Degrees of Freedom of a Communication Channel: Using DOF Singular Values

Ram Somaraju and Jochen Trumpf

Abstract

A fundamental problem in any communication system is: given a communication channel between a transmitter and a receiver, how many "independent" signals can be exchanged between them? Arbitrary communication channels that can be described by linear compact channel operators mapping between normed spaces are examined in this paper. The (well-known) notions of degrees of freedom at level ϵ and essential dimension of such channels are developed in this general setting. We argue that the degrees of freedom at level ϵ and the essential dimension fundamentally limit the number of independent signals that can be exchanged between the transmitter and the receiver. We also generalise the concept of singular values of compact operators to be applicable to compact operators defined on arbitrary normed spaces which do not necessarily carry a Hilbert space structure. We show how these generalised singular values, which we call Degrees of Freedom (DOF) singular values, can be used to calculate the degrees of freedom at level ϵ and the essential dimension of compact operators that describe communication channels. We describe physically realistic channels that require such general channel models.

Index Terms

Operator Channels, Degrees of Freedom, Singular Values, Essential Dimension

Degrees of Freedom of a Communication Channel: Using DOF Singular Values

I. Introduction

The basic consideration in this paper can be stated as follows: given an arbitrary communication channel, is it possible to evaluate the number of independent sub-channels or modes available for communication. Though this question is not generally examined explicitly, it plays an important role in various information theoretic problems.

A rigorous proof of Shannon's famous capacity result [1] for continuous-time band-limited white Gaussian noise channels requires a calculation of the number of approximately time-limited and band-limited sub-channels (see e.g. [2, ch. 8] and [3,4]). This result can be generalised to dispersive/non-white Gaussian channels using the *water-filling* formula [1,2]. In order to use this formula, one needs to diagonalise the channel operator and allocate power to the different sub-channels or modes based on the singular values of the corresponding sub-channel. One therefore needs to calculate the modes and the power transferred (square of the singular values) on each one of these sub-channels to calculate the channel capacity.

The water-filling formula has been used extensively in order to calculate the capacity of channels that use different forms of diversity. In particular, the capacity of multiple-input multiple-output (MIMO) antenna systems has been calculated using this water-filling formula for various conditions imposed on the transmitting and the receiving antennas (see e.g. [5] and references therein). Water-filling type formulas have been used for other multi-access schemes such as OFDM-MIMO [6] and CDMA [7] (see also Tulino [8, sec 1.2] and references therein). More recently, several papers have examined the number of degrees of freedom[1] available in spatial channels [9]–[13]. Questions of this nature have also been studied in other contexts such as optics [14] and spatial sampling of electromagnetic waves [15,16].

Both types of results, the modes of communication used for the water-filling formula and the number of degrees of freedom of spatial channels use the singular

[1]Note that other terms such as *modes of communication*, *essential dimension* etc. have been used instead of degrees of freedom in some of these papers.

value decomposition (SVD) theorem. One can use SVD to diagonalise the channel operator and the magnitude of the singular values determines the power transferred on each of the sub-channels. The magnitude of these singular values can therefore be used to calculate the number of degrees of freedom of the channel (see e.g. [9,12]). However, the SVD theorem is only applicable to compact operators defined on Hilbert spaces. An implicit and valid assumption that is used in these papers is that the operators describing the communication channels are defined on Hilbert spaces. These results can therefore not be generalised directly to communication systems that are modeled by operators defined on normed spaces that do not admit an inner product structure. There are several instances of practical channels that can not be modeled using operators defined on inner-product spaces (see Section II-A for examples). In this paper, we develop a general theory that enables one to evaluate the number of degrees of freedom of such systems.

We wish to examine if it is possible to evaluate the number of parallel sub-channels available in general communication systems that can be described using linear compact operators. Any communication channel is subject to various physical constraints such as noise at the receiver or finite power available for transmission. If the channel can be modeled via a linear compact operator, then these constraints ensure that only finitely many independent channels are available for communication. Roughly speaking, we call the number of such channels the number of degrees of freedom of the communication system (see Section III for a precise definition). Note that if the channel is modeled using a linear operator that is not compact then it will in fact have infinitely many parallel sub-channels, or some channels that can transfer an infinite amount of power (see Theorem 3.10 below and the discussion following it). It could hence be argued that the theory presented in this paper is the most general theory needed to model physically realistic channels.

We give novel definitions for the terms degrees of freedom and essential dimension in the following section. Even though these terms have been used interchangeably in the literature, we distinguish between the two. The essential dimension of a channel is useful for channels

that have numbers of degrees of freedom that are essentially independent of the receiver noise level (e.g. the time-width/band-width limited channels in Slepian's work [17]). Also, we generalise the notion of singular values to compact operators defined on normed spaces and explain how these generalised singular values, which we call Degrees of Freedom (DOF) singular values can be used to compute degrees of freedom and the essential dimension.

A. Channel Model

We assume that a communication channel between a transmitter and a receiver can be modeled as follows. Let X be a linear vector space of functions that the transmitter can generate and let Y be a linear vector space of functions that the receiver can measure. We assume the existence of a linear operator $T : X \to Y$ that maps each signal generated by a transmitter to a signal that a receiver can measure. We also assume that there is a norm $\| \cdot \|_X$ on X and a norm $\| \cdot \|_Y$ on Y. This model is very general and can be applied to various situations of practical relevance.

For instance, consider a MIMO communication system wherein the transmitter symbol waveform shape on each antenna is a raised cosine. In this case we can think of the space of transmitter functions X to be (more precisely, to be parametrised by) the n-dimensional complex space $\mathbb{C}^n$ that determines the phase and amplitude of the raised cosine waveform on each antenna. Here n is the number of transmitting antennas. Also, we can think of the space of receiver functions as $\mathbb{C}^m$, where m is the number of receiving antennas. T in this context is a channel matrix, representing the linearized channel operator that depends on the scatterers in the environment.

Alternatively, consider a MIMO communication system in which the transmitter symbols are not fixed but can be any waveform of time. Suppose the symbol time is fixed to t_s seconds. In this case, we can think of the space of transmitter functions, X, as the space $\mathcal{L}^2([0, t_s], \mathbb{C}^n)$ of $\mathbb{C}^n$-valued square integrable functions defined on $[0, t_s]$. Similarly, we can think of the space of receiver functions, Y, as the space $\mathcal{L}^2([0, t_s], \mathbb{C}^m)$. Again, T is the channel operator.

Irrespective of the precise form of the underlying spaces X and Y, we always call elements of X transmitter functions and the elements of Y receiver functions. Also, we call the space X the space of transmitter functions and the space Y the space of receiver functions. In particular, we do not distinguish between the two different physical situations: a) the elements of X are

functions of time and b) the elements of X are vectors in some finite dimensional space. This should cause no confusion and we use this convention for the remainder of this document.

We now restrict ourselves to situations where there is a source constraint $\| \cdot \|_X \leq P$ that can be imposed on the space of transmitter functions X, and where the operator T is compact. Roughly speaking, the norm on the space of transmitter functions X captures the physical restriction that the transmitter functions can not be *arbitrarily big*, while the norm on the space of receiver functions can be interpreted as a measure of how *big* the received signals are compared to a pre-specified noise level. We therefore try to find how many linearly independent signals can be generated at the receiver that are *big* enough by transmitter functions that are not *too big*. The compactness of the operator T ensures that only finitely many independent signals can be received (see Section II-A for examples of such channels). This vague idea is clarified further in the following two sections.

B. Outline

The remainder of this paper is organised as follows: in the next section we consider a finite dimensional example and motivate the definition of degrees of freedom. We also discuss several examples of practical communication systems to which the theory developed in this paper may be applied. Section III presents the main results of this paper as well as formal definitions of degrees of freedom, essential dimension and DOF singular values. Conclusions are presented in Section IV. Detailed proofs of the theorems in this paper are presented in the Appendix.

Most of the material presented in this paper forms part of the first author's PhD thesis [18].

II. MOTIVATION

We motivate our definition of degrees of freedom at level ϵ for compact operators on normed spaces by considering linear operators on finite dimensional spaces. Consider a communication channel that uses n transmitting antennas and m receiving antennas which can be mathematically modeled as follows. Let the current on the n transmitting antennas be given by $\mathbf{x} \in \mathbb{C}^n$. This current on the transmitting antennas generates a current $\mathbf{y} \in \mathbb{C}^m$ in the m receiving antennas according to the equation

$$\mathbf{y} = \mathbf{H}\mathbf{x}.$$

Here, $\mathbf{H} \in \mathbb{C}^{m \times n}$ is the channel matrix. We can define the operator $T : \mathbb{C}^n \to \mathbb{C}^m$ by $\mathbf{x} \mapsto \mathbf{y} = \mathbf{H}\mathbf{x}$. Also,

for $n = 1, 2, \ldots$, $\| \cdot \| = \sqrt{(\cdot)^*(\cdot)}$, with $(\cdot)^*$ denoting the complex conjugate transpose, is the standard norm in $\mathbb{C}^n$. In this context, the norm determines the power of the signal on the antennas.

The singular value decomposition theorem tells us that there exist sets of orthonormal basis vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subset \mathbb{C}^n$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\} \subset \mathbb{C}^m$ such that the matrix representation for T in these bases is diagonal. Let $\mathbf{H}_d$ be such a matrix with the basis vectors ordered such that the diagonal elements (i.e. the singular values of T) are in non-increasing order. A simple examination of the diagonal matrix proves that for all $\epsilon > 0$ there exist a number N and a set of linearly independent vectors $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\} \subset \mathbb{C}^m$ such that for all $\mathbf{x} \in \overline{B}_{1, \mathbb{C}^n}(0)^2$

$$\inf_{a_1, \ldots, a_N} \left\| \mathbf{H}_d \mathbf{x} - \sum_{i=1}^{N} a_i \mathbf{y}_i \right\| \leq \epsilon.$$

For a given ϵ, call the smallest number that satisfies the above condition $\mathcal{N}(\epsilon)$. Note that the vectors $\mathbf{y}_1, \ldots, \mathbf{y}_N$ span the space of all linear combinations of the left singular vectors of T whose corresponding singular values are greater than or equal to ϵ.

