Eigenvalue Estimates and Mutual Information for the Linear Time-Varying Channel

Brendan Farrell, *Member, IEEE*, and Thomas Strohmer

Abstract—We consider linear time-varying channels with additive white Gaussian noise. For a large class of such channels we derive rigorous estimates of the eigenvalues of the correlation matrix of the effective channel in terms of the sampled time-varying transfer function and, thus, provide a theoretical justification for a relationship that has been frequently observed in the literature. We then use this eigenvalue estimate to derive an estimate of the mutual information of the channel. Our approach is constructive and is based on a careful balance of the tradeoff between approximate operator diagonalization, signal dimension loss, and accuracy of eigenvalue estimates.

Index Terms—Approximate diagonalization, eigenvalue estimates, mutual information, time-varying channel, Weyl–Heisenberg system.

I. INTRODUCTION

A. Motivation

THE linear, time-invariant (LTI) channel with impulse response h

$$r(t) = \int h(t - \tau)s(\tau)d\tau \tag{1}$$

and additive white Gaussian noise with variance σ^2 has normalized capacity

$$\frac{1}{2W} \int_{-W}^{W} \log\left(1 + \frac{|\hat{h}(\omega)|^2}{\sigma^2}\right) d\omega \tag{2}$$

for signals band-limited to $[-W, W]$. This classical result is, of course, due to Shannon [1], and is probably the most fundamental result in information theory. We refer to [2] for the mathematical steps and the information-theoretic details for establishing (2).

The linear, time-varying (LTV) channel is given by

$$r(t) = \int h(t, t - \tau)s(\tau)d\tau. \tag{3}$$

Motivated by Shannon's groundbreaking result, it has been a longstanding desire of engineers and mathematicians to derive a characterization of the capacity of time-varying channels in terms of the associated time-varying transfer function, analogous to (2). While such a characterization seems still quite out of reach for the general case, our aim in this paper is to get one step closer to this ambitious goal. The mathematical foundation for Shannon's famous result is the fact that in the time-invariant case the (generalized) eigenvalues of the channel matrix are directly related to samples of the transfer function. Thus, it is natural to ask to what extent such a relationship carries over to the time-varying case, which is what we plan to answer in this paper.

For information-theoretic studies of some special cases of time-varying channels we refer the reader to [3] and its vast list of references. In this paper we focus on the class of time-varying channels whose spreading function decays at an exponential rate both in time and frequency. This channel class is motivated by physical properties of channel propagation and includes for instance underspread channels [4], [5].

B. Contributions

A precise formulation of the results of this paper requires several steps of preparation. Therefore, we delay the rigorous presentation of our results to later sections, and instead give an informal description of our contributions.

The main result of our paper shows that the eigenvalues of the correlation matrix of the effective channel can be well approximated via sampling values of the autocorrelation of the time-varying transfer function. We derive rigorous bounds for the accuracy of this approximation. Our approach is constructive and is based on a careful balance of the tradeoff between approximate matrix diagonalization, signal dimension loss, and accuracy of eigenvalue estimates. While the proof of the eigenvalue estimate is quite delicate, this will come as no surprise to the expert in pseudodifferential operator theory, since characterizing the spectrum of a pseudodifferential operator (which is essentially an operator of the form (3)) via its symbol has always been a difficult task.

We then show how this eigenvalue estimate can be used to derive an estimate of the mutual information of these channels. Recall that for the time-invariant case the mutual information (and thus, in turn, the capacity) is *precisely* captured by the sampled Fourier transform of the autocorrelation of the impulse response, as the time interval is extended to infinity. Building on our eigenvalue estimates, we rigorously relate the mutual information to samples of the Fourier transform of the "twisted auto-convolved" spreading function.

C. Remarks on the Proof Strategy

A few comments on the proof strategy seem in order. Two different types of signal sets will play an important role: Weyl–Heisenberg signals and prolate spheroidal wave functions. The reader may wonder why we do not stick with just one of these two types. The reason is that each of the two has some major advantages, but also some significant limitations. Thus, by introducing both types, Weyl–Heisenberg signals and prolate spheroidal wave functions (PSWFs), we can fully utilize the positive properties of each set, while mitigating its negative properties with the other set.

For the eigenvalue estimate we rely on a set of well localized Weyl–Heisenberg signals whose span is close to the span of the PSWFs in a sense that will be formalized in the proof. While the PSWFs are optimally localized in an L^2-sense, their lack of sufficient temporal decay (except for the first few PSWFs) prohibits us from linking the eigenvalues of $\mathbf{A}^*\mathbf{A}$, the correlation matrix of the effective channel, to the associated time-varying transfer function. The off-diagonal entries of the resulting matrix would have at best linear decay, which is simply insufficient for any reasonable estimate. On the other hand, the excellent localization properties of the Weyl–Heisenberg set yield an approximate diagonalization of the channel, so that the off-diagonal entries of $\mathbf{A}^*\mathbf{A}$ decay exponentially, which allows us to obtain a rather tight eigenvalue estimate.

The mutual information will depend on the type and number of transmission signals. We use a signaling set consisting of about $2TW$ mutually orthonormal W-bandlimited signals which are "essentially localized" to a time interval of length T. The associated signal space, rigorously defined in Definition 2.1, will be denoted by $L^2(T, W, \varepsilon)$. It is not difficult to construct a linear independent, well-localized set of Weyl–Heisenberg signals. However, due to the infamous Balian–Low theorem (see Section III-A) such a set will be necessarily incomplete in $L^2(\mathbb{R})$, which in turn implies that the number of Weyl–Heisenberg signals inside $L^2(T, W, \varepsilon)$ is somewhat smaller than $2TW + 1$, the approximate dimension of $L^2(T, W, \varepsilon)$. This dimension loss makes a direct estimate of the mutual information somewhat cumbersome. And that is where PSWFs come into play. We (approximately) represent $L^2(T, W, \varepsilon)$ via the PSWFs, and then quantify the (small) dimension loss between the Weyl–Heisenberg set and the PSWFs. Combining this estimate with our eigenvalue estimate enables us then to estimate the mutual information in terms of the time-varying transfer function.

D. Connections to Prior Work

Our work is related to previous research on two aspects of time-varying channels. Previous authors have discussed diagonalizing the channel and giving the capacity in terms of singular values [6]–[8], and other authors have focused on determining transmission signals with various useful properties [9], [10]. Our paper is probably closest in spirit to [11], where the authors derive estimates for the noncoherent capacity for certain time-varying channels by carefully combining signal design with approximate diagonalization.

Much of the mathematical approach to time-varying channels from a time-frequency analysis perspective originated with Kozek [9], [12], [13]. While he addresses issues such as the composition and estimation of time-varying channel operators and the time-frequency localization of transmission signals, his focus is a WSSUS model. Here we work with a deterministic channel.

The remainder of the paper is organized as follows. At the end of this section we introduce mathematical tools and notation used throughout the paper. Section II describes our setup, the channel model and the signal model. We derive the eigenvalue estimate in Section III and present the estimate of the mutual information in Section IV.

E. Mathematical Tools and Notation

Let f be a function in $L^2(\mathbb{R})$. The *modulation operator* M_ω is defined by

$$M_\omega f(t) = e^{2\pi i \omega \cdot t} f(t) \qquad (4)$$

and the *translation operator* T_x is defined by

$$T_x f(t) = f(t - x) \qquad (5)$$

for all $f \in L^2(\mathbb{R})$. The *Fourier transform* of a function $f \in L^2(\mathbb{R})$ is given by

$$(\mathcal{F}f)(\omega) = \int f(t) e^{-2\pi i \omega t} dt. \qquad (6)$$

We also write $\hat{f}$ for $\mathcal{F}f$. The Fourier transform of a function in two variables is defined by extending (6) in the usual way to two dimensions. Sometimes we need to take the Fourier transform of a function $f(t_1, t_2)$ with respect to the first or the second variable only. In this case we write $\mathcal{F}_1 f$ or $\mathcal{F}_2 f$, respectively. When no interval is given, integration is over all of $\mathbb{R}$. For a complex-valued function f, we denote its complex conjugate by $\bar{f}$. The eigenvalues of a matrix $\mathbf{A}$ are denoted by $\lambda_j(\mathbf{A})$.

