



- (51) International Patent Classification:
H04L 25/03 (2006.01) H04B 7/155 (2006.01)
- (21) International Application Number:
PCT/EP2016/057039
- (22) International Filing Date:
31 March 2016 (31.03.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (71) Applicant: HUAWEI TECHNOLOGIES CO., LTD.
[CN/CN]; Huawei Administration Building Bantian, Long-gang District, Shenzhen, Guangdong 518129 (CN).
- (72) Inventors; and
(71) Applicants (for US only): ESTELLA AGUERRI, Inaki [ES/DE]; c/o Huawei Technologies Duesseldorf GmbH, Riesstr.25, 80992 Munich (DE). BENAMMAR, Meryem [FR/DE]; c/o Huawei Technologies Duesseldorf GmbH, Riesstr.25, 80992 Munich (DE). ZAIDI, Abdellatif [FR/DE]; c/o Huawei Technologies Duesseldorf GmbH, Riesstr.25, 80992 Munich (DE).
- (74) Agent: KREUZ, Georg; Huawei Technologies Duesseldorf GmbH, Riesstr. 8, 80992 Munich (DE).

- (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, 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, KN, KP, KR, 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, 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).

Published:
— with international search report (Art. 21(3))

(54) Title: RADIO SYSTEM AND DECODING DEVICE FOR DISTRIBUTED COMPRESSION

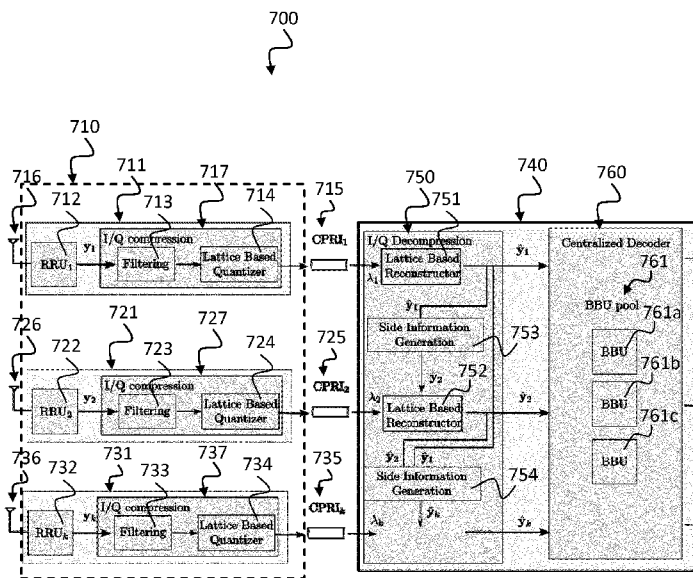


Fig. 7

(57) Abstract: The disclosure relates to a radio system (710) comprising a plurality of radio units (711, 721, 731), each radio unit, comprising: a receive interface (712, 722, 732), configured to receive at least one radio signal, y_k , over a multiple-input multiple-output, MIMO, radio channel; compression parameters from a decoding device (740); a compressor (717, 727, 737) configured to generate a codeword signal, λ_k , based on the compression parameters by encoding the at least one radio signal y_k with a lattice based distributed code; and a transmit interface (715, 725, 735), configured to transmit the codeword signal λ_k to the decoding device (740). The disclosure further relates to a decoding device (740), comprising: a receive interface, configured to receive from a plurality of radio units (711, 721, 731) a plurality of signals, λ_k , over a plurality of links, each signal λ_k corresponding to a compressed radio signal y_k , wherein the compressed radio signal y_k carries a plurality of messages, ω_i , from a plurality of users; a decompressor (750), configured to decompress the plurality of signals λ_k with a nested lattice based distributed code to provide a plurality of decompressed radio signals, \hat{y}_k a decoder (760), configured to decode the plurality of decompressed radio signals, \hat{y}_k , to recover the plurality of messages;

a controller, configured to determine compression parameters based on the decoder structure and based on a correlation of the received signals; and a transmit interface configured to transmit the compression parameters to each of the plurality of radio units (711, 721, 731).

Radio system and decoding device for distributed compression

TECHNICAL FIELD

5 The present disclosure relates to a radio system and a decoding device using lattice based distributed code. In particular, the present disclosure relates to methods for distributed compression using nested lattices techniques.

BACKGROUND

10

Interference is one of the most limiting factors of communication in networks. In certain cases, its effect can be mitigated through the use of distributed network architectures, formed by distributed radio units (RU) – or remote radio units (RRUs) – and a centralized processor (CP). Interference can be mitigated by enabling the centralized processing of the signals received by multiple RRUs at the CP. This topology has some other appreciable features, such as low cost deployment of BSs and flexible network utilization. In this architecture, the RRUs implement only the radio functionalities, i.e., transmission, reception, analog-to-digital conversion (ADC) and digital-to-analog conversion (DAC), while the CP centralizes the based band processing units (BBUs). Alternative architectures in Next Generation Fronthaul Interfaces (NGFI), consider RRUs which also share some base band functionalities with the BBU at the CP.

15
20

The cloud radio access network (C-RAN) architecture as shown for example in Fig. 1, is an example of such distributed network topology proposed for 4.5G and 5G, in which base stations (BSs), containing the RRUs, are connected to a cloud-computing central processor (CP) via finite capacity link containing the BBUs. Fig. 3 shows a schematic diagram illustrating a Cloud Radio Access Setup 300. In the uplink of such a scenario, there is a plurality of received signals received at a plurality of antennas 313, 323, whether they are in the same BS (not depicted in Fig. 3) or in different BSs 310, 320 (as depicted in Fig. 3). The signals are forwarded to the CP 301, which has a BBU pool 302. At each antenna the received signal is processed by a RRU unit 312, 322, and the IQ samples are forwarded to the BBU 303, 304.

25
30

Distributed architecture networks can also be considered to reduce the hardware requirement, e.g., in Massive MIMO stations. Fig. 2 shows a schematic diagram

35

illustrating a Massive MIMO station 200. There are a plurality of received signals received at a plurality of antennas 224, 225, 234, 235, which may be in the same device 210 (as depicted in Fig. 2) or in different devices (not depicted in Fig. 2). The signal received at each antenna is processed at the RRU 222, 223, 232, 233 and forwarded over the CPRI links 204, 205 to a remote base band unit (BBU) pool 201 having a plurality of BBUs 202, 203

Digital base band signal transmission between RRUs and BBUS are usually carried with a Common Public Radio Interface (CPRI) link 305, 306. In order to reduce the amount of data transmitted, between RRUs 312, 322 and BBUs 303, 304, compression is applied. In the baseline solution 400, as depicted in Fig. 4, compression and decompression are performed independently for each antenna. That is, if the signals from the RRUs have to be forwarded, they are independently compressed and decompressed, independently of the other signals at the other RRUs. In the downlink, the process would be the reverse.

Fig. 4 shows a schematic diagram illustrating a multi-antenna network system according to a baseline solution 400. A first path includes first transceiver 410 with BBU 411, compression 412, decompression 413, CPRI link 415, second transceiver 430 with RRU 433, compression 431 and decompression 432. A second path, independent from first path includes third transceiver 420 with BBU 421, compression 422, decompression 423, CPRI link 425, fourth transceiver 440 with RRU 443, compression 441 and decompression 442. In the baseline solution independent compression 412, 431, 422, 441 and decompression 413, 432, 423, 442 is performed for each antenna 434, 444. In Fig. 4 the uplink and downlink directions are provided. Note that the two decompression and compression operations are done independently with respect to each other.

The signals received at the plurality of antennas are highly correlated, and consequently, so are the digitized IQ samples at the RRUs after the ADC. Distributed compression can be utilized to take into account the statistical correlation between the IQ samples at the plurality of RRUs. For example, Wyner-Ziv type codes can be utilized, that consider the signals at the other RRUs as side information. However, distributed codes are mostly theoretical and lack efficient constructions.

On the other hand, in general, both baseline and distributed compression solutions independently design the compressor block and the decoder block. That is, the compression is designed so as to minimize the distortion between the received signals at

the RUs and their decompressed versions, without taking into account the specific structure of the decoder at the CP. However, there are cases in which compression should be designed to maximize the information transmission. This falls within the information bottleneck class of problems.

5

SUMMARY

It is the object of the invention to provide a design for compression techniques in radio networks which minimizes distortion and maximizes information transmission.

10

This object is achieved by the features of the independent claims. Further implementation forms are apparent from the dependent claims, the description and the figures.

A basic idea of the invention is to apply a novel concept for optimizing a distributed compression block for a given decoder at the CP, and implementing the solution using nested lattices. A practically feasible distributed compression scheme is provided that is based on nested lattices and accounts for the specific structure of the CP decoder. The presented idea provides a significant modification of the transmitted signal between RRU and CP with respect to existing modulation formats (e.g., those used in LTE), that is, in the transmission over the CPRI link.

20

Since the presented idea is related to the design of the radio units, e.g., Base station in an uplink cellular system, data provided to antennas in a Massive MIMO setup, it can be included and explicitly described in communication standards. It will be shown that, by optimizing the compression taking into account the specific structure of the decoder, one can obtain improvements over solutions in which the compressor and decoder blocks are designed independently.

25

The compression and decompression techniques described herein may be implemented in wireless communication networks, in particular communication networks based on mobile communication standards such as LTE, in particular LTE-A and/or OFDM. The transmission and reception devices described herein may further be implemented in a base station (NodeB, eNodeB) or a mobile device (or mobile station or User Equipment (UE)). The described devices may include integrated circuits and/or passives and may be manufactured according to various technologies. For example, the circuits may be

30

designed as logic integrated circuits, analog integrated circuits, mixed signal integrated circuits, optical circuits, memory circuits and/or integrated passives.

5 The compression and decompression techniques described herein may be configured to transmit and/or receive radio signals. Radio signals may be or may include radio frequency signals radiated by a radio transmitting device (or radio transmitter or sender) with a radio frequency lying in a range of about 3 Hz to 300 GHz. The frequency range may correspond to frequencies of alternating current electrical signals used to produce and detect radio waves.

10

The compression and decompression techniques described herein may be designed in accordance to mobile communication standards such as e.g. the Long Term Evolution (LTE) standard or the advanced version LTE-A thereof. LTE (Long Term Evolution), marketed as 4G, 5G LTE and beyond, is a standard for wireless communication of high-
15 speed data for mobile phones and data terminals.

In order to describe the invention in detail, the following terms, abbreviations and notations will be used:

20	BS:	Base Station, eNodeB
	UE:	User Equipment, e.g. a mobile device or a machine type communication Device
	RU:	Radio Unit
	CP:	Central Processor
25	C-RAN:	Cloud Radio Access Network
	PtP:	Point-to-Point
	LSWZ:	Lattice Successive Wyner Ziv
	CPRI:	Common Public Radio Interface
	SNR:	Signal-to-Noise Ratio
30	MIMO:	Multiple-Input Multiple-Output
	MMSE:	Minimum Mean Square Error
	5G:	5 th generation according to 3GPP standardization
	LTE:	Long Term Evolution
	UL:	Uplink
35	DL:	Downlink

C: front-haul capacity
EVM: Error Vector Magnitude

5 According to a first aspect, the invention relates to a radio system comprising a plurality of radio units, each radio unit, comprising: a receive interface, configured to receive at least one radio signal, y_k , over a multiple-input multiple-output, MIMO, radio channel; compression parameters from a decoding device; a compressor configured to generate a codeword signal, λ_k , based on the compression parameters by encoding the at least one radio signal y_k with a lattice based distributed code; and a transmit interface, configured
10 to transmit the codeword signal λ_k to the decoding device.

The radio units of the radio system can be at different locations and not necessarily in the same physical device. For example in the C-RAN scenario, there are various base stations (BSs) at different locations all connected to the same CP. However, in one
15 implementation, the radio units of the radio system can be arranged in the same physical device.

When using compression parameters from a decoding device for encoding the radio signal with a lattice based distributed code, the decompression and compression
20 operations can be performed with respect to each other, thereby producing a codeword that minimizes distortion and maximizes information transmission.

In a first possible implementation form of the radio system according to the first aspect, the distributed compression code and the compression parameters are known at the radio
25 system. The distributed compression code can be formed by the lattice based distributed code.

This provides the advantage of generating distributedly compression codewords that allow, when decompressed with the appropriate decompressor, to account for the
30 correlation of the compressed signals at different RRUs and to maximize the information transfer between the compressed signal and the reconstruction at the decoding device, by jointly optimizing compression, decompression and decoding parameters.

In a second possible implementation form of the radio system according to the first aspect, the compression parameters comprise parameters of a plurality of nested lattices forming the lattice based distributed code. For example, the compression parameters comprise second order moments of the plurality of nested lattices.

5

This provides the advantage that the nested lattices can be efficiently described, minimizing the amount of information that has to be exchanged between the controller and the compressor to configure the compression method at the radio system.

10 In a third possible implementation form of the radio system according to the first or the second implementation form of the first aspect, the compressor comprises: a dither module configured to add a dither to the radio signal y_k ; a quantization module, configured to quantize the dithered radio signal based on the plurality of nested lattices; and a modulo reduction module, configured to modulo reduce the quantized dithered radio
15 signal to generate the codeword signal, λ_k .

This provides the advantage that the compressor (and also the decompressor) involve simple filtering, whitening and modulo reduction operations and the lattice quantization and modulo reduction operations can be implemented with efficient low complexity
20 operations.

In a fourth possible implementation form of the radio system according to the third implementation form of the first aspect, the compressor is configured to compress the radio signal y_k based on a lattice operation comprising a quantization and a modulo
25 reduction.

This provides the advantage that the lattice operation can be efficiently implemented with low complexity and fast operations and that the resulting lattice codewords can be decompressed at the decompressor in a way that allows to exploit the correlation with the
30 signals compressed at the different RRUs.

According to a second aspect, the invention relates to a decoding device, comprising: a receive interface, configured to receive from a plurality of radio devices a plurality of signals, λ_k , over a plurality of links, each signal λ_k corresponding to a compressed radio

signal y_k , wherein the compressed radio signal y_k carries a plurality of messages, ω_l , from a plurality of users; a decompressor, configured to decompress the plurality of signals λ_k with a nested lattice based distributed code to provide a plurality of decompressed radio signals, \hat{y}_k ; a decoder, configured to recover the plurality of

5 messages, ω_l , from the plurality of decompressed radio signals \hat{y}_k ; a controller, configured to determine compression parameters based on the decoder structure and based on a correlation of the received signals; and a transmit interface configured to transmit the compression parameters to each of the plurality of radio devices.

10 When determining compression parameters based on the decoder structure and based on a correlation of the received signals, the decompression and compression operations can be performed with respect to each other, thereby producing a codeword that minimizes distortion and maximizes information transmission.

15 In a first possible implementation form of the decoding device according to the second aspect, the decompressor comprises a plurality of lattice based reconstruction modules, configured to successively decompress the plurality of signals λ_k .

This provides the advantage that the plurality of signals can be decompressed one by one, thereby saving computational complexity.

20

In a second possible implementation form of the decoding device according to the first implementation form of the second aspect, a decompression result of a reconstruction module of the plurality of lattice based reconstruction modules depends on a

25 decompression result of a preceding reconstruction module of the plurality of lattice based reconstruction modules.

This provides the advantage that an iterative decompression scheme can be applied which has a low computational complexity and can be easily implemented.

30

In a third possible implementation form of the decoding device according to any of the first or the second implementation forms of the second aspect, a decompression result of a reconstruction module of the plurality of lattice based reconstruction modules depends on

side information generated from decompression results of preceding reconstruction modules of the plurality of lattice based reconstruction modules.

5 This provides the advantage that the information transfer between the transmitted signal and the reconstruction at the CP can be maximized by exploiting the correlation between the received signals to generate effective side information and improve decompression.

10 In a fourth possible implementation form of the decoding device according to the third implementation form of the second aspect, the decompressor comprises an estimator, in particular a linear filtering estimator, configured to generate the side information based on estimation using the decompression results of the preceding reconstruction modules of the plurality of lattice based reconstruction modules. The linear filtering estimator can be for example an optimum filter estimator, e.g. a linear minimum mean squares estimator, LMMSE. However, the LMMSE is just one example on how to design the filter to generate
15 the side information. The side information may be generated by any arbitrary function that generates the side information sequence using the plurality of previous decompression results.

20 This provides the advantage of high accuracy when applying the linear filtering estimator.

In a fifth possible implementation form of the decoding device according to any of the third or the fourth implementation forms of the second aspect, each reconstruction module comprises a combiner, configured to combine a respective signal λ_k with the corresponding side information to generate the respective decompressed radio signal, \hat{y}_k
25

This provides the advantage that by combining the respective signal with the corresponding side information increases accuracy of decompression.

30 In a sixth possible implementation form of the decoding device according to the fifth implementation form of the second aspect, the k-th combiner is based on the following lattice operation: $\hat{y}_k = [\lambda_k - \tilde{y}_k - \mathbf{u}_k] \bmod \Lambda_k + \tilde{y}_k$, where λ_k is the k-th signal, \tilde{y}_k is the k-th side information, Λ_k is the k-th lattice, \mathbf{u}_k is a k-th dither and \hat{y}_k is the k-th decompressed radio signal. Note, that this combining operation to decompress needs that

the compression operation is $\lambda_k = [y_k + u_k] \bmod \Lambda_k$. There may be alternative methods based on lattice codes to do this step, and this is a particular embodiment.