A simple examination of the diagonal matrix tells us that $\mathcal{N}(\epsilon)$ is equal to the number of singular values of T that are greater than ϵ and is hence clearly independent of the bases chosen. This leads us to our definition for degrees of freedom in finite dimensional spaces.

Definition 2.1: Let $T : \mathbb{C}^n \to \mathbb{C}^m$ be a linear operator and let $\epsilon > 0$ be given. Then the number of degrees of freedom at level ϵ for T is the smallest number $\mathcal{N}(\epsilon)$ such that there exists a set of vectors $\mathbf{y}_1, \ldots, \mathbf{y}_N \in \mathbb{C}^m$ such that for all $\mathbf{x} \in \overline{B}_{1, \mathbb{C}^n}(0)$

$$\inf_{a_1, \ldots, a_N} \left\| T\mathbf{x} - \sum_{i=1}^{N} a_i \mathbf{y}_i \right\| \leq \epsilon.$$

This definition is appropriate for the number of degrees of freedom because for a MIMO system the norm $\| \cdot \|$ represents the power in the signal. Suppose we wish to transmit N linearly independent signals from the transmitter to the receiver, and the total power available for transmission is bounded. Suppose further that the received signal is measured in the presence of noise. By requiring that $\mathbf{x} \in \overline{B}_{1, \mathbb{C}^n}(0)$ we are constraining the power available for transmission. We model the noise by assuming that any two signals at the receiver can be distinguished if the power of the difference between the signals is greater than some level ϵ. Similar ideas have been used for instance by Bucci *et. al.* [16] (see also [4,10,17]). According to this definition, the number

of degrees of freedom is equal to the number of linearly independent signals that the receiver can distinguish under the assumptions of a transmit power constraint and a receiver noise level represented by ϵ. Note that we are making the implicit assumption that the power P is 1 in the above definition. This does not cause a problem because we can always scale the norm in order to consider situations where $P \neq 1$.

The above definition was motivated using the singular value decomposition theorem in finite dimensional spaces. It can therefore be easily generalised to infinite dimensional Hilbert spaces using the corresponding singular value decomposition in infinite dimensional Hilbert spaces (see eg. [16,18][3]). However, the singular value decomposition can only be used for operators defined on Hilbert spaces. It cannot be used for operators defined on general normed spaces. Observe that the definition for degrees of freedom above only depends on the norm $\| \cdot \|$ and not on the assumption that the underlying spaces $\mathbb{C}^n$ and $\mathbb{C}^m$ are Hilbert spaces. It will be shown in this paper that the above definition can be extended to compact operators defined on arbitrary normed spaces.

Now consider the situation where the singular values of the operator T show a step like behavior. For instance, suppose the singular values are $\{1, 0.9, 0.85, 0.5, 0.1, 0.05, .0005\}$. In this particular case, the number of degrees of freedom is essentially independent of the actual value of ϵ chosen because for $\epsilon \in (0.1, 0.5)$, the number of degrees of freedom is a constant. This range (0.1-0.5) is big compared to the total range (0.0005 - 1.0) which contains all the singular values. Such a situation arises in several important cases (see eg. [4,9,14,16,17]). It would be useful to have a general way in which one can specify a number of degrees of freedom of a channel that is independent of the arbitrarily chosen level ϵ. In this paper we provide a novel definition for such a number and call it the essential dimension of the channel. This definition is sufficiently general to be applicable to a variety of channels and quantifies the essential dimension of any channel that can be described using a compact operator.

A. Examples

As explained in section I-A, we assume that a communication channel can be described using the triple X, Y and T. Here X is the space of transmitter functions, Y is the space of receiver functions and T is the channel operator and is assumed to be compact. As explained earlier in this section, if the spaces X and Y are Hilbert spaces and if the operator T is a linear compact operator

---

[2]Given a normed space X, $r \geq 0$ and $x \in X$, $\overline{B}_{r,X}(x)$ denotes the closed ball of radius r centered at $x \in X$.

[3]Also compare with the time-bandwidth problem in [4,17].

then the well known theory of singular values of Hilbert space operators can be used to determine the number of degrees of freedom of such channels. However, if either one of the spaces X or Y is not an inner product space then one cannot use this theory.

There are several practical channels that are best described using abstract spaces that do not admit an inner product structure. In this subsection, we consider three examples of such channels. In the first example, the measurement technique used in the receiver restricts the space of receiver functions. In the second one, the modulation technique used means that the constraints on the space of transmitter functions are best described using a norm that is not compatible with an inner product. The final example discusses a physical channel that naturally admits a norm on the space of transmitter functions that is described using a vector product and therefore does not admit an inner-product structure.

Example 2.1: In any practical digital communication system, the receiver is designed to receive a finite set of transmitted signals. Suppose the transmitted signal is generated from a source alphabet $\{t_1, \ldots, t_N\}$ and for simplicity assume that in a noiseless system each element from the source alphabet $t_i, 1 \leq i \leq N$, generates a signal $r_i, 1 \leq i \leq N$, at the receiver. In the corresponding noisy system, the fundamental problem is to determine which element from the source alphabet was transmitted given the signal $r = r_i + n$ was received. Here, n is the noise in the system. One common approach to solving this problem is to define some metric $d(\cdot, \cdot)$ that measures the distance between two receiver signals and to calculate

$$r' = \underset{\{r_i, 1 \leq i \leq n\}}{\operatorname{argmin}} \; d(r, r_i).$$

One concludes that the element from the source alphabet that corresponds to r' is (most likely) the transmitted signal. Generally, this metric $d(\cdot, \cdot)$ determines the abstract space Y of receiver function.

Now consider a MIMO antenna system with n transmitting and m receiving antennas. Suppose that the receiver measures the signals on the m receiving antennas for a period of τ seconds. One can describe the received signal by a function $\mathbf{y}(t)$, where $\mathbf{y} : [0, \tau] \to \mathbb{C}^m$. In order to implement the receiver one can use a matched filter if the shapes of all noiseless receiver signals are known. In this case the distance between two received signals can be described using the metric

$$d(\mathbf{y}_1, \mathbf{y}_2) = \left(\int_0^\tau (\mathbf{y}_1(t) - \mathbf{y}_2(t))^* (\mathbf{y}_1(t) - \mathbf{y}_2(t)) dt \right)^{1/2}$$

One can describe the space of receiver functions using the Hilbert space $\mathcal{L}^2([0, \tau], \mathbb{C}^m)$ with the inner product

defined by

$$\langle \mathbf{y}_1, \mathbf{y}_2 \rangle := \int_0^\tau \mathbf{y}_1^*(t) \mathbf{y}_2(t) dt.$$

This is the common approach used in information theory.

However, it is generally easier to measure just the amplitude of the received signal on each of the m antennas. In fact, in a rapidly changing environment it might not be possible to build an effective matched filter and therefore there is no benefit in measuring the square of the received signal. In this case the distance between any two signals can be described using the metric

$$d(\mathbf{y}_1, \mathbf{y}_2) = \int_0^\tau |\mathbf{y}_1(t) - \mathbf{y}_2(t)| dt.$$

Here, one can describe the space of receiver functions using the Banach space $\mathcal{L}^1([0, \tau], \mathbb{C}^m)$ with the norm defined by

$$\|\mathbf{y}\| := \int_0^\tau |\mathbf{y}(t)| dt.$$

This channel therefore is best described using a normed space as opposed to an inner product space to model the set of receiver signals.

Example 2.2: Consider a multi-carrier communication system that uses some form of amplitude or angle modulation to transmit information. Suppose that there are n carriers and that the vector $\phi = [\phi_1, \ldots, \phi_n]$ determines the modulating signal on each of the carriers. We can think of the modulating waveforms as the space of transmitter functions X[4].

If amplitude modulation is used then the vector ϕ determines the total power used for modulation. If the total power available for transmission is bounded then one might have an inequality of the form

$$\sum_{i=1}^n |\phi_i|^2 \leq P.$$

We can therefore describe the space of transmitter functions using the standard Euclidian space $\mathbb{R}^n$ with inner product

$$\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \mathbf{x}_1^T \mathbf{x}_2.$$

Now consider the case where angle modulation is used. In this case all the transmitted signals have the same power and the total power available for transmission places no restrictions on the space of transmitter functions. However, the space of transmitter functions can be subjected to other forms of constraints. For

---

[4]In this case we do not consider the actual signal on the transmitting antenna (i.e. carrier + modulation) to be the transmitter function. Cf. the discussion in Subsection I-A.

instance, if frequency modulation is used then the maximum frequency deviation used might be bounded by some number b to minimise co-channel interference (see e.g. [19, p. 110,513]). Similarly if phase modulation is used the maximum phase variation has to be less than $\pm\pi$. This bound may also depend on other practical considerations such as linearity of the modulator. In this case one might constrain the space of transmitter functions as

$$\sup_{1\leq i\leq n} |\phi_i| < b.$$

The space of transmitter functions of this channel is best described using the n-dimensional Banach space $\mathbb{R}^n_\infty$ with norm

$$\|\mathbf{x}\| = \sup_{1\leq i\leq n} |\mathbf{x}_i|.$$

Example 2.3: In this final example we examine spatial waveform channels (SWCs) [18]. In SWCs we assume that a current flows in a volume in space and generates an electromagnetic field in a receiver volume that is measured [10,15,16,18]. Such channels have been used to model MIMO systems previously [10,12,13,15,16,18]. If a current flows in a volume in space that has a finite conductivity, power is lost from the transmitting volume in two forms. Firstly, power is lost as heat and secondly power is radiated as electromagnetic energy. So the total power lost can be described using the set of equations

$$\begin{aligned} P_{total} &= P_{rad} + P_{lost} \\ P_{lost} &= \int_V \mathbf{J}^*(\mathbf{r})\mathbf{J}(\mathbf{r})d\mathbf{r} \\ P_{rad} &= \int_\Omega \mathbf{E}^*(\mathbf{r}) \times \mathbf{H}(\mathbf{r})d\Omega \end{aligned}$$

Here, V is some volume that contains the transmitting antennas, $\mathbf{J}$ is the current density in the volume V and Ω is some sufficiently smooth surface the interior of which contains V with $d\Omega$ denoting a surface area element. Also $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic fields generated by the current density $\mathbf{J}$ and $\cdot \times \cdot$ denotes the vector product in $\mathbb{R}^3$.

Because of the vector product in the last equation above, the total power lost defines a norm on the space of square-integrable functions that does not admit an inner-product structure [18]. The theory developed in this paper is used to calculate the degrees of freedom of such spatial waveform channels in [18].

III. Main Results

In this section we outline the main results of this paper. All the proofs of theorems are given in the Appendix.

A. Degrees of Freedom for Compact Operators

The definition of degrees of freedom at level ϵ for compact operators on normed spaces is identical to the finite dimensional counterpart (Definition 2.1) discussed in the previous section with $\mathbb{C}^n$ and $\mathbb{C}^m$ replaced by general normed spaces. The following theorem ensures that the definition makes sense even in the infinite dimensional setting.

Theorem 3.1: Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $T: X \to Y$ is a compact operator. Then for all $\epsilon > 0$ there exist[5] $N \in \mathbb{Z}_0^+$ and a set $\{\psi_i\}_{i=1}^N \subset Y$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1,\ldots,a_N} \left\| Tx - \sum_{i=1}^N a_i\psi_i \right\|_Y \leq \epsilon.$$

Note that for $N = 0$ the set $\{\psi_i\}_{i=1}^N$ is empty and the sum in the above expression is void. We will use the following definition for the number of degrees of freedom at level ϵ for compact operators on normed spaces.

Definition 3.1 (Degrees of freedom at level ϵ): Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $T: X \to Y$ is a compact operator. Then the number of degrees of freedom of T at level ϵ is the smallest $N \in \mathbb{Z}_0^+$ such that there exists a set of vectors $\{\psi_1,\ldots,\psi_N\} \subset Y$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1,\ldots,a_N} \left\| Tx - \sum_{i=1}^N a_i\psi_i \right\|_Y \leq \epsilon.$$

This definition has exactly the same interpretation as in the finite dimensional case: if there is some constraint $\|\cdot\|_X \leq 1$ on the space of source functions and if the receiver can only measure signals that satisfy $\|\cdot\|_Y > \epsilon$, then the number of degrees of freedom is the maximum number of linearly independent signals that the receiver can measure under these constraints.

This definition however is a descriptive one and can not be used to calculate the number of degrees of freedom for a given compact operator because the proof of Theorem 3.1 is not constructive. In the finite dimensional case we can calculate the degrees of freedom by calculating the singular values. However, as far as we are aware, there is no known generalisation of singular values for compact operators on arbitrary normed spaces[6]. In the following subsection we will

[5]$\mathbb{Z}$, $\mathbb{Z}_0^+$ and $\mathbb{Z}^+$ are respectively the sets of integers, non-negative integers and positive integers.

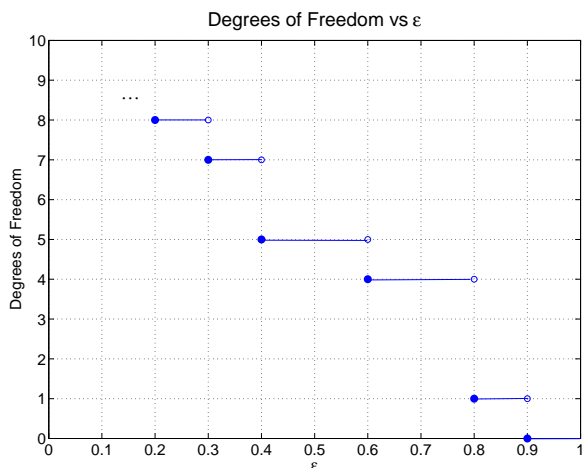[6]A generalisation to compact operators on Hilbert spaces is of course classical and well known.

Fig. 1. Degrees of Freedom of a Compact Operator

propose such a generalisation. In fact, we will use the degrees of freedom to generalise the concept of singular values and call such singular values DOF singular values. We will discuss the problem of computing degrees of freedom using DOF singular values in subsection III-D below.

Next, we establish some useful properties of degrees of freedom that will help motivate the definition of DOF singular values given in the next subsection.

Theorem 3.2: Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $T : X \to Y$ is a compact operator. Let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom of T at level ϵ. Then

1) $\mathcal{N}(\epsilon) = 0$ for all $\epsilon \geq \|T\|$.
2) Unless T is identically zero, there exists an $\epsilon_0 > 0$ such that $\mathcal{N}(\epsilon) \geq 1$ for all $0 < \epsilon < \epsilon_0$.
3) $\mathcal{N}(\epsilon)$ is a non-increasing, upper semicontinuous function of ϵ.
4) In any finite interval $(\epsilon_1, \epsilon_2) \subset \mathbb{R}$, with $0 < \epsilon_1 < \epsilon_2$, $\mathcal{N}(\epsilon)$ has only finitely many discontinuities, i.e. $\mathcal{N}(\epsilon)$ only takes finitely many non-negative integer values in any finite ϵ interval.

The following two examples show that as ϵ goes to zero, $\mathcal{N}(\epsilon)$ need not be finite nor go to infinity.

Example 3.1: Let l^1 be the Banach space of all real-valued sequences with finite l^1 norm and let $(e_1, e_2, \dots)$ be the standard Schauder basis for l^1. Define the operator $T : l^1 \to l^1$ by $e_n \mapsto e_1$ for all $n \in \mathbb{Z}^+$. This operator is well-defined and compact and $\mathcal{N}(\epsilon) \leq 1$ for all $\epsilon > 0$.

Example 3.2: Let l^1 and $(e_1, e_2, \dots)$ be defined as in the previous example. Define $T : l^1 \to l^1$ by $e_n \mapsto \frac{1}{n}e_n$ for all $n \in \mathbb{Z}^+$. Again T is well-defined and compact but $\lim_{\epsilon \to 0} \mathcal{N}(\epsilon) = \infty$.

Figure 1 shows a typical example of degrees of freedom at level ϵ for some compact operator that satisfies all the properties in the above theorem.

B. DOF Singular Values

We will identify the discontinuities in the number of degrees of freedom of T at level ϵ with the DOF singular values of T.

Definition 3.2 (DOF Singular Values): Suppose X and Y are normed spaces and $T : X \to Y$ is a compact operator. Let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom of T at level ϵ. Then ϵ_m is the mth DOF singular value of T if

$$\sup_{\epsilon > \epsilon_m} \mathcal{N}(\epsilon) = m - 1 \quad \text{and}$$
$$\inf_{\epsilon < \epsilon_m} \mathcal{N}(\epsilon) = M \geq m.$$

Further, if $m < M$ then for all $m < n \leq M$, $\epsilon_n := \epsilon_m$ is the nth DOF singular value of T.

Note that by Theorem 3.2, part 3 we have $\mathcal{N}(\epsilon_m) \leq m - 1$ with equality if (but not only if) ϵ_m is not a repeated DOF singular value.

Let the degrees of freedom of some operator T be as shown in Figure 1. Then the DOF singular values, ϵ_m, of T identify the jumps in the degrees of freedom. So, $\epsilon_1 = 0.9, \epsilon_2 = \epsilon_3 = \epsilon_4 = 0.8, \epsilon_5 = 0.6, \dots$

Another way of understanding the connection between the number of degrees of freedom at level ϵ and DOF singular values is as follows.

Proposition 3.3: Suppose X and Y are normed spaces and $T : X \to Y$ is a compact operator. Let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom of T at level ϵ. Then $\mathcal{N}(\epsilon)$ is equal to the number of DOF singular values that are greater than ϵ.

The intuition behind the definition for DOF singular values needs further clarification. In the finite dimensional case, if σ_p is the pth singular value of some operator $T : \mathbb{C}^n \to \mathbb{C}^m$, then there exist corresponding left and right singular vectors $v_p \in \mathbb{C}^n$ and $u_p \in \mathbb{C}^m$ such that v_p is of unit norm, $T v_p = u_p$ and the norm of u_p is σ_p. This is not necessarily true for arbitrary compact operators on normed spaces as the following example proves.

Example 3.3: Let l^1 and $(e_1, e_2, \dots)$ be defined as in Example 3.1. Define the operator $T : l^1 \to l^1$ by $T e_n = (1 - \frac{1}{n})e_1$ for all $n \in \mathbb{Z}^+$. Then T is well-defined and compact. Also, the number of degrees of freedom of T at level ϵ is

$$\mathcal{N}(\epsilon) = \begin{cases} 0 & \text{if } \epsilon \geq 1, \\ 1 & \text{if } \epsilon < 1. \end{cases}$$

So $\epsilon_1 = 1$. However, for any vector x in the unit sphere in l^1, $\|Tx\|_{l^1} < 1$.

The above example motivates the slightly more complicated statement in the following theorem which explains

the intuition behind the definition of DOF singular values.

Theorem 3.4: Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $T : X \to Y$ is a compact operator. Let ϵ_m be a DOF singular value of the operator T. Then for all $\theta > 0$ there exists a $\phi \in X$, $\|\phi\|_X = 1$, such that

$$\epsilon_m + \theta \geq \|T\phi\|_Y \geq \epsilon_m - \theta.$$

The above theorem shows how the DOF singular values are related to the traditionally accepted notion of singular values of compact operators on Hilbert spaces. In general, they are values the operator restricted to the unit sphere can get arbitrarily close to in norm. However, we still need to prove that in the special case of Hilbert spaces the new definition for DOF singular values agrees with the traditionally accepted definition for singular values.

Recall that if $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$ respectively and if $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator then the Hilbert adjoint operator for T is defined as the unique operator $T^* : \mathcal{H}_2 \to \mathcal{H}_1$ that satisfies [20, Sec. 3.9]

$$\langle Tx, y \rangle_{\mathcal{H}_2} = \langle x, T^*y \rangle_{\mathcal{H}_1}$$

for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$. The singular values of T are defined to be the square roots of the eigenvalues of the operator $T^*T : \mathcal{H}_1 \to \mathcal{H}_1$. We will refer to these as Hilbert space singular values to distinguish them from DOF singular values. Note that we always count repeated eigenvalues or (DOF) singular values repeatedly. The following two theorems establish the connection between Hilbert space singular values and the number of degrees of freedom at level ϵ. The theorems are important in their own right because they show that there are two other equivalent ways of calculating the degrees of freedom of a Hilbert space operator.