The Weyl pseudodifferential operator $\mathbf{L}_\sigma$ is defined as

$$\mathbf{L}_\sigma f(t) = \int \int \hat{\sigma}(\omega, x) e^{-\pi i x \omega} T_{-x} M_\omega f(t) \, d\omega \, dx. \qquad (7)$$

Here, σ is called the *symbol* and its Fourier transform, $\hat{\sigma}$, is called the *spreading function*. We can express the composition of two pseudodifferential operators $\mathbf{L}_\sigma$, $\mathbf{L}_\tau$ in terms of their symbols. There holds $\mathbf{L}_\sigma \mathbf{L}_\tau = \mathbf{L}_{\sigma \sharp \tau}$, where $\sigma \sharp \tau = \mathcal{F}^{-1}(\hat{\sigma} \natural \hat{\tau})$ denotes the *twisted product* of σ and τ, and

$$(\hat{\sigma} \natural \hat{\tau})(\omega, x) =$$
$$\int \int \hat{\sigma}(\omega', x') \hat{\tau}(\omega - \omega', x - x') e^{-\pi i (x\omega' - \omega x')} d\omega' dx'$$

is called the *twisted convolution* of $\hat{\sigma}$ and $\hat{\tau}$, see [14]. This can be seen as a generalization of the composition rule of two time-invariant operators via ordinary convolution.

We set $\mathcal{S} = \bar{\sigma} \sharp \sigma$, which is the Fourier transform of the "twisted autocorrelation" of $\hat{\sigma}$. Since $\mathcal{S}$ takes values in $\mathbb{R}$, $\mathcal{S}^+(u, v)$ is defined by $\mathcal{S}^+(u, v) = \max(\mathcal{S}(u, v), 0)$.

II. CHANNEL MODEL AND SIGNAL MODEL

We first derive an equivalent representation of the channel model (3). We set $\sigma(t,\omega) = \mathcal{F}_2 h(t,\cdot)$. Several manipulations and applications of the Fourier transform yield [14]

$$\int h(t, t-\tau)s(\tau)d\tau = \int\int \hat{\sigma}(\omega, x)M_\omega T_{-x}s(t)d\omega dx. \quad (8)$$

This allows us to equivalently express the linear time-varying channel as a pseudodifferential operator

$$\mathbf{L}_\sigma s(t) = \int\int \hat{\sigma}(\omega, x)e^{-\pi i x\omega}T_{-x}M_\omega s(t)d\omega dx. \quad (9)$$

The integral in (9) has the interpretation that the received signal is a weighted sum of shifted and modulated copies of the original signal. Using the Weyl form allows us to express the channel as an operator that has further useful relationships to other forms that will be helpful in our proof. See [14] for further background on such operators.

Our model is now given by the following steps and is illustrated in (10)–(14). First the random variable $x \in \mathbb{C}^N, x \sim \mathcal{NC}(0, I_N)$ is mapped to a set of orthonormal transmission signals ϕ_i as coefficients (10). The signal passes through the channel given by $\mathbf{L}_\sigma$ (11) and is corrupted by AWGN (12). The received signal is mapped to a sequence of random variables y by taking the inner product with the detection signals ψ_k^r (13)

$$\mathcal{NC}(0, I_N) \sim x \overset{\Phi}{\rightarrow} \sum_{i=1}^N x_i\phi_i \quad (10)$$

$$\overset{\mathbf{L}_\sigma}{\rightarrow} \mathbf{L}_\sigma \sum_{i=1}^N x_i\phi_i \quad (11)$$

$$\overset{\oplus \text{ noise}}{\rightarrow} \mathbf{L}_\sigma \sum_{i=1}^N x_i\phi_i + n \quad (12)$$

$$\overset{\mathbf{C}}{\rightarrow} \{\langle \mathbf{L}_\sigma \sum_{i=1}^N x_i\phi_i + n, \psi_k^r\rangle\}_{k\in\mathbb{Z}} \quad (13)$$

$$= y. \quad (14)$$

The reader will have noticed that we use a different set of signals at the transmitter and the receiver. The mutual information between x and y, $\mathcal{I}(x, y)$ depends on the transmission signals $\{\phi_i\}_{i=1}^N$ and the number of transmission signals, but as long as $\{\psi_k^r\}_{k\in\mathbb{Z}}$ is an orthonormal basis for $L^2(\mathbb{R})$ or a tight frame, then $\mathcal{I}(x, y)$ is independent of the receive signals. It is clear that the transmission signals $\{\phi_i\}_{i=1}^N$ should form a linearly independent set. As already briefly indicated, later the Balian–Low theorem will force us to select the linearly independent set of transmission signals from a set of functions that is also *incomplete* in $L^2(\mathbb{R})$. Obviously, this implies a dimension loss of the signal space which manifests itself in an additional error term in our main estimate of the mutual information. An *additional* dimension loss would occur if we also used an incomplete signaling set at the receiver. However, at the receiver we are not restricted to linearly independent signaling sets (thus, the Balian–Low theorem is no longer an obstacle), and,

therefore, we will use a different, and, in fact, overcomplete, signaling set at the receiver.

Now we introduce and discuss our requirement on the transmission signals. We require that they are L^2-localized to a time-frequency rectangle, which we formalize with the following definition.

Definition 2.1: We define the space $L^2(T, W, \varepsilon)$ by

$$L^2(T, W, \varepsilon) =$$
$$\left\{ f \in L^2(\mathbb{R}) : \int_0^T |f(t)|^2 dt \geq (1-\varepsilon^2)\|f\|_{L^2(\mathbb{R})} \right.$$
$$\left. \text{and} \quad \int_{-W}^W |\hat{f}(\omega)|^2 d\omega \geq (1-\varepsilon^2)\|f\|_{L^2(\mathbb{R})} \right\}.$$

Given the intervals $[0, T], [-W, W]$ we denote by $\{\varphi_n\}_{n=0}^\infty$ the associated PSWFs similar to [15], [25].[1] Let $\mathbf{P}$ be the orthogonal projection onto the span of $\varphi_0, \ldots, \varphi_{2TW}$. By Theorem 12 in [15] for every $f \in L^2(T, W, \varepsilon)$

$$\|f - \mathbf{P}f\|_2 \leq 7\varepsilon\|f\|_2. \quad (15)$$

In other words, $L^2(T, W, \varepsilon)$ is well approximated by the first $2TW + 1$ elements of the PSWFs and $L^2(T, W, \varepsilon)$ is essentially $(2TW + 1)$-dimensional.

There are several reasons for restricting our transmission signals to this space. Firstly, any real-world communication signal has finite duration and (essentially) finite bandwidth. The above model is a standard way to describe this property in a mathematically meaningful way [15]. Secondly, for time-varying channels it is more insightful to have expressions for eigenvalue estimates or mutual information for finite time intervals (and of course finite bandwidth) than for infinite time, as is also reflected in the papers [2], [6], [11]. Thus, it is useful to require some form of time-frequency localization of the transmission signals. We note that we could have chosen the signal space with somewhat different localization conditions, such as for instance using exactly time-limited signals. However, our symmetric localization condition in Definition 2.1 lends itself to a somewhat shorter proof (admittedly, in spite of the overall length of our proof, the reader might find that using the term "shorter" is not appropriate here).

III. EIGENVALUE ESTIMATES FOR TIME-VARYING CHANNELS

A. Weyl–Heisenberg Systems, Time-Frequency Localization and Mutual Information

We assume that the reader is familiar with frame theory and refer to [14] for background.

Definition 3.1: For a given function $\phi \in L^2(\mathbb{R})$ (the window function) and given parameters $a, b > 0$, we denote the associated *Gabor system* or *Weyl–Heisenberg system* by $(\phi, a, b) := \{M_{bl}T_{ak}\phi\}_{k,l\in\mathbb{Z}}, a, b \in \mathbb{R}^+$. The *redundancy* of this system is $\frac{1}{ab}$. (Note that $ab \leq 1$ is necessary for (ϕ, a, b) to be a frame for $L^2(\mathbb{R})$ [14].)

---

[1]The minor and trivial difference to [15], [25] is that we consider $[0, T]$ and not $[-T, T]$.

Proposition 3.2: Let $g_s(t) = (2s)^{-1/4}e^{-\frac{\pi}{s}t^2}$, and set $\psi_s = \mathbf{S}^{-1/2}g_s$, where $\mathbf{S}$ is the frame operator corresponding to $(g_s, \frac{a}{\rho}, \frac{b}{\rho})$. Then $(\psi_s, \frac{\rho}{b}, \frac{\rho}{a}) = (\psi_s, \rho a, \rho b)$ $(ab = 1$ and $\rho > 1)$ is an orthonormal system and there exist constants $C > 0$ and $0 < D < 1$ such that

$$|\psi_s(t)| \le Ce^{-D\frac{\pi}{s}|t|} \quad \forall t \in \mathbb{R}$$
$$|\widehat{\psi_s}(\omega)| \le Ce^{-D\pi s|\omega|} \quad \forall \omega \in \mathbb{R}.$$

Proof: A fundamental theorem due to Lyubarskii, Seip and Wallsten states that $(g_s, \frac{b}{\rho}, \frac{a}{\rho})$ is a frame for $L^2(\mathbb{R})$ if and only if $\frac{ab}{\rho^2} < 1$ [17]–[19]. By Theorem 5.1.6 and Corollary 7.3.2 in [14], $(\mathbf{S}^{-1/2}g_s, \frac{b}{\rho}, \frac{a}{\rho}) = (\psi_s, \frac{b}{\rho}, \frac{a}{\rho})$ is a tight frame for $L^2(\mathbb{R})$ with frame constant ρ^2. Now we use the Weyl–Heisenberg biorthogonality relations [20]–[22], which state that if $\mathbf{S}_{g,\gamma} = \sum_{k,l\in\mathbb{Z}}\langle \cdot, M_{\beta l}T_{\alpha k}g\rangle M_{\beta l}T_{\alpha k}\gamma = I$ on $L^2(\mathbb{R})$, then $\langle\gamma, M_{l/\alpha}T_{k/\beta}g\rangle = \alpha\beta\delta_{k,0}\delta_{l,0}$. A ready consequence of this essential theorem is that $(\psi_s, \frac{\rho}{a}, \frac{\rho}{b}) = (\psi_s, \rho b, \rho a)$ $(ab = 1)$ is an orthonormal set [14]. Note that $(\psi_s, \rho b, \rho a)$ does not span $L^2(\mathbb{R})$). By Theorem 5 in [23], up to a factor $0 < D < 1$, the exponential decay of g_s and $\hat{g}_s$ is preserved in ψ_s and $\hat{\psi}_s$. Finally, Theorem IV.2 in [24] implies that if ψ_1 is the window function for the orthonormal set based on the initial window g_1, then ψ_s is the corresponding window function for g_s. ∎

Let ψ_s, a and b be as in the previous proposition. We construct our signals by setting $a = \frac{\beta}{\alpha}$, $b = \frac{\alpha}{\beta}$ and $s = (\frac{\alpha}{\beta})^2$. The signals are then defined by:

D1) $\psi^t = \psi_{(\frac{\alpha}{\beta})^2}$;
D2) $\psi^r = \frac{1}{\rho}\psi_{(\frac{\alpha}{\beta})^2}$;
D3) $\psi^t_{k,l} = M_{\rho\frac{\alpha}{\beta}l}T_{\rho\frac{\beta}{\alpha}k}\psi^t$.

Here, t stands for "transmit" and r stands for "receive".

Definition 3.3: A function $f \in L^2(\mathbb{R})$ is *exponentially localized* to the region $[0, T] \times [-W, W]$ if there exist constants c_1, C_1, c_2 and C_2 such that

$$|f(t)| \le C_1 e^{-c_1|t|} \text{ and } |\hat{f}(\omega)| \le C_2 e^{-c_2|\omega|} \quad (16)$$

for all $t \notin [0, T]$ and all $\omega \notin [-W, W]$.

The Balian–Low theorem [14] precludes the existence of an orthonormal Weyl–Heisenberg basis (ϕ, a, b) for $L^2(\mathbb{R})$ with well-localized window function. In particular, ϕ and $\hat{\phi}$ could never have exponential decay. On the other hand [as, for instance, Proposition (3.2) shows] it is not difficult to construct an orthonormal system that is incomplete in $L^2(\mathbb{R})$ or an overcomplete system (ϕ, a, b) with a ϕ that is exponentially well localized in time and frequency. Thus, the Balian–Low theorem is the reason why we use a signaling set at the transmitter drawn from an incomplete system for $L^2(\mathbb{R})$ (implying $\rho > 1$) and an overcomplete signaling set at the receiver.

While mutual information is not the main topic of this section, we take the opportunity to address a nontrivial aspect associated with mutual information that arises from using a tight frame instead of an orthonormal basis as receive functions. If we used an orthonormal basis at the receiver, then the noise covariance matrix, $\mathbf{C}_N\mathbf{C}_N^*$ in the proof below, would be a multiple of the identity, and this proposition would be simple and standard. Using a unit-norm tight frame rather than an orthonormal basis does not change the eigenvalues, but it does make the property addressed in the proposition below more delicate. The exponential localization at the receiver and the $L^2(T, W, \varepsilon)$-property at the transmitter, however, deliver the necessary approximations for this proposition to hold.

Proposition 3.4: Let $\{\phi_{kl}\}_{(k,l)\in\mathcal{J}}$, $|\mathcal{J}| < \infty$, be orthonormal transmission signals contained in $L^2(T, W, \varepsilon)$, and let $\{\psi^r_{kl}\}_{k,l\in\mathbb{Z}}$ be a tight frame of exponentially localized receiver signals (with frame bound 1). Let $x \sim \mathcal{NC}(0, I_{|\mathcal{J}|})$ and

$$y_{kl} = \langle\mathbf{L}_\sigma\sum_{k'l'\in\mathcal{J}}x_{k'l'}\phi_{k'l'} + n, \psi^r_{kl}\rangle, \quad \text{for } k, l \in \mathbb{Z}$$

where $n(t)$ is AWGN of variance η^2. Denote

$$\mathbf{A}_{klk'l'} = \langle\mathbf{L}_\sigma\phi_{k'l'}, \psi^r_{kl}\rangle. \quad (17)$$

Then

$$\mathcal{I}(x; y) = \sum_{i=1}^{|\mathcal{J}|}\log\left(1 + \frac{\lambda_i(\mathbf{A}\mathbf{A}^*)}{\eta^2}\right). \quad (18)$$

Proof: Let $\Phi : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be the orthogonal projection onto $\text{span}\{\phi_{kl}\}_{(k,l)\in\mathcal{J}}$, and let $\mathbf{C} : L^2(\mathbb{R}) \to l^2(\mathbb{Z}^2)$ and $\mathbf{C}_N : L^2(\mathbb{R}) \to \mathbb{C}^{(2N+1)^2}$ be the coefficient operators given by $\mathbf{C}f = \{\langle f, \psi^r_{kl}\rangle\}_{k,l\in\mathbb{Z}}$ and $\mathbf{C}_Nf = \{\langle f, \psi^r_{kl}\rangle\}_{|k|,|l|\le N}$ for $N \in \mathbb{N}$. The mutual information $\mathcal{I}(x; y)$ is

$$\mathcal{I}(x; y)$$
$$= \lim_{N\to\infty}\{\log\det(\mathbf{C}_N\mathbf{L}_\sigma\Phi\mathbf{L}_\sigma^*\mathbf{C}_N^* + \eta^2\mathbf{C}_N\mathbf{C}_N^*)$$
$$\qquad - \log\det(\eta^2\mathbf{C}_N\mathbf{C}_N^*)\}.$$

Assume $\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\Phi$ has rank k, and arrange all eigenvalues in nonincreasing order. We must show that

$$\lim_{N\to\infty}\lambda_i(\mathbf{C}_N\mathbf{L}_\sigma\Phi\mathbf{L}_\sigma^*\mathbf{C}_N^* + \mathbf{C}_N\mathbf{C}_N^*)\lambda_i(\mathbf{C}\mathbf{L}_\sigma\Phi\mathbf{L}_\sigma\mathbf{C}^*) + 1$$

for $i = 1, \ldots, k$. Note that $\mathbf{C}\mathbf{L}_\sigma\Phi\mathbf{L}_\sigma^*\mathbf{C}^*$ and $\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\Phi$ have the same nonzero eigenvalues.

Since $\hat{\sigma}$ decays exponentially in both variables and each $\phi_{k'l'} \in L^2(T, W, \varepsilon)$, using the Cauchy-Schwartz inequality shows that each $\mathbf{L}_\sigma\phi_{k,l}$ is exponentially localized a time-frequency rectangle slightly larger than $[0, T] \times [-W, W]$. Thus, the range of $\mathbf{L}_\sigma\Phi$ is exponentially localized in time and frequency, and so any eigenvectors of $\mathbf{L}_\sigma\Phi\mathbf{L}_\sigma^*$ corresponding to nonzero eigenvalues, since they belong to the range of $\mathbf{L}_\sigma\Phi$, are similarly exponentially localized, which holds as well for $\mathbf{C}_N\mathbf{L}_\sigma\Phi\mathbf{L}_\sigma^*\mathbf{C}_N^*$ for all N. In particular, for all f in the range of $\mathbf{L}_\sigma\Phi$, there exist positive constants c, C such that

$$\|\mathbf{C}_N^*\mathbf{C}_Nf - f\|_{L^2(\mathbb{R})} \le Ce^{-cN}.$$

Let $u_i^{(N)}$ be an eigenvector of $\mathbf{C}_N \mathbf{L}_\sigma \Phi \mathbf{L}_\sigma^* \mathbf{C}_N^*$ corresponding to the nonzero eigenvalue λ_i. Then $u_i^{(N)} = \mathbf{C}_N f_i^{(N)}$ for some $f_i^{(N)}$ in the range of $\mathbf{L}_\sigma \Phi$. Now

$$\lim_{N \to \infty} (\mathbf{C}_N f_i^{(N)})^* \mathbf{C}_N \mathbf{C}_N^* (\mathbf{C}_N f_i^{(N)}) \tag{19}$$

$$= \lim_{N \to \infty} \langle f_i^{(N)}, \mathbf{C}_N \mathbf{C}_N^* \mathbf{C}_N \mathbf{C}_N^* f_i^{(N)} \rangle \tag{20}$$

$$= \lim_{N \to \infty} \langle f_i^{(N)}, f_i^{(N)} \rangle$$

$$= 1. \tag{21}$$

The convergence in lines (20) and (21) is exponential. While exponential convergence is not necessary, without sufficient localization of all the functions involved, convergence at all does not hold a priori for (20) and (21). For $i = 1, \ldots, k$

$$\lim_{N \to \infty} \lambda_i (\mathbf{C}_N \mathbf{L}_\sigma \Phi \mathbf{L}_\sigma^* \mathbf{C}_N^* + \eta^2 \mathbf{C}_N \mathbf{C}_N^*)$$

$$= \lim_{N \to \infty} \lambda_i (\mathbf{C}_N \mathbf{L}_\sigma \Phi \mathbf{L}_\sigma^* \mathbf{C}_N^*) + \eta^2.$$

The remaining eigenvectors of $\mathbf{C}_N \mathbf{C}_N^*$ are in the kernel of $\mathbf{C}_N \mathbf{L}_\sigma \Phi \mathbf{L}_\sigma^* \mathbf{C}_N^*$. Thus

$$\lim_{N \to \infty} \left\{ \sum_{i=1}^{(2N+1)^2} \log(\lambda_i (\mathbf{C}_N \mathbf{L}_\sigma \Phi \mathbf{L}_\sigma^* \mathbf{C}_N^* + \eta^2 \mathbf{C}_N \mathbf{C}_N^*)) \right.$$

$$\left. - \sum_{i=1}^{(2N+1)^2} \log(\lambda_i (\eta^2 \mathbf{C}_N \mathbf{C}_N^*)) \right\}$$

$$= \lim_{N \to \infty} \sum_{i=1}^{k} \log(\lambda_i (\mathbf{C}_N \mathbf{L}_\sigma \Phi \mathbf{L}_\sigma^* \mathbf{C}_N^*) + \eta^2)$$

$$- \lim_{N \to \infty} \sum_{i=1}^{k} \log(\lambda_i (\eta^2 \mathbf{C}_N \mathbf{C}_N^*)) \tag{22}$$

$$= \lim_{N \to \infty} \sum_{i=1}^{k} \log\left(1 + \frac{\lambda_i (\mathbf{C}_N \mathbf{L}_\sigma \Phi \mathbf{L}_\sigma^* \mathbf{C}_N^*)}{\eta^2}\right) \tag{23}$$

$$= \sum_{i=1}^{k} \log\left(1 + \frac{\lambda_i (\Phi^* \mathbf{L}_\sigma^* \mathbf{L}_\sigma \Phi^*)}{\eta^2}\right)$$

$$= \sum_{i=1}^{k} \log\left(1 + \frac{\lambda_i (\mathbf{A}^* \mathbf{A})}{\eta^2}\right)$$

$$= \sum_{i=1}^{|\mathcal{J}|} \log\left(1 + \frac{\lambda_i (\mathbf{A}^* \mathbf{A})}{\eta^2}\right)$$

where lines (22) and (23) are consequences of the first half of the proof. ∎

B. Eigenvalue Estimates

We are ready to give a rigorous formulation of our main result, which states that the eigenvalues of the correlation matrix $\mathbf{A}^* \mathbf{A}$ can be well approximated by samples of $\mathcal{S}$, the twisted autocorrelation of the time-varying transfer function.

Theorem 3.5 (Eigenvalue Estimate): Assume the same setup as in Proposition 3.4. Furthermore, suppose that

$$|\hat{\sigma}(\omega, x)| \leq C e^{-\beta |\omega| - \alpha |x|}. \tag{24}$$

Let $\mathcal{S} = \overline{\sigma} \sharp \sigma$. Then for $j = 1, \ldots, |\mathcal{J}|$, there exists an index pair (k, l) such that

$$\left| \lambda_j (\mathbf{A}^* \mathbf{A}) - \mathcal{S}\left(\rho \frac{\beta}{\alpha} k, \rho \frac{\alpha}{\beta} l\right) \right| \leq \mathcal{O}\left(e^{-\frac{\rho}{2}(\beta + \alpha)} + \frac{1}{(\alpha \beta D)^2} \right). \tag{25}$$

Remark: Our decay condition (24) on the spreading function comprises the standard conditions of exponential decay of delay spread and compact support of the Doppler spread [5]. Moreover, we could have imposed an *underspread* condition on the spreading function, see [16] for various notions of underspread channels. It is not hard to see that condition (41) includes (or can be easily adapted to) several forms of underspread channels. This would result in somewhat different constants in the error estimate at the cost of a slightly longer proof, but the essence of the theorem would remain the same. Furthermore, one can replace the exponential decay condition by some form of (practically less justified) polynomial decay and show that the error term in (25) would then decrease at a corresponding polynomial rate.

To prove Theorem 3.5 we cannot use PSWFs, but instead introduce exponentially localized signals. The reason is that the PSWFs decay linearly [25] and, thus, do not permit the bounds obtained in the main two lemmas of this section. This is heuristically explained by the fact that the PSWFs are the approximate eigenfunctions of the operator that restricts in time and frequency, which is a much different operator than a time-varying channel, for which the exponentially localized signals are approximate eigenfunctions. This is seen formally in the off-diagonal decay in the matrix $\mathbf{A}$ in Proposition 3.7 below. However, since both sets of signals are localized, the spaces that they span are close, which is a point that we formalize later in the proof of Theorem 4.1. Thus, the general idea is the standard linear algebra approach of working with the same space, but switching to a basis that allows for approximate diagonalization.

We first need an auxiliary result.

Lemma 3.6: For $f, g \in L^2(\mathbb{R})$, let $\mathcal{W}(f, g)$ and $\mathcal{A}(f, g)$ denote their cross-ambiguity and cross-Wigner distributions [14]. If $|\psi(x)| \leq C e^{-c_1 |x|}$ and $|\hat{\psi}(\omega)| \leq C e^{-c_2 |\omega|}$ for $c_1, c_2 > 0$, then

$$|\mathcal{W}(\psi, \psi)(x, \omega)| \leq C^2 e^{-\frac{1}{4}(c_1 |x| + c_2 |\omega|)}$$

and

$$|\mathcal{A}(\psi, \psi)(x, \omega)| \leq C^2 e^{-\frac{1}{4}(c_1 |x| + c_2 |\omega|)}.$$

Proof: The proof is contained in the proof of Theorem 2.4 in [26], when one views both distributions as short-time Fourier transforms, as explained in [14]. ∎

A key ingredient in our proof of Theorem 3.5 is the following lemma, which shows that the entries of the matrix $\mathbf{A}$ defined in (27) decay exponentially fast as we move away from the main diagonal. The approximate diagonalization of $\mathbf{A}$ via a properly designed Weyl–Heisenberg systems is well known in a qualitative sense [11], [13], [27]. What is new in the following lemma is that we give a precise quantitative formulation of this statement. This quantitative version is important in the subsequent

steps, where it will give rise to explicit and rigorous bounds on the approximation of the eigenvalues of $\mathbf{A}^*\mathbf{A}$ by samples of the twisted autocorrelation $\mathcal{S}$ of the time-varying transfer function.

Lemma 3.7: Assume that

$$|\hat{\sigma}(\omega, x)| \leq Ce^{-\beta|\omega|-\alpha|x|} \tag{26}$$

that the signals are given according to properties D1–D3 above and that

$$\mathbf{A}_{klk'l'} = \langle \mathbf{L}_\sigma \psi^t_{k'l'}, \psi^r_{kl} \rangle. \tag{27}$$

Then

$$|\mathbf{A}_{klk'l'}| \leq C(e^{-\alpha\rho|\frac{1}{\rho^2}l-l'|} + e^{-\frac{\pi}{4}D(\frac{\alpha}{\beta})^2\rho|\frac{1}{\rho^2}l-l'|})$$
$$\times (e^{-\beta\rho|\frac{1}{\rho^2}k-k'|} + e^{-\frac{\pi}{4}D(\frac{\beta}{\alpha})^2\rho|\frac{1}{\rho^2}k-k'|}).$$

Proof: The following two essential identities hold for pseudodifferential operators, cf. [14]

$$\langle \mathbf{L}_\sigma f, g \rangle = \langle \sigma, \mathcal{W}(g, f) \rangle \tag{28}$$
$$|\langle \mathbf{L}_\sigma T_u M_\eta f, T_v M_\gamma g \rangle| = |(\hat{\sigma} * \mathcal{A}(f, g))(u-v, \eta-\gamma)|. \tag{29}$$

The system is given by $\psi_{(\frac{\alpha}{\beta})^2} = \mathbf{S}^{-1/2} g_{(\frac{\alpha}{\beta})^2}$, where $g_{(\frac{\alpha}{\beta})^2}(t) = (2(\frac{\alpha}{\beta})^2)^{-1/4} e^{-\pi(\frac{\beta}{\alpha})^2 t^2}$, and by Proposition (3.2)

$$|\psi_{(\frac{\alpha}{\beta})^2}(t)| \leq Ce^{-\pi(\frac{\beta}{\alpha})^2 D|t|} \tag{30}$$
$$|\hat{\psi}_{(\frac{\alpha}{\beta})^2}(\omega)| \leq Ce^{-\pi(\frac{\alpha}{\beta})^2 D|\omega|}. \tag{31}$$

Lemma 3.6 implies

$$|\mathcal{A}(\psi_{(\frac{\alpha}{\beta})^2}, \psi_{(\frac{\alpha}{\beta})^2})(x, \omega)| \leq Ce^{-\frac{\pi}{4}D(\frac{\alpha}{\beta})^2|x|-\frac{\pi}{4}D(\frac{\beta}{\alpha})^2|\omega|}$$

$$|\mathbf{A}_{k,l,k',l'}|$$
$$= |\langle \mathbf{L}_\sigma \psi^t_{k'l'}, \psi^r_{kl} \rangle|$$
$$= |\langle \mathbf{L}_\sigma M_{\rho\frac{\alpha}{\beta}l'} T_{\rho\frac{\beta}{\alpha}k'} \psi, M_{\frac{1}{\rho}\frac{\alpha}{\beta}l} T_{\frac{1}{\rho}\frac{\beta}{\alpha}k} \psi \rangle|$$
$$= |(\hat{\sigma} * \mathcal{A}(\psi, \psi))(\frac{\beta}{\alpha}(\frac{1}{\rho}k - \rho k'), \frac{\alpha}{\beta}(\frac{1}{\rho}l - \rho l'))|$$
$$= |\int\int \hat{\sigma}(\omega, x) \mathcal{A}(\psi, \psi)$$
$$(\frac{\beta}{\alpha}(\frac{1}{\rho}k - \rho k') - x, \frac{\alpha}{\beta}(\frac{1}{\rho}l - \rho l') - \omega) dx d\omega|$$
$$\leq C\int\int e^{-\beta|\omega|-\alpha|x|}$$
$$e^{-\frac{\pi}{4}(\frac{\beta}{\alpha})^2 D|\frac{\beta}{\alpha}(\frac{1}{\rho}k-\rho k')-x|-\frac{\pi}{4}(\frac{\alpha}{\beta})^2 D|\frac{\alpha}{\beta}(\frac{1}{\rho}l-\rho l')-\omega|} d\omega dx$$
$$= C\int e^{-\beta|\omega|-\frac{\pi}{4}(\frac{\alpha}{\beta})^2 D|\frac{\alpha}{\beta}(\frac{1}{\rho}l-\rho l')-\omega|} d\omega$$
$$\times \int e^{-\alpha|x|-\frac{\pi}{4}(\frac{\beta}{\alpha})^2 D|\frac{\beta}{\alpha}(\frac{1}{\rho}k-\rho k')-x|} dx$$
$$\leq C(e^{-\alpha\rho|\frac{1}{\rho^2}l-l'|} + e^{-\frac{\pi}{4}\frac{\alpha}{\beta}D\rho|\frac{1}{\rho^2}l-l'|})$$
$$\times (e^{-\beta\rho|\frac{1}{\rho^2}k-k'|} + e^{-\frac{\pi}{4}\frac{\beta}{\alpha}D\rho|\frac{1}{\rho^2}k-k'|}) \tag{32}$$

where we have used the bound

$$\int e^{-c_1|y|} e^{-c_2|X-y|} dy \leq C(e^{-c_1|X|} + e^{-c_2|X|}).$$

∎

The following lemma shows that the eigenvalues of $\mathbf{A}^*\mathbf{A}$ are well approximated by its diagonal entries.

Lemma 3.8: Assume again the hypotheses of Proposition 3.7. Then for $j = 1, \ldots, |\mathcal{J}|$, there exists an index pair (k, l) such that

$$|\lambda_j(\mathbf{A}^*\mathbf{A}) - (\mathbf{A}^*\mathbf{A})_{klkl}| \leq \mathcal{O}(e^{-\frac{\rho}{2}(\beta+\alpha)}). \tag{33}$$

Proof:

$$(\mathbf{A}^*\mathbf{A})_{klk'l'} = \sum_{j,j'\in\mathbb{Z}} \overline{\mathbf{A}_{jj'kl}} \mathbf{A}_{jj'k'l'}$$
$$= \sum_{j,j'\in\mathbb{Z}} \overline{\langle \mathbf{L}_\sigma \psi^t_{kl}, \psi^r_{jj'} \rangle} \langle \mathbf{L}_\sigma \psi^t_{k'l'}, \psi^r_{jj'} \rangle$$
$$= \sum_{j,j'\in\mathbb{Z}} \overline{\langle \mathbf{L}_\sigma \psi^t_{kl}, \psi^r_{jj'} \rangle} \langle \psi^r_{jj'}, \mathbf{L}_\sigma \psi^t_{k'l'} \rangle$$
$$= \langle \mathbf{L}_\sigma \psi^t_{k'l'}, \mathbf{L}_\sigma \psi^t_{kl} \rangle$$
$$= \langle \mathbf{L}_\mathcal{S} \psi^t_{k'l'}, \psi^t_{kl} \rangle$$

where $\mathcal{S} = \overline{\sigma}\sharp\sigma$ was defined in Section II. Using the estimate from the proof of Lemma 3.7, we have that $|\hat{\mathcal{S}}(\omega, x)| \leq \frac{C}{\alpha\beta} e^{-\beta|\omega|-\alpha|x|}$. Using the identity in (29)

$$|\langle \mathbf{L}_\mathcal{S} \psi^t_{k'l'}, \psi^t_{kl} \rangle|$$
$$= |(\hat{\mathcal{S}} * \mathcal{A}(\psi^t, \psi^t))(\rho\frac{\beta}{\alpha}(k'-k), \rho\frac{\alpha}{\beta}(l'-l))|$$
$$\leq C(e^{-\alpha\rho|l-l'|} + e^{-\frac{\pi}{4}\frac{\alpha}{\beta}D\rho|l-l'|})$$
$$\times (e^{-\beta\rho|k-k'|} + e^{-\frac{\pi}{4}\frac{\beta}{\alpha}D\rho|k-k'|}).$$

Next

$$\sum_{\substack{k=-K,\ldots,K,k\neq k' \\ l=-L,\ldots,L,l\neq l'}} |(\mathbf{A}^*\mathbf{A})_{klk'l'}|$$
$$\leq C \sum_{k=-K,\ldots,K,k\neq k'} (e^{-\beta\rho|k-k'|} + e^{-\frac{\pi}{4}D\frac{\beta}{\alpha}\rho|k-k'|})$$
$$\times \sum_{l=-L,\ldots,L,l\neq l'} (e^{-\alpha\rho|l-l'|} + e^{-\frac{\pi}{4}D\frac{\alpha}{\beta}\rho|l-l'|})$$
$$= \mathcal{O}\left(\left(\frac{e^{-\beta\rho}}{1-e^{-\beta\rho}} + \frac{e^{-\frac{\pi}{4}D\frac{\beta}{\alpha}\rho}}{1-e^{-\frac{\pi}{4}D\frac{\beta}{\alpha}\rho}}\right)\right.$$
$$\left.\times \left(\frac{e^{-\alpha\rho}}{1-e^{-\alpha\rho}} + \frac{e^{-\frac{\pi}{4}D\frac{\alpha}{\beta}\rho}}{1-e^{-\frac{\pi}{4}D\frac{\alpha}{\beta}\rho}}\right)\right) \tag{34}$$
$$= \mathcal{O}\left(\frac{e^{-\rho(\beta+\alpha)}}{(1-e^{-\frac{\pi}{4}D\frac{\beta}{\alpha}\rho})(1-e^{-\frac{\pi}{4}D\frac{\alpha}{\beta}\rho})}\right)$$
$$= \mathcal{O}\left(e^{-\frac{\rho}{2}(\beta+\alpha)}\right). \tag{35}$$

We now have an estimate on the off-diagonal sums of the matrix $\mathbf{A}^*\mathbf{A}$ and may apply the Gershgorin disc theorem to obtain the claim. ∎

Having established that the spectrum of $\mathbf{A}^*\mathbf{A}$ is very close to its diagonal entries, we next show that in turn the diagonal of $\mathbf{A}^*\mathbf{A}$ is well approximated by the samples of the associated twisted autocorrelation $\mathcal{S}$.

Lemma 3.9: Assume again the hypotheses of Proposition 3.4 and that

$$|\hat{\sigma}(\omega, x)| \le Ce^{-\beta|\omega|-\alpha|x|}.$$

Let $\mathcal{S} = \overline{\sigma} \sharp \sigma$. Then

$$|(\mathbf{A}^*\mathbf{A})_{klkl} - \mathcal{S}(\rho\frac{\beta}{\alpha}k, \rho\frac{\alpha}{\beta}l)| = \mathcal{O}\left(\frac{1}{(\alpha\beta D)^2}\right).$$

Proof: We first look at $(\mathbf{A}^*\mathbf{A})_{klkl}$. The diagonal entries of $\mathbf{A}^*\mathbf{A}$ are

$$(\mathbf{A}^*\mathbf{A})_{klkl} = \sum_{k'l' \in \mathbb{Z}^2} |\langle \mathbf{L}_\sigma \psi_{kl}^t, \psi_{k'l'}^r \rangle|^2 \tag{36}$$

$$= \|\mathbf{L}_\sigma \psi_{kl}^t\|_2^2 \tag{37}$$

since $(\psi, \frac{1}{\rho}a, \frac{1}{\rho}b)$ is a tight Weyl–Heisenberg frame [Proposition (3.2)]

$$\|\mathbf{L}_\sigma \psi^t\|_2^2 \tag{38}$$

$$= \langle \mathbf{L}_\sigma \psi_{kl}^t, \mathbf{L}_\sigma \psi_{kl}^t \rangle \tag{39}$$

$$= \langle \overline{\sigma} \sharp \sigma, \mathcal{W}(\psi_{kl}^t, \psi_{kl}^t) \rangle$$

$$= \int_{\mathbb{R}^2} \mathcal{S}(x, \omega) \mathcal{W}(\psi^t, \psi^t)(x - \rho\frac{\beta}{\alpha}k, \omega - \rho\frac{\alpha}{\beta}l) d\omega dx. \tag{40}$$

Setting $\mathcal{S}' = \partial_x \partial_\omega \mathcal{S}$, by the Riemann–Lebesgue Lemma

$$\|\mathcal{S}'\|_\infty \le \iint |\hat{\mathcal{S}}(\omega, x)| d\omega dx$$

$$\le C \iint \frac{1}{\alpha\beta} e^{-\frac{\beta}{2}|\omega|-\frac{\alpha}{2}|x|} d\omega dx$$

$$= \frac{C}{(\alpha\beta)^2}.$$

We use Lemma 3.6 and the fact that $\iint \mathcal{W}(\psi, \psi)(\omega, x) d\omega dx = \|\psi\|_2^2 = 1$, cf. [14]

$$|\|\mathbf{L}_\sigma \psi^t\|_2^2 - \mathcal{S}(\rho\frac{\beta}{\alpha}k, \rho\frac{\alpha}{\beta}l)|$$

$$= |\int_{\mathbb{R}^2} \mathcal{S}(x, \omega) \mathcal{W}(\psi, \psi)(x - \rho\frac{\beta}{\alpha}k, \omega - \rho\frac{\alpha}{\beta}l) d\omega dx$$

$$- \mathcal{S}(\rho\frac{\alpha}{\beta}l, \rho\frac{\beta}{\alpha}k)|$$

$$= |\int_{\mathbb{R}^2} \mathcal{S}(x + \rho\frac{\beta}{\alpha}k, \omega + \rho\frac{\alpha}{\beta}l) \mathcal{W}(\psi, \psi)(x, \omega) d\omega dx$$

$$- \mathcal{S}(\rho\frac{\alpha}{\beta}l, \rho\frac{\beta}{\alpha}k)|$$

$$= |\int_{\mathbb{R}^2} [\mathcal{S}(x + \rho\frac{\beta}{\alpha}k, \omega + \rho\frac{\alpha}{\beta}l) - \mathcal{S}(\rho\frac{\alpha}{\beta}l, \rho\frac{\beta}{\alpha}k)]$$

$$\mathcal{W}(\psi, \psi)(x, \omega) d\omega dx|$$

$$\le \|\mathcal{S}'\|_\infty \int_{\mathbb{R}^2} (|x| + |\omega|) \mathcal{W}(\psi, \psi)(x, \omega)| d\omega dx$$

$$\le C\frac{1}{(\alpha\beta)^2} \int_{\mathbb{R}^2} (|x| + |\omega|) e^{-\frac{\pi s D}{4}|x|-\frac{\pi}{4s}D|\omega|} d\omega dx$$

$$= C\frac{1}{(\alpha\beta D)^2}.$$

These two bounds prove the lemma. ∎

Proof of Theorem 3.5: The estimate (25) follows now readily by applying the triangle inequality to the left-hand-side of (25), and then using Lemma 3.8 and Lemma 3.9. ∎

Remark: In the proof of this theorem we rely on using Weyl–Heisenberg systems. Instead we could have resorted to orthonormal Wilson bases [14], which do not suffer from the Balian–Low Theorem. However, it would have resulted in a less elegant relationship between eigenvalues and samples of $\mathcal{S}$. In particular, (28) and (29) would have to be replaced by more complicated expressions.

IV. FROM ESTIMATING EIGENVALUES TO ESTIMATING MUTUAL INFORMATION

For the time-invariant case, the mutual information is precisely captured by samples of the Fourier transform of the auto-correlation of the impulse response when one allows $T \to \infty$. At the core of this relationship is the fact that the (generalized) eigenvalues of the channel are directly linked to samples of the transfer function. It turns out that for our class of time-varying channels a similar connection is true in an approximate sense. Using the eigenvalue estimate from the previous section we will show that one can obtain an estimate of the mutual information via samples of the Fourier transform of the "twisted auto-convolved" spreading function. This is the contents of the following theorem.

Theorem 4.1 (Mutual Information Estimate): Assume that the spreading function $\hat{\sigma}$ in the system model satisfies

$$|\hat{\sigma}(\omega, x)| \le Ce^{-\beta|\omega|-\alpha|x|} \tag{41}$$

and the AWGN $n(t)$ has variance η^2. Let $\Phi_{T,W} = \{\phi_k\}_{k=1}^N$ be a set of orthonormal functions contained in $L^2(T, W, \varepsilon)$, where $N = (1 - \delta)(2TW + 1)$ for some $0 \le \delta < 1$. Let $\mathcal{I}_{\Phi_{T,W}}(x, y)$ denote the resulting mutual information of the system given in lines (10)–(14). Then there exist constants $0 < D, 1 < \rho$ and small constants $0 \le \delta_1, \delta_2$ such that

$$\left| \mathcal{I}_{\Phi_{T,W}}(x, y) - \sum_{k=0, l=-L}^{K, L} \log\left(1 + \frac{\mathcal{S}^+(\rho\frac{\beta}{\alpha}k, \rho\frac{\alpha}{\beta}l)}{\eta^2}\right) \right| \tag{42}$$

$$\le (2TW + 1)\left(\log\left(1 + \mathcal{O}\left(e^{-\frac{\rho}{2}(\beta+\alpha)} + \frac{1}{(\alpha\beta D)^2}\right)\right) \right) \tag{43}$$

$$+ \log\left(1 + \left(\frac{14\varepsilon}{\eta^2} + \frac{(14\varepsilon + \delta)}{\eta^2}\right)\|\mathcal{S}\|_{L^\infty(\mathbb{R})}\right) \tag{44}$$

$$+ \log\left(1 + \left(\frac{14\varepsilon}{\eta^2} + \frac{1 - \frac{(1-49\varepsilon^2)}{\rho^2} + \frac{\delta_1}{\rho}\frac{\beta}{\alpha} + \frac{\delta_2}{\rho}\frac{\alpha}{\beta}}{\eta^2}\right)\|\mathcal{S}\|_{L^\infty(\mathbb{R})}\right) \tag{45}$$

where $K = \frac{T}{\rho}\frac{\alpha}{\beta} - \delta_1$ and $L = \frac{2W}{\rho}\frac{\beta}{\alpha} - \delta_2$. The parameters D and ρ have the relationship that $D \to 0$ as $\rho \to 1$ and $\rho \to \infty$ as $D \to 1$. The numbers δ_1 and δ_2 depend on the parameters α, β and ε, but remain small as T and W increase.

Before we proceed to the proof of this theorem, it seems prudent to comment on the statement of this theorem and the various elements that come into play here.

Remark 1: In a nutshell, our theorem shows that

$$\mathcal{I}_{\Phi_{T,W}}(x, y) \approx \sum_{k=0, l=-L}^{K, L} \log\left(1 + \frac{\mathcal{S}^+(\rho\frac{\beta}{\alpha}k, \rho\frac{\alpha}{\beta}l)}{\eta^2}\right)$$

and quantifies rigorously in which sense this approximation is true. The error due to estimating the mutual information from the samples is given in (43) and is the conceptually more important one for this paper. The error in (44) results from the transition from the system $\Phi_{T,W}$ in $L^2(T, W, \varepsilon)$ to the PSWFs, and the error (45) is due to the fact that the number of the constructed Weyl–Heisenberg signals used is less than the number of PSWFs corresponding to the time-frequency region.

Remark 2: The factor ρ is necessary for our construction and is greater than 1, see Proposition 3.2 and the subsequent discussion. While taking ρ very close to 1 would make the error in (45) very small, it would increase the error in (43). We can, however, take ρ to be fairly close to 1, such as $\rho = 5/4$. This issue of the tradeoff between time-frequency localization and loss of dimensions in signal space has also been pointed out in [11].

We need the following lemma for the proof of Theorem 4.1.

Lemma 4.2: Let $\mathcal{S} = \overline{\sigma}\sharp\sigma$ and $S^+(x, \omega) = (S(x, \omega))^+$. Then

$$\left| \log(1 + \lambda_{k,l}(\mathbf{A}^*\mathbf{A})) - \log(1 + \mathcal{S}^+(\rho\frac{\beta}{\alpha}k, \rho\frac{\alpha}{\beta}l)) \right|$$
$$= \log\left(1 + \mathcal{O}\left(e^{-\frac{\rho}{2}(\beta+\alpha)} + \frac{1}{(\alpha\beta D)^2}\right)\right).$$

Proof: Using Lemmas 3.8 and 3.9

$$\left| \log(1 + \lambda_{k,l}(\mathbf{A}^*\mathbf{A})) - \log(1 + \mathcal{S}^+(\rho\frac{\beta}{\alpha}k, \rho\frac{\alpha}{\beta}l)) \right|$$
$$\leq \left| \log(1 + \lambda_{k,l}(\mathbf{A}^*\mathbf{A})) - \log(1 + (\mathbf{A}^*\mathbf{A})_{klkl}) \right|$$
$$+ \left| \log(1 + (\mathbf{A}^*\mathbf{A})_{klkl}) - \log(1 + \mathcal{S}^+(\rho\frac{\beta}{\alpha}k, \rho\frac{\alpha}{\beta}l)) \right|$$
$$= \log\left(1 + \mathcal{O}\left(e^{-\frac{\rho}{2}(\beta+\alpha)} + \frac{1}{(\alpha\beta D)^2}\right)\right).$$

∎

Proof of Theorem 4.1: Let $\mathbf{P}$ denote the projection of $L^2(\mathbb{R})$ onto the span of the $2TW + 1$ PSWFs corresponding to $[0, T] \times [-W, W]$. From (15) we obtain

$$\|\mathbf{P}f\|^2_{L^2(\mathbb{R})} \geq 1 - 49\varepsilon^2\|f\|^2_{L^2(\mathbb{R})} \qquad (46)$$

for all $f \in L^2(T, W, \varepsilon)$. We write $\mathbf{P}_\Phi$ for the projection onto the set $\{\phi_1, \ldots, \phi_N\}$ and G for the Gram matrix of $\{\mathbf{P}\phi_1, \ldots, \mathbf{P}\phi_N\}$, i.e.,

$$G_{i,j} = \langle \mathbf{P}\phi_j, \mathbf{P}\phi_i \rangle \quad i, j = 1, \ldots, N. \qquad (47)$$

Then $\text{rank}(\mathbf{P}\mathbf{P}_\Phi) = \text{rank}(G)$. Note that the diagonal entries of G are positive and, since $\{\phi_1, \ldots, \phi_N\}$ are orthonormal, that the eigenvalues of G have absolute value at most 1. By inequality (46)

$$\sum_{j=1}^{N} G_{j,j} = \sum_{j=1}^{N} \|\mathbf{P}\phi_j\|^2$$
$$\geq N(1 - 49\varepsilon^2)$$

so that $\text{rank}(G) \geq N(1 - 49\varepsilon^2)$. Therefore

$$\text{rank}(\mathbf{P}^\perp\mathbf{P}_\Phi) \leq \text{rank}(\mathbf{P}) - \text{rank}(\mathbf{P}\mathbf{P}_\Phi)$$
$$\leq (2TW + 1) - (1 - 49\varepsilon^2)N$$

and

$$\|\mathbf{P}_\Phi\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi - \mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\|_{\text{HS}}$$
$$= \| - \mathbf{P}_\Phi^\perp\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi + \mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi - \mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\|_{\text{HS}}$$
$$= \| - \mathbf{P}_\Phi^\perp\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi + \mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}$$
$$\qquad\qquad - \mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi^\perp - \mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\|_{\text{HS}}$$
$$\leq \|\mathbf{P}_\Phi^\perp\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi\|_{\text{HS}} + \|\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi^\perp\|_{HS}$$
$$\leq 2\|\mathbf{P}_\Phi^\perp\mathbf{P}\|_{\text{HS}}\|\mathbf{L}_\sigma\|^2$$
$$\leq 2\,\text{rank}(\mathbf{P}_\Phi^\perp\mathbf{P})\|\mathbf{L}_\sigma\|^2$$
$$\leq 2((2TW + 1) - (1 - 49\varepsilon^2)N)\|\mathbf{L}_\sigma\|^2.$$

If $2TW + 1 > N$, then set $\lambda_j(\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}_\Phi) = 0$ for $N < j \leq 2TW + 1$. Let π be a permutation of the integers $1, \ldots, 2TW + 1$. Then

$$\left| \sum_{j=1}^{2TW+1} \log\left(1 + \frac{\lambda_j(\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}_\Phi)}{\eta^2}\right) \right.$$
$$\left. - \sum_{j=1}^{2TW+1} \log\left(1 + \frac{\lambda_{\pi(j)}(\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P})}{\eta^2}\right) \right|$$
$$\leq \sum_{j=1}^{2TW+1} \log\left(1 + \frac{|\lambda_j(\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}_\Phi) - \lambda_{\pi(j)}(\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P})|}{\eta^2}\right)$$
$$\leq \sum_{j=1}^{2TW+1} \log\left(1 + \frac{|\lambda_j(\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}_\Phi) - \lambda_j(\mathbf{P}_\Phi\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi)|}{\eta^2}\right.$$
$$\left. + \frac{|\lambda_j(\mathbf{P}_\Phi\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi) - \lambda_{\pi(j)}(\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P})|}{\eta^2}\right).$$

We consider the first eigenvalue difference in the expression above. Applying Theorem A.46 in [28] we obtain

$$|\lambda_j(\mathbf{P}_\Phi\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi) - \lambda_j(\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P})|$$
$$\leq \|\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}_\Phi - \mathbf{P}_\Phi\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi\|. \quad (48)$$

Let $\mathbf{P}_\Phi f = u + v$ where $u \in \text{range}\,\mathbf{P}$ and $v \in \text{range}\,\mathbf{P}^\perp$. Then

$$\|(\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}_\Phi - \mathbf{P}_\Phi\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}\mathbf{P}_\Phi)f\|$$
$$= \|\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*(u + v) - \mathbf{P}_\Phi\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*u\|$$
$$\leq \|(\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^* - \mathbf{P}_\Phi\mathbf{P}\mathbf{L}_\sigma\mathbf{L}_\sigma^*)u\| + \|\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*v\|$$
$$= \|\mathbf{P}_\Phi\mathbf{P}^\perp\mathbf{L}_\sigma\mathbf{L}_\sigma^*u\| + \|\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*v\|$$
$$\leq \|\mathbf{P}_\Phi\mathbf{P}^\perp\|\|\mathbf{L}_\sigma\|^2\|u\| + \|\mathbf{P}_\Phi\mathbf{L}_\sigma\mathbf{L}_\sigma^*\mathbf{P}^\perp\mathbf{P}_\Phi f\|$$
$$\leq 7\varepsilon\|\mathbf{L}_\sigma\|^2\|f\| + 7\varepsilon\|\mathbf{L}_\sigma\|^2\|f\|$$
$$\leq 14\varepsilon\|\mathbf{L}_\sigma\|^2\|f\|$$

where we have used (15) in the penultimate step. Hence

$$|\lambda_j(\mathbf{P}_\Phi \mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P} \mathbf{P}_\Phi) - \lambda_j(\mathbf{P}_\Phi \mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P})| \le 14\varepsilon \|\mathbf{L}_\sigma\|^2. \tag{49}$$

Concerning the second difference of eigenvalues recall that according to Theorem A.37 of [28] there exists a permutation π such that

$$\sum_{j=1}^{2TW+1} |\lambda_j(\mathbf{P}_\Phi \mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P} \mathbf{P}_\Phi) - \lambda_{\pi(j)}(\mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P})|^2$$

$$\le \|\mathbf{P}_\Phi \mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P} \mathbf{P}_\Phi - \mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P}\|_{\mathrm{HS}}^2$$
$$\le 4((2TW+1) - (1 - 49\varepsilon^2)N)^2 \|\mathbf{L}_\sigma\|^4. \tag{50}$$

Using (49), (50) and the concavity of the log function, we compute

$$\sum_{j=1}^{2TW+1} \log \left(1 + \frac{|\lambda_j(\mathbf{P}_\Phi \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P}_\Phi) - \lambda_j(\mathbf{P}_\Phi \mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P} \mathbf{P}_\Phi)|}{\eta^2} \right.$$
$$\left. + \frac{|\lambda_j(\mathbf{P}_\Phi \mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P} \mathbf{P}_\Phi) - \lambda_{\pi(j)}(\mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P})|}{\eta^2} \right)$$

$$\le \sum_{j=1}^{2TW+1} \log \left(1 + \frac{14\varepsilon \|\mathbf{L}_\sigma\|^2}{\eta^2} \right.$$
$$\left. + \frac{|\lambda_j(\mathbf{P}_\Phi \mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P} \mathbf{P}_\Phi) - \lambda_{\pi(j)}(\mathbf{P}_\Phi \mathbf{P} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P})|}{\eta^2} \right)$$

$$\le \sum_{j=1}^{2TW+1} \log \left(1 + \frac{14\varepsilon \|\mathbf{L}_\sigma\|^2}{\eta^2} \right.$$
$$\left. + \frac{2((2TW+1) - (1 - 49\varepsilon^2)N)\|\mathbf{L}_\sigma\|^2}{\eta^2(2TW+1)} \right).$$

We will return to (50) twice, taking N to be the cardinality of $\Phi_{T,W}$ and of our constructed set.

We look at the system $\{\psi_{k,l}^t\}$ from Section III-A. The signal $\psi_{k,l}^t$ is exponentially localized around the point $(\rho \frac{\alpha}{\beta} l, \rho \frac{\beta}{\alpha} k)$. We select those signals that are contained in $L^2(T, W, \varepsilon)$. For some positive constants δ_1 and δ_2, these are those signals with indices $0 \le k \le \frac{T}{\rho} \frac{\alpha}{\beta} - \delta_1$ and $0 \le |l| \le \frac{W}{\rho} \frac{\beta}{\alpha} - \delta_2$. We set $K = \frac{T}{\rho} \frac{\alpha}{\beta} - \delta_1$ and $L = \frac{W}{\rho} \frac{\beta}{\alpha} - \delta_2$. We denote by $\mathbf{P}_{K,L}$ the projection operator from $L^2(R)$ onto the span of $\{\psi_{k,l}^t\}_{k=0,l=-L}^{K,L}$. Now we use (50) twice: once with $N = (1-\delta)(2TW+1)$ for the cardinality of the set $\Phi_{T,W}$, as assumed in the statement of the theorem, and once for $\Psi_{K,L}$, where the cardinality satisfies

$$K(2L+1) \ge \frac{2TW+1}{\rho^2} - \delta_1 \frac{2W}{\rho} \frac{\beta}{\alpha} - \delta_2 \frac{T}{\rho} \frac{\alpha}{\beta}.$$

The arguments above then yield

$$\left| \sum_{j=1}^{2TW+1} \log \left(1 + \frac{\lambda_j(\mathbf{P}_\Phi \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P}_\Phi)}{\eta^2} \right) \right. \tag{51}$$

$$\left. - \sum_{j=1}^{2TW+1} \log \left(1 + \frac{\lambda_{\pi(j)}(\mathbf{P}_{K,L} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P}_{K,L})}{\eta^2} \right) \right|$$

$$\le (2TW+1)\left(\log \left(1 + \left(\frac{14\varepsilon}{\eta^2} + \frac{2(49\varepsilon^2 + \delta)}{\eta^2} \right) \|\mathbf{L}_\sigma\|^2 \right) \right.$$

$$\left. + \log \left(1 + \left(\frac{14\varepsilon}{\eta^2} + \frac{2(1 - \frac{(1-49\varepsilon^2)}{\rho^2} + \frac{\delta_1}{\rho} \frac{\beta}{\alpha} + \frac{\delta_2}{\rho} \frac{\alpha}{\beta})}{\eta^2} \right) \|\mathbf{L}_\sigma\|^2 \right) \right). \tag{52}$$

The estimation of the eigenvalues $\mathbf{P}_{K,L} \mathbf{L}_\sigma \mathbf{L}_\sigma^* \mathbf{P}_{K,L}$ is given by the Lemmas 3.8 and 4.2. Applying these two lemmas together with inequality (52) complete the proof of the theorem. ∎

REFERENCES

[1] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, 1949.

[2] R. G. Gallager, *Information Theory and Reliable Communication*. Hoboken, NJ: Wiley, 1968.

[3] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[4] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun.*, vol. COM-11, no. 4, pp. 360–393, Apr. 1963.

[5] W. C. Jakes Jr., *Microwave Mobile Communications*. Hoboken, NJ: Wiley, 1974.

[6] M. Medard and R. G. Gallager, "The issue of spreading in multipath time-varying channels," in *Proc. IEEE Vehicular Technology Conf.*, 1995, pp. 1–5.

[7] S. N. Diggavi, "Analysis of multicarrier transmission in time-varying channels," in *Proc. IEEE Int. Conf. Communications*, Jun. 1997, vol. 3, pp. 1191–1195.

[8] S. Barbarossa and A. Scaglione, "On the capacity of linear time-varying channels," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Mar. 1999, vol. 5, pp. 2627–2630.

[9] W. Kozek and A. F. Molisch, "Nonorthogonal pulse shapes for multicarrier communications in doubly dispersive channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 10, pp. 1579–1589, Oct. 1998.

[10] K. Liu, T. Kadous, and A. Sayeed, "Orthogonal time-frequency signaling over doubly dispersive channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2583–2603, Nov. 2004.

[11] G. Durisi, H. Bölcskei, and S. Shamai (Shitz), "Capacity of underspread noncoherent WSSUS fading channels under peak signal constraints," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 156–160.

[12] W. Kozek, "On the transfer function calculus for underspread ltv channels," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 219–223, Jan. 1997.

[13] W. Kozek and A. F. Molisch, "On the eigenstructure of underspread wssus channels," in *Proc. 1st IEEE Workshop on Signal Processing Advances in Wireless Communications*, Paris, France, Apr. 1997, pp. 325–328.

[14] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Boston, MA: Birkhäuser, 2001.

[15] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty – III: The dimension of the space of essentially time- and band-limited signals," *Bell Syst. Tech. J.*, vol. 41, no. 1, pp. 1295–1336, 1962.

[16] G. Matz and F. Hlawatsch, "Time-frequency transfer function calculus (symbolic calclulus) of linear time-varying systems (linear operators) based on a generalized underspread theory," *J. Math. Phys.*, vol. 39, pp. 4041–4070, Aug. 1998.

[17] Y. I. Lyubarskii, "Frames in the Bargmann space of entire functions," *Adv. Soviet Math.*, vol. 429, pp. 107–113, 1992.

[18] K. Seip, "Density theorems for sampling and interpolation in the Bargmann-Fock space," *Bull. AMS*, vol. 26, pp. 322–328, 1992.

[19] K. Seip and R. Wallsten, "Density theorems for sampling and interpolation in the Bargmann-Fock space II," *J. Reine angewandte Math.*, vol. 429, pp. 107–113, 1992.

[20] J. Wexler and S. Raz, "Discrete Gabor expansions," *Signal Process.*, vol. 21, no. 3, pp. 207–221, Nov. 1990.

[21] A. J. E. M. Janssen, "Duality and biorthogonality for Weyl–Heisenberg frames," *J. Four. Anal. Appl.*, vol. 1, no. 4, pp. 403–436, 1995.

[22] I. Daubechies, H. Landau, and Z. Landau, "Gabor time-frequency lattices and the Wexler-Raz identity," *J. Four. Anal. Appl.*, vol. 1, no. 4, pp. 437–478, 1995.

[23] H. Bölcskei and A. J. E. M. Janssen, "Gabor frames, unimodularity, and window decay," *J. Four. Anal. Appl.*, vol. 6, no. 3, pp. 255–276, 2000.

[24] T. Strohmer and S. Beaver, "Optimal OFDM design for time-frequency dispersive channels," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1111–1122, Jul. 2003.

[25] D. Slepian, "Some comments on Fourier analysis, uncertainty and modelling," *SIAM Rev.*, vol. 25, pp. 379–393, 1983.

[26] T. Strohmer, "Approximation of dual Gabor frames, window decay, and wireless communications," *Appl. Comput. Harmon. Anal.*, vol. 11, no. 2, pp. 243–262, 2001.

[27] T. Strohmer, "Pseudodifferential operators and Banach algebras in mobile communications," *Appl. Comput. Harmon. Anal.*, vol. 20, no. 2, pp. 237–249, 2006.

[28] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, second ed. New York: Springer, 2010, Springer Series in Statistics.

Brendan Farrell (M'08) biography not available at the time of publication.

Thomas Strohmer, biography not available at the time of publication.