(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 August 2017 (31.08.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/147439 A1

(51) International Patent Classification:
H04B 1/10 (2006.01) *H04L 12/28* (2006.01)
H04B 1/46 (2006.01) *H04L 12/56* (2006.01)
H04B 1/69 (2011.01) *H04W 24/08* (2009.01)
H04B 1/707 (2011.01)

(21) International Application Number:
PCT/US2017/019376

(22) International Filing Date:
24 February 2017 (24.02.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/299,985 25 February 2016 (25.02.2016) US
62/303,318 3 March 2016 (03.03.2016) US

(71) Applicant: **COHERE TECHNOLOGIES** [US/US]; 2550 Walsh Avenue, Suite 150, Santa Clara, CA 95051 (US).

(72) Inventors: **HADANI, Ronny**; 2550 Walsh Avenue, Suite 150, Santa Clara, CA 95051 (US). **RAKIB, Shlomo, Selim**; 2550 Walsh Avenue, Suite 150, Santa Clara, CA 95051 (US). **MONK, Anton**; 2550 Walsh Avenue, Suite 150, Santa Clara, CA 95051 (US). **TSATSANIS, Michail**; 2550 Walsh Avenue, Suite 150, Santa Clara, CA 95051 (US). **HEBRON, Yoav**; 2550 Walsh Avenue, Suite 150, Santa Clara, CA 95051 (US).

(74) Agent: **SATHE, Vinay**; Perkins Coie LLP, P.O. Box 1247, Seattle, WA 98111-1247 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,

[Continued on next page]

(54) Title: REFERENCE SIGNAL PACKING FOR WIRELESS COMMUNICATIONS

(57) Abstract: In a wireless communication network, pilot signals are transmitted over a wireless communication channel by determining a maximum delay spread for a transmission channel, determining a maximum Doppler frequency spread for the transmission channel, and allocating a set of transmission resources in a time-frequency domain to a number of pilot signals based on the maximum delay spread and the maximum Doppler frequency spread.

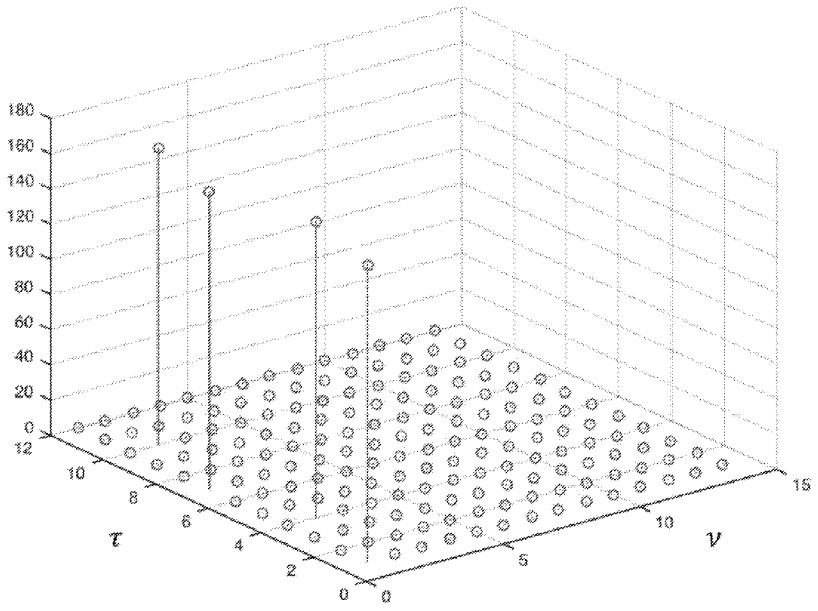


FIG. 29

WO 2017/147439 A1

LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

REFERENCE SIGNAL PACKING FOR WIRELESS COMMUNICATIONS

CROSS-REFERENCE TO RELATED APPLICATION(S)

[1] This patent document claims the benefit of priority from U. S. Provisional Patent Applications 62/303,318, filed March 3, 2016, and 62/299,985, filed February 25, 2016. All of the aforementioned patent applications are incorporated by reference herein in their entirety.

TECHNICAL FIELD

[2] This document relates to the field of telecommunications, in particular, estimation and compensation of impairments in telecommunications data channels.

BACKGROUND

[3] Due to an explosive growth in the number of wireless user devices and the amount of wireless data that these devices can generate or consume, current wireless communication networks are fast running out of bandwidth to accommodate such a high growth in data traffic and provide high quality of service to users.

[4] Various efforts are underway in the telecommunication industry to come up with next generation of wireless technologies that can keep up with the demand on performance of wireless devices and networks.

SUMMARY

[5] Various techniques for pilot packing, e.g., assigning transmission resources to a number of pilot signals for transmission, are disclosed. The disclosed techniques provide various operational advantages, including, for example, pilot packing by staggering pilots to achieve improved separation among the pilots, providing a number of pilot that is commensurate with the target delay-Doppler spread to be combated in the wireless channel, thereby optimally using the available transmission bandwidth, and so on.

[6] In one example aspect, a wireless communication method is disclosed. Using the method, pilot signals are transmitted over a wireless communication channel by

determining a maximum delay spread for a transmission channel, determining a maximum Doppler frequency spread for the transmission channel, and allocating a set of transmission resources in a time-frequency domain to a number of pilot signals based on the maximum delay spread and the maximum Doppler frequency spread.

[7] In another aspect, a method of wireless communication includes determining a maximum delay spread for a transmission channel, determining a maximum Doppler frequency spread for the transmission channel, determining a number of pilot signals that can be transmitted using a set of two-dimensional transmission resources at least based on the maximum delay spread and the maximum Doppler frequency spread, allocating the set of transmission resources from a two-dimensional set of resources to the number of pilot, and transmitting the pilot signals over a wireless communication channel using transmission resources.

[8] In yet another aspect, a wireless communication apparatus comprising a memory, a processor and a transmitter is disclosed. The wireless communication apparatus may implement any of the above-described methods and other associated techniques described in the present document.

[9] These, and other aspects, are described in greater detail in the present document.

BRIEF DESCRIPTION OF THE DRAWINGS

[10] Figure 1 shows an example trajectory of Time Varying Impulse Response for Accelerating Reflector.

[11] Figure 2 shows an example of Delay-Doppler Representation for an Accelerating Reflector Channel.

[12] Figure 3 depicts example Levels of Abstraction: Signaling over the (i) actual channel with a signaling waveform (ii) the time-frequency Domain (iii) the delay-Doppler Domain.

[13] Figure 4 shows examples of notation Used to Denote Signals at Various Stages of Transmitter and Receiver.

[14] Figure 5 depicts an example of a conceptual Implementation of the Heisenberg Transform in the Transmitter and the Wigner Transform in the Receiver.

[15] Figure 6 shows an example of cross-correlation between $g_{tr}(t)$ and $g_r(t)$ for OFDM Systems.

[16] Figure 7 shows an example of Information Symbols in the Information (Delay-Doppler) Domain (Right), and Corresponding Basis Functions in the Time-Frequency Domain (Left).

[17] Figure 8 shows a One Dimensional Multipath Channel Example: (i) Sampled Frequency Response at $\Delta f = 1$ Hz (ii) Periodic Fourier Transform with Period $1/\Delta f = 1$ sec (iii) Sampled Fourier Transform with Period $1/\Delta f$ and Resolution $1/M\Delta f$.

[18] Figure 9 shows a One Dimensional Doppler Channel Example: (i) Sampled Frequency Response at $\mathrm{T_s} = 1$ sec (ii) Periodic Fourier Transform with Period $1/\mathrm{T_s} = 1$ Hz (iii) Sampled Fourier Transform with Period $1/\mathrm{T_s}$ and Resolution $1/\mathrm{NT_s}$.

[19] Figure 10 depicts an example of a time-Varying Channel Response in the Time-Frequency Domain

[20] Figure 11 depicts an example SDFT of Channel response – (τ, ν) Delay-Doppler Domain.

[21] Figure 12 depicts an example SFFT of Channel Response – Sampled (τ, ν) Delay-Doppler Domain.

[22] Figure 13 depicts an example of Transformation of the Time-Frequency Plane to the Doppler-Delay Plane.

[23] Figure 14 depicts an example of a Discrete Impulse in the OTFS Domain Used for Channel Estimation.

[24] Figure 15 shows an example of Different Basis Functions, Assigned to Different Users, Span the Whole Time-Frequency Frame.

[25] Figure 16 shows an example embodiment of multiplexing three users in the Time-Frequency Domain.

[26] Figure 17 shows an example embodiment of Multiplexing three users in the Time-Frequency Domain with Interleaving.

[27] Figure 18 shows an example of an OTFS architecture block diagram.

[28] Figure 19 shows an example of a t-f pilot lattice (1902) superimposed on a data lattice (1904) when N=14, M=2.

[29] Figure 20 shows examples of OTFS based pilots. [a] shows 10 pilots on the Delay-Doppler plane associated with the data plane, [b] shows the samples of the real portion of $P3 = \delta\left(\frac{0.1}{df}, \frac{0.0357}{dt}\right)$ on the t-f data lattice of Figure 19, [c] shows the representation of the same 10 pilots on the Delay-Doppler plane associated with the coarser pilot lattice of Figure 19 (N=14, M=2), and d shows the samples of the real portion of P3 on the t-f pilot lattice.

[30] Figure 21 shows an example of 4x2 pilots in a staggered structure on Delay Doppler plane.

[31] Figure 22 shows an example of a pilot lattice with N=28 and M=1 superimposed on the data lattice.

[32] Figure 23 shows an example embodiment with 12 pilot lattices with N=28, M=12 each, superimposed on the data lattice.

[33] Figure 24 shows an example embodiment with 2 pilot lattices with N=28 and M=2 superimposed on the data lattice.

[34] Figure 25 shows an example of UL (2502) and DL (2504) pilot sample points on a data lattice (2506).

[35] Figure 26 shows an example of packing 10x4=40 pilots on the Delay-Doppler plane associated with the pilot lattice of N=28, M=1 (2504 points of the lattice of Figure 25)

[36] Figure 27 shows an example of a torus.

[37] Figure 28 shows an example of a sampled Delay-Doppler data plane showing the absolute value of an instantiation of the time-domain cyclic shift portion of the 4 LTE UL DM RSs.

[38] Figure 29 shows an example staggered version of the 4 LTE DM RSs shown in Figure 28.

[39] Figure 30 shows a graphical depiction of an example of average SNR of estimated 4 ETU-50 channels (using MMSE interpolation) for the 4 DM RSs of Figure 28 (3004) and Figure 29 (3002) when receiver input SNR is 50dB.

[40] Figure 31 shows an example of pilots' sample points (3102) on a data lattice (3104).

[41] Figure 32 shows an example communication network in which the disclosed technology can be embodied.

[42] Figure 33 shows a flowchart of an example method of wireless communication.

[43] Figure 34 is a block diagram of an example of a wireless communication apparatus.

[44] Figure 35 shows a flowchart of an example method of wireless communication.

[45] Figure 36 is a block diagram of an example of a wireless communication apparatus that can be used for embodying some techniques disclosed in this patent document.

DETAILED DESCRIPTION

[46] Section headings are used in this document to help improve readability and do not limit scope of the technology discussed in each section only to that section. Furthermore, for ease of explanation, a number of simplifying assumptions have been made. Although these simplifying assumptions are intended to help convey ideas, they are not intended to be limiting. Some of these simplifying assumptions are:

[47] 1. Introduction

[48] 4G wireless networks have served the public well, providing ubiquitous access to the internet and enabling the explosion of mobile apps, smartphones and sophisticated data intensive applications like mobile video. This continues an honorable tradition in the

evolution of cellular technologies, where each new generation brings enormous benefits to the public, enabling astonishing gains in productivity, convenience, and quality of life.

[49] Looking ahead to the demands that the ever increasing and diverse data usage is putting on the network, it is becoming clear to the industry that current 4G networks will not be able to support the foreseen needs in the near term future. The data traffic volume has been and continues to increase exponentially. AT&T reports that its network has seen an increase in data traffic of 100,000% in the period 2007-2015. Looking into the future, new applications like immersive reality, and remote robotic operation (tactile internet) as well as the expansion of mobile video are expected to overwhelm the carrying capacity of current systems. One of the goals of 5G system design is to be able to economically scale the network to 750 Gbps per sq. Km in dense urban settings, something that is not possible with today's technology.

[50] Beyond the sheer volume of data, the quality of data delivery will need to improve in next generation systems. The public has become accustomed to the ubiquity of wireless networks and is demanding a wireline experience when untethered. This translates to a requirement of 50+ Mbps everywhere (at the cell edge), which will require advanced interference mitigation technologies to be achieved.

[51] Another aspect of the quality of user experience is mobility. Current systems' throughput is dramatically reduced with increased mobile speeds due to Doppler effects which evaporate MIMO capacity gains. Future 5G systems aim to not only increase supported speeds up to 500 Km/h for high speed trains and aviation, but also support a host of new automotive applications for vehicle-to-vehicle and vehicle-to-infrastructure communications.

[52] While the support of increased and higher quality data traffic is necessary for the network to continue supporting the user needs, carriers are also exploring new applications that will enable new revenues and innovative use cases. The example of automotive and smart infrastructure applications discussed above is one of several. Others include the deployment of public safety ultra-reliable networks, the use of cellular networks to support the sunset of the PSTN, etc. The biggest revenue opportunity however, is arguably the deployment of large number of internet connected devices, also known as the

internet of things (IoT). Current networks however are not designed to support a very large number of connected devices with very low traffic per device.

[53] In summary, current LTE networks cannot achieve the cost/performance targets required to support the above objectives, necessitating a new generation of networks involving advanced PHY technologies. There are numerous technical challenges that will have to be overcome in 5G networks as discussed next.

[54] **1.1 4G Technical Challenged**

[55] In order to enable machine-to-machine communications and the realization of the internet of things, the spectral efficiency for short bursts will have to be improved, as well as the energy consumption of these devices (allowing for 10 years operation on the equivalent of 2 AA batteries). In current LTE systems, the network synchronization requirements place a burden on the devices to be almost continuously on. In addition, the efficiency goes down as the utilization per UE (user equipment, or mobile device) goes down. The PHY requirements for strict synchronization between UE and eNB (Evolved Node B, or LTE base station) will have to be relaxed, enabling a re-designing of the MAC for IoT connections that will simplify transitions from idle state to connected state.

[56] Another important use case for cellular IoT (CIoT) is deep building penetration to sensors and other devices, requiring an additional 20dB or more of dynamic range. 5G CIoT solutions should be able to coexist with the traditional high-throughput applications by dynamically adjusting parameters based on application context.

[57] The path to higher spectral efficiency points towards a larger number of antennas. A lot of research work has gone into full dimension and massive MIMO architectures with promising results. However, the benefits of larger MIMO systems may be hindered by the increased overhead for training, channel estimation and channel tracking for each antenna. A PHY that is robust to channel variations will be needed as well as innovative ways to reduce the channel estimation overhead.

[58] Robustness to time variations is usually connected to the challenges present in high Doppler use cases such as in vehicle-to-infrastructure and vehicle-to-vehicle automotive applications. With the expected use of spectrum up to 60GHz for 5G

applications, this Doppler impact will be an order of magnitude greater than with current solutions. The ability to handle mobility at these higher frequencies would be extremely valuable.

[59] 1.2 OTFS Based Solution

[60] OTFS is a modulation technique that modulates each information (e.g., QAM) symbol onto one of a set of two dimensional (2D) orthogonal basis functions that span the bandwidth and time duration of the transmission burst or packet. The modulation basis function set is specifically derived to best represent the dynamics of the time varying multipath channel.

[61] OTFS transforms the time-varying multipath channel into a time invariant delay-Doppler two dimensional convolution channel. In this way, it eliminates the difficulties in tracking time-varying fading, for example in high speed vehicle communications.

[62] OTFS increases the coherence time of the channel by orders of magnitude. It simplifies signaling over the channel using well studied AWGN codes over the average channel SNR. More importantly, it enables linear scaling of throughput with the number of antennas in moving vehicle applications due to the inherently accurate and efficient estimation of channel state information (CSI). In addition, since the delay-doppler channel representation is very compact, OTFS enables massive MIMO and beamforming with CSI at the transmitter for four, eight, and more antennas in moving vehicle applications. The CSI information needed in OTFS is a fraction of what is needed to track a time varying channel.

[63] In deep building penetration use cases, one QAM symbol may be spread over multiple time and/or frequency points. This is a key technique to increase processing gain and in building penetration capabilities for CIoT deployment and PSTN replacement applications. Spreading in the OTFS domain allows spreading over wider bandwidth and time durations while maintaining a stationary channel that does not need to be tracked over time.

[64] Loose synchronization: CoMP and network MIMO techniques have stringent clock synchronization requirements for the cooperating eNBs. If clock frequencies are not well synchronized, the UE will receive each signal from each eNB with an apparent "Doppler" shift. OTFS's reliable signaling over severe Doppler channels can enable CoMP deployments while minimizing the associated synchronization difficulties.

[65] These benefits of OTFS will become apparent once the basic concepts behind OTFS are understood. There is a rich mathematical foundation of OTFS that leads to several variations; for example it can be combined with OFDM or with multicarrier filter banks. In this paper we navigate the challenges of balancing generality with ease of understanding as follows:

[66] In Section 2 we start by describing the wireless Doppler multipath channel and its effects on multicarrier modulation.

[67] In Section 3, we develop OTFS as a modulation that matches the characteristics of the time varying channel. We show OTFS as consisting of two processing steps:

[68] A step that allows transmission over the time frequency plane, via orthogonal waveforms generated by translations in time and/or frequency. In this way, the (time-varying) channel response is sampled over points of the time-frequency plane.

[69] A pre-processing step using carefully crafted orthogonal functions employed over the time-frequency plane, which translate the time-varying channel in the time-frequency plane, to a time-invariant one in the new information domain defined by these orthogonal functions.

[70] In Section 4 we develop some more intuition on the new modulation scheme by exploring the behavior of the channel in the new modulation domain in terms of coherence, time and frequency resolution etc.

[71] In Sections 5 and 6 we explore aspects of channel estimation in the new information domain and multiplexing multiple users respectively, while in Section 7 we address complexity and implementation issues.

[72] In Sections 8, we provide some performance results and we put the OTFS modulation in the context of cellular systems, discuss its attributes and its benefits for 5G systems.

[73] **2. The Wireless Channel**

[74] The multipath fading channel is commonly modeled in the baseband as a convolution channel with a time varying impulse response

$$r(t) = \int \hbar(\tau, t)s(t-\tau)d\tau \qquad (1)$$

[75] where $s(t)$ and $r(t)$ represent the complex baseband channel input and output respectively and where $\hbar(\tau, t)$ is the complex baseband time varying channel response.

[76] This representation, while general, does not give us insight into the behavior and variations of the time varying impulse response. A more useful and insightful model, which is also commonly used for Doppler multipath doubly fading channels is

$$r(t) = \iint h(\tau, \nu)e^{j2\pi\nu(t-\tau)}s(t-\tau)d\nu d\tau \qquad (2)$$

[77] In this representation, the received signal is a superposition of reflected copies of the transmitted signal, where each copy is delayed by the path delay τ, frequency shifted by the Doppler shift ν and weighted by the time-invariant delay-Doppler impulse response $h(\tau, \nu)$ for that τ and ν. In addition to the intuitive nature of this representation, Eq. (2) maintains the generality of Eq. (1). In other words it can represent complex Doppler trajectories, like accelerating vehicles, reflectors etc. This can be seen if we express the time varying impulse response as a Fourier expansion with respect to the time variable t

$$\hbar(\tau, t) = \int h(\tau, \nu)e^{j2\pi\nu t}dt \qquad (3)$$

[78] Substituting (3) in (1) we obtain Eq. (2) after some manipulation[1]. More specifically, we obtain $y(t) = \iint e^{j2\pi\nu\tau}h(\tau, \nu)e^{j2\pi\nu(t-\tau)}x(t-\tau)d\nu d\tau$ which differs from the

[1] More specifically we obtain $y(t) = \iint e^{j2\pi\nu\tau}h(\tau, \nu)e^{j2\pi\nu(t-\tau)}x(t-\tau)d\nu d\tau$ which differs from **Error! Reference source not found.** by an exponential factor; however, we can

above equations by an exponential factor; however, we can absorb the exponential factor in the definition of the impulse response $h(\tau, \nu)$ making the two representations equivalent.

[79] As an example, Figure 1 shows the time-varying impulse response for an accelerating reflector in the (τ, t) coordinate system, while Figure 2 shows the same channel represented as a time invariant impulse response in the (τ, ν) coordinate system.

[80] An important feature revealed by these two figures is how compact the (τ, ν) representation is compared to the (τ, t) representation. This has important implications for channel estimation, equalization and tracking as will be discussed later.

[81] Notice that while $h(\tau, \nu)$ is, in fact, time-invariant, the operation on $s(t)$ is still time varying, as can be seen by the effect of the explicit complex exponential function of time in Eq. (2). The technical efforts in this paper are focused on developing a modulation scheme based on appropriate choice of orthogonal basis functions that render the effects of this channel truly time-invariant in the domain defined by those basis functions. Let us motivate those efforts with a high level outline of the structure of the proposed scheme here.

[82] Let us consider a set of orthonormal basis functions $\phi_{\tau,\nu}(t)$ indexed by τ, ν which are orthogonal to translation and modulation, i.e.,

$$\phi_{\tau,\nu}(t - \tau_0) = \phi_{\tau+\tau_0,\nu}(t)$$
$$e^{j2\pi\nu_0 t}\phi_{\tau,\nu}(t) = \phi_{\tau,\nu-\nu_0}(t) \qquad (4)$$

[83] and let us consider the transmitted signal as a superposition of these basis functions

$$s(t) = \iint x(\tau, \nu)\, \phi_{\tau,\nu}(t) d\tau d\nu \qquad (5)$$

---

absorb the exponential factor in the definition of the impulse response $h(\tau, \nu)$ making the two representations equivalent.

[84] where the weights $x(\tau, \nu)$ represent the information bearing signal to be transmitted. After the transmitted signal of (5) goes through the time varying channel of Eq. (2) we obtain a superposition of delayed and modulated versions of the basis functions, which due to (4) results in

$$\begin{aligned} r(t) &= \iint h(\tau, \nu) e^{j2\pi\nu(t-\tau)} s(t-\tau) d\nu d\tau \\ &= \iint \phi_{\tau,\nu}(t)\{h(\tau, \nu) * x(\tau, \nu)\} d\tau d\nu \end{aligned} \qquad (6)$$

[85] where $*$ denotes two dimensional convolution. Eq. *(6)* can be thought of as a generalization of the derivation of the convolution relationship for linear time invariant systems, using one dimensional exponentials as basis functions. Notice that the term in brackets can be recovered at the receiver by matched filtering against each basis function $\phi_{\tau,\nu}(t)$. In this way a two dimensional channel relationship is established in the (τ, ν) domain $y(\tau, \nu) = h(\tau, \nu) * x(\tau, \nu)$, where $y(\tau, \nu)$ is the receiver two dimensional matched filter output. Notice also, that in this domain the channel is described by a time invariant two-dimensional convolution.

[86] A final different interpretation of the wireless channel will also be useful in what follows. Let us consider $s(t)$ and $r(t)$ as elements of the Hilbert space of square integrable functions $\mathcal{H}$. Then Eq. (2) can be interpreted as a linear operator on $\mathcal{H}$ acting on the input $s(t)$, parametrized by the impulse response $h(\tau, \nu)$, and producing the output $r(t)$

$$r = \Pi_h(s): \qquad s(t) \in \mathcal{H} \xrightarrow{\Pi_h(\cdot)} r(t) \in \mathcal{H} \qquad (7)$$

[87] Notice that although the operator is linear, it is not time-invariant. In the no Doppler case, i.e., if $h(\nu, \tau) = h(0, \tau)\delta(\nu)$, then Eq. (2) reduces to a time invariant convolution. Also notice that while for time invariant systems the impulse response is parameterized by one dimension, in the time varying case we have a two dimensional impulse response. While in the time invariant case the convolution operator produces a superposition of delays of the input $s(t)$, (hence the parameterization is along the one dimensional delay axis) in the time varying case we have a superposition of delay-and-modulate operations as seen in Eq. (2) (hence the parameterization is along the two

dimensional delay and Doppler axes). This is a major difference which makes the time varying representation non-commutative (in contrast to the convolution operation which is commutative), and complicates the treatment of time varying systems.

[88] The important point of Eq. (7) is that the operator $\Pi_h(\cdot)$ can be compactly parametrized in a two dimensional space $h(\tau, \nu)$, providing an efficient, time invariant description of the channel. Typical channel delay spreads and Doppler spreads are a very small fraction of the symbol duration and subcarrier spacing of multicarrier systems.

[89] In the mathematics literature, the representation of time varying systems of (2) and (7) is called the Heisenberg representation [1]. It can actually be shown that every linear operator (7) can be parameterized by some impulse response as in (2).

[90] **3. OTFS modulation over the Doppler multipath channel**

[91] The time variation of the channel introduces significant difficulties in wireless communications related to channel acquisition, tracking, equalization and transmission of channel state information (CSI) to the transmit side for beamforming and MIMO processing. In this paper, we develop a modulation domain based on a set of orthonormal basis functions over which we can transmit the information symbols, and over which the information symbols experience a static, time invariant, two dimensional channel for the duration of the packet or burst transmission. In that modulation domain, the channel coherence time is increased by orders of magnitude and the issues associated with channel fading in the time or frequency domain in SISO or MIMO systems are significantly reduced.

[92] Orthogonal Time Frequency Space (OTFS) modulation is comprised of a cascade of two transformations. The first transformation maps the two dimensional plane where the information symbols reside (and which we call the delay-Doppler plane) to the time frequency plane. The second one transforms the time frequency domain to the waveform time domain where actual transmitted signal is constructed. This transform can be thought of as a generalization of multicarrier modulation schemes.

[93] Figure 3 provides a pictorial view of the two transformations that constitute the OTFS modulation. It shows at a high level the signal processing steps that are required at

the transmitter and receiver. It also includes the parameters that define each step, which will become apparent as we further expose each step. Further, Figure 4 shows a block diagram of the different processing stages at the transmitter and receiver and establishes the notation that will be used for the various signals.

[94] We start our description with the transform which relates the waveform domain to the time-frequency domain.

[95] **3.1 The Heisenberg Transform**

[96] Our purpose in this section is to construct an appropriate transmit waveform which carries information provided by symbols on a grid in the time-frequency plane. Our intent in developing this modulation scheme is to transform the channel operation to an equivalent operation on the time-frequency domain with two important properties:

[97] The channel is orthogonalized on the time-frequency grid.

[98] The channel time variation is simplified on the time-frequency grid and can be addressed with an additional transform.

[99] Fortunately, these goals can be accomplished with a scheme that is very close to well-known multicarrier modulation techniques, as explained next. We will start with a general framework for multicarrier modulation and then give examples of OFDM and multicarrier filter bank implementations.

[100] Let us consider the following components of a time frequency modulation:

[101] A lattice or grid on the time frequency plane, that is a sampling of the time axis with sampling period T and the frequency axis with sampling period Δf.

$$\Lambda = \{(nT, m\Delta f),\ n, m \in \mathbb{Z}\} \qquad (8)$$

[102] A packet burst with total duration NT secs and total bandwidth $M\Delta f$ Hz

[103] A set of modulation symbols $X[n, m],\ n = 0, \ldots, N-1,\ m = 0, \ldots, M-1$ we wish to transmit over this burst

[104] A transmit pulse $g_{tr}(t)$ with the property[2] of being orthogonal to translations by T and modulations by Δf

$$\begin{aligned} &< g_{tr}(t), g_{tr}(t-nT)e^{j2\pi m\Delta f(t-nT)} > \\ &= \int g_{tr}^{*}(t)g_{r}(t-nT)e^{j2\pi m\Delta f(t-nT)}dt \\ &= \delta(m)\delta(n) \end{aligned} \qquad (9)$$

[105] Given the above components, the time-frequency modulator is a Heisenberg operator on the lattice Λ, that is, it maps the two dimensional symbols $X[n.m]$ to a transmitted waveform, via a superposition of delay-and-modulate operations on the pulse waveform $g_{tr}(t)$

$$s(t) = \sum_{m=-M/2}^{M/2-1} \sum_{n=0}^{N-1} X[n,m]g_{tr}(t-nT)e^{j2\pi m\Delta f(t-nT)} \qquad (10)$$

[106] More formally

$$x = \Pi_X(g_{tr}) \quad : \qquad g_{tr}(t) \in \mathcal{H} \xrightarrow{\Pi_X(\cdot)} y(t) \in \mathcal{H} \qquad (11)$$

[107] where we denote by $\Pi_X(\cdot)$ the "discrete" Heisenberg operator, parameterized by discrete values $X[n,m]$.

[108] Notice the similarity of (11) with the channel equation (7). This is not by coincidence, but rather because we apply a modulation effect that mimics the channel effect, so that the end effect of the cascade of modulation and channel is more tractable at the receiver. It is not uncommon practice; for example, linear modulation (aimed at time invariant channels) is in its simplest form a convolution of the transmit pulse $g(t)$ with a delta train of QAM information symbols sampled at the Baud rate T.

$$s(t) = \sum_{n=0}^{N-1} X[n]g(t-nT) \qquad (12)$$

[109] In our case, aimed at the time varying channel, we convolve-and-modulate the transmit pulse (c.f. the channel Eq. (2)) with a two dimensional delta train which samples the time frequency domain at a certain Baud rate and subcarrier spacing.

---

[2] This orthogonality property is required if the receiver uses the same pulse as the transmitter. We will generalize it to a bi-orthogonality property in later sections.

[110] The sampling rate in the time-frequency domain is related to the bandwidth and time duration of the pulse $g_{tr}(t)$ namely its time-frequency localization. In order for the orthogonality condition of (9) to hold for a frequency spacing Δf, the time spacing must be $T \geq 1/\Delta f$. The critical sampling case of $T = 1/\Delta f$ is generally not practical and refers to limiting cases, for example to OFDM systems with cyclic prefix length equal to zero or to filter banks with $g_{tr}(t)$ equal to the ideal Nyquist pulse.

[111] Some examples are as follows:

[112] Example 1: OFDM Modulation: Let us consider an OFDM system with M subcarriers, symbol length T_{OFDM}, cyclic prefix length T_{CP} and subcarrier spacing $1/T_{OFDM}$. If we substitute in Equation (10) symbol duration $T = T_{OFDM} + T_{CP}$, number of symbols $N = 1$, subcarrier spacing $\Delta f = 1/T_{OFDM}$ and $g_{tr}(t)$ a square window that limits the duration of the subcarriers to the symbol length T

$$g_{tr}(t) = \begin{cases} 1/\sqrt{T - T_{CP}}, & -T_{CP} < t < T - T_{CP} \\ 0, & else \end{cases} \qquad (13)$$

[113] then we obtain the OFDM formula[3]

$$x(t) = \sum_{m=-M/2}^{M/2-1} X[n,m] g_{tr}(t) e^{j2\pi m \Delta f t} \qquad (14)$$

[114] Example 2: Single Carrier Modulation: Equation (10) reduces to single carrier modulation if we substitute $M = 1$ subcarrier, T equal to the Baud period and $g_{tr}(t)$ equal to a square root raised cosine Nyquist pulse.

[115] Example 3: Multicarrier Filter Banks (MCFB): Equation (10) describes a MCFB if $g_{tr}(t)$ is a square root raised cosine Nyquist pulse with excess bandwith α, T is equal to the Baud period and $\Delta f = (1 + \alpha)/T$.

[116] Expressing the modulation operation as a Heisenberg transform as in Eq. (11) may be counterintuitive. We usually think of modulation as a transformation of the modulation symbols $X[m,n]$ to a transmit waveform $s(t)$. The Heisenberg transform instead, uses $X[m,n]$ as weights/parameters of an operator that produces $s(t)$ when

[3] Technically, the pulse of Eq. (13) is not orthonormal but is orthogonal to the receive filter (where the CP samples are discarded) as we will see shortly.

applied to the prototype transmit filter response $g_{tr}(t)$ – c.f. Eq. (11). While counterintuitive, this formulation is useful in pursuing an abstraction of the modulation-channel-demodulation cascade effects in a two dimensional domain where the channel can be described as time invariant.

[117] We next turn our attention to the processing on the receiver side needed to go back from the waveform domain to the time-frequency domain. Since the received signal has undergone the cascade of two Heisenberg transforms (one by the modulation effect and one by the channel effect), it is natural to inquire what the end-to-end effect of this cascade is. The answer to this question is given by the following result:

[118] **Proposition 1:** Let two Heisenberg transforms as defined by Eqs. (7), (2) be parametrized by impulse responses $h_1(\tau, \nu)$, $h_2(\tau, \nu)$ and be applied in cascade to a waveform $g(t) \in \mathcal{H}$. Then

$$\Pi_{h_2}\left(\Pi_{h_1}(g(t))\right) = \Pi_h(g(t)) \qquad (15)$$

[119] where $h(\tau, \nu) = h_2(\tau, \nu) \odot h_1(\tau, \nu)$ is the "twisted" convolution of $h_1(\tau, \nu)$, $h_2(\tau, \nu)$ defined by the following convolve-and-modulate operation

$$h(\tau, \nu) = \iint h_2(\tau', \nu') h_1(\tau - \tau', \nu - \nu') e^{j2\pi\nu'(\tau - \tau')t}\, d\tau' d\nu' \qquad (16)$$

[120] **Proof:** See Appendix 0.

[121] Applying the above result to the cascade of the modulation and channel Heisenberg transforms of (11) and (7), we can show that the received signal is given by the Heisenberg transform

$$r(t) = \Pi_f(g_{tr}(t)) + v(t) = \iint f(\tau, \nu) e^{j2\pi\nu(t-\tau)} g_{tr}(t - \tau) d\nu d\tau + v(t) \qquad (17)$$

[122] where $v(t)$ is additive noise and $f(\tau, \nu)$, the impulse response of the combined transform, is given by the twisted convolution of $X[n, m]$ and $h(\tau, \nu)$

$$\begin{aligned} f(\tau, \nu) &= h(\tau, \nu) \odot X[n, m] \\ &= \sum_{m=-M/2}^{M/2-1} \sum_{n=0}^{N-1} X[n, m] h(\tau - nT, \nu - m\Delta f) e^{j2\pi(\nu - m\Delta f)nT} \end{aligned} \qquad (18)$$

[123] This result can be considered an extension of the single carrier modulation case, where the received signal through a time invariant channel is given by the convolution of the QAM symbols with a composite pulse, that pulse being the convolution of the transmitter pulse and the channel impulse response.

[124] With this result established we are ready to examine the receiver processing steps.

[125] **3.2 Receiver processing and the Wigner transform**

[126] Typical communication system design dictates that the receiver performs a matched filtering operation, taking the inner product of the received waveform with the transmitter pulse, appropriately delayed or otherwise distorted by the channel. In our case, we have used a collection of delayed and modulated transmit pulses, and we need to perform a matched filter on each one of them. Figure 5 provides a conceptual view of this processing. On the transmitter, we modulate a set of M subcarriers for each symbol we transmit, while on the receiver we perform matched filtering on each of those subcarrier pulses. We define a receiver pulse $g_r(t)$ and take the inner product with a collection of delayed and modulated versions of it. The receiver pulse $g_r(t)$ is in many cases identical to the transmitter pulse, but we keep the separate notation to cover some cases where it is not (most notably in OFDM where the CP samples have to be discarded).

[127] While this approach will yield the sufficient statistics for data detection in the case of an ideal channel, a concern can be raised here for the case of non-ideal channel effects. In this case, the sufficient statistics for symbol detection are obtained by matched filtering with the channel-distorted, information-carrying pulses (assuming that the additive noise is white and Gaussian). In many well designed multicarrier systems however (e.g., OFDM and MCFB), the channel distorted version of each subcarrier signal is only a scalar version of the transmitted signal, allowing for a matched filter design that is independent of the channel and uses the original transmitted subcarrier pulse. We will make these statements more precise shortly and examine the required conditions for this to be true.

[128] Figure 5 is only a conceptual illustration and does not point to the actual implementation of the receiver. Typically this matched filtering is implemented in the digital

domain using an FFT or a polyphase transform for OFDM and MCFB respectively. In this paper we are rather more interested in the theoretical understanding of this modulation. To this end, we will consider a generalization of this matched filtering by taking the inner product $< g_r(t-\tau)e^{j2\pi\nu(t-\tau)}, r(t) >$ of the received waveform with the delayed and modulated versions of the receiver pulse for arbitrary time and frequency offset (τ, ν). While this is not a practical implementation, it allows us to view the operations of Figure 5 as a two dimensional sampling of this more general inner product.

[129] Let us define the inner product

$$A_{g_r,r}(\tau,\nu) = < g_r(t-\tau)e^{j2\pi\nu(t-\tau)}, r(t) > = \int g_r^*(t-\tau)e^{-j2\pi\nu(t-\tau)}r(t)dt \quad (19)$$

[130] The function $A_{g_r,r}(\tau,\nu)$ is known as the cross-ambiguity function in the radar and math communities and yields the matched filter output if sampled at $\tau = nT,\ \nu = m\Delta f$ (on the lattice Λ), i.e.,

$$Y[n,m] = A_{g_r,r}(\tau,\nu)|_{\tau=nT,\nu=m\Delta f} \quad (20)$$

[131] In the math community, the ambiguity function is related to the inverse of the Heisenberg transform, namely the Wigner transform. Figure 5 provides an intuitive feel for that, as the receiver appears to invert the operations of the transmitter[4].

[132] The key question here is what the relationship is between the matched filter output $Y[n,m]$ (or more generally $Y(\tau,\nu)$) and the transmitter input $X[n,m]$. We have already established in (17) that the input to the matched filter $r(t)$ can be expressed as a Heisenberg representation with impulse response $f(\tau,\nu)$ (plus noise). The output of the matched filter then has two contributions

$$Y(\tau,\nu) = A_{g_r,r}(\tau,\nu) = A_{g_r,[\Pi_f(g_{tr})+\nu]}(\tau,\nu) = A_{g_r,\Pi_f(g_{tr})}(\tau,\nu) + A_{g_r,\nu}(\tau,\nu) \quad (21)$$

---

[4] More formally, if we take the cross-ambiguity or the transmit and receive pulses $A_{g_r,g_{tr}}(\tau,\nu)$, and use it as the impulse response of the Heisenberg operator, then we obtain the orthogonal cross-projection operator

$$\Pi_{A_{g_r,g_{tr}}}(y(t)) = g_{tr}(t) < g_r(t), y(t) >$$

In words, the coefficients that come out of the matched filter, if used in a Heisenberg representation, will provide the best approximation to the original $y(t)$ in the sense of minimum square error.

[133] The last term is the contribution of noise, which we will denote $V(\tau,\nu) = A_{g_r,v}(\tau,\nu)$. The first term on the right hand side is the matched filter output to the (noiseless) input comprising of a superposition of delayed and modulated versions of the transmit pulse. We next establish that this term can be expressed as the twisted convolution of the two dimensional impulse response $f(\tau,\nu)$ with the cross-ambiguity function (or two dimensional cross correlation) of the transmit and receive pulses.

[134] The following theorem summarizes the key result.

[135] **Theorem 1:** (Fundamental time-frequency domain channel equation). If the received signal can be expressed as

$$\Pi_f(g_{tr}(t)) = \iint f(\tau,\nu)e^{j2\pi\nu(t-\tau)}g_{tr}(t-\tau)d\nu d\tau \qquad (22)$$

[136] Then the cross-ambiguity of that signal with the receive pulse $g_{tr}(t)$ can be expressed as

$$A_{g_r,\Pi_f(g_{tr})}(\tau,\nu) = f(\tau,\nu) \odot A_{g_r,g_{tr}}(\tau,\nu) \qquad (23)$$

□

[137] **Proof:** See Appendix 0.

[138] Recall from (18) that $f(\tau,\nu) = h(\tau,\nu) \odot X[n,m]$, that is, the composite impulse response is itself a twisted convolution of the channel response and the modulation sumbols.

[139] Substituting $f(\tau,\nu)$ from (18) into (21) we obtain the end-to-end channel description in the time frequency domain

$$\begin{aligned} Y(\tau,\nu) &= A_{g_r,\Pi_r(g_{tr})}(\tau,\nu) + V(\tau,\nu) \\ &= h(\tau,\nu) \odot X[n,m] \odot A_{g_r,g_{tr}}(\tau,\nu) + V(\tau,\nu) \end{aligned} \qquad (24)$$

[140] where $V(\tau,\nu)$ is the additive noise term. Eq. (24) provides an abstraction of the time varying channel on the time-frequency plane. It states that the matched filter output at any time and frequency point (τ,ν) is given by the delay-Doppler impulse response of the channel twist-convolved with the impulse response of the modulation

operator twist-convolved with the cross-ambiguity (or two dimensional cross correlation) function of the transmit and receive pulses.

[141] Evaluating Eq. (24) on the lattice Λ we obtain the matched filter output modulation symbol estimates

$$\hat{X}[m,n] = Y[n,m] = Y(\tau,\nu)\big|_{\tau=nT,\nu=m\Delta f} \qquad (25)$$

[142] In order to get more intuition on Equations (24), (25) let us first consider the case of an ideal channel, i.e., $h(\tau,\nu) = \delta(\tau)\delta(\nu)$. In this case by direct substitution we get the convolution relationship

$$Y[n,m] = \sum_{m'=-M/2}^{M/2-1} \sum_{n'=0}^{N-1} X[n',m'] A_{g_r,g_{tr}}\big((n-n')T, (m-m')\Delta f\big) + V[m,n] \qquad (26)$$

[143] In order to simplify Eq. (26) we will use the orthogonality properties of the ambiguity function. Since we use a different transmit and receive pulses we will modify the orthogonality condition on the design of the transmit pulse we stated in *(9)* to a bi-orthogonality condition

$$\begin{aligned} &< g_{tr}(t), g_r(t-nT)e^{j2\pi m\Delta f(t-nT)} > \\ &\quad = \int g_{tr}^*(t) g_r(t-nT) e^{j2\pi m\Delta f(t-nT)} dt \\ &\quad = \delta(m)\delta(n) \end{aligned} \qquad (27)$$

[144] Under this condition, only one term survives in (26) and we obtain

$$Y[n,m] = X[n,m] + V[n,m] \qquad (28)$$

[145] where $V[n,m]$ is the additive white noise. Eq. (28) shows that the matched filter output does recover the transmitted symbols (plus noise) under ideal channel conditions. Of more interest of course is the case of non-ideal time varying channel effects. We next show that even in this case, the channel orthogonalization is maintained (no intersymbol or intercarrier interference), while the channel complex gain distortion has a closed form expression.

[146] The following theorem summarizes the result as a generalization of (28).

[147] **Theorem 2:** (End-to-end time-frequency domain channel equation):

[148] If $h(\tau,\nu)$ has finite support bounded by (τ_{max},ν_{max}) and if $A_{g_r,g_{tr}}(\tau,\nu)=0$ for $\tau\in(nT-\tau_{max},nT+\tau_{max})$, $\nu\in(m\Delta f-\nu_{max},m\Delta f+\nu_{max})$, that is, the ambiguity function bi-orthogonality property of (27) is true in a neighborhood of each grid point $(m\Delta f,nT)$ of the lattice Λ at least as large as the support of the channel response $h(\tau,\nu)$, then the following equation holds

$$Y[n,m]=H[n,m]X[n,m]$$
$$H[n,m]=\iint h(\tau,\nu)e^{j2\pi\nu nT}e^{-j2\pi(\nu+m\Delta f)\tau}d\nu d\tau \qquad (29)$$

[149] If the ambiguity function is only approximately bi-orthogonal in the neighborhood of Λ (by continuity), then (29) is only approximately true.

[150] **Proof:** See Appendix 0.

[151] Eq. (29) is a fundamental equation that describes the channel behavior in the time-frequency domain. It is the basis for understanding the nature of the channel and its variations along the time and frequency dimensions.

[152] Some observations are now in order on Eq. (29). As mentioned before, there is no interference across $X[n,m]$ in either time n or frequency m.

[153] The end-to-end channel distortion in the modulation domain is a (complex) scalar that needs to be equalized

[154] If there is no Doppler, i.e. $h(\tau,\nu)=h(\tau,0)\delta(\nu)$, then Eq. (29) becomes

$$\begin{aligned}Y[n,m]&=X[n,m]\int h(\tau,0)e^{-j2\pi m\Delta f\tau}d\tau \\ &=X[n,m]H(0,m\Delta f)\end{aligned} \qquad (30)$$

[155] which is the well-known multicarrier result, that each subcarrier symbol is multiplied by the frequency response of the time invariant channel evaluated at the frequency of that subcarrier.

[156] If there is no multipath, i.e. $h(\tau,\nu) = h(0,\nu)\delta(\tau)$, then Eq. (29) becomes

$$Y[n,m] = X[n,m]\int h(\nu,0)e^{j2\pi\nu nT}d\tau \qquad (31)$$

[157] Notice that the fading each subcarrier experiences as a function of time nT has a complicated expression as a weighted superposition of exponentials. This is a major complication in the design of wireless systems with mobility like LTE; it necessitates the transmission of pilots and the continuous tracking of the channel, which becomes more difficult the higher the vehicle speed or Doppler bandwidth is.

[158] We close this section with some examples of this general framework.

[159] Example 3: (OFDM modulation). In this case the fundamental transmit pulse is given by *(13)* and the fundamental receive pulse is

$$g_r(t) = \begin{cases} 0 & -T_{CP} < t < 0 \\ \dfrac{1}{\sqrt{T - T_{CP}}} & 0 < t < T - T_{CP} \\ 0 & else \end{cases} \qquad (32)$$

[160] i.e., the receiver zeroes out the CP samples and applies a square window to the symbols comprising the OFDM symbol. It is worth noting that in this case, the bi-orthogonality property holds exactly along the time dimension. Figure 6 shows the cross correlation between the transmit and receive pulses of *(13*) and (32). Notice that the cross correlation is exactly equal to one and zero in the vicinity of zero and $\pm T$ respectively, while holding those values for the duration of T_{CP}. Hence, as long as the support of the channel on the time dimension is less than T_{CP} the bi-orthogonality condition is satisfied along the time dimension. Across the frequency dimension the condition is only approximate, as the ambiguity takes the form of a sinc function as a function of frequency and the nulls are not identically zero for the whole support of the Doppler spread.

[161] Example 4: (MCFB modulation). In the case of multicarrier filter banks $g_{tr}(\mathrm{t}) = g_r(\mathrm{t}) = g(t)$. There are several designs for the fundamental pulse $g(t)$. A square root raised cosine pulse provides good localization along the frequency dimension at the expense of less localization along the time dimension. If T is much larger than the support

of the channel in the time dimension, then each subchannel sees a flat channel and the bi-orthogonality property holds approximately.

[162] In summary, in this section we described the one of the two transforms that define OTFS. We explained how the transmitter and receiver apply appropriate operators on the fundamental transmit and receive pulses and orthogonalize the channel according to Eq. (29). We further saw via examples how the choice of the fundamental pulse affect the time and frequency localization of the transmitted modulation symbols and the quality of the channel orthogonalization that is achieved. However, Eq. (29) shows that the channel in this domain, while free of intersymbol interference, suffers from fading across both the time and the frequency dimensions via a complicated superposition of linear phase factors.

[163] In the next section we will start from Eq. (29) and describe the second transform that defines OTFS; we will show how that transform defines an information domain where the channel does not fade in either dimension.

[164] 3.3 The 2D OTFS Transform

[165] Notice that the time-frequency response $H[n,m]$ in (29) is related to the channel delay-Doppler response $h(\tau,\nu)$ by an expression that resembles a Fourier transform. However, there are two important differences: (i) the transform is two dimensional (along delay and Doppler) and (ii) the exponentials defining the transforms for the two dimensions have opposing signs. Despite these difficulties, Eq. (29) points in the direction of using complex exponentials as basis functions on which to modulate the information symbols; and only transmit on the time-frequency domain the superposition of those modulated complex exponential bases. This is the approach we will pursue in this section.

[166] This is akin to the SC-FDMA modulation scheme, where in the frequency domain we transmit a superposition of modulated exponentials (the output of the DFT preprocessing block). The reason we pursue this direction is to exploit Fourier transform properties and translate a multiplicative channel in one Fourier domain to a convolution channel in the other Fourier domain.

[167] Given the difficulties of Eq. (29) mentioned above we need to develop a suitable version of Fourier transform and associated sampling theory results. Let us start with the following definitions:

[168] **Definition 1:** Symplectic Discrete Fourier Transform: Given a square summable two dimensional sequence $X[m,n] \in \mathbb{C}(\Lambda)$ we define

$$x(\tau,\nu) = \sum_{m,n} X[n,m] e^{-j2\pi(\nu nT - \tau m\Delta f)} \triangleq SDFT(X[n,m]) \quad (33)$$

[169] Notice that the above 2D Fourier transform (known as the Symplectic Discrete Fourier Transform in the math community) differs from the more well known Cartesian Fourier transform in that the exponential functions across each of the two dimensions have opposing signs. This is necessary in this case, as it matches the behavior of the channel equation.

[170] Further notice that the resulting $x(\tau,\nu)$ is periodic with periods $(1/\Delta f, 1/T)$. This transform defines a new two dimensional plane, which we will call the delay-Doppler plane, and which can represent a max delay of $1/\Delta f$ and a max Doppler of $1/T$. A one dimensional periodic function is also called a function on a circle, while a 2D periodic function is called a function on a torus (or donut). In this case $x(\tau,\nu)$ is defined on a torus Z with circumferences (dimensions) $(1/\Delta f, 1/T)$.

[171] The periodicity of $x(\tau,\nu)$ (or sampling rate of the time-frequency plane) also defines a lattice on the delay-Doppler plane, which we will call the reciprocal lattice

$$\Lambda^{\perp} = \left\{ \left(m\frac{1}{\Delta f}, n\frac{1}{T} \right), \ n,m \in \mathbb{Z} \right\} \quad (34)$$

[172] The points on the reciprocal lattice have the property of making the exponent in (33), an integer multiple of 2π.

[173] The inverse transform is given by:

$$X[n,m] = \frac{1}{c}\int_0^{\frac{1}{\Delta f}}\int_0^{\frac{1}{T}} x(\tau,\nu)e^{j2\pi(\nu nT-\tau m\Delta f)}d\nu\, d\tau \triangleq SDFT^{-1}(x(\tau,\nu)) \quad (35)$$

[174] where $c = T\Delta f$.

[175] We next define a sampled version of $x(\tau,\nu)$. In particular, we wish to take M samples on the delay dimension (spaced at $1/\mathrm{M}\Delta f$) and N samples on the Doppler dimension (spaced at $1/\mathrm{NT}$). More formally we define a denser version of the reciprocal lattice

$$\Lambda_0^{\perp} = \left\{\left(m\frac{1}{M\Delta f}, n\frac{1}{NT}\right),\ \ n,m \in \mathbb{Z}\right\} \quad (36)$$

[176] So that $\Lambda^{\perp} \subseteq \Lambda_0^{\perp}$. We define discrete periodic functions on this dense lattice with period $(1/\Delta f, 1/T)$, or equivalently we define functions on a discrete torus with these dimensions

$$\mathrm{Z}_0^{\perp} = \left\{\left(m\frac{1}{M\Delta f}, n\frac{1}{NT}\right), \quad m = 0,\ldots,M-1, \quad n = 0,\ldots N-1,\right\} \quad (37)$$

[177] These functions are related via Fourier transform relationships to discrete periodic functions on the lattice Λ, or equivalently, functions on the discrete torus

$$Z_0 = \{(nT, m\Delta f), \quad m = 0,\ldots,M-1, \quad n = 0,\ldots N-1,\} \quad (38)$$

[178] We wish to develop an expression for sampling Eq. (33) on the lattice of (37). First, we start with the following definition.

[179] **Definition 2:** Symplectic Finite Fourier Transform: If $X_p[k,l]$ is periodic with period (N,M), then we define

$$x_p[m,n] = \sum_{k=0}^{N-1}\sum_{l=-\frac{M}{2}}^{\frac{M}{2}-1} X_p[k,l]e^{-j2\pi(\frac{nk}{N}-\frac{ml}{\mathrm{M}})} \triangleq SFFT(X[k,l]) \quad (39)$$

[180] Notice that $x_p[m,n]$ is also periodic with period $[M,N]$ or equivalently, it is defined on the discrete torus $Z_0^{\perp}$. Formally, the $SFFT(X[n,m])$ is a linear transformation from $\mathbb{C}(Z_0) \rightarrow \mathbb{C}(Z_0^{\perp})$.

[181] Let us now consider generating $x_p[m,n]$ as a sampled version of (33), i.e., $x_p[m,n] = x[m,n] = x(\tau,\nu)|_{\tau=\frac{m}{M\Delta f},\nu=\frac{n}{NT}}$. Then we can show that (39) still holds where $X_p[m,n]$ is a periodization of $X[n,m]$ with period (N,M)

$$X_p[n,m] = \sum_{l,k=-\infty}^{\infty} X[n-kN, m-lM] \qquad (40)$$

[182] This is similar to the well-known result that sampling in one Fourier domain creates aliasing in the other domain.

[183] The inverse discrete (symplectic) Fourier transform is given by

$$X_p[n,m] = \frac{1}{MN}\sum_{l,k} x[l,k]e^{j2\pi\left(\frac{nk}{N}-\frac{ml}{M}\right)} \triangleq SFFT^{-1}(x[l,k]) \qquad (41)$$

[184] where $l = 0,\ldots,M-1,\ k = 0,\ldots,N-1$. If the support of $X[n,m]$ is time-frequency limited to Z_0 (no aliasing in (40)), then $X_p[n,m] = X[n,m]$ for $n,m \in Z_0$, and the inverse transform (41) recovers the original signal.

[185] In the math community, the SDFT is called “discrete” because it represents a signal using a discrete set of exponentials, while the SFFT is called “finite” because it represents a signal using a finite set of exponentials.

[186] Arguably the most important property of the symplectic Fourier transform is that it transforms a multiplicative channel effect in one domain to a circular convolution effect in the transformed domain. This is summarized in the following proposition:

[187] **Proposition 2:** Let $X_1[n,m] \in \mathbb{C}(Z_0)$, $X_2[n,m] \in \mathbb{C}(Z_0)$ be periodic 2D sequences. Then

$$SFFT(X_1[n,m] * X_2[n,m]) = SFFT(X_1[n,m]) \cdot SFFT(X_2[n,m]) \qquad (42)$$

[188] where $*$ denotes two dimensional circular convolution.

[189] **Proof:** See Appendix 0.

[190] With this framework established we are ready to define the OTFS modulation.

[191] **Discrete OTFS modulation**: Consider a set of NM QAM information symbols arranged on a 2D grid $x[l,k]$, $k=0,\ldots,N-1$, $l=0,\ldots,M-1$ we wish to transmit. We will consider $x[l,k]$ to be two dimensional periodic with period $[N,M]$. Further, assume a multicarrier modulation system defined by

[192] A lattice on the time frequency plane, that is a sampling of the time axis with sampling period T and the frequency axis with sampling period Δf (c.f. Eq. *(8)*).

[193] A packet burst with total duration NT secs and total bandwidth $M\Delta f$ Hz.

[194] Transmit and receive pulses $g_{tr}(t)$, $g_{tr}(t) \in L_2(\mathbb{R})$ satisfying the bi-orthogonality property of (27).

[195] A transmit windowing square summable function $W_{tr}[n,m] \in \mathbb{C}(\Lambda)$ multiplying the modulation symbols in the time-frequency domain

[196] A set of modulation symbols $X[n,m]$, $n=0,\ldots,N-1$, $m=0,\ldots,M-1$ related to the information symbols $x[k,l]$ by a set of basis functions $b_{k,l}[n,m]$

$$X[n,m] = \frac{1}{MN} W_{tr}[n,m] \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} x[l,k]\, b_{k,l}[n,m] \qquad (43)$$

$$b_{k,l}[n,m] = e^{j2\pi(\frac{ml}{M} - \frac{nk}{N})}$$

[197] where the basis functions $b_{k,l}[n,m]$ are related to the inverse symplectic Fourier transform (c.f., Eq. (41))

[198] Given the above components, we define the discrete OTFS modulation via the following two steps

$$X[n,m] = W_{tr}[n,m]\, SFFT^{-1}(x[k,l]) \qquad (44)$$

$$s(t) = \Pi_X(g_{tr}(t))$$

[199] The first equation in (44) describes the OTFS transform, which combines an inverse symplectic transform with a widowing operation. The second equation describes

the transmission of the modulation symbols $X[n,m]$ via a Heisenberg transform of $g_{tr}(t)$ parameterized by $X[n,m]$. More explicit formulas for the modulation steps are given by Equations (41) and *(10)*.

[200] While the expression of the OTFS modulation via the symplectic Fourier transform reveals important properties, it is easier to understand the modulation via Eq. (43), that is, transmitting each information symbol $x[k,l]$ by modulating a 2D basis function $b_{k,l}[n,m]$ on the time-frequency plane.

[201] Figure 7 visualizes this interpretation by isolating each symbol in the information domain and showing its contribution to the time-frequency modulation domain. Of course the transmitted signal is the superposition of all the symbols on the right (in the information domain) or all the basis functions on the left (in the modulation domain).

[202] Figure 7 uses the trivial window $W_{tr}[n,m] = 1$ for all $n = 0, \ldots, N-1$, $m = -\frac{M}{2}, \ldots \frac{M}{2} - 1$ and zero else. This may seem superfluous but there is a technical reason for this window: recall that $SFFT^{-1}(x[k,l])$ is a periodic sequence that extends to infinite time and bandwidth. By applying the window we limit the modulation symbols to the available finite time and bandwidth. The window in general could extend beyond the period of the information symbols $[M,N]$ and could have a shape different from a rectangular pulse. This would be akin to adding cyclic prefix/suffix in the dimensions of both time and frequency with or without shaping. The choice of window has implications on the shape and resolution of the channel response in the information domain as we will discuss later. It also has implications on the receiver processing as the potential cyclic prefix/suffix has to either be removed or otherwise handled as we see next.

[203] **Discrete OTFS demodulation:** Let us assume that the transmitted signal $s(t)$ undergoes channel distortion according to (7), (2) yielding $r(t)$ at the receiver. Further, let the receiver employ a receive windowing square summable function $W_r[n,m]$. Then, the demodulation operation consists of the following steps:

[204] Matched filtering with the receive pulse, or more formally, evaluating the ambiguity function on Λ (Wigner transform) to obtain estimates of the time-frequency modulation symbols

$$Y[n,m] = A_{g_r,y}(\tau,\nu)\big|_{\tau=nT,\nu=m\Delta f} \qquad (45)$$

[205] windowing and periodization of $Y[n,m]$

$$Y_w[n,m] = W_r[n,m]Y[n,m]$$
$$Y_p[n,m] = \sum_{k,l=-\infty}^{\infty} Y_w[n-kN, m-l\mathrm{M}] \qquad (46)$$

[206] and applying the symplectic Fourier transform on the periodic sequence $Y_p[n,m]$

$$\hat{x}[l,k] = y[l,k] = SFFT(Y_p[n,m]) \qquad (47)$$

[207] The first step of the demodulation operation can be interpreted as a matched filtering operation on the time-frequency domain as we discussed earlier. The second step is there to ensure that the input to the SFFT is a periodic sequence. If the trivial window is used, this step can be skipped. The third step can also be interpreted as a projection of the time-frequency modulation symbols on the orthogonal basis functions

$$\hat{x}[l,k] = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \hat{X}(n,m)\,\mathrm{b}^*_{k,l}(n,m)$$
$$\mathrm{b}^*_{k,l}(n,m) = e^{-j2\pi(\frac{lm}{L}-\frac{kn}{K})} \qquad (48)$$

[208] The discrete OTFS modulation defined above points to efficient implementation via discrete-and-periodic FFT type processing. However, it does not provide insight into the time and bandwidth resolution of these operations in the context of two dimensional Fourier sampling theory. We next introduce the continouse OTFS modulation and relate the more practical discrete OTFS as a sampled version of the continuous modulation.

[209] **Continuous OTFS modulation**: Consider a two dimensional periodic function $x(\tau,\nu)$ with period $[1/\Delta f, 1/T]$ we wish to transmit; the choice of the period may seem arbitrary at this point, but it will become clear after the discussion in the next section. Further, assume a multicarrier modulation system defined by

[210] A lattice on the time frequency plane, that is a sampling of the time axis with sampling period T and the frequency axis with sampling period Δf (c.f. Eq. (8)).

[211] Transmit and receive pulses $g_{tr}(t)$, $g_{tr}(t) \in L_2(\mathbb{R})$ satisfying the bi-orthogonality property of (27)

[212] A transmit windowing function $W_{tr}[n,m] \in \mathbb{C}(\Lambda)$ multiplying the modulation symbols in the time-frequency domain

[213] Given the above components, we define the continuous OTFS modulation via the following two steps

$$X[n,m] = W_{tr}[n,m]\, SDFT^{-1}(x(\tau,\nu))$$
$$s(t) = \Pi_X(g_{tr}(t)) \qquad (49)$$

[214] The first equation describes the inverse discrete time-frequency symplectic Fourier transform [c.f. Eq. (35)] and the windowing function, while the second equation describes the transmission of the modulation symbols via a Heisenberg transform [c.f. Eq. (10)].

[215] **Continuous OTFS demodulation:** Let us assume that the transmitted signal $s(t)$ undergoes channel distortion according to (7), (2) yielding $r(t)$ at the receiver. Further, let the receiver employ a receive windowing function $W_r[n,m] \in \mathbb{C}(\Lambda)$. Then, the demodulation operation consists of two steps:

[216] Evaluating the ambiguity function on Λ (Wigner transform) to obtain estimates of the time-frequency modulation symbols

$$Y[n,m] = A_{g_r,y}(\tau,\nu)\big|_{\tau=nT,\nu=m\Delta f} \qquad (50)$$

[217] Windowing and applying the symplectic Fourier transform on the modulation symbols

$$\hat{x}(\tau,\nu) = SDFT(W_r[n,m]Y[n,m]) \qquad (51)$$

[218] Notice that in (50), (51) there is no periodization of $Y[n,m]$, since the SDFT is defined on aperiodic square summable sequences. The periodization step needed in

discrete OTFS can be understood as follows. Suppose we wish to recover the transmitted information symbols by performing a continuous OTFS demodulation and then sampling on the delay-Doppler grid

$$\hat{x}(l,k) = \hat{x}(\tau,\nu)|_{\tau=\frac{m}{M\Delta f},\nu=\frac{n}{NT}}$$

[219] Since performing a continuous symplectic Fourier transform is not practical we consider whether the same result can be obtained using SFFT. The answer is that SFFT processing will produce exactly the samples we are looking for if the input sequence is first periodized (aliased) – see also (39) (40).

[220] We have now described all the steps of the OTFS modulation as depicted in Figure 3. We have also discussed how the Wigner transform at the receiver inverts the Heisenberg transform at the transmitter [c.f. Eqs. (26), (28)], and similarly for the forward and inverse symplectic Fourier transforms. The key question is what form the end-to-end signal relationship takes when a non-ideal channel is between the transmitter and receiver. The answer to this question is addressed next.

[221] **3.4 Channel Equation in the OTFS Domain**

[222] The main result in this section shows how the time varying channel in (2), (7), is transformed to a time invariant convolution channel in the delay Doppler domain.

[223] **Proposition 3:** Consider a set of NM QAM information symbols arranged in a 2D periodic sequence $x[l,k]$ with period $[M,N]$. The sequence $x[k,l]$ undergoes the following transformations:

[224] It is modulated using the discrete OTFS modulation of Eq. (44).

[225] It is distorted by the delay-Doppler channel of Eqs.(2), (7).

[226] It is demodulated by the discrete OTFS demodulation of Eqs. (45), (47).

[227] The estimated sequence $\hat{x}[l,k]$ obtained after demodulation is given by the two dimensional periodic convolution

$$\hat{x}[l,k] \simeq \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x[m,n]h_w(\frac{l-m}{M\Delta f},\frac{k-n}{NT}) \qquad (52)$$

[228] of the input QAM sequence $x[m,n]$ and a sampled version of the windowed impulse response $h_w(\cdot)$,

$$h_w\left(\frac{l-m}{M\Delta f},\frac{k-n}{NT}\right) = h_w(\tau',\nu')|_{\tau'=\frac{l-m}{M\Delta f},\nu'=\frac{k-n}{NT}} \quad (53)$$

[229] where $h_w(\tau',\nu')$ denotes the circular convolution of the channel response with a windowing function[5]

$$h_w(\tau',\nu') = \iint e^{-j2\pi\nu\tau}h(\tau,\nu)w(\tau'-\tau,\nu'-\nu)d\tau d\nu \quad (54)$$

[230] where the windowing function $w(\tau,\nu)$ is the symplectic Fourier transform of the time-frequency window $W[n,m]$

$$w(\tau,\nu) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} W[n,m]e^{-j2\pi(\nu nT-\tau m\Delta f)} \quad (55)$$

[231] and where $W[n,m]$ is the product of the transmit and receive window.

$$W[n,m] = W_{tr}[n,m]W_r[n,m] \quad (56)$$

□

[232] **Proof:** See Appendix 0.

[233] In many cases, the windows in the transmitter and receiver are matched, i.e., $W_{tr}[n,m] = W_0[n,m]$ and $W_r[n,m] = W_0^*[n,m]$, hence $W[n,m] = |W_0[n,m]|^2$.

[234] The window effect is to produce a blurred version of the original channel with a resolution that depends on the span of the frequency and time samples available as will be discussed in the next section. If we consider the rectangular (or trivial) window, i.e., $W[n,m] = 1,\ n = 0,\dots,N-1,\ m = -M/2,\dots,M/2-1$ and zero else, then its SDFT $w(\tau,\nu)$ in (55) is the two dimensional Dirichlet kernel with bandwidth inversely proportional to N and M.

[5] To be precise, in the window $w(\tau,\nu)$ is circularly convolved with a slightly modified version of the channel impulse response $e^{-j2\pi\nu\tau}h(\tau,\nu)$ (by a complex exponential) as can be seen in the equation.

[235] There are several other uses of the window function. The system can be designed with a window function aimed at randomizing the phases of the transmitted symbols, akin to how QAM symbol phases are randomized in WiFi and Multimedia-Over-Coax communication systems. This randomization may be more important for pilot symbols than data carrying symbols. For example, if neighboring cells use different window functions, the problem of pilot contamination is avoided.

[236] A different use of the window is the ability to implement random access systems over OTFS using spread spectrum/CDMA type techniques as will be discussed later.

[237] **4. Channel Time/Frequency coherence and OTFS resolution**

[238] In this section we examine certain OTFS design issues, like the choice of data frame length, bandwidth, symbol length and number of subcarriers. We study the tradeoffs among these parameters and gain more insight on the capabilities of OTFS technology.

[239] Since OTFS is based on Fourier representation theory similar spectral analysis concepts apply like frequency resolution vs Fourier transform length, sidelobes vs windowing shape etc. One difference that can be a source of confusion comes from the naming of the two Fourier transform domains in the current framework.

[240] OTFS transforms the time-frequency domain to the delay-Doppler domain creating the Fourier pairs: (i) time $\Leftrightarrow$ Doppler and (ii) frequency $\Leftrightarrow$ delay. The "spectral" resolution of interest here therefore is either on the Doppler or on the delay dimensions.

[241] These issues can be easier clarified with an example. Let us consider a time-invariant multipath channel (zero Doppler) with frequency response $H(f,0)$ for all t. In the first plot of Figure 8 we show the real part of $H(f,0)$ as well as a sampled version of it on a grid of $M = 8$ subcarriers. The second plot of Figure 8 shows the SDFT of the sampled $H(m\Delta f,0)$, i.e., $h(\tau,0)$ along the delay dimension. Notice that taking this frequency response to the "delay" domain reveals the structure of this multipath channel, that is, the existence of two reflectors with equal power in this example. Further, notice that the delay dimension of the SDFT is periodic with period $1/\Delta f$ as expected due to the nature of the discrete Fourier transform. Finally, in the third plot of Figure 8 we show the SFFT of the

frequency response, which as expected is a sampled version of the SDFT of the second plot. Notice that the SFFT has $M = 8$ points in each period $1/\Delta f$ leading to a resolution in the delay domain of $1/M\Delta f = 1/BW$.

[242] In the current example, the reflectors are separated by more than $1/M\Delta f$ and are resolvable. If they were not, then the system would experience a flat channel within the bandwidth of observation, and in the delay domain the two reflectors would have been blurred into one.

[243] Figure 9 shows similar results for a flat Doppler channel with time varying frequency response $H(0,t)$ for all f. The first plot shows the the response as a function of time, while the second plot shown the SDFT along the Doppler dimension. Finally the third plot shows the SFFT, that is the sampled version of the transform. Notice that the SDFT is periodic with period $1/T$ while the SFFT is periodic with period $1/T$ and has resolution of $1/NT$.

[244] The conclusion one can draw from Figure 9 is that as long as there is sufficient variability of the channel within the observation time NT, that is as long as reflectors have Doppler frequency difference larger than $1/NT$, the OTFS system will resolve these reflectors and will produce an equivalent channel in the delay-Doppler domain that is not fading. In other words, OTFS can take a channel that inherently has a coherence time of only T and produce an equivalent channel in the delay Doppler domain that has coherence time NT. This is an important property of OTFS as it can increase the coherence time of the channel by orders of magnitude and enable MIMO processing and beamforming under Doppler channel conditions.

[245] The two one-dimensional channel examples we have examined are special cases of the more general two-dimensional channel of Figure 10. The time-frequency response and its sampled version are shown in this figure, where the sampling period is $(T, \Delta f)$. Figure 11 shows the SDFT of this sampled response which is periodic with period $(1/T, 1/\Delta f)$, across the Doppler and delay dimensions respectively.

[246] Let us now examine the Nyquist sampling requirements for this channel response. $1/T$ is generally on the order of Δf (for an OFDM system with zero length CP it

is exactly $1/T = \Delta f$) so the period of the channel response in Figure 11 is approximately $(\Delta f, T)$, and aliasing can be avoided as long as the support of the channel response is less than $\pm\Delta f/2$ in the Doppler dimension and $\pm T/2$ in the delay dimension.

[247] Figure 12 shows the SFFT, that is, the sampled version of Figure 11. The resolution of Figure 11 is $1/NT, 1/M\Delta f$ across the Doppler and delay dimensions respectively.

[248] We summarize the sampling aspects of the OTFS modulation in Figure 13. The OTFS modulation consists of two steps shown in this figure:

[249] A Heisenberg transform translates a time-varying convolution channel in the waveform domain to an orthogonal but still time varying channel in the time frequency domain. For a total bandwidth BW and M subcarriers the frequency resolution is $\Delta f = BW/M$. For a total frame duration T_f and N symbols the time resolution is $T = T_f/N$.

[250] A SFFT transform translates the time-varying channel in the time-frequency domain to a time invariant one in the delay-Doppler domain. The Doppler resolution is $1/T_f$ and the delay resolution is $1/BW$.

[251] The choice of window can provide a tradeoff between main lobe width (resolution) and side lobe suppression, as in classical spectral analysis.

[252] 5 Channel Estimation in the OTFS Domain

[253] There is a variety of different ways a channel estimation scheme could be designed for an OTFS system, and a variety of different implementation options and details. In the section we will only present a high level summary and highlight the key concepts.

[254] A straightforward way to perform channel estimation entails transmitting a soudning OTFS frame containing a discrete delta function in the OTFS domain or equivalently a set of unmodulated carriers in the time frequency domain. From a practical standpoint, the carriers may be modulated with known, say BPSK, symbols which are removed at the receiver as is common in many OFDM systems. This approach could be

considered an extension of the channel estimation symbols used in WiFi and Multimedia-Over-Coax modems. Figure 14 shows an OTFS symbol containing such an impulse.

[255] This approach may however be wasteful as the extend of the channel response is only a fraction of the full extend of the OTFS frame $(1/T, 1/\Delta f)$. For example, in LTE systems $1/T \approx 15$ KHz while the maximum Doppler shift $f_{d,max}$ is typically one to two orders of magnitude smaller. Similarly $1/\Delta f \approx 67$ usec, while maximum delay spread τ_{max} is again one to two orders of magnitude less. We therefore can have a much smaller region of the OTFS frame devoted to channel estimation while the rest of the frame carries useful data. More specifically, for a channel with support $(\pm f_{d,max}, \pm\tau_{max})$ we need an OTFS subframe of length $(2f_{d,max}/T, 2\tau_{max}/\Delta f)$.

[256] In the case of multiuser transmission, each UE can have its own channel estimation subframe positioned in different parts of the OTFS frame. This is akin to multiplexing of multiple users when transmitting Uplink Sounding Reference Signals in LTE. The difference is that OTFS benefits from the virtuous effects of its two dimensional nature. For example, if τ_{max} is 5% of the extend of the delay dimension and $f_{d,max}$ is 5% of the Doppler dimesion, the channel estimation subframe need only be 5% x 5% = 0.25% of the OTFS frame.

[257] Notice that although the channel estimation symbols are limited to a small part of the OTFS frame, they actually sound the whole time-frequency domain via the corresponding basis functions associated with these symbols.

[258] A different approach to channel estimation is to devote pilot symbols on a subgrid in the time-frequency domain. This is akin to CRS pilots in downlink LTE subframes. The key question in this approach is the determination of the density of pilots that is sufficient for channel estimation without introducing aliasing. Assume that the pilots occupy the subgrid $(n_0T, m_0\Delta f)$ for some integers n_0, m_0. Recall that for this grid the SDFT will be periodic with period $(1/n_0T, 1/m_0\Delta f)$. Then, applying the aliasing results discussed earlier to this grid, we obtain an alias free Nyquist channel support region of $(\pm f_{d,max}, \pm\tau_{max}) = (\pm 1/2n_0T, \pm 1/2m_0\Delta f)$. The density of the pilots can then be determined from this relation given the maximum support of the channel. The pilot subgrid

should extend to the whole time-frequency frame, so that the resolution of the channel is not compromised.

[259] 6 OTFS-Access: Multiplexing More than one User

[260] There is a variety of ways to multiplex several uplink or downlink transmissions in one OTFS frame. This is a rich topic whose full treatment is outside the scope of this paper. Here we will briefly review the following multiplexing methods:

- Multiplexing in the OTFS delay-Doppler domain
- Multiplexing in the time-frequency domain
- Multiplexing in the code speading domain
- Multiplexing in the spatial domain

[261] **Multiplexing in the delay-Doppler domain:** This is the most natural multiplexing scheme for downlink transmissions. Different sets of OTFS basis functions, or sets of information symbols or resource blocks are given to different users. Given the orthogonality of the basis functions, the users can be separated at the UE receiver. The UE need only demodulate the portion of the OTFS frame that is assigned to it.

[262] This approach is similar to the allocation of PRBs to different UEs in LTE. One difference is that in OTFS, even a small subframe or resource block in the OTFS domain will be transmitted over the whole time-frequency frame via the basis functions and will experience the average channel response. Figure 15 illustrates this point by showing two different basis functions belonging to different users. Because of this, there is no compromise on channel resolution for each user, regardless of the resource block or subframe size.

[263] In the uplink direction, transmissions from different users experience different channel responses. Hence, the different subframes in the OTFS domain will experience a different convolution channel. This can potentially introduce inter-user interference at the edges where two user subframes are adjacent, and would require guard gaps to eliminate it. In order to avoid this overhead, a different multiplexing scheme can be used in the uplink as explained next.

[264] **Multiplexing in the time-frequency domain:** In this approach, resource blocks or subframes are allocated to different users in the time-freqeuncy domain. Figure 16 illustrates this for a three user case. In this figure, User 1 (blue, 1602) occupies the whole frame length but only half the available subcarriers. Users 2 and 3 (red, 1604, and black, 1606, respectively) occupy the other half subcarriers, and divide the total length of the frame between them.

[265] Notice that in this case, each user employs a slightly different version of the OTFS modulation described in Section 3. One difference is that each user i performs an SFFT on a subframe (N_i, M_i), $N_i \leq N$, $M_i \leq M$. This reduces the resolution of the channel, or in other words reduces the extent of the time-frequency plane in which each user will experience its channel variation. On the other side, this also gives the scheduler the opportunity to schedule users in parts of the time-frequency plane where their channel is best.

[266] If we wish to extract the maximum diversity of the channel and allocate users across the whole time-frequency frame, we can multiplex users via interleaving. In this case, one user occupies a subsampled grid of the time-frequency frame, while another user occupies another subsampled grid adjacent to it. Figure 17 shows the same three users as before but interleaved on the subcarrier dimension. Of course, interleaving is possible in the time dimension as well, and/or in both dimensions. The degree of interleaving, or subsampling the grip per user is only limited by the spread of the channel that we need to handle.

[267] **Multiplexing in the time-frequency spreading code domain:** Let us assume that we wish to design a random access PHY and MAC layer where users can access the network without having to undergo elaborate RACH and other synchronization procedures. There have been several discussions on the need for such a system to support Internet of Things (IoT) deployments. OTFS can support such a system by employing a spread-spectrum approach. Each user is assigned a different two-dimensional window function that is designed as a randomizer. The windows of different users are designed to be nearly orthogonal to each other and nearly orthogonal to time and frequency shifts. Each user then only transmits on one or a few basis functions and uses

the window as a means to randomize interference and provide processing gain. This can result in a much simplified system that may be attractive for low cost, short burst type of IoT applications.

[268] **Multiplexing in the spatial domain:** Finally, like other OFDM multicarrier systems, a multi-antenna OTFS system can support multiple users transmitting on the same basis functions across the whole time-frequency frame. The users are separated by appropriate transmitter and receiver beamforming operations. A detailed treatment of MIMO-OTFS architectures however is outside the scope of this paper.

[269] **7. Implementation Issues**

[270] OTFS is a novel modulation technique with numerous benefits and a strong mathematical foundation. From an implementation standpoint, its added benefit is the compatibility with OFDM and the need for only incremental change in the transmitter and receiver architecture.

[271] Recall that OTFS consists of two steps. The Heisenberg transform (which takes the time-frequency domain to the waveform domain) is already implemented in today's systems in the form of OFDM/OFDMA. In the formulation of this paper, this corresponds to a prototype filter $g(t)$ which is a square pulse. Other filtered OFDM and filter bank variations have been proposed for 5G, which can can also be accommodated in this general framework with different choices of $g(t)$.

[272] The second step of OTFS is the two dimensional Fourier transform (SFFT). This can be thought of as a pre- and post-processing step at the transmitter and receiver respectively as illustrated in Figure 18. In that sense it is similar, from an implementation standpoint, to the SC-FDMA pre-processing step.

[273] From a complexity comparison standpoint, we can calculate that for a frame of N OFDM symbols of M subcarriers, SC-FDMA adds N DFTs of M point each (assuming worse case M subcarriers given to a single user). The additional complexity of SC-FDMA is then $NMlog_2(M)$ over the baseline OFDM architecture. For OTFS, the 2D SFFT has complexity $NMlog_2(NM) = NMlog_2(M) + NMlog_2(N)$, so the term $NMlog_2(N)$ is the OTFS additional complexity compared to SC-FDMA. For an LTE subframe with $M = 1200$

subcarriers and $N = 14$ symbols, the additional complexity is 37% more compared to the additional complexity of SC-FDMA

[274] Notice also that from an architectural and implementation standpoint, OTFS augments the PHY capabilities of an existing LTE modem architecture and does not introduce co-existence and compatibility issues.

[275] **8. Example Benefits of OTFS modulation**

[276] The OTFS modulation has numerous benefits that tie into the challenges that 5G systems are trying to overcome. Arguably, the biggest benefit and the main reason to study this modulation is its ability to communicate over a channel that randomly fades within the time-frequency frame and still provide a stationary, deterministic and non-fading channel interaction between the transmitter and the receiver. In the OTFS domain all information symbols experience the same channel and same SNR.

[277] Further, OTFS best utilizes the fades and power fluctuations in the received signal to maximize capacity. To illustrate this point assume that the channel consists of two reflectors which introduce peaks and valleys in the channel response either across time or across frequency or both. An OFDM system can theoretically address this problem by allocating power resources according to the waterfilling principle. However, due to practical difficulties such approaches are not pursued in wireless OFDM systems, leading to wasteful parts of the time-frequency frame having excess received energy, followed by other parts with too low received energy. An OTFS system would resolve the two reflectors and the receiver equalizer would employ coherent combining of the energy of the two reflectors, providing a non-fading channel with the same SNR for each symbol. It therefore provides a channel interaction that is designed to maximize capacity under the transmit assumption of equal power allocation across symbols (which is common in existing wireless systems), using only standard AWGN codes.

[278] In addition, OTFS provides a domain in which the channel can be characterized in a very compact form. This has significant implications for addressing the channel estimation bottlenecks that plague current multi-antenna systems and can be a key enabling technology for addressing similar problems in future massive MIMO systems.

[279] One benefit of OTFS is its ability to easily handle extreme Doppler channels. We have verified in the field 2x2 and 4x4, two and four stream MIMO transmission respectively in 90 Km/h moving vehicle setups. This is not only useful in vehicle-to-vehicle, high speed train and other 5G applications that are Doppler intensive, but can also be an enabling technology for mm wave systems where Doppler effects will be significantly amplified.

[280] Further, OTFS provides a natural way to apply spreading codes and deliver processing gain, and spread-spectrum based CDMA random access to multicarrier systems. It eliminates the time and frequency fades common to multicarrier systems and simplifies the receiver maximal ratio combining subsystem. The processing gain can address the challenge of deep building penetration needed for IoT and PSTN replacement applications, while the CDMA multiple access scheme can address the battery life challenges and short burst efficiency needed for IOT deployments.

[281] Last but not least, the compact channel estimation process that OTFS provides can be essential to the successful deployment of advanced technologies like Cooperative Multipoint (Co-MP) and distributed interference mitigation or network MIMO.

[282] **Appendix 0**

[283] **Proof of Proposition 1**: Let

$$g_1(t) = \iint h_1(\tau,\nu)e^{j2\pi\nu(t-\tau)}g(t-\tau)d\nu d\tau \qquad (57)$$

$$g_2(t) = \iint h_2(\tau,\nu)e^{j2\pi\nu(t-\tau)}g_1(t-\tau)d\nu d\tau \qquad (58)$$

[284] Substituting (58) into (57) we obtain after some manipulation

$$g_2(t) = \iint f(\tau,\nu)e^{j2\pi\nu(t-\tau)}g(t-\tau)d\nu d\tau \qquad (59)$$

[285] with $f(\tau,\nu)$ given by (16).

□

[286] **Proof of Theorem 1:** The theorem can be proven by straightforward but tedious substitution of the left hand side of (23); by definition

$$A_{g_r,\Pi_f(g_{tr})}(\tau,\nu) = < g_r(t-\tau)e^{j2\pi\nu t}, \Pi_f(g_{tr}) >$$
$$= \int g_r^*(t-\tau)e^{-j2\pi\nu t}\Pi_f(g_{tr}(t))dt$$
$$= \int g_r^*(t \qquad (60)$$
$$-\tau)e^{-j2\pi\nu t}\iint f(\tau',\nu')e^{j2\pi\nu'(t-\tau')}g_{tr}(t-\tau')d\nu' d\tau'\, dt$$

[287] By changing the order of integration and the variable of integration $(t-\tau') \rightarrow t$ we obtain

$$A_{g_r,\Pi_f(g_{tr})}(\tau,\nu)$$
$$= \iint f(\tau',\nu')e^{j2\pi\nu'(t-\tau')}\int g_r^*(t-\tau)g_{tr}(t$$
$$-\tau')e^{-j2\pi\nu t}dt\, d\nu' d\tau' \qquad (61)$$
$$= \iint f(\tau',\nu')e^{j2\pi\nu'(\tau-\tau')}\, A_{g_r,g_{tr}}(\tau-\tau',\nu$$
$$-\nu')e^{j2\pi\nu'(\tau-\tau')}d\nu' d\tau'$$

[288] where

$$A_{g_r,g_{tr}}(\tau-\tau',\nu-\nu') = \int g_r^*(t-(\tau-\tau'))g_{tr}(t)e^{-j2\pi(\nu-\nu')t-(\tau-\tau')}dt \qquad (62)$$

[289] Notice that the right second line of (61) is exactly the right hand side of (23), which is what we wanted to prove. □

[290] **Proof of Theorem 2:** Substituting into (23) and evaluating on the lattice Λ we obtain:

$$\hat{X}[m,n] = \sum_{m'=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{n'=0}^{N-1} X[m',n'] \qquad (63)$$
$$\times \iint h(\tau-nT,\nu-m\Delta f)\, A_{g_r,g_{tr}}(nT-\tau,m\Delta f-\nu)e^{j1\pi\nu(nT-\tau)t} + V[m,n]$$

[291] Using the bi-orthogonality condition in (63) only one term survives in the right hand side and we obtain the desired result of (29).

[292] **Proof of Proposition 2:** Based on the definition of SFFT, it is not hard to verify that a delay translates into a linear phase

$$SFFT(X_2[n-k,m-l]) = SFFT(X_2[n,m])e^{-j2\pi(\frac{nk}{N}-\frac{ml}{M})} \qquad (64)$$

[293] Based on this result we can evaluate the SFFT of a circular convolution

$$SFFT\left(\sum_{k=0}^{N-1}\sum_{l=-\frac{M}{2}}^{\frac{M}{2}-1} X_1[k,l]X_2[(n-k)\ mod\ N,(-l)\ mod\ M]\right)$$
$$= \sum_{k=0}^{N-1}\sum_{l=-\frac{M}{2}}^{\frac{M}{2}-1} X_1[k,l]SFFT(X_2[n,m])\ e^{-j2\pi(\frac{nk}{N}-\frac{ml}{M})} \qquad (65)$$
$$= SFFT(X_1[n,m])SFFT(X_2[n,m])$$

[294] yielding the desired result.

[295] **Proof of Proposition 3:** We have already proven that on the time-frequency domain we have a multiplicative frequency selective channel given by (29). This result, combined with the interchange of convolution and multiplication property of the symplectic Fourier transform [c.f. Proposition 1 and Eq. (42)] leads to the desired result.

[296] In particular, if we substitute $Y(n,m)$ in the demodulation equation (48) from the time-frequency channel equation (29) and $X[n,m]$ in (29) from the modulation equation (43) we get a (complicated) end-to-end expression

$$\hat{x}[k,l] = \frac{1}{MN}\sum_{k'=0}^{N-1}\sum_{l'=0}^{M-1} x[k',l']\iint h(\tau,\nu)e^{-j2\pi\nu\tau}\times$$
$$\times\left[\sum_{m=0}^{L-1}\sum_{n=0}^{K-1} W(n,m)\ e^{-j2\pi nT(\frac{k-k'}{NT}-\nu)}e^{j2\pi m\Delta f(\frac{l-l'}{M\Delta f}-\tau)}\right]d\nu d\tau \qquad (66)$$

[297] Recognizing the factor in brackets as the discrete symplectic Fourier transform of $W(n,m)$ we have

$$\hat{x}[k,l] = \frac{1}{MN}\sum_{k'=0}^{N-1}\sum_{l'=0}^{M-1} x[k',l'] \iint h(\tau,\nu) e^{-j2\pi\nu\tau} w(\frac{l-l'}{M\Delta f} - \tau, \frac{k-k'}{NT} - \nu)\, d\nu d\tau \qquad (67)$$

[298] Further recognizing the double integral as a convolution of the channel impulse response (multiplied by an exponential) with the transformed window we obtain

$$\hat{x}[k,l] = \frac{1}{MN}\sum_{k'=0}^{N-1}\sum_{l'=0}^{M-1} x[k',l'] h_w(\frac{l-l'}{M\Delta f} - \tau, \frac{k-k'}{NT} - \nu) \qquad (68)$$

[299] which is the desired result.

[300] **9. Reference signals**

[301] Unless otherwise specifically mentioned, the terms reference signals and pilot signals are used interchangeably in the present document.

[302] **9.1 The OTFS-based reference signals**

[303] Assume the time-frequency (t-f) lattice defined by the following discrete points:

$$\Lambda_{t,f}^{D} = \mathbb{Z}dt \oplus \mathbb{Z}df = \{(Kdt, Ldf) : K, L \in \mathbb{Z}\} \qquad (69)$$

[304] Where dt (in sec) and df (in Hz) are the physical distances between the lattice points in the time and frequency dimensions respectfully, and K and L are integers[6]. We will call this lattice the data lattice as most of the points on this lattice will be occupied by data samples. The reference signals (pilots) will occupy a subset of the data lattice. When the pilot samples occupy a regular subset of the data lattice they form a regular (coarser) pilot lattice defined by:

$$\Lambda_{t,f}^{P} = \mathbb{Z}Ndt \oplus \mathbb{Z}Mdf \quad N, M \in \mathbb{Z} \geq 1 \qquad (70)$$

[6] In OFDM terminology, df may be the subcarrier spacing and dt may be the OFDM symbol time.

[305] As an example, for $N = 14,\ M = 2,$ the t-f plane will look as shown in Figure 19.

[306] As shown in this document, the data lattice (69) is associated, through the symplectic Fourier transform, with a Delay-Doppler (τ, ν) torus which has the following delay and Doppler circumferences respectfully:

$$C_\tau^D = 1/df\ ,\ C_\nu^D = 1/dt$$

[307] The Delay-Doppler torus associated with the (coarser) pilot lattice (70) is a torus with the following smaller circumferences

$$C_\tau^P = 1/(Mdf)\ ,\ C_\nu^P = 1/(Ndt)$$

[308] It can be shown that a 2-D function $x(\tau, \nu)$ on the continuous Delay-Doppler torus associated with the t-f lattice defined in (69) can be transformed to a 2-D discrete function $X[i, j]$ on the t-f lattice using an inverse symplectic discrete Fourier transform defined as

$$X[i,j] = SDFT^{-1}\big(x(\tau,\nu)\big) \triangleq \frac{1}{dtdf}\int_0^{\frac{1}{\mathrm{df}}}\int_0^{\frac{1}{dt}} x(\tau,\nu)e^{j2\pi(\nu idt - \tau jdf)}d\nu\, d\tau \tag{71}$$

[309] If the p^{th} pilot $x_p(\tau, \nu)$ is chosen as the delta function $\delta(\tau_p, \nu_p)$ on the Delay-Doppler torus, then the representation of this pilot on the time-frequency lattice defined in (69) will be

$$X_p[i,j] = \frac{1}{dtdf}e^{j2\pi(\nu_p idt - \tau_p jdf)}\ \ i, j \in \mathbb{Z} \tag{72}$$

[310] Figure 20 shows an example of positioning 10 OTFS-based pilots on the Delay-Doppler plane and how one of the pilots look after it goes through the inverse symplectic discrete Fourier transform (71).

[311] Each OTFS-based reference signal is a delta function placed on the Delay-Doppler torus at a different point (τ_p, ν_p). The sum of these delta functions is then transformed to the t-f plane using the inverse symplectic discrete Fourier transform, and a

subset of the samples in the t-f plane are selected to be transmitted. Stated differently, an OTFS-based reference signal is a symplectic exponential which is restricted to a subset of points on the data lattice.

[312] If P pilots are sent using a subset of the data lattice which forms a regular lattice as represented in (70), the t-f samples of the n pilots will be:

$$X[i,j] = \frac{1}{dtdf}\sum_{p=1}^{P} e^{j2\pi(\nu_p idt - \tau_p jdf)} \;\; i = kN, j = lM; \;\; k,l \in \mathbb{Z} \tag{73}$$

[313] Where N and M are fixed positive integers representing the size of the coarser pilot lattice.

[314] When the pilot samples in the t-f plane form a regular lattice, the number of pilots that can be packed in that lattice can be calculated from the circumferences of the torus associated with the pilot lattice and the maximum delay and Doppler spreads of the channels that each of the pilots is expected to experience. To avoid leakage between the pilots, the maximum number of pilots that can be packed in each dimension is the circumference of the torus divided by the maximum spread. Noting the delay and Doppler spreads as Δ_τ and Δ_ν respectfully, the maximum number of pilots that can be packed in each of the dimensions is:

$$\begin{aligned} N^P_{\max_\tau} &= \lfloor C^P_\tau / \Delta_\tau \rfloor \\ N^P_{\max_\nu} &= \lfloor C^P_\nu / \Delta_\nu \rfloor \end{aligned} \tag{74}$$

[315] As an example, for channels with average maximum delay spread of $5\,\mu s$ and a maximum Doppler frequency of 50Hz (maximum Doppler spread of 100Hz), a torus with a delay circumference of $C^P_\tau = 67\;\mu s$ and Doppler circumference of $C^P_\nu = 200\,Hz$ can support up to 13 pilots in the delay dimension and 2 pilots in the Doppler dimension, for a total of 26 pilots. How close a system can get to the maximum achievable pilot packing will depend on the pilot observation window. A finite observation window of the pilots translates to convolving the pilots in the Delay-Doppler plane with the symplectic Fourier transform of the window (which, in the case of a rectangular window, is a two-dimensional

sinc function). Hence, a larger observation window will result in lower leakage between the received pilots, which will enable:

- improved accuracy of the channel estimation and
- Tighter packing of pilots, up to the maximum number stated in (74) for infinitely large window.

[316] Note that staggering the pilots (e.g., not placing them on a rectangular grid in the Delay-Doppler plane) may improve the separation between the pilots and hence could provide better, or denser, packing. Figure 21 shows an example of staggered pilots. As can be seen from the figure, all the even pilots are staggered. An example of how staggering the LTE UL DM reference signals can improve the channel estimation is shown in the present document.

[317] **9.2 Reference Signal Packing**

[318] Packing orthogonal OTFS-based pilots can be done using one of the following schemes:

[319] Delay-Doppler Pilot Packing (DDPP): Arranging the pilots in the Delay-Doppler plane keeping the distances large enough to minimize the leakage of the received pilots onto each other after going through the worst case delay and Doppler shifts of the channels each of the pilots may experience, and taking into consideration the impact of the (sub) sampling of the pilots in the t-f domain, and the pilot observation window.

- Time-Frequency Pilot Packing (TFPP): Assigning for each pilot (with no overlap) a coarse enough lattice that can support the relevant channel (with a regular pilot lattice, the circumferences of the associated torus have to be larger than the largest expected delay and Doppler shifts through the channel).
- Mixed Pilot Packing (MPP): A combination of DDPP and TFPP.

[320] OTFS-based pilots can be packed very efficiently, and hence can support the simultaneous transmission of a very large number of orthogonal pilots without using a significant percentage of the channel capacity.

[321] Here are a few pilot packing examples:

[322] Example 1 (DDPP):

[323] Assume the following:

- Data lattice parameters (LTE numerology):
  - $dt = 1/14\ ms$
  - $df = 15\ kHz$
- Channel parameters:
  - Delay spread: $5\ \mu s$ (ETU)
  - Max Doppler frequency: $50\ Hz$ ($100\ Hz$ spread)

[324] If we chose a pilot lattice with $N = 28$ ($Ndt = 2ms$) and $M = 1$, as shown in Figure 22, the circumferences of the pilot torus will be

$$C_\tau^P = 66.67us\ ,\ C_\nu^P = 500\ Hz$$

[325] The maximum number of pilots that can be supported in this configuration is $13x5 = 65$. This is by placing 13 pilots in the delay dimension of the torus (spaced $5.13\ \mu s$ apart) and 5 replicas of these pilots in the Doppler dimension of the torus (spaced $100\ Hz$ apart). In practice with $10\ MHz$ channel bandwidth and allowing for a reasonable size window in the time dimension this configuration can support at least 40 pilots ($10x4$).

[326] Example 2 (TFPP):

[327] For the same data lattice as in Example 1 assume the following channel parameters:

- Delay spread: $5\ \mu s$ (ETU)
- Max Doppler frequency: $200\ Hz$ ($400\ Hz$ spread)

[328] If we split the pilot lattice of Example 1 into 12 different pilot lattices represented by the different color diamonds in Figure 23, then each of these pilot lattices will have $N = 28$ ($Ndt = 2ms$) and $M = 12$ ($Mdf = 180KHz$). The circumferences of all the tori associated with these lattices will be:

$$C_\tau^P = 5.56\ us\ ,\ C_\nu^P = 500\ Hz$$

[329] As can be seen from the circumferences, each pilot lattice can support only a single pilot, for a total of 12 pilots that can be supported by the 12 pilot lattices of Figure 23.

[330] Example 3 (MPP):

[331] If we split the lattice of Example 1 into two lattices as shown in Figure 24, each of the lattices will have $Ndt = 2\,ms$ and $Mdf = 30\,kHz$. The tori associated with the two lattices will both have the following circumferences:

$$C_\tau^P = 33.33\,\mu s\,,\, C_\nu^P = 500\,Hz$$

[332] Assuming the same channel parameters as in Example 1, the maximum number of pilots that can be supported by each of these tori is $6x5 = 30$ for a total of 60 pilots on the two lattices. A practical number with $10\,MHz$ channel bandwidth and a reasonable size window in the time dimension is expected to be 40 $(4x5x2)$ or 50 $(5x5x2)$ pilots.

[333] Another example of MPP is shown later in the document.

[334] The advantage of using DDPP is that it provides:

- More flexibility in supporting different channel delay and Doppler spreads. With DDPP the pilots can be placed anywhere on the continuous torus whereas when multiplexing the pilots in the time-frequency plane the options are limited to using discrete lattices.
- Lower latency than TFPP when the pilots are used for demodulating data, since in TFPP the lattice used for each pilot is coarser, and hence the average time between the data and the last pilot used for interpolation (the pilot following the data) is larger.

[335] TFPP has an advantage when trying to use a short pilot observation window as an equivalent quality of the channel estimation as is achieved with DDPP can be achieved with TFPP using a shorter observation window.

[336] 9.3 Potential Use of OTFS Based Reference Signals in LTE

[337] To support massive MIMO using channel reciprocity, all active UEs need to send pilots on the UL, so that the eNodeB can predict the channel for pre-coding its DL transmissions to these UEs. This requires supporting a large number of pilots.

[338] One way to support a large number of pilots is to send OTFS-based reference signals using the resources allocated to the Sounding Reference Signals (SRSs). The SRSs in the LTE system are transmitted on the last symbol of the UL sub-frame. In TDD mode the SRSs can be scheduled with the shortest configuration period being 5 sub-frames ($5\ ms$). With this configuration, the SRSs use a pilot lattice with $N = 70, M = 1$ (see section 9.1). The torus associated with this lattice has the following delay and Doppler circumferences respectfully:

$$C_\tau^P = 66.67\ \mu s, C_\nu^P = 200\ Hz$$

[339] Assuming, as an example, an ETU channel with maximum Doppler frequency of up to $10\ Hz$, the maximum number of pilots that can be supported on this lattice is $13x10 = 130$. With a practical finite observation window, the number of pilots that can be supported with good enough channel prediction of $5\ ms$ (the distance between the pilots) into the future will be smaller, but is still expected to be very large.

[340] 9.4 Examples of Reference Signals for 5G Communications

[341] The proposed structure of the reference signals supports pre-coding of the downlink (DL) transmissions using channel reciprocity in the presence of time varying channels. We refer to pre-coding as a generalized beamforming scheme for supporting multi-layer transmission in a MIMO system.

[342] For the 5G reference signals it is proposed:

- To dedicate a subset of the time-frequency data lattice to reference signals
- To use the OTFS-based reference signals described in section 9.1.
- To scramble the reference signals by multiplying their time-frequency samples by 2-D chirp sequences (e.g. 2-D Zadoff-Chu), for the purpose of limiting the inter-cell interference between the reference signals. The 2-D sequences will have a much richer selection of sequences with good cross correlation characteristics than single dimension sequences.

- To pack all the reference signals (pilots) required for the operation of the system on the lattice dedicated to the reference signals (except maybe for demodulation reference signals, when needed, that could be sent with the data).
- To have the eNodeB (base station) transmit the DL reference signals continuously on the dedicated time-frequency DL pilot lattice.
- To have each UE (subscriber device) transmit its uplink (UL) reference signals on the dedicated UL pilot lattice before the eNodeB starts to pre-code its transmissions to the UE.

[343] Separating the pilots from the data enables starting the transmission of the pilots before the data transmission starts, which enables the receiver to use a large pilot observation window resulting in higher channel observation resolution. The higher channel observation resolution enables:

- A better channel estimation
- A better pilot packing (due to reduced leakage between the received pilots)
- Improved predictability of the channel, which will improve the precoding in the presence of Doppler spreads,

[344] all without impacting data transmission latency.

[345] To enable using the channel reciprocity for pre-coding, the channel response information has to be current during the DL transmission time. To achieve that, the eNodeB has to receive pilots from all active UEs on a regular basis so that the eNodeB has an up to date channel information whenever it needs to transmit to a UE. It can be shown that using the proposed pilot lattice that meets the conditions in (74) with a long enough observation window, provides good channels prediction that can be used for pre-coding.

[346] The number of pilots that need to be supported in the UL is equal to the number of active UEs in a cell and at the edges of the neighboring cells (to minimize interference between pilots and to allow support of interference cancelation in the DL) times the number of spatially multiplexed layers per UE. For the purpose of pilot

transmissions, an active UE will be a UE that started sending or receiving data (alternatively it can be the time it wakes up to start sending or receiving data). At that point the UE will start sending the pilots. Until the eNodeB collects enough pilots from the UE to support pre-coding the eNodeB will send data to that UE without pre-coding. Also, to limit the number of UEs that need to send pilots continuously, the support for pre-coding per UE can be configurable. In that case the UE could send demodulation pilots only with the data.

[347] To support pre-coding on the DL, the proposal is to have the active UEs transmit their pilots on a regular basis. This will allow the eNodeB to collect a history of pilot information from all the active UEs. When a packet needs to be transmitted to a specific UE, the eNodeB can use the pilot history of the specific UE to calculate the pre-coder, and apply it to the transmitted packet. In addition to using the pilots for pre-coding DL transmissions, the eNodeB can use the regularly transmitted UL pilots to estimate the channel for demodulating packets transmitted on the UL. Using pilot history will also help improve the separation of the desired pilot from the other pilots and the quality of the channel estimation for demodulating the received signal. It is assumed that the transmissions on the UL are either not pre-coded or that the eNodeB has knowledge of the pre-coders used.

[348] In the DL, assuming the pre-coding is good enough, the UEs will not need demodulation reference signals (DM RSs) on the pre-coded layers. With that assumption, the eNodeB will only need to send reference signals on the spatial layers that do not use pre-coding. Hence the number of pilots on the DL will be much smaller than on the UL. The proposal is to send all the DL pilots on a regular basis. These pilots will be used by the UE both as DM RSs (for the non pre-coded transmissions) and for measuring the Observed Time Difference Of Arrival (OTDOA). If it is perceived that DM reference signals are still needed after the pre-coding, then the DM reference signals can be sent with the data.

[349] The number of pilots that need to be supported in the DL is equal to the number of non pre-coded layers per cell times the number of neighboring cells. This is to prevent the pilots from interfering with the pilots of the neighboring cells and to support measuring the OTDOA.

[350] 9.4.1 Downlink Reference Signals

[351] For the LTE numerology it is proposed to use a pilot lattice with $N = 28$ and $M = 1$. This will support up to 40 pilots for ETU channels with average maximum Doppler frequency of $50\ Hz$ (ETU-50), as shown in Example 1 in section 9.2. If a smaller number of pilots are needed this lattice will support higher Doppler spreads (e.g. with 20 pilots it can support a maximum Doppler frequency of $100\ Hz$) and vice versa.

[352] If the subcarrier spacing changes to $150\ KHz$, the data lattice parameters will be:

- $dt = 1/140\ ms$
- $df = 150\ KHz$

[353] This numerology also supports 40 pilots for ETU-50 channels using the same pilot lattice $(N = 28, M = 1)$. In this case all the pilots will be packed in the Doppler dimension.

[354] The DL pilots will be transmitted continuously on the pilot lattice. The UEs should collect a long enough history of the pilots to support good enough channel estimation for the purpose of receiving non pre-coded (data or control) transmissions from the eNodeB, and for improving the measured TOA.

[355] 9.4.2 Uplink Reference Signals

[356] For the LTE numerology it is proposed to use one or more adjacent pilot lattices with $N = 28$ and $M = 1$, and/or one or more adjacent pilot lattices with $N = 14$ and $M = 1$. Each of the first lattices will support up to 40 pilots for ETU-50 (as shown in Example 1 in section 9.2), and each of the second lattices will support 80 pilots. A good example is using one lattice with $N = 28$, $M = 1$ and one lattice with $N = 14, M = 1$, in combination with a DL pilot lattice of $N = 28, M = 1$. This example, demonstrated in Figure 25, supports 120 UL pilots and 40 DL pilots for ETU channels with an average Doppler frequency of $50\ Hz$. Figure 26 shows the representation of 40 equally spaced pilots on the Delay-Doppler plane that is associated with the pilot lattice of $N = 28, M = 1$.

[357] The pilot structure of Figure 25 supports both symmetric and asymmetric DL/UL transmissions. Note that with this pilot structure a switching guard period (GP) is required after every DL sub-frame. Hence, the more asymmetric the transmissions are the more switching guard periods (GPs) will be required. If the downlink-to-uplink switch-point periodicity is N sub-frames, then $(N-1)$ GPs will be required per N sub-frames.

[358] The pilot structure of Figure 25 adds overhead of 14.3% $(2/14)$ for supporting 160 pilots. This is an overhead of 0.09% per pilot. Note that this overhead per pilot depends on the delay and Doppler spreads of the channels. The number of supported pilots in this pilot configuration will be doubled and the overhead per pilot will be cut by half for a maximum Doppler frequency of $25\ Hz$ (instead of $50\ Hz$).

[359] **9.5 Comparison with LTE Pilot Packing**

[360] As shown in section 9.4.2, the reference signal structure proposed in section 9.3 can accommodate 40 pilots on the DL and 120 pilots on the UL, all supporting an ETU-50 channel. The DL pilots use 3.6% of the total PHY resources (data lattice), and the UL pilots occupy 10.7% of the total PHY resources.

[361] In LTE, the cell specific reference signals occupy 14.3% of the DL PHY resources. With these reference signals LTE supports up to 4 non pre-coded DL spatial layers. On the UL, to support 8 spatial layers, the UEs can be configured (in TDD mode) to send SRSs with a configuration period of 5ms. In this mode, to support ETU channels, a total of 8 SRSs can be supported. These SRSs occupy 1.43% of the UL PHY resources. These reference signals can't support any significant Doppler spread in ETU channels.

[362] The following Table 1 shows a summary of the comparison between the proposed reference signals and the LTE reference signals. Note that for supporting lower Doppler channels than shown in the table for the OTFS RSS, the number of OTFS RSs could either be increased proportionally to the decrease in the Doppler spread or the overhead could decrease proportionally. As an example, for ETU-5 channels the overhead of the OTFS RSs could decrease 10 times (to around 0.02% per RS) while still supporting 140 pilots on the DL and 20 pilots on the UL.

Table 1

Parameter	LTE RSs	OTFS RSs	Comments
# of DL non pre-coded RSs	4	40	
# of UL non pre-coded SRSs	8	120	Assuming SRSs configuration period is 5ms
Overhead DL RSs (from DL portion)	14.3% (3.6% per RS)	7.1% (0.18% per RS)	Assuming symmetric UL/DL
Overhead UL SRSs (from UL portion)	1.43% (0.18% per RS)	21.4% (0.18% per RS)	Assuming symmetric UL/DL
Supported channels	ETU, no Doppler	ETU-50	

[363] **Appendix A - Mathematical background**

[364] A function g of a discrete variable ndt where $n \in \mathbb{Z}$ (the set of integer numbers) and $dt \in \mathbb{R}$ (the set of real numbers) is a function on the one dimensional lattice $\Lambda_t = \mathbb{Z}dt = \{ndt: n \in \mathbb{Z}, dt \in \mathbb{R}\}$. It is well known that the Fourier transform of the discrete function $g(ndt)$ is a continuous periodic function with period $1/dt$. The discrete Fourier transform transforms the function $g(ndt)$ to a continuous function $G(f)$ that resides on $f = [0,1/dt)$. Since the discrete Fourier transform translates a multiplication of two functions on the lattice Λ_t to a circular convolution, it is convenient to refer to $G(f)$ as residing on a circle with a circumference of $1/dt$.

[365] Similar to the one dimensional case, it can be shown that the discrete symplectic Fourier transform (a twisted version of the two dimensional discrete Fourier transform) transforms a function g of two discrete variables to a function of two continuous periodic variables. Assume that the function g resides on the following lattice:

$$\Lambda_{t,f} = \mathbb{Z}dt \oplus \mathbb{Z}df = \{(ndt, mdf): n, m \in \mathbb{Z}, dt, df \in \mathbb{R}\} \quad (75)$$

[366] The discrete symplectic Fourier transform of $g(ndt, mdf)$ is given by:

$$SF(g)(\tau, \nu) = G(\tau, \nu) = dtdf \cdot \sum_{n,m} e^{-2\pi j(\nu ndt - \tau mdf)} g(ndt, ndf) \quad (76)$$

[367] The function $G(\tau, \nu)$ resides on a two-dimensional plane $\tau \times \nu = [0,1/df)x[0,1/dt)$ or equivalently on a torus with circumferences $1/df$ in the τ dimension and $1/dt$ in the ν dimension. This torus is referred to as the torus associated with the lattice $\Lambda_{t,f}$. An example is depicted in Figure 27.

[368] **Appendix B - Effect of Staggering the LTE UL DM Reference Signals**

[369] The uplink demodulation reference signals in LTE are defined by the time-domain cyclic shift $\tau(\lambda)$ of the base sequence $r_{uv}(k)$ according to

$$r_{uv}(mL_{RS} + k, \tau) = w_m(\lambda)e^{j\tau(\lambda)k}r_{uv}(k) \mid 0 \leq k \leq L_{RS} - 1, m = 0,1 \quad (77)$$

[370] Where L_{RS} is the length of the reference signal sequence (in number of subcarriers), λ is the spatial layer index, $[w_0(\lambda)\ w_1(\lambda)] = [\pm 1\ \pm 1]$ is the Orthogonal Cover Code (OCC), and $r_{uv}(k)$ is the base (Zadoff-Chu) sequence. The term $e^{j\tau(\lambda)k}$ represents the layer dependent cyclic shift which separates the pilots of the different layers. For rank 4 UL transmission the 4 UL reference signals can be represented in the Delay-Doppler plane as shown in Figure 28. As can be seen from Figure 28, all 4 reference signals have the same Doppler shift. Staggering the reference signals as shown in Figure 29 enables better estimation of the channel when using a small number of PRBs (small observation window) as shown in Figure 30.

[371] **Appendix C MPP Examples**

[372] Using the same numerology and channel parameters of example 1 in section 9.2, chose the lattice points as shown in Figure 31. The red columns form the same torus as in example 1 and can support a maximum of 65 pilots. The remaining red points can be viewed as 13 coarse lattices, each with N=28 and M=12. The circumferences of torus associated with this lattice are

$$C_\tau^P = 5.56us\ ,\ C_\nu^P = 500\ Hz$$

[373] This torus supports 1x5 pilots, so the total number of pilots that can be supported by these 13 coarse lattices is 5x13=65. Hence the total maximum number of pilots that can be supported by the pilots' sample points in Figure 31 is 130. A more

practical number is 40 on the first lattice (same as in Example 1) and $4x13 = 52$ on the coarser lattices, for a total of 92 pilots.

[374] For the same example we can partition the pilots' sample points differently, into 25 lattices with $N = 28$ and $M = 12$. Each such lattice supports a maximum of $1x5$ pilots for a total of $5x25 = 125$ pilots and a more practical number of $4x25 = 100$ pilots.

[375] Figure 32 shows an example communication network in which the disclosed technology can be embodied. The network 3200 may include a base station transmitter that transmits wireless signals s(t) (downlink signals) to one or more receivers r(t) which may be located in a variety of locations, including inside or outside a building and in a moving vehicle. The receivers may transmit uplink transmissions to the base station, typically located near the wireless transmitter.

[376] Figure 33 shows a flowchart of an example method 3300 of wireless communication. The method 3300 includes the following operations: (3302) determining a maximum delay spread for a transmission channel, (3304) determining a maximum Doppler frequency spread for the transmission channel, (3306) allocating a set of transmission resources in a time-frequency domain to a number of pilot signals based on the maximum delay spread and the maximum Doppler frequency spread, and (3308) transmitting the pilot signals over a wireless communication channel using transmission resources. Various examples and options are disclosed in the present document, in particular, in Section 9.

[377] In some embodiments, each pilot signal may correspond to a delta function in the delay-Doppler domain.

[378] In some embodiments, the allocating operation 3306 may include staggering transmission resources for the number of pilots with respect to each other such that at least some pilots occupy transmission resources that do not occur on a rectangular grid in the delay-Doppler domain. In some embodiments, every other pilot signal position may be staggered from n original position on the rectangular grid. For example, in various embodiments, all even-numbered (or odd-numbered) pilot signals may be staggered.

[379] In some embodiments, the set of transmission resources in the time-frequency domain occupied by any given pilot signal corresponds to a lattice comprising time instances uniformly distributed along a time axis and having a first step size and frequencies uniformly distributed along a frequency axis and having a second step size. It will be understood that the step sizes in the time-frequency domain are different from frequency domain spacing of pilot signals.

[380] In some embodiments, the set of transmission resources in the time-frequency domain occupied by the pilot signal correspond to a lattice comprising time instances non-uniformly distributed along a time axis.

[381] In some embodiments, the set of transmission resources in the time-frequency domain occupied by at least one pilot signal correspond to a lattice comprising frequencies that are non-uniformly distributed along a frequency axis.

[382] In some embodiments, the set of transmission resources in the time-frequency domain occupied by at least one pilot signal are non-overlapping with another set of resources in the time-frequency domain over which user data is transmitted by the wireless communication device.

[383] In some embodiments, the operation 3308 of transmitting includes transmitting the pilot signal to a given user equipment prior to transmitting data to the user equipment.

[384] In some embodiments, the pilot signals may be generated by scrambling a basis signal using a two-dimensional (2-D) chirp sequence. In some embodiments, the pilot signals may be generated by cyclically shifting by a different amount a root 2-D Zadoff-Chu sequence. The shift may be performed in the time domain and/or in the frequency domain.

[385] In some embodiments, the transmission 3308 may be performed on a continuous basis from the transmitter to a user equipment, regardless of there is data transmission going on from the transmitter to the UE. In some embodiments, data may be pre-coded prior to the transmission.

[386] In some embodiments, each pilot signal generated by the wireless device using the method 3300 may occupy non-overlapping and distinct transmission resources.

[387] In some embodiments, the wireless communication device includes a base station, the method 3300 further including generating at least two pilot signals occupying two sets of transmission resources are non-overlapping in the delay-Doppler domain. In some embodiments, the at least two pilot signals use non-overlapping delay domain resources. In some embodiments, the at least two pilot signals use non-overlapping Doppler-domain resources.

[388] In some embodiments, the wireless communication device includes a user equipment, and wherein the set of transmission resources are specified to the wireless communication device in a upper layer message.

[389] Figure 34 is a block diagram of an example of a wireless communication apparatus 3400 that includes a memory 3402 for storing instructions, a processor 3404 and a transmitter 3406. The transmitter 3406 is communicatively coupled with the processor 3404 and the memory 3402. The memory 3402 stores instructions for the processor 3404 to generate a pilot signal according to the methods described here (e.g., method 3300 and method 3500). The transmitter 3404 transmits the pilot signal over a wireless communication channel using transmission resources that are designated for pilot signal transmission.

[390] Figure 35 shows a flowchart of an example method 3500 of wireless communication. The method 3500 includes the following operations: determining (3502) a maximum delay spread for a transmission channel, determining (3504) a maximum Doppler frequency spread for the transmission channel, determining (3506) a number of pilot signals that can be transmitted using a set of two-dimensional transmission resources at least based on the maximum delay spread and the maximum Doppler frequency spread, allocating (3508) the set of transmission resources from a two-dimensional set of resources to the number of pilot, and transmitting (3510) the pilot signals over a wireless communication channel using transmission resources. Various examples and options are disclosed in the present document, in particular, in Section 9.

[391] In some embodiments, the operation 3506 may include determining the number of pilot signals based on one or more of a number of receivers to send the pilot signals to, a number of transmission layers used for transmissions to the receivers, a

number of receivers that are also transmitting pilot signals, and possible interference from another cell's pilot signals. As previously described, a target observation window may be determined in the time-frequency domain based on the desired resolution and observation time.

[392] In some embodiments, pilot signals may be staggered. Some examples are shown and described with respect to Figure 29 and Figure 30. The staggering may be achieve by re-locating pilots from a position on a grid to a position along one of the dimensions (time or frequency) to maximize the separation from the non-staggered pilots.

[393] Figure 36 is a block diagram of an example of a wireless communication apparatus that can be used for embodying some techniques disclosed in this patent document. The apparatus 3600 may be used to implement method 3300 or 3500. The apparatus 3600 includes a processor 3602, a memory 3604 that stores processor-executable instructions and data during computations performed by the processor. The apparatus 3600 includes reception and/or transmission circuitry 3606, e.g., including radio frequency operations for receiving or transmitting signals.

[394] It will be appreciated that various techniques are disclosed for pilot packing in an OTFS-based communication network.

[395] The disclosed and other embodiments and the functional operations described in this document can be implemented in digital electronic circuitry, or in computer software, firmware, or hardware, including the structures disclosed in this document and their structural equivalents, or in combinations of one or more of them. The disclosed and other embodiments can be implemented as one or more computer program products, i.e., one or more modules of computer program instructions encoded on a computer readable medium for execution by, or to control the operation of, data processing apparatus. The computer readable medium can be a machine-readable storage device, a machine-readable storage substrate, a memory device, a composition of matter effecting a machine-readable propagated signal, or a combination of one or more them. The term "data processing apparatus" encompasses all apparatus, devices, and machines for processing data, including by way of example a programmable processor, a computer, or multiple processors or computers. The apparatus can include, in addition to hardware, code that

creates an execution environment for the computer program in question, e.g., code that constitutes processor firmware, a protocol stack, a database management system, an operating system, or a combination of one or more of them. A propagated signal is an artificially generated signal, e.g., a machine-generated electrical, optical, or electromagnetic signal, that is generated to encode information for transmission to suitable receiver apparatus.

[396] A computer program (also known as a program, software, software application, script, or code) can be written in any form of programming language, including compiled or interpreted languages, and it can be deployed in any form, including as a stand-alone program or as a module, component, subroutine, or other unit suitable for use in a computing environment. A computer program does not necessarily correspond to a file in a file system. A program can be stored in a portion of a file that holds other programs or data (e.g., one or more scripts stored in a markup language document), in a single file dedicated to the program in question, or in multiple coordinated files (e.g., files that store one or more modules, sub programs, or portions of code). A computer program can be deployed to be executed on one computer or on multiple computers that are located at one site or distributed across multiple sites and interconnected by a communication network.

[397] The processes and logic flows described in this document can be performed by one or more programmable processors executing one or more computer programs to perform functions by operating on input data and generating output. The processes and logic flows can also be performed by, and apparatus can also be implemented as, special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application specific integrated circuit).

[398] Processors suitable for the execution of a computer program include, by way of example, both general and special purpose microprocessors, and any one or more processors of any kind of digital computer. Generally, a processor will receive instructions and data from a read only memory or a random access memory or both. The essential elements of a computer are a processor for performing instructions and one or more memory devices for storing instructions and data. Generally, a computer will also include,

or be operatively coupled to receive data from or transfer data to, or both, one or more mass storage devices for storing data, e.g., magnetic, magneto optical disks, or optical disks. However, a computer need not have such devices. Computer readable media suitable for storing computer program instructions and data include all forms of non volatile memory, media and memory devices, including by way of example semiconductor memory devices, e.g., EPROM, EEPROM, and flash memory devices; magnetic disks, e.g., internal hard disks or removable disks; magneto optical disks; and CD ROM and DVD-ROM disks. The processor and the memory can be supplemented by, or incorporated in, special purpose logic circuitry.

[399] While this document contains many specifics, these should not be construed as limitations on the scope of an invention that is claimed or of what may be claimed, but rather as descriptions of features specific to particular embodiments. Certain features that are described in this document in the context of separate embodiments can also be implemented in combination in a single embodiment. Conversely, various features that are described in the context of a single embodiment can also be implemented in multiple embodiments separately or in any suitable sub-combination. Moreover, although features may be described above as acting in certain combinations and even initially claimed as such, one or more features from a claimed combination can in some cases be excised from the combination, and the claimed combination may be directed to a sub-combination or a variation of a sub-combination. Similarly, while operations are depicted in the drawings in a particular order, this should not be understood as requiring that such operations be performed in the particular order shown or in sequential order, or that all illustrated operations be performed, to achieve desirable results.

[400] Only a few examples and implementations are disclosed. Variations, modifications, and enhancements to the described examples and implementations and other implementations can be made based on what is disclosed.

Claims

1. A wireless communication method, implemented by a wireless communication device, comprising:

determining a maximum delay spread for a transmission channel;

determining a maximum Doppler frequency spread for the transmission channel;

allocating a set of transmission resources in a time-frequency domain to a number of pilot signals based on the maximum delay spread and the maximum Doppler frequency spread; and

transmitting the pilot signals over a wireless communication channel using transmission resources.

2. The method of claim 1, wherein the allocating the set of transmission resources includes:

staggering transmission resources for the number of pilots with respect to each other such that at least some pilots occupy transmission resources that do not occur on a rectangular grid in the delay-Doppler domain.

3. The method of claim 2, wherein the staggering includes staggering every other pilot.

4. The method of claim 1, wherein the set of transmission resources in the time-frequency domain occupied by any given pilot signal corresponds to a lattice comprising time instances uniformly distributed along a time axis and having a first step size and frequencies uniformly distributed along a frequency axis and having a second step size.

5. The method of claim 1, wherein the set of transmission resources in the time-frequency domain occupied by the pilot signal correspond to a lattice comprising time instances non-uniformly distributed along a time axis.

6. The method of claim 1, wherein the set of transmission resources in the time-frequency domain occupied by at least one pilot signal correspond to a lattice comprising frequencies that are non-uniformly distributed along a frequency axis.

7. The method of claim 1, wherein the set of transmission resources in the time-frequency domain occupied by at least one pilot signal are non-overlapping with another set of resources in the time-frequency domain over which user data is transmitted by the wireless communication device.

8. The method of claim 1, wherein the transmitting the pilot signal includes transmitting the pilot signal to a given user equipment prior to transmitting data to the user equipment.

9. The method of claim 1, wherein the generating the pilot signal includes:
scrambling a basis signal using a two-dimensional (2-D) chirp sequence.

10. The method of claim 1, wherein each pilot signal corresponds to a delta function in a delay-Doppler domain.

11. The method of claim 1, wherein each pilot signal corresponds to a different cyclic shift in a time domain and./or frequency domain of a root 2-D Zadoff-Chu sequence.

12. The method of claim 1, wherein the transmitting the pilot signal is performed continuously, regardless of data transmissions.

13. The method of claim 1, wherein the wireless communication device includes a base station, the method further including pre-coding data prior to data transmissions.

14. The method of claim 1, wherein the wireless communication device includes a base station, the method further including generating at least two pilot signals occupying two sets of transmission resources non-overlapping in the time-frequency domain.

15. The method of claim 14, further including:

individually transmitting the at least two pilot signals to two different user equipment at time instances that are non-overlapping with each other.

16. The method of claim 14, further including:

individually transmitting the at least two pilot signals from two different user equipment at time instances that are non-overlapping with each other.

17. The method of claim 1, wherein the wireless communication device includes a base station, the method further including generating at least two pilot signals occupying two sets of transmission resources are non-overlapping in the delay-Doppler domain.

18. The method of claim 14, wherein the at least two pilot signals use non-overlapping delay domain resources.

19. The method of claim 14, wherein the at least two pilot signals use non-overlapping Doppler-domain resources.

20. The method of claim 1, wherein the wireless communication device includes a user equipment, and wherein the set of transmission resources are specified to the wireless communication device in a upper layer message.

21. A wireless communication method, implemented by a wireless communication device, comprising:

determining a maximum delay spread for a transmission channel;

determining a maximum Doppler frequency spread for the transmission channel;

determining a number of pilot signals that can be transmitted using a set of two-dimensional transmission resources at least based on the maximum delay spread and the maximum Doppler frequency spread;

allocating the set of transmission resources from a two-dimensional set of resources to the number of pilot; and

transmitting the pilot signals over a wireless communication channel using transmission resources.

22. The method of claim 21, wherein the determining the number of pilot signals further includes determining the number of pilot signals based on one or more of a number of receivers to send the pilot signals to, a number of transmission layers used for transmissions to the receivers, a number of receivers that are also transmitting pilot signals, and possible interference from another cell's pilot signals.

23. The method of claim 21, wherein the allocating the pilot signals includes determining an observation window in a time-frequency domain for the pilot signals.

24. The method of claim 21, wherein the allocating the set of transmission resources includes:

staggering transmission resources for the number of pilots with respect to each other such that at least some pilots occupy transmission resources that do not occur on a rectangular grid in the delay-Doppler domain.

25. The method of claim 24, wherein the staggering is performed by shifting locations of staggered pilots from non-staggered pilots to maximize a distance in a dimension of the shift.

26. A wireless communication apparatus, comprising:

a memory storing instructions;

a processor; and

a transmitter communicatively coupled to the memory and the processor; wherein the memory stores instructions for causing the processor to implement a method recited in any of claims 1 to 25.

27. An apparatus for receiving reference signals disclosed in the present document, processing the received reference signals, and modifying an operation of the apparatus based on the processing of the reference signal.

28. A method, apparatus or computer program product disclosed in the present document.

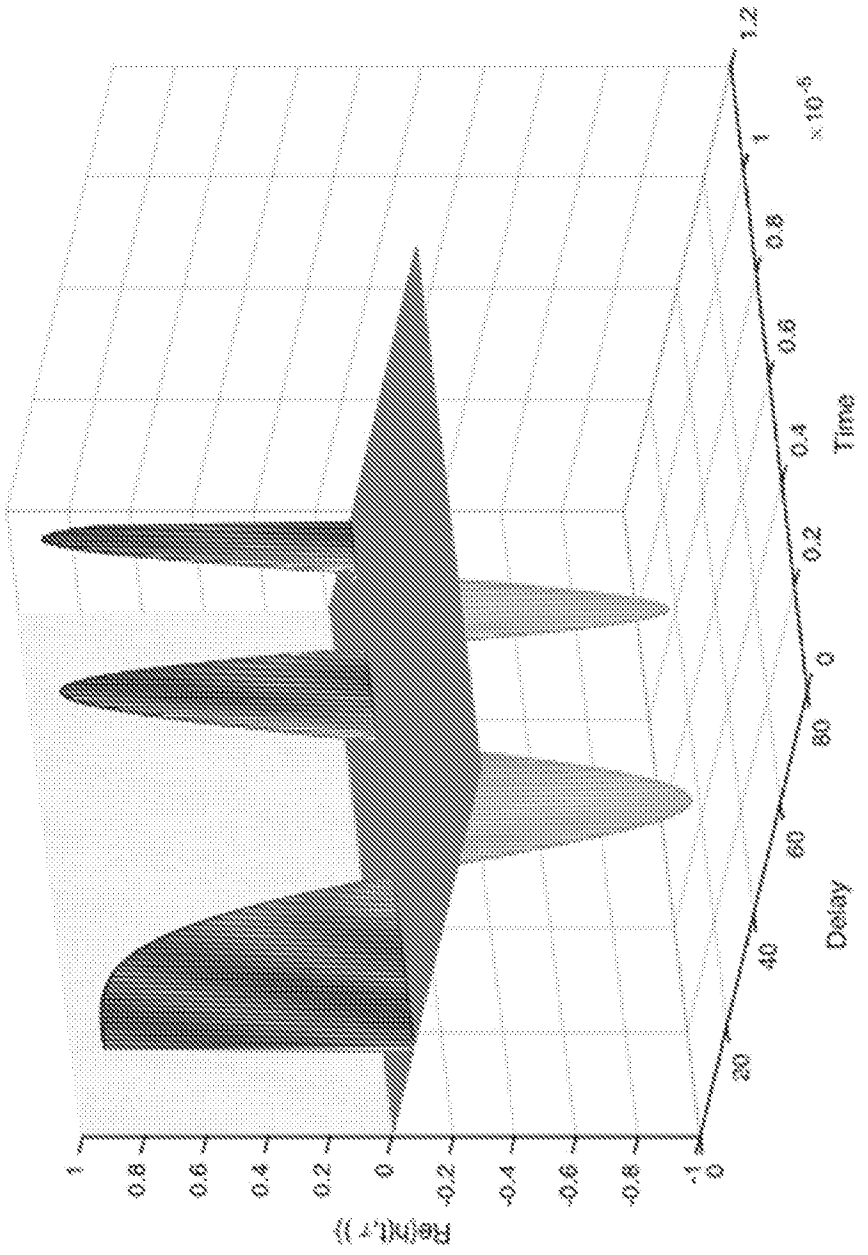

Re(N(L,τ))
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
20
40
60
80
Delay
0
0.2
0.4
0.6
0.8
1
1.2
×10^-6
Time

FIG. 1

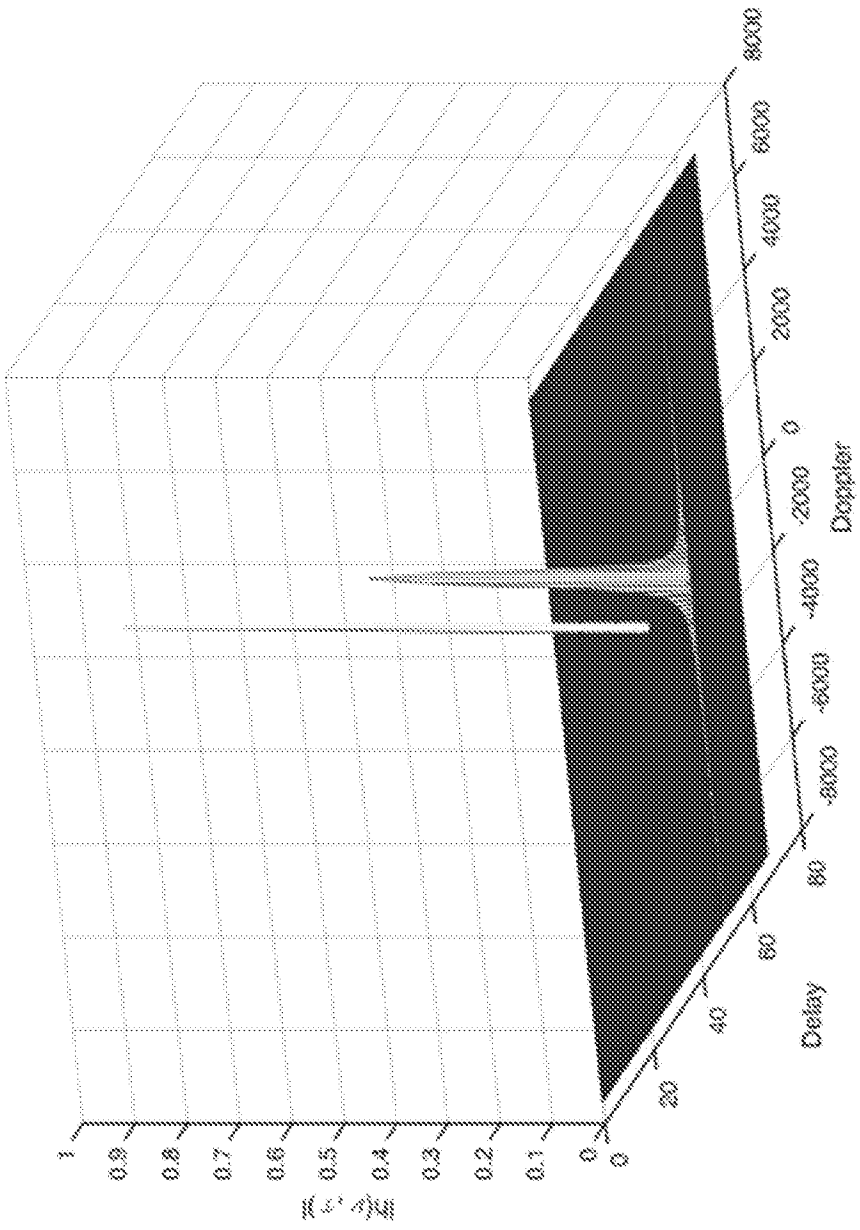

8000
6000
4000
2000
0
-2000
-4000
-6000
-8000
Doppler
0
20
40
60
80
Delay
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

FIG. 2

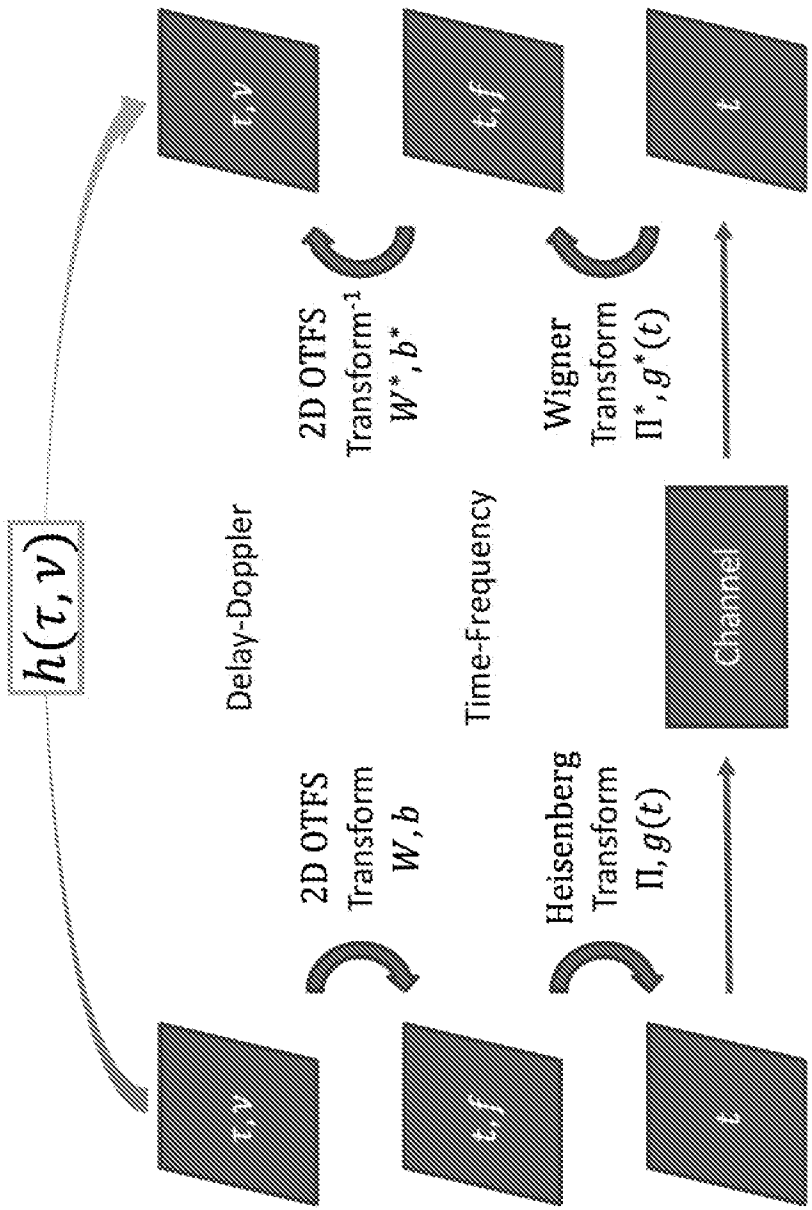

$h(\tau, \nu)$
Delay-Doppler
Time-Frequency
2D OTFS Transform
W, b
2D OTFS Transform^{-1}
W^*, b^*
Heisenberg Transform
$\Pi, g(t)$
Wigner Transform
$\Pi^*, g^*(t)$
Channel
τ, ν
t, f
t

FIG. 3

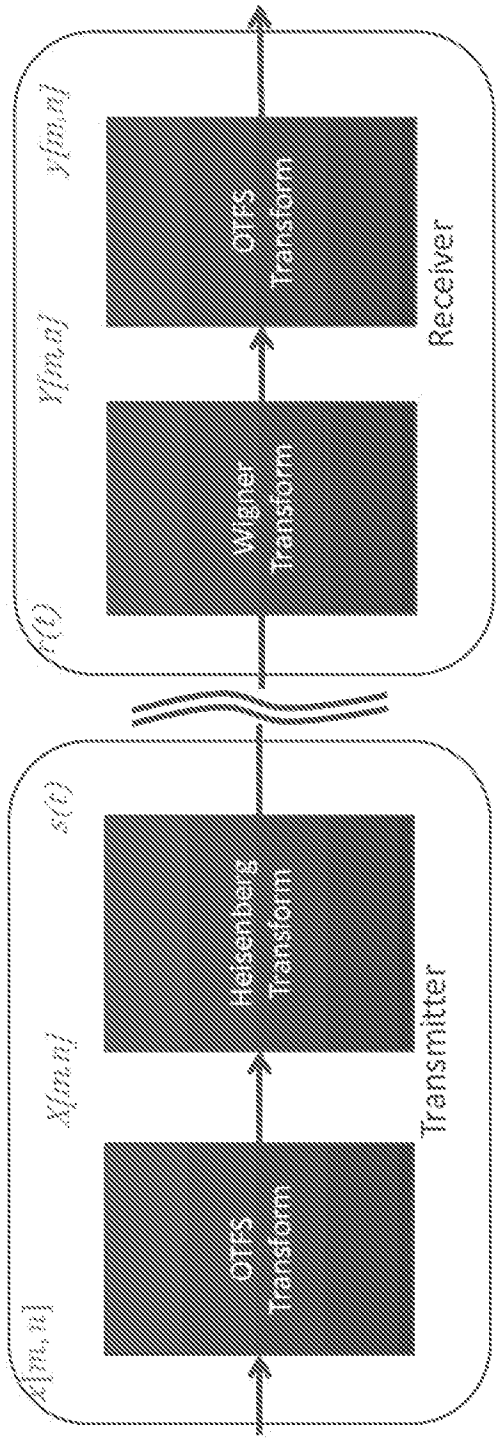

x[m,n]
X[m,n]
s(t)
OTFS Transform
Heisenberg Transform
Transmitter
r(t)
Y[m,n]
y[m,n]
Wigner Transform
OTFS Transform
Receiver

FIG. 4

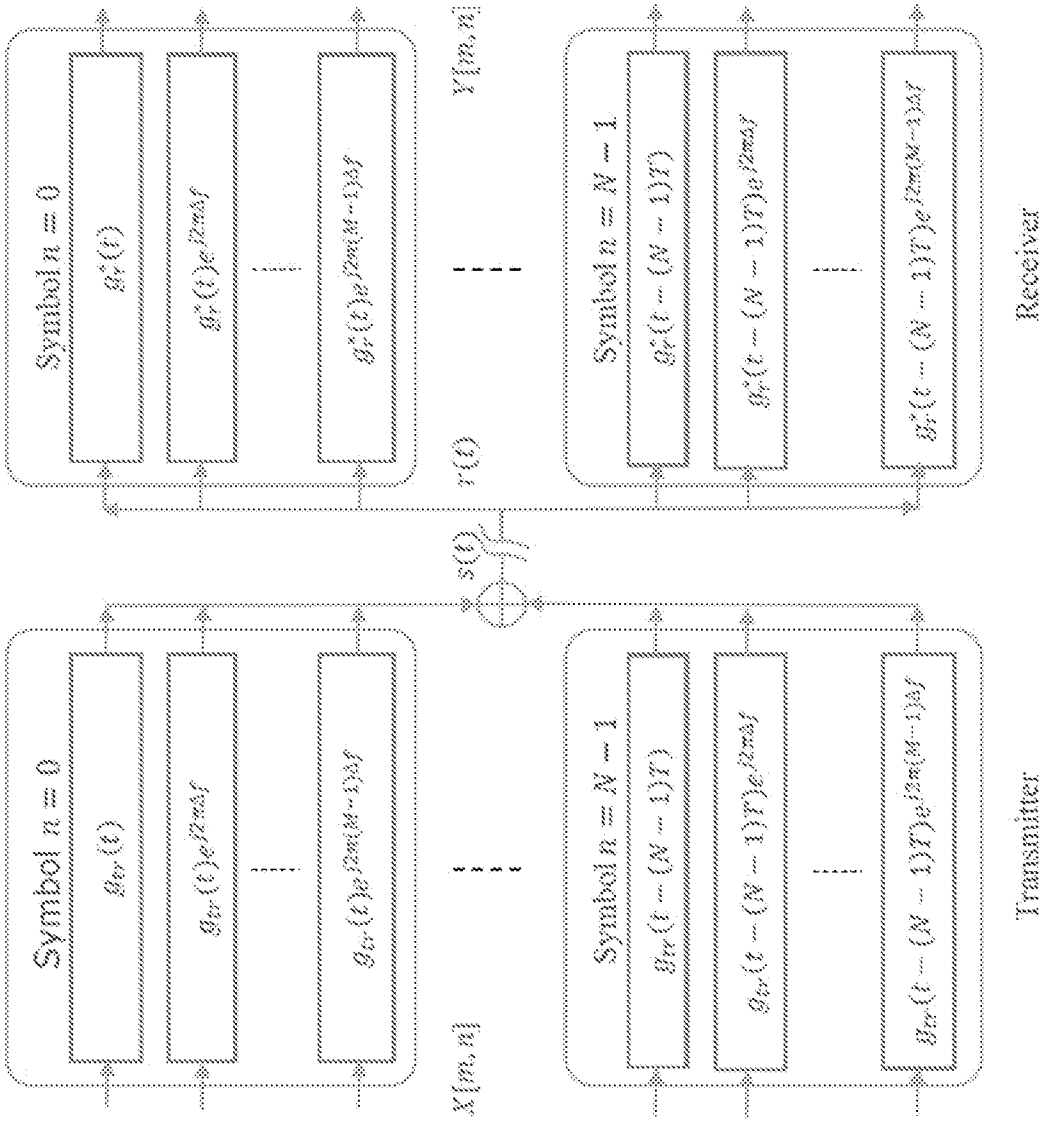

Symbol $n = 0$
$g_{tr}(t)$
$g_{tr}(t)e^{j2\pi\Delta f t}$
$g_{tr}(t)e^{j2\pi(M-1)\Delta f t}$
Symbol $n = N-1$
$g_{tr}(t-(N-1)T)$
$g_{tr}(t-(N-1)T)e^{j2\pi\Delta f t}$
$g_{tr}(t-(N-1)T)e^{j2\pi(M-1)\Delta f t}$
$X[m,n]$
$s(t)$
$r(t)$
$Y[m,n]$
Symbol $n = 0$
$g_r(t)$
$g_r(t)e^{j2\pi\Delta f t}$
$g_r(t)e^{j2\pi(M-1)\Delta f t}$
Symbol $n = N-1$
$g_r(t-(N-1)T)$
$g_r(t-(N-1)T)e^{j2\pi\Delta f t}$
$g_r(t-(N-1)T)e^{j2\pi(M-1)\Delta f t}$
Transmitter
Receiver

FIG. 5

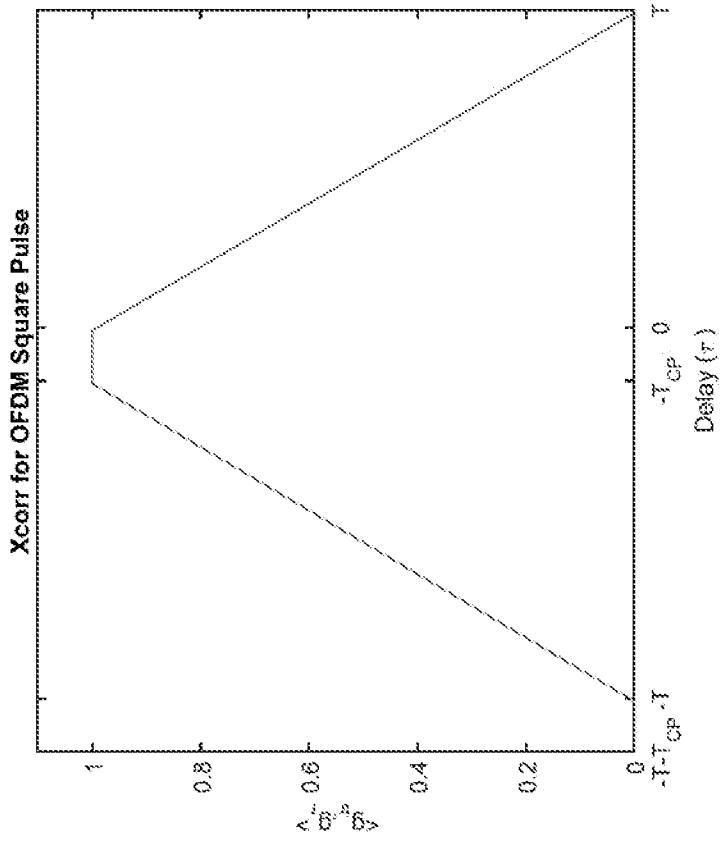


FIG. 6

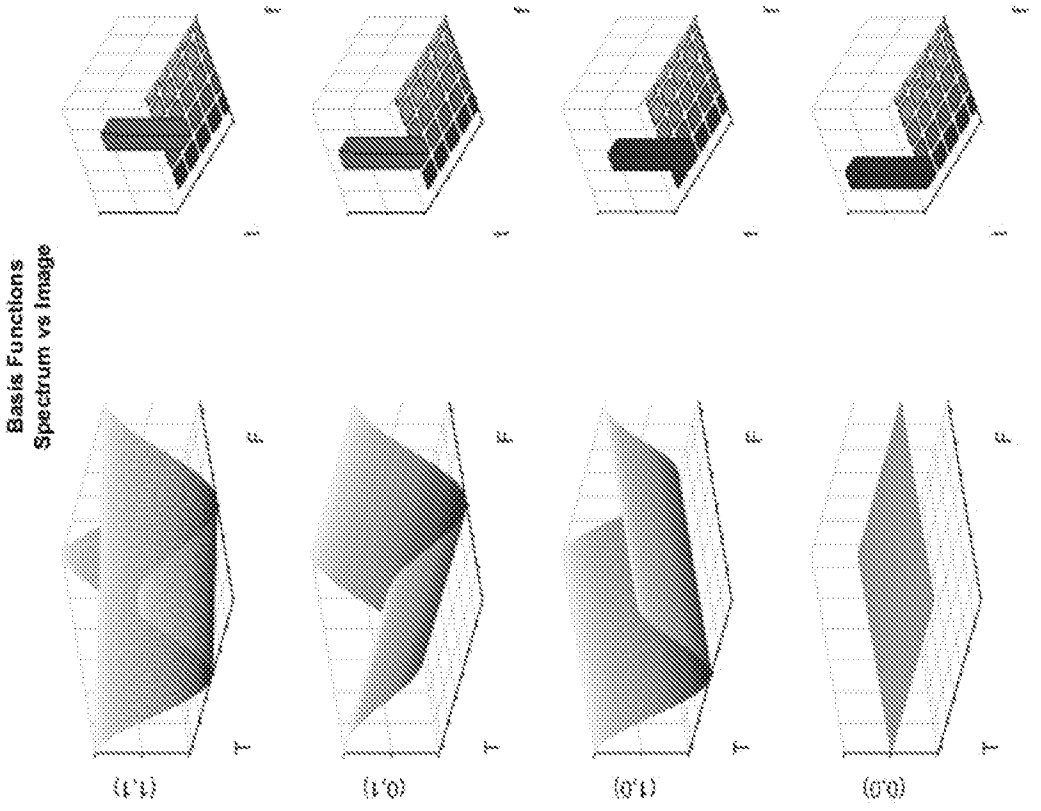


FIG. 7

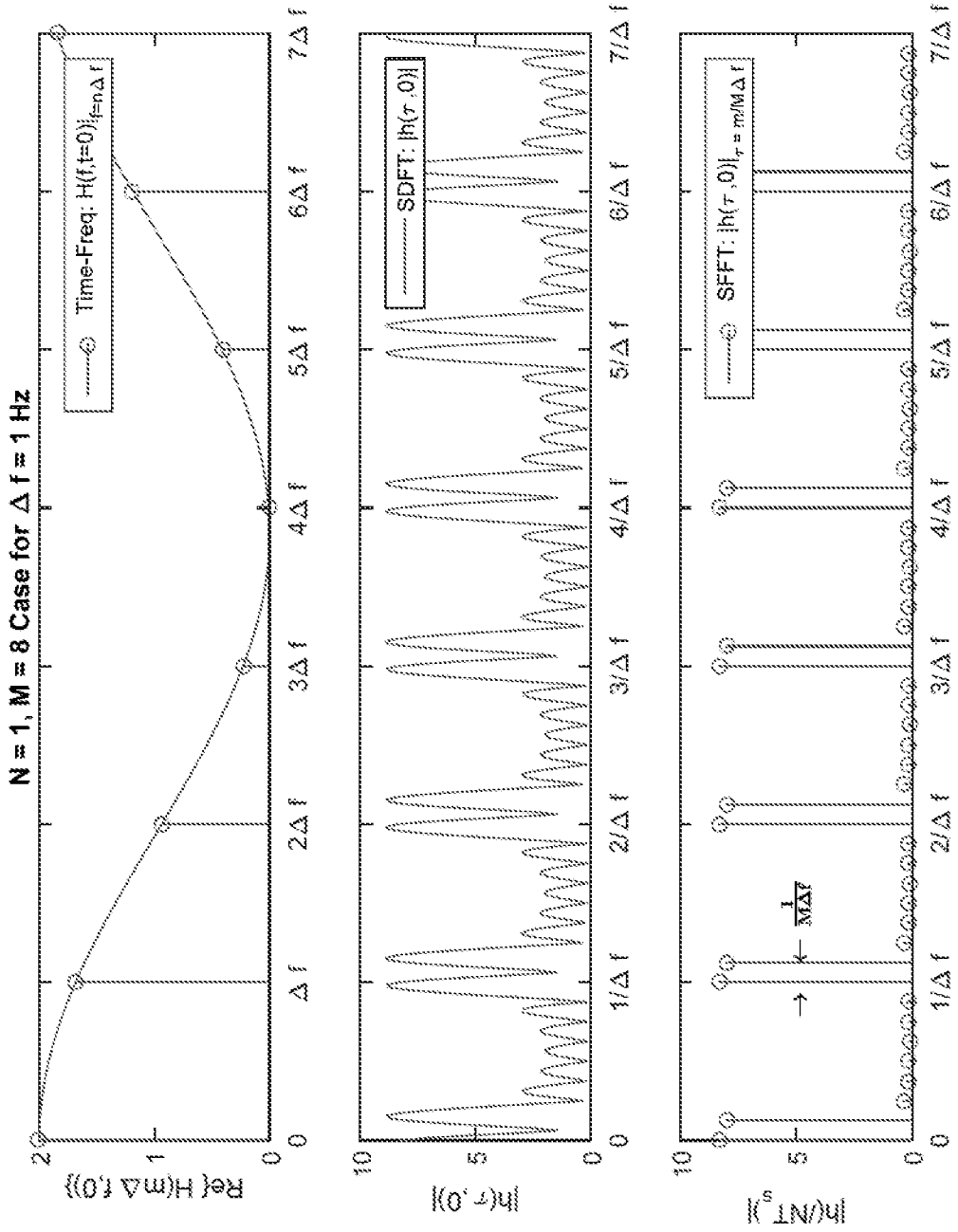

N = 1, M = 8 Case for Δ f = 1 Hz
Time-Freq: H(f,t=0)|f=nΔ f
Re{ H(mΔ f,0)}
SDFT: |h(τ ,0)|
|h(τ ,0)|
SFFT: |h(τ ,0)|τ = m/MΔ f
|h(/NT_s)|
1/MΔf
0
Δ f
2Δ f
3Δ f
4Δ f
5Δ f
6Δ f
7Δ f
1/Δ f
2/Δ f
3/Δ f
4/Δ f
5/Δ f
6/Δ f
7/Δ f
2
1
10
5

FIG. 8

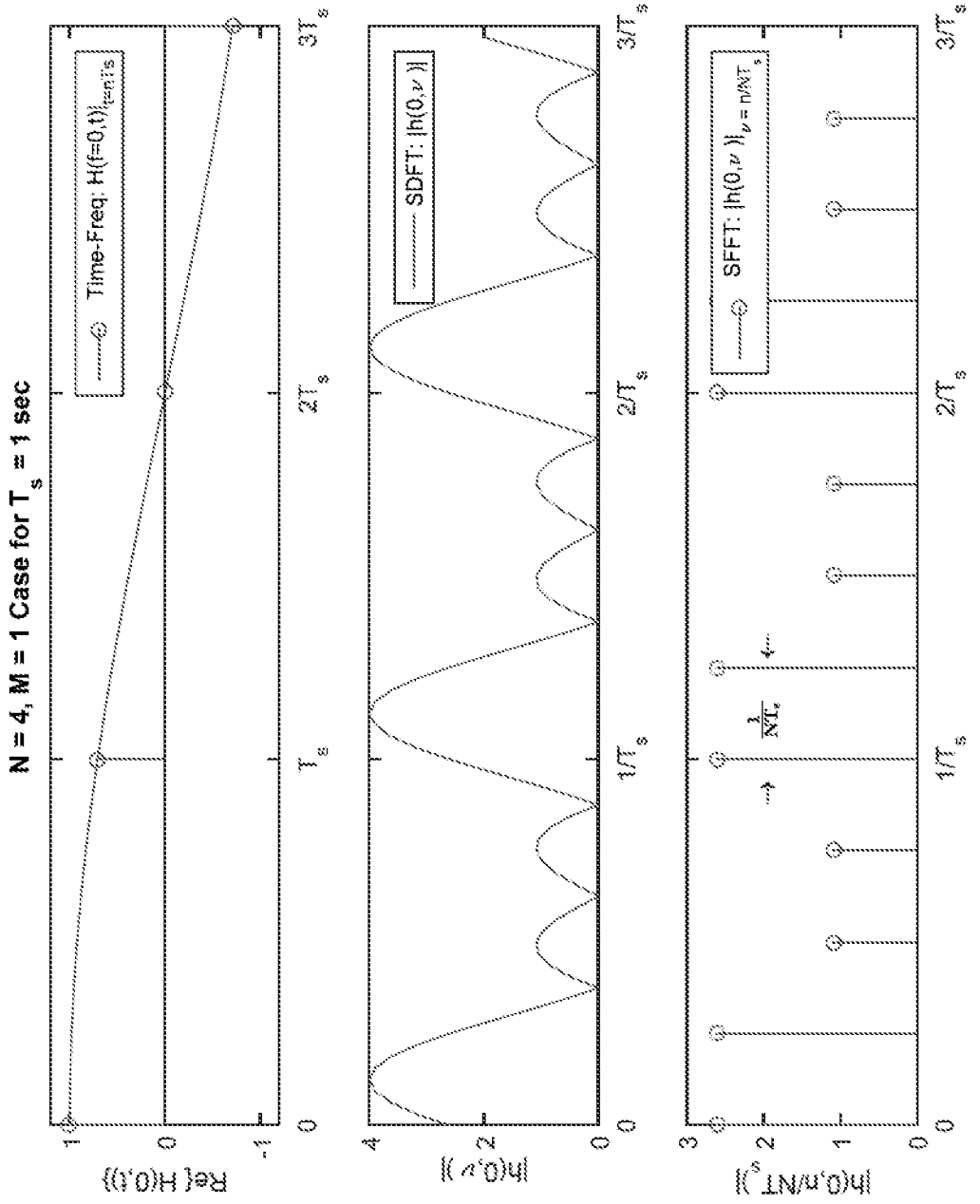

N = 4, M = 1 Case for T_s = 1 sec
Time-Freq: H(f=0,t)|t=n/Ts
Re{ H(0,t) }
SDFT: |h(0,ν)|
|h(0,ν)|
SFFT: |h(0,ν)|ν = n/NTs
|h(0,n/NTs)|
1/Ts
2/Ts
3/Ts
Ts
2Ts
3Ts
1/NTs

FIG. 9

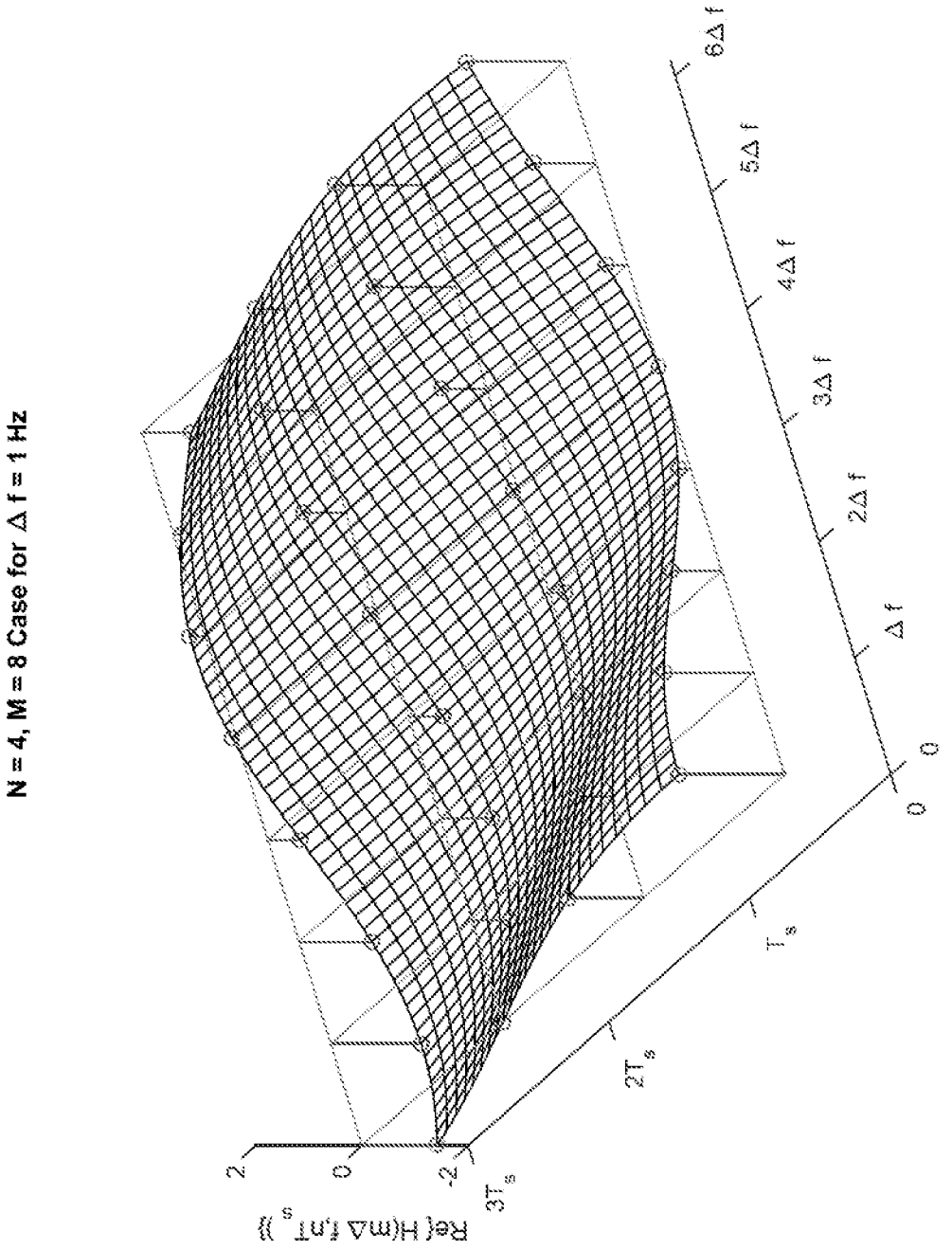

N = 4, M = 8 Case for Δ f = 1 Hz
Re{ H(mΔ f,nT_s)}
2
0
-2
3T_s
2T_s
T_s
0
0
Δ f
2Δ f
3Δ f
4Δ f
5Δ f
6Δ f

FIG. 10

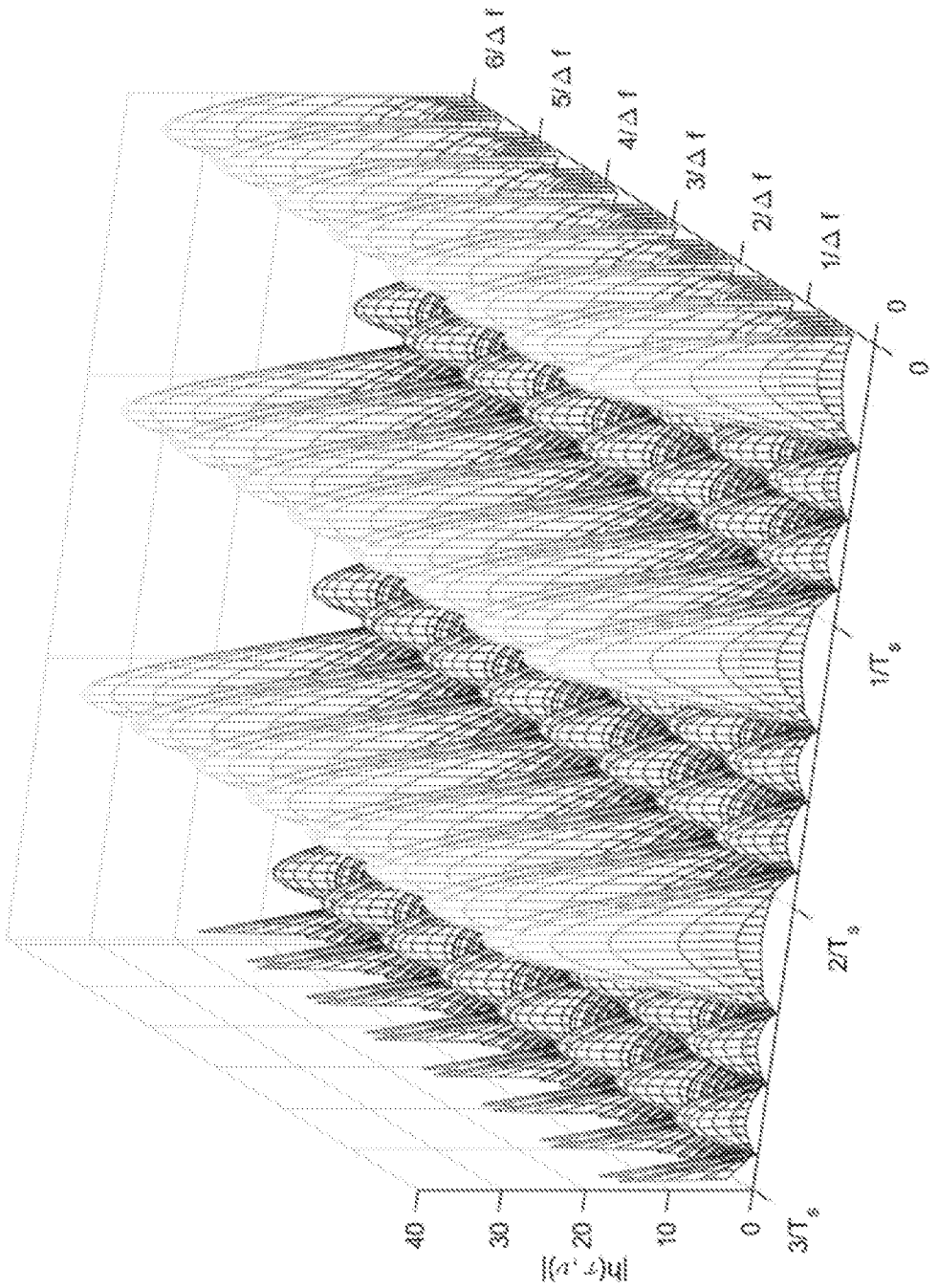


FIG. 11

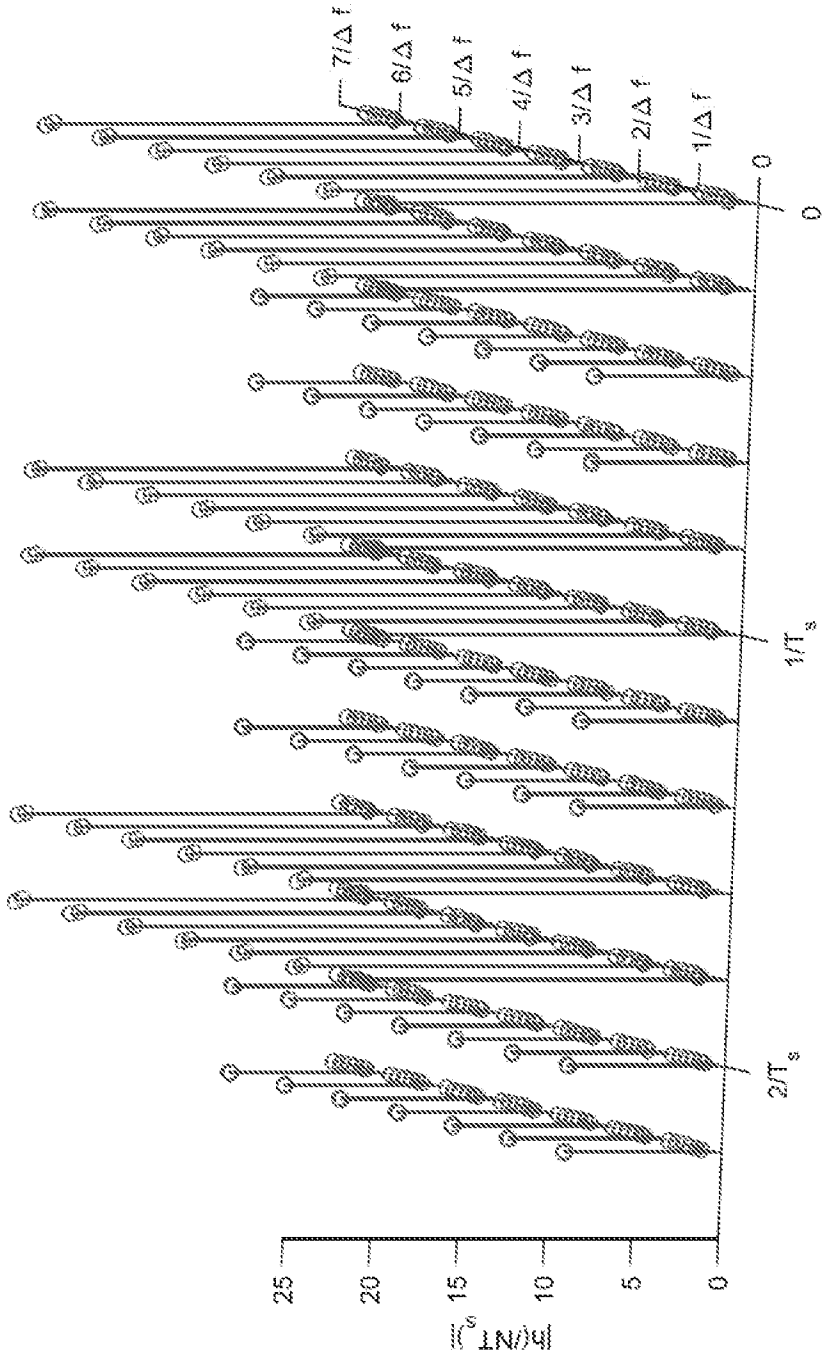


FIG. 12

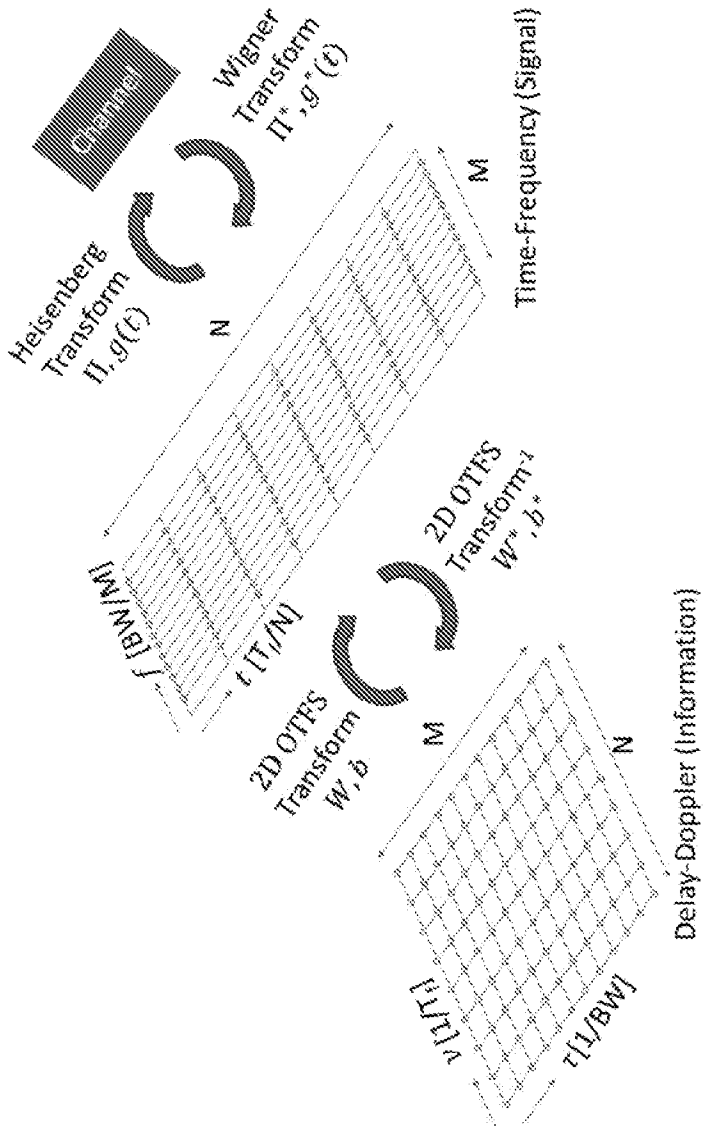

Heisenberg Transform
Π, g(t)
Wigner Transform
Π*, g*(t)
N
M
Time-Frequency (Signal)
f [BW/M]
t [T/N]
2D OTFS Transform
W, b
2D OTFS Transform
W*, b*
M
N
Delay-Doppler (Information)
ν [1/T]
τ [1/BW]

FIG. 13

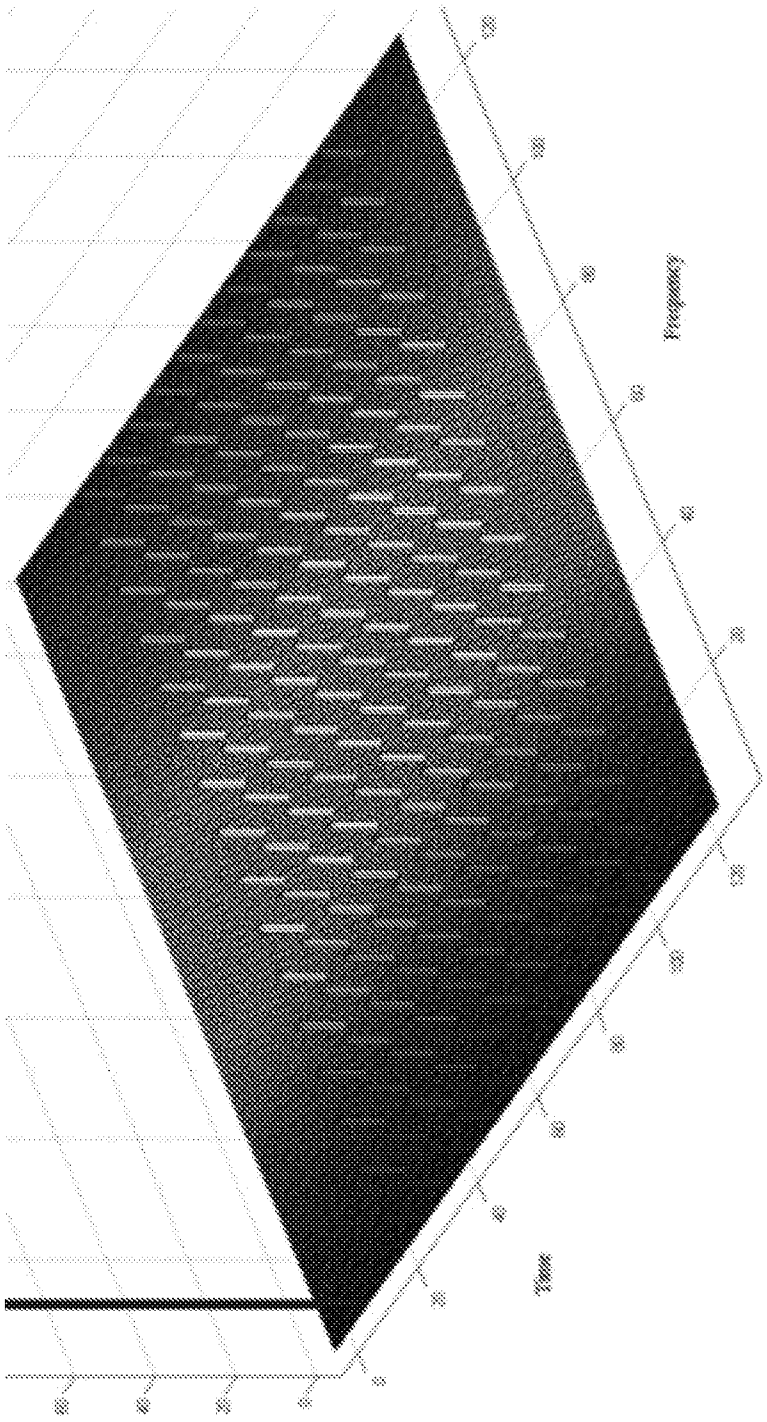


FIG. 14

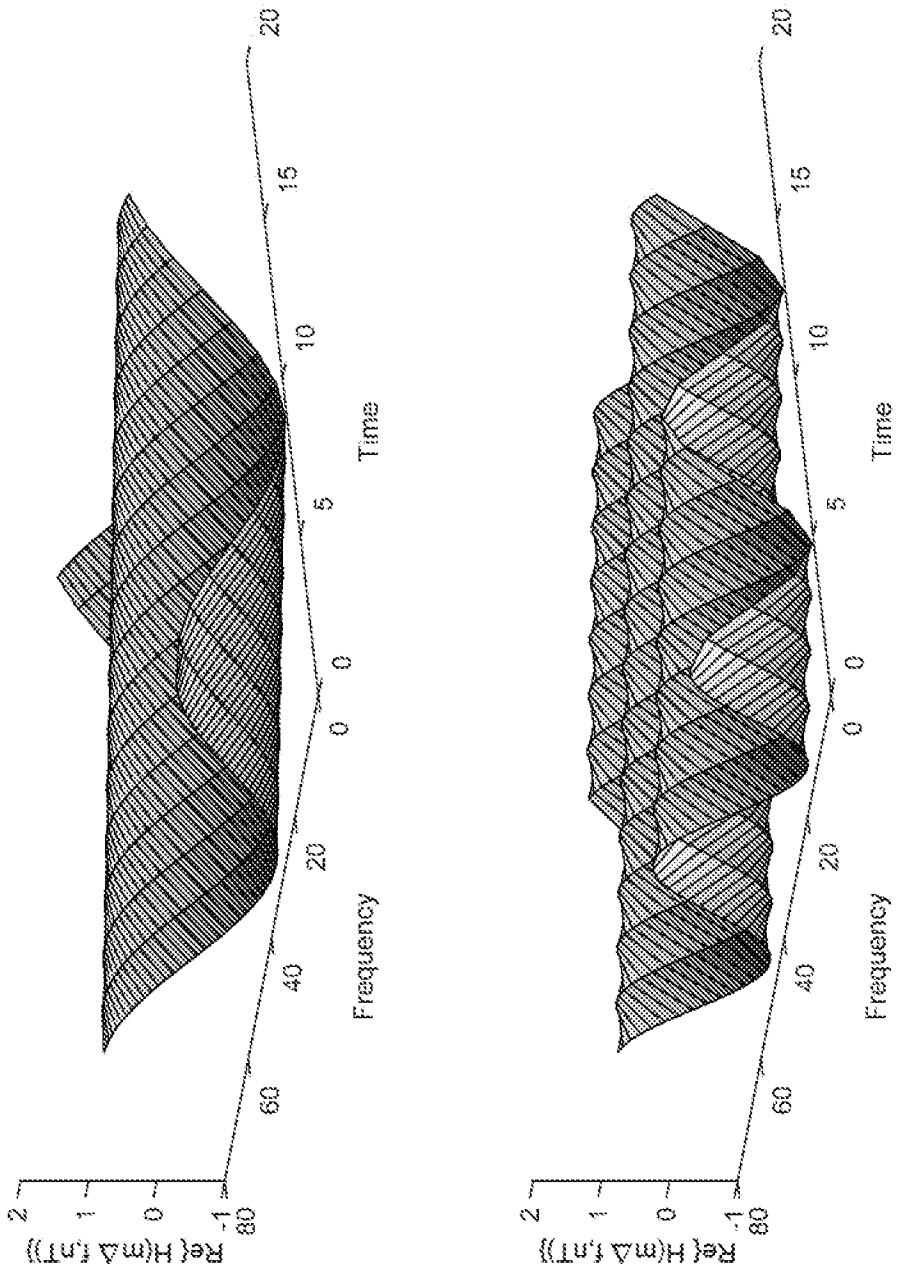

Re{ H(mΔ f,nT)}
2
1
0
-1
80
60
40
20
0
0
5
10
15
20
Frequency
Time
Re{ H(mΔ f,nT)}
2
1
0
-1
80
60
40
20
0
0
5
10
15
20
Frequency
Time

FIG. 15

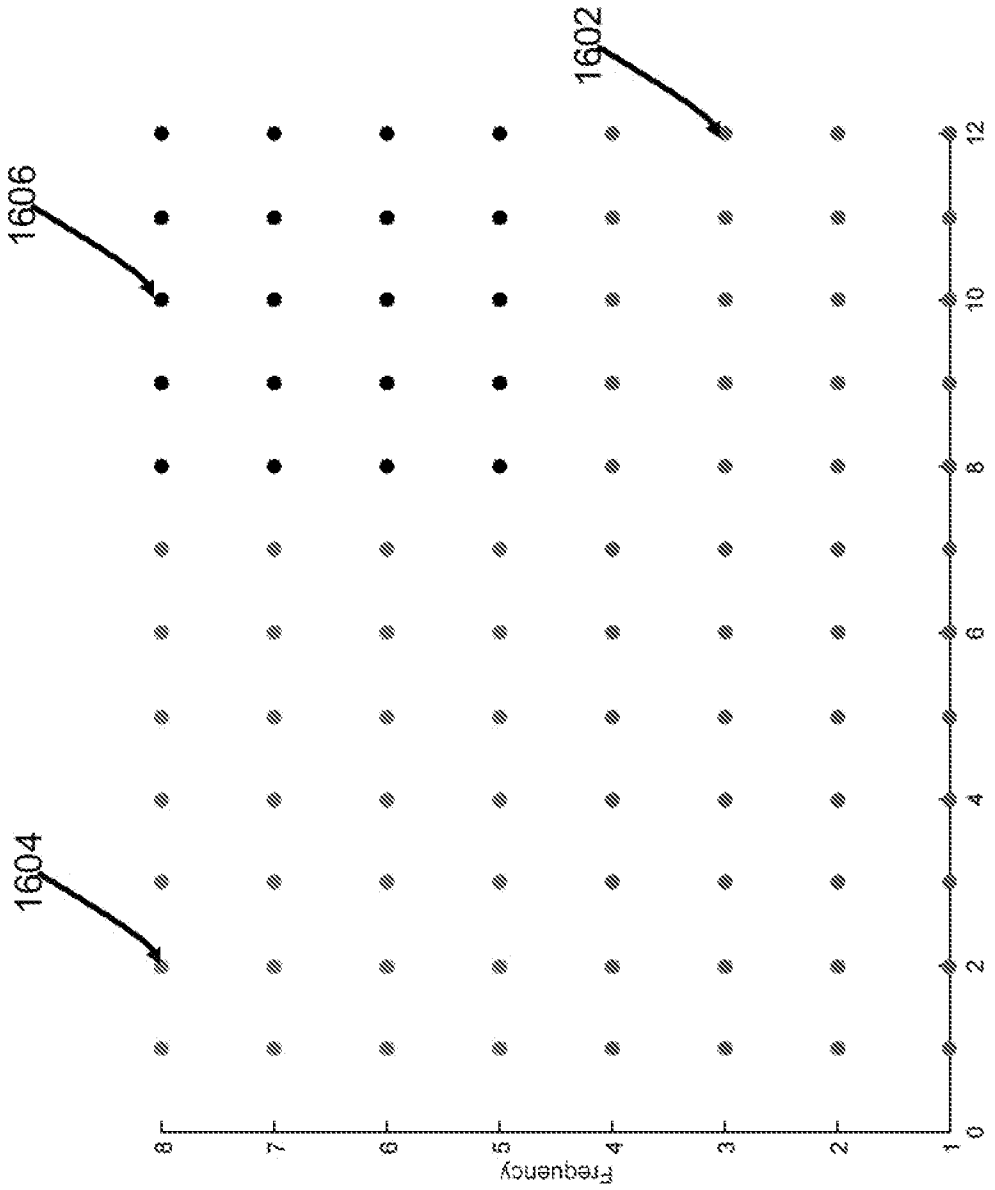


FIG. 16

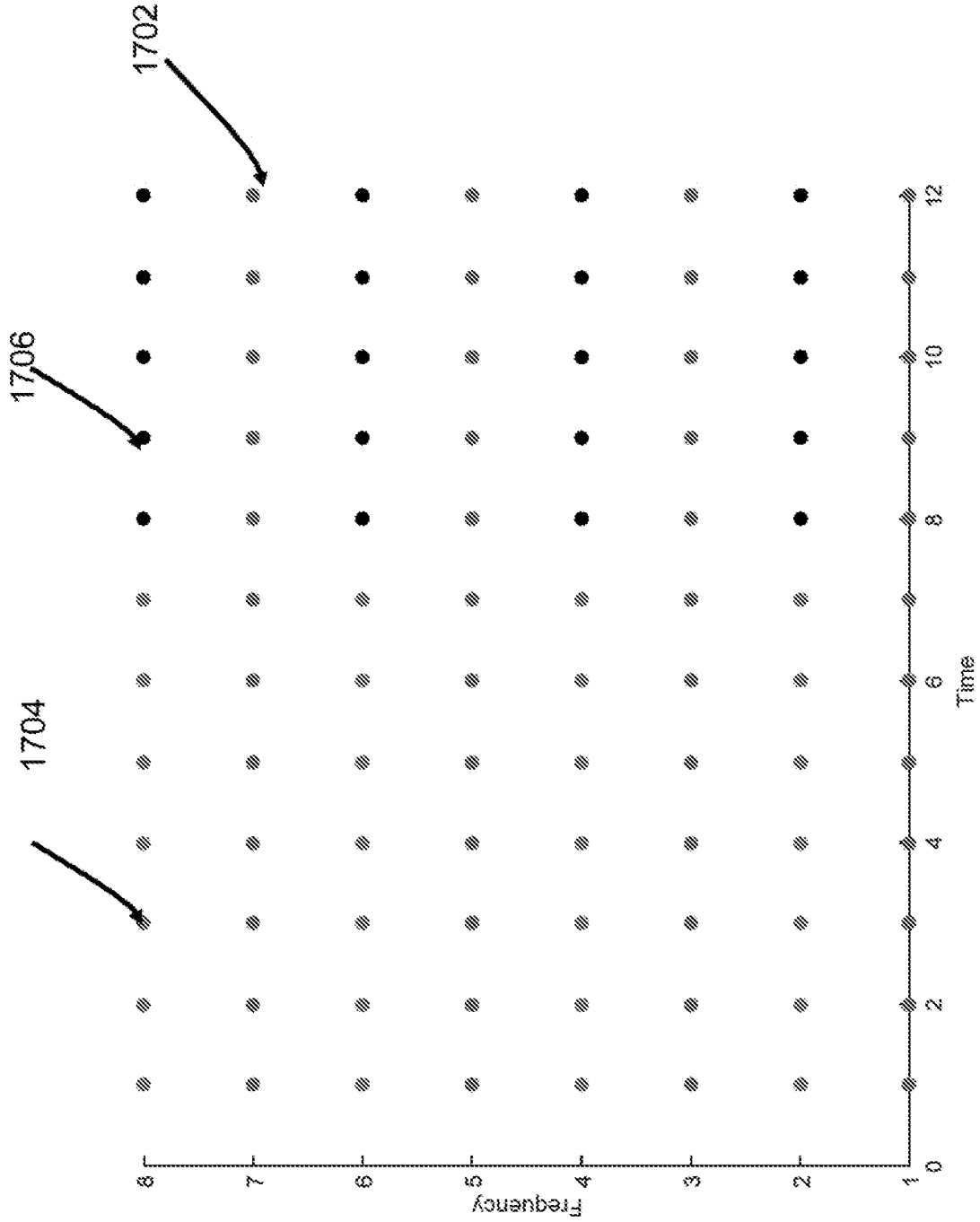


FIG. 17

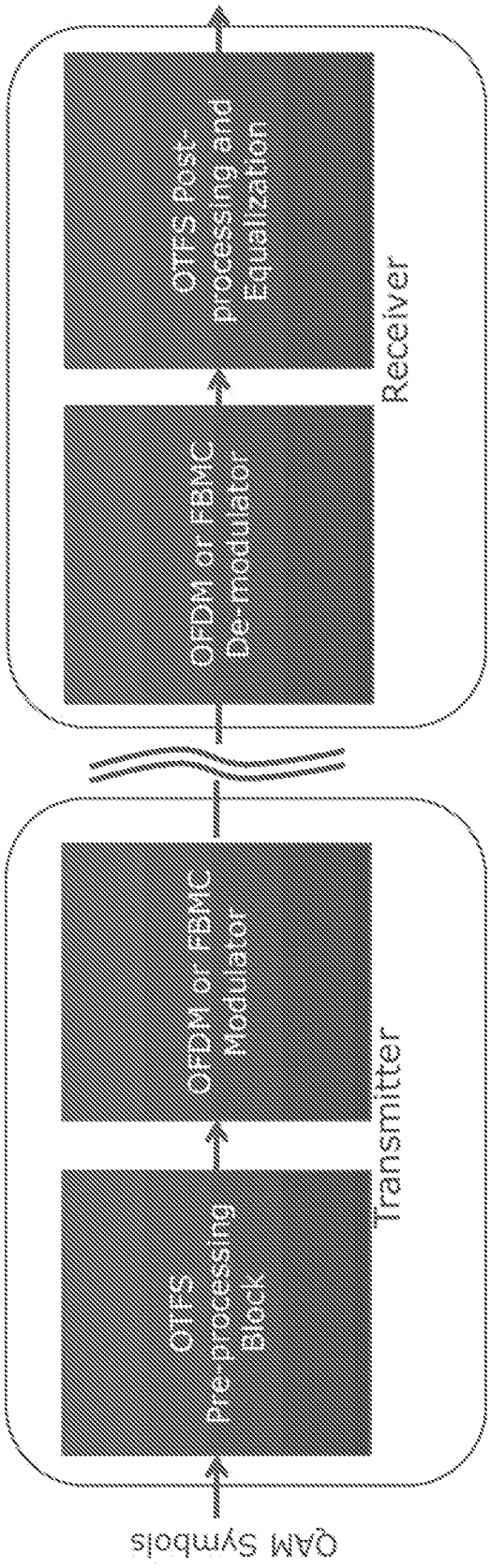

QAM Symbols
OTFS Pre-processing Block
OFDM or FBMC Modulator
Transmitter
OFDM or FBMC De-modulator
OTFS Post-processing and Equalization
Receiver

FIG. 18

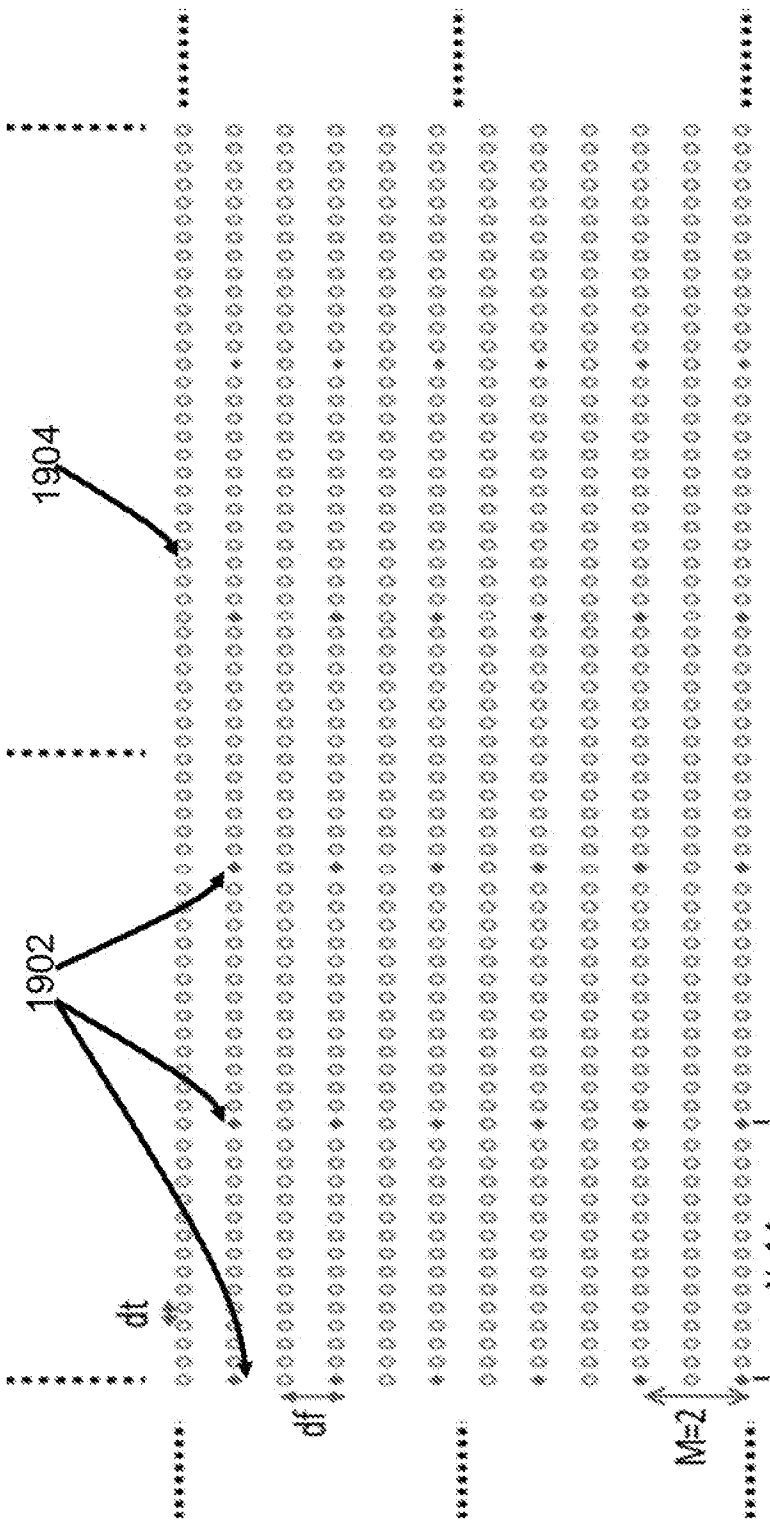


FIG. 19

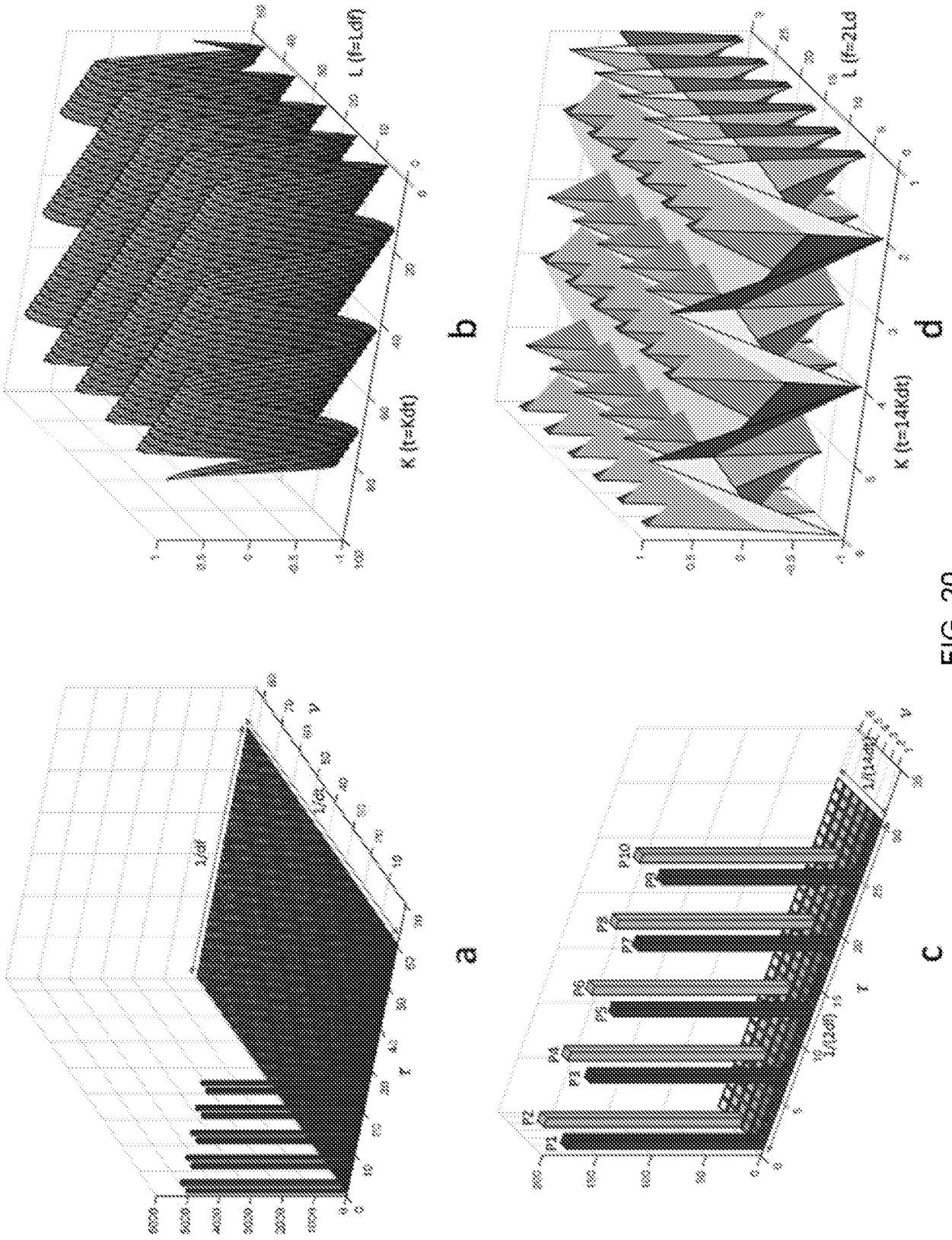


FIG. 20

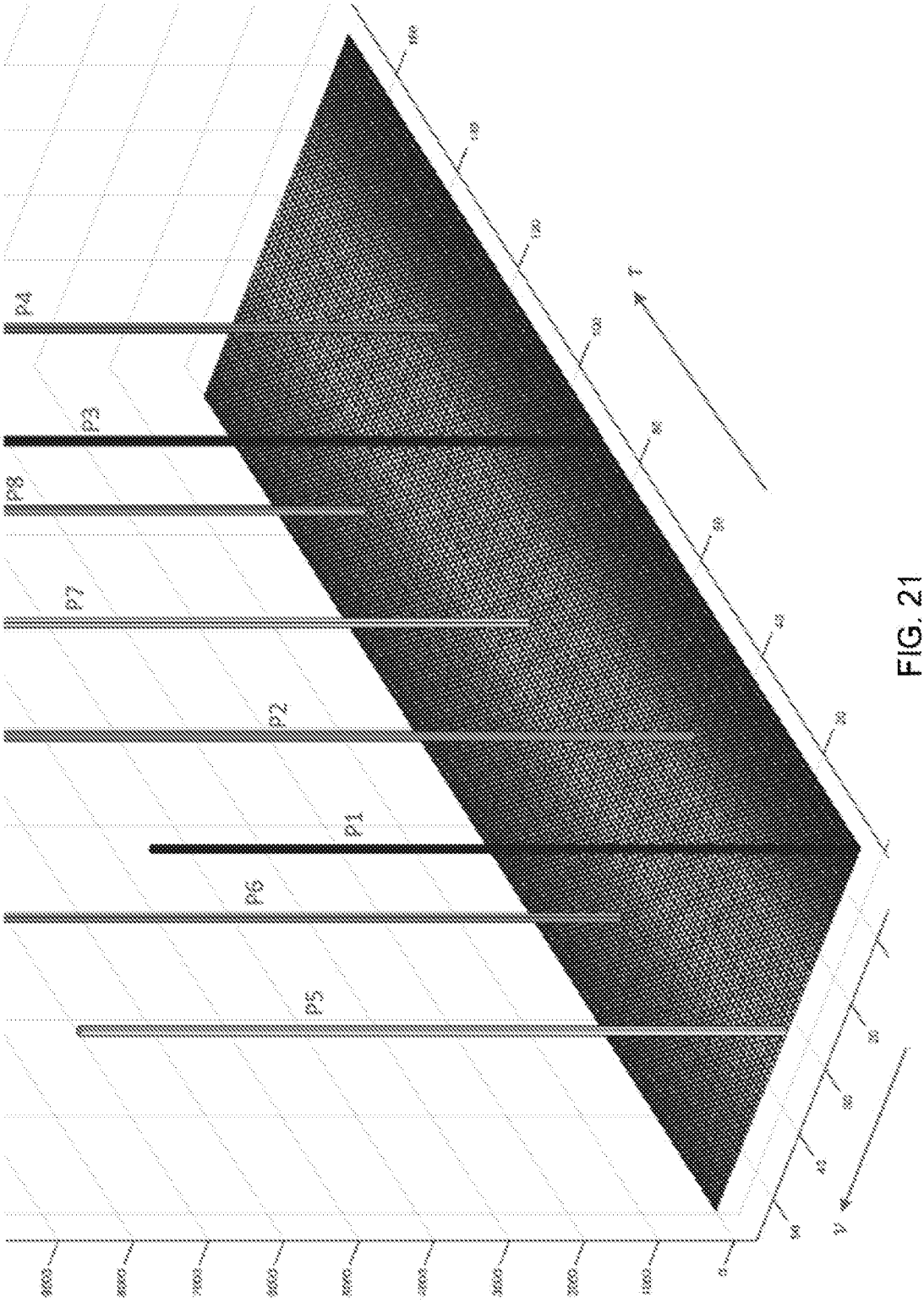


FIG. 21

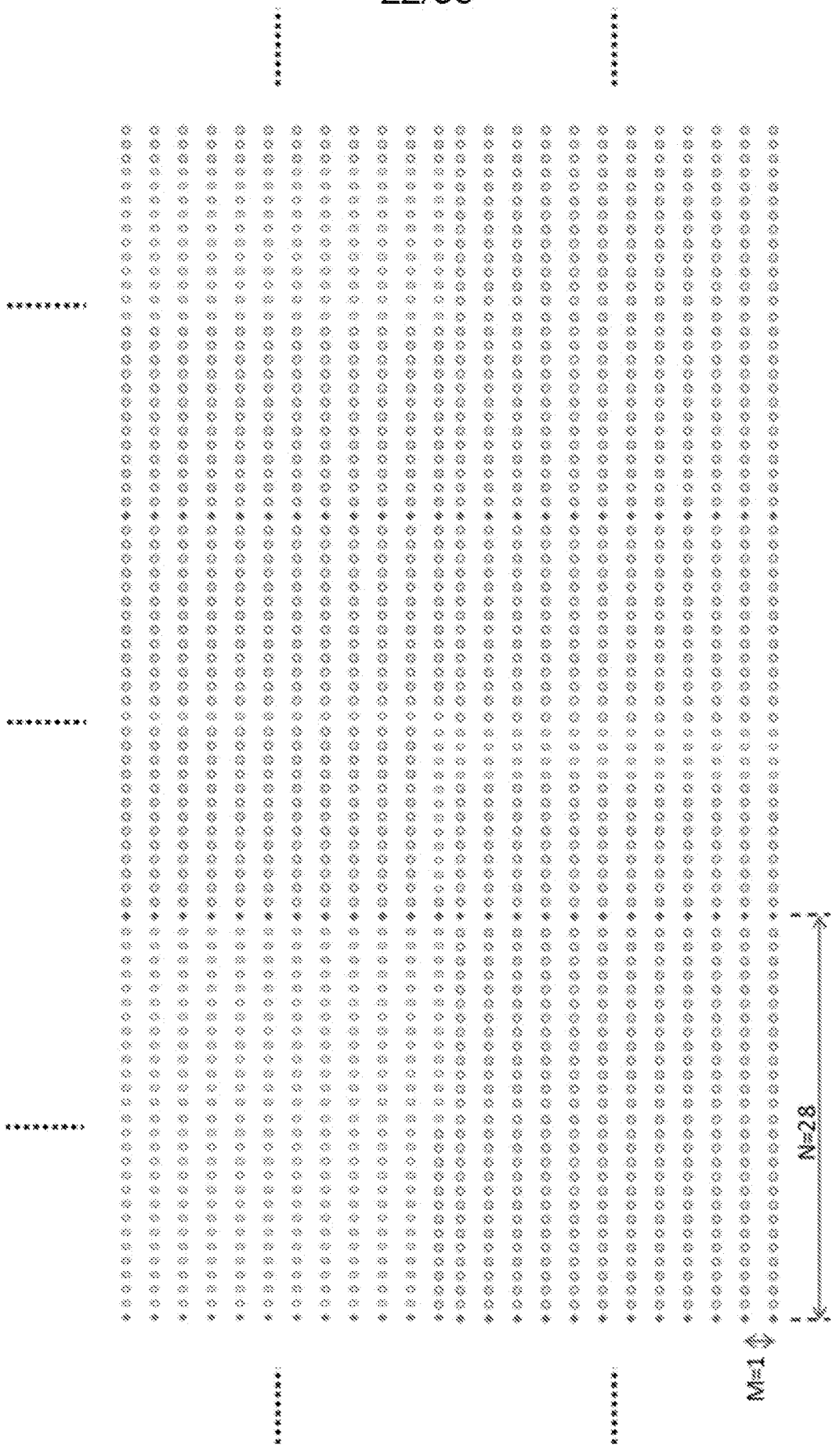


FIG. 22

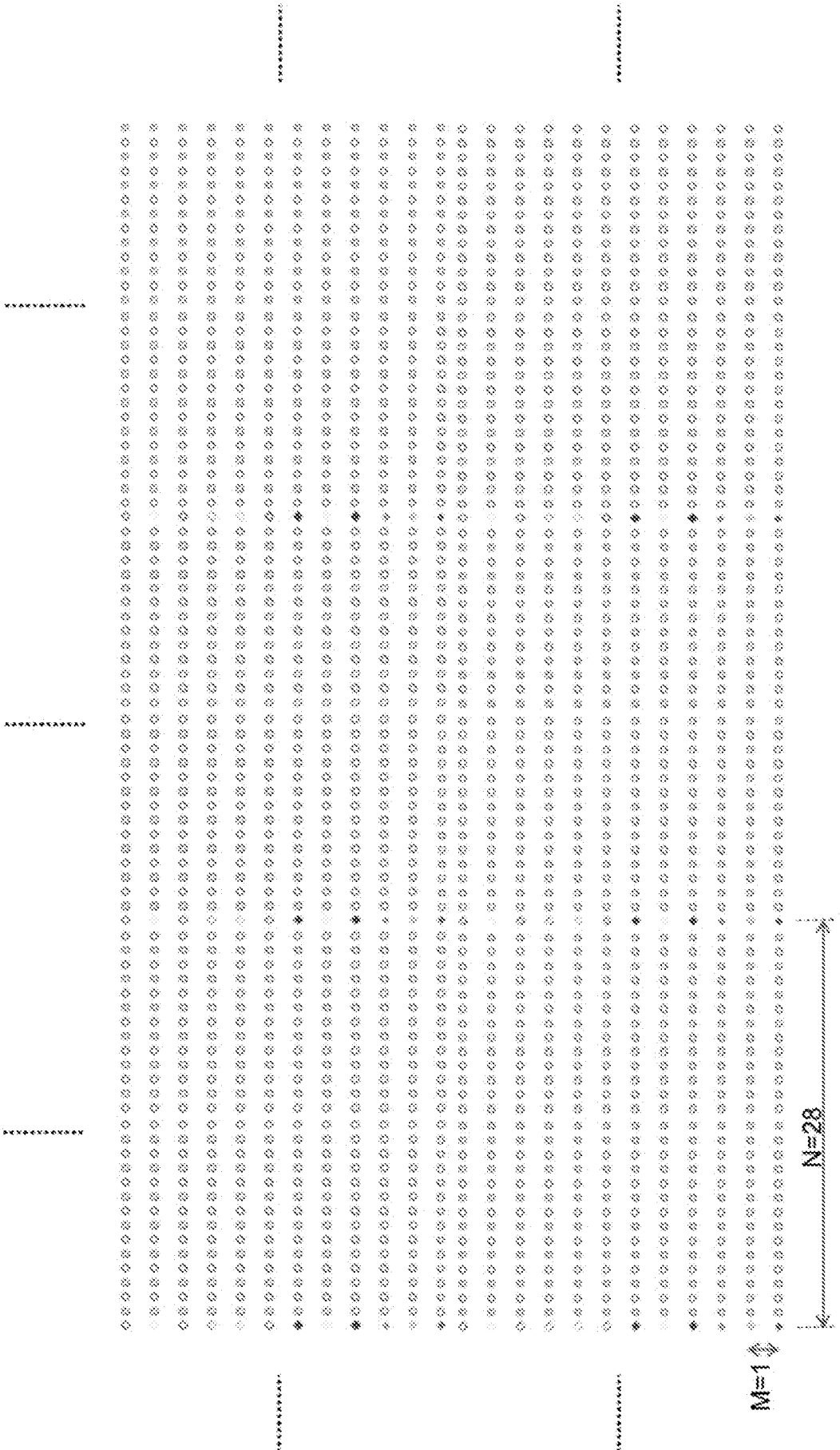


FIG. 23

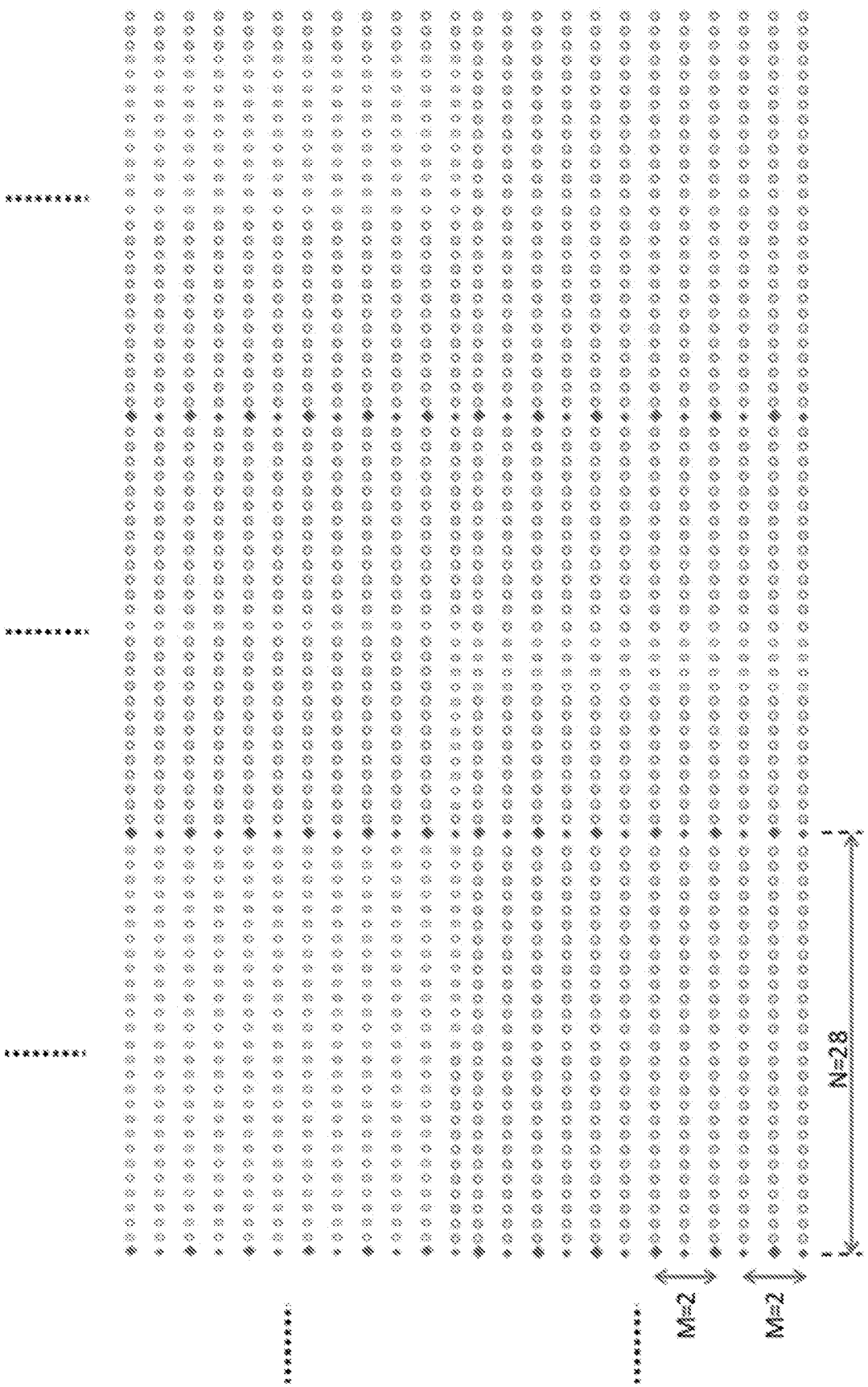


FIG. 24

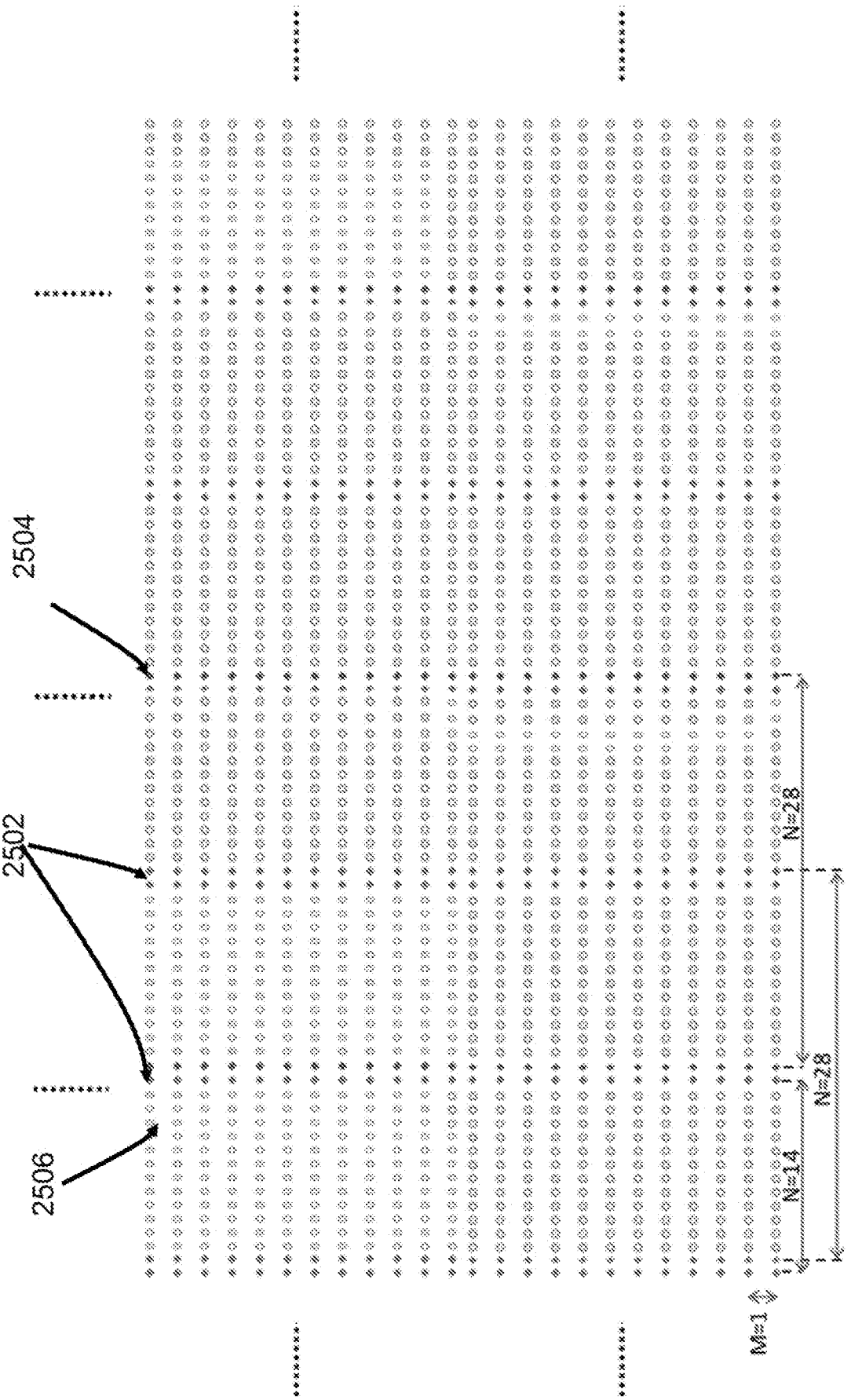


FIG. 25

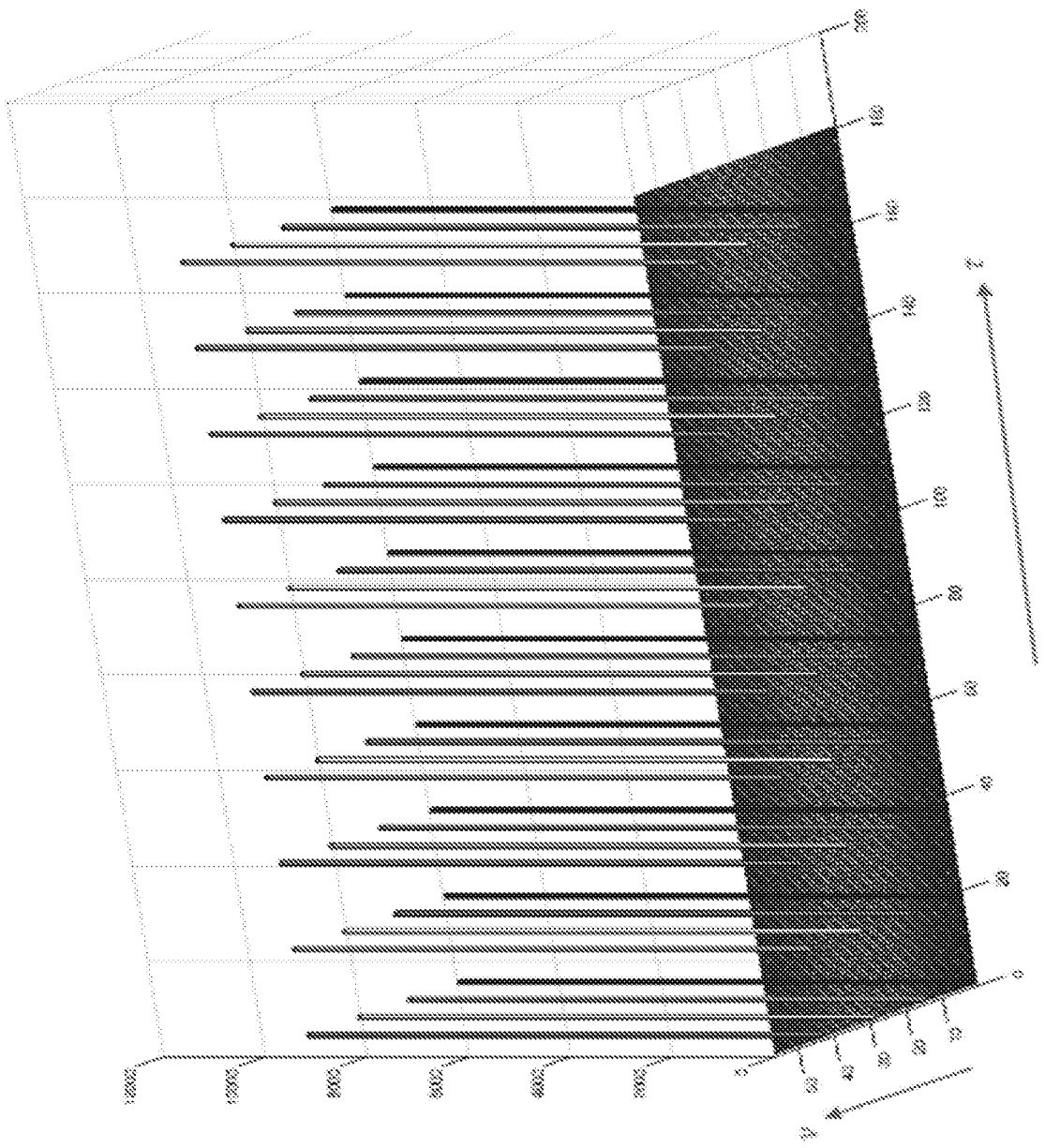


FIG. 26

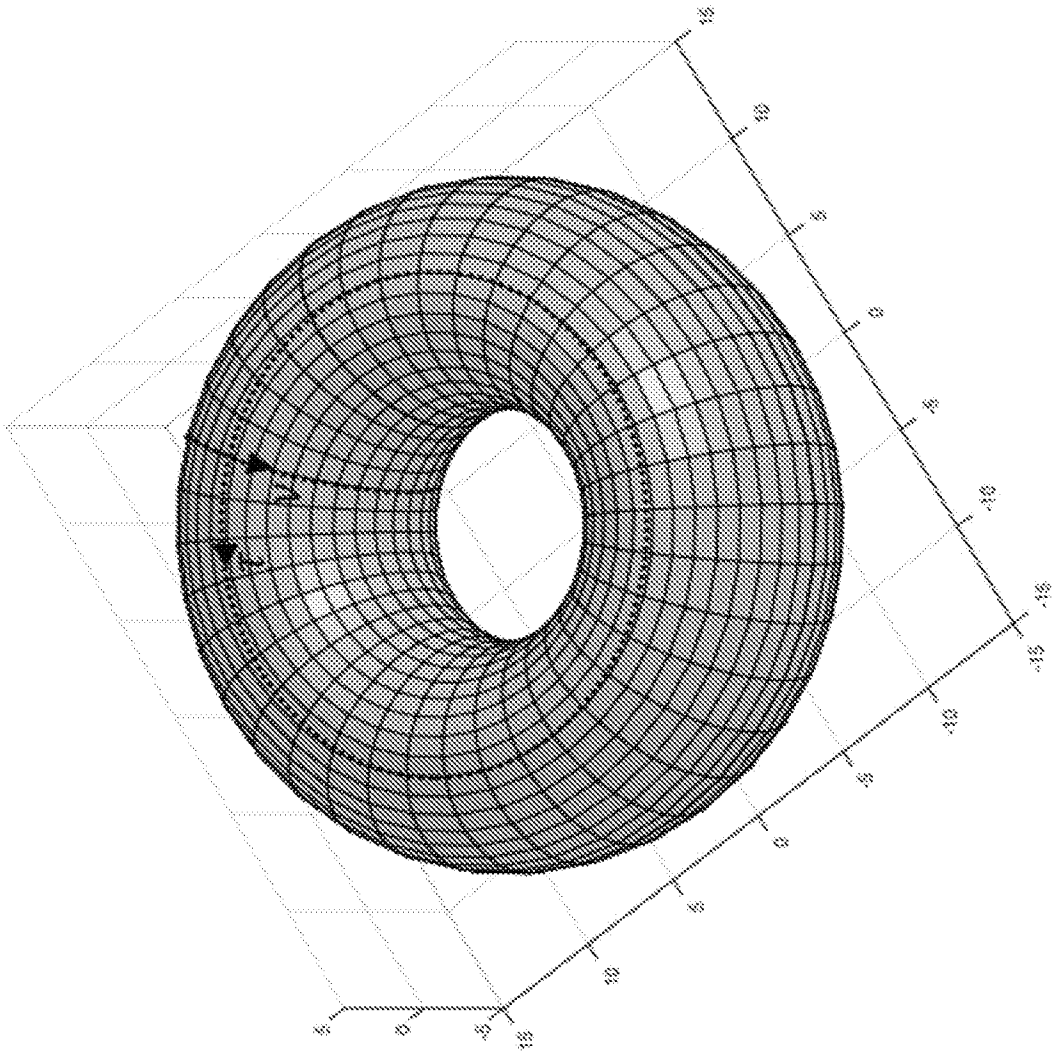


FIG. 27

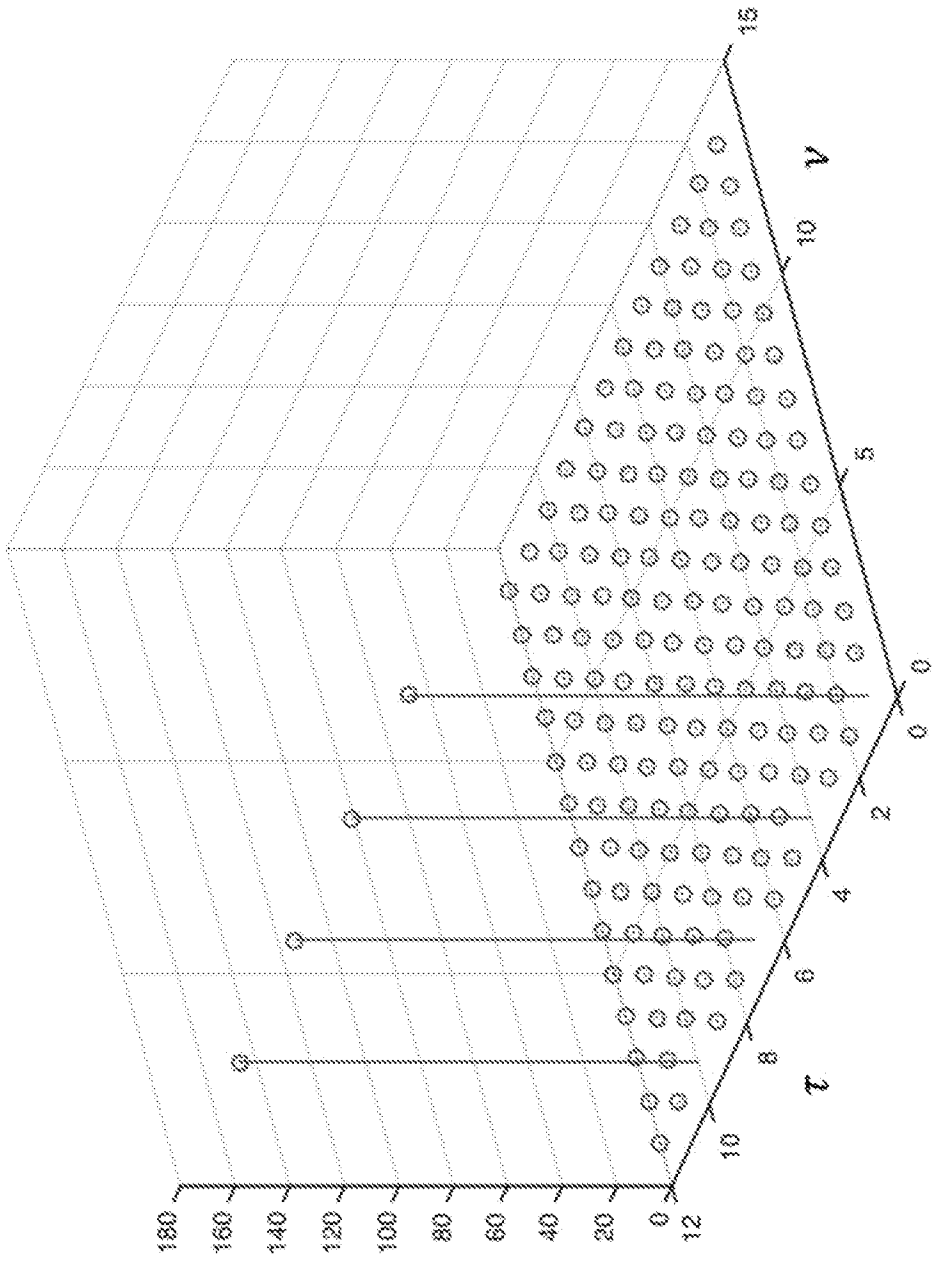


FIG. 28

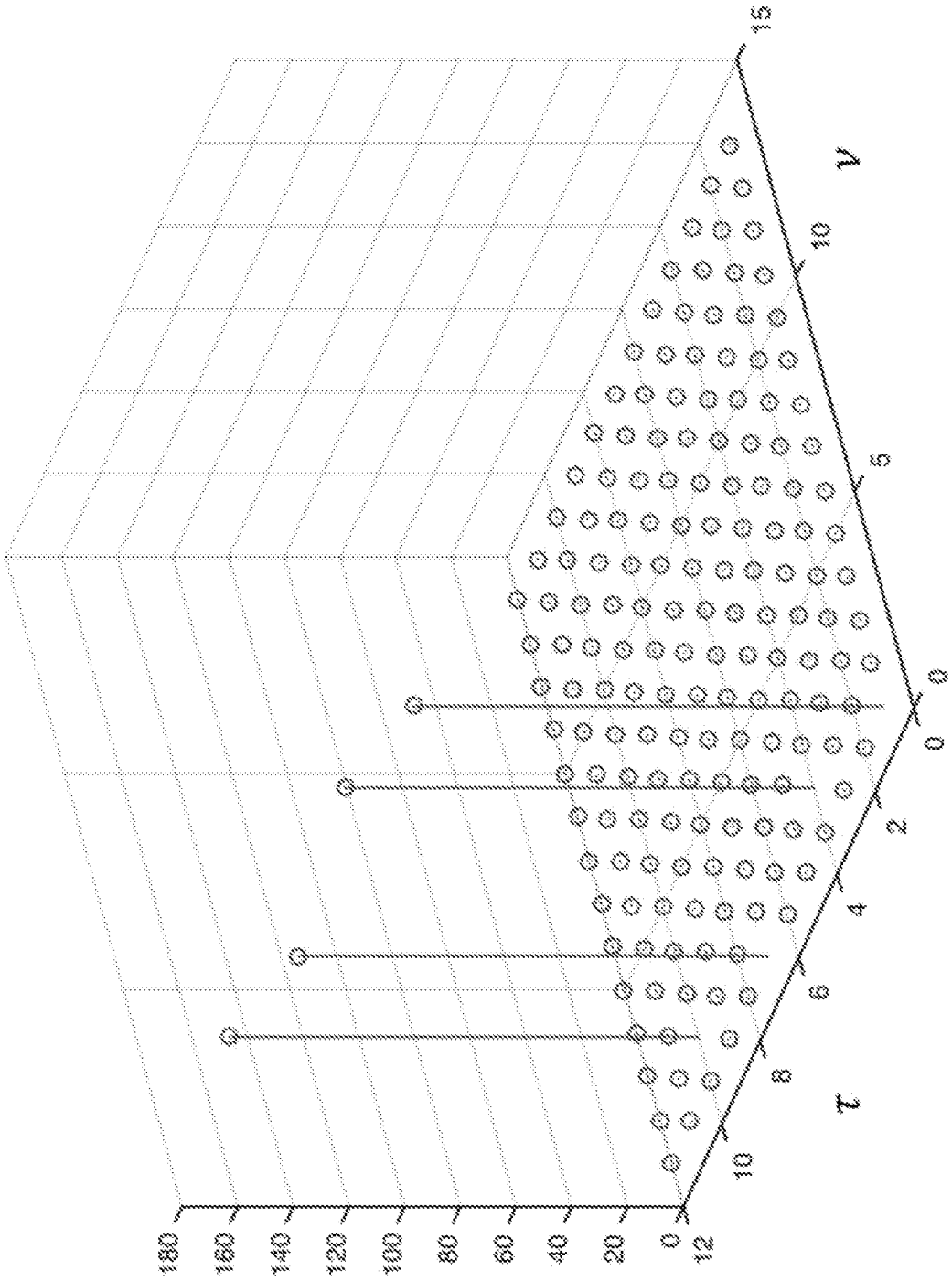


FIG. 29

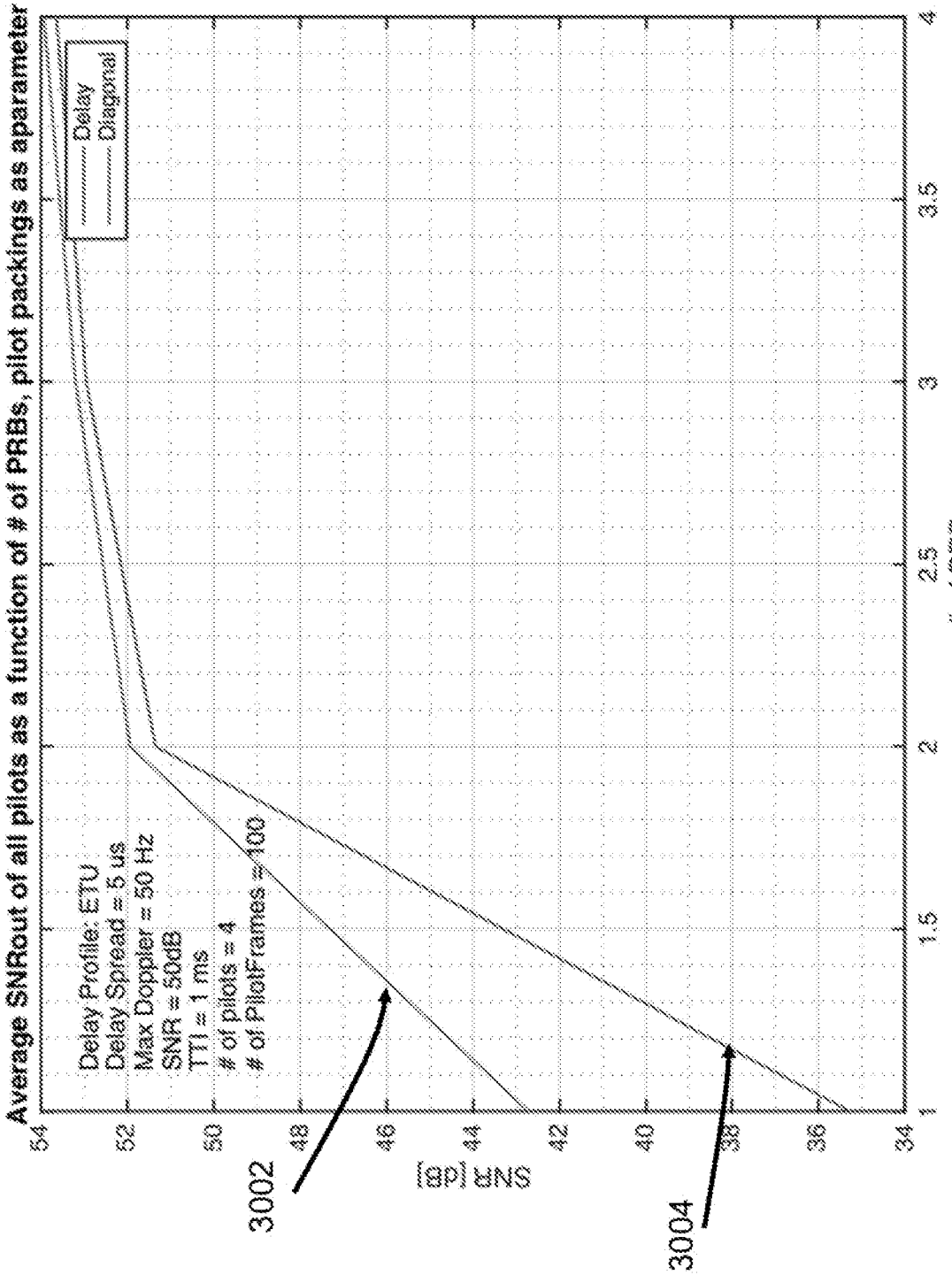

Average SNRout of all pilots as a function of # of PRBs, pilot packings as aparameter
Delay
Diagonal
Delay Profile: ETU
Delay Spread = 5 us
Max Doppler = 50 Hz
SNR = 50dB
TTI = 1 ms
# of pilots = 4
# of PilotFrames = 100
3002
3004
SNR [dB]
# of PRBs
54
52
50
48
46
44
42
40
38
36
34
1
1.5
2
2.5
3
3.5
4

FIG. 30

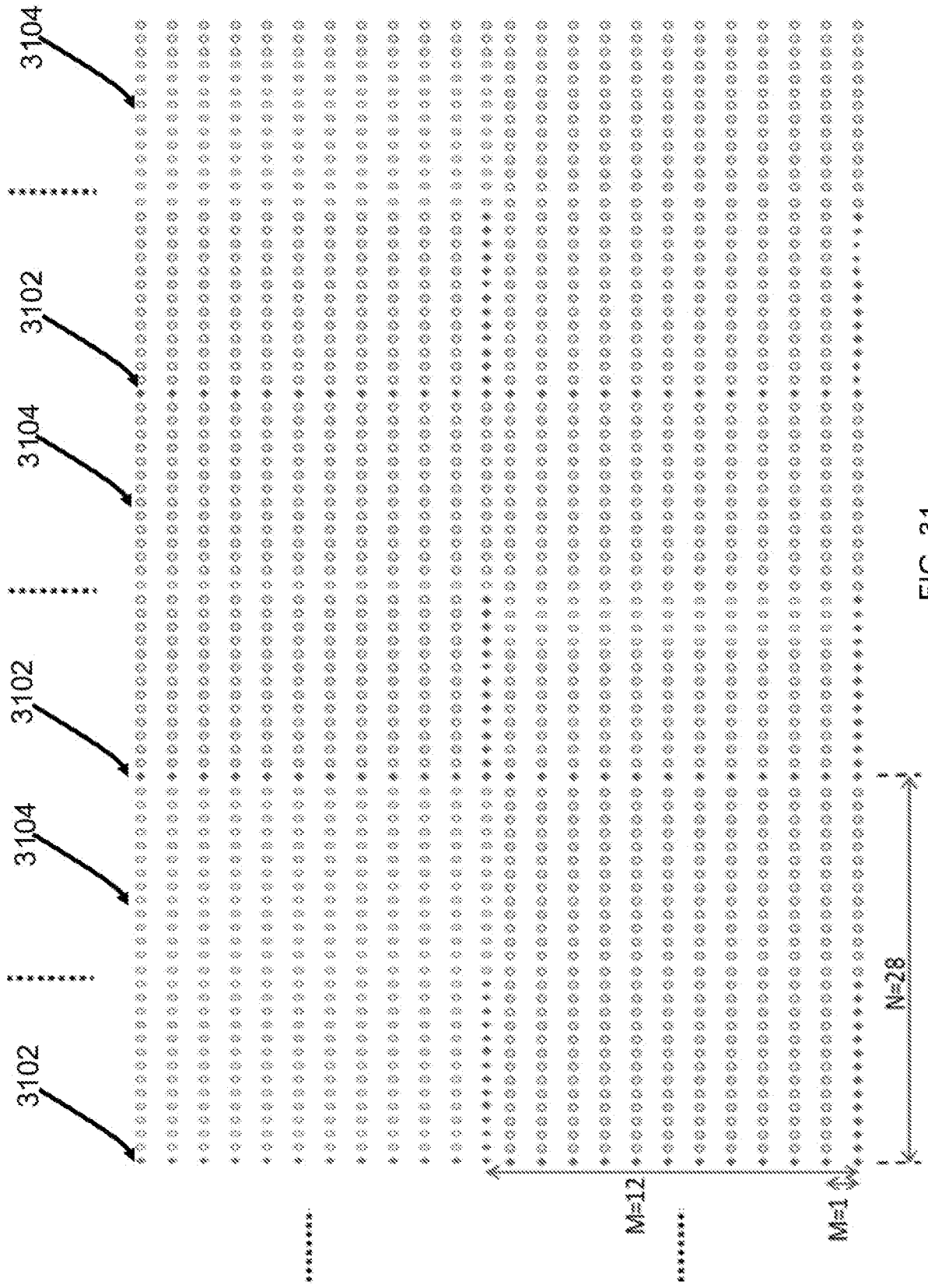


FIG. 31

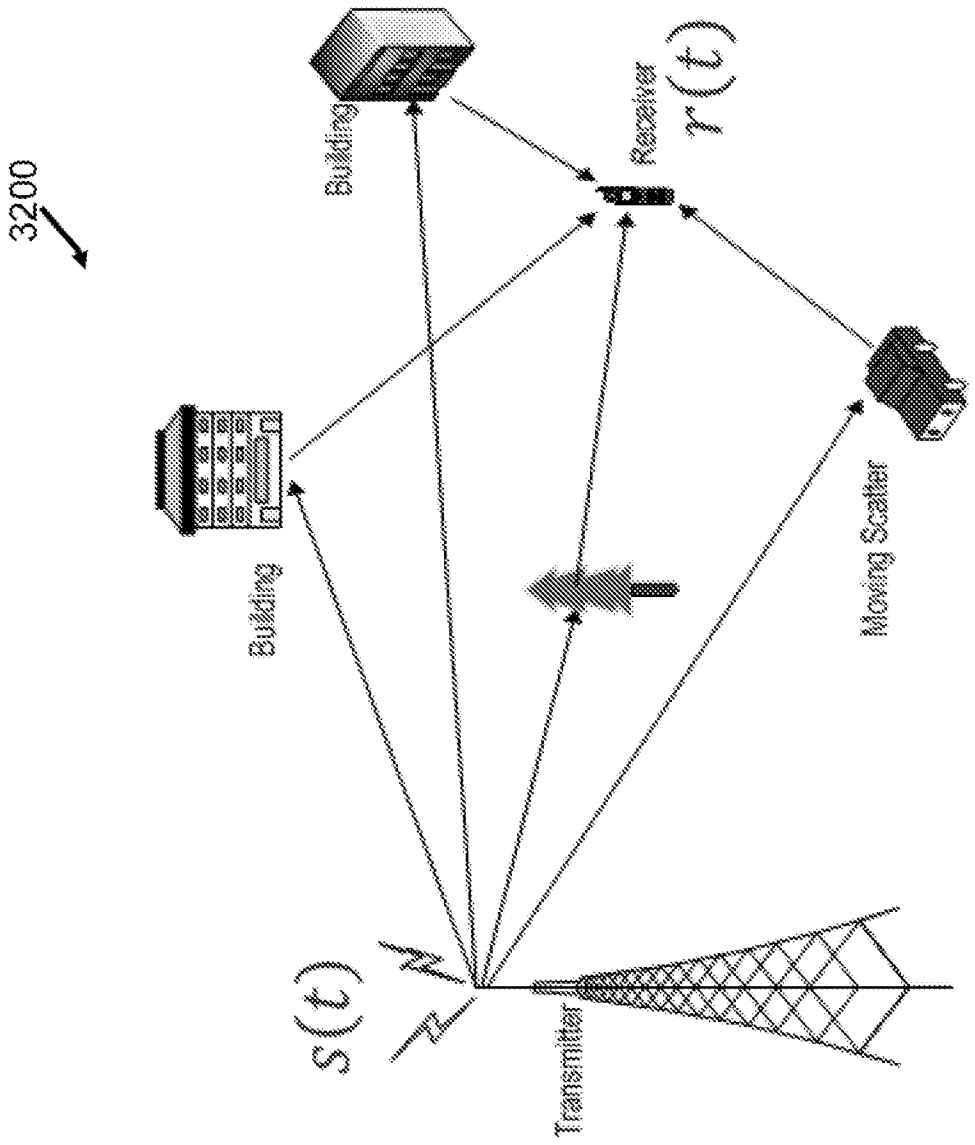

3200
Building
Receiver
$r(t)$
Building
Moving Scatter
$s(t)$
Transmitter

FIG. 32

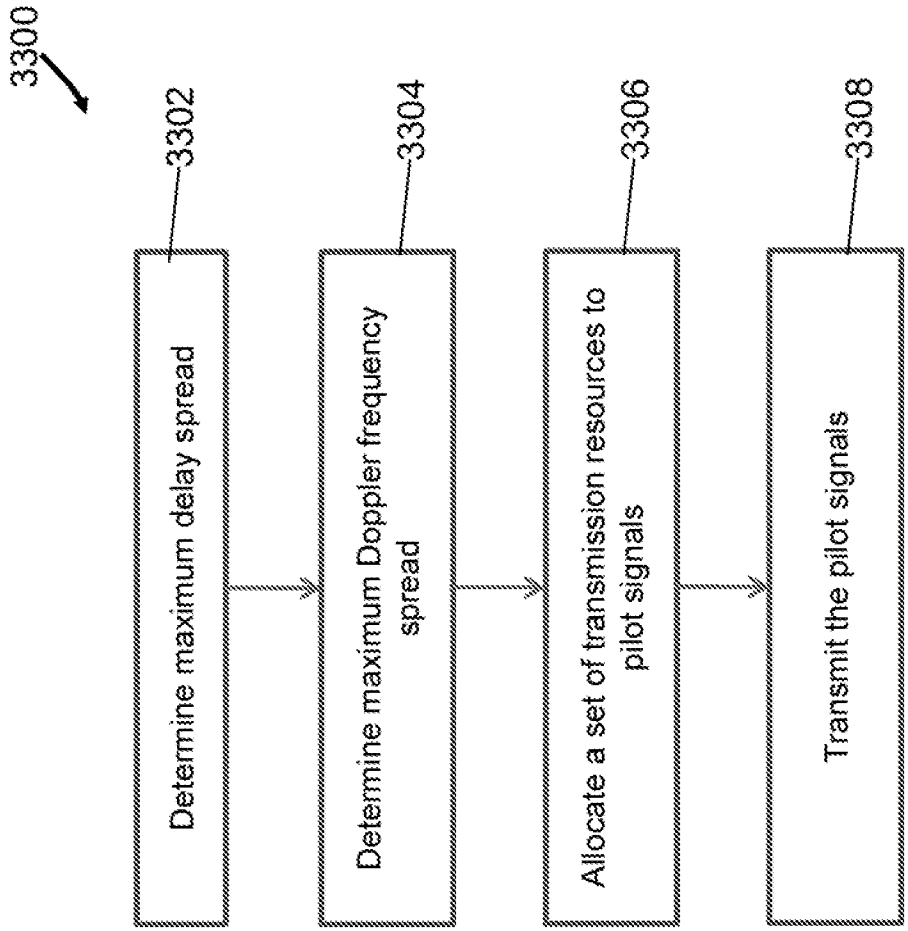

3300
3302
Determine maximum delay spread
3304
Determine maximum Doppler frequency spread
3306
Allocate a set of transmission resources to pilot signals
3308
Transmit the pilot signals

FIG. 33

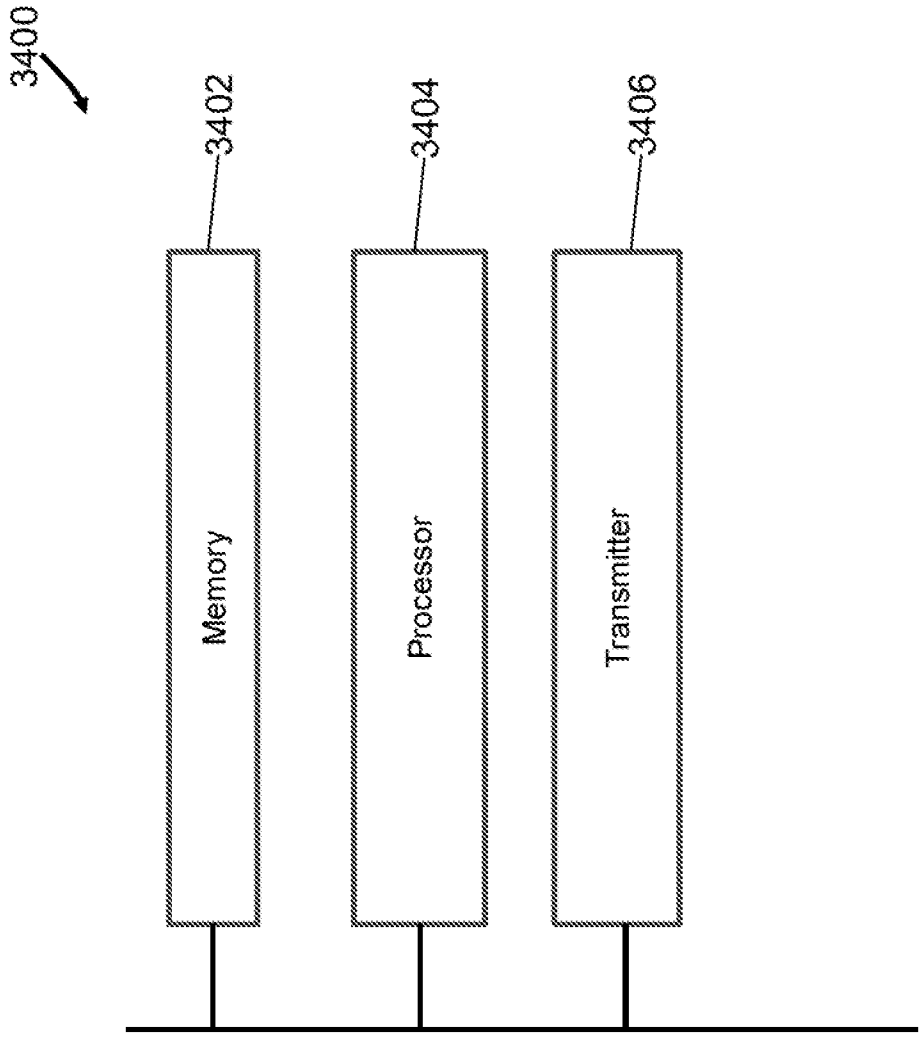


FIG. 34

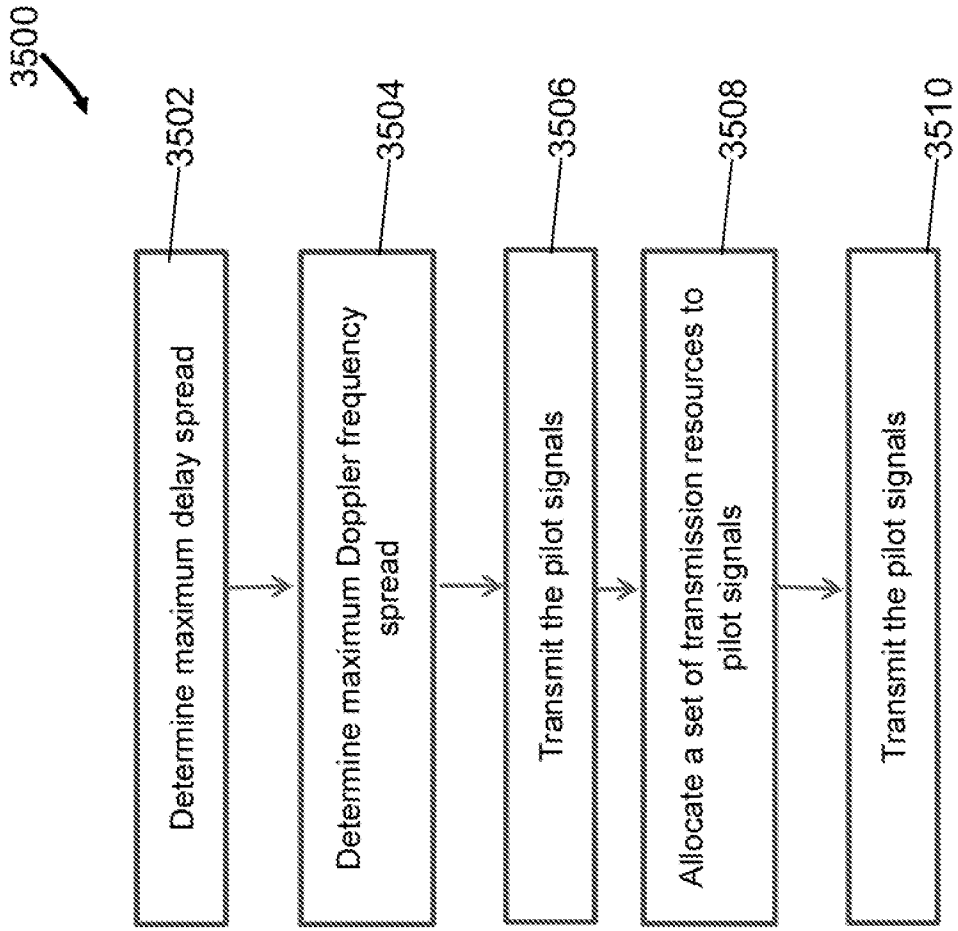

3500
3502
Determine maximum delay spread
3504
Determine maximum Doppler frequency spread
3506
Transmit the pilot signals
3508
Allocate a set of transmission resources to pilot signals
3510
Transmit the pilot signals

FIG. 35

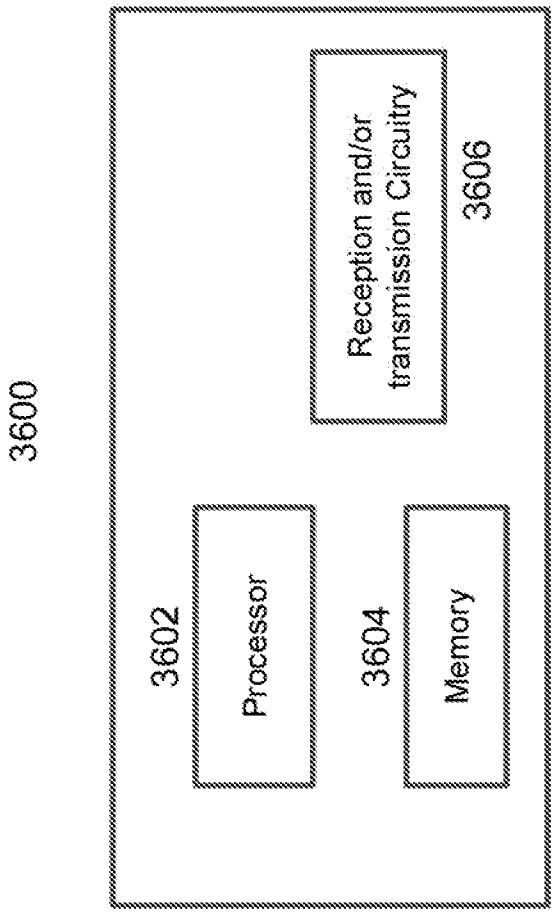

3600
3602
Processor
3604
Memory
Reception and/or transmission Circuitry
3606

FIG. 36

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/019376

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - H04B 1/10; H04B 1/46; H04B 1/69; H04B 1/707; H04L 12/28; H04L 12/56; H04W 24/08 (2017.01)
CPC - H04B 1/713; H04B 1/71632; H04B 2001/6912; H04L 1/0007; H04L 5/0007; H04L 5/0048; H04L 27/10; H04L 27/103; H04L 27/261; H04W 24/08; H04W 28/06 (2017.02)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 370/252; 370/394; 375/139; 375/149; 375/350; 455/81 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 2004/0131110 A1 (ALARD et al.) 08 July 2004 (08.07.2004) entire document	1, 4, 8, 10, 21, 23, 26-28 2, 3, 5-7, 9, 11-20, 22, 24, 25
Y	US 2006/0133381 A1 (WANG) 22 June 2006 (22.06.2006) entire document	2, 3, 5-7, 12, 14-19, 24, 25
Y	US 2006/0276143 A1 (ANGLIN, JR.) 07 December 2006 (07.12.2006) entire document	9, 20
Y	US 2012/0269232 A1 (HISCOCK) 25 October 2012 (25.10.2012) entire document	11, 13, 15
Y	US 2014/0269357 A1 (RESEARCH IN MOTION LIMITED) 18 September 2014 (18.09.2014) entire document	22
A	US 2008/0137788 A1 (BANG et al.) 12 June 2008 (12.06.2008) entire document	1-28

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 April 2017	05 MAY 2017
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774