5 This provides the advantage that the lattice operation can be efficiently implemented by standard hardware or software modules such as addition, multiplication and modulo operation.

10 In a seventh possible implementation form of the decoding device according to the second aspects as such or according to any of the preceding implementation forms of the second aspect, the controller is configured to determine the compression parameters based on a decoding metric depending on the decoding operation of the decoder, the decompression operation of the decompressor and a correlation of the received signals.

15 This provides the advantage that the compression parameters can be optimized in two directions, i.e. for the decoding operation of the decoder and the decompression operation of the decompressor.

20 In an eighth possible implementation form of the decoding device according to the seventh implementation form of the second aspect, the decoding metric depends on at least one of the following parameters: a number of the plurality of users, channel characteristics of the plurality of links, signal to noise ratios of the plurality of links, a quantization and/or statistics of a quantization error due to the nested lattice distributed code.

25 This provides the advantage that by using dependencies on such a variety of parameters, the coding and compression performance can be improved.

30 In a ninth possible implementation form of the decoding device according to any of the seventh or the eighth implementation forms of the second aspect, the controller is configured to maximize the decoding metric to determine optimal compression parameters.

This provides the advantage that by using optimal compression parameters, data throughput can be increased for a fixed capacity of the links between the plurality of RRU

and decoding device. Likewise, the required capacity of the links between RRUs and CP can also be reduced and a certain data throughput can be guaranteed.

5 According to a third aspect, the invention relates to a method to use lattice-based compression and decompression based on nested lattices for the information bottleneck problem.

This method provides the advantage of a modular compression block to forward the received signals with maximum information transfer.

10

According to a fourth aspect, the invention relates to a method to jointly design the compression parameters and decoding parameters for the aim of maximizing a given performance metric.

15 This method provides the advantage that users can transmit with any code and any centralized decoder can be applied on the output signal of the block.

According to a fifth aspect, the invention relates to a method for successive decoding of the lattice compression codewords and effective side information generation.

20

This method provides the advantage that the compression is adjusted to the respective decoder. Therefore its implementation requires only partial update of the system.

25 According to a sixth aspect, the invention relates to a method to account for the correlation of the signals received at other antennas when using successive decompression using effective side information sequences.

This method provides the advantage that the method operates on each received signal independently without requiring the received signals to be available at the same position.
30 Hence, the method significantly improves upon standard point to point compression.

BRIEF DESCRIPTION OF THE DRAWINGS

Further embodiments of the invention will be described with respect to the following figures, in which:

Fig. 1 shows a block diagram illustrating a system model of a Cloud Radio Access
5 Network (C-RAN) 100;

Fig. 2 shows a schematic diagram illustrating a Massive MIMO station 200;

Fig. 3 shows a schematic diagram illustrating a Cloud Radio Access Setup 300;
10

Fig. 4 shows a schematic diagram illustrating a multi-terminal network system according to a baseline solution 400;

Fig. 5 shows a schematic diagram of a method 500 using decoder aware compression design according to an implementation form;
15

Fig. 6 shows a schematic diagram of a multi-terminal network system 600 according to an implementation form;

Fig. 7 shows a schematic diagram of a multi-terminal network system using lattice based compression and decompression 700 according to an implementation form;
20

Fig. 8 shows a block diagram of a compressor 800 according to an implementation form;

Fig. 9 shows a block diagram of a decompressor 900 according to an implementation form;
25

Fig. 10 shows a schematic diagram of a C-RAN network 1000 according to an implementation form;
30

Fig. 11 shows a performance diagram illustrating average throughput 1100 versus SNR in a C-RAN network for compression with different lattice codes according to an implementation form;

Fig. 12 shows a performance diagram illustrating average throughput 1100 versus front-haul capacity in a C-RAN network for compression with different lattice codes according to an implementation form;

5 Fig. 13 shows a performance diagram illustrating average EVM 1300 versus front-haul capacity in a C-RAN network for compression with different lattice codes according to an implementation form;

10 Fig. 14 shows a performance diagram illustrating average EVM 1300 versus front-haul capacity gain in a C-RAN network for compression with different lattice codes according to an implementation form;

Fig. 15 shows a schematic diagram of a Massive MIMO station with dedicated CPRI links 1500 according to an implementation form;

15

Fig. 16 shows a performance diagram illustrating average throughput 1600 versus SNR in a Massive MIMO network for compression with different lattice codes according to an implementation form;

20 Fig. 17 shows a performance diagram illustrating average throughput 1700 versus front-haul capacity in a Massive MIMO network for compression with different lattice codes according to an implementation form;

25 Fig. 18 shows a performance diagram illustrating average EVM 1800 versus front-haul capacity in a Massive MIMO network for compression with different lattice codes according to an implementation form;

30 Fig. 19 shows a performance diagram illustrating average EVM 1900 versus front-haul capacity gain in a Massive MIMO network for compression with different lattice codes according to an implementation form;

Fig. 20 shows a performance diagram illustrating average EVM 2000 versus front-haul capacity in a Massive MIMO network with shared CPRI for compression with different lattice codes according to an implementation form; and

35

Fig. 21 shows a performance diagram illustrating average EVM 2000 versus SNR per user in a Massive MIMO network with different lattice codes according to an implementation form.

5

DETAILED DESCRIPTION OF EMBODIMENTS

In the following detailed description, reference is made to the accompanying drawings, which form a part thereof, and in which is shown by way of illustration specific aspects in which the disclosure may be practiced. It is understood that other aspects may be utilized and structural or logical changes may be made without departing from the scope of the present disclosure. The following detailed description, therefore, is not to be taken in a limiting sense, and the scope of the present disclosure is defined by the appended claims.

10 It is understood that comments made in connection with a described method may also hold true for a corresponding device or system configured to perform the method and vice versa. For example, if a specific method step is described, a corresponding device may include a unit to perform the described method step, even if such unit is not explicitly described or illustrated in the figures. Further, it is understood that the features of the various exemplary aspects described herein may be combined with each other, unless specifically noted otherwise.

Fig. 5 shows a block diagram of a method 500 using decoder aware compression design according to an implementation form. The method 500 to optimize the parameters is illustrated in Figure 5. It uses the inputs 501: decompression method 502, compression method 503, system parameters 504, decoding method 505, performance metric 506 and provides the outputs 520: optimal compression parameters 521 and optimal decoding parameters 522. The method 500 comprises the following steps: Characterize the Effective Channel 511: this step characterizes the distribution of the output after decompression in terms of the compression parameters, for given lattice compression and decompression methods and system parameters (including SNR, noise distribution, channel coefficients, number of UE, number of RU); Compression Parameter Optimization 512: A given a performance metric depending on the effective channel, decoding method and system parameters, e.g., mutual information, bit-error rate, the

25
30

compression and decoding parameters are chosen to optimize this metric. In the embodiments below particular examples for this method are shown.

5 Fig. 6 shows a schematic diagram of a multi-terminal network system 600 according to an implementation form.

The multi-terminal network system 600 illustrates an end-to-end design of a lattice based distributed codes for the information bottleneck problem of maximizing the end-to-end information transfer, which takes into account the centralized decoder applied at the CP and the correlation of the signals at different RU in its design. The main aspects are: Use
10 of nested lattices for the multi-terminal information bottleneck problem, that is, the method to encode and decode for a given nested lattice code for distributed compression for the purpose of end-to-end maximization of information transfer; and a method for the joint design of the compression block and decoder structure. That implies, for example, the
15 lattice design, effective side information generation, filtering and equalization.

The end-to-end multi-terminal network system 600 includes the following blocks: B1) Lattice Based Compression 610: Use nested lattices to compress the received signal at each RU accounting for statistical correlation with the signals at other RUs. A method is
20 provided to compress at each RU for a given nested lattice code. B2) Lattice Based Decompression 622: Use lattice decoding at the CP to decompress the signals. We provide a method to decompress the signals successively. At each step, previous signals are used to generate an effective side information sequence, which is used in decompression. B3) Centralized Decoding 626: Use a given multi-user decoder (e.g.,
25 linear filtering followed by MMSE decoding based on nearest-decoding or Zero-Forcing) to recover the messages from the decompressed signals. The decompressed signals form an effective channel from which the messages are recovered. We provide a method to account for the signal degradation caused by the compression. B4) Joint Parameter Design 628: Compression and decoding processing are jointly optimized based on a
30 metric that depends both on the decoder and the compression.

In the following, a first embodiment of implementation of the proposed blocks and optimization method detailed above is described for a setup consisting of L single antenna UE 601, 603, 605 and a device 610 including K radio units describing a nested lattice

compression with centralized decoding. It is assumed that the L single antenna users transmit with Gaussian codewords $x \sim N(0, P_{I_n})$. Similarly, also finite constellation codewords as M-QAM, or lattice codebooks can be used. These embodiments are denoted as LSWZ.

5

The Lattice Based Compression 610 according to B1) can be implemented as described in the following: The lattice construction and compression method employed at each RU may be based on a lattice scheme by using a Nested Codebook construction and Compression 613, 614, 615 at k -th RU. A lattice Λ of dimension n is a discrete additive subgroup of \mathbb{R}^n . It can be written in terms of a lattice Generator matrix G as $\Lambda = \{\lambda = Gi: i \in \mathbb{Z}^n\}$. The fundamental Voronoi region $V(\Lambda)$ is the set of point in \mathbb{R}^n that are closer to the zero vector than any other lattice point. The reduction modulo Λ is defined as $[x] \bmod \Lambda = x - Q_\Lambda(x)$, where $Q_\Lambda(x): \mathbb{R}^n \rightarrow \Lambda$ is a lattice quantizer. A lattice Λ_1 is nested inside another lattice Λ if $\Lambda_1 \subset \Lambda$. At each BS, let Λ_k and $\Lambda_{q,k}$ be a pair of n dimensional nested lattices satisfying $\Lambda_k \subseteq \Lambda_{q,k}$ with their respective generator matrices G_k and $G_{q,k}$. The pair of lattices form a codebook $\mathcal{L}_k = V(\Lambda_k) \cap \Lambda_{q,k}$ of rate C_k and of length n codewords $\lambda_k \in \mathcal{L}_k$. The second moment of each lattice can be described as $\sigma^2(\Lambda_k)$ and $\sigma^2(\Lambda_{q,k})$. The k -th RU may compress the received signal $y_k = h_k X + z_k$ at rate C_k by adding a dither u_k uniformly distributed over the Voronoi region of the finite lattice $\Lambda_{q,k}$, and applying the following lattice operation:

20

$$\lambda_k = \left[Q_{\Lambda_{q,k}}(y_k + u_k) \right] \bmod \Lambda_k.$$

The index corresponding to the codeword $\lambda_k \in \mathcal{L}_k$ may be forwarded to the CP 620.

A codebook is a set of codewords. The number of codewords in a codebook \mathcal{C} formed by a pair of nested lattices $\Lambda \subset \Lambda_f$ of dimension n is:

25

$$R = \frac{1}{n} \log |\mathcal{C}| = \frac{1}{n} \log \frac{\text{Vol}(\Lambda)}{\text{Vol}(\Lambda_f)}.$$

In this case, the rate of the codebook is C_k , that is, it has 2^{nC_k} . In this case it coincides with the capacity of the error free link of capacity C_k since the compression codewords are transmitted over link k .

30

The Lattice Based Compression 610 according to B1) can be implemented in one physical device including the multiple compressors 613, 614, 615, e.g. according to the Massive

MIMO station as described above with respect to Fig. 2. Each of the radio units may include one or more antennas 611. Alternatively, the Lattice Based Compression 610 according to B1) can be implemented as one system of physical devices coupled with respect to each other, where each physical device includes one or more of the
 5 compressors 613, 614, 615 depicted in Fig. 6, e.g. according to the Cloud Radio Access Network as described above with respect to Fig. 3. Each of the radio units may include one or more antennas 611.

The Lattice Based Decompression 622 according to B2) can be implemented as
 10 described in the following: After receiving the K indices and remapping to the compression codewords $(\lambda_1, \dots, \lambda_K)$, the CP 620 successively reconstructs y_k as \hat{y}_k starting from the first RU. Note that the order in which the RU are decoded can be optimized, however, here for simplicity we assume that each RU is labelled according to the order in which they are decompressed. Effective side information generation 623 can be implemented as
 15 follows: First, given $k - 1$ reconstructed signals have already been decompressed as $\hat{y}_1, \dots, \hat{y}_{k-1}$, the CP 620 computes an estimate (using the linear MMSE) of each sample $y_k(t)$ to generate an effective side information sequence as:

$$\tilde{y}_k(t) = \text{LMMSE}[y_k(t)|\hat{y}_1(t), \dots, \hat{y}_{k-1}(t)],$$

which can be calculated with the linear filtering of the decompressed sequences

$$20 \quad \tilde{y}_k(t) = \mathbf{h}_k (\mathbf{H}_{k-1}^T \Sigma_{k-1}^{-1} \mathbf{H}_{k-1} + \text{SNR}^{-1} \mathbf{I})^{-1} [\hat{y}_1(t), \dots, \hat{y}_{k-1}(t)],$$

where

$$\mathbf{H}_{k-1} = [\mathbf{h}_1^T, \dots, \mathbf{h}_{k-1}^T]^T$$

and

$$\Sigma_{k-1} = \text{diag}(\sigma^2(\Lambda_{q,1}), \dots, \sigma^2(\Lambda_{q,k-1})).$$

Decompression 622 of k -th compressed signal at CP 620 can be implemented as follows:

25 Using the effective side information sequence, the CP 620 reconstructs y_k using the effective side information with the following lattice operation:

$$\hat{y}_k = [\lambda_k - \mathbf{u}_k - \tilde{y}_k] \bmod \Lambda_k + \tilde{y}_k.$$

Once the K compression codewords $\hat{y}_1, \dots, \hat{y}_K$ have been reconstructed, the effective channel of the compression can be characterized with the distribution of $\hat{\mathbf{Y}} = \mathbf{H}\mathbf{X} + \mathbf{Z} + \mathbf{Q}$,

where Q is due to the compression, has covariance matrix Σ and each of its component is uniformly distributed over the Voronoi region of $\Lambda_{q,k}$ and

$$\Sigma = \begin{bmatrix} \sigma^2(\Lambda_{q,1}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2(\Lambda_{q,K}) \end{bmatrix}.$$

5 The Centralized Decoding 626 according to B3) can be implemented as described in the following: In this block, the CP 620 recovers the transmitted messages from the decompressed signal accounting for the compression noise. The CP 620 may have a decoding structure of successive interference MMSE nearest neighbor receiver with noise prediction accounting for the error due to the compression. This decoder 626 works as
10 follows. Once the K compression codewords have been reconstructed, the central unit has an effective channel modeled as $\hat{Y} = HX + Z + Q$; the CP 620 applies successive interference cancellation in order to recover all the transmitted messages. First, the central unit performs linear MMSE estimation of X from \tilde{Y} using the filter designed accounting for the compression noise:

$$15 \quad \mathbf{B} = \text{SNR} \mathbf{H}_K^T (\mathbf{I}_K + \Sigma + \text{SNR} \mathbf{H}_K \mathbf{H}_K^T)^{-1}.$$

Note that this step accounts for the compression noise distribution. Then the CP 620 computes the MMSE error matrix:

$$\mathbf{K}_E = \text{SNR} \left(\mathbf{I}_K - \text{SNR} \mathbf{H}_K^T (\mathbf{I}_K + \Sigma + \text{SNR} \mathbf{H}_K \mathbf{H}_K^T)^{-1} \mathbf{H}_K \right),$$

and computes its unique Cholesky decomposition such that $\mathbf{K}_E = \text{SNR} \mathbf{L} \mathbf{L}^T$, where \mathbf{L} is a
20 lower triangular matrix with strictly positive diagonal entries. The CP 620 decodes each message successively as follows: to decode user 1, the CP 620 filters the decompressed signal with the first column of \mathbf{B} , $\mathbf{b}^{(1)}$, $\bar{\mathbf{y}}_1 = \mathbf{b}^{(1)} \tilde{\mathbf{Y}}$ and decodes the first user symbol x_1 using a nearest neighbour decoding. If w_1 is correctly decoded, the central unit computes the effective noise:

$$25 \quad \hat{\mathbf{w}}_1 = \frac{\mathbf{b}^{(1)} \tilde{\mathbf{Y}} - x_1}{l_{11} \sqrt{\text{SNR}}},$$

where l_{11} is the first diagonal entry of \mathbf{L} . Then the receiver uses $\hat{\mathbf{w}}_1$ to reduce the noise in the channel, where \mathbf{l}_1 is the first column of \mathbf{L} as:

$$\bar{\mathbf{y}}_2 = \mathbf{b}^{(1)}\hat{\mathbf{Y}} - \sqrt{SNR}\mathbf{g}_1\hat{\mathbf{w}}_1,$$

and recovers the message from the second user \mathbf{x}_2 similarly to \mathbf{x}_1 using a nearest neighbour decoding.

- 5 Many other decoding strategies can be considered, such as MMSE filtering followed by independent nearest neighbor decoding or Zero-Forcing Decoding followed by nearest neighbor decoding. For example, linear MMSE filtering followed by nearest neighbor decoding decodes each message by applying nearest neighbor in parallel to the filtered decompressed signal: $\bar{\mathbf{y}}_l = \mathbf{b}^{(l)}\hat{\mathbf{Y}}$.

10

The Joint Parameter design 628 according to B4) can be implemented as described in the following: In this block the method to jointly optimize the decoding structure and the compression parameters is applied. The performance metric of interest is the throughput, which can be characterized as follows:

$$R_1 + \dots + R_L \leq \log \left(\mathbf{I}_K - \text{SNR} \mathbf{H}_K^T (\mathbf{I}_K + \boldsymbol{\Sigma} + \text{SNR} \mathbf{H}_K \mathbf{H}_K^T)^{-1} \mathbf{H}_K \right)$$

- 15 subject to successful decompression, which happens if

$$C_k > \log \left(\frac{(E[|\mathbf{y}_k - \tilde{\mathbf{y}}_k|^2])}{\sigma^2(\Lambda_{q,k})} + 1 \right) + \log(2\pi e G(\Lambda_k))$$

In the embodiment under consideration, filtering in decoding has already been designed as the MMSE receiver, accounting for the compression noise distribution. Then, the compression parameters are designed in order to maximize the end-to-end maximum rate transmission, and are given as

20

$$\sigma^2(\Lambda_k) = \frac{v(\mathbf{y}_k|\hat{\mathbf{y}}_1^{k-1})}{\frac{2^{C_k}}{2\pi e G(\Lambda_k)} - 1},$$

$$\sigma^2(\Lambda_{q,k}) = v(\mathbf{y}_k|\hat{\mathbf{y}}_1^{k-1}) \left(1 + \frac{1}{\frac{2^{C_k}}{2\pi e G(\Lambda_k)} - 1} \right)$$

where

$$v(\mathbf{y}_k | \hat{\mathbf{y}}_1^{k-1}) = \mathbf{h}_k (\mathbf{H}_{k-1}^T \boldsymbol{\Sigma}_{k-1}^{-1} \mathbf{H}_{k-1} + \text{SNR}^{-1} \mathbf{I})^{-1} \mathbf{h}_k^T$$

is the error in reconstructing \mathbf{y}_k given $\hat{\mathbf{y}}_1, \dots, \hat{\mathbf{y}}_{k-1}$ with a linear MMSE estimator.

For different receivers, the end-to-end function to optimize changes. The decoding
 5 constrains have to be incorporated and a joint designed performed. For example, if linear
 MMSE processing is considered, the achievable sum-rate is given by

$$R_1 + \dots + R_L \leq \sum \log(1 + \text{SINR}_l)$$

where

$$\text{SINR}_l = \frac{1}{\left[\mathbf{I}_K - \text{SNR} \mathbf{H}_K^T (\mathbf{I}_K + \boldsymbol{\Sigma} + \text{SNR} \mathbf{H}_K \mathbf{H}_K^T)^{-1} \mathbf{H}_K \right]_{ll}} - 1$$

subject to successful decompression, which happens if

$$C_k > \log \left(\frac{(E[|\mathbf{y}_k - \tilde{\mathbf{y}}_k|^2])}{\sigma^2(\Lambda_{q,k})} + 1 \right) + \log(2 \pi e G(\Lambda_k))$$

Again, $\sigma^2(\Lambda_{q,k})$ and $\sigma^2(\Lambda_k)$ have to be chosen to minimize the sum-rate.

10

In the following a second embodiment of implementation of the proposed blocks and
 optimization method detailed above is described for a setup consisting of N_r receiver
 antennas per RU describing a multivariate lattice compression followed by decoding.

In this embodiment there are N_r receiver antennas per RU and a signal \mathbf{Y}_k is received.

15 Each RU then considers N_r pairs of independent nested codebooks $\Lambda_{k,r}$ and $\Lambda_{q,k,r}$ which
 are constructed as in the embodiment 1 described above. Before compressing, the N_r
 received samples are linearly transformed with a matrix \mathbf{U}_k of size $N_r \times N_r$ as $\boldsymbol{\Theta}_k = \mathbf{U}_k \mathbf{Y}_k$.
 Then, each transformed component is independently compressed by the RU considering
 the signals available at the CP at the moment of decoding, as done in Embodiment 1. The
 20 received successively recovers the compressed signals. The previously decoded signals
 already decoded signals $\hat{\boldsymbol{\Theta}}_{\mathcal{R}}$ are used as side information, where we note by \mathcal{R} the set
 containing the indices of the signals already decompressed, i.e., $\mathcal{R} = \{1, \dots, k-1\}$. Then,
 the k -th BS applies independent lattice compression on each component. After lattice

compression and decompression, the effective channel can be characterized as $\widehat{\mathbf{Y}}_k = \mathbf{H}_k \mathbf{X} + \mathbf{Z}_k + \mathbf{Q}_k$, where \mathbf{Q}_k is due to the compression, has covariance matrix $\mathbf{\Omega}_k$ and each of its component is uniformly distributed over the Voronoi region of $\Lambda_{q,k}$ and

$$\mathbf{\Omega}_k = \mathbf{U}_k \begin{bmatrix} \eta_{k,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \eta_{k,N_r} \end{bmatrix}^{-1} \mathbf{U}_k^T,$$

5 where the compression parameters are given as

$$\sigma^2(\Lambda_{kj}^r) = \lambda_{k,j} + \eta_{k,j}^{-1},$$

and

$$\sigma^2(\Lambda_{q,k,j}^r) = \eta_{k,j}^{-1},$$

where $\mathbf{S}_k = \text{diag}(\lambda_{k,1}, \dots, \lambda_{k,N_r})$ and \mathbf{U}_k are obtained from the SVD

$$10 \quad \mathbf{\Sigma}_{\Theta_k | \widehat{\Theta}_{\mathcal{R}}} : \bar{\mathbf{H}}_{\mathcal{R}} \mathbf{\Sigma}_{\mathbf{x} | \Theta_{\mathcal{R}}} \bar{\mathbf{H}}_{\mathcal{R}}^T + \mathbf{I} = \mathbf{U}_k \mathbf{S}_k \mathbf{U}_k^T,$$

where

$$\mathbf{\Sigma}_{\mathbf{x} | \widehat{\Theta}_{\mathcal{R}}} = \mathbf{\Sigma}_{\mathbf{x}} - \mathbf{\Sigma}_{\mathbf{x}} \bar{\mathbf{H}}_{\mathcal{R}}^T (\bar{\mathbf{H}}_{\mathcal{R}} \mathbf{\Sigma}_{\mathbf{x}} \bar{\mathbf{H}}_{\mathcal{R}}^T + \mathbf{\Sigma}_{\mathcal{R}})^{-1} \bar{\mathbf{H}}_{\mathcal{R}}^T \mathbf{\Sigma}_{\mathbf{x}}$$

and

$$\mathbf{\Sigma}_k = \mathbf{I} + \mathbf{\Omega}_k,$$

15 and

$$\mathbf{\Omega}_{\mathcal{R}} = \text{diag}(\mathbf{\Omega}_{\pi(1)}, \dots, \mathbf{\Omega}_{\pi(k-1)}).$$

Similarly, to block B4) 628, now, compression parameters may be designed jointly depending on the decoder.

20

For any given metric, this problem can be casted as an optimization problem over the lattice parameters and examples of its solutions are given below with respect to Fig. 15.

Fig. 7 shows a schematic diagram of a multi-terminal network system 700 using lattice based compression and decompression according to an implementation form. The multi-

25

terminal network system 700 is an implementation of the multi-terminal network system 600 described above with respect to Fig. 6.

5 The multi-terminal network system 700 includes a radio system 710 and a decoding device 740. The radio system 710 includes a plurality of radio units 711, 721, 731 transmitting signals over respective CPRI links 715, 725, 735 to the decoding device 740.

10 Each radio unit includes a receive interface 712, 722, 732, compression parameters received from the decoding device 740, a compressor 717, 727, 737 and a transmit interface, e.g. an antenna port 715, 725, 735. The receive interface 712, 722, 732 is configured to receive at least one radio signal, y_k , over a multiple-input multiple-output, MIMO, radio channel. The compressor 717, 727, 737 is configured to generate a codeword signal, λ_k , based on the compression parameters by encoding the at least one radio signal y_k with a lattice based distributed code. The transmit interface 715, 725, 735
15 is configured to transmit the codeword signal λ_k to the decoding device 740.

These radio units 711, 721, 731 can be implemented in one physical device 710, e.g. according to the Massive MIMO station as described above with respect to Fig. 2. Each of the radio units 711, 721, 731 may be coupled to one or more antennas 716, 726, 736.
20 Alternatively, the radio units 711, 721, 731 can be implemented as a system of physical devices, where each physical device includes one or more of the radio units 711, 721, 731, e.g. according to the Cloud Radio Access Network as described above with respect to Fig. 3. Each of the radio units 711, 721, 731 may be coupled to one or more antennas 716, 726, 736.

25 The compression parameters may comprise parameters of a plurality of nested lattices forming the lattice based distributed code. The lattice based distributed code may be known at the radio system 710 and the compression parameters may comprise second order moments of the plurality of nested lattices. The compressor may comprise a dither module 801 configured to add a dither to the radio signal y_k ; a quantization module 802,
30 configured to quantize the dithered radio signal based on the plurality of nested lattices; and a modulo reduction module 803, configured to modulo reduce the quantized dithered radio signal to generate the codeword signal, λ_k , e.g. as described below with respect to

Fig. 8. The compressor 717, 727, 737 may be configured to compress the radio signal y_k based on a lattice operation comprising a quantization and a modulo reduction.

The decoding device 740 includes a receive interface, a decompressor 750, a decoder
5 760, a controller and a transmit interface. The receive interface is configured to receive from a plurality of radio units 711, 721, 731 a plurality of signals, λ_k , over a plurality of links. Each signal λ_k corresponds to a compressed radio signal y_k . The compressed radio signal y_k carries a plurality of messages, ω_l , from a plurality of users. The decompressor 750 is configured to decompress the plurality of signals λ_k with a nested lattice based
10 distributed code to provide a plurality of decompressed radio signals, \hat{y}_k . The decoder 760 is configured to decode the plurality of decompressed radio signals, \hat{y}_k , to recover the plurality of messages. The controller is configured to determine compression parameters based on the decoder structure and based on a correlation of the received signals. The transmit interface is configured to transmit the compression parameters to
15 each of the plurality of radio units 711, 721, 731. The decompressor 750 may comprise a plurality of lattice based reconstruction modules 751, 752, configured to successively decompress the plurality of signals λ_k .

A decompression result of a reconstruction module 751, 752 of the plurality of lattice
20 based reconstruction modules may depend on a decompression result of a preceding reconstruction module of the plurality of lattice based reconstruction modules.

A decompression result of a reconstruction module 751, 752 of the plurality of lattice
based reconstruction modules may depend on side information 753, 754 generated from
25 decompression results of preceding reconstruction modules of the plurality of lattice based reconstruction modules.

The decompressor 750 may comprise an optimum filter estimator, in particular a linear
minimum mean squares estimator, LMMSE, configured to generate the side information
30 based on optimum filter estimation using the decompression results of the preceding reconstruction modules of the plurality of lattice based reconstruction modules 751, 752.

Each reconstruction module 751, 752 may comprise a combiner, configured to combine a respective signal λ_k with the corresponding side information to generate the respective decompressed radio signal, \hat{y}_k . The k-th combiner may be based on the following lattice operation: $\hat{y}_k = [\lambda_k - \tilde{y}_k - u_k] \bmod \Lambda_k + \tilde{y}_k$, where λ_k is the k-th signal, \tilde{y}_k is the k-th side information, Λ_k is the k-th lattice, u_k is a k-th dither and \hat{y}_k is the k-th decompressed radio signal, for example as described above with respect to Fig. 6.

The controller may be configured to determine the compression parameters based on a decoding metric depending on the decoding operation of the decoder 760, the decompression operation of the decompressor 750 and a correlation of the received signals.

The decoding metric may depend on at least one of the following parameters: a number of the plurality of users, channel characteristics of the plurality of links, signal to noise ratios of the plurality of links, a quantization and/or statistics of a quantization error due to the nested lattice distributed code.

The controller may be configured to maximize the decoding metric to determine optimal compression parameters.

Compression may be performed at each of the K RRUs independently, whereas decompression may be performed with K decompressions, but utilizing a successive decompression. At decompression k, the quantization codeword may be utilized, as well as the effective side information sequence \tilde{y}_k , generated from the previously decompressed signals in the previous k-1 decompressors. The use of nested lattice compression and decompression allows to combine at the decompressor units the quantization codewords and the side information sequence.

Successive decompression at the central processor may be implemented as follows: K signals are compressed, forwarded and then decompressed one after the other using K decompressors, one at each step. At each decompression step, the previously reconstructed baseband signals, $\hat{y}_1, \dots, \hat{y}_{k-1}$ are used to generate a side information sequence, which is provided to the decompressor to facilitate the decompression operation of y_k .

The successive decomposition combining utilizing the corresponding quantization codeword and the effective side information sequence procedure allows to exploit correlation, and to reduce the degradation in the reconstruction. The use of distributed
5 nested lattice codes allows constructing compression codes that can combine the side information sequence and the quantization codeword at the compressor.

The design of the nested lattice compression parameters accounts for the decoder structure, whereas in the baseline algorithms as described above with respect to Fig. 4,
10 there is only minimization of the reconstruction error between the received signal and its reconstruction.

The correlation of the signal is essentially the covariance matrix (or other characterizations such as the autocorrelation functions) of the received signals at different
15 antennas. In general, the covariance can be given as $C_y = E[YY^H]$ where $Y = [y_1, \dots, y_K]$. This statistical relation can be estimated at the CP with various methods, for example in a training or by other classical means.

20 Fig. 8 shows a block diagram of a compressor 800 according to an implementation form and Fig. 9 shows a block diagram of a decompressor 900 according to an implementation form.

25 The compressor 800 includes a dither module 801 configured to add a dither to the radio signal y_k ; a quantization module 802, configured to quantize the dithered radio signal based on the plurality of nested lattices; and a modulo reduction module 803, configured to modulo reduce the quantized dithered radio signal to generate the codeword signal, λ_k . The decompressor 900 includes a lattice decomposition module 910 including a dither
30 module 901 for adding a dither u_k to the codeword λ_k and a modulo reduction module 902 for performing a modulo reduction operation on the dithered codeword. The decompressor 900 further includes a side-information module 903 for adding 904, 905 side information to the dithered codeword or to the result of the modulo reduction operation 902.

Fig. 10 shows a schematic diagram of a C-RAN network 1000 according to an implementation form. The C-RAN network 1000 includes a central unit 1001, a plurality of Relaying units 1002, 1003, 1004 which relay radio signals to UEs 1010, 1011, 1012, 1013.

5

Figure 10 and the further figures describe several applications and their performance. In the simulations, the following scenario is applied: Flat fading channel coefficients:

$\mathbf{h}_{k,l} \sim N(0,1)$ and 2000 realizations; Equal limited capacity backhaul links of capacity

$C_1 = \dots = C_k = C$; Additive Noise power $N = 1$; and Signal to Noise Ratio: $\text{SNR} = \frac{P}{N}$.

10

In the C-RAN network 1000 single antenna transmitters and single antenna base stations (acting as RU) are connected to a central processor 1001 by means of error-free fronthaul links of limited capacity C_k . The performance of the disclosed scheme is evaluated with respect to the state of the art point-to-point compressors. In particular, the average

15

achievable sum-rate is illustrated in the following scenario: Number of transmitters: $L = 10$; and Receive BS: $K = 7$.

Fig. 11 shows a performance diagram illustrating average throughput 1100 versus SNR in a C-RAN network for compression with different lattice codes according to an

20

implementation form and Fig. 12 shows a performance diagram illustrating average throughput 1100 versus front-haul capacity in a C-RAN network for compression with different lattice codes according to an implementation form. Figure 11 depicts the average throughput with respect to the SNR when the compressor uses spheric lattices 1102, 1103 and finite dimension lattices cubic lattice 1104, 1105, i.e., $\Lambda = Z$. The Cut-Set 1101 is

25

depicted as reference. It can be observed that the LSWZ outperforms PtP compression in all regimes. Similarly, Figure 12 shows the achievable throughput under different fronthaul capacity values for spheric lattices 1202, 1205 and cubic lattices 1203, 1204. The Cut-Set 1201 is depicted as reference.

30

Fig. 13 shows a performance diagram illustrating average EVM 1300 versus front-haul capacity in a C-RAN network for compression with different lattice codes according to an implementation form and Fig. 14 shows a performance diagram illustrating average EVM 1300 versus front-haul capacity gain in a C-RAN network for compression with different lattice codes according to an implementation form. Figure 13 depicts the average EVM,

defined as the average distortion at which the transmitted symbols can be reconstructed from the compressed signal

$$EVM = \frac{1}{L} \sum E \left[(\mathbf{x}_l - E[\mathbf{x}_l | \hat{\mathbf{Y}}])^2 \right]$$

when the compressor uses spheric lattices 1302 and finite dimension lattices cubic lattice 1301, i.e., $\Lambda = Z$, while Figure 14 shows the relative gains of the disclosed method for spheric lattices 1402 and cubic lattices 1401 with respect to point to point compression. It can be observed that the LSWZ achieves high gains with respect to standard compression in terms of average EVM. The achievable throughput under different front-haul capacity values shows similar behavior..

- 10 Fig. 15 shows a schematic diagram of a Massive MIMO station with dedicated CPRI links 1500 according to an implementation form.

Here, a Massive MIMO setup with K antennas 1504 serving L users 1520, 1521, 1522 is considered. Each antenna is assigned a dedicated CPRI link 1505 which connects it to the central processor 1506. In order to reduce complexity, each antenna 1504 compresses the received signal by using a respective compressor 1511, 1512, 1513, 1514, 1515, 1516 without access to the signals at the other antennas. The CP 1506 includes a joint decompression unit, i.e. a decompressor as described above with respect to Fig. 6 or 7, and a centralized decoder, i.e. a decoder as described above with respect to Fig. 6 or 7.

The CP 1506 uses a successive interference MMSE receiver and optimizes the Lattice parameters as in the CRAN setup, i.e. according to:

$$\sigma^2(\Lambda_k) = \frac{v(\mathbf{y}_k | \hat{\mathbf{y}}_1^{k-1})}{\frac{2^{C_k}}{2\pi e G(\Lambda_k)} - 1},$$

$$\sigma^2(\Lambda_{q,k}) = v(\mathbf{y}_k | \hat{\mathbf{y}}_1^{k-1}) \left(1 + \frac{1}{\frac{2^{C_k}}{2\pi e G(\Lambda_k)} - 1} \right)$$

The performance of the disclosed scheme is evaluated with respect to the state of the art point-to-point compressors. In particular, the average achievable sum-rate is considered in the scenario with Number of transmitters $L = 20$; and receive antennas $K = 100$.

- 5 Fig. 16 shows a performance diagram illustrating average throughput 1600 versus SNR in a Massive MIMO network for compression with different lattice codes according to an implementation form and Fig. 17 shows a performance diagram illustrating average throughput 1700 versus front-haul capacity in a Massive MIMO network for compression with different lattice codes according to an implementation form. Figure 16 depicts the
- 10 huge performance improvements are possible in the Massive MIMO setup by using the disclosed method. This is observed in the average throughput with respect to the SNR when the compressor uses Spheric lattices 1602, 1604 and finite dimension lattices cubic lattice 1603, 1605, i.e., $\Lambda = Z$. It can be seen that the LSWZ 1602, 1603 outperforms PtP compression 1604, 1605 in all regimes, The Cut-Set 1601 is depicted as reference.
- 15 Similarly, Figure 17 shows the achievable throughput under different fronthaul capacity values for spheric lattices 1703, 1705 and cubic lattices 1702, 1704. The Cut-Set 1701 is depicted as reference.

- 20 Fig. 18 shows a performance diagram illustrating average EVM 1800 versus front-haul capacity in a Massive MIMO network for compression with different lattice codes according to an implementation form and Fig. 19 shows a performance diagram illustrating average EVM 1900 versus front-haul capacity gain in a Massive MIMO network for compression with different lattice codes according to an implementation form. Figure 18 depicts the average EVM when the disclosed method is used with Spheric lattices
- 25 1802 and finite dimension lattices cubic lattice 1801, i.e., $\Lambda = Z$. It can be seen that the LSWZ outperforms PtP compression in all regimes, Similarly, Figure 19 shows the gains in EVM achieved with the disclosed method for spheric lattices 1902 and cubic lattices 1901 with respect to point-to-point compression.

- 30 Fig. 20 shows a performance diagram illustrating average EVM 2000 versus front-haul capacity in a Massive MIMO network with shared CPRI for compression with different lattice codes according to an implementation form and Fig. 21 shows a performance diagram illustrating average EVM 2000 versus SNR per user in a Massive MIMO network with different lattice codes according to an implementation form. A single MIMO RU
- 35 scenario with $K = 2$ antennas serving $L = 4$ users is considered. It is assumed that both

antennas share a CPRI link and that the signals at both antennas can be combined before compression. The achievable rate can be determined as an optimization problem and finding the optimal compression parameters when the decoder is an MMSE-SIC decoder and an linear MMSE decoder. The performance is compared with the case in which the
5 compression parameters are chosen to minimize the distortion between the received signal and the reconstruction signal. In all cases, the linear transformation mapping before applying compression is chosen as the eigenvectors of matrix $HH^H = USU^H$.

Figure 20 and Figure 21 show the average with respect to the available fronthaul and the
10 SNR per UE. The following curves achieved by optimizing the optimizing the lattice parameters are depicted: MMSE-SIC Opt 2002, 2103: maximum information transfer assuming that there is a MMSE successive interference cancellation decoder at the CP; MMSE-SIC Dist 2003, 2102: MMSE-SIC decoder and compression designed to minimize distortion between received and reconstructed signals at the CP; MMSE Opt 2105:
15 maximum information transfer assuming that there is a linear MMSE decoder at the CP; MMSE Dist 2005, 2104: MMSE decoder and compression designed to minimize distortion between received and reconstructed signals at the CP; and MMSE with ML par 2006, 2106: MMSE decoder with compression designed assuming that there is an ML decoder at the CP. This is the state of the art solutions. The Cut-Set 2001, 2101 is depicted as
20 reference.

A loss in performance can be observed if compression is not designed taking into account the received at the CP. If MMSE-SIC is used, the performance suffers if the compression is designed to minimize distortion. Similarly, a loss in performance is observed if an ML
25 receiver is assumed for compression design while the true decoder is a suboptimal linear MMSE decoder. This loss is significant at high SNR regimes, as observed in Figure 21.

The present disclosure relates to the case of a compression and decoding design based on nested lattice based compression and decompression and the joint design of
30 compression and decoding parameters to optimize a certain metric. The present disclosure also supports a method to encode the lattices to account for the correlation of the signal received at different radio units. A method to successively reconstruct the compressed signals, to generate an effective side information sequence, and an effective channel is used in the decoding process are also provided. Degradation due to

compression is also incorporated at the decoder. It has been shown that gains over the situation in which compressor and the decoder are designed independently can be achieved.

5 The present disclosure also supports a computer program product including computer executable code or computer executable instructions that, when executed, causes at least one computer to execute the performing and computing steps described herein, in particular the steps of the methods described above. Such a computer program product may include a readable non-transitory storage medium storing program code thereon for
10 use by a computer. The program code may perform the performing and computing steps described herein, in particular the methods described above.

While a particular feature or aspect of the disclosure may have been disclosed with respect to only one of several implementations, such feature or aspect may be combined
15 with one or more other features or aspects of the other implementations as may be desired and advantageous for any given or particular application. Furthermore, to the extent that the terms "include", "have", "with", or other variants thereof are used in either the detailed description or the claims, such terms are intended to be inclusive in a manner similar to the term "comprise". Also, the terms "exemplary", "for example" and "e.g." are
20 merely meant as an example, rather than the best or optimal. The terms "coupled" and "connected", along with derivatives may have been used. It should be understood that these terms may have been used to indicate that two elements cooperate or interact with each other regardless whether they are in direct physical or electrical contact, or they are not in direct contact with each other.

25 Although specific aspects have been illustrated and described herein, it will be appreciated by those of ordinary skill in the art that a variety of alternate and/or equivalent implementations may be substituted for the specific aspects shown and described without departing from the scope of the present disclosure. This application is intended to cover
30 any adaptations or variations of the specific aspects discussed herein.

Although the elements in the following claims are recited in a particular sequence with corresponding labeling, unless the claim recitations otherwise imply a particular sequence for implementing some or all of those elements, those elements are not necessarily
35 intended to be limited to being implemented in that particular sequence.

Many alternatives, modifications, and variations will be apparent to those skilled in the art in light of the above teachings. Of course, those skilled in the art readily recognize that there are numerous applications of the invention beyond those described herein. While
5 the present invention has been described with reference to one or more particular embodiments, those skilled in the art recognize that many changes may be made thereto without departing from the scope of the present invention. It is therefore to be understood that within the scope of the appended claims and their equivalents, the invention may be practiced otherwise than as specifically described herein.

CLAIMS:

1. A radio system (710) comprising a plurality of radio units (711, 721, 731), each radio unit, comprising:
 - 5 a receive interface (712, 722, 732), configured to receive
at least one radio signal, y_k , over a multiple-input multiple-output, MIMO, radio channel;
compression parameters from a decoding device (740);
a compressor (717, 727, 737) configured to generate a codeword signal, λ_k , based
10 on the compression parameters by encoding the at least one radio signal y_k with a lattice based distributed code; and
a transmit interface (715, 725, 735), configured to transmit the codeword signal λ_k to the decoding device (740).
2. The radio system (710) of claim 1,
 - 15 wherein the compression parameters comprise parameters of a plurality of nested lattices forming the lattice based distributed code.
3. The radio system (710) of claim 2,
Wherein the lattice based distributed code is known at the radio system (710) and the compression parameters comprise second order moments of the plurality of nested
20 lattices.
4. The radio system (710) of claim 2 or 3, wherein the compressor comprises:
 - a dither module (801) configured to add a dither to the radio signal y_k ;
 - a quantization module (802), configured to quantize the dithered radio signal based on the plurality of nested lattices; and

a modulo reduction module (803), configured to modulo reduce the quantized dithered radio signal to generate the codeword signal, λ_k .

5. The radio system (710) of claim 4,

wherein the compressor (717, 727, 737) is configured to compress the radio signal y_k based on a lattice operation comprising a quantization and a modulo reduction.

6. A decoding device (740), comprising:

a receive interface, configured to receive from a plurality of radio units (711, 721, 731) a plurality of signals, λ_k , over a plurality of links, each signal λ_k corresponding to a compressed radio signal y_k , wherein the compressed radio signal y_k carries a plurality of messages, ω_l , from a plurality of users;

a decompressor (750), configured to decompress the plurality of signals λ_k with a nested lattice based distributed code to provide a plurality of decompressed radio signals, \hat{y}_k ;

a decoder (760), configured to recover the plurality of messages, ω_l , from the plurality of decompressed radio signals \hat{y}_k ;

a controller, configured to determine compression parameters based on the decoder structure and based on a correlation of the received signals; and

a transmit interface configured to transmit the compression parameters to each of the plurality of radio units (711, 721, 731).

7. The decoding device (740) of claim 6,

wherein the decompressor (750) comprises a plurality of lattice based reconstruction modules (751, 752), configured to successively decompress the plurality of signals λ_k .

8. The decoding device (740) of claim 7,

wherein a decompression result of a reconstruction module (751, 752) of the plurality of lattice based reconstruction modules depends on a decompression result of a preceding reconstruction module of the plurality of lattice based reconstruction modules.

9. The decoding device (740) of claim 7 or 8,

5 wherein a decompression result of a reconstruction module (751, 752) of the plurality of lattice based reconstruction modules depends on side information (753, 754) generated from decompression results of preceding reconstruction modules of the plurality of lattice based reconstruction modules.

10. The decoding device (740) of claim 9,

10 wherein the decompressor (750) comprises an estimator, in particular a linear filtering estimator, configured to generate the side information based on estimation using the decompression results of the preceding reconstruction modules of the plurality of lattice based reconstruction modules (751, 752).

11. The decoding device (740) of claim 9 or 10,

15 wherein each reconstruction module (751, 752) comprises a combiner, configured to combine a respective signal λ_k with the corresponding side information to generate the respective decompressed radio signal, \hat{y}_k .

12. The decoding device (740) of claim 11,

wherein the k-th combiner is based on the following lattice operation:

20
$$\hat{y}_k = [\lambda_k - \tilde{y}_k - \mathbf{u}_k] \bmod \Lambda_k + \tilde{y}_k$$

where λ_k is the k-th signal, \tilde{y}_k is the k-th side information, Λ_k is the k-th lattice, u_k is a k-th dither and \hat{y}_k is the k-th decompressed radio signal.

13. The decoding device (740) of one of claims 6 to 12,

25 wherein the controller is configured to determine the compression parameters based on a decoding metric depending on the decoding operation of the decoder (760), the decompression operation of the decompressor (750) and a correlation of the received signals.

14. The decoding device of claim 13,

wherein the decoding metric depends on at least one of the following parameters:

a number of the plurality of users, channel characteristics of the plurality of links,
signal to noise ratios of the plurality of links, a quantization and/or statistics of a
5 quantization error due to the nested lattice distributed code.

15. The decoding device (740) of claim 13 or 14,

wherein the controller is configured to maximize the decoding metric to determine
optimal compression parameters.

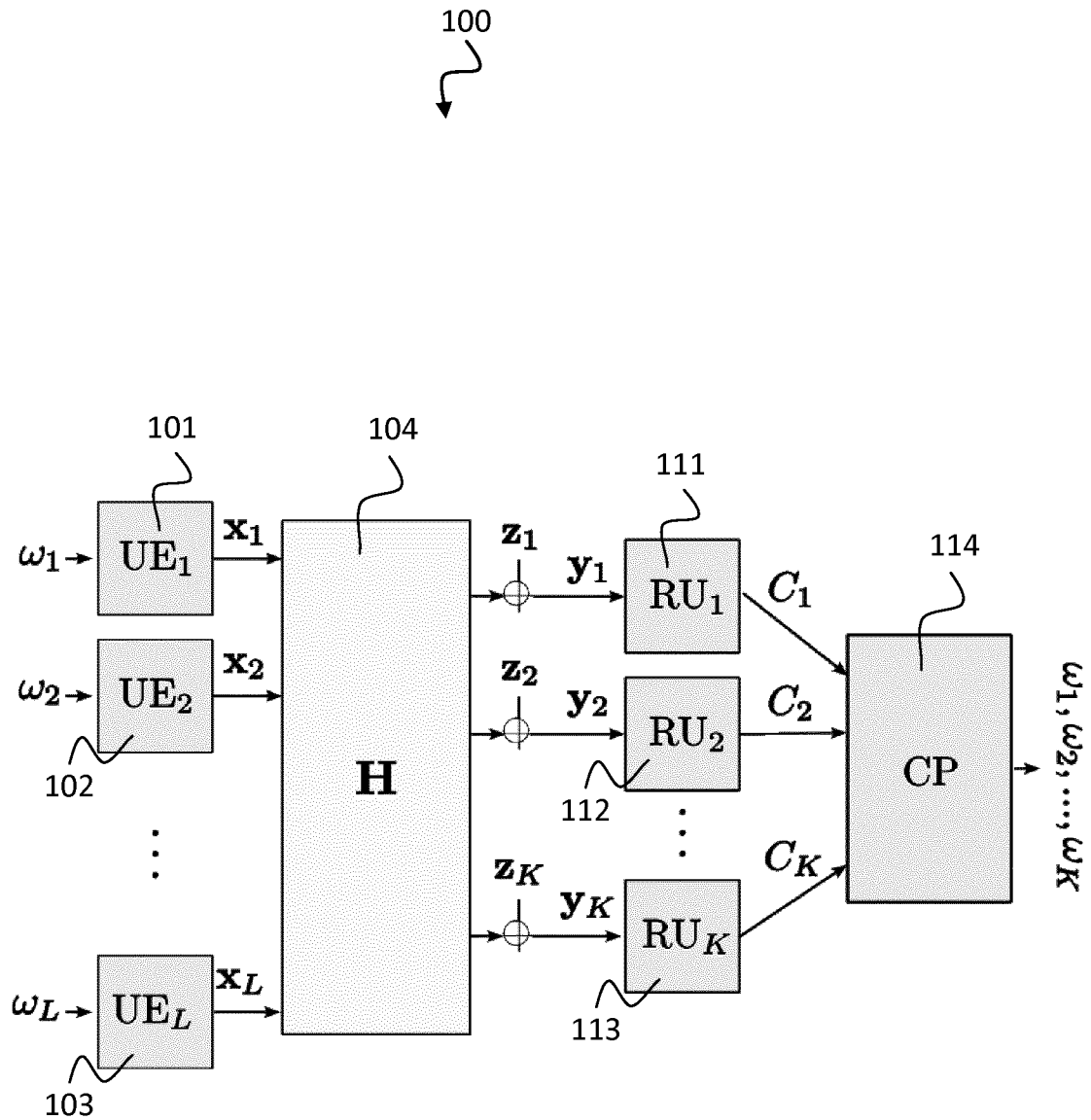


Fig. 1

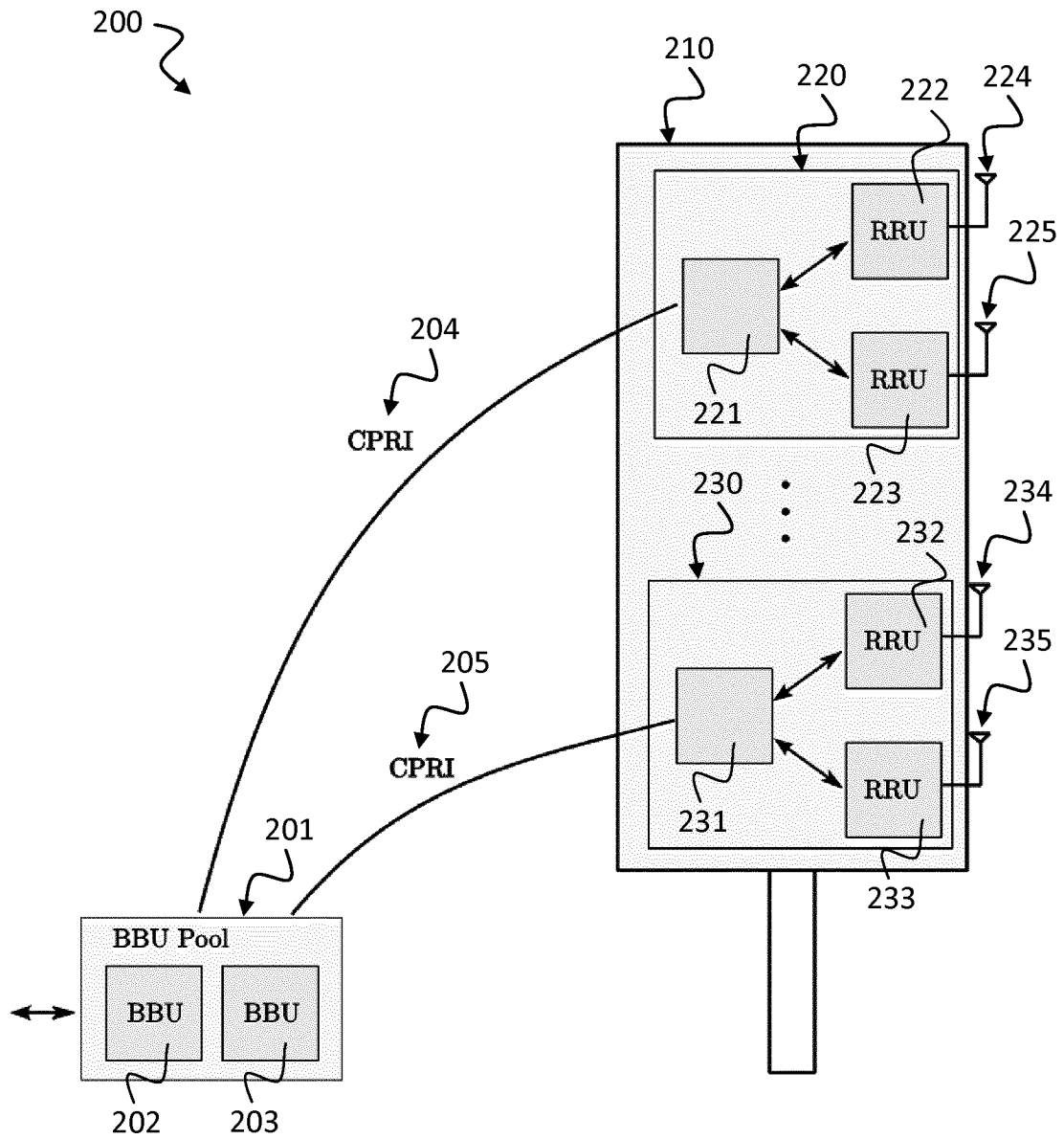


Fig. 2

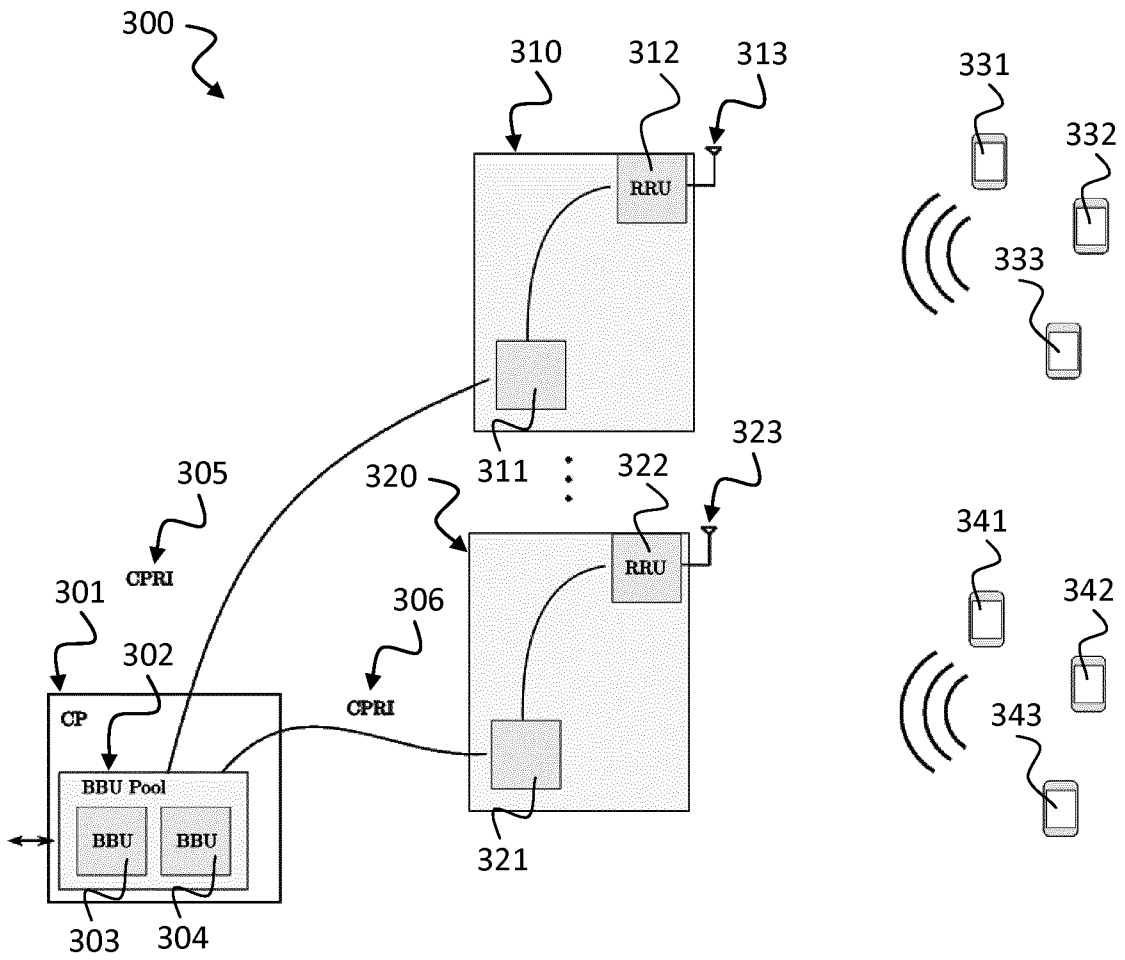


Fig. 3

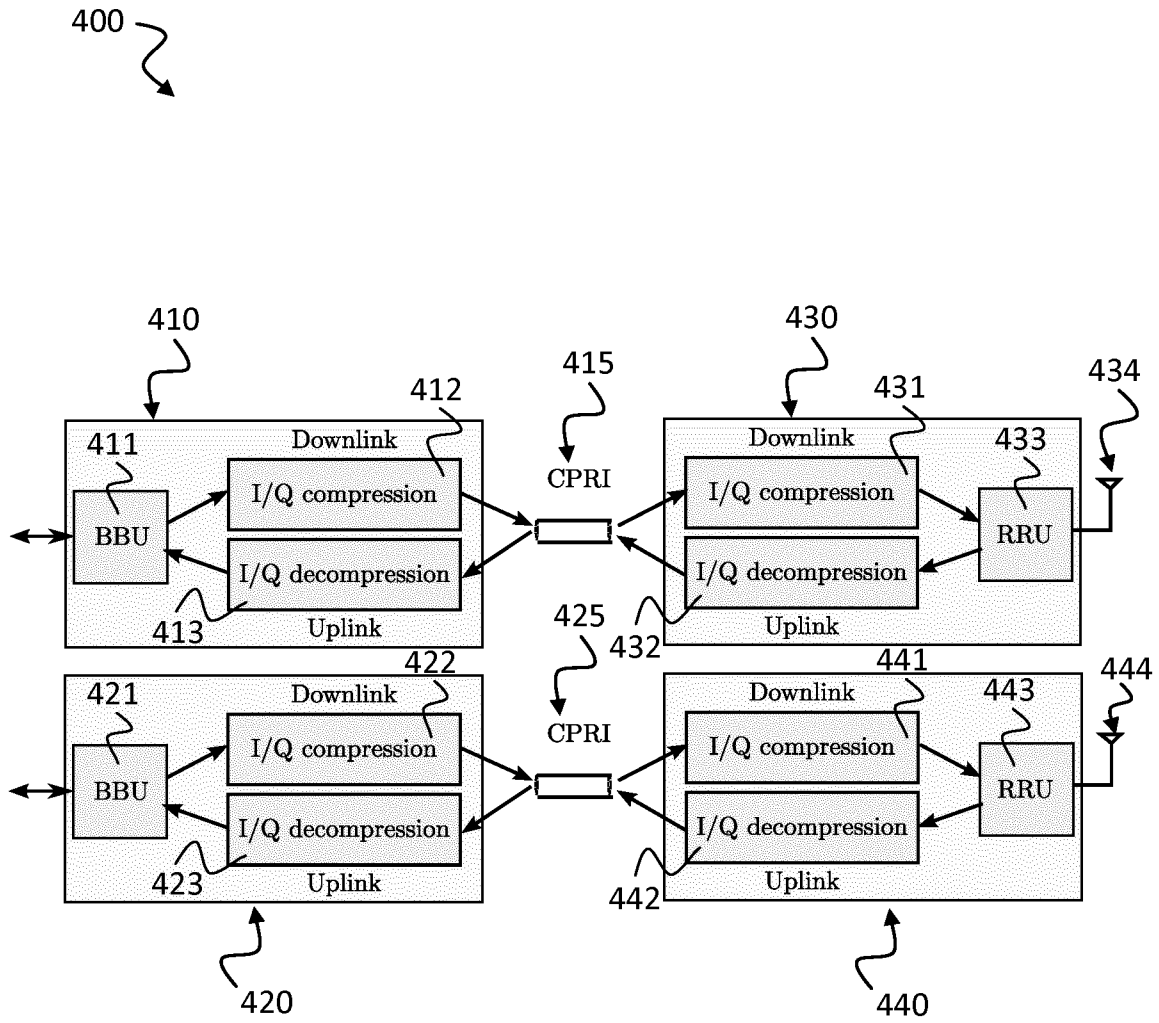


Fig. 4

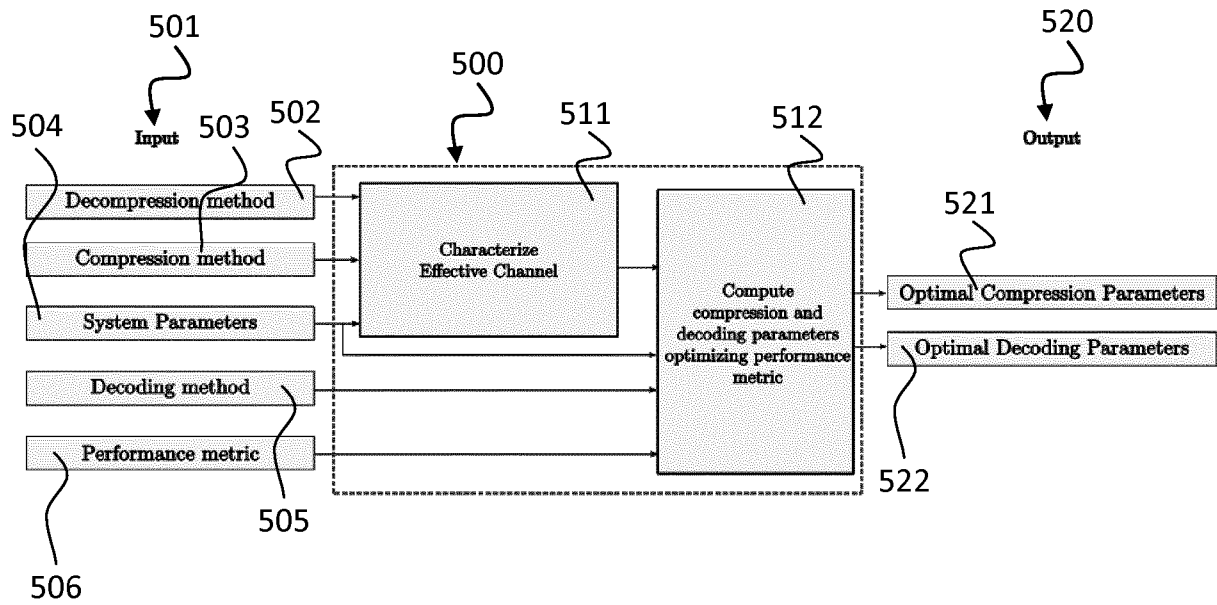


Fig. 5

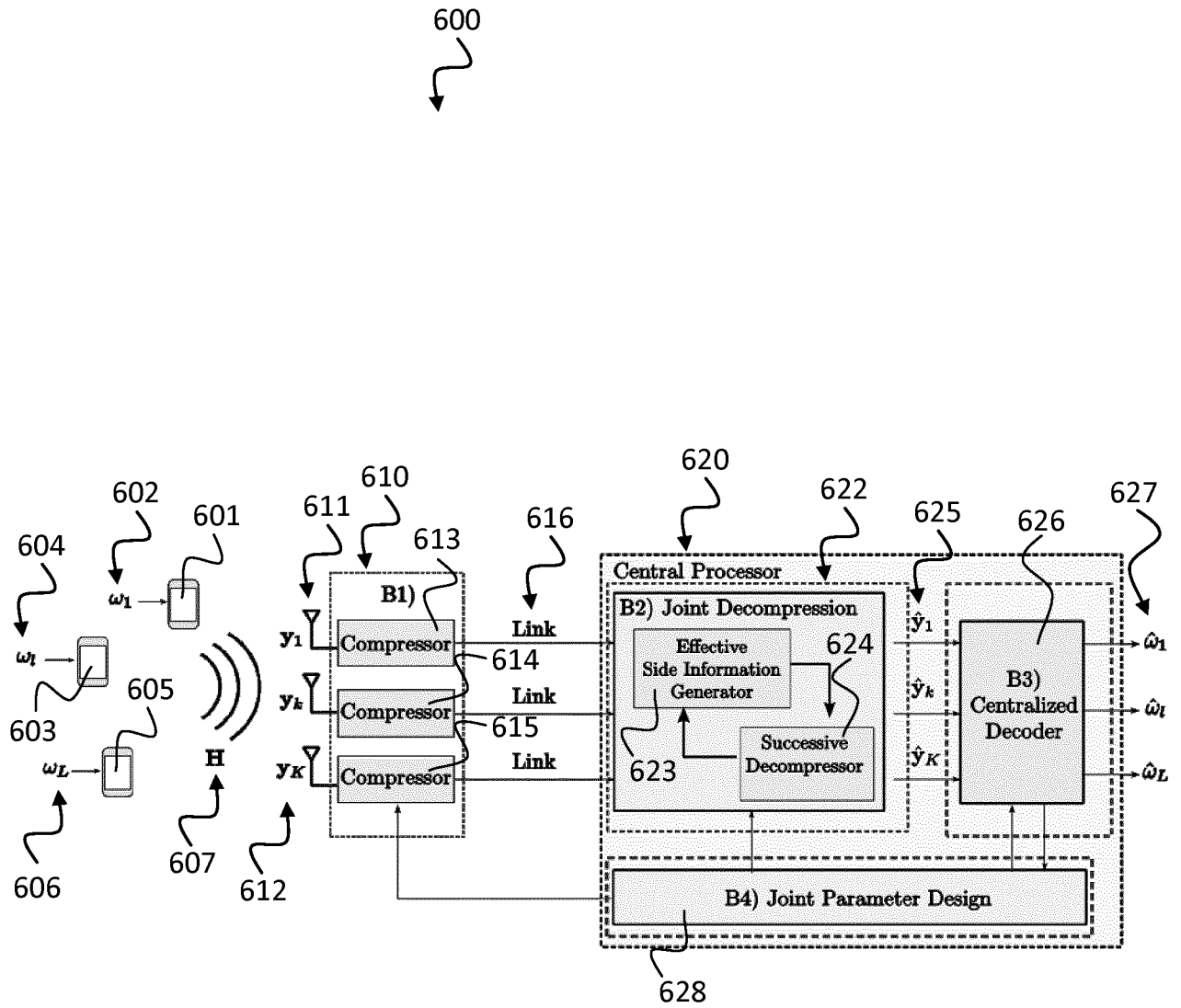


Fig. 6

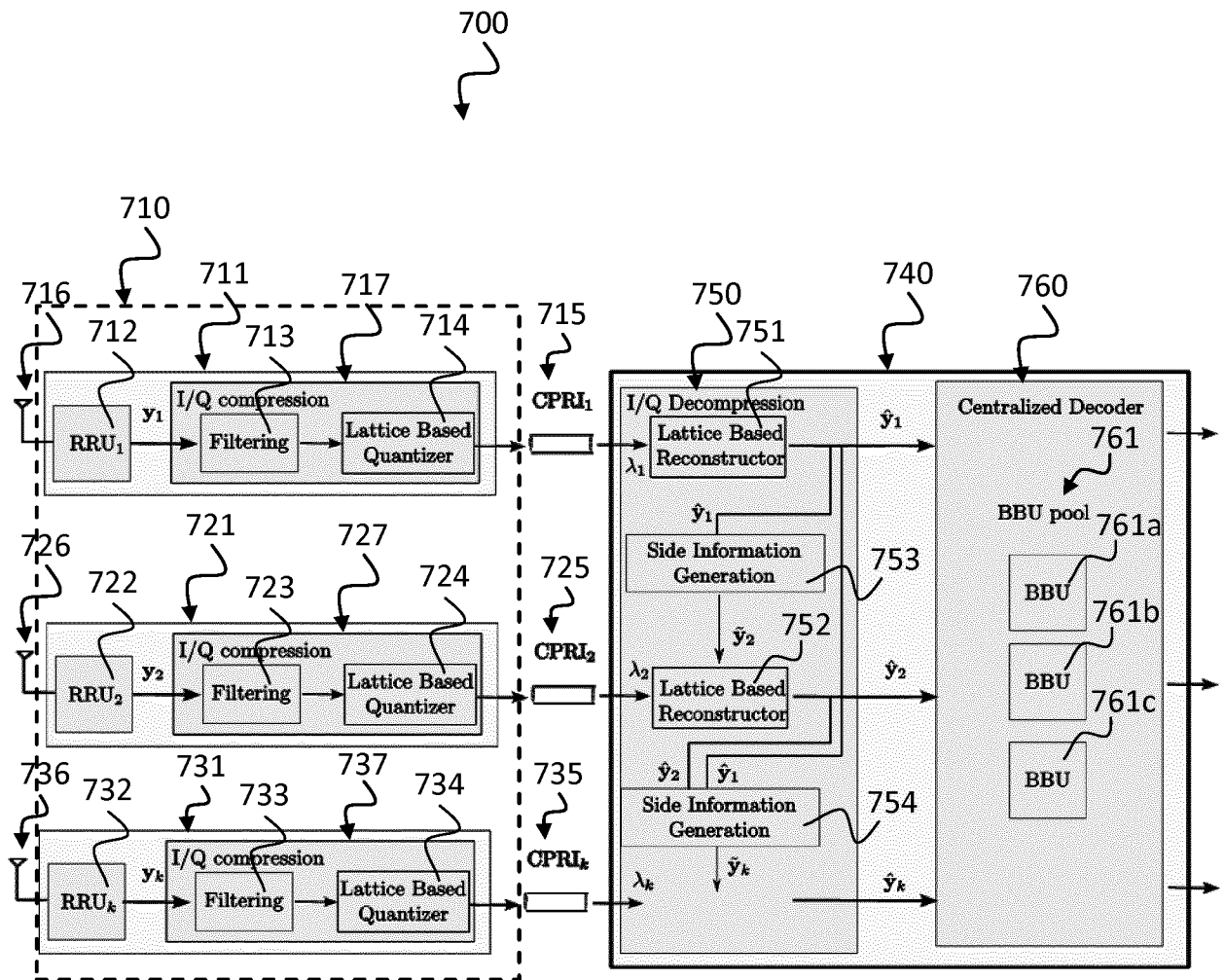


Fig. 7

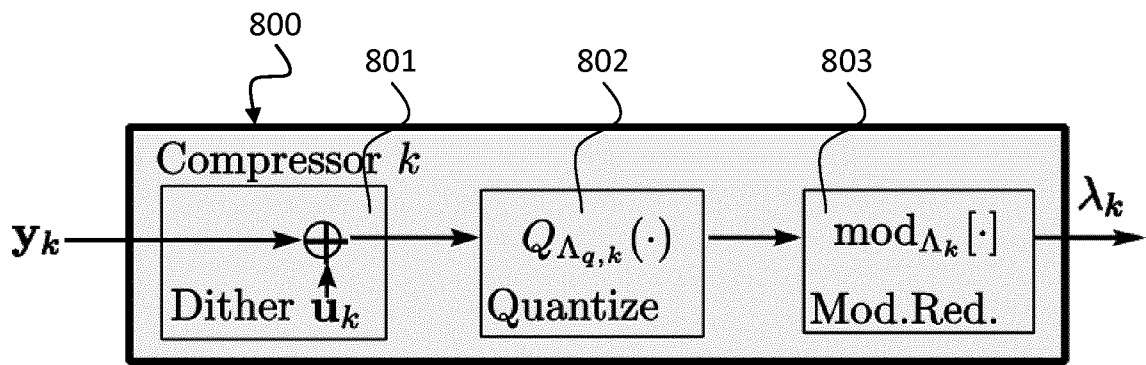


Fig. 8

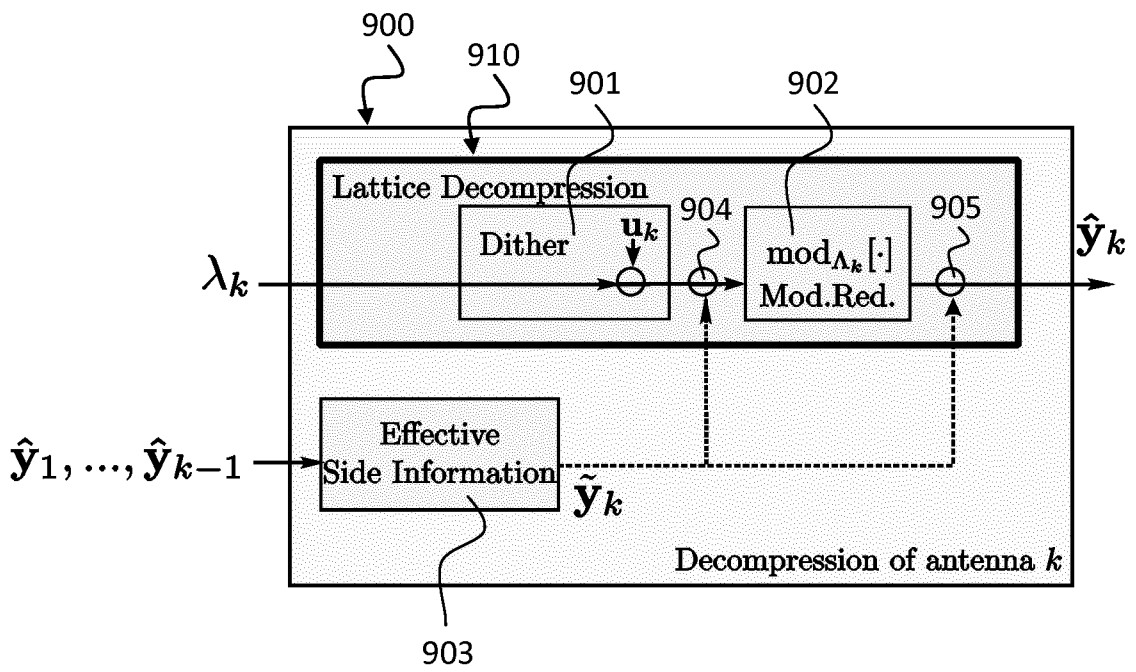


Fig. 9

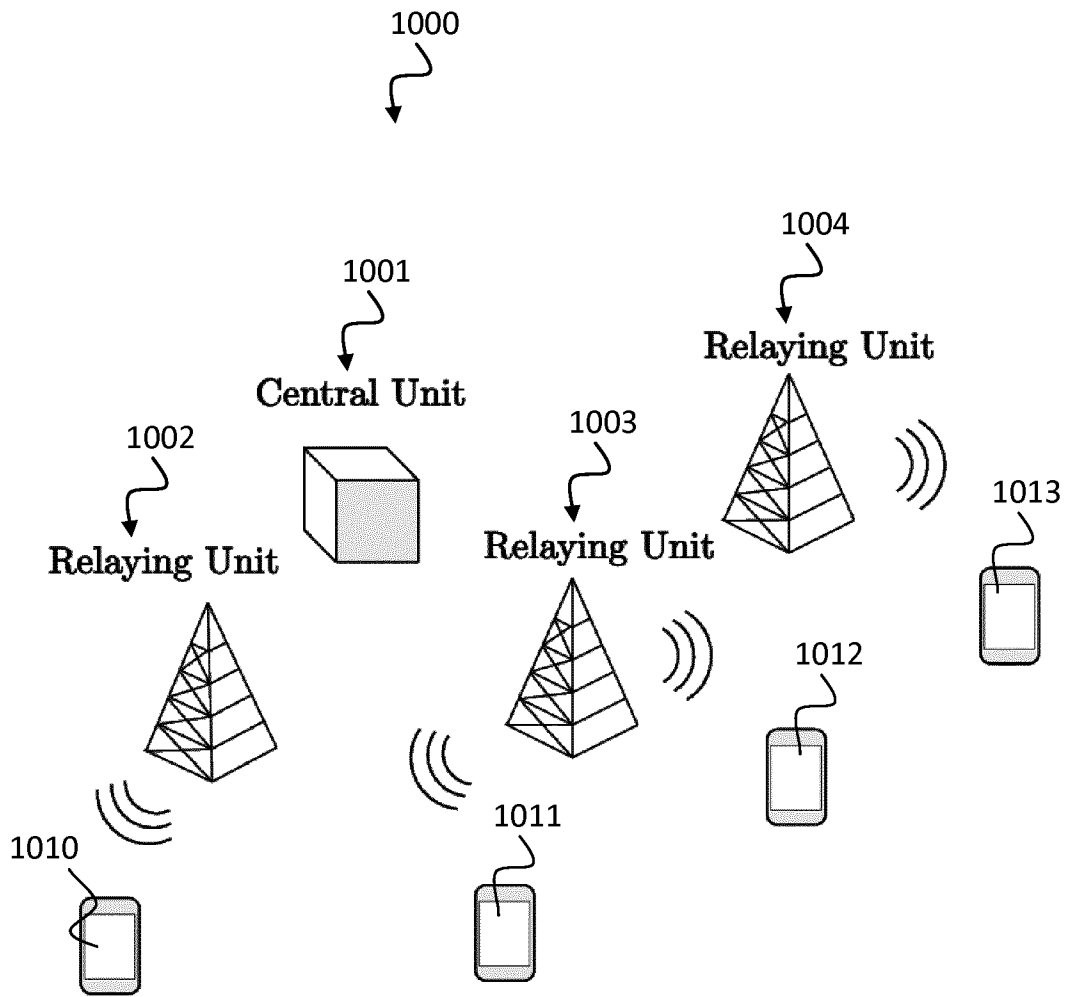


Fig. 10

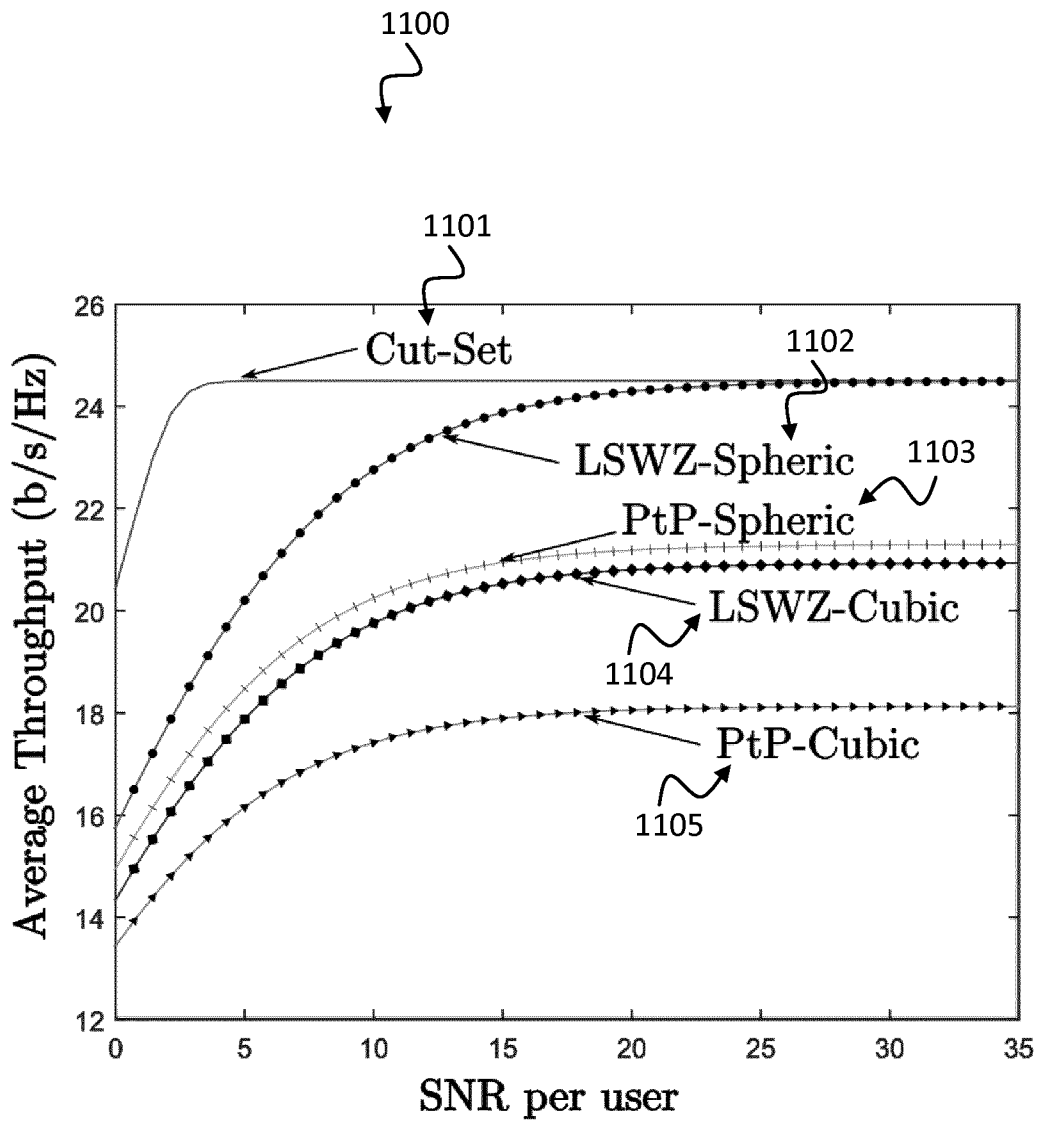


Fig. 11

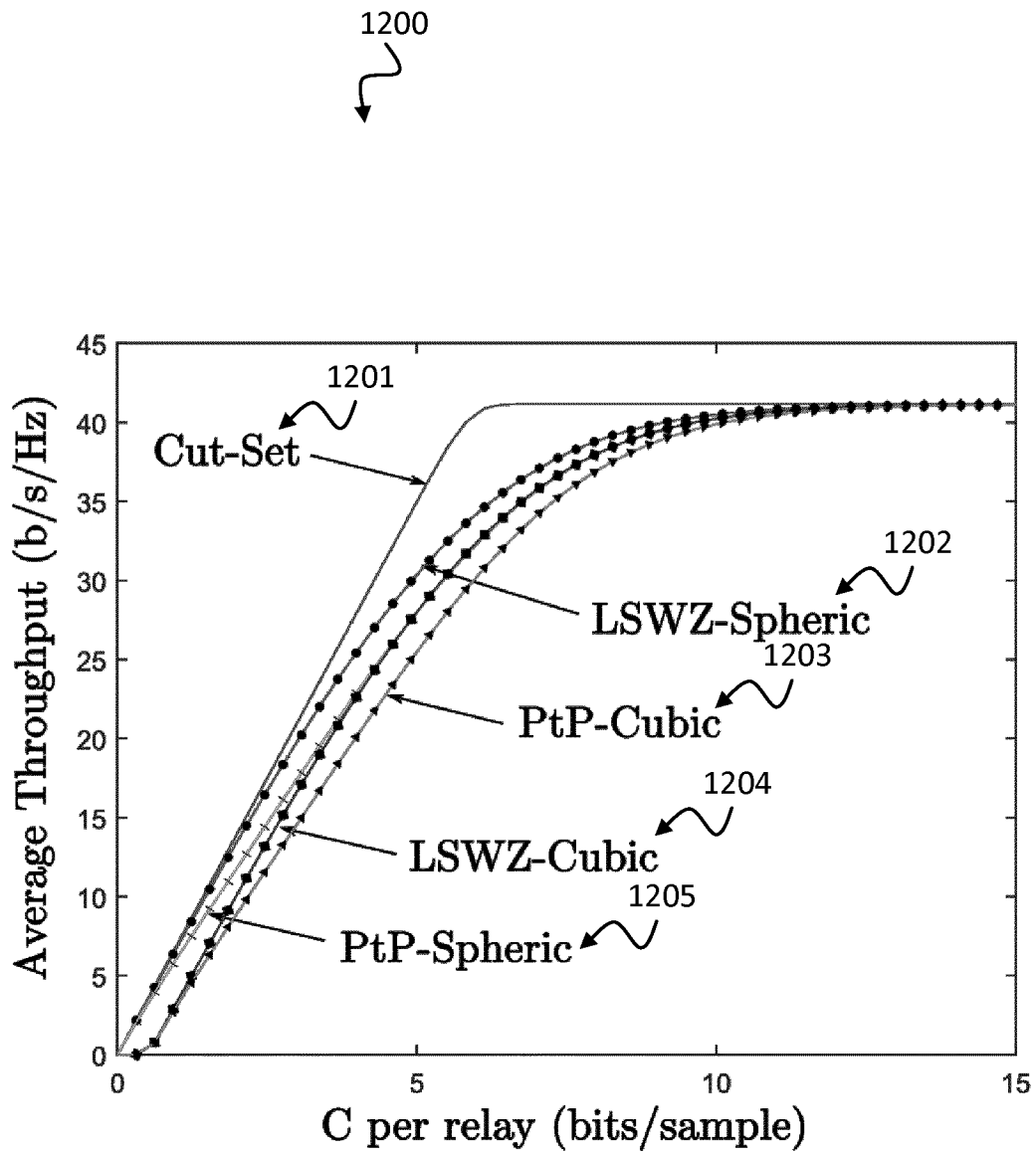


Fig. 12

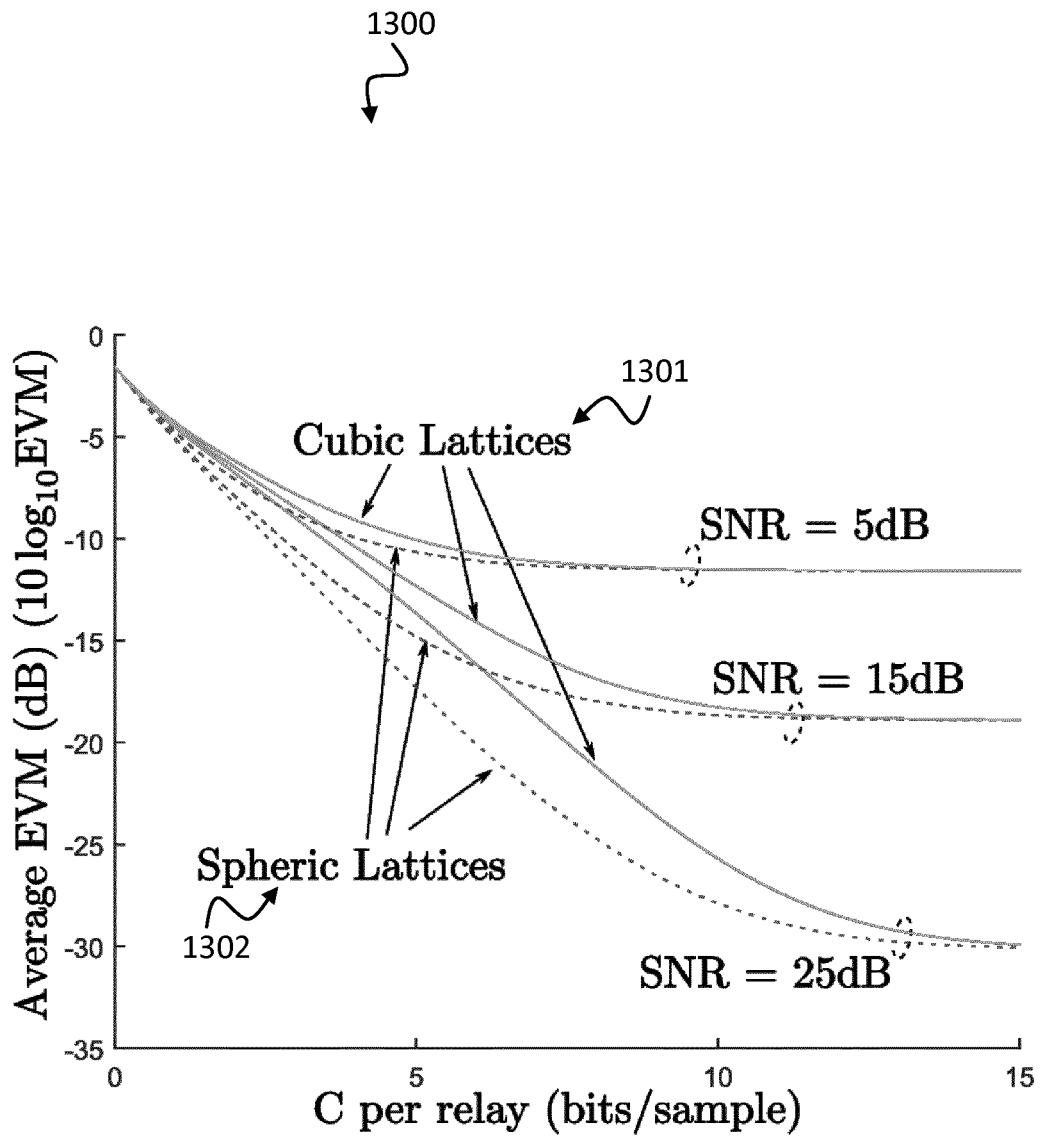


Fig. 13

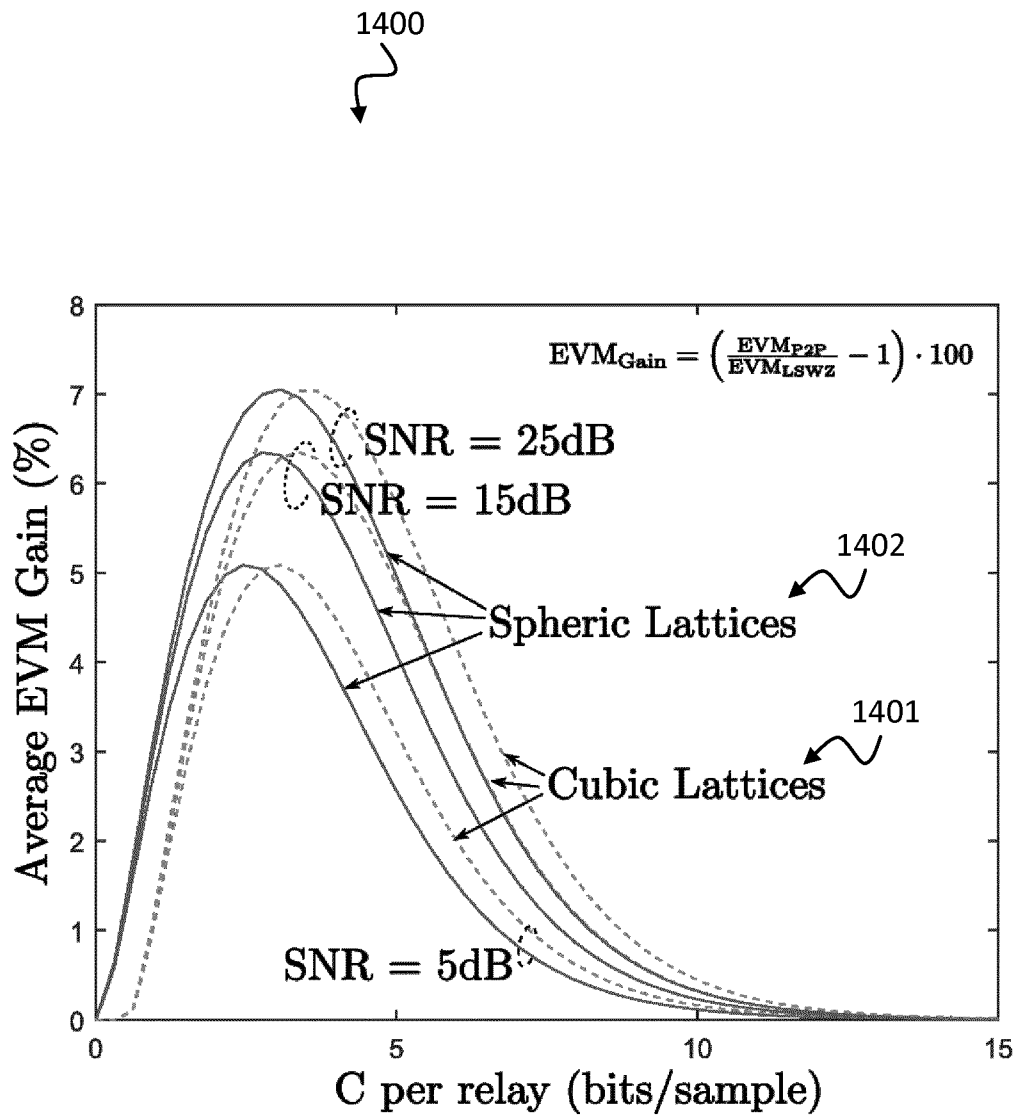


Fig. 14

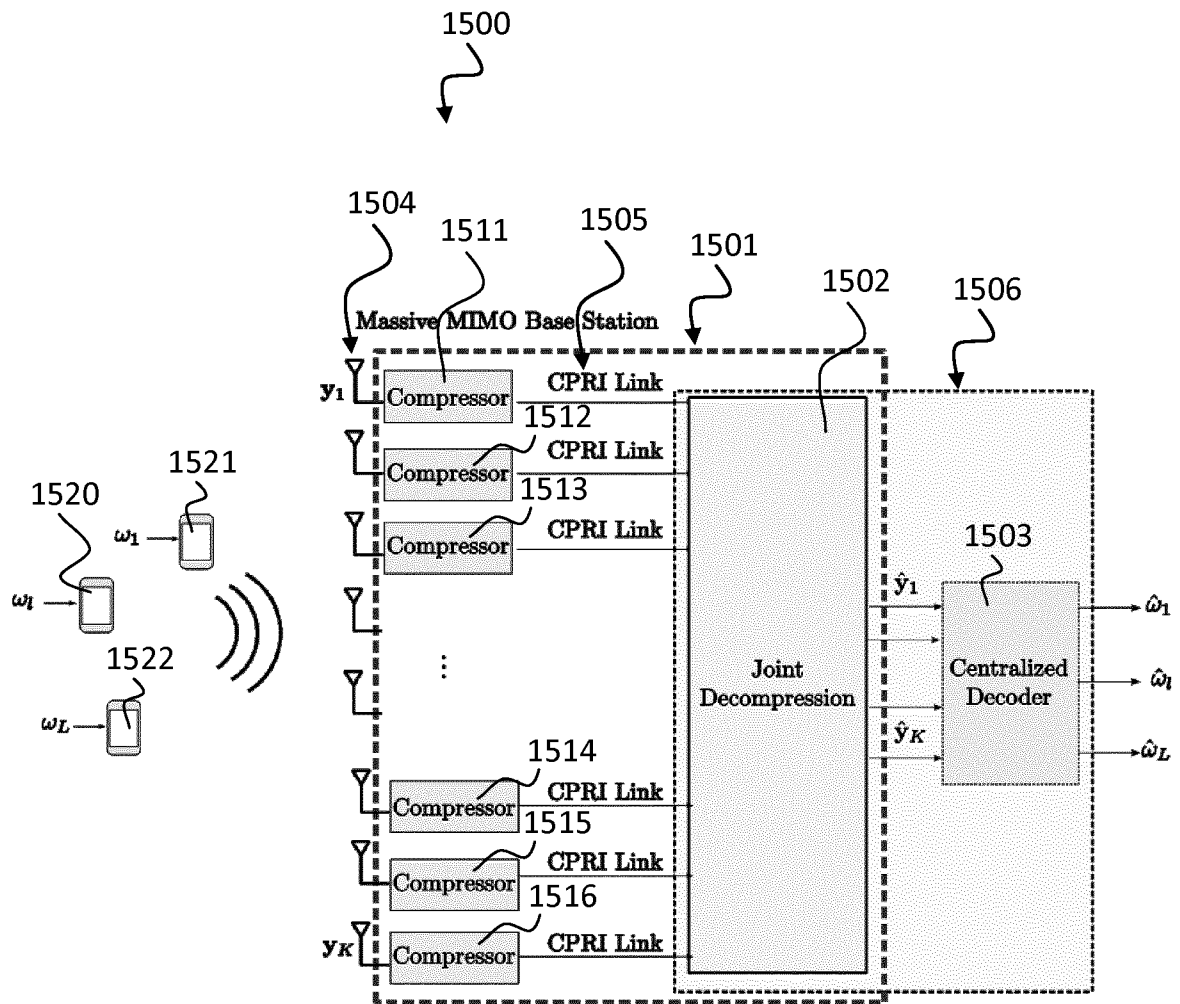


Fig. 15

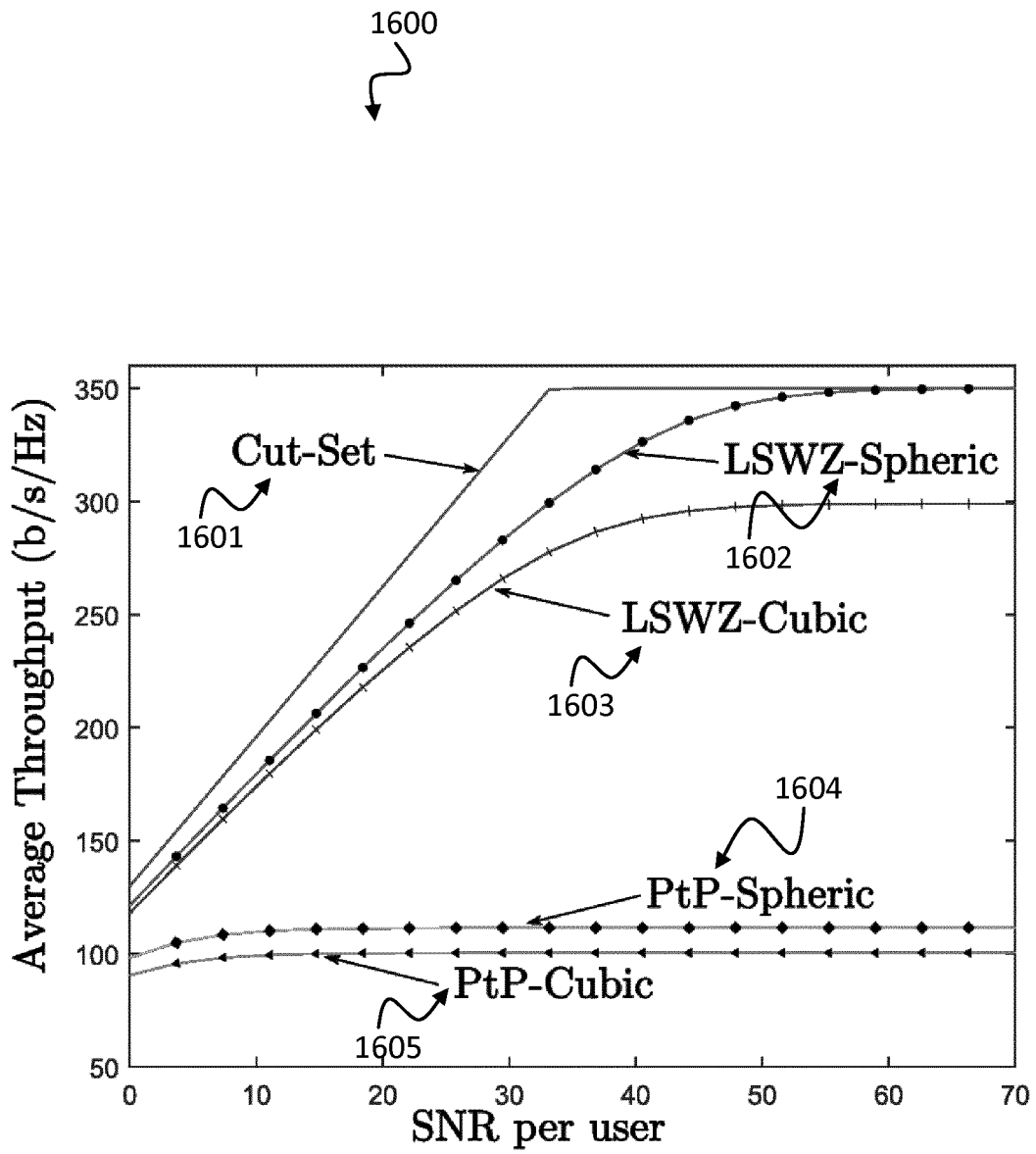


Fig. 16

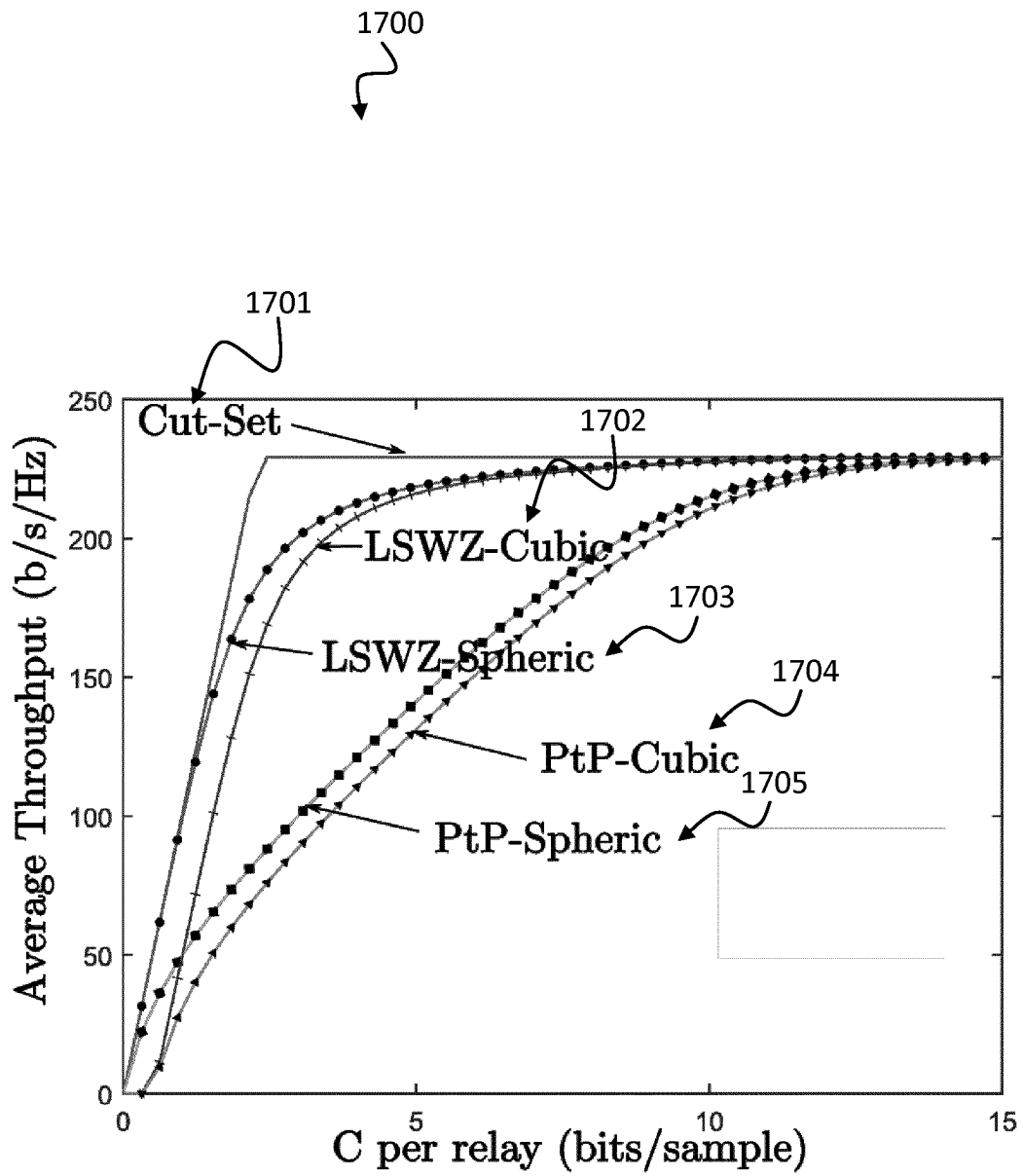


Fig. 17

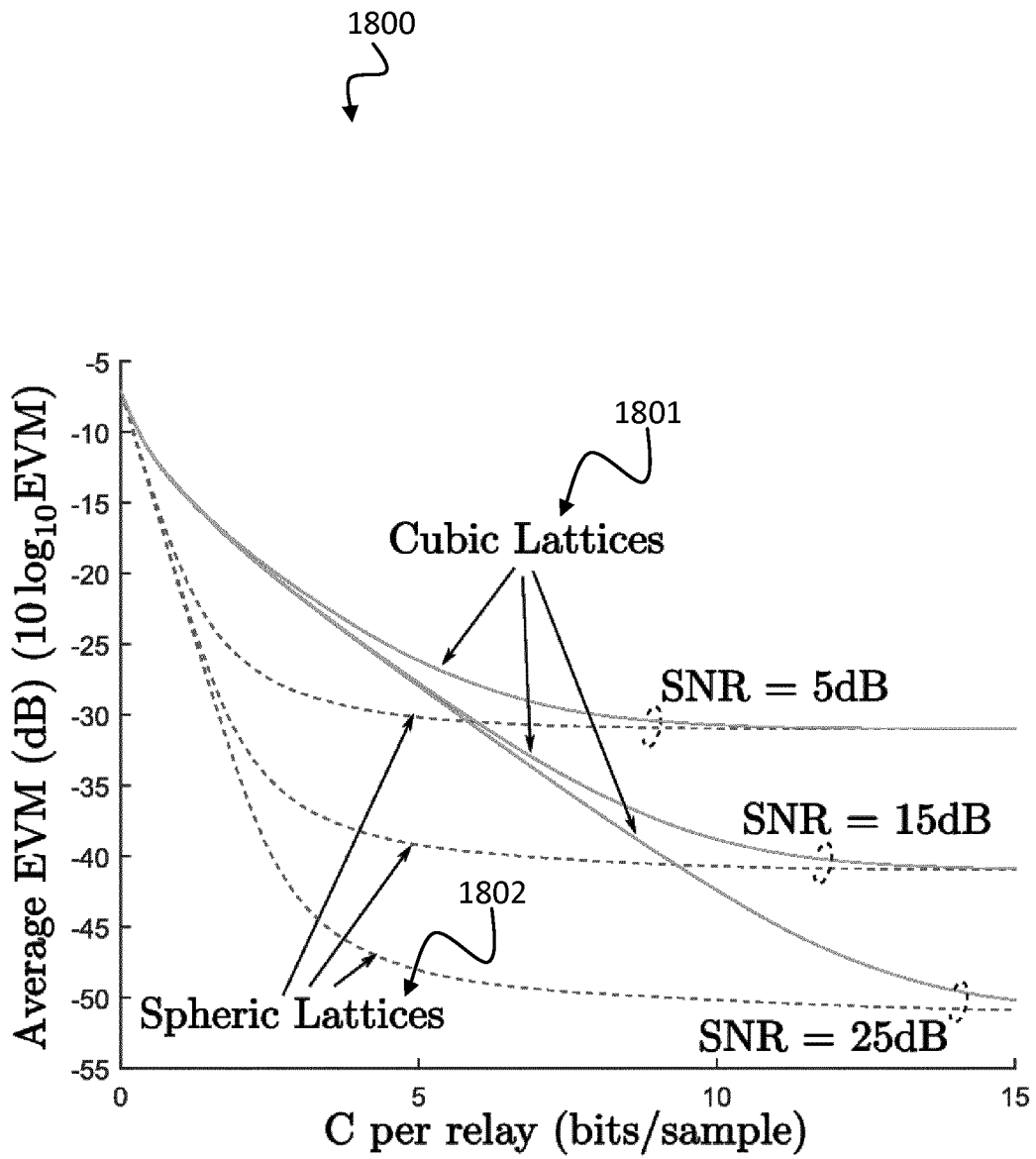


Fig. 18

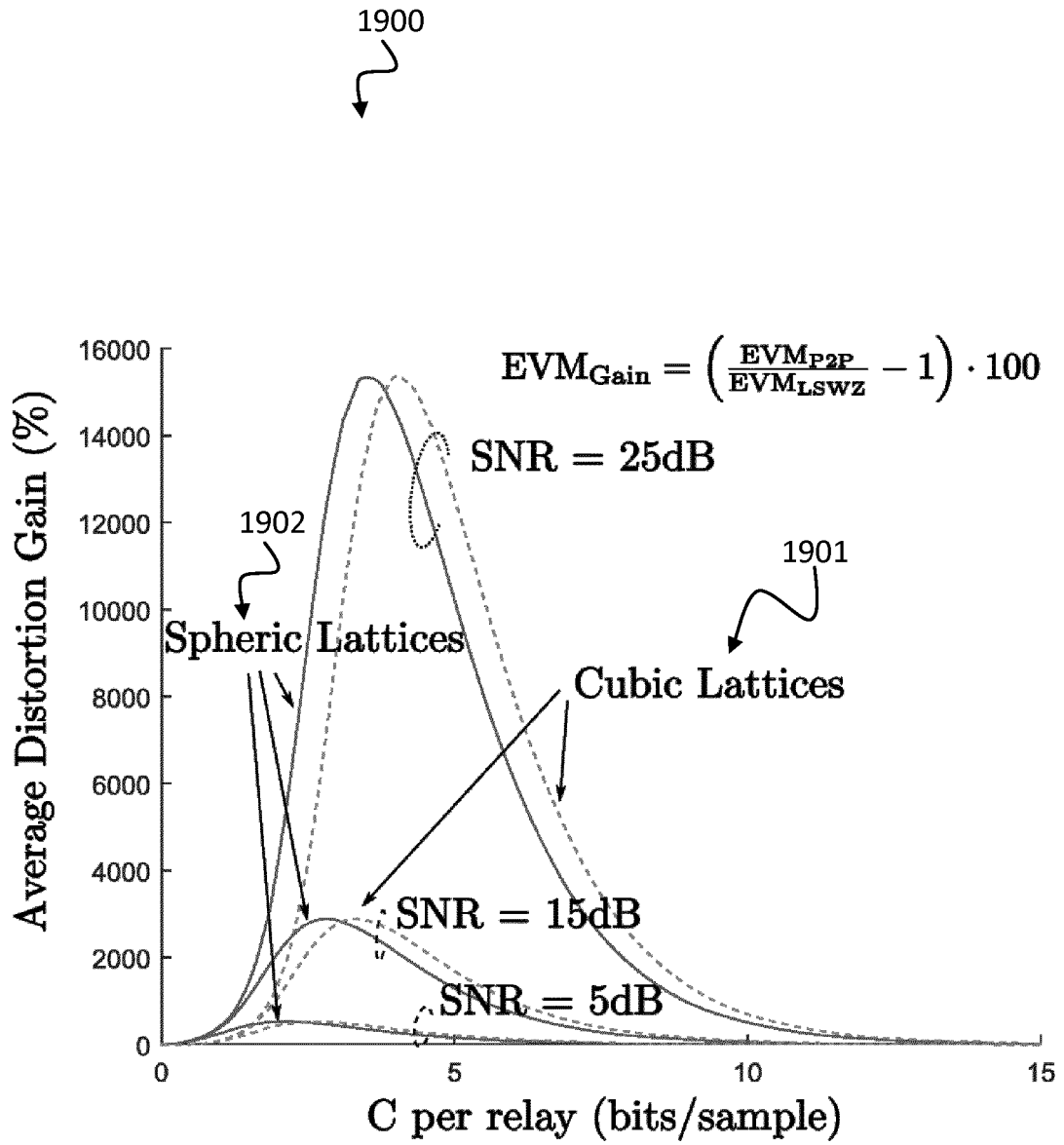


Fig. 19

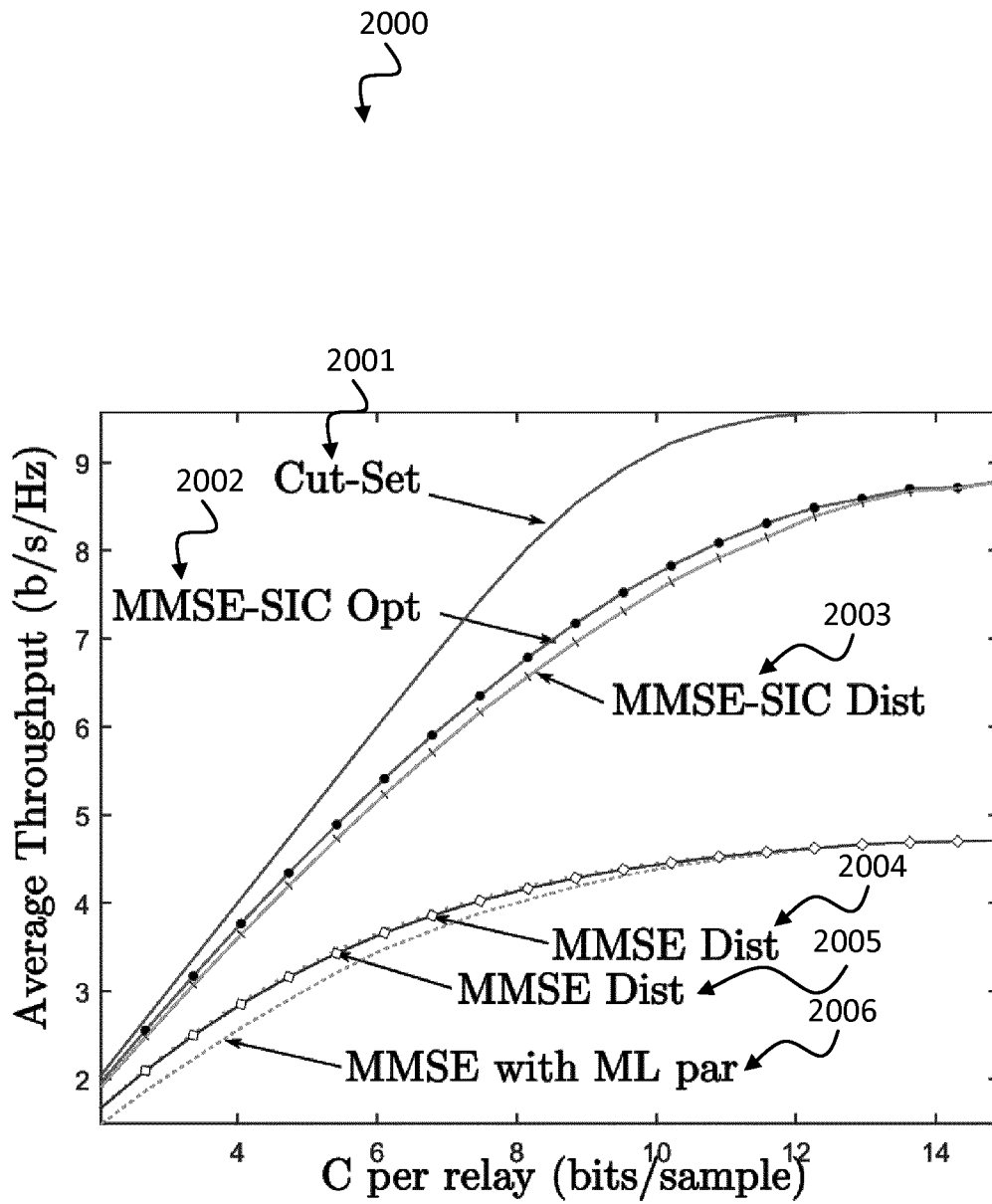


Fig. 20

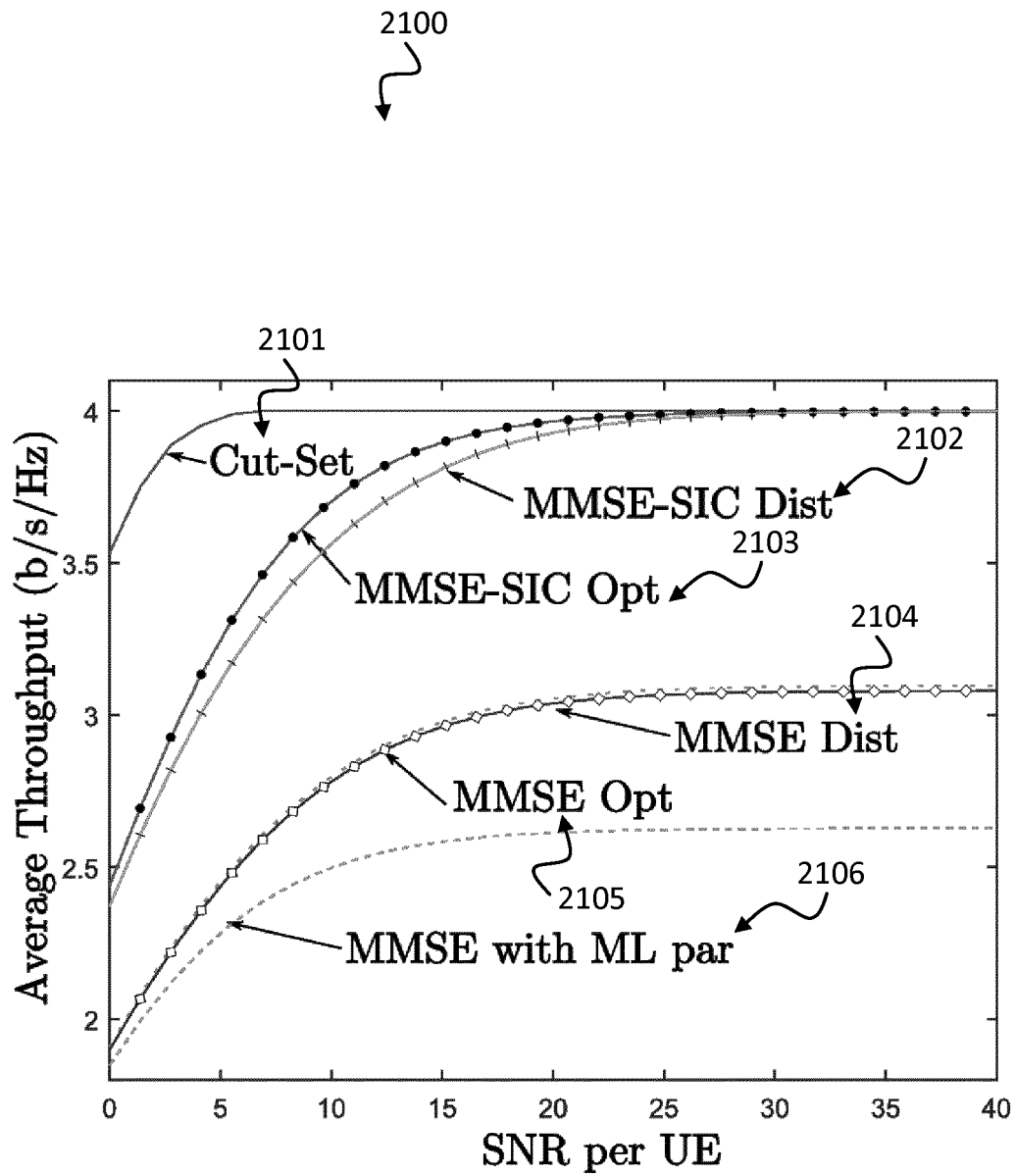


Fig. 21

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/057039

A. CLASSIFICATION OF SUBJECT MATTER
 INV. H04L25/03 H04B7/155
 ADD.
 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)
 H04L H04B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
 EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AGUERRI INAKI ESTELLA ET AL: "Partial compute-compress-and-forward for limited backhaul uplink multicell processing", 2015 53RD ANNUAL ALLERTON CONFERENCE ON COMMUNICATION, CONTROL, AND COMPUTING (ALLERTON), IEEE, 29 September 2015 (2015-09-29), pages 1264-1271, XP032889809, DOI: 10.1109/ALLERTON.2015.7447153 [retrieved on 2016-04-04] the whole document -----	1-15

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 21 November 2016	Date of mailing of the international search report 29/11/2016
--	---

Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Baltersee, Jens
--	--