Signal Representations with Minimum ℓ_∞-Norm

Christoph Studer, Wotao Yin, and Richard G. Baraniuk

Rice University, Houston, TX, USA
e-mail: {studer, wotao.yin, richb}@rice.edu

Abstract—Maximum (or ℓ_∞) norm minimization subject to an underdetermined system of linear equations finds use in a large number of practical applications, such as vector quantization, peak-to-average power ratio (PAPR) (or "crest factor") reduction in wireless communication systems, approximate neighbor search, robotics, and control. In this paper, we analyze the fundamental properties of signal representations with minimum ℓ_∞-norm. In particular, we develop bounds on the maximum magnitude of such representations using the uncertainty principle (UP) introduced by Lyubarskii and Vershynin, 2010, and we characterize the limits of ℓ_∞-norm-based PAPR reduction. Our results show that matrices satisfying the UP, such as randomly subsampled Fourier or i.i.d. Gaussian matrices, enable the efficient computation of so-called *democratic representations*, which have both provably small ℓ_∞-norm and low PAPR.

I. Introduction

In this paper, we analyze properties of the solution $\dot{\mathbf{x}} \in \mathbb{C}^N$ to the following (convex) ℓ_∞-norm minimization problem:

$$(\mathrm{P}_\infty^\varepsilon) \quad \underset{\tilde{\mathbf{x}} \in \mathbb{C}^N}{\text{minimize}} \ \|\tilde{\mathbf{x}}\|_\infty \quad \text{subject to} \ \|\mathbf{y} - \mathbf{D}\tilde{\mathbf{x}}\|_2 \leq \varepsilon.$$

Here, the vector $\mathbf{y} \in \mathbb{C}^M$ denotes the *signal* to be represented, $\mathbf{D} \in \mathbb{C}^{M \times N}$ corresponds to an overcomplete matrix (dictionary) with $M < N$, and the real-valued parameter $\varepsilon \geq 0$ controls the accuracy of the *signal representation* $\dot{\mathbf{x}}$.

A. Application Examples

The ℓ_∞-norm minimization problem $(\mathrm{P}_\infty^\varepsilon)$ features prominently in a variety of practical applications, including:

- *Vector quantization:* As shown in [1], certain matrices $\mathbf{D}$ enable one to compute representations $\dot{\mathbf{x}}$ whose magnitudes are roughly of the order $1/\sqrt{N}$. For such signal representations, each entry is of approximately the same importance. Hence, element-wise quantization affects all entries of $\dot{\mathbf{x}}$ equally, which renders them less susceptible to quantization noise compared to a direct quantization of the signal vector $\mathbf{y}$. Moreover, the corruption of a few entries of $\dot{\mathbf{x}}$ results in only a small error and, hence,

computing $\mathbf{y} = \mathbf{D}\dot{\mathbf{x}}$ after, e.g., unreliable transmission, provides a robust estimate of the signal vector $\mathbf{y}$ [1].

- *Peak-to-average power ratio (PAPR) reduction:* Wireless communication systems employing orthogonal frequency division multiplexing (OFDM) require linear and power-inefficient radio-frequency (RF) components (e.g., power amplifiers), as OFDM signals typically exhibit a large dynamic range [2]. By allocating certain unused OFDM tones, known as tone reservation [3], or by exploiting the excess of degrees-of-freedom in large-scale multi-antenna wireless systems [4], one can transmit signal representations with small ℓ_∞-norm to reduce the PAPR (often also called "crest factor"). Transmitting such representations in OFDM-based communication systems can substantially alleviate the need for expensive and power-inefficient RF components.

- *Approximate nearest neighbor search:* Signal representations obtained from $(\mathrm{P}_\infty^\varepsilon)$ also find use in identifying approximate nearest neighbors in high-dimensional spaces [5]. The underlying idea is to compute a representation $\dot{\mathbf{x}}$ for the query vector $\mathbf{y}$, which resembles to an antipodal signal with most coefficients corresponding to $\{-\alpha, +\alpha\}$, for some $\alpha > 0$. This property of the magnitudes of $\dot{\mathbf{x}}$ can then be used to efficiently find approximate nearest vectors in a N-dimensional Hamming space.

- *Robotics and control:* Kinematically redundant robots or manipulators admit an infinite number of inverse solutions. Certain applications require a solution that minimizes the *maximum* force, acceleration, torque, or joint velocity, for example, rather than minimizing the energy or power. In such situations, one is typically interested in solving problems of the form $(\mathrm{P}_\infty^\varepsilon)$ rather than solving minimum-energy problems; corresponding applications have been described in, e.g., [6]–[8].

Note that the problem $(\mathrm{P}_\infty^\varepsilon)$ with $\varepsilon = 0$ can also be used to recover antipodal solutions, i.e., vectors with coefficients belonging to $\{-\alpha, +\alpha\}$, from an underdetermined system of linear equations $\mathbf{y} = \mathbf{D}\mathbf{x}$ provided that certain conditions on the matrix $\mathbf{D}$ are met (see, e.g., [9], [10]). In this paper, however, we focus on the computation of *signal representations* having minimal ℓ_∞-norm and small dynamic range, rather than on the *recovery* a given vector $\mathbf{x}$ from $\mathbf{y} = \mathbf{D}\mathbf{x}$.

B. Relevant Prior Art

Initial results for minimizing the maximum amplitude of continuos, real-valued signals subject to linear constraints

reach back to the 1960s where Neustadt [11] studied the so-called *minimum-effort control problem*. In 1971, Cadzow proposed a practicable algorithm for the minimum-effort problem [6], where he proposed to solve the following (convex) ℓ_∞-norm minimization problem:

$$(\text{P}_\infty) \quad \underset{\tilde{\mathbf{x}} \in \mathbb{R}^N}{\text{minimize}} \, \|\tilde{\mathbf{x}}\|_\infty \quad \text{subject to } \mathbf{y} = \mathbf{D}\tilde{\mathbf{x}},$$

which coincides with $(\text{P}_\infty^\varepsilon)$ for $\varepsilon = 0$. Specifically, it was shown in [6] that for most matrices $\mathbf{D}$, a dominant portion of the magