al., 2009] Proposition 5.9 (iii), the canonical geodesic curve between two absolutely continuous measures consists of absolutely continuous measures, hence frames. This is because absolutely continuous measures are not supported in a linear subspace, since the Lebesgue measure of a linear subspace is zero. So by Proposition 2.1.3, an absolutely continuous measure must be a probabilistic frame.

That implies, there is a path within the set of frames from any given probabilistic frame $\nu$ to the frame $\mu_r$. This is what we wanted to show. □

## 2.3 Wasserstein Openness of the Set of Probabilistic p-Frames.

In order to show a general openness result we return to general p-frames in this section. We start with the proof of Proposition 2.1.2 from the introduction.

*Proof of Proposition 2.1.2.* First, note that by Cauchy-Schwarz, the integral on the right is well-defined for all $\mu \in \mathscr{P}_p(\mathbb{R}^n)$. Now, for any unit vector $\mathbf{x} \in \mathbb{S}^{n-1}$

$$|\langle \mathbf{x}, \mathbf{v} \rangle|^p = \operatorname{dist}^p(\mathbf{v}, \mathbf{x}^\perp) = \inf_{\mathbf{y} \in \mathbf{x}^\perp} |\mathbf{y} - \mathbf{v}|^p = |\pi_{\mathbf{x}^\perp}(\mathbf{v}) - \mathbf{v}|^p. \tag{2.3.1}$$

By definition

$$W_p^p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = \inf_{\gamma \in \Gamma(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu)} \int_{\mathbb{R}^{2n}} |\mathbf{v} - \mathbf{y}|^p \; d\gamma(\mathbf{v}, \mathbf{y}),$$

but that minimum is taken on when pushing forward the mass with the orthogonal projection onto $\mathbf{x}^\perp$, since that way every point moves minimal distance to the target, hence

$$W_p^p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = \int_{\mathbb{R}^n} |\mathbf{v} - \mathbf{y}|^p \; d(\mathrm{Id} \times \pi_{\mathbf{x}^\perp})_\# \mu = \int_{\mathbb{R}^n} |\mathbf{v} - \pi_{\mathbf{x}^\perp}(\mathbf{v})|^p \; d\mu(\mathbf{v}).$$

Hence

$$W_p^p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{v} \rangle|^p \; d\mu(\mathbf{v}).$$

□

We note that a stronger minimization property is true. Since $\operatorname{supp}((\pi_{\mathbf{x}^\perp})_\# \mu) \subset \mathbf{x}^\perp$ we have

$$\inf_{\{\gamma \in \Gamma(\mu, \nu) \colon \operatorname{supp}(\nu) \subset \mathbf{x}^\perp\}} \int_{\mathbb{R}^{2n}} |\mathbf{v} - \mathbf{y}|^p \; d\gamma(\mathbf{v}, \mathbf{y}) \leq W_p^p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu).$$

Since on the other hand for each point $\mathbf{v}$ the orthogonal projection $\pi_{\mathbf{x}^\perp}(\mathbf{v})$ minimizes the distance to $\mathbf{x}^\perp$ it is obvious that the push-forward of $\mu$ with $\pi_{\mathbf{x}^\perp}$ minimizes among all measures supported in $\mathbf{x}^\perp$, hence we see

$$\inf_{\{\gamma \in \Gamma(\mu, \nu) \colon \operatorname{supp}(\nu) \subset \mathbf{x}^\perp\}} \int_{\mathbb{R}^{2n}} |\mathbf{v} - \mathbf{y}|^p \; d\gamma(\mathbf{v}, \mathbf{y}) = W_p^p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu).$$

The proposition would also follow from an extended version of Theorem 2.1.4 that includes the limiting case of semi-definite symmetric operators. This is true, but we do not prove it here.

*Proof of Proposition 2.1.3.* Suppose $\mu \in \mathscr{P}_p(\mathbb{R}^n)$ is a probabilistic p-frame with bounds $0 < A \leq B$, then for any unit vector $\mathbf{x}$,

$$\int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{y} \rangle|^p \, d\mu(\mathbf{y}) \geq A > 0.$$

Recall by Proposition 2.1.2 we have $W_p^p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{y} \rangle|^p d\mu(\mathbf{y})$, so that for any unit vector $\mathbf{x}$, $W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) \geq A^{1/p} > 0$.

Conversely, in the proof of Proposition 2.3.2, we show that $W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu)$ depends continuously $\mathbf{x}$. Since $\mathbb{S}^{n-1}$ is compact $W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu)$ takes on its minimum at some point, say $\mathbf{x}_{\min} \in \mathbb{S}^{n-1}$. Then for any $\mathbf{x} \in \mathbb{S}^{n-1}$,

$$0 < W_p^p(\mu, (\pi_{\mathbf{x}_{\min}^\perp})_\# \mu) \leq W_p^p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{y} \rangle|^p d\mu(\mathbf{y}).$$

Therefore, for any $\mathbf{x} \in \mathbb{R}^n$,

$$W_p^p(\mu, (\pi_{\mathbf{x}_{\min}^\perp})_\# \mu) \|\mathbf{x}\|^p \leq \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{y} \rangle|^p d\mu(\mathbf{y}) \leq M_p(\mu) \|\mathbf{x}\|^p,$$

where the last inequality comes from the Cauchy-Schwarz inequality. That is to say, $\mu$ is a p-frame with bounds $W_p^p(\mu, (\pi_{\mathbf{x}_{\min}^\perp})_\# \mu)$ and $M_p(\mu)$. □

Then we have another equivalence of p-frames in the following corollary.

**Corollary 2.3.1.** *$\mu \in \mathscr{P}_p(\mathbb{R}^n)$ is a probabilistic p-frame, if and only if span$\{$supp$(\mu)\} = \mathbb{R}^n$.*

*Proof.* Note that *span*$\{$*supp*$(\mu)\} = \mathbb{R}^n$ is equivalent to $W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) > 0$ for all unit vectors $\mathbf{x} \in \mathbb{S}^{n-1}$. Then by Proposition 2.1.3, we complete the proof. □

**Proposition 2.3.2** (Openness of the Set of Probabilistic p-Frames)**.** *The set of probabilistic p-frames is open in the $W_p$ topology.*

*Proof.* Suppose $\mu \in \mathscr{P}_p(\mathbb{R}^n)$. Then the minimal cost to transport $\mu$ to a probability on a given 1-codimensional subspace, say $\mathbf{x}^\perp$, is attained by pushing forward $\mu$ to $\mathbf{x}^\perp$ via the orthogonal projection $\pi_{\mathbf{x}^\perp} : \mathbb{R}^n \to \mathbf{x}^\perp$. This is clear, since for any $p \geq 1$, the transport distance for any point in $\mathbb{R}^n$ in the support of $\mu$ to $\mathbf{x}^\perp$ is minimal under the projection. Hence

$$W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = \inf_{\substack{\nu \in \mathscr{P}_p(\mathbb{R}^n) \\ \text{supp } \nu \subset \mathbf{x}^\perp}} W_p(\mu, \nu)$$

represents the p-Wasserstein distance of $\mu$ to the set of measures supported in $\mathbf{x}^{\perp}$. If $\mu$ is a probabilistic p-frame (with constants $0 < A \leq B < \infty$), it cannot be supported in the hyperplane $\mathbf{x}^{\perp}$, since then we would have the contradiction $0 < A \leq \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{v}\rangle|^p \, d\mu(\mathbf{v}) = 0$. Hence $W_p(\mu, (\pi_{\mathbf{x}^{\perp}})_{\#}\mu) > 0$. Now for a fixed probabilistic p-frame, say $\mu$, the p-Wasserstein distance to $\mathbf{x}^{\perp}$ depends continuously on the subspace $\mathbf{x}^{\perp}$, that is on $\mathbf{x}$, in the topology induced by the p-Wasserstein metric. This is relatively straightforward to see. Firstly, given two vectors $\mathbf{x}, \mathbf{y}$, the triangle inequality implies

$$|W_p(\mu, (\pi_{\mathbf{x}^{\perp}})_{\#}\mu) - W_p(\mu, (\pi_{\mathbf{y}^{\perp}})_{\#}\mu)| \leq W_p((\pi_{\mathbf{x}^{\perp}})_{\#}\mu, (\pi_{\mathbf{y}^{\perp}})_{\#}\mu)$$

To estimate the Wasserstein distance on the right, consider the push-forward $D_{\#}\mu$ of $\mu$ under the diagonal map $D(x) = (x, x) \in \mathbb{R}^{2n}$. Then pushing this measure forward with $\pi_{\mathbf{x}^{\perp}} \times \pi_{\mathbf{y}^{\perp}}$ results in a coupling between $(\pi_{\mathbf{x}^{\perp}})_{\#}\mu$ and $(\pi_{\mathbf{y}^{\perp}})_{\#}\mu$. Hence we have an estimate

$$W_p^p((\pi_{\mathbf{x}^{\perp}})_{\#}\mu, (\pi_{\mathbf{y}^{\perp}})_{\#}\mu) \leq \int_{\mathbb{R}^n} \|\pi_{\mathbf{x}^{\perp}}(\mathbf{z}) - \pi_{\mathbf{y}^{\perp}}(\mathbf{z})\|^p \, d\mu(\mathbf{z})$$

To simplify write $\pi_{\mathbf{x}^{\perp}}(\mathbf{z}) = \mathbf{z} - \langle \mathbf{z}, \mathbf{x}\rangle \mathbf{x}$

$$\int_{\mathbb{R}^n} \|\pi_{\mathbf{x}^{\perp}}(\mathbf{z}) - \pi_{\mathbf{y}^{\perp}}(\mathbf{z})\|^p \, d\mu = \int_{\mathbb{R}^n} \|\langle \mathbf{z}, \mathbf{x}\rangle \mathbf{x} - \langle \mathbf{z}, \mathbf{y}\rangle \mathbf{y}\|^p \, d\mu.$$

Now put $\mathbf{y} = \mathbf{x} + \hat{\mathbf{y}}$ to get $\langle \mathbf{z}, \mathbf{y}\rangle \mathbf{y} = \langle \mathbf{z}, \mathbf{x} + \hat{\mathbf{y}}\rangle(\mathbf{x} + \hat{\mathbf{y}}) = \langle \mathbf{z}, \mathbf{x}\rangle \mathbf{x} + \langle \mathbf{z}, \hat{\mathbf{y}}\rangle \mathbf{x} + \langle \mathbf{z}, \mathbf{x} + \hat{\mathbf{y}}\rangle \hat{\mathbf{y}}$. Hence

$$\int_{\mathbb{R}^n} \|\langle \mathbf{z}, \mathbf{x}\rangle \mathbf{x} - \langle \mathbf{z}, \mathbf{y}\rangle \mathbf{y}\|^p \, d\mu = \int_{\mathbb{R}^n} \|\langle \mathbf{z}, \hat{\mathbf{y}}\rangle \mathbf{x} + \langle \mathbf{z}, \mathbf{x} + \hat{\mathbf{y}}\rangle \hat{\mathbf{y}}\|^p \, d\mu$$

Use Minkowski's inequality followed by Cauchy-Schwarz while recalling that $\|\mathbf{x}\| = 1$

$$\leq 2^{p-1} \int_{\mathbb{R}^n} (\|\langle \mathbf{z}, \hat{\mathbf{y}}\rangle \mathbf{x}\|^p + \|\langle \mathbf{z}, \mathbf{x} + \hat{\mathbf{y}}\rangle \hat{\mathbf{y}}\|^p) d\mu \leq 2^{p-1} \|\hat{\mathbf{y}}\|^p (1 + \|\mathbf{x} + \hat{\mathbf{y}}\|^p) \int_{\mathbb{R}^n} \|\mathbf{z}\|^p \, d\mu.$$

Since $\mathbf{y} = \mathbf{x} + \hat{\mathbf{y}}$ is a unit vector, we obtain

$$W_p^p((\pi_{\mathbf{x}^{\perp}})_{\#}\mu, (\pi_{\mathbf{y}^{\perp}})_{\#}\mu) \leq 2^p \|\hat{\mathbf{y}}\|^p \int_{\mathbb{R}^n} \|\mathbf{z}\|^p \, d\mu = 2^p M_p(\mu) \|\mathbf{y} - \mathbf{x}\|^p,$$

which is the continuity of the p-Wasserstein distance for projections. To conclude the argument we note that the space of 1-codimensional subspaces in $\mathbb{R}^n$, identified with $\mathbb{RP}^{n-1}$ (via $\mathbf{x}^{\perp} \leftrightarrow \mathbf{x}$)

is compact. Compactness implies that the continuous function $\mathbf{x} \mapsto W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu)$ takes on its minimum at some point, say $\mathbf{x}_{\min}$. We already noticed that $W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) > 0$ for any subspace $\mathbf{x}^\perp$ of codimesion 1, hence

$$0 < c := W_p(\mu, (\pi_{\mathbf{x}_{\min}^\perp})_\# \mu) \leq W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu)$$

for all $\mathbf{x} \in \mathbb{S}^{n-1}$. Then if $\nu \in \mathscr{P}_p(\mathbb{R}^n)$ is so that $W_p(\mu, \nu) < W_p(\mu, (\pi_{\mathbf{x}_{\min}^\perp})_\# \mu) \leq W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu)$, we must have $supp(\nu) \not\subset \mathbf{x}^\perp$ for any unit vector $\mathbf{x}$. Then $W_p(\nu, (\pi_{\mathbf{x}^\perp})_\# \nu) > 0$ for any unit vector $\mathbf{x}$ and hence $\nu$ is a p-frame. In particular for a given p-frame $\mu$ the set $\{\nu \in \mathscr{P}_p(\mathbb{R}^n) : \; W_p(\mu, \nu) < W_p(\mu, (\pi_{\mathbf{x}_{\min}^\perp})_\# \mu)\}$ is an open ball of p-frames centered at the p-frame $\mu$ in the p-Wasserstein topology. □

For $p \neq 2$ the Wasserstein distances $W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu)$ may be difficult to determine, on the other hand there are some conversions:

**Corollary 2.3.3.** *For any unit-vector* $\mathbf{x} \in \mathbb{S}^{n-1}$ *and any* $p \geq 1$ *we have*

$$W_p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = W_p((\pi_{\mathbf{x}})_\# \mu, \delta_{\mathbf{0}}) \tag{2.3.2}$$

*and*

$$W_2^2(\mu, \delta_{\mathbf{0}}) = W_2^2(\mu, (\pi_{\mathbf{x}})_\# \mu) + W_2^2((\pi_{\mathbf{x}})_\# \mu, \delta_{\mathbf{0}}).$$

*Proof.* All one needs to notice is $|\langle \mathbf{x}, \mathbf{v} \rangle|^p = |\pi_{\mathbf{x}} \mathbf{v}|^p = \text{dist}^p(\pi_{\mathbf{x}} \mathbf{v}, 0)$, together with $|\langle \mathbf{x}, \mathbf{v} \rangle|^p = |\pi_{\mathbf{x}^\perp} \mathbf{v} - \mathbf{v}|^p$ from 2.3.1 gives:

$$\begin{aligned} W_p^p((\pi_{\mathbf{x}})_\# \mu, \delta_{\mathbf{0}}) = \int_{\mathbb{R}} \text{dist}^p(v, 0) \; d(\pi_{\mathbf{x}})_\# \mu(v) = \int_{\mathbb{R}^n} \text{dist}^p(\pi_{\mathbf{x}} \mathbf{v}, 0) \; d\mu(\mathbf{v}) = \\ = \int_{\mathbb{R}^n} |\mathbf{v} - \pi_{\mathbf{x}^\perp}(\mathbf{v})|^p \; d\mu(\mathbf{v}) = W_p^p(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu). \end{aligned} \tag{2.3.3}$$

The second statement is the Theorem of Pythagoras. □

If specifically a unit vector $\mathbf{x}$ in Proposition 2.1.2 is an eigenvector of $\mathbf{S}_\mu$, then the corresponding eigenvalue is given by $W_2^2(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu)$.

Expanding a vector $\mathbf{x} = (x_1, ..., x_n)$ in an eigen-basis $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ of $\mathbf{S}_\mu$ we get

**Corollary 2.3.4.** *If* $\mathbf{x} = (x_1, ..., x_n)$ *is a unit vector in eigen-coordinates then*

$$W_2^2(\mu, (\pi_{\mathbf{x}^\perp})_\# \mu) = \sum_{i=1}^n x_i^2 \cdot W_2^2(\mu, (\pi_{\mathbf{e}_i^\perp})_\# \mu).$$

*In particular if* $\mathbf{S}_\mu$ *is positive definite, then the vectors* $\frac{\mathbf{x}}{W_2(\mu,(\pi_{\mathbf{x}^\perp})_\#\mu)}$ *where* $\mathbf{x}$ *is a unit vector lie on the ellipsoid*

$$\sum_{i=1}^n x_i^2 \cdot W_2^2(\mu, (\pi_{\mathbf{e}_i^\perp})_\# \mu) = 1.$$

## 2.4 Wasserstein Distances for Frames: from Standard Estimates to Uniqueness

The theorems proved in this section are adapted and refined carry-overs of [Gelbrich, 1990] and particularly [Cuesta-Albertos et al., 1996] where instead of frame operators covariance operators are considered. There is little difference in the key arguments. However, our perspective is frame theoretic. One advantage is that the criterion for equality of the lower estimate for the Wasserstein distance in 2.1.4 is more direct and easier. This is because a frame operator is positive definite, while the covariance operator generally is not. In particular, we can avoid involving centered measures. This difference will become clear at the end of this section after building the background and statements.

In what follows, we need frame operator transforms under (linear) push-forwards. This is shown in several papers for example [Loukili and Maslouhi, 2020], we add the argument for the reader's convenience. Let $\mathbf{T}$ be a linear transformation of $\mathbb{R}^n$, and $\mu$ be a probabilistic frame, then the frame operator of $\mathbf{T}_\# \mu$ is given by

$$\begin{aligned} \langle \mathbf{x}, \mathbf{S}_{\mathbf{T}_\#\mu} \mathbf{x} \rangle = \int \langle \mathbf{x}, \mathbf{y} \rangle^2 \, d\mathbf{T}_\#\mu(\mathbf{y}) = \int \langle \mathbf{x}, \mathbf{T}\mathbf{y} \rangle^2 \, d\mu(\mathbf{y}) = \\ \int \langle \mathbf{T}^t \mathbf{x}, \mathbf{y} \rangle^2 \, d\mu(\mathbf{y}) = \langle \mathbf{T}^t \mathbf{x}, \mathbf{S}_\mu \mathbf{T}^t \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{T}\mathbf{S}_\mu \mathbf{T}^t \mathbf{x} \rangle. \end{aligned} \tag{2.4.1}$$

Since this identity holds for all $\mathbf{x} \in \mathbb{R}^n$ we have $\mathbf{S}_{\mathbf{T}_\#\mu} = \mathbf{T} \cdot \mathbf{S}_\mu \cdot \mathbf{T}^t$. Obviously $\mathbf{S}_{\mathbf{T}_\#\mu}$ is positive semi-definite and symmetric, since $\mathbf{S}_\mu$ is symmetric and positive definite. If $\mathbf{T}$ is invertible then $\mathbf{S}_{\mathbf{T}_\#\mu}$ is as well.

**Proposition 2.4.1.** *For any fixed* $\mathbf{S} \in \mathbb{S}_{++}^n$ *the congruence map* $f_{\mathbf{S}} : \mathbb{S}_+^n \to \mathbb{S}_+^n$ *given by* $f_{\mathbf{S}}(\mathbf{M}) :=$

**MSM**, *is bijective and its inverse is given by* $f_{\mathbf{S}}^{-1}(\mathbf{T}) = \mathbf{A}(\mathbf{S},\mathbf{T})$. *In particular* $\mathbf{A}^{-1}(\mathbf{T},\mathbf{S}) = \mathbf{A}(\mathbf{S},\mathbf{T})$.

*Proof.* Note that the image of $f_{\mathbf{S}}$ is always a symmetric positive semi-definite matrix. For given $\mathbf{T} \in \mathbb{S}_+^n$ let us solve $f_{\mathbf{S}}(\mathbf{M}) = \mathbf{T}$, i.e. solve $\mathbf{MSM} = \mathbf{T}$ for $\mathbf{M}$. Since $\mathbf{S} \in \mathbb{S}_{++}^n$ we can rewrite the previous equation as

$$\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2} = \mathbf{S}^{1/2}\mathbf{MSM}\mathbf{S}^{1/2} = \mathbf{S}^{1/2}\mathbf{M}\mathbf{S}^{1/2}\mathbf{S}^{1/2}\mathbf{M}\mathbf{S}^{1/2} = (\mathbf{S}^{1/2}\mathbf{M}\mathbf{S}^{1/2})^2.$$

Since $\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2} \in \mathbb{S}_+^n$ taking its root and solving for $\mathbf{M}$ gives

$$\mathbf{M} = \mathbf{S}^{-1/2}(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}\mathbf{S}^{-1/2} \in \mathbb{S}_+^n.$$

That the map is bijective follows from $f_{\mathbf{S}} \circ f_{\mathbf{S}}^{-1}(\mathbf{T}) = f_{\mathbf{S}}(\mathbf{A}(\mathbf{S},\mathbf{T})) = \mathbf{T}$ and $f_{\mathbf{S}}^{-1} \circ f_{\mathbf{S}}(\mathbf{M}) = A(\mathbf{S},\mathbf{MSM}) = \mathbf{M}$. The last identity follows directly from $\mathbf{S}^{1/2}\mathbf{MSM}\mathbf{S}^{1/2} = (\mathbf{S}^{1/2}\mathbf{M}\mathbf{S}^{1/2})^2$.

For the last statement one easily verifies that $f_{\mathbf{S}}(\mathbf{A}^{-1}(\mathbf{T},\mathbf{S})) = \mathbf{T}$, because $f_{\mathbf{S}}$ is a bijection the claim follows. □

Given two probabilistic frames $\mu,\nu$ with frame operators $\mathbf{S}_\mu$ and $\mathbf{S}_\nu$ let us write $\mathbf{A}_{\mu,\nu} := \mathbf{A}(\mathbf{S}_\mu,\mathbf{S}_\nu)$. Recall, the *center of mass* or *mean* of a measure $\mu$ is the vector $\mathbf{m}_\mu = \int_{\mathbb{R}^n} \mathbf{v}\, d\mu(\mathbf{v})$. Then the *centered measure* of $\mu$ is given by $\overline{\mu}(A) := \mu(A + \mathbf{m}_\mu)$ for any Borel set A. Recall the covariance matrix of $\mu$ is given by $\mathbf{\Sigma}_\mu = \mathbf{S}_{\overline{\mu}}$. Note that this is generally an abuse of language because $\mathbf{\Sigma}_\mu$ is not necessarily invertible, i.e. $\mathbf{S}_{\overline{\mu}}$ is not necessarily definite. In particular, a centered probabilistic frame is not necessarily a probabilistic frame and in this case $\mathbf{S}_{\overline{\mu}}^{-1/2}$, respectively $\mathbf{\Sigma}_\mu^{-1/2}$, will be the Moore-Penrose inverse. If $\mathbf{\Pi}_\mu$ is the (matrix version of the) orthogonal projection onto Im $\mathbf{S}_\mu$, then the Moore-Penrose inverse has the property $\mathbf{\Pi}_\mu = \mathbf{S}_\mu\mathbf{S}_\mu^{-1} = \mathbf{S}_\mu^{-1}\mathbf{S}_\mu$, see [Gelbrich, 1990]. With that in mind we have

$$\mathbf{A}_{\overline{\mu},\overline{\nu}} = \mathbf{A}(\mathbf{\Sigma}_\mu,\mathbf{\Sigma}_\nu) = \mathbf{\Sigma}_\mu^{-1/2}(\mathbf{\Sigma}_\mu^{1/2}\mathbf{\Sigma}_\nu\mathbf{\Sigma}_\mu^{1/2})^{1/2}\mathbf{\Sigma}_\mu^{-1/2}.$$

These matrices have somewhat surprising properties that may not seem obvious at first glance. We would like to mention that the first part of the formula is shown for a special case in [Loukili and Maslouhi, 2020].

**Lemma 2.4.2.** *Let* $\mu,\nu \in \mathscr{P}_2(\mathbb{R}^n)$, *not necessarily frames, then:*

1. *If* $\mathbf{S} \in \mathbb{S}_+^n$ *then* $\mathbf{A}_{\mu,\mathbf{S}_\#\mu} = \mathbf{\Pi}_\mu \mathbf{S} \mathbf{\Pi}_\mu$*, and if* $\mu$ *is a frame then* $\mathbf{A}_{\mu,\mathbf{S}_\#\mu} = \mathbf{S}$.

2. *If* $\nu = (\mathbf{A}_{\mu,\nu})_\#\mu$*, then* $(\mathbf{\Pi}_{\overline{\mu}})_\#\overline{\nu} = (\mathbf{A}_{\overline{\mu},\overline{\nu}})_\#\overline{\mu}$.

*Proof.* For the first statement, since $\mathbf{S}_\mu^{1/2}\mathbf{S}\mathbf{S}_\mu\mathbf{S}\mathbf{S}_\mu^{1/2} = (\mathbf{S}_\mu^{1/2}\mathbf{S}\mathbf{S}_\mu^{1/2})^2$, by symmetry of $\mathbf{S}_\mu^{1/2}$ and the fact that $\text{Im } \mathbf{S}_\mu = \text{Im } \mathbf{S}_\mu^{1/2}$, we have

$$\mathbf{A}_{\mu,\mathbf{S}_\#\mu} = \mathbf{S}_\mu^{-1/2}(\mathbf{S}_\mu^{1/2}\mathbf{S}\mathbf{S}_\mu\mathbf{S}\mathbf{S}_\mu^{1/2})^{1/2}\mathbf{S}_\mu^{-1/2} = \mathbf{S}_\mu^{-1/2}\mathbf{S}_\mu^{1/2}\mathbf{S}\mathbf{S}_\mu^{1/2}\mathbf{S}_\mu^{-1/2} = \mathbf{\Pi}_\mu\mathbf{S}\mathbf{\Pi}_\mu.$$

If $\mu$ is a frame, then $\mathbf{S}_\mu \in \mathbb{S}_{++}^n$, hence $\mathbf{\Pi}_\mu = \mathbf{Id}$.

For the second identity recall that $\overline{(\mathbf{A}_{\mu,\nu})_\#\mu} = (\mathbf{A}_{\mu,\nu})_\#\overline{\mu}$. Using $(\mathbf{\Pi}_{\overline{\mu}})_\#\overline{\mu} = \overline{\mu}$ and the previous formula we get

$$\begin{aligned}(\mathbf{A}_{\overline{\mu},\overline{\nu}})_\#\overline{\mu} &= (\mathbf{A}_{\overline{\mu},\overline{(\mathbf{A}_{\mu,\nu})_\#\mu}})_\#\overline{\mu} = (\mathbf{\Pi}_{\overline{\mu}}\mathbf{A}_{\mu,\nu}\mathbf{\Pi}_{\overline{\mu}})_\#\overline{\mu}\\ &= (\mathbf{\Pi}_{\overline{\mu}})_\#(\mathbf{A}_{\mu,\nu})_\#(\mathbf{\Pi}_{\overline{\mu}})_\#\overline{\mu} = (\mathbf{\Pi}_{\overline{\mu}})_\#(\mathbf{A}_{\mu,\nu})_\#\overline{\mu} = (\mathbf{\Pi}_{\overline{\mu}})_\#\overline{(\mathbf{A}_{\mu,\nu})_\#\mu}.\end{aligned}$$

□

Statement 2 of the Lemma is expected from properties that will become clearer below. In fact, part 2 verifies that the equality condition, $\nu = (\mathbf{A}_{\mu,\nu})_\#\mu$, for the respective inequalities in propositions 2.4.3 and 2.4.6 implies the one for the centered measures, $(\mathbf{\Pi}_{\overline{\mu}})_\#\overline{\nu} = (\mathbf{A}_{\overline{\mu},\overline{\nu}})_\#\overline{\mu}$, as used in the inequalities of [Gelbrich, 1990] and [Cuesta-Albertos et al., 1996].

**Proposition 2.4.3.** *Suppose* $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^n)$*. For any unit vector* $\mathbf{x}$ *we have*

$$W_2^2((\pi_\mathbf{x})_\#\mu, (\pi_\mathbf{x})_\#\nu) \geq \left(W_2(\mu, (\pi_{\mathbf{x}^\perp})_\#\mu) - W_2(\nu, (\pi_{\mathbf{x}^\perp})_\#\nu)\right)^2, \tag{2.4.2}$$

*and if* $\{e_1, ..., e_n\}$ *is an orthonormal basis then*

$$W_2^2(\mu,\nu) \geq \sum_{i=1}^n \left(W_2(\mu, (\pi_{e_i^\perp})_\#\mu) - W_2(\nu, (\pi_{e_i^\perp})_\#\nu)\right)^2. \tag{2.4.3}$$

*Furthermore, equality holds if* $\nu = \mathbf{T}_\#\mu$ *where* $\mathbf{T}$ *is given by a symmetric, positive semi-definite matrix with eigenbasis* $\{\mathbf{e}_i\}$ *of* $\mathbf{T}$.

*Proof.* Abbreviating $\Gamma := \Gamma(\mu, \nu)$ one has

$$\begin{aligned} W_2^2(\mu,\nu) &= \inf_{\gamma\in\Gamma} \int_{\mathbb{R}^n\times\mathbb{R}^n} \|\mathbf{x}-\mathbf{y}\|^2 \, d\gamma = \inf_{\gamma\in\Gamma} \sum_{i=1}^n \int_{\mathbb{R}^n\times\mathbb{R}^n} |x_i - y_i|^2 \, d\gamma \\ &= \inf_{\gamma\in\Gamma} \sum_{i=1}^n \int_{\mathbb{R}\times\mathbb{R}} |x-y|^2 \, d(\pi_{e_i}\times\pi_{e_i})_{\#}\gamma \geq \sum_{i=1}^n W_2^2((\pi_{e_i})_{\#}\mu, (\pi_{e_i})_{\#}\nu) \end{aligned} \tag{2.4.4}$$

Now if $\gamma_i \in \Gamma((\pi_{e_i})_{\#}\mu, (\pi_{e_i})_{\#}\nu)$ minimizes $W_2^2((\pi_{e_i})_{\#}\mu, (\pi_{e_i})_{\#}\nu)$ then by the reverse triangle inequality (in $L^2$):

$$\begin{aligned} W_2^2((\pi_{e_i})_{\#}\mu, (\pi_{e_i})_{\#}\nu) &= \int_{\mathbb{R}\times\mathbb{R}} |x-y|^2 \, d\gamma_i \\ &\geq \left( \left( \int_{\mathbb{R}} |x|^2 \, d(\pi_{e_i})_{\#}\mu \right)^{1/2} - \left( \int_{\mathbb{R}} |y|^2 \, d(\pi_{e_i})_{\#}\nu \right)^{1/2} \right)^2 \\ &= \left( \left( \int_{\mathbb{R}^n} \langle \mathbf{x}, e_i\rangle^2 \, d\mu \right)^{1/2} - \left( \int_{\mathbb{R}^n} \langle \mathbf{y}, e_i\rangle^2 \, d\nu \right)^{1/2} \right)^2 \\ &= \left( W_2(\mu, (\pi_{e_i^\perp})_{\#}\mu) - W_2(\nu, (\pi_{e_i^\perp})_{\#}\nu) \right)^2, \end{aligned} \tag{2.4.5}$$

one obtains

$$W_2^2(\mu,\nu) \geq \sum_{i=1}^n \left( W_2(\mu, (\pi_{e_i^\perp})_{\#}\mu) - W_2(\nu, (\pi_{e_i^\perp})_{\#}\nu) \right)^2.$$

Equation 2.4.2 is obtained by replacing $e_i$ with $\mathbf{x}$ in the above proof.

Writing square terms out in inequality 2.4.5 and using the marginals of $\gamma_i$ equality amounts to have equality in the following inequality:

$$\int_{\mathbb{R}\times\mathbb{R}} xy \, d\gamma_i \leq \left( \int_{\mathbb{R}} x^2 \, d(\pi_{e_i})_{\#}\mu \right)^{1/2} \left( \int_{\mathbb{R}} y^2 \, d(\pi_{e_i})_{\#}\nu \right)^{1/2}$$

Like for the Cauchy-Schwarz inequality, this inequality is an equality if $y = \lambda_i x$ for some $\lambda_i \geq 0$ and the marginal distributions agree (almost everywhere). In this case $\gamma_i$ is supported on a graph of a line and hence a push-forward $\gamma_i = (Id \times \lambda_i Id)_{\#}(\pi_{\mathbf{e}_i})_{\#}\mu$ where $Id$ is the identity map on $\mathbb{R}$. If this is the case inequality 2.4.4 is also an equality for $\nu = \mathbf{T}_{\#}\mu$ where $\mathbf{T}$ is linear maximizing the Cauchy-Schwarz inequalities in the respective eigendirections, i.e., the push forward is by the map given by a positive semi-definite symmetric matrix with eigenvalues $\lambda_i \geq 0$ in direction $\{e_i\}$. If directions with $\lambda_i = 0$ appear we have a projection. More specifically,

$$\begin{aligned} W_2^2(\mu, \mathbf{T}_\#\mu) &= \int_{\mathbb{R}^n} \|\mathbf{x} - \mathbf{T}\mathbf{x}\|^2 \; d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} \| \sum_{i=1}^n \langle \mathbf{x}, \mathbf{e}_i \rangle (\mathbf{e}_i - \mathbf{T}\mathbf{e}_i) \|^2 \; d\mu(\mathbf{x}) \\ &= \sum_{i=1}^n \int_{\mathbb{R}^n} |\langle \mathbf{x}, e_i \rangle (1 - \lambda_i)|^2 \; d\mu(\mathbf{x}) = \sum_{i=1}^n \int_{\mathbb{R}\times\mathbb{R}} |x - y|^2 \; d\gamma_i \\ &= \sum_{i=1}^n W_2^2((\pi_{e_i})_\#\mu, (\pi_{e_i})_\#\nu) = \sum_{i=1}^n \left( W_2(\mu, (\pi_{e_i{}^\perp})_\#\mu) - W_2(\nu, (\pi_{e_i{}^\perp})_\#\nu) \right)^2 . \end{aligned}$$

□

With Proposition 2.4.3, we can show the continuity of the frame map (see also [Wickman and Okoudjou, 2023]).

**Corollary 2.4.4.** *The frame map $S : \mathscr{P}_2 \to \mathbb{S}_+^n$ is continuous with respect to the Wasserstein topology, and hence weak-$*$ topology, on $\mathscr{P}_2(\mathbb{R}^n)$. More precisely $\|\mathbf{S}_\mu^{1/2} - \mathbf{S}_\nu^{1/2}\|_{op} \le W_2(\mu, \nu)$ with respect to the operator norm $\|\cdot\|_{op}$. In particular $\|\mathbf{S}^{1/2} - \mathbf{T}^{1/2}\|_{op} \le W_2(\mathscr{P}_\mathbf{S}, \mathscr{P}_\mathbf{T}) = d_W(\mathbf{S}, \mathbf{T})$.*

*Proof.* Take $\mu, \nu \in \mathscr{P}_2$ with frame operators $\mathbf{S}_\mu$ and $\mathbf{S}_\nu$ respectively. Let $\{\mathbf{e}_i\}$ an orthogonal basis in which (the symmetric matrix) $\mathbf{S}_\mu^{1/2} - \mathbf{S}_\nu^{1/2}$ is diagonal. Then $\sup_{\mathbf{y}\in\mathbf{S^{n-1}}} |\mathbf{y}^t(\mathbf{S}_\mu^{1/2} - \mathbf{S}_\nu^{1/2})\mathbf{y}| = |\mathbf{x}^t(\mathbf{S}_\mu^{1/2} - \mathbf{S}_\nu^{1/2})\mathbf{x}|$ for some unit vector $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$, so that particularly $|x_i| \le 1$. Now

$$\begin{aligned} \|\mathbf{S}_\mu^{1/2} - \mathbf{S}_\nu^{1/2}\|^2 &= \sup_{\|\mathbf{y}\|=\mathbf{1}} |\mathbf{y}^t(\mathbf{S}_\mu^{1/2} - \mathbf{S}_\nu^{1/2})\mathbf{y}|^2 \\ &\le \sum_{i=1}^n x_i^4 \sum_{i=1}^n (\mathbf{e}_i^t(\mathbf{S}_\mu^{1/2} - \mathbf{S}_\nu^{1/2})\mathbf{e}_i)^2 \le \sum_{i=1}^n (\langle \mathbf{e}_i, \mathbf{S}_\mu^{1/2}\mathbf{e}_i \rangle - \langle \mathbf{e}_i, \mathbf{S}_\nu^{1/2}\mathbf{e}_i \rangle)^2 \\ &= \sum_{i=1}^n \left( W_2(\mu, (\pi_{\mathbf{e}_i{}^\perp})_\#\mu) - W_2(\nu, (\pi_{\mathbf{e}_i{}^\perp})_\#\nu) \right)^2 \le W_2^2(\mu, \nu) \end{aligned}$$

The first inequality uses Cauchy-Schwarz and the last step is estimate 2.4.3. We see $f(\mu) := \mathbf{S}_\mu^{1/2}$ is continuous, so that $f^2$ is continuous too. The last statement follows from the definition of $W_2(\mathscr{P}_\mathbf{S}, \mathscr{P}_\mathbf{T}) = d(\mathbf{S}, \mathbf{T})$ in the introduction. That shows the claim. □

Given a probability $\mu$ the p-th (central) moment $M_p(\mu)$ is given by $\int_{\mathbb{R}^n} \|x\|^p \; d\mu(x)$. Right from the definitions, one easily confirms the well-known formula

$$M_2(\mu) = \sum_{i=1}^n W_2^2(\mu, (\pi_{\mathbf{e}_i^\perp})_\#\mu) = \mathrm{tr}\; \mathbf{S}_\mu \tag{2.4.6}$$

for any orthogonal-basis $\{e_i\}$ of $\mathbb{R}^n$. Indeed

$$\operatorname{tr} \mathbf{S}_\mu = \sum_{i=1}^n \langle \mathbf{e}_i, \mathbf{S}_\mu \mathbf{e}_i \rangle = \sum_{i=1}^n \int_{\mathbb{R}^n} \langle \mathbf{e}_i, \mathbf{v} \rangle^2 d\mu = \int_{\mathbb{R}^n} \|\mathbf{v}\|^2 d\mu(\mathbf{v}) = M_2(\mu).$$

Specifying a particular symmetric positive definite map in the previous proposition gives Gelbrich's estimate, see [Gelbrich, 1990]. Here we have Gelbrichs estimate again for frame operators instead of covariance operators. Again the proof is formally the same as the proof of Theorem 2.1 in [Cuesta-Albertos et al., 1996].

**Corollary 2.4.5** (Gelbrich's bound for frame operators [Gelbrich, 1990])**.** *Let $\mu, \nu \in \mathscr{P}_{++}$ with respective frame operators $\mathbf{S}_\mu$ and $\mathbf{S}_\nu$, then*

$$W_2^2(\mu,\nu) \geq \operatorname{tr}(\mathbf{S}_\mu + \mathbf{S}_\nu - 2(\mathbf{S}_\mu^{1/2}\mathbf{S}_\nu\mathbf{S}_\mu^{1/2})^{1/2}) = \operatorname{tr} \mathbf{S}_\mu(\mathbf{Id} - \mathbf{A}_{\mu,\nu})^2. \tag{2.4.7}$$

*Equality holds if $\nu = \mathbf{A}(\mathbf{S}_\mu, \mathbf{S}_\nu)_{\#}\mu$.*

*Proof.* Given inequality 2.4.3 of proposition 2.4.3 the statement will follow from the formula

$$\operatorname{tr} (\mathbf{S}_\mu^{1/2}\mathbf{S}_\nu\mathbf{S}_\mu^{1/2})^{1/2} = \sum_{i=1}^n \langle \mathbf{e}_i, \mathbf{S}_\nu \mathbf{e}_i \rangle^{1/2} \langle \mathbf{e}_i, \mathbf{S}_\mu \mathbf{e}_i \rangle^{1/2}.$$

for some orthogonal basis $\{\mathbf{e}_i\}$ of $\mathbb{R}^n$. Note, that the right hand side of the stated identity immediately follows from the right hand side of 2.4.3 using $W_2(\mu, (\pi_{e_i^\perp})_{\#}\mu) = \langle \mathbf{e}_i, \mathbf{S}_\nu \mathbf{e}_i \rangle^{1/2}$ and the respective formula for $\nu$. From proposition 2.4.1 we recall $\mathbf{A}_{\mu,\nu} = \mathbf{A}(\mathbf{S}_\mu, \mathbf{S}_\nu)$ symmetric positive definite, so that $\mathbf{S}_\nu = \mathbf{A}_{\mu,\nu}\mathbf{S}_\mu\mathbf{A}_{\mu,\nu}$. Let $\{\mathbf{e}_i\}$ be an eigen-basis for $\mathbf{A}_{\mu,\nu}$ with corresponding set of (positive) eigenvalues $\{\lambda_i\}$, then:

$$\langle \mathbf{e}_i, \mathbf{S}_\nu \mathbf{e}_i \rangle = \langle \mathbf{e}_i, (\mathbf{A}_{\mu,\nu}\ \mathbf{S}_\mu \mathbf{A}_{\mu,\nu})\mathbf{e}_i \rangle = \langle \mathbf{A}_{\mu,\nu}\mathbf{e}_i, \mathbf{S}_\mu \mathbf{A}_{\mu,\nu}\mathbf{e}_i \rangle = \lambda_i^2 \langle \mathbf{e}_i, \mathbf{S}_\mu \mathbf{e}_i \rangle.$$

Taking roots on both sides and using $\mathbf{A}_{\mu,\nu} = \mathbf{S}_\mu^{-1/2}(\mathbf{S}_\mu^{1/2}\mathbf{S}_\nu\mathbf{S}_\mu^{1/2})^{1/2}\mathbf{S}_\mu^{-1/2}$ from Proposition 2.4.1, formal properties of the trace give the sought identity:

$$\begin{aligned}\operatorname{tr} (\mathbf{S}_\mu^{1/2}\mathbf{S}_\nu\mathbf{S}_\mu^{1/2})^{1/2} &= \operatorname{tr} (\mathbf{S}_\mu^{1/2}\mathbf{A}_{\mu,\nu}\mathbf{S}_\mu^{1/2}) = \operatorname{tr} (\mathbf{S}_\mu \mathbf{A}_{\mu,\nu}) = \\ &= \sum_{i=1}^n \lambda_i \langle \mathbf{e}_i, \mathbf{S}_\mu \mathbf{e}_i \rangle = \sum_{i=1}^n \langle \mathbf{e}_i, \mathbf{S}_\nu \mathbf{e}_i \rangle^{1/2} \langle \mathbf{e}_i, \mathbf{S}_\mu \mathbf{e}_i \rangle^{1/2}.\end{aligned}$$

Putting things together we obtain the stated estimate

$$W_2^2(\mu,\nu) \geq \operatorname{tr}(\mathbf{S}_\mu + \mathbf{S}_\nu - 2(\mathbf{S}_\mu^{1/2}\mathbf{S}_\nu\mathbf{S}_\mu^{1/2})^{1/2}) = \sum_{i=1}^n (1-\lambda_i)^2 \langle \mathbf{e}_i, \mathbf{S}_\mu \mathbf{e}_i \rangle$$
$$= \sum_{i=1}^n \langle \mathbf{e}_i, (\mathbf{Id} - \mathbf{A}_{\mu,\nu})\mathbf{S}_\mu(\mathbf{Id} - \mathbf{A}_{\mu,\nu})\mathbf{e}_i \rangle = \operatorname{tr}\, \mathbf{S}_\mu(\mathbf{Id} - \mathbf{A}_{\mu,\nu})^2.$$

Then by proposition 2.4.3 we have equality, if $\nu = \mathbf{A}(\mathbf{S}_\mu, \mathbf{S}_\nu)_\#\mu$. □

We use this optimization result below to compare traces of certain matrices.

**Proposition 2.4.6.** *Let $\mathbf{S}, \mathbf{T} \in \mathbf{S}^n_{++}$ be frame operators. Then for every $\mu \in \mathscr{P}_\mathbf{S}$ the push-forward $\mathbf{A}(\mathbf{S},\mathbf{T})_\#\mu$ is the unique frame with frame operator $\mathbf{T}$, so that $W_2(\mu, \mathbf{A}(\mathbf{S},\mathbf{T})_\#\mu) = W_2(\mu, \mathscr{P}_\mathbf{T})$.*

*Proof.* By Theorem 2.1.4, we know that for any given $\mathbf{S}, \mathbf{A} \in \mathbb{S}^n_{++}$ and $\mu \in \mathscr{P}_\mathbf{S}$,

$$W_2^2(\mu, \mathbf{A}_\#\mu) = W_2^2(\mu, \mathscr{P}_{\mathbf{ASA}}) = \operatorname{tr}\, \mathbf{S}(\mathbf{Id} - \mathbf{A})^2$$

Now we apply the above theorem to $\mathbf{A} \in \mathbb{S}^n_{++}$ that solves $\mathbf{T} = \mathbf{ASA}$ given by

$$\mathbf{A} = \mathbf{A}(\mathbf{S},\mathbf{T}) := \mathbf{S}^{-1/2}(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}\mathbf{S}^{-1/2}.$$

We now extend an argument that was presented for the special case of $\mathbf{T} = \mathbf{Id}$ by [Loukili and Maslouhi, 2020] to show the uniqueness. Consider the push forward $\mathbf{A}(\mathbf{S},\mathbf{T})_\#\mu$ and assume $\nu$ with frame operator $\mathbf{T}$ minimizes the 2-Wasserstein distance to $\mu$. Then, if $\gamma$ is a coupling between $\nu$ and $\mu$ so that $W_2(\mu,\nu) = W_2(\mu, \mathscr{P}_\mathbf{T})$, its push forward by $\mathbf{Id} \times \mathbf{A}(\mathbf{S},\mathbf{T})$ is a coupling between $\nu$ and $\mathbf{A}(\mathbf{S},\mathbf{T})_\#\mu$ with respective frame operator

$$\mathbf{S}_{(\mathbf{Id}\times\mathbf{A}(\mathbf{S},\mathbf{T}))_\#\gamma} = \begin{bmatrix} \mathbf{Id} & 0 \\ 0 & \mathbf{A}(\mathbf{S},\mathbf{T}) \end{bmatrix} \begin{bmatrix} \mathbf{T} & \mathbf{T}\cdot\mathbf{A}(\mathbf{T},\mathbf{S}) \\ (\mathbf{T}\cdot\mathbf{A}(\mathbf{T},\mathbf{S}))^t & \mathbf{S} \end{bmatrix} \begin{bmatrix} \mathbf{Id} & 0 \\ 0 & \mathbf{A}(\mathbf{S},\mathbf{T}) \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{T} & \mathbf{T}\cdot\mathbf{A}(\mathbf{T},\mathbf{S})\cdot\mathbf{A}(\mathbf{S},\mathbf{T}) \\ (\mathbf{T}\cdot\mathbf{A}(\mathbf{T},\mathbf{S})\cdot\mathbf{A}(\mathbf{S},\mathbf{T}))^t & \mathbf{A}(\mathbf{S},\mathbf{T})\cdot\mathbf{S}\cdot\mathbf{A}(\mathbf{S},\mathbf{T}) \end{bmatrix} = \begin{bmatrix} \mathbf{T} & \mathbf{T} \\ \mathbf{T} & \mathbf{T} \end{bmatrix}$$

so that

$$W_2^2(\mathbf{A}(\mathbf{S},\mathbf{T})_\#\mu, \nu) \leq \operatorname{tr}(\mathbf{T} + \mathbf{T} - 2\mathbf{T}) = 0.$$

Hence $\mathbf{A}(\mathbf{S},\mathbf{T})_{\#}\mu = \nu$.

□

*Proof of Theorem 2.1.4.* Push-forward with a continuous map is continuous, and in particular, if the push-forward is by a linear $\mathbf{A} \in \mathbb{S}^n_{++}$, then push-forward with $\mathbf{A}^{-1} \in \mathbb{S}^n_{++}$ provides a continuous inverse. Hence the push-forward $\mathbf{A}_{\#} : \mathscr{P}_{\mathbf{S}} \to \mathscr{P}_{\mathbf{ASA}}$ defines a homeomorphism.

The equation for the Wasserstein distance follows from Corollary 2.4.5. The particular shape of that formula for push-forwards with $\mathbf{A} \in \mathbb{S}^n_{++}$ follows from the first identity in Lemma 2.4.2, i.e., $\mathbf{A}_{\mu,\mathbf{A}_{\#}\mu} = \mathbf{A}$. Finally, the push-forward with $\mathbf{A} \in \mathbb{S}^n_{++}$ is the unique minimizer as shown in Proposition 2.4.6, again using the first statement of Lemma 2.4.2 to adapt to the situation stated in the Theorem. Therefore, $W_2(\mu,\nu) > W_2(\mu,\mathbf{A}_{\#}\mu)$ for any $\nu \in \mathscr{P}_{\mathbf{ASA}}$ so that $\nu \neq \mathbf{A}_{\#}\mu$. □

We now derive the matrix optimizing problem of Olkin and Pukelsheim [Olkin and Pukelsheim, 1982] which we need afterwards in two instances. Given two probabilistic frames $\mu$ with frame operator $\mathbf{S}$ and $\nu$ with frame operator $\mathbf{T}$ respectively, then

$$\begin{aligned} W_2^2(\mu,\nu) &= \inf_{\gamma} \int_{\mathbb{R}^{2n}} \|\mathbf{x}-\mathbf{y}\|^2 d\gamma(\mathbf{x},\mathbf{y}) \\ &= \int_{\mathbb{R}^n} \|\mathbf{x}\|^2 \, d\mu + \int_{\mathbb{R}^n} \|\mathbf{y}\|^2 \, d\nu - 2\sup_{\gamma} \int_{\mathbb{R}^{2n}} \langle \mathbf{x},\mathbf{y}\rangle \, d\gamma(\mathbf{x},\mathbf{y}) \end{aligned}$$

The frame operator of any $\gamma \in \Gamma(\mu,\nu)$ is $\mathbf{S}_\gamma = \int_{\mathbb{R}^{2n}} (\mathbf{x},\mathbf{y})\cdot(\mathbf{x},\mathbf{y})^t \, d\gamma(\mathbf{x},\mathbf{y})$, that converted into matrix form is

$$\mathbf{S}_\gamma = \begin{bmatrix} \mathbf{S} & \mathbf{\Psi} \\ \mathbf{\Psi}^t & \mathbf{T} \end{bmatrix}, \quad \text{where } \mathbf{\Psi} = \int_{\mathbb{R}^{2n}} \mathbf{x}\cdot\mathbf{y}^t \, d\gamma(\mathbf{x},\mathbf{y}).$$

Note, that

$$\mathrm{tr}\mathbf{\Psi} = \mathrm{tr} \int_{\mathbb{R}^{2n}} \mathbf{x}\cdot\mathbf{y}^t \, d\gamma(\mathbf{x},\mathbf{y}) = \int_{\mathbb{R}^{2n}} \langle \mathbf{x},\mathbf{y}\rangle \, d\gamma(\mathbf{x},\mathbf{y}),$$

so that the previous equation for the Wasserstein distance implies for any coupling $\gamma \in \Gamma(\mu,\nu)$:

$$W_2^2(\mu,\mathscr{P}_{\mathbf{T}}) \le \mathrm{tr}(\mathbf{S}+\mathbf{T}-2\mathbf{\Psi}).$$

The following matrix optimization problem is to determine the above matrix among the set of semi-definite matrices with fixed diagonal entries $\mathbf{S}$ and $\mathbf{T}$, so that tr $\mathbf{\Psi}$ is maximal. An optimal matrix is given by the frame matrix of a push forward and turns the estimate for the Wasserstein distance into an equality. The solution of the matrix problem is going back to Olkin and Pukelsheim [Olkin and Pukelsheim, 1982], see also Gelbrich [Gelbrich, 1990].

**Proposition 2.4.7.** *Given* $\mathbf{S}, \mathbf{T} \in \mathscr{P}_{++}$*, then the* $\mathbf{\Psi}$ *in*

$$\mathbf{S}_\gamma = \begin{bmatrix} \mathbf{S} & \mathbf{\Psi} \\ \mathbf{\Psi}^t & \mathbf{T} \end{bmatrix} \in \mathbb{S}_+^{2n}$$

*so that* tr $\mathbf{\Psi}$ *is maximal is given by* $\mathbf{\Psi} = \mathbf{SA}(\mathbf{S}, \mathbf{T})$.

*Proof.* From the discussion before the proposition, it is clear that any $\mathbf{\Psi}$, so that

$$W_2^2(\mu, \mathscr{P}_{\mathbf{T}}) = \mathrm{tr}(\mathbf{S} + \mathbf{T} - 2\mathbf{\Psi})$$

provides a semi-definite matrix of the form stated in the proposition with a maximaltr $\mathbf{\Psi}$. From Proposition 2.4.6, we have seen that $\mathbf{A}(\mathbf{S}, \mathbf{T})_{\#}\mu$ is the unique frame with frame operator $\mathbf{T}$ so that $W_2(\mu, \mathscr{P}_{\mathbf{T}}) = W_2(\mu, \mathbf{A}(\mathbf{S}, \mathbf{T})_{\#}\mu)$, where the push forward of $\mu$ by $\mathbf{Id} \times \mathbf{A}(\mathbf{S}, \mathbf{T})$ provides an optimal coupling between any $\mu$ with frame operator $\mathbf{S}$ and $\mathbf{A}(\mathbf{S}, \mathbf{T})_{\#}\mu$ with frame operator $\mathbf{T}$. Calculating the frame operator of the push forward translates into

$$\mathbf{S}_\gamma = \begin{bmatrix} \mathbf{Id} \\ \mathbf{A}(\mathbf{S}, \mathbf{T}) \end{bmatrix} \mathbf{S} \begin{bmatrix} \mathbf{Id} & \mathbf{A}(\mathbf{S}, \mathbf{T}) \end{bmatrix} = \begin{bmatrix} \mathbf{S} & \mathbf{SA}(\mathbf{S}, \mathbf{T}) \\ (\mathbf{SA}(\mathbf{S}, \mathbf{T}))^t & \mathbf{T} \end{bmatrix},$$

so that indeed

$$W_2^2(\mu, \mathbf{A}(\mathbf{S}, \mathbf{T})_{\#}\mu) = \mathrm{tr}(\mathbf{S} + \mathbf{T} - 2\mathbf{SA}(\mathbf{S}, \mathbf{T})) = \mathrm{tr}(\mathbf{S} + \mathbf{T} - 2(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}).$$

□

We are now in a position to show Proposition 2.1.5.

*Proof of Proposition 2.1.5.* Recall that $\mathbf{\Psi}$ so that tr $\mathbf{\Psi}$ is maximal under the condition

$$\begin{bmatrix} \mathbf{S} & \mathbf{\Psi} \\ \mathbf{\Psi}^t & \mathbf{T} \end{bmatrix} \in \mathbb{S}_+^{2n}$$

is given by $\Psi = \mathbf{S}^{1/2}(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}\mathbf{S}^{-1/2}$ with maximal value $\mathrm{tr}(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}$, alternatively see [Gelbrich, 1990], or [Olkin and Pukelsheim, 1982]. For the matrix $\mathbf{\Psi} = \mathbf{S}^{1/2}\mathbf{T}^{1/2}$ the identity $\mathbf{\Psi}\mathbf{T}^{-1}\mathbf{\Psi}^t = \mathbf{S}$ holds, hence $\mathbf{S} - \mathbf{\Psi}\mathbf{T}^{-1}\mathbf{\Psi}^t \geq 0$. By Lemma 1 in [Olkin and Pukelsheim, 1982]

$$\begin{bmatrix} \mathbf{S} & \mathbf{S}^{1/2}\mathbf{T}^{1/2} \\ (\mathbf{S}^{1/2}\mathbf{T}^{1/2})^t & \mathbf{T} \end{bmatrix} \in \mathbb{S}^{2n}_+.$$

Hence by the main result in [Olkin and Pukelsheim, 1982], or our own results above, we must have have $\mathrm{tr}\ \mathbf{S}^{1/2}\mathbf{T}^{1/2} \leq \mathrm{tr}\ (\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}$ and hence from the proof of Proposition 2.4.6, we conclude

$$\begin{aligned} W_2^2(\mathscr{P}_\mathbf{S}, \mathscr{P}_\mathbf{T}) &= \mathrm{tr}(\mathbf{S} + \mathbf{T} - 2(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}) \\ &\leq \mathrm{tr}(\mathbf{S} + \mathbf{T} - 2(\mathbf{S}^{1/2}\mathbf{T}^{1/2})) = \mathrm{tr}(\mathbf{S}^{1/2} - \mathbf{T}^{1/2})^2 = \|\mathbf{S}^{1/2} - \mathbf{T}^{1/2}\|_F^2. \end{aligned}$$

We add the arguments to show that $d_W$ is a metric on $\mathbb{S}^n_+$, but treat the case $\mathbb{S}^n_{++}$ first. Clearly $W_2(\mathscr{P}_\mathbf{T}, \mathscr{P}_\mathbf{S}) \geq 0$ and equality happens if and only if $\mathbf{T} = \mathbf{S}$. Symmetry is clear as well since $W_2$ is a metric. For the triangle inequality let $\mathbf{P} \in \mathbb{S}^n_{++}$ and consider $\mu \in \mathscr{P}_p(\mathbb{R}^n)$, then

$$\begin{aligned} W_2(\mathscr{P}_\mathbf{T}, \mathscr{P}_\mathbf{S}) &\leq W_2(\mathbf{A}(\mathbf{P}, \mathbf{T})_{\#}\mu, \mathbf{A}(\mathbf{P}, \mathbf{S})_{\#}\mu) \\ &\leq W_2(\mathbf{A}(\mathbf{P}, \mathbf{T})_{\#}\mu, \mu) + W_2(\mu, \mathbf{A}(\mathbf{P}, \mathbf{S})_{\#}\mu) \\ &= W_2(\mathscr{P}_\mathbf{T}, \mathscr{P}_\mathbf{P}) + W_2(\mathscr{P}_\mathbf{P}, \mathscr{P}_\mathbf{S}). \end{aligned}$$

Regarding the equivalences of topologies, note that all norms on a finite dimensional vector space are equivalent. Now consider the metrics $d_{op}(\mathbf{S}, \mathbf{T}) := \|\mathbf{S}^{1/2} - \mathbf{T}^{1/2}\|_{op}$, $d_F(\mathbf{S}, \mathbf{T}) := \|\mathbf{S}^{1/2} - \mathbf{T}^{1/2}\|_F$ respectively together with the lower estimate from Corollary 2.4.4 to finish the estimate.

As for the symmetric representation of the metric note that transporting from measures with frame operator $\mathbf{T}$ to $\mathbf{S}$ instead of the other way round leads to an optimal matrix 2.4 so that $\mathbf{\Psi} = \mathbf{T}^{-1/2}(\mathbf{T}^{1/2}\mathbf{S}\mathbf{T}^{1/2})^{1/2}\mathbf{T}^{-1/2}$ and the entries of $\mathbf{T}$ and $\mathbf{S}$ are exchanged. Clearly, from Proposition 2.4.6

$$\mathrm{tr}(\mathbf{T} + \mathbf{S} - 2(\mathbf{T}^{1/2}\mathbf{S}\mathbf{T}^{1/2})^{1/2}) = W_2^2(\mathscr{P}_\mathbf{S}, \mathscr{P}_\mathbf{T}) = \mathrm{tr}(\mathbf{S} + \mathbf{T} - 2(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}).$$

In particular tr $(\mathbf{T}^{1/2}\mathbf{S}\mathbf{T}^{1/2})^{1/2} =$ tr $(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}$ so that

$$d_W(\mathbf{S},\mathbf{T}) = \mathrm{tr}(\mathbf{S}+\mathbf{T}-(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2}-(\mathbf{T}^{1/2}\mathbf{S}\mathbf{T}^{1/2})^{1/2}).$$

Finally using tr $(\mathbf{T}^{1/2}\mathbf{S}\mathbf{T}^{1/2})^{1/2} =$ tr $(\mathbf{T}\mathbf{A}(\mathbf{T},\mathbf{S}))$ and tr $(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2} =$ tr $(\mathbf{S}\mathbf{A}(\mathbf{S},\mathbf{T}))$ we may rewrite this as

$$d_W(\mathbf{S},\mathbf{T}) = \mathrm{tr}\ (\mathbf{S}(\mathbf{Id}-\mathbf{A}(\mathbf{S},\mathbf{T}))+\mathbf{T}(\mathbf{Id}-\mathbf{A}(\mathbf{T},\mathbf{S}))).$$

The distance $d_W$ extends from $\mathbb{S}^n_{++}$ to $\mathbb{S}^n_{+}$ for continuity reasons. Indeed since the function

$$(\mathbf{S},\mathbf{T}) \mapsto \mathrm{tr}(\mathbf{S}+\mathbf{T}-2(\mathbf{S}^{1/2}\mathbf{T}\mathbf{S}^{1/2})^{1/2})$$

is well-defined and continuous on $\mathbb{S}^n_{+}\times\mathbb{S}^n_{+}$ and $\mathbb{S}^n_{+}$ is the closure of $\mathbb{S}^n_{++}$ the metric properties hold on $\mathbb{S}^n_{+}$. □

## 2.5 Connection to Covariance and Centered Measures.

Recall the *center of mass* or *mean* of $\mu$ is the vector $\mathbf{m}_\mu = \int_{\mathbb{R}^n} \mathbf{v}\ d\mu(\mathbf{v})$. We obtain the *centered measure* $\overline{\mu}(A) := \mu(A+\mathbf{m}_\mu) = (\tau_{\mathbf{m}_\mu})_\#\mu$, where $\tau_{\mathbf{v}}(\mathbf{x}) := \mathbf{x}-\mathbf{v}$. For any unit vector $\mathbf{x}$ put $m_{\mathbf{x}} := \langle \mathbf{m}_\mu, \mathbf{x}\rangle$. Note, that if $\mu \in \mathscr{P}_2(\mathbb{R}^n)$, then $\overline{\mu}$ is a frame if and only if supp $\mu$ is not contained in any hyperplane. For if supp $\mu$ would be contained in a hyperplane, then supp $\overline{\mu}$ is contained in a linear subspace. We add some identities as below.

For linear $\mathbf{T}$ we have $\mathbf{m}_{\mathbf{T}_\#\mu} = \mathbf{T}(\mathbf{m}_\mu)$ since

$$\mathbf{m}_{\mathbf{T}_\#\mu} = \int \mathbf{v}\ d\mathbf{T}_\#\mu = \int \mathbf{T}\mathbf{v}\ d\mu = \mathbf{T}\int \mathbf{v}\ d\mu = \mathbf{T}(\mathbf{m}_\mu).$$

Again for linear $\mathbf{T}$ and any measure $\mu$ we have

$$\mathbf{T}_\#(\tau_{\mathbf{v}})_\#\mu = (\tau_{\mathbf{T}^{-1}\mathbf{v}})_\#\mathbf{T}_\#\mu.$$

Indeed, if $A \subset \mathbb{R}$ is Borel, we have

$$\mathbf{T}_{\#}(\tau_{\mathbf{v}})_{\#}\mu(A) = \mu(\mathbf{T}^{-1}(A+\mathbf{v})) = (\tau_{\mathbf{T}^{-1}\mathbf{v}})_{\#}\mu(\mathbf{T}^{-1}A) = (\tau_{\mathbf{T}^{-1}\mathbf{v}})_{\#}\mathbf{T}_{\#}\mu(A).$$

The frame operator of $\overline{\mu}$ is the (co)variance operator of $\mu$, as above denoted by $\mathbf{\Sigma}_\mu$. Because of

$$\begin{aligned}\langle \mathbf{x}, \mathbf{S}_{\overline{\mu}}\mathbf{x}\rangle &= \int \langle \mathbf{x}, \mathbf{y}\rangle^2 \; d\mu(\mathbf{y}+\mathbf{m}_\mu) = \int \langle \mathbf{x}, \mathbf{y}-\mathbf{m}_\mu\rangle^2 \; d\mu(\mathbf{y}) = \\ &\int \langle \mathbf{x}, \mathbf{y}\rangle^2 \; d\mu(\mathbf{y}) - 2\langle \mathbf{x}, \mathbf{m}_\mu\rangle \int \langle \mathbf{x}, \mathbf{y}\rangle \; d\mu(\mathbf{y}) + \langle \mathbf{x}, \mathbf{m}_\mu\rangle^2 \\ &= \int \langle \mathbf{x}, \mathbf{y}\rangle^2 \; d\mu(\mathbf{y}) - \langle \mathbf{x}, \mathbf{m}_\mu\rangle^2 = \langle \mathbf{x}, (\mathbf{S}_\mu - \mathbf{m}_\mu\mathbf{m}_\mu^t)\mathbf{x}\rangle,\end{aligned} \tag{2.5.1}$$

we have $\mathbf{S}_{\overline{\mu}} = \mathbf{S}_\mu - \mathbf{m}_\mu\mathbf{m}_\mu^t$. Note that the vectors $\mathbf{m}_\mu$ and matrices $\mathbf{m}_\mu\mathbf{m}_\mu^t$, $\mathbf{\Sigma}_\mu$ depend continuously on $\mu$. That follows from the next proposition for $\mathbf{m}_\mu$. Then continuity of $\mathbf{m}_\mu\mathbf{m}_\mu^t$ is an easy consequence so that $\mathbf{\Sigma}_\mu = \mathbf{S}_\mu - \mathbf{m}_\mu\mathbf{m}_\mu^t$ is also continuous.

**Proposition 2.5.1.**

$$W_2^2(\overline{\mu}, \overline{\nu}) + \|\mathbf{m}_\mu - \mathbf{m}_\nu\|^2 = W_2^2(\mu, \nu).$$

*Proof.* Suppose $\overline{\gamma} \in \Gamma(\overline{\mu}, \overline{\nu})$ is a minimizing coupling, i.e. so that $W_2^2(\overline{\mu}, \overline{\nu}) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x}-\mathbf{y}\|^2 \; d\overline{\gamma}$, then with $\gamma := \tau_{\mathbf{m}_\mu} \times \tau_{\mathbf{m}_\nu})_{\#}\overline{\gamma} \in \Gamma(\mu, \nu)$ we have

$$\begin{aligned}W_2^2(\overline{\mu}, \overline{\nu}) &= \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x}-\mathbf{y}\|^2 \; d\overline{\gamma} = \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x}-\mathbf{m}_\mu - (\mathbf{y}-\mathbf{m}_\nu)\|^2 \; d\gamma = \\ &\int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x}-\mathbf{y}\|^2 \; d\gamma \; - \|\mathbf{m}_\mu - \mathbf{m}_\nu\|^2 \geq W_2^2(\mu,\nu) - \|\mathbf{m}_\mu - \mathbf{m}_\nu\|^2.\end{aligned}$$

The last identity follows from writing

$$\|\mathbf{x}-\mathbf{y}-(\mathbf{m}_\mu - \mathbf{m}_\nu)\|^2 = \langle x-y-(\mathbf{m}_\mu - \mathbf{m}_\nu), x-y-(\mathbf{m}_\mu - \mathbf{m}_\nu)\rangle,$$

then using linearity followed by definition of means $\mathbf{m}_\mu$ and $\mathbf{m}_\nu$. To finish, run the same argument backward, starting with a minimizing coupling $\gamma \in \Gamma(\mu, \nu)$. □

Using this, the following result of [Cuesta-Albertos et al., 1996] follows directly from inequality 2.4.3 applied to the centered measures

**Corollary 2.5.2.** *If* $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ *is an orthonormal basis then*

$$W_2^2(\mu,\nu) \geq \|\mathbf{m}_\mu - \mathbf{m}_\nu\|^2 + \sum_{i=1}^n \left(W_2(\overline{\mu}, (\pi_{\mathbf{e}_i^\perp})_\# \overline{\mu}) - W_2(\overline{\nu}, (\pi_{\mathbf{e}_i^\perp})_\# \overline{\nu})\right)^2. \tag{2.5.2}$$

*Moreover, equality holds in case* $(\mathbf{\Pi}_{\overline{\mu}})_\# \overline{\nu} = \mathbf{T}_\# \overline{\mu}$ *where* $\mathbf{T}$ *is a symmetric, positive definite linear map with respect to an eigenbasis* $\{\mathbf{e}_i\}$ *of* $\mathbf{T}$.

The projection matrix $\mathbf{\Pi}_{\overline{\mu}}$ appears since centered probabilistic frames are not necessary probabilistic frames. It is interesting to compare the two estimates, centered and not centered, and the respective conditions for equality. In fact, if equality holds for the direct estimate, then the target frame is a push-forward $\nu = (\mathbf{A}_{\mu,\nu})_\# \mu$ that depends only on two frame operators, i.e. $\mathbf{S}_\mu$ and the target frame operator $\mathbf{S}_\nu$. In this case by the second identity of Lemma 2.4.2 we have $(\mathbf{\Pi}_{\overline{\mu}})_\# \overline{\nu} = (\mathbf{A}_{\overline{\mu},\overline{\nu}})_\# \overline{\mu}$. Hence we automatically get the identity for the Wasserstein distance between the respective centered measures. The opposite direction does not hold, for the reason that the push-forward maps are positive definite:

**Example**. Let $x \in \mathbb{R}\backslash\{0\}$, then consider $\mu = \delta_x$ and $\nu = \delta_{-x}$. Then $\overline{\mu} = \delta_0 = \overline{\nu}$ but there is no $\lambda > 0$, so that $\lambda_\# \mu = \nu$. For the same reason, absolutely continuous measures such as the equal distribution $\mu$ on $[a, 3a]$ with $a > 0$ so that $m_\mu = 2a$. Now put $\nu = (\tau_{2m_\mu})_\# \mu$, in other words $\nu(A) = \mu(A + 4a)$ for a Borel set $A \subset \mathbb{R}$, cannot be moved to one other by positive scaling.

Given two frames $\mu$ and $\nu$ the condition $(\mathbf{\Pi}_{\overline{\mu}})_\# \overline{\nu} = (\mathbf{A}_{\overline{\mu},\overline{\nu}})_\# \overline{\mu}$ does not guarantee a symmetric positive definite linear map that pushes $\mu$ to $\nu$. Indeed, if the projection $\mathbf{\Pi}_{\overline{\mu}}$ is not trivial and does not act as the identity on $\overline{\nu}$, i.e. if supp $\nu \nsubseteq$ Im $\mathbf{\Sigma}_\mu$ then such a push-forward cannot be found since it would have to contain a projection part which then makes the liner map singular.

# Chapter 3

# Probabilistic Frame Perturbations

The Paley–Wiener Theorem is a classical result about the stability of a basis in a Banach space, claiming that if a sequence is close to a basis, then this sequence is a basis. Similar results are extended to frames in Hilbert spaces. As the extension of finite frames in $\mathbb{R}^n$, probabilistic frames are probability measures on $\mathbb{R}^n$ with finite second moments and the support of which span $\mathbb{R}^n$. This section generalizes the Paley–Wiener theorem to the probabilistic frame setting. We claim that if a probability measure is close to a probabilistic frame in some sense, this probability measure is also a frame. See [Chen, 2023] for more details.

## 3.1 The Paley–Wiener Theorem for Frames

First proposed by Paley and Wiener in [Paley and Wiener, 1934], the Paley–Wiener theorem is a classical result about the stability of a basis in a Hilbert space, which claims that if a sequence is close to an orthonormal basis in a Hilbert space, this sequence also forms a basis. Later, Boas noticed that Paley and Wiener's proof still holds in a Banach space [Boas Jr, 1940]:

**Theorem 3.1.1** (Theorem 1 in [Boas Jr, 1940], Theorem 10 in [Young, 2001])**.** *Let $\{\mathbf{x}_i\}_{i=1}^{\infty}$ be a basis for a Banach space $\mathcal{X}$ with norm $\|\cdot\|$. Suppose $\{\mathbf{y}_i\}_{i=1}^{\infty}$ is a sequence in $\mathcal{X}$ such that*

$$\Big\|\sum_{i=1}^{n} c_i(\mathbf{x}_i - \mathbf{y}_i)\Big\| \leq \lambda \Big\|\sum_{i=1}^{n} c_i \mathbf{x}_i\Big\|$$

*for some constant $0 \leq \lambda < 1$ and all choices of scalars $c_1, \dots, c_n$ $(n = 1, 2, 3, \dots)$. Then $\{\mathbf{y}_i\}_{i=1}^{\infty}$ is* *a basis for the Banach space $\mathscr{X}$ and equivalent to* $\{\mathbf{x}_i\}_{i=1}^{\infty}$ [1].

Since then, many variations of this stability theorem have been generalized, such as the perturbation theory of basis in a Banach space [Singer, 1970, p. 84-109] and entire functions of exponential type [Young, 2001, p. 85]. Christensen first generalized the Paley–Wiener Theorem to study the stability of frames in Hilbert spaces by the following theorem [Christensen, 1995b]:

**Theorem 3.1.2** (Theorem 1 in [Christensen, 1995b]). *Let* $\{\mathbf{f}_i\}_{i=1}^{\infty}$ *be a frame for a Hilbert space* $\mathscr{H}$ *with bounds* $0 < A \leq B < \infty$ *and* $\{\mathbf{g}_i\}_{i=1}^{\infty}$ *a sequence in* $\mathscr{H}$. *If there exist constants* $\lambda, \delta \geq 0$ *such that* $\lambda + \frac{\delta}{\sqrt{A}} < 1$ *and*

$$\Big\|\sum_{i=1}^{n} c_i(\mathbf{f}_i - \mathbf{g}_i)\Big\| \leq \lambda\Big\|\sum_{i=1}^{n} c_i\mathbf{f}_i\Big\| + \delta\Big[\sum_{i=1}^{n}|c_i|^2\Big]^{1/2}$$

*for all scalars* $c_1, \ldots, c_n (n = 1, 2, 3, \ldots)$. *Then* $\{\mathbf{g}_i\}_{i=1}^{\infty}$ *is a frame for* $\mathscr{H}$ *with bounds*

$$A\Big(1 - (\lambda + \frac{\delta}{\sqrt{A}})\Big)^2 \text{ and } B\Big(1 + \lambda + \frac{\delta}{\sqrt{B}}\Big)^2.$$

Later, Casazza and Christensen improved Theorem 3.1.2 by adding one more term related to the sequence $\{\mathbf{g}_i\}_{i=1}^{\infty}$ on the right-hand side of the inequality [Cazassa and Christensen, 1997]:

**Theorem 3.1.3** (Theorem 2 in [Cazassa and Christensen, 1997]). *Let* $\{\mathbf{f}_i\}_{i=1}^{\infty}$ *be a frame for a Hilbert space* $\mathscr{H}$ *with bounds* $0 < A \leq B < \infty$. *Let* $\{\mathbf{g}_i\}_{i=1}^{\infty}$ *be a sequence in* $\mathscr{H}$ *and assume that there exist constants* $\lambda_1, \lambda_2, \delta \geq 0$ *such that max* $(\lambda_1 + \frac{\delta}{\sqrt{A}}, \lambda_2) < 1$ *and*

$$\Big\|\sum_{i=1}^{n} c_i(\mathbf{f}_i - \mathbf{g}_i)\Big\| \leq \lambda_1\Big\|\sum_{i=1}^{n} c_i\mathbf{f}_i\Big\| + \lambda_2\Big\|\sum_{i=1}^{n} c_i\mathbf{g}_i\Big\| + \delta\Big[\sum_{i=1}^{n}|c_i|^2\Big]^{1/2}$$

*for all scalars* $c_1, \ldots, c_n (n = 1, 2, 3, \ldots)$. *Then* $\{\mathbf{g}_i\}_{i=1}^{\infty}$ *is a frame for* $\mathscr{H}$ *with bounds*

$$A\Big(1 - \frac{\lambda_1 + \lambda_2 + \frac{\delta}{\sqrt{A}}}{1 + \lambda_2}\Big)^2 \text{ and } B\Big(1 + \frac{\lambda_1 + \lambda_2 + \frac{\delta}{\sqrt{B}}}{1 - \lambda_2}\Big)^2.$$

From then on, Paley-Wiener type theorems have been studied for many other topics related to frames, such as Banach frames [Christensen and Heil, 1997], frames containing Riesz basis [Casazza and Christensen, 1998], frame sequence [Christensen et al., 2000], sequences with reconstruction properties in a Banach space [Casazza and Christensen, 2008], von Neumann–Schatten

[1] Equivalence of basis $\{\mathbf{x}_i\}_{i=1}^{\infty}$ and $\{\mathbf{y}_i\}_{i=1}^{\infty}$ for the Banach space $\mathscr{X}$ means that there exists a bounded and invertible operator $T$ on $\mathscr{X}$ such that $T\mathbf{x}_i = \mathbf{y}_i$, for any $i$.

dual frames [Arefijamaal and Sadeghi, 2016], operator represented frames [Christensen and Hasannasab, 2017], g–frames [Sun, 2007], continuous frames on quaternionic Hilbert spaces [Khokulan and Thirulogasanthar, 2019], approximately dual frames [Javanshiri et al., 2022], frames for metric spaces [Krishna and Johnson, 2022], Hilbert–Schmidt frames and sequences [Poria, 2017, Zhang and Li, 2022]. Especially [Chen et al., 2014] introduced dual frames in the perturbation condition, which differs from previous conditions to preserve Hilbert frames:

**Theorem 3.1.4** (Theorem 2.1 in [Chen et al., 2014])**.** *Let* $\{\mathbf{f}_i\}_{i=1}^{\infty}$ *be a frame for the Hilbert space* $\mathcal{H}$ *with bounds* $0 < A \leq B < \infty$ *and* $\{\mathbf{h}_i\}_{i=1}^{\infty}$ *a dual frame of* $\{\mathbf{f}_i\}_{i=1}^{\infty}$ *with upper frame bound* $0 < D < \infty$*. Suppose* $\{\mathbf{g}_i\}_{i=1}^{\infty}$ *is a sequence in* $\mathcal{H}$ *such that*

$$\alpha := \sum_{i=1}^{\infty} \|\mathbf{f}_i - \mathbf{g}_i\|^2 < \infty, \ \beta := \sum_{i=1}^{\infty} \|\mathbf{f}_i - \mathbf{g}_i\| \|\mathbf{h}_i\| < 1.$$

*Then* $\{\mathbf{g}_i\}_{i=1}^{\infty}$ *is a frame in* $\mathcal{H}$ *with bounds* $\frac{(1-\beta)^2}{D}$ *and* $B(1 + \sqrt{\frac{\alpha}{B}})$.

This section extends the frame perturbation theory to probabilistic frames for $\mathbb{R}^n$. The primary tool we use is the invertible linear operator theory on Banach spaces, which is introduced in Section 3.2. In Section 3.3, we generalize the Paley-Wiener theorem of frames to probabilistic frames by replacing the sum and $l_2$ sequences with integration and continuous functions with compact support. In Section 3.4, we give a sufficient perturbation condition by including the probabilistic dual frames, which generalizes Theorem 3.1.4. Especially we consider the particular case of the canonical probabilistic dual frame. In the end, we give an example of the stability of finite frames in $\mathbb{R}^n$ by applying the results we obtain.

## 3.2 Invertibility of Linear Operators on Banach Spaces

This section introduces the invertibility of linear operators on Banach spaces. We refer to [Cazassa and Christensen, 1997] for more details. It is well-known that a bound linear operator $U$ on a Banach space $\mathcal{X}$ is invertible if $\|I - U\| < 1$ and

$$\|U^{-1}\| \leq \frac{1}{1 - \|I - U\|},$$

where $I$ is identity operator in $\mathcal{X}$. Casazza and Christensen generalized the above result to the following lemma.

**Lemma 3.2.1** (Lemma 1 in [Cazassa and Christensen, 1997])**.** *Let $\mathcal{X}$, $\mathcal{Y}$ be Banach spaces and $U : \mathcal{X} \to \mathcal{X}$ a linear operator on $\mathcal{X}$. If there exist $\lambda_1, \lambda_2 \in [0, 1)$ such that for any $\mathbf{x} \in \mathcal{X}$,*

$$\|U\mathbf{x} - \mathbf{x}\| \leq \lambda_1 \|\mathbf{x}\| + \lambda_2 \|U\mathbf{x}\|.$$

*Then $U$ is bounded invertible, and for any $\mathbf{x} \in \mathcal{X}$,*

$$\frac{1 - \lambda_1}{1 + \lambda_2}\|\mathbf{x}\| \leq \|U\mathbf{x}\| \leq \frac{1 + \lambda_1}{1 - \lambda_2}\|\mathbf{x}\|, \ \frac{1 - \lambda_2}{1 + \lambda_1}\|\mathbf{x}\| \leq \|U^{-1}\mathbf{x}\| \leq \frac{1 + \lambda_2}{1 - \lambda_1}\|\mathbf{x}\|.$$

**Corollary 3.2.2** (Remark of Corollary 1 in [Cazassa and Christensen, 1997])**.** *Suppose $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces. Let $U : \mathcal{X} \to \mathcal{Y}$ be a bounded linear operator, $\mathcal{X}_0$ a dense subspace of $\mathcal{X}$, and $V : \mathcal{X} \to \mathcal{Y}$ a linear mapping. If for any $\mathbf{x} \in \mathcal{X}_0$,*

$$\|U\mathbf{x} - V\mathbf{x}\| \leq \lambda_1 \|U\mathbf{x}\| + \lambda_2 \|V\mathbf{x}\| + \delta \|\mathbf{x}\|,$$

*where $\lambda_1, \lambda_2, \delta \in [0, 1)$. Then $V$ has a unique extension to a bounded linear operator (of the same norm) from $\mathcal{X}$ to $\mathcal{Y}$, and the extension still satisfies the inequality.*

## 3.3 The Paley-Wiener Theorem for Probabilistic Frames

In this section, we generalize the Paley-Wiener theorem to probabilistic frames by employing Casazza and Christensen's criteria for the invertibility of linear operators. Recall that if $\nu \in \mathscr{P}_2(\mathbb{R}^n)$, $\nu$ is a Bessel probability measure with bound $M_2(\nu) := \int_{\mathbb{R}^n} \|y\|^2 d\nu(y)$. Let $C_c(\mathbb{R}^n)$ be the set of continuous functions on $\mathbb{R}^n$ with compact support. Then we have the main perturbation theorem about probabilistic frames in the following that generalizes Theorem 3.1.3.

**Theorem 3.3.1.** *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$ with bounds $0 < A \leq B < \infty$ and $\nu \in \mathscr{P}_2(\mathbb{R}^n)$.*

*If there exist constants $\lambda_1, \lambda_2, \delta \geq 0$ such that max $(\lambda_1 + \frac{\delta}{\sqrt{A}}, \lambda_2) < 1$ and for all $w \in C_c(\mathbb{R}^n)$,*

$$\Big\| \int_{\mathbb{R}^n} w(\mathbf{x})\mathbf{x} d\mu(\mathbf{x}) - \int_{\mathbb{R}^n} w(\mathbf{y})\mathbf{y} d\nu(\mathbf{y}) \Big\| \\ \leq \lambda_1 \Big\| \int_{\mathbb{R}^n} w(\mathbf{x})\mathbf{x} d\mu(\mathbf{x}) \Big\| + \lambda_2 \Big\| \int_{\mathbb{R}^n} w(\mathbf{y})\mathbf{y} d\nu(\mathbf{y}) \Big\| + \delta \|w\|_{L^2(\mu)}.$$

*Then $\nu$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $\frac{A^2(1-(\lambda_1+\frac{\delta}{\sqrt{A}}))^2}{(1+\lambda_2)^2 M_2(\nu)}$ and $M_2(\nu)$.*

*Proof.* Since $\nu \in \mathscr{P}_2(\mathbb{R}^n)$, then $\nu$ is Bessel with bound $M_2(\nu)$. Now let us get the lower frame bound. Define $U : L^2(\mu) \to \mathbb{R}^n$ and $T : L^2(\mu) \to \mathbb{R}^n$ in the following way

$$U(w) := \int_{\mathbb{R}^n} w(\mathbf{x})\mathbf{x} d\mu(\mathbf{x}),\ T(w) := \int_{\mathbb{R}^n} w(\mathbf{y})\mathbf{y} d\nu(\mathbf{y}),$$

where $w \in C_c(\mathbb{R}^n)$. Clearly $U$ is bounded linear and $\|U\|^2 \leq M_2(\mu)$. Furthermore, $T$ is well-defined. Since $C_c(\mathbb{R}^n)$ is dense in $L^2(\mu)$, then by the theorem condition and Corollary 3.2.2, we know that $T$ could be extended uniquely to a bounded linear operator that is still denoted by $T$, and for any $w \in L^2(\mu)$,

$$\Big\|U(w) - T(w)\Big\| \leq \lambda_1 \Big\|U(w)\Big\| + \lambda_2 \Big\|T(w)\Big\| + \delta \|w\|_{L^2(\mu)}. \tag{3.3.1}$$

Therefore, for any $w \in L^2(\mu)$,

$$\|T(w)\| \leq \|U(w)\| + \|U(w) - T(w)\| \leq \big((\lambda_1 + 1)\big\|U\| + \delta\big)\|w\|_{L^2(\mu)} + \lambda_2 \|T(w)\|.$$

Thus $T$ is well-defined(bounded) and $\|T\| \leq \frac{(\lambda_1+1)\|U\|+\delta}{1-\lambda_2} < +\infty$. Define $U^+ : \mathbb{R}^n \to L^2(\mu)$ by

$$(U^+\mathbf{x})(\cdot) := (U^*(UU^*)^{-1}\mathbf{x})(\cdot) = (U^*(\mathbf{S}_\mu^{-1}\mathbf{x}))(\cdot) = \big\langle \mathbf{S}_\mu^{-1}\mathbf{x}, \cdot \big\rangle \in L^2(\mu).$$

where $U^*$ is the adjoint operator of $U$, $UU^* = \mathbf{S}_\mu$. Then

$$\|U^+\mathbf{x}\|^2_{L^2(\mu)} = \int_{\mathbb{R}^n} \big\langle \mathbf{S}_\mu^{-1}\mathbf{x}, \mathbf{y} \big\rangle^2 d\mu(\mathbf{y}) = \int_{\mathbb{R}^n} \big\langle \mathbf{x}, \mathbf{S}_\mu^{-1}\mathbf{y} \big\rangle^2 d\mu(\mathbf{y}) \\ = \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{y} \rangle^2 d(\mathbf{S}_\mu^{-1}{}_{\#}\mu)(\mathbf{y}) \leq \frac{1}{A}\|\mathbf{x}\|^2.$$

where the last inequality follows from $\mathbf{S}_\mu^{-1}{}_{\#}\mu$ being the probabilistic (dual) frame with bounds $\frac{1}{B}$

and $\frac{1}{A}$. For every $\mathbf{x} \in \mathbb{R}^n$, replacing $w$ in Equation (3.3.1) with $U^+\mathbf{x}$ leads to

$$\begin{aligned}\Big\|\mathbf{x} - T(U^+\mathbf{x})\Big\| &\leq \lambda_1\|\mathbf{x}\| + \lambda_2\|T(U^+\mathbf{x})\Big\| + \delta\|U^+\mathbf{x}\|_{L^2(\mu)} \\ &\leq (\lambda_1 + \frac{\delta}{\sqrt{A}})\|\mathbf{x}\| + \lambda_2\|T(U^+\mathbf{x})\Big\|.\end{aligned}$$

Since $\max(\lambda_1 + \frac{\delta}{\sqrt{A}}, \lambda_2) < 1$, by Lemma 3.2.1, we know that $TU^+$ is invertible, and

$$\|(TU^+)^{-1}\| \leq \frac{1+\lambda_2}{1-(\lambda_1 + \frac{\delta}{\sqrt{A}})}.$$

Note that any $\mathbf{x} \in \mathbb{R}^n$ could be written as

$$\mathbf{x} = TU^+(TU^+)^{-1}\mathbf{x} = \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}(TU^+)^{-1}\mathbf{x}, \mathbf{y}\right\rangle \mathbf{y} d\nu(\mathbf{y}).$$

Therefore,

$$\begin{aligned}\|\mathbf{x}\|^4 = \langle \mathbf{x}, \mathbf{x}\rangle^2 &= \Big| \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}(TU^+)^{-1}\mathbf{x}, \mathbf{y}\right\rangle \langle \mathbf{x}, \mathbf{y}\rangle \, d\nu(\mathbf{y})\Big|^2 \\ &\leq \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}(TU^+)^{-1}\mathbf{x}, \mathbf{y}\right\rangle^2 d\nu(\mathbf{y}) \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{y}\rangle^2 \, d\nu(\mathbf{y}) \\ &\leq \|\mathbf{S}_\mu^{-1}\|_2^2 \; \|(TU^+)^{-1}\|^2 \; \|\mathbf{x}\|^2 \int_{\mathbb{R}^n} \|\mathbf{y}\|^2 d\nu(\mathbf{y}) \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{y}\rangle^2 \, d\nu(\mathbf{y}) \\ &\leq \frac{M_2(\nu)}{A^2}\Big(\frac{1+\lambda_2}{1-(\lambda_1 + \frac{\delta}{\sqrt{A}})}\Big)^2 \|\mathbf{x}\|^2 \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{y}\rangle^2 \, d\nu(\mathbf{y}),\end{aligned}$$

where the first inequality comes from the Cauchy–Schwarz inequality in $L^2(\nu)$ and $\|\mathbf{S}_\mu^{-1}\|_2$ is the 2-matrix norm of $\mathbf{S}_\mu^{-1}$ with $\|\mathbf{S}_\mu^{-1}\|_2 \leq \frac{1}{A}$. Thus for any $\mathbf{x} \in \mathbb{R}^n$,

$$\frac{A^2(1-(\lambda_1 + \frac{\delta}{\sqrt{A}}))^2}{(1+\lambda_2)^2 M_2(\nu)}\|\mathbf{x}\|^2 \leq \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{y}\rangle^2 \, d\nu(\mathbf{y}) \leq M_2(\nu) \; \|\mathbf{x}\|^2.$$

That is to say, $\nu$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $\frac{A^2(1-(\lambda_1+\frac{\delta}{\sqrt{A}}))^2}{(1+\lambda_2)^2 M_2(\nu)}$ and $M_2(\nu)$. □

As a particular case in Theorem 3.3.1, $\lambda_1 = \lambda$ and $\lambda_2 = 0$ give rise to the following proposition, which is a generalization of Theorem 3.1.2.

**Proposition 3.3.2.** *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$ with bounds $0 < A \leq B < \infty$ and*

$\nu \in \mathscr{P}_2(\mathbb{R}^n)$. *Suppose there exist constants* $\lambda, \delta \geq 0$ *such that* $\lambda + \frac{\delta}{\sqrt{A}} < 1$ *and for all* $w \in C_c(\mathbb{R}^n)$

$$\Big\| \int_{\mathbb{R}^n} w(\mathbf{x})\mathbf{x} d\mu(\mathbf{x}) - \int_{\mathbb{R}^n} w(\mathbf{y})\mathbf{y} d\nu(\mathbf{y}) \Big\| \leq \lambda \Big\| \int_{\mathbb{R}^n} w(\mathbf{x})\mathbf{x} d\mu(\mathbf{x}) \Big\| + \delta \|w\|_{L^2(\mu)}.$$

*Then* $\nu$ *is a probabilistic frame for* $\mathbb{R}^n$ *with bounds* $\frac{A^2(1-(\lambda+\frac{\delta}{\sqrt{A}}))^2}{M_2(\nu)}$ *and* $M_2(\nu)$.

**Remark 3.3.1.** *Since* $\frac{\delta}{\sqrt{A}} \leq \frac{\sqrt{B}\delta}{A}$, *then the condition max* $(\lambda_1 + \frac{\delta}{\sqrt{A}}, \lambda_2) < 1$ *in Theorem 3.3.1 and* $\lambda + \frac{\delta}{\sqrt{A}} < 1$ *in Proposition 3.3.2 could be replaced by max* $(\lambda_1 + \frac{\sqrt{B}\delta}{A}, \lambda_2) < 1$ *and* $\lambda + \frac{\sqrt{B}\delta}{A} < 1$, *and the lower frame bounds for* $\nu$ *are* $\frac{A^2(1-(\lambda_1+\frac{\sqrt{B}\delta}{A}))^2}{(1+\lambda_2)^2 M_2(\nu)}$ *and* $\frac{A^2(1-(\lambda+\frac{\sqrt{B}\delta}{A}))^2}{M_2(\nu)}$ *respectively. This is due to another way to get* $\|U^+\mathbf{x}\|_{L^2(\mu)}$ *by the fact that* $\mu$ *is a probabilistic frame with bounds* $A$ *and* $B$*:*

$$\|U^+\mathbf{x}\|_{L^2(\mu)}^2 = \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}\mathbf{x}, \mathbf{y} \right\rangle^2 d\mu(\mathbf{y}) \leq B\|\mathbf{S}_\mu^{-1}\mathbf{x}\|^2 \leq B\|\mathbf{S}_\mu^{-1}\|_2^2 \|\mathbf{x}\|^2 \leq \frac{B}{A^2}\|\mathbf{x}\|^2.$$

The following lemma is inspired by a particular case $\lambda = 0, \delta = \sqrt{R}$ in Proposition 3.3.2 but just formulated in the "adjoint" form of the "synthesis" operator of sighed measure "$\mu - \nu$." However, an easier way to prove it is to apply the definition of probabilistic frames.

**Lemma 3.3.3** (Sweetie's Lemma)**.** *Let* $\mu$ *be a probabilistic frame for* $\mathbb{R}^n$ *with bounds* $0 < A \leq B < \infty$ *and* $\nu \in \mathscr{P}(\mathbb{R}^n)$. *Suppose there exists a constant* $R$ *where* $0 < R < A$, *such that for any* $\mathbf{x} \in \mathbb{R}^n$,

$$\Big| \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{y} \rangle^2 d\mu(\mathbf{y}) - \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{z} \rangle^2 d\nu(\mathbf{z}) \Big| \leq R\|\mathbf{x}\|^2.$$

*Then* $\nu$ *is a probabilistic frame for* $\mathbb{R}^n$ *with bounds* $A - R$ *and* $B + R$.

*Proof.* From the lemma condition, we know that for any $\mathbf{x} \in \mathbb{R}^n$,

$$\int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{y} \rangle^2 d\mu(\mathbf{y}) - R\|\mathbf{x}\|^2 \leq \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{z} \rangle^2 d\nu(\mathbf{z}) \leq \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{y} \rangle^2 d\mu(\mathbf{y}) + R\|\mathbf{x}\|^2.$$

Since $\mu$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $A$ and $B$, then for any $\mathbf{x} \in \mathbb{R}^n$,

$$(A - R)\|\mathbf{x}\|^2 \leq \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{z} \rangle^2 d\nu(\mathbf{z}) \leq (B + R)\|\mathbf{x}\|^2.$$

That is to say, $\nu$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $A - R$ and $B + R$. ☐

**Remark 3.3.2.** *Lemma 3.3.3 could be improved to any coupling $\gamma \in \Gamma(\mu, \nu)$. If for any $\mathbf{x} \in \mathbb{R}^n$,*

$$\Big| \int_{\mathbb{R}^n \times \mathbb{R}^n} \langle \mathbf{x}, \mathbf{y} \rangle^2 - \langle \mathbf{x}, \mathbf{z} \rangle^2 \, d\gamma(\mathbf{y}, \mathbf{z}) \Big| \leq R \|\mathbf{x}\|^2,$$

*then $\nu$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $A - R$ and $B + R$. This is due to $\gamma \in \Gamma(\mu, \nu)$ and*

$$\Big| \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{y} \rangle^2 \, d\mu(\mathbf{y}) - \int_{\mathbb{R}^n} \langle \mathbf{x}, \mathbf{z} \rangle^2 \, d\nu(\mathbf{z}) \Big| = \Big| \int_{\mathbb{R}^n \times \mathbb{R}^n} \langle \mathbf{x}, \mathbf{y} \rangle^2 - \langle \mathbf{x}, \mathbf{z} \rangle^2 \, d\gamma(\mathbf{y}, \mathbf{z}) \Big|.$$

**Remark 3.3.3.** *The test function in Lemma 3.3.3 could be improved to continuous functions $C(\mathbb{R}^n)$. If*

$$\sup_{w \in C(\mathbb{R}^n)} \Big| \int_{\mathbb{R}^n} w(\mathbf{y}) d\mu(\mathbf{y}) - \int_{\mathbb{R}^n} w(\mathbf{y}) d\nu(\mathbf{y}) \Big| \leq R,$$

*then $v$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $A - R$ and $B + R$. The proof is clear by taking the test functions to be $w_{\mathbf{x}}(\mathbf{y}) = \left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \mathbf{y} \right\rangle^2$ where $\mathbf{x}$ is nonzero.*

## 3.4 Perturbations Including Probabilistic Dual Frames

In this section, we generalize Theorem 3.1.4 to the probabilistic frames setting: we give a sufficient perturbation condition where the probabilistic dual frames are used in the following theorem without the quadratic close condition $\alpha < +\infty$.

**Theorem 3.4.1.** *Suppose $\mu$ is a probabilistic frame for $\mathbb{R}^n$ and $\nu$ is a probabilistic dual frame of $\mu$ with respect to $\gamma_{12} \in \Gamma(\mu, \nu)$. Let $\eta \in \mathscr{P}_2(\mathbb{R}^n)$ and $\gamma_{23} \in \Gamma(\nu, \eta)$. Then there exists $\tilde{\pi} \in \mathscr{P}(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n)$ with marginals $\gamma_{12}$ and $\gamma_{23}$, and if*

$$\sigma := \int_{\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x} - \mathbf{z}\| \|\mathbf{y}\| d\tilde{\pi}(\mathbf{x}, \mathbf{y}, \mathbf{z}) < 1,$$

*then $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $\frac{(1-\sigma)^2}{M_2(\nu)}$ and $M_2(\eta)$. And if the upper frame bound for $\nu$ is $D > 0$, then the frame bounds for $\eta$ are $\frac{(1-\sigma)^2}{D}$ and $M_2(\eta)$.*

*Proof.* By Lemma 1.2.3(Gluing Lemma), there exists $\tilde{\pi} \in \mathscr{P}(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n)$ such that ${\pi_{xy}}_{\#}\tilde{\pi} = \gamma_{12}$, ${\pi_{yz}}_{\#}\tilde{\pi} = \gamma_{23}$ where $\pi_{xy}$ and $\pi_{yz}$ are projections to $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{y}, \mathbf{z})$ coordinates. Since $\eta \in \mathscr{P}_2(\mathbb{R}^n)$, then $\eta$ is Bessel with bound $M_2(\eta)$. Next let us show the lower frame bound. Define a linear operator $L : \mathbb{R}^n \to \mathbb{R}^n$ by

$$L(\mathbf{f}) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \langle \mathbf{f}, \mathbf{y} \rangle \, \mathbf{z} \; d\gamma_{23}(\mathbf{y}, \mathbf{z}), \text{ for any } \mathbf{f} \in \mathbb{R}^n.$$

Since $\nu$ is the probabilistic dual frame of $\mu$ with respect to $\gamma_{12} \in \Gamma(\mu, \nu)$, then

$$\mathbf{f} = \int_{\mathbb{R}^n \times \mathbb{R}^n} \langle \mathbf{f}, \mathbf{y} \rangle \, \mathbf{x} \; d\gamma_{12}(\mathbf{x}, \mathbf{y}), \text{ for any } \mathbf{f} \in \mathbb{R}^n.$$

Therefore,

$$\begin{aligned} \|\mathbf{f} - L(\mathbf{f})\| &= \Big\| \int_{\mathbb{R}^n \times \mathbb{R}^n} \langle \mathbf{f}, \mathbf{y} \rangle \, \mathbf{x} \; d\gamma_{12}(\mathbf{x}, \mathbf{y}) - \int_{\mathbb{R}^n \times \mathbb{R}^n} \langle \mathbf{f}, \mathbf{y} \rangle \, \mathbf{z} \; d\gamma_{23}(\mathbf{y}, \mathbf{z}) \Big\| \\ &= \Big\| \int_{\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n} \langle \mathbf{f}, \mathbf{y} \rangle \, (\mathbf{x} - \mathbf{z}) \; d\tilde{\pi}(\mathbf{x}, \mathbf{y}, \mathbf{z}) \Big\| \leq \sigma \|\mathbf{f}\| < \|\mathbf{f}\|. \end{aligned}$$

Thus $L : \mathbb{R}^n \to \mathbb{R}^n$ is invertible and $\|L^{-1}\| \leq \frac{1}{1-\sigma}$. Note that for any $\mathbf{f} \in \mathbb{R}^n$,

$$\mathbf{f} = LL^{-1}(\mathbf{f}) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\langle L^{-1}\mathbf{f}, \mathbf{y} \right\rangle \mathbf{z} \; d\gamma_{23}(\mathbf{y}, \mathbf{z}).$$

Therefore,

$$\begin{aligned} \|\mathbf{f}\|^4 = \langle \mathbf{f}, \mathbf{f} \rangle^2 &= \Big| \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\langle L^{-1}\mathbf{f}, \mathbf{y} \right\rangle \langle \mathbf{f}, \mathbf{z} \rangle \;\; d\gamma_{23}(\mathbf{y}, \mathbf{z}) \Big|^2 \\ &\leq \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\langle L^{-1}\mathbf{f}, \mathbf{y} \right\rangle^2 d\gamma_{23}(\mathbf{y}, \mathbf{z}) \; \int_{\mathbb{R}^n \times \mathbb{R}^n} \langle \mathbf{f}, \mathbf{z} \rangle^2 \, d\gamma_{23}(\mathbf{y}, \mathbf{z}) \\ &= \int_{\mathbb{R}^n} \left\langle L^{-1}\mathbf{f}, \mathbf{y} \right\rangle^2 d\nu(\mathbf{y}) \; \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{z} \rangle^2 \, d\eta(\mathbf{z}) \leq \frac{M_2(\nu)}{(1-\sigma)^2} \|\mathbf{f}\|^2 \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{z} \rangle^2 \, d\eta(\mathbf{z}), \end{aligned}$$

where the first inequality is due to Cauchy Schwarz inequality in $L^2(\gamma_{23})$ and the third equality comes from $\gamma_{23} \in \Gamma(\nu, \eta)$. Thus for any $\mathbf{f} \in \mathbb{R}^n$,

$$\frac{(1-\sigma)^2}{M_2(\nu)} \|\mathbf{f}\|^2 \leq \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{z} \rangle^2 \, d\eta(\mathbf{z}) \leq M_2(\eta) \; \|\mathbf{f}\|^2.$$

Therefore, $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $\frac{(1-\sigma)^2}{M_2(\nu)}$ and $M_2(\eta)$. If the upper frame bound for the probabilistic dual frame $\nu$ is $0 < D < \infty$, then

$$\|\mathbf{f}\|^4 \leq \int_{\mathbb{R}^n} \left\langle L^{-1}\mathbf{f}, \mathbf{y} \right\rangle^2 d\nu(\mathbf{y}) \; \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{z} \rangle^2 \, d\eta(\mathbf{z}) \leq \frac{D\|\mathbf{f}\|^2}{(1-\sigma)^2} \; \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{z} \rangle^2 \, d\eta(\mathbf{z}).$$

In this case, the frame bounds for $\eta$ are $\frac{(1-\sigma)^2}{D}$ and $M_2(\eta)$. $\square$

If the probabilistic dual frame of $\mu$ is given by the canonical probabilistic dual frame ${\mathbf{S}_\mu^{-1}}_\#\mu$, we have the following corollary.

**Corollary 3.4.2.** *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$ with bounds $0 < A \leq B < \infty$, and $\eta \in \mathscr{P}_2(\mathbb{R}^n)$. If*

$$\hat{\sigma} := \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{S}_\mu^{-1}\mathbf{x}\| \|\mathbf{x} - \mathbf{z}\| d\mu(\mathbf{x}) d\eta(\mathbf{z}) < 1 \tag{3.4.1}$$

*then $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $A(1-\hat{\sigma})^2$ and $M_2(\eta)$.*

*Proof.* Note that ${\mathbf{S}_\mu^{-1}}_\#\mu$ is the canonical probabilistic dual frame of $\mu$ with respect to $\gamma_{12} := (\mathbf{Id} \times \mathbf{S}_\mu^{-1})_\#\mu \in \Gamma(\mu, {\mathbf{S}_\mu^{-1}}_\#\mu)$. Let $\gamma_{23}$ be the product measure $\gamma_{23} := {\mathbf{S}_\mu^{-1}}_\#\mu \otimes \eta \in \Gamma({\mathbf{S}_\mu^{-1}}_\#\mu, \eta)$. Then by the disintegration theorem and Lemma 1.2.3(Gluing Lemma), the transport coupling with marginals $\gamma_{12}$ and $\gamma_{23}$ is given by $\tilde{\pi} := \gamma_{12} \otimes \eta \in \mathscr{P}(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n)$. Thus

$$\hat{\sigma} := \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x} - \mathbf{z}\| \|\mathbf{S}_\mu^{-1}\mathbf{x}\| d\mu(\mathbf{x}) d\eta(\mathbf{z}) = \int_{\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x} - \mathbf{z}\| \|\mathbf{y}\| d\tilde{\pi}(\mathbf{x}, \mathbf{y}, \mathbf{z}).$$

Since $\hat{\sigma} < 1$ and the upper frame bound for the dual frame ${\mathbf{S}_\mu^{-1}}_\#\mu$ is $\frac{1}{A}$, then by Theorem 3.4.1, $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $A(1-\hat{\sigma})^2$ and $M_2(\eta)$. $\square$

**Remark 3.4.1.** *Since $\|\mathbf{S}_\mu^{-1}\mathbf{x}\| \leq \|\mathbf{S}_\mu^{-1}\|_2 \|\mathbf{x}\| \leq \frac{1}{A}\|\mathbf{x}\|$, if*

$$\hat{\epsilon} := \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x}\| \ \|\mathbf{x} - \mathbf{z}\| d\mu(\mathbf{x}) d\eta(\mathbf{z}) < A, \tag{3.4.2}$$

*then $\hat{\sigma} \leq \frac{\hat{\epsilon}}{A} < 1$ and thus $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $A(1-\hat{\sigma})^2$ and $M_2(\eta)$*

Indeed, Corollary 3.4.2 and Remark 3.4.1 could be generalized to any $\gamma \in \Gamma(\mu, \eta)$.

**Proposition 3.4.3.** *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$ with bounds $0 < A \leq B < \infty$ and $\eta \in \mathscr{P}_2(\mathbb{R}^n)$. Let $\gamma \in \Gamma(\mu, \eta)$ be any coupling with marginal $\mu$ and $\eta$. Suppose*

$$\epsilon := \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x}\| \ \|\mathbf{x} - \mathbf{z}\| d\gamma(\mathbf{x}, \mathbf{z}) < A,$$

*then $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $\frac{(A-\epsilon)^2}{B}$ and $M_2(\eta)$. And if*

$$\chi := \int_{\mathbb{R}^n\times\mathbb{R}^n} \|\mathbf{S}_\mu^{-1}\mathbf{x}\| \, \|\mathbf{x}-\mathbf{z}\| d\gamma(\mathbf{x},\mathbf{z}) < 1,$$

*then $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $\frac{A^2(1-\chi)^2}{B}$ and $M_2(\eta)$.*

*Proof.* Since $\eta \in \mathscr{P}_2(\mathbb{R}^n)$, $\eta$ is Bessel with bound $M_2(\eta)$. Since ${\mathbf{S}_\mu^{-1}}_{\#}\mu$ is the canonical probabilistic dual frame of $\mu$ with respect to $(\mathbf{Id}\times\mathbf{S}_\mu^{-1})_{\#}\mu \in \Gamma(\mu, {\mathbf{S}_\mu^{-1}}_{\#}\mu)$, then

$$\mathbf{f} = \int_{\mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{S}_\mu^{-1}\mathbf{x}\right\rangle \mathbf{x} \, d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}\mathbf{f}, \mathbf{x}\right\rangle \mathbf{x} \, d\mu(\mathbf{x}), \text{ for any } \mathbf{f}\in\mathbb{R}^n.$$

Define a linear operator $L:\mathbb{R}^n\to\mathbb{R}^n$ by

$$L(\mathbf{f}) = \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}\mathbf{f}, \mathbf{x}\right\rangle \mathbf{z} \, d\gamma(\mathbf{x},\mathbf{z}), \text{ for any } \mathbf{f}\in\mathbb{R}^n.$$

Therefore,

$$\begin{aligned}\|\mathbf{f}-L(\mathbf{f})\| &= \left\| \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}\mathbf{f}, \mathbf{x}\right\rangle \mathbf{x} \, d\mu(\mathbf{x}) - \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}\mathbf{f}, \mathbf{x}\right\rangle \mathbf{z} \, d\gamma(\mathbf{x},\mathbf{z})\right\| \\ &= \left\| \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}\mathbf{f}, \mathbf{x}\right\rangle (\mathbf{x}-\mathbf{z}) \, d\gamma(\mathbf{x},\mathbf{z})\right\| \le \epsilon \, \|\mathbf{S}_\mu^{-1}\|_2 \, \|\mathbf{f}\| \le \frac{\epsilon}{A}\|\mathbf{f}\|.\end{aligned}$$

Thus $L:\mathbb{R}^n\to\mathbb{R}^n$ is invertible and $\|L^{-1}\| \le \frac{1}{1-\epsilon/A}$. Then for any $\mathbf{f}\in\mathbb{R}^n$,

$$\mathbf{f} = LL^{-1}(\mathbf{f}) = \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}L^{-1}\mathbf{f}, \mathbf{x}\right\rangle \mathbf{z} \, d\gamma(\mathbf{x},\mathbf{z}).$$

Therefore,

$$\begin{aligned}\|\mathbf{f}\|^4 = \left\langle \mathbf{f}, \mathbf{f}\right\rangle^2 &= \left| \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}L^{-1}\mathbf{f}, \mathbf{x}\right\rangle \left\langle \mathbf{f}, \mathbf{z}\right\rangle \, d\gamma(\mathbf{x},\mathbf{z})\right|^2 \\ &\le \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}L^{-1}\mathbf{f}, \mathbf{x}\right\rangle^2 d\gamma(\mathbf{x},\mathbf{z}) \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{z}\right\rangle^2 d\gamma(\mathbf{x},\mathbf{z}) \\ &= \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1}L^{-1}\mathbf{f}, \mathbf{x}\right\rangle^2 d\mu(\mathbf{x}) \int_{\mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{z}\right\rangle^2 d\eta(\mathbf{z}) \\ &\le B\|\mathbf{S}_\mu^{-1}\|_2^2 \, \|L^{-1}\|^2 \, \|\mathbf{f}\|^2 \int_{\mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{z}\right\rangle^2 d\eta(\mathbf{z}) \le \frac{B\|\mathbf{f}\|^2}{A^2(1-\frac{\epsilon}{A})^2} \int_{\mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{z}\right\rangle^2 d\eta(\mathbf{z}),\end{aligned}$$

where the first inequality is due to Cauchy-Schwarz inequality in $L_2(\gamma)$ and the second inequality comes from $\mu$ being a probabilistic frame with upper frame bound $B$. Thus for any $\mathbf{f} \in \mathbb{R}^n$,

$$\frac{A^2(1-\frac{\epsilon}{A})^2}{B}\|\mathbf{f}\|^2 \leq \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{z}\rangle^2 \, d\eta(\mathbf{z}) \leq \|\mathbf{f}\|^2 M_2(\eta).$$

Therefore, $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $\frac{(A-\epsilon)^2}{B}$ and $M_2(\eta)$. Furthermore, if

$$\chi := \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x}-\mathbf{z}\|\|\mathbf{S}_\mu^{-1}\mathbf{x}\| d\gamma(\mathbf{x},\mathbf{z}) < 1,$$

then

$$\|\mathbf{f} - L(\mathbf{f})\| = \left\| \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{S}_\mu^{-1}\mathbf{x} \right\rangle (\mathbf{x}-\mathbf{z}) \; d\gamma(\mathbf{x},\mathbf{z}) \right\| \leq \chi \; \|\mathbf{f}\|.$$

Therefore, $\|I - L\| \leq \chi < 1$ implies $L$ is invertible and $\|L^{-1}\| \leq \frac{1}{1-\chi}$. Similar proof shows that $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $\frac{A^2(1-\chi)^2}{B}$ and $M_2(\eta)$. $\square$

**Remark 3.4.2.** $\frac{(A-\epsilon)^2}{B}$ *is smaller than* $\frac{A^2(1-\chi)^2}{B}$ *since* $\|\mathbf{S}_\mu^{-1}\mathbf{x}\| \leq \frac{1}{A}\|\mathbf{x}\|$.

The key step in the proof of Proposition 3.4.3 is to use the canonical probabilistic dual frame to give a constructive formula for $\mathbf{f}$. Another way to construct $\mathbf{f}$ is to use the canonical Parseval probabilistic frame ${\mathbf{S}_\mu^{-1/2}}_{\#}\mu$, i.e., for any $\mathbf{f} \in \mathbb{R}^n$,

$$\mathbf{f} = \int_{\mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{S}_\mu^{-1/2}\mathbf{x} \right\rangle \mathbf{S}_\mu^{-1/2}\mathbf{x} \; d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1/2}\mathbf{f}, \mathbf{x} \right\rangle \mathbf{S}_\mu^{-1/2}\mathbf{x} \; d\mu(\mathbf{x}).$$

According to this reconstruction formula, we get the last proposition in this section. Since it is similar to the previous proposition, we omit some details in the proof.

**Proposition 3.4.4.** *Let* $\mu$ *be a probabilistic frame for* $\mathbb{R}^n$ *with bounds* $0 < A \leq B < \infty$ *and* $\eta \in \mathscr{P}_2(\mathbb{R}^n)$. *Let* $\gamma \in \Gamma(\mu, \eta)$ *be any coupling with marginal* $\mu$ *and* $\eta$. *Suppose*

$$\tau := \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x}\|\|\mathbf{S}_\mu^{-1/2}\mathbf{x} - \mathbf{z}\| d\gamma(\mathbf{x},\mathbf{z}) < \sqrt{A},$$

*then* $\eta$ *is a probabilistic frame for* $\mathbb{R}^n$ *with bounds* $\frac{(\sqrt{A}-\tau)^2}{B}$ *and* $M_2(\eta)$.

*Proof.* Clearly $\eta$ is Bessel with bound $M_2(\eta)$. Since ${\mathbf{S}_\mu^{-1/2}}_{\#}\mu$ is a probabilistic Parseval frame, then

$$\mathbf{f} = \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1/2}\mathbf{f}, \mathbf{x} \right\rangle \mathbf{S}_\mu^{-1/2}\mathbf{x} \; d\mu(\mathbf{x}), \text{ for any } \mathbf{f} \in \mathbb{R}^n.$$

Define a linear operator $L : \mathbb{R}^n \to \mathbb{R}^n$ by

$$L(\mathbf{f}) = \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1/2}\mathbf{f}, \mathbf{x}\right\rangle \mathbf{z}\; d\gamma(\mathbf{x},\mathbf{z}), \text{ for any } \mathbf{f}\in\mathbb{R}^n.$$

Therefore,

$$\|\mathbf{f} - L(\mathbf{f})\| \le \int_{\mathbb{R}^n\times\mathbb{R}^n} \|\mathbf{x}\|\|\mathbf{S}_\mu^{-1/2}\mathbf{x} - \mathbf{z}\| d\gamma(\mathbf{x},\mathbf{z})\; \|\mathbf{S}_\mu^{-1/2}\mathbf{f}\| \le \frac{\tau}{\sqrt{A}}\|\mathbf{f}\|.$$

Thus $L : \mathbb{R}^n \to \mathbb{R}^n$ is invertible and $\|L^{-1}\| \le \frac{1}{1-\frac{\tau}{\sqrt{A}}}$. Then for any $\mathbf{f}\in\mathbb{R}^n$,

$$\mathbf{f} = LL^{-1}(\mathbf{f}) = \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1/2}L^{-1}\mathbf{f}, \mathbf{x}\right\rangle \mathbf{z}\; d\gamma(\mathbf{x},\mathbf{z}).$$

Therefore,

$$\begin{aligned}\|\mathbf{f}\|^4 = \langle \mathbf{f},\mathbf{f}\rangle^2 &= \Big| \int_{\mathbb{R}^n\times\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1/2}L^{-1}\mathbf{f}, \mathbf{x}\right\rangle \langle \mathbf{f},\mathbf{z}\rangle \; d\gamma(\mathbf{x},\mathbf{z})\Big|^2 \\ &\le \int_{\mathbb{R}^n} \left\langle \mathbf{S}_\mu^{-1/2}L^{-1}\mathbf{f}, \mathbf{x}\right\rangle^2 d\mu(\mathbf{x}) \; \int_{\mathbb{R}^n} \langle \mathbf{f},\mathbf{z}\rangle^2 d\eta(\mathbf{z}) \le \frac{B\|\mathbf{f}\|^2}{A(1-\frac{\tau}{\sqrt{A}})^2} \int_{\mathbb{R}^n} \langle \mathbf{f},\mathbf{z}\rangle^2 d\eta(\mathbf{z}).\end{aligned}$$

Thus for any $\mathbf{f}\in\mathbb{R}^n$,

$$\frac{A(1-\frac{\tau}{\sqrt{A}})^2}{B}\|\mathbf{f}\|^2 \le \int_{\mathbb{R}^n} \langle \mathbf{f},\mathbf{z}\rangle^2 d\eta(\mathbf{z}) \le M_2(\eta)\; \|\mathbf{f}\|^2.$$

Therefore, $\eta$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $\frac{(\sqrt{A}-\tau)^2}{B}$ and $M_2(\eta)$. □

The following example applies Proposition 3.4.3 and Proposition 3.4.4 to the case of finite frames for $\mathbb{R}^n$ by taking $\mu = \frac{1}{N}\sum_{i=1}^{N}\delta_{\mathbf{f}_i}$, $\nu = \frac{1}{M}\sum_{j=1}^{M}\delta_{\mathbf{g}_j}$, and $\gamma = \mu\otimes\nu \in \Gamma(\mu,\nu)$.

**Example 3.4.1.** *Let $M, N \ge n$, and $\{\mathbf{f}_i\}_{i=1}^N$ a frame for $\mathbb{R}^n$ with bounds $0 < A \le B < \infty$ and frame operator $\mathbf{S} = \sum_{i=1}^{N}\mathbf{f}_i\mathbf{f}_i^T$. If $\{\mathbf{g}_j\}_{j=1}^M$ is a sequence in $\mathbb{R}^n$ such that*

$$\epsilon := \frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}\|\mathbf{f}_i\|\|\mathbf{f}_i - \mathbf{g}_j\| < \frac{A}{N},$$

*then* $\{\mathbf{g}_j\}_{j=1}^M$ *is a frame with bounds* $\frac{M(A/N-\epsilon)^2}{B/N}$ *and* $\sum\limits_{j=1}^{M} \|\mathbf{g}_i\|^2$*; If* $\{\mathbf{g}_j\}_{j=1}^M$ *is such that*

$$\chi := \frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} \|\mathbf{S}^{-1}\mathbf{f}_i\| \ \|\mathbf{f}_i - \mathbf{g}_j\| < 1,$$

*then* $\{\mathbf{g}_j\}_{j=1}^M$ *is a frame with bounds* $\frac{MA^2(1-\chi)^2}{BN}$ *and* $\sum\limits_{j=1}^{M} \|\mathbf{g}_i\|^2$*; If* $\{\mathbf{g}_j\}_{j=1}^M$ *is such that*

$$\tau := \frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} \|\mathbf{f}_i\| \|\mathbf{S}^{-1/2}\mathbf{f}_i - \mathbf{g}_j\| < \sqrt{\frac{A}{N}},$$

*then* $\{\mathbf{g}_j\}_{j=1}^M$ *is a frame for* $\mathbb{R}^n$ *with bounds* $\frac{M(\sqrt{A/N}-\tau)^2}{B/N}$ *and* $\sum\limits_{j=1}^{M} \|\mathbf{g}_i\|^2$.

## 3.5 Quadratic Closeness of Probabilistic Frames

Recall that from Corollary 2.1.6, we show that the set of probabilistic frames is open in the 2-Wasserstein metric. In this section, we further show that one could obtain the frame bounds. It is well-known that if $\{\mathbf{f}_i\}$ is a frame for Hilbert space $\mathscr{H}$ with bounds $A$ and $B$, and $\{\mathbf{g}_i\}$ is a sequence such that

$$K := \sum_{i=1}^{\infty} \|\mathbf{f}_i - \mathbf{g}_i\|^2 < A,$$

then $\{\mathbf{g}_i\}$ is a frame for $\mathscr{H}$ with bounds $(\sqrt{A} - \sqrt{K})^2$ and $(\sqrt{B} + \sqrt{K})^2$ [Christensen, 1995a]. We get a similar result for probabilistic frames.

**Proposition 3.5.1.** *Let* $\mu$ *be a probabilistic frame with bounds* $A$ *and* $B$ *and* $\nu \in \mathscr{P}(\mathbb{R}^n)$*. Let* $\gamma \in \Gamma(\mu, \nu)$ *be any coupling with marginal* $\mu$ *and* $\nu$*. Suppose*

$$\lambda := \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x} - \mathbf{y}\|^2 d\gamma(\mathbf{x}, \mathbf{y}) < A.$$

*Then* $\nu$ *is a probabilistic frame with bounds* $(\sqrt{A} - \sqrt{\lambda})^2$ *and* $M_2(\nu)$*.*

*Proof.* Since

$$\lambda = \int_{\mathbb{R}^n} \|\mathbf{x}\|^2 d\mu(\mathbf{x}) + \int_{\mathbb{R}^n} \|\mathbf{y}\|^2 d\nu(\mathbf{y}) - 2 \int_{\mathbb{R}^n \times \mathbb{R}^n} \langle \mathbf{x}, \mathbf{x} \rangle \, d\gamma(\mathbf{x}, \mathbf{y}) < A < +\infty.$$

Then $M_2(\nu) = \int_{\mathbb{R}^n} \|\mathbf{y}\|^2 d\nu(\mathbf{y}) < +\infty$. Therefore, $\nu \in \mathscr{P}_2(\mathbb{R}^n)$ and $\nu$ is Bessel with bound $M_2(\nu)$. Next, let us show the lower frame bound. Note that $\mathbf{S}_{\mu}^{-1}{}_{\#}\mu$ is the probabilistic dual frame with bounds $\frac{1}{B}$ and $\frac{1}{A}$. Then for any $\mathbf{f} \in \mathbb{R}^n$,

$$\mathbf{f} = \int_{\mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{S}_{\mu}^{-1}\mathbf{x} \right\rangle \mathbf{x} d\mu(\mathbf{x}) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{S}_{\mu}^{-1}\mathbf{x} \right\rangle \mathbf{x} d\gamma(\mathbf{x}, \mathbf{y}).$$

For any $\mathbf{f} \in \mathbb{R}^n$, define a linear operator $L : \mathbb{R}^n \to \mathbb{R}^n$ by

$$L(\mathbf{f}) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{S}_{\mu}^{-1}\mathbf{x} \right\rangle \mathbf{y} d\gamma(\mathbf{x}, \mathbf{y}).$$

Therefore,

$$\begin{aligned}
\|\mathbf{f} - L(\mathbf{f})\|^2 &= \left\| \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\langle \mathbf{f}, \mathbf{S}_{\mu}^{-1}\mathbf{x} \right\rangle (\mathbf{x} - \mathbf{y}) \ d\gamma(\mathbf{x}, \mathbf{y}) \right\|^2 \\
&\leq \left( \int_{\mathbb{R}^n \times \mathbb{R}^n} | \left\langle \mathbf{f}, \mathbf{S}_{\mu}^{-1}\mathbf{x} \right\rangle | \ \|\mathbf{x} - \mathbf{y}\| \ d\gamma(\mathbf{x}, \mathbf{y}) \right)^2 \\
&\leq \int_{\mathbb{R}^n} | \left\langle \mathbf{f}, \mathbf{S}_{\mu}^{-1}\mathbf{x} \right\rangle |^2 d\mu(\mathbf{x}) \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\mathbf{x} - \mathbf{y}\|^2 d\gamma(\mathbf{x}, \mathbf{y}) \leq \frac{\lambda}{A} \|\mathbf{f}\|^2,
\end{aligned}$$

where the first two inequalities are due to triangle inequality and Cauchy Schwarz inequality, and the last inequality follows from the fact that $\mathbf{S}_{\mu}^{-1}{}_{\#}\mu$ is a probabilistic frame with upper bound $\frac{1}{A}$. Thus $L : \mathbb{R}^n \to \mathbb{R}^n$ is invertible and

$$\|L^{-1}\| \leq \frac{1}{1 - \sqrt{\frac{\lambda}{A}}}.$$

Then for any $\mathbf{f} \in \mathbb{R}^n$,

$$\mathbf{f} = LL^{-1}(\mathbf{f}) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\langle L^{-1}\mathbf{f}, \mathbf{S}_{\mu}^{-1}\mathbf{x} \right\rangle \mathbf{y} d\gamma(\mathbf{x}, \mathbf{y}).$$

Therefore,

$$\begin{aligned}
\|\mathbf{f}\|^4 = \langle \mathbf{f}, \mathbf{f} \rangle^2 &= \left| \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\langle L^{-1}\mathbf{f}, \mathbf{S}_{\mu}^{-1}\mathbf{x} \right\rangle \langle \mathbf{f}, \mathbf{y} \rangle \, d\gamma(\mathbf{x}, \mathbf{y}) \right|^2 \\
&\leq \int_{\mathbb{R}^n} \left\langle L^{-1}\mathbf{f}, \mathbf{S}_{\mu}^{-1}\mathbf{x} \right\rangle^2 d\mu(\mathbf{x}) \ \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{y} \rangle^2 \, d\nu(\mathbf{y}) \\
&\leq \frac{\|L^{-1}\mathbf{f}\|^2}{A} \ \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{y} \rangle^2 \, d\nu(\mathbf{y}) \\
&\leq \frac{\|\mathbf{f}\|^2}{A(1 - \sqrt{\frac{\lambda}{A}})^2} \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{y} \rangle^2 \, d\nu(\mathbf{y}),
\end{aligned}$$

where the first inequality is because of the Cauchy-Schwarz inequality and the second one follows from ${\mathbf{S}_\mu^{-1}}_{\#}\mu$ being a probabilistic frame with upper bound $\frac{1}{A}$. Thus for any $\mathbf{f} \in \mathbb{R}^n$,

$$A(1 - \sqrt{\frac{\lambda}{A}})^2 \|\mathbf{f}\|^2 \leq \int_{\mathbb{R}^n} \langle \mathbf{f}, \mathbf{y} \rangle^2 \, d\nu(\mathbf{y}) \leq M_2(\nu) \|\mathbf{f}\|^2.$$

Therefore, $\nu$ is a probabilistic frame for $\mathbb{R}^n$ with bounds $(\sqrt{A} - \sqrt{\lambda})^2$ and $M_2(\nu)$. □

If $\nu \in \mathscr{P}_2(\mathbb{R}^n)$ and $\gamma \in \Gamma(\mu, \nu)$ is the optimal coupling with respect to $W_2(\mu, \nu)$, we have the following corollary which ends this section.

**Corollary 3.5.2.** *Let $\mu$ be a probabilistic frame with bounds bounds $A$ and $B$ and $\nu \in \mathscr{P}_2(\mathbb{R}^n)$. Suppose*

$$W_2(\mu, \nu) < \sqrt{A},$$

*then $\nu$ is a probabilistic frame with bounds $(\sqrt{A} - W_2(\mu, \nu))^2$ and $M_2(\nu)$.*

*Proof.* Take $\gamma \in \Gamma(\mu, \nu)$ as the optimal transport coupling in Proposition 3.5.1. □

# Chapter 4

# Discussion

By the following two theorems, it is well-known that the closest Parseval frame to a given frame in $\mathbb{R}^n$ is the canonical Parseval frame, and the closest Parseval probabilistic frame to a given probabilistic frame for $\mathbb{R}^n$ is also the canonical probabilistic Parseval frame. It is natural to consider this minimization problem for probabilistic p-frames in the $p$–Wasserstein topology where $p \geq 1$. In addition, restricting probabilistic frames to the unit sphere $\mathbb{S}^{n-1}$ is also an interesting topic.

**Theorem 4.0.1** ( [Casazza and Kutyniok, 2007])**.** *If $\mathbf{f} = \{\mathbf{f}_i\}_{i=1}^N$ is a frame for $\mathbb{R}^n$, then the minimizing problem*

$$min\Big\{d(\mathbf{f},\mathbf{g}) = \sum_{i=1}^{N} \|\mathbf{f}_i - \mathbf{g}_i\|^2 : \mathbf{g} = \{\mathbf{g}_i\}_{i=1}^N \subset \mathbb{R}^n,\ \mathbf{g}\ is\ Parseval\Big\}$$

*admits a unique solution given by $\mathbf{f}^* = \{\mathbf{S}^{-1/2}\mathbf{f}_i\}_{i=1}^N$ where $\mathbf{S}$ is defined by*

$$\mathbf{S}(\mathbf{x}) = \sum_{i=1}^{N} \langle \mathbf{x}, \mathbf{f}_i \rangle \, \mathbf{f}_i,\ \ \mathbf{x} \in \mathbb{R}^n$$

**Theorem 4.0.2** ( [Cheng and Okoudjou, 2019,Loukili and Maslouhi, 2020])**.** *Let $\mu$ be a probabilistic frame for $\mathbb{R}^n$, then $I_{\mathscr{P}}(\mu, \mathbb{R}^n) := \inf_{\nu \in \mathscr{T}_{\mathscr{P}}(\mathbb{R}^n)} W_2(\mu, \nu)$ admits an unique optimizer $\mu^* = \mathbf{S}_\mu^{-\frac{1}{2}}{}_{\#}\mu$, and*

$$I_{\mathscr{P}}(\mu, \mathbb{R}^n) = W_2(\mu, \mu^*) = \sqrt{\text{tr } (\mathbf{S}_\mu^{\frac{1}{2}} - \mathbf{Id})^2},$$

*where $\mathscr{T}_{\mathscr{P}}(\mathbb{R}^n)$ be the set of Parseval probabilistic frames in $\mathbb{R}^n$.*

## 4.1 Minimization of Probabilistic p-Frames in $W_p$

Recall that $\mu \in \mathscr{P}_p(\mathbb{R}^n)$ is called probabilistic p-frame if there exist $0 < A \leq B$ such that for any $\mathbf{x} \in \mathbb{R}^n$,

$$A\|\mathbf{x}\|^p \leq \int_{\mathbb{R}^n} |\langle \mathbf{x}, \mathbf{y} \rangle|^p \, d\mu(\mathbf{y}) \leq B\|\mathbf{x}\|^p.$$

Furthermore, $\mu$ is called a tight frame if $A = B$, and Parseval if $A = B = 1$.

One natural question is finding the closest probabilistic Parseval p-frame to a given probabilistic p-frame in the $p$-Wasserstein topology where $p \geq 1$.

**Problem 4.1.1.** *Let $\mu$ be a probabilistic p-frame for $\mathbb{R}^n$ where $p \geq 1$. Then consider the following minimization problem*

$$\mathit{inf}_{\nu \in \mathscr{T}^p_{\mathscr{P}}(\mathbb{R}^n)} W_p(\mu, \nu),$$

*where $\mathscr{T}^p_{\mathscr{P}}(\mathbb{R}^n)$ is the set of Parseval probabilistic p-frames for $\mathbb{R}^n$.*

## 4.2 Minimization of Probabilistic Frames on the Unit Sphere

When designing frames, people like to sample points from the unit sphere $\mathbb{S}^{n-1}$ to get unit-norm frames, which leads to the following definition about probabilistic frame on $\mathbb{S}^{n-1}$.

**Definition 4.2.1** ( [Ehler, 2012])**.** *$\mu \in \mathscr{P}(\mathbb{S}^{n-1})$ is a probabilistic frame on $\mathbb{S}^{n-1}$ if there exist constants $0 < A \leq B < \infty$ such that for any $\mathbf{x} \in \mathbb{R}^n$,*

$$A\|\mathbf{x}\|^2 \leq \int_{\mathbb{S}^{n-1}} |\langle \mathbf{x}, \mathbf{y} \rangle|^2 d\mu(\mathbf{y}) \leq B\|\mathbf{x}\|^2.$$

*$\mu$ is said to be a tight probabilistic frame if $A = B$, and Parseval if $A = B = 1$.*

For a given probabilistic frame $\mu$ on $\mathbb{S}^{n-1}$, the following minimization problem about finding the closest probabilistic Parseval frame on $\mathbb{S}^{n-1}$ is still open:

$$I_{\mathscr{P}}(\mu, \mathbb{S}^{n-1}) := \inf_{\nu \in \mathscr{T}_{\mathscr{P}}(\mathbb{S}^{n-1})} W_2(\mu, \nu),$$

where $\mathscr{T}_{\mathscr{P}}(\mathbb{S}^{d-1})$ is the set of probabilistic Parseval frames on $\mathbb{S}^{n-1}$. The difficulty is due to the absence of linear structures on the unit sphere $\mathbb{S}^{d-1}$. It is clear that the optimizer for $I_{\mathscr{P}}(\mu, \mathbb{S}^{n-1})$ exists since $\mathscr{T}_{\mathscr{P}}(\mathbb{S}^{n-1})$ is weakly compact by the following lemma and corollary.

**Lemma 4.2.2.** $\mathscr{T}_{\mathscr{P}}(\mathbb{S}^{n-1})$ *is convex and weakly compact.*

*Proof.* It is clear that $\mathscr{T}_{\mathscr{P}}(\mathbb{S}^{n-1})$ is convex. Since $\mathscr{P}(\mathbb{S}^{n-1})$ is compact in the weak topology, it suffices to show that $\mathscr{T}_{\mathscr{P}}(\mathbb{S}^{n-1})$ is weakly closed. Let $\{\nu_n\}_{n=1}^{\infty} \subset \mathscr{T}_{\mathscr{P}}(\mathbb{S}^{n-1})$ be a sequence such that $\{\nu_n\}_{n=1}^{\infty}$ converges weakly to $\nu \in \mathscr{P}(\mathbb{S}^{n-1})$. Since the $(i,j)$–th entry of the frame operator $\mathbf{S}_{\nu_n}$ is $\delta_{ij}$, then

$$\int_{\mathbb{S}^{n-1}} y_i y_j d\nu(\mathbf{y}) = \lim_{n\to\infty} \int_{\mathbb{S}^{n-1}} y_i y_j d\nu_n(\mathbf{y}) = \delta_{ij}.$$

Therefore, $\mathbf{S}_\nu = \mathbf{Id}$ and thus $\nu$ is a probabilistic Parseval frame on $\mathbb{S}^{n-1}$, which implies $\mathscr{T}_{\mathscr{P}}(\mathbb{S}^{n-1})$ is weakly closed and thus weakly compact. □

**Corollary 4.2.3.** $I_{\mathscr{P}}(\mu, \mathbb{S}^{n-1})$ *admits an optimizer in* $\mathscr{T}_{\mathscr{P}}(\mathbb{S}^{n-1})$.

It is interesting to find the optimizer for $I_{\mathscr{P}}(\mu, \mathbb{S}^{n-1})$. Note that the canonical probabilistic Parseval frame $\mathbf{S}_\mu^{-\frac{1}{2}}{}_{\#}\mu$ may not be supported on the unit sphere. Therefore, $\mathbf{S}_\mu^{-\frac{1}{2}}{}_{\#}\mu$ may not be an optimizer. On the other hand, since $\mathscr{T}_{\mathscr{P}}(\mathbb{S}^{n-1}) \subset \mathscr{T}_{\mathscr{P}}(\mathbb{R}^n)$, then $I_{\mathscr{P}}(\mu, \mathbb{R}^n)$ is a lower bound of $I_{\mathscr{P}}(\mu, \mathbb{S}^{n-1})$.

**Corollary 4.2.4.** $I_{\mathscr{P}}(\mu, \mathbb{S}^{n-1}) \geq I_{\mathscr{P}}(\mu, \mathbb{R}^n) = W_2(\mu, \mathbf{S}_\mu^{-\frac{1}{2}}{}_{\#}\mu) = \sqrt{\text{tr } (\mathbf{S}_\mu^{\frac{1}{2}} - \mathbf{Id})^2}$.

# Bibliography


[Arefijamaal and Sadeghi, 2016] Arefijamaal, A. A. and Sadeghi, G. (2016). von Neumann–Schatten dual frames and their perturbations. *Results in Mathematics*, 69:431–441.

[Boas Jr, 1940] Boas Jr, R. (1940). General expansion theorems. *Proceedings of the National Academy of Sciences*, 26(2):139–143.

[Bodmann and Casazza, 2010] Bodmann, B. G. and Casazza, P. G. (2010). The road to equal-norm parseval frames. *Journal of Functional Analysis*, 258(2):397–420.

[Bourgain, 1986] Bourgain, J. (1986). On high dimensional maximal functions associated to convex bodies. *American Journal of Mathematics*, 108(6):1467–1476.

[Cahill et al., 2017] Cahill, J., Mixon, D. G., and Strawn, N. (2017). Connectivity and irreducibility of algebraic varieties of finite unit norm tight frames. *SIAM Journal on Applied Algebra and Geometry*, 1(1):38–72.

[Casazza, 2000] Casazza, P. G. (2000). The art of frame theory. *Taiwanese journal of mathematics*, 4(2):129–201.

[Casazza, 2001] Casazza, P. G. (2001). Modern tools for weyl-heisenberg (gabor) frame theory. In *Advances in Imaging and Electron Physics*, volume 115, pages 1–127. Elsevier.

[Casazza, 2013] Casazza, P. G. (2013). The Kadison–Singer and Paulsen problems in finite frame theory. *Finite frames: theory and applications*, pages 381–413.

[Casazza and Christensen, 1998] Casazza, P. G. and Christensen, O. (1998). Frames containing a Riesz basis and preservation of this property under perturbations. *SIAM Journal on Mathematical Analysis*, 29(1):266–278.

[Casazza and Christensen, 2008] Casazza, P. G. and Christensen, O. (2008). The reconstruction property in Banach spaces and a perturbation theorem. *Canadian Mathematical Bulletin*, 51(3):348–358.

[Casazza and Kutyniok, 2007] Casazza, P. G. and Kutyniok, G. (2007). A generalization of gram–schmidt orthogonalization generating all parseval frames. *Advances in Computational Mathematics*, 27(1):65–78.

[Casazza and Kutyniok, 2012] Casazza, P. G. and Kutyniok, G. (2012). *Finite frames: Theory and applications*. Springer Science & Business Media.

[Casazza and Lynch, 2016] Casazza, P. G. and Lynch, R. G. (2016). A brief introduction to Hilbert space frame theory and its applications. *Finite Frame Theory: A Complete Introduction to Overcompleteness*, 93(1):2.

[Cazassa and Christensen, 1997] Cazassa, P. G. and Christensen, O. (1997). Perturbation of operators and applications to frame theory. *Journal of Fourier Analysis and Applications*, 3(5):543–557.

[Chen, 2023] Chen, D. (2023). Paley-wiener theorem for probabilistic frames. *arXiv preprint arXiv:2310.17830*.

[Chen et al., 2014] Chen, D. Y., Li, L., and Zheng, B. T. (2014). Perturbations of frames. *Acta Mathematica Sinica, English Series*, 30(7):1089–1108.

[Cheng, 2018] Cheng, D. (2018). *Frames and subspaces.* PhD thesis, University of Missouri–Columbia.

[Cheng and Okoudjou, 2019] Cheng, D. and Okoudjou, K. A. (2019). Optimal properties of the canonical tight probabilistic frame. *Numerical Functional Analysis and Optimization*, 40(2):216–240.

[Christensen, 1995a] Christensen, O. (1995a). Frame perturbations. *Proceedings of the American Mathematical Society*, 123(4):1217–1220.

[Christensen, 1995b] Christensen, O. (1995b). A Paley–Wiener theorem for frames. *Proceedings of the American Mathematical Society*, 123(7):2199–2201.

[Christensen, 2016] Christensen, O. (2016). An introduction to frames and Riesz bases. *Applied and Numerical Harmonic Analysis*.

[Christensen and Hasannasab, 2017] Christensen, O. and Hasannasab, M. (2017). Operator representations of frames: boundedness, duality, and stability. *Integral Equations and Operator Theory*, 88:483–499.

[Christensen and Heil, 1997] Christensen, O. and Heil, C. (1997). Perturbations of Banach frames and atomic decompositions. *Mathematische Nachrichten*, 185(1):33–47.

[Christensen et al., 2000] Christensen, O., Lennard, C., and Lewis, C. (2000). Perturbation of frames for a subspace of a Hilbert space. *The Rocky Mountain Journal of Mathematics*, pages 1237–1249.

[Cuesta-Albertos et al., 1996] Cuesta-Albertos, J. A., Matrán-Bea, C., and Tuero-Diaz, A. (1996). On lower bounds for the l2-wasserstein metric in a hilbert space. *Journal of Theoretical Probability*, 9(2):263–284.

[Daubechies, 1992] Daubechies, I. (1992). *Ten lectures on wavelets*. SIAM.

[Duffin and Schaeffer, 1952] Duffin, R. J. and Schaeffer, A. C. (1952). A class of nonharmonic Fourier series. *Transactions of the American Mathematical Society*, 72(2):341–366.

[Ehler, 2012] Ehler, M. (2012). Random tight frames. *Journal of Fourier Analysis and Applications*, 18(1):1–20.

[Ehler and Galanis, 2011] Ehler, M. and Galanis, J. (2011). Frame theory in directional statistics. *Statistics & probability letters*, 81(8):1046–1051.

[Ehler and Okoudjou, 2012] Ehler, M. and Okoudjou, K. A. (2012). Minimization of the probabilistic p-frame potential. *Journal of Statistical Planning and Inference*, 142(3):645–659.

[Ehler and Okoudjou, 2013] Ehler, M. and Okoudjou, K. A. (2013). Probabilistic frames: an overview. *Finite frames*, pages 415–436.

[Eldar, 2003] Eldar, Y. C. (2003). Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *Journal of Fourier Analysis and Applications*, 9:77–96.

[Fickus et al., 2013] Fickus, M., Massar, M. L., and Mixon, D. G. (2013). Finite frames and filter banks. *Finite Frames: Theory and Applications*, pages 337–379.

[Figalli and Glaudo, 2021] Figalli, A. and Glaudo, F. (2021). *An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows.* EMS Press.

[Gelbrich, 1990] Gelbrich, M. (1990). On a formula for the $l^2$ wasserstein metric between measures on euclidean and hilbert spaces. *Mathematische Nachrichten*, 147(1):185–203.

[Gröchenig, 2001] Gröchenig, K. (2001). *Foundations of time-frequency analysis.* Springer Science & Business Media.

[Han, 2007] Han, D. (2007). *Frames for undergraduates*, volume 40. American Mathematical Society.

[Javanshiri et al., 2022] Javanshiri, H., Hajiabootorabi, M., and Mardanbeigi, M. R. (2022). The effect of perturbations of frames on their alternate and approximately dual frames. *Mathematical Methods in the Applied Sciences*, 45(4):2058–2071.

[Khokulan and Thirulogasanthar, 2019] Khokulan, M. and Thirulogasanthar, K. (2019). Perturbation of continuous frames on quaternionic Hilbert spaces. *arXiv preprint arXiv:1905.09393.*

[Krishna and Johnson, 2022] Krishna, K. M. and Johnson, P. S. (2022). Frames for metric spaces. *Results in Mathematics*, 77(1):49.

[Kutyniok and Labate, 2012] Kutyniok, G. and Labate, D. (2012). *Multiscale analysis for multivariate data.* Springer.

[Lau and Okoudjou, 2015] Lau, C. W. and Okoudjou, K. A. (2015). Scalable probabilistic frames. *ArXiv preprint arXiv:1501.07321.*

[Li et al., 2016] Li, D., Leng, J., Huang, T., and Xu, Y. (2016). Some equalities and inequalities for probabilistic frames. *Journal of Inequalities and Applications*, 2016(1):1–11.

[Loukili and Maslouhi, 2020] Loukili, S. and Maslouhi, M. (2020). A minimization problem for probabilistic frames. *Applied and Computational Harmonic Analysis*, 49(2):558–572.

[Maslouhi and Loukili, 2019] Maslouhi, M. and Loukili, S. (2019). Probabilistic tight frames and representation of positive operator-valued measures. *Applied and Computational Harmonic Analysis*, 47(1):212–225.

[Naidu and Murthy, 2020] Naidu, R. R. and Murthy, C. R. (2020). Construction of unimodular tight frames for compressed sensing using majorization-minimization. *Signal Processing*, 172:107516.

[Olkin and Pukelsheim, 1982] Olkin, I. and Pukelsheim, F. (1982). The distance between two random vectors with given dispersion matrices. *Linear Algebra and its Applications*, 48:257–263.

[Paley and Wiener, 1934] Paley, R. E. A. C. and Wiener, N. (1934). *Fourier transforms in the complex domain*, volume 19. American Mathematical Society.

[Poria, 2017] Poria, A. (2017). Approximation of the inverse frame operator and stability of Hilbert–Schmidt frames. *Mediterranean Journal of Mathematics*, 14(4):153.

[Singer, 1970] Singer, I. (1970). *Bases in Banach spaces I.* Springer.

[Strohmer, 2001] Strohmer, T. (2001). Approximation of dual gabor frames, window decay, and wireless communications. *Applied and Computational Harmonic Analysis*, 11(2):243–262.

[Strohmer and Heath Jr, 2003] Strohmer, T. and Heath Jr, R. W. (2003). Grassmannian frames with applications to coding and communication. *Applied and Computational Harmonic analysis*, 14(3):257–275.

[Sun, 2007] Sun, W. (2007). Stability of g-frames. *Journal of Mathematical Analysis and Applications*, 326(2):858–868.

[V.D. and Pajor, 1989] V.D., M. and Pajor, A. (1989). Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space. *Lecture Notes in Mathematics*, 1376:64–104.

[Villani, 2003] Villani, C. (2003). *Topics in Optimal Transportation*, volume 58. American Mathematical Soc.

[Villani et al., 2009] Villani, C. et al. (2009). *Optimal transport: old and new*, volume 338. Springer.

[Wickman and Okoudjou, 2017] Wickman, C. and Okoudjou, K. (2017). Duality and geodesics for probabilistic frames. *Linear Algebra and Its Applications*, 532:198–221.

[Wickman and Okoudjou, 2023] Wickman, C. and Okoudjou, K. A. (2023). Gradient flows for probabilistic frame potentials in the Wasserstein space. *SIAM Journal on Mathematical Analysis*, 55(3):2324–2346.

[Wickman, 2014] Wickman, C. G. (2014). *An optimal transport approach to some problems in frame theory.* PhD thesis, University of Maryland, College Park.

[Young, 2001] Young, R. M. (2001). *An Introduction to Non-Harmonic Fourier Series, Revised Edition, 93.* Elsevier.

[Zhang and Li, 2022] Zhang, X.-L. and Li, Y.-Z. (2022). Portraits and perturbations of Hilbert–Schmidt frame sequences. *Bulletin of the Malaysian Mathematical Sciences Society*, 45(6):3197–3223.