Theorem 3.5: Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator. Then for all $\epsilon > 0$ there exist an $N \in \mathbb{Z}_0^+$ and a set of N mutually orthogonal vectors $\{\phi_i\}_{i=1}^N \subset \mathcal{H}_1$ such that if

$$x \in \mathcal{H}_1, \ \|x\|_{\mathcal{H}_1} \leq 1 \text{ and } \langle x, \phi_i \rangle_{\mathcal{H}_1} = 0$$

then

$$\|Tx\|_{\mathcal{H}_2} \leq \epsilon.$$

Moreover, the smallest N that satisfies the above condition for a given ϵ is equal to the number of Hilbert space singular values of T that are greater than ϵ.

Theorem 3.6: Suppose that $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator. Then the number of degrees of freedom at level ϵ is equal to the number of Hilbert space singular values of T that are greater than ϵ.

As a corollary of Theorem 3.6 we get the following result.

Corollary 3.1: Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator. Suppose $\{\epsilon_m\}$ are the DOF singular values of T and $\{\sigma_m\}$ are the possibly repeated Hilbert space singular values of T written in non-increasing order. Then

$$\sigma_m = \epsilon_m$$

for all $m \in \mathbb{Z}^+$.

This corollary, reassuringly, proves that the DOF singular values are in fact generalisations of the traditionally accepted notion of Hilbert space singular values. We will therefore use the terms DOF singular values and singular values interchangeably unless specified otherwise for the remainder of this paper.

In Hilbert spaces we have three characterizations for degrees of freedom: 1) as in Definition 3.2, 2) as in Theorem 3.6 in terms of singular values and 3) as in Theorem 3.5 in terms of mutually orthogonal functions in the domain.

We have used the first two characterisations in the generalisation to normed spaces. However, the final characterisation is more difficult to generalise. It would be extremely useful to generalise the final characterisation because, for the Hilbert space case, the functions ϕ_i in Theorem 3.5 are in some sense the best functions to transmit (see e.g. [14]). One could possibly replace the mutual orthogonality by almost orthogonality using the Riesz lemma (see e.g. [20, pp. 78]).

Lemma 3.7 (Riesz's lemma): Let Y and Z be subspaces of a normed space X and suppose that Y is closed and is a proper subspace of Z. Then for all $\theta \in (0, 1)$ there exists a $z \in Z$, $\|z\| = 1$, such that for all $y \in Y$

$$\|y - z\| \geq \theta.$$

The following conjecture is still an open question.

Conjecture 3.1: Let X and Y be reflexive Banach spaces and let $T : X \to Y$ be compact. Given any $\epsilon > 0$ and some $\theta \in (0, 1)$, there exists a finite set of vectors $\{\phi_i\}_{i=1}^N \subset X$ such that for all $x \in X$, $\|x\|_X \leq 1$,

$$\inf_{a_1,\ldots,a_N} \left\| x - \sum_{i=1}^N a_i \phi_i \right\|_X \geq \theta \tag{1}$$

implies

$$\|Tx\|_Y \leq \epsilon.$$

Comparing with Theorem 3.5, condition (1) is analogous to requiring that x be orthogonal to all the ϕ_i. The conjecture is definitely not true unless we impose additional

conditions such as reflexivity on X and/or Y as the next example proves.

Example 3.4: Let l^1, $(e_1, e_2, \ldots)$ and the compact operator $T : l^1 \to l^1$ be defined as in Example 3.1. Now let $\epsilon < 1$. For any $x = \sum_n \alpha_n e_n \in l^1$, if $\|x\| = 1$ and if $\alpha_n \geq 0$ for all n then $\|Tx\| = \|x\| = 1 > \epsilon$. Hence no finite set of vectors can satisfy the conditions in the conjecture.

In the following subsection, we use degrees of freedom and DOF singular values to define the essential dimension of a communication channel.

C. Essential Dimension for Compact Operators

The definition for degrees of freedom given in Section III-A depends on the arbitrarily chosen number ϵ and therefore this definition does not give a unique number for a given channel. The physical intuition behind choosing this arbitrary small number ϵ is nicely explained in Xu and Janaswamy [12]. In that paper $\epsilon = \sigma^2$ denotes the noise level at the receiver and the authors state that the number of degrees of freedom fundamentally depends on this noise level.

However, in several important cases the number of degrees of freedom of a channel is essentially independent of this arbitrarily chosen positive number [4,9,11,13,14, 16]. This is due to the fact that in these cases the singular values of the channel operator show a step like behavior. Therefore, for a range of values of ϵ, the number of degrees of freedom at level ϵ is constant. This leads us to the concept of essential dimensionality[7] which is only a function of the channel and not the arbitrarily chosen positive number ϵ. Some of the properties that one might require from the essential dimension of a channel operator are:

1) It must be uniquely defined for a given operator T.
2) The definition must be applicable to a general class of operators under consideration so that comparisons can be made between different operators.[8]
3) It must in some sense *represent* the number of degrees of freedom at level ϵ.

The last requirement above needs further clarification. Obviously the essential dimension of T can not in general be equal to the number of degrees of freedom at level ϵ because the latter is a function of ϵ. However, if

the singular values of T plotted in non-increasing order change suddenly from being large to being small then the number of degrees of freedom at the "knee" in this graph is the essential dimension of T. The following definition for the essential dimension tries to identify this "knee" in the set of DOF singular values.

Each level ϵ defines a unique number of degrees of freedom $\mathcal{N}(\epsilon)$ for a given compact operator T. So for each positive integer $n \in \mathbb{Z}^+$ we can calculate $E(n) = \mu(\{\epsilon : n = \mathcal{N}(\epsilon)\})$. Here $\mu(\cdot)$ is the Lebesgue measure. The function $E(n)$ is well defined because of the properties of DOF singular values discussed in Theorem 3.2. We can now define the essential dimension of T as follows.

Definition 3.3: The essential dimension of a compact operator T is

$$\text{EssDim}(T) = \text{argmax}\{E(n) : n \in \mathbb{Z}^+\}$$

where $E(n)$ is defined as above. If argmax above is not unique then choose the smallest n of all the n that maximise $E(n)$ as the essential dimension.

In this definition we are simply calculating the maximum range of values of the arbitrarily chosen ϵ over which the number of degrees of freedom of an operator does not change. It uniquely determines the essential dimension of all compact operators. Further, it is equal to the number of degrees of freedom at level ϵ for the maximum range of ϵ. Choosing this value for the number of degrees of freedom in order to model communication systems has the big advantage that it is independent of the noise level at the receiver. Further, if for a given noise level the number of degrees of freedom is greater than the essential dimension then one can be sure that even if the noise level varies by a significant amount the number of degrees of freedom will always be greater than the essential dimension.

The essential dimension of T is the smallest number of DOF singular values of T after which the change in two consecutive singular values is a maximum. One could also look at how the DOF singular values are changing gradually and the above definition is a special case of the following notion of essential dimension of order n, namely the case where $n = 1$.

Definition 3.4: Let X, Y be normed spaces and let $T : X \to Y$ be a compact operator. Let $\{\epsilon_m\}$ be the set of DOF singular values of T numbered in non-increasing order. Then define the essential dimension of T of order n to be N if n is even and

$$\epsilon_{N-n/2} - \epsilon_{N+n/2} \geq \epsilon_{M-n/2} - \epsilon_{M+n/2}$$

for all $M \neq N$. If there are several N that satisfy the above condition then choose the smallest such N. If n

---

[7]Note that the term "essential dimension" has been used instead of "degrees of freedom" in several papers. As far as we are aware, this is the first time an explicit distinction is being made between the two terms.

[8]This requirement is in contrast to the essential dimension definition in [17] that is only applicable to the time-bandwidth problem.
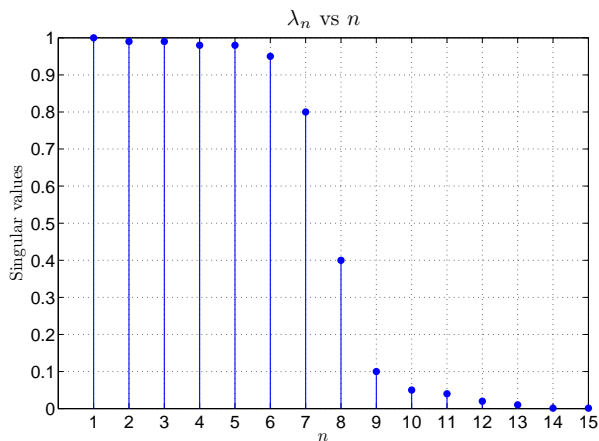
Fig. 2. Singular values of an Operator

is odd then choose the smallest N that satisfies

$$\epsilon_{N-(n-1)/2} - \epsilon_{N+(n+1)/2} \geq \epsilon_{M-(n-1)/2} - \epsilon_{M+(n+1)/2}$$

for all $M \neq N$.

A simple example illustrates the concepts of essential dimensionality and degrees of freedom.

Example 3.5: Figure 2 shows the singular values of some operator T. For this operator the number of degrees of freedom at level 0.75 is 7 and at level 0.1 is 8.

The essential dimension of the channel is 7. This is because for $\epsilon \in [0.4, 0.8)$, $\mathcal{N}(\epsilon) = 7$. Therefore $E(7) = 0.4$ which is greater than $E(n)$ for all $n \neq 7$. The essential dimension of order 2 is 8 because $\epsilon_7 - \epsilon_9 = 0.7$ which is greater than $\epsilon_{M-1} - \epsilon_{M+1}$ for all $M \neq 8$.

D. Computing DOF singular values

Both, degrees of freedom and essential dimension for a communication channel, can be evaluated if the DOF singular values of the operator T describing the channel are known. However, no known method exists for computing these singular values for general compact operators. In this section, we develop a numerical method, based on finite dimensional approximations, that could be used to calculate DOF singular values.

Theorem 3.8: Suppose X and Y are normed spaces and $T : X \to Y$ is a compact operator. Also suppose that X has a complete Schauder basis $\{\phi_1, \phi_2, \ldots\}$ and let $S_n = \text{span}\{\phi_1, \ldots, \phi_n\}$. Let $T_n = T|_{S_n} : S_n \to Y$, $n \in \mathbb{Z}^+$. If ϵ_m, the mth singular value of T, exists then for n large enough $\epsilon_{m,n}$, the mth singular value of T_n, will exist and

$$\lim_{n \to \infty} \epsilon_{m,n} = \epsilon_m.$$

If $\epsilon_{m,n}$ exists then it is a lower bound for ϵ_m.

The theorem shows that if the domain of the operator has some complete Schauder basis then we can calculate

the DOF singular values of the operator restricted to finite dimensional subspaces and as the subspaces get bigger we will approach the singular values of the original operator. Moreover, the theorem also proves that the singular values of the finite dimensional operators provide lower bounds for the original DOF singular values. We, however, still need a practical method of calculating the singular values of linear operators defined on finite dimensional normed spaces.

Let X, Y be two finite dimensional Banach spaces and let $T : X \to Y$ be a linear operator. Suppose $\epsilon_1, \ldots, \epsilon_n$ are the DOF singular values of T and denote $B_1 = \{x \in X : \|x\|_X \leq 1\}$. We know that for all $\epsilon \geq \epsilon_{p+1}$, $\mathcal{N}(\epsilon) \leq p$. Hence for each $\epsilon \geq \epsilon_{p+1}$ there exists a set $\{\psi_i\}_{i=1}^p \subset Y$ such that

$$\sup_{x \in B_1} \inf_{a_1, \ldots, a_p} \left\| Tx - \sum_{i=1}^p a_i \psi_i \right\|_Y \leq \epsilon.$$

Let $\Psi_{p,\epsilon}$ denote the set of all sets $\{\psi_i : \|\psi_i\|_Y \leq 1\}_{i=1}^p \subset Y$ that satisfy the above inequality for a given $\epsilon \geq \epsilon_{p+1}$ and let

$$\Psi_p = \bigcup_{\epsilon \geq \epsilon_{p+1}} \Psi_{p,\epsilon}.$$

With this notation we can now prove that the DOF singular values of a linear operator defined on a finite dimensional normed space can be expressed as the solution of an optimisation problem.

Theorem 3.9: Let X, Y be two finite dimensional Banach spaces and let $T : X \to Y$ be a linear operator. Also let B_1 be the closed unit ball in X and suppose Ψ_p is defined as explained above. Then

$$\sup_{x \in B_1} \|Tx\|_Y = \epsilon_1$$

and for all $p \in \mathbb{Z}^+$

$$\inf_{\{\psi_i\}_{i=1}^p \in \Psi_p} \sup_{x \in B_1} \inf_{a_1, \ldots, a_p} \left\| Tx - \sum_{i=1}^p a_i \psi_i \right\|_Y = \epsilon_{p+1}.$$

Given the "correct" set of functions ψ_i, the above theorem characterises the singular values in terms of a maximisation problem over a finite dimensional domain. It is however difficult to check whether a given set of functions $\{\psi_i\}_{i=1}^p$ is an element of Ψ_p. We therefore propose the following algorithm to calculate bounds on the DOF singular values.

Suppose X, Y, $T : X \to Y$, $\epsilon_1, \ldots, \epsilon_n$ and B_1 are defined as in Theorem 3.9. Let

$$\epsilon_1' = \sup_{x \in B_1} \|Tx\|_Y.$$

Because $B_1 \subset X$ is a compact set and $\|\cdot\|_Y$ and T are continuous, there exists an $x_1 \in B_1$ such that $\|Tx_1\|_Y = \epsilon_1'$. Choose $\psi_1 = Tx_1$.

Now suppose $\psi_1, \ldots, \psi_p$ have been chosen. Then let

$$\epsilon'_{p+1} = \sup_{x \in B_1} \inf_{a_1, \ldots, a_p} \left\| Tx - \sum_{i=1}^{p} a_i \psi_i \right\|_Y. \qquad (2)$$

Again, because $B_1 \subset X$ is a compact set and $\| \cdot \|_Y$ and T are continuous, there exists an $x_{p+1} \in B_1$ such that x_{p+1} attains the maximum in the above equation. Choose $\psi_{p+1} = Tx_{p+1}$. Comparing with Theorem 3.9 we note that ϵ'_{p+1} is an upper bound for ϵ_{p+1}. It is an open question as to whether $\epsilon'_{p+1} = \epsilon_{p+1}$.

In this algorithm, instead of searching over all possible sets in Ψ_p we select a special set that is in some sense (it consists of images of the $x \in B_1$ that attain the maximum in equation (2)) the best possible set to use. This choice is essential because otherwise the calculation of DOF singular values becomes too cumbersome (one needs to find the set Ψ_p before calculating ϵ_{p+1}). Note however, that the above algorithm gives the correct value for ϵ_1.

The theory presented here has been used to compute the DOF singular values and degrees of freedom in spatial waveform channels of the type discussed in Example 2.3. The results of these computations are presented in Somaraju [18]. Due to space constraints, these results are not further discussed in this paper.

E. Non-compactness of channel operators

Throughout this paper we have exclusively dealt with channels that can be modeled using compact operators. We have done so because of the following result.

Theorem 3.10: (Converse to Theorem 3.1) Suppose X and Y are normed spaces with norms $\| \cdot \|_X$ and $\| \cdot \|_Y$, respectively, and $T : X \to Y$ is a bounded linear operator. If for all $\epsilon > 0$ there exist $N \in \mathbb{Z}_0^+$ and a set $\{\psi_i\}_{i=1}^{N} \subset Y$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1, \ldots, a_N} \left\| Tx - \sum_{i=1}^{N} a_i \psi_i \right\|_Y \leq \epsilon$$

then T is compact.
So any bounded channel operator with finitely many sub-channels must be compact. Indeed, if one can find a channel that is not described by a compact operator, then it will have infinitely many sub-channels and will therefore have infinite capacity. Also, if the channel is described by an operator that is linear but unbounded then there will obviously exist sub-channels over which arbitrarily large gains can be obtained.[9]

---

[9]It could hence be argued that non-compact channel operators are unphysical, however, we will leave it to the reader to make this judgement.

IV. CONCLUSION

In this paper we assume that a communication channel can be modeled by a normed space X of transmitter functions that a transmitter can generate, a normed space Y of functions that a receiver can measure and an operator $T : X \to Y$ that maps the transmitter functions to functions measured by the receiver. We then introduce the concepts of degrees of freedom at level ϵ, essential dimension and DOF singular values of such channel operators in the case where they are compact. One can give a physical interpretation for degrees of freedom as follows: if there is some constraint $\| \cdot \|_X \leq 1$ on the space of source functions and if the receiver can only measure signals that satisfy $\| \cdot \|_Y > \epsilon$ then the number of degrees of freedom is the number of linearly independent signals that the receiver can measure under the given constraints. If the degrees of freedom are largely independent of the level ϵ then it makes sense to talk about the essential dimension of the channel. The essential dimension of the channel is the smallest number of degrees of freedom of the channel that is the same for the largest range of levels ϵ. We show how one can use the number of degrees of freedom at level ϵ to generalise the Hilbert space concept of singular values to arbitrary normed spaces. We also provide a simple algorithm that can be used to approximately calculate these DOF singular values. Finally, we prove that if the operator describing the channel is not compact then it must either have infinite gain or have an infinite number of degrees of freedom. The general theory developed in this paper is applied to spatial waveform channels in Somaraju [18].

APPENDIX

Proofs of Theorems: *Theorem 3.1.* Suppose X and Y are normed spaces with norms $\| \cdot \|_X$ and $\| \cdot \|_Y$, respectively, and $T : X \to Y$ is a compact operator. Then for all $\epsilon > 0$ there exist $N \in \mathbb{Z}_0^+$ and a set $\{\psi_i\}_{i=1}^{N} \subset Y$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1, \ldots, a_N} \left\| Tx - \sum_{i=1}^{N} a_i \psi_i \right\|_Y \leq \epsilon. \qquad (3)$$

Proof: The proof is by contradiction. Let $\epsilon > 0$ be given. Suppose no such N exists.

Let $x_1 \in \overline{B}_{1,X}(0)$ be any vector. Choose $\psi_1 = Tx_1$. Suppose that $\{x_1, \ldots, x_N\}$ and $\{\psi_1, \ldots, \psi_N\}$ have been chosen. Then, by our assumption, there exists an $x_{N+1} \in \overline{B}_{1,X}(0)$ such that

$$\inf_{a_1, \ldots, a_N} \left\| Tx_{N+1} - \sum_{i=1}^{N} a_i \psi_i \right\|_Y > \epsilon. \qquad (4)$$

Choose $\psi_{N+1} = Tx_{N+1}$. By induction, for $M \leq N$ we have

$$\|Tx_{N+1} - Tx_M\|_Y > \epsilon.$$

This follows from (4) by setting $a_i = 0$, $i \leq N$, $i \neq M$, and $a_M = 1$. Therefore, using the Cauchy criterion, the sequence $\{Tx_n\}_{n=1}^\infty$ chosen by induction cannot have a convergent subsequence. This is the required contradiction because $\{x_n\}_{n=1}^\infty$ is a bounded sequence and T is compact. ∎

Theorem 3.2. Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $T : X \to Y$ is a compact operator. Let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom of T at level ϵ. Then

1) $\mathcal{N}(\epsilon) = 0$ for all $\epsilon \geq \|T\|$.
2) Unless T is identically zero, there exists an $\epsilon_0 > 0$ such that $\mathcal{N}(\epsilon) \geq 1$ for all $0 < \epsilon < \epsilon_0$.
3) $\mathcal{N}(\epsilon)$ is a non-increasing, upper semicontinuous function of ϵ.
4) In any finite interval $(\epsilon_1, \epsilon_2) \subset \mathbb{R}$, with $0 < \epsilon_1 < \epsilon_2$, $\mathcal{N}(\epsilon)$ has only finitely many discontinuities, i.e. $\mathcal{N}(\epsilon)$ only takes finitely many non-negative integer values in any finite ϵ interval.

Proof:

1) Because T is compact it is bounded, and therefore $\|T\| < \infty$. Suppose $\epsilon \geq \|T\|$ then $\|Tx\|_Y \leq \|T\| \leq \epsilon$ for all $x \in \overline{B}_{1,X}(0)$. Therefore $\mathcal{N}(\epsilon) = 0$.
2) If $\|T\| > 0$ there exists an $x \in X$, $\|x\|_X \leq 1$ such that $\|Tx\|_Y > 0$. Set $\epsilon_0 := \|Tx\|_Y$. Then for all $0 < \epsilon < \epsilon_0$, $\mathcal{N}(\epsilon) \geq 1$.
3) Suppose $0 < \epsilon_1 < \epsilon_2$. Then there exist functions $\psi_1, \ldots, \psi_{\mathcal{N}(\epsilon_1)}$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1,\ldots,a_{\mathcal{N}(\epsilon_1)}} \left\| Tx - \sum_{i=1}^{\mathcal{N}(\epsilon_1)} a_i \psi_i \right\|_Y < \epsilon_1 < \epsilon_2$$

Therefore $\mathcal{N}(\epsilon_2) \leq \mathcal{N}(\epsilon_1)$ from the definition of the number of degrees of freedom at level ϵ, i.e. $\mathcal{N}(\epsilon)$ is non-increasing. In particular we have

$$\lim_{\epsilon \searrow \epsilon_1} \mathcal{N}(\epsilon) \leq \mathcal{N}(\epsilon_1).$$

Assume that the above inequality is strict. Then there exists an $N \in \mathbb{Z}_0^+$, $N < \mathcal{N}(\epsilon_1)$, and for all $\theta > 0$ there exists a set $\{\psi_i^\theta\}_{i=1}^N \subset Y$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1,\ldots,a_N} \left\| Tx - \sum_{i=1}^N a_i \psi_i^\theta \right\|_Y \leq \epsilon_1 + \theta. \quad (5)$$

On the other hand, since $\mathcal{N}(\epsilon_1) > N$, for all sets $\{\psi_i\}_{i=1}^N \subset Y$ there exists an $x \in \overline{B}_{1,X}(0)$ such

that

$$\mu := \inf_{a_1,\ldots,a_N} \left\| Tx - \sum_{i=1}^N a_i \psi_i \right\|_Y > \epsilon_1. \quad (6)$$

But (5) contradicts (6) for $\theta := \frac{1}{2}(\mu - \epsilon_1)$. Hence $\lim_{\epsilon \searrow \epsilon_1} \mathcal{N}(\epsilon) = \mathcal{N}(\epsilon_1)$ and $\mathcal{N}(\epsilon)$ is upper semi-continuous.

4) This follows from Parts 1 and 3. ∎

Proposition 3.3. Suppose X and Y are normed spaces and $T : X \to Y$ is a compact operator. Let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom of T at level ϵ. Then $\mathcal{N}(\epsilon)$ is equal to the number of DOF singular values that are greater than ϵ.

Proof: This follows from careful counting of the numbers of degrees of freedom at level ϵ including repeated counting according to the height of any occurring "jumps". ∎

Theorem 3.4. Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $T : X \to Y$ is a compact operator. Let ϵ_m be a DOF singular value of the operator T. Then for all $\theta > 0$ there exists a $\phi \in X$, $\|\phi\|_X = 1$, such that

$$\epsilon_m + \theta \geq \|T\phi\|_Y \geq \epsilon_m - \theta.$$

Proof: The proof is by contradiction. Assume that there exists a $\theta > 0$ such that for all $\phi \in X$, $\|\phi\|_X = 1$, we have $\|T\phi\|_Y \notin [\epsilon_m - \theta, \epsilon_m + \theta]$. Let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom at level ϵ of the operator T. From the definition of degrees of freedom at level ϵ we have

$$\mathcal{N}(\epsilon_m + \theta) \leq m - 1, \quad (7)$$
$$\mathcal{N}(\epsilon_m - \theta) \geq m. \quad (8)$$

By (7), there exist vectors $\psi_1, \ldots, \psi_{m-1} \in Y$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1,\ldots,a_{m-1}} \left\| Tx - \sum_{i=1}^{m-1} a_i \psi_i \right\| \leq \epsilon_m + \theta.$$

By our assumption on $\|T\phi\|_Y$,

$$\inf_{a_1,\ldots,a_{m-1}} \left\| T\phi - \sum_{i=1}^{m-1} a_i \psi_i \right\| \leq \epsilon_m - \theta.$$

This follows from consideration of the case $a_1 = \cdots = a_{m-1} = 0$. Hence $\mathcal{N}(\epsilon_m - \theta) \leq m - 1$ since scaling ϕ to non-unit norm is equivalent to scaling all the a_i. This contradicts inequality (8). Therefore there exists a ϕ that satisfies the conditions of the theorem. ∎

Theorem 3.5. Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator. Then for all

$\epsilon > 0$ there exist an $N \in \mathbb{Z}_0^+$ and a set of N mutually orthogonal vectors $\{\phi_i\}_{i=1}^N \subset \mathcal{H}_1$ such that if

$$x \in \mathcal{H}_1, \ \|x\|_{\mathcal{H}_1} \le 1 \text{ and } \langle x, \phi_i \rangle_{\mathcal{H}_1} = 0$$

then

$$\|Tx\|_{\mathcal{H}_2} \le \epsilon.$$

Moreover, the smallest N that satisfies the above condition for a given ϵ is equal to the number of Hilbert space singular values of T that are greater than ϵ.

Proof: We first prove that such an N is given by the number of Hilbert space singular values of T that are greater than ϵ and then prove that this is the smallest such N.

Let $\epsilon > 0$ be given. Because T is compact, we can use the singular value decomposition theorem which says [21, p. 261]

$$T \cdot = \sum_i \sigma_i \langle \cdot, \phi_i \rangle_{\mathcal{H}_1} \psi_i. \tag{9}$$

Here, σ_i, ϕ_i and ψ_i with $i \in \mathbb{Z}^+$ are the Hilbert space singular values and left and right singular vectors of T, respectively. We assume w.l.o.g. that the Hilbert space singular values are ordered in non-increasing order. We denote by $N_1 \in \mathbb{Z}^+$ the number of Hilbert space singular values of T that are greater than ϵ, i.e. $\sigma_i > \epsilon$ if and only if $i \le N_1$.

Now, if x is orthogonal to $\phi_i, i = 1, \ldots, N_1$ and if $\|x\|_{\mathcal{H}_1} \le 1$ then from equation (9)

$$
\begin{aligned}
\|Tx\|_{\mathcal{H}_2}^2 &= \sum_{i=1}^{\infty} \sigma_i^2 |\langle x, \phi_i \rangle_{\mathcal{H}_1}|^2 \|\psi_i\|_{\mathcal{H}_2}^2 \\
&\le \epsilon^2 \sum_{i=N_1+1}^{\infty} |\langle x, \phi_i \rangle_{\mathcal{H}_1}|^2 \\
&\le \epsilon^2.
\end{aligned}
$$

For $N < N_1$, the linear span of any set $\{\varphi_i\}_{i=1}^N \subset \mathcal{H}_1$ has a non-trivial orthogonal complement in the span of $\{\phi_i\}_{i=1}^{N_1}$. Any vector x in this complement with $\|x\|_{\mathcal{H}_1} = 1$ fullfills the conditions of the theorem but $\|Tx\|_{\mathcal{H}_2} > \epsilon$ by equation (9). ∎

Theorem 3.6. Suppose that $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator. Then the number of degrees of freedom at level ϵ is equal to the number of Hilbert space singular values of T that are greater than ϵ.

Proof: As in the prove of the previous theorem, let $N_1 \in \mathbb{Z}^+$ denote the number of Hilbert space singular values of T that are greater than ϵ. Let σ_i, ϕ_i and ψ_i with $i \in \mathbb{Z}^+$ denote the Hilbert space singular values in non-increasing order and the left and right singular vectors

of T, respectively. Let $N_2 \in \mathbb{Z}^+$ denote the number of degrees of freedom of T at level ϵ.

We first prove that $N_1 \ge N_2$. If x is in the unit ball in $\mathcal{H}_1$ then we can write $x = \sum_{i=1}^{\infty} \langle x, \phi_i \rangle_{\mathcal{H}_1} \phi_i + x_r$. Here x_r is the remainder term that is orthogonal to all the ϕ_i. From equation (9) and $\sigma_i \le \epsilon$ for $i > N_1$ it follows that

$$\left\| Tx - \sum_{i=1}^{N_1} \sigma_i \langle x, \phi_i \rangle_{\mathcal{H}_1} \psi_i \right\|_{\mathcal{H}_2} \le \epsilon$$

and hence $N_1 \ge N_2$ by the definition of the number of degrees of freedom at level ϵ (set $a_i = \sigma_i \langle x, \phi_i \rangle_{\mathcal{H}_1}$ in that definition).

To prove that $N_1 \le N_2$ assume that $N_1 > N_2$ to arrive at a contradiction. Then there exists a set $\{\psi_i'\}_{i=1}^{N_2} \subset \mathcal{H}_2$ such that

$$\inf_{a_1, \ldots, a_{N_2}} \left\| Tx - \sum_{i=1}^{N_2} a_i \psi_i' \right\|_{\mathcal{H}_2} \le \epsilon$$

for all $x \in \mathcal{H}_1$, $\|x\|_{\mathcal{H}_1} \le 1$. Because we assume that $N_1 > N_2$, there exists a $y \in \text{span}\{\psi_1, \ldots, \psi_{N_1}\}$ which is orthogonal to all the ψ_i'. Let $y = \sum_{i=1}^{N_1} b_i \psi_i$. Then $y = Tx$ where $x = \sum_{i=1}^{N_1} \frac{b_i}{\sigma_i} \phi_i$ by equation (9). We can assume w.l.o.g. that the b_i are normalised so that $\|x\|_{\mathcal{H}_1} = 1$. If this is done then

$$
\begin{aligned}
\inf_{a_1, \ldots, a_{N_2}} \left\| Tx - \sum_{i=1}^{N_2} a_i \psi_i' \right\|_{\mathcal{H}_2}^2 &= \|y\|_{\mathcal{H}_2}^2 \tag{10} \\
&= \sum_{i=1}^{N_1} b_i^2 \tag{11} \\
&> \sum_{i=1}^{N_1} \frac{b_i^2}{\sigma_i^2} \epsilon^2 \tag{12} \\
&= \epsilon^2. \tag{13}
\end{aligned}
$$

In the above we get equation (10) from the fact that y is orthogonal to all the ψ_i', inequality (12) from $\sigma_i > \epsilon$ for $i \le N_1$ and equation (13) from $\|x\|_{\mathcal{H}_1} = 1$. The inequality (10)–(13) is the required contradiction. This proves that $N_1 \le N_2$ and hence $N_1 = N_2$. ∎

Corollary 3.1. Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator. Suppose $\{\epsilon_m\}$ are the DOF singular values of T and $\{\sigma_m\}$ are the possibly repeated Hilbert space singular values of T written in non-increasing order. Then

$$\sigma_m = \epsilon_m$$

for all $m \in \mathbb{Z}^+$.

Proof: This follows immediately from Theorem 3.6 and Proposition 3.3 by a simple counting argument. ∎

Theorem 3.8. Suppose X and Y are normed spaces and $T : X \to Y$ is a compact operator. Also suppose that X has a complete Schauder basis $\{\phi_1, \phi_2, \ldots\}$ and let $S_n = \text{span}\{\phi_1, \ldots, \phi_n\}$. Let $T_n = T|_{S_n} : S_n \to Y$, $n \in \mathbb{Z}^+$. If ϵ_m, the mth singular value of T, exists then for n large enough $\epsilon_{m,n}$, the mth singular value of T_n, will exist and

$$\lim_{n \to \infty} \epsilon_{m,n} = \epsilon_m.$$

If $\epsilon_{m,n}$ exists then it is a lower bound for ϵ_m.

Proof Outline: The crux of the argument used to prove the theorem is as follows. Assume $\epsilon > 0$ is given and let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom at level ϵ for the operator T. By definition there exist functions $\{\psi_1, \ldots, \psi_{\mathcal{N}(\epsilon)}\} \subset Y$ such that for all $x \in X$, $\|x\|_X \leq 1$, Tx can be approximated to level ϵ by a linear combination of the ψ_i and further, no set of functions $\{\psi'_1, \ldots, \psi'_N\} \subset Y$ can approximate all the Tx if $N < \mathcal{N}(\epsilon)$. Equivalently, there is a vector in the closed unit ball in X whose image under T can be approximated by a vector in $\text{span}\{\psi_1, \ldots, \psi_{\mathcal{N}(\epsilon)}\}$ but not by any vector in $\text{span}\{\psi'_1, \ldots, \psi'_N\}$.

So we take the inverse image of an ϵ-net of points in $\text{span}\{\psi_1, \ldots, \psi_{\mathcal{N}(\epsilon)}\}$ and choose n large enough so that all the inverse images are close to S_n. We can do this because the ϕ_i form a complete Schauder basis for X. We then show that there exists a vector in S_n such that its image under T cannot be approximated by a linear combination of $\psi'_1, \ldots, \psi'_N$ for $N < \mathcal{N}(\epsilon)$. This will prove that the number of degrees of freedom at level ϵ of T_n approaches that of T and consequently so do the singular values. The details are as follows.

Proof: We will prove this theorem in two parts. Assume that ϵ_m exists. In part a) we will prove that if $\epsilon_{m,N}$ exists for some $N \in \mathbb{Z}^+$ then $\epsilon_{m,n}$ exists for all $n > N$, and the $\epsilon_{m,n}$ form a non-decreasing sequence indexed by n that is bounded from above by ϵ_m. In part b) we prove by contradiction that $\epsilon_{m,n}$ exists for some $n \in \mathbb{Z}^+$ and that $\epsilon_{m,n}$ must converge to ϵ_m.

We will use the following notation in the proof:

$$\text{span}_\epsilon\{\psi_1, \ldots, \psi_N\} =$$

$$\left\{ y \in Y : \inf_{a_1, \ldots, a_N} \left\| y - \sum_{i=1}^{N} a_i \psi_i \right\|_Y \leq \epsilon \right\}$$

and $B_r = \{x \in X : \|x\|_X \leq r\}$.

Part a: Let T and T_n be defined as in the theorem and let $\mathcal{N}(\epsilon)$ and $\mathcal{N}_n(\epsilon)$ be the numbers of degrees of freedom at level ϵ of T and T_n, respectively. Assume that ϵ_{m,n_1} exists and let $n_2 > n_1$.

Then for all sets $\{\psi_1, \ldots, \psi_{\mathcal{N}_{n_1}(\epsilon)-1}\} \subset Y$ there is a $\xi \in S_{n_1} \cap B_1$ such that

$$T_{n_1}\xi = T\xi \notin \text{span}_\epsilon\{\psi_1, \ldots, \psi_{\mathcal{N}_{n_1}(\epsilon)-1}\}.$$

Because $S_{n_1} \subset S_{n_2}$ we have $\xi \in S_{n_2} \cap B_1$ and

$$T_{n_2}\xi = T\xi \notin \text{span}_\epsilon\{\psi_1, \ldots, \psi_{\mathcal{N}_{n_1}(\epsilon)-1}\}.$$

Therefore for all $\epsilon > 0$

$$\mathcal{N}_{n_2}(\epsilon) \geq \mathcal{N}_{n_1}(\epsilon). \tag{14}$$

Because

$$\inf_{\epsilon < \epsilon_{m,n_1}} \mathcal{N}_{n_1}(\epsilon) \geq m \tag{15}$$

we have $\mathcal{N}_{n_2}(\epsilon) \geq \mathcal{N}_{n_1}(\epsilon) \geq m$ for $\epsilon < \epsilon_{m,n_1}$. Hence ϵ_{m,n_2} must exist.

From the definition of DOF singular values we have inequality (15) and

$$\sup_{\epsilon > \epsilon_{m,n_2}} \mathcal{N}_{n_2}(\epsilon) \leq m - 1$$

If $\epsilon_{m,n_1} > \epsilon_{m,n_2}$ then there exists an ϵ' such that $\epsilon_{m,n_1} > \epsilon' > \epsilon_{m,n_2}$. Therefore,

$$\mathcal{N}_{n_1}(\epsilon') \geq m > m - 1 \geq \mathcal{N}_{n_2}(\epsilon').$$

This contradicts inequality (14). Therefore $\epsilon_{m,n_1} \leq \epsilon_{m,n_2}$.

The same line of arguments as above can be used to show that if both ϵ_m and $\epsilon_{m,n}$ exist then $\epsilon_{m,n} \leq \epsilon_m$. Recall that we have assumed at the beginning that ϵ_m exists. Therefore, if $\epsilon_{m,N}$ exists for some $N \in \mathbb{Z}^+$ then $\epsilon_{m,n}$ is a non-decreasing sequence in $n \geq N$ that is bounded from above by ϵ_m.

Part b: By part a), if $\epsilon_{m,n}$ exists for $n \geq n_1$ then, because $\epsilon_{m,n}$ is a bounded monotonic sequence in n it must converge to some $\epsilon'_m \leq \epsilon_m$.

Now there are two situations to consider. Firstly, $\epsilon_{m,n}$ might not exist for any $n \in \mathbb{Z}^+$. Secondly, $\epsilon_{m,n}$ might exist for some n but the limit ϵ'_m might be strictly less than ϵ_m. We consider the two situations separately and arrive at the same set of inequalities in both situations. We then derive a contradiction from that set.

Situation 1: Assume that $\epsilon_{m,n}$ does not exist for any $n \in \mathbb{Z}^+$. Then

$$\mathcal{N}_n(\epsilon) \leq m - 1 \tag{16}$$

for all $n \in \mathbb{Z}^+$ and $\epsilon > 0$. Using the definition of degrees of freedom for T there exist constants $\alpha < \beta < \epsilon_m$ such that

$$\mathcal{N}_n(\alpha) \leq m - 1 \quad \text{for all } n \in \mathbb{Z}^+ \quad \text{and} \tag{17}$$

$$\mathcal{N}(\beta) \geq m. \tag{18}$$

Situation 2: Assume that $\epsilon'_m < \epsilon_m$. From the definition of DOF singular values we know

$$\sup_{\epsilon > \epsilon_{m,n}} \mathcal{N}_n(\epsilon) \le m - 1 \quad \text{for all } n \in \mathbb{Z}^+ \quad \text{and}$$
$$\inf_{\epsilon < \epsilon_m} \mathcal{N}(\epsilon) \ge m.$$

Because $\epsilon_{m,n} \le \epsilon'_m$, we know that there exist numbers α and β, $\epsilon'_m < \alpha < \beta < \epsilon_m$ such that

$$\mathcal{N}_n(\alpha) \le m - 1 \quad \text{for all } n \in \mathbb{Z}^+ \quad \text{and} \quad (19)$$
$$\mathcal{N}(\beta) \ge m. \quad (20)$$

These are the same conditions as (17) and (18). Therefore, in both situations we need to prove that the inequalities (19) and (20) cannot be simultaneously true.

Because T is compact, TB_1 is totally bounded [20, ch. 8]. Therefore, TB_1 has a finite ϵ-net for all $\epsilon > 0$. Hence there exists a set of vectors $\{\xi_1, \ldots, \xi_P\} \subset B_1$ such that for all $y \in TB_1$ there exists a p, $1 \le p \le P$ with

$$\|T\xi_p - y\|_Y < \frac{\beta - \alpha}{2}. \quad (21)$$

Now, because $\{\phi_1, \phi_2, \ldots\}$ is a complete Schauder basis for X and because $P < \infty$, there exists a number N such that for all $n > N$ and for all p, $1 \le p \le P$, there exists a $\xi_{p,n} \in S_n \cap B_1$ such that

$$\|\xi_{p,n} - \xi_p\|_X < \frac{\beta - \alpha}{2\|T\|}. \quad (22)$$

Therefore, for all $y \in TB_1$ and for all $n > N$ there exists a $p, 1 \le p \le P$ and a $\xi_{p,n} \in S_n \cap B_1$ such that

$$
\begin{aligned}
\|T\xi_{p,n} - y\|_Y &= \|T\xi_{p,n} - T\xi_p + T\xi_p - y\|_Y \\
&\le \|T\xi_{p,n} - T\xi_p\|_Y + \|T\xi_p - y\|_Y \\
&< \|T(\xi_{p,n} - \xi_p)\|_Y + \frac{\beta - \alpha}{2} \\
&< \|T\|\frac{\beta - \alpha}{2\|T\|} + \frac{\beta - \alpha}{2} \\
&= \beta - \alpha. \quad (23)
\end{aligned}
$$

We get the first inequality above from the triangle inequality, the second one from inequality (21) and the final one from inequality (22). From inequality (19) and the definition of the number of degrees of freedom, we know that for all $n \in \mathbb{Z}^+$ there exists a set of vectors $\{\psi_{1,n}, \ldots, \psi_{m-1,n}\} \subset Y$ such that

$$y \in \text{span}_\alpha\{\psi_{1,n}, \ldots, \psi_{m-1,n}\} \quad (24)$$

for all $y \in T(S_n \cap B_1)$.

But, from the definition of the number of degrees of freedom and inequality (20) we know that for all $n \in \mathbb{Z}^+$ and all sets of vectors $\{\psi_{1,n}, \ldots, \psi_{m-1,n}\}$ there exists a vector $\psi \in TB_1$ such that

$$\psi \notin \text{span}_\beta\{\psi_{1,n}, \ldots, \psi_{m-1,n}\}.$$

From inequality (23) we know that for all $n > N$ there exists a $\xi_{p,n} \in S_n \cap B_1$ such that

$$\|T\xi_{p,n} - \psi\| < \beta - \alpha.$$

Therefore, for all $n > N$ there exists a $\xi_{p,n} \in S_n \cap B_1$ such that

$$T\xi_{p,n} \notin \text{span}_\alpha\{\psi_{1,n}, \ldots, \psi_{m-1,n}\}. \quad (25)$$

This directly contradicts condition (24). Therefore, if ϵ_m exists then $\epsilon_{m,n}$ exists for n large enough and

$$\lim_{n \to \infty} \epsilon_{m,n} = \epsilon_m.$$

∎

Theorem 3.9. Let X, Y be two finite dimensional Banach spaces and let $T: X \to Y$ be a linear operator. Also let B_1 be the closed unit ball in X and suppose Ψ_p is defined as in Section III-D. Then

$$\sup_{x \in B_1} \|Tx\|_Y = \epsilon_1$$

and for all $p \in \mathbb{Z}^+$

$$\inf_{\{\psi_i\}_{i=1}^p \in \Psi_p} \sup_{x \in B_1} \inf_{a_1, \ldots, a_p} \left\| Tx - \sum_{i=1}^p a_i \psi_i \right\|_Y = \epsilon_{p+1}.$$

Proof: Let ϵ'_{p+1} denote the left hand side of the above equation. Assume $\epsilon'_{p+1} < \epsilon_{p+1}$. Then there exists a set $\{\psi_i\}_{i=1}^p \in \Psi_p$ such that

$$\epsilon''_{p+1} := \sup_{x \in B_1} \inf_{a_1, \ldots, a_p} \left\| Tx - \sum_{i=1}^p a_i \psi_i \right\| < \epsilon_{p+1}.$$

By definition this implies $\mathcal{N}(\epsilon''_{p+1}) \le p$, a contradiction to $\inf_{\epsilon < \epsilon_{p+1}} \mathcal{N}(\epsilon) \ge p + 1$. Hence $\epsilon'_{p+1} \ge \epsilon_{p+1}$. Now assume $\epsilon'_{p+1} > \epsilon_{p+1}$. Let $\epsilon \in (\epsilon_{p+1}, \epsilon'_{p+1})$. From $\sup_{\epsilon > \epsilon_{p+1}} \mathcal{N}(\epsilon) = p$ it follows $\mathcal{N}(\epsilon) \le p$. Hence there exists a set $\{\psi_i\}_{i=1}^p \subset Y$ such that

$$\sup_{x \in B_1} \inf_{a_1, \ldots, a_p} \left\| Tx - \sum_{i=1}^p a_i \psi_i \right\| \le \epsilon < \epsilon'_{p+1}.$$

Therefore $\{\psi_i\}_{i=1}^p \in \Psi_{p,\epsilon} \subset \Psi_p$ and

$$\epsilon'_{p+1} > \inf_{\{\psi_i\}_{i=1}^p \in \Psi_p} \sup_{x \in B_1} \inf_{a_1, \ldots, a_p} \left\| Tx - \sum_{i=1}^p a_i \psi_i \right\|,$$

a contradiction. Hence $\epsilon'_{p+1} = \epsilon_{p+1}$. ∎

Theorem 3.10. (Converse to Theorem 3.1) Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $T: X \to Y$ is a bounded linear operator. If for all $\epsilon > 0$ there exist $N \in \mathbb{Z}_0^+$ and a set $\{\psi_i\}_{i=1}^N \subset Y$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1, \ldots, a_N} \left\| Tx - \sum_{i=1}^N a_i \psi_i \right\|_Y \le \epsilon$$

then T is compact.

Proof: We prove that T is compact by showing that the set $T(\overline{B}_{1,X}(0))$ is totally bounded. Let $\delta > 0$ be given. Then there exist an $N \in \mathbb{Z}_0^+$ and a set $\{\psi_i\}_{i=1}^N \subset Y$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1,\ldots,a_N} \left\| Tx - \sum_{i=1}^N a_i \psi_i \right\|_Y \leq \frac{\delta}{4}. \tag{26}$$

For any given $x \in \overline{B}_{1,X}(0)$ we can choose a_i^x, $i = 1, \ldots, N$ such that

$$\left\| Tx - \sum_{i=1}^N a_i^x \psi_i \right\|_Y$$

$$\leq \inf_{a_1,\ldots,a_N} \left\| Tx - \sum_{i=1}^N a_i \psi_i \right\|_Y + \frac{\delta}{4} \tag{27}$$

$$\leq \frac{\delta}{2}. \tag{28}$$

Here, the last inequality follows from (26). Also, because we can choose $a_i = 0$ for $i = 1, \ldots, N$, for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{a_1,\ldots,a_N} \left\| Tx - \sum_{i=1}^N a_i \psi_i \right\|_Y \leq \|Tx\|_Y. \tag{29}$$

Substituting inequality (29) into (27) and using the triangle inequality, we get

$$\left\| \sum_{i=1}^N a_i^x \psi_i \right\|_Y \leq 2\|Tx\|_Y + \frac{\delta}{4} \leq 2\|T\| + \frac{\delta}{4}. \tag{30}$$

We get the last inequality from the boundedness of T. Because the span of $\psi_1, \ldots, \psi_N$ is finite dimensional and because of the uniform bound (30), there exists a finite set of elements $\{y_1, \ldots, y_M\} \subset Y$ such that for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{i=1,\ldots,M} \left\| y_i - \sum_{j=1}^N a_j^x \psi_j \right\|_Y \leq \frac{\delta}{2}. \tag{31}$$

From inequalities (31) and (28) and the triangle inequality we get for all $x \in \overline{B}_{1,X}(0)$

$$\inf_{i=1,\ldots,M} \|y_i - Tx\|_Y \leq \delta. \tag{32}$$

Therefore, the y_i, $i = 1, \ldots, M$ form a finite δ-net for $T(\overline{B}_{1,X}(0))$ and therefore $T(\overline{B}_{1,X}(0))$ is totally bounded. Hence, T is compact. ∎

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[2] R. Gallagher, *Information Theory and Reliable Communication*. New York, USA: John Wiley & Sons, 1968.

[3] S. Verdú, "Fifty years of Shannon theory," *IEEE Transactions on Information Theory*, vol. 44, no. 6, p. 2057, 1998.

[4] H. Landau and H. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty - III: The dimension of the space of essentially time- and band-limited signals," *The Bell System Technical Journal*, vol. 41, pp. 1295–1336, Jul 1962.

[5] E. Biglieri and G. Taricco, *Transmission and Reception with Multiple Antennas: Theoretical Foundations*, ser. Foundations and Trends in Communications and Information Theory. Now Publishers, 2004.

[6] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Transactions on Communications*, vol. 50, no. 2, p. 225, 2002.

[7] A. Grant and P. D. Alexander, "Random sequence multisets for synchronous code-division multiple-access channels," *IEEE Transactions on Information Theory*, vol. 44, no. 7, p. 2832, 1998.

[8] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*, ser. Foundations and Trends in Communications and Information Theory. Now Publishers, 2004.

[9] A. S. Y. Poon, R. W. Brodersen, and D. N. C. Tse, "Degrees of freedom in multiple-antenna channels: A signal space approach," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 523–536, February 2005.

[10] L. Hanlen and M. Fu, "Wireless communication systems with spatial diversity: A volumetric model," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 133–142, January 2006.

[11] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Transactions on Signal Processing*, vol. 55, pp. 2542–2556, 2007.

[12] J. Xu and R. Janaswamy, "Electromagnetic degrees of freedom in 2-D scattering environments," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 12, pp. 3882–3894, December 2006.

[13] M. D. Migliore, "On the role of the number of degrees of freedom of the field in MIMO channels," *IEEE Transactions on Antennas Propagation*, vol. 54, no. 2, pp. 620–628, February 2006.

[14] D. A. Miller, "Communicating with waves between volumes: evaluating orthogonal spatial channels and limits on coupling strengths," *Applied Optics*, vol. 39, no. 11, pp. 1681–1699, April 2000.

[15] O. M. Bucci and G. Franceschetti, "On spatial bandwidth of scattered fields," *IEEE Transactions on Antennas and Propagation*, vol. 35, no. 12, pp. 1445–1455, December 1987.

[16] ——, "On the degrees of freedom of scattered fields," *IEEE Transactions on Antennas Propagation*, vol. 37, no. 7, pp. 318–326, July 1989.

[17] D. Slepian, "On bandwidth," *Proc. IEEE*, vol. 64, no. 3, pp. 292–300, Mar. 1976.

[18] R. Somaraju, "Essential dimension and degrees of freedom for spatial waveform channels," Ph.D. dissertation, The Australian National University, 2008.

[19] S. Haykin, *Communication Systems*. Wiley; 4th edition, 2000.

[20] E. Kreyszig, *Introductory functional analysis with applications*. John Wiley & Sons, 1989.

[21] T. Kato, *Perturbation Theory for Linear Operators*. Springer, 1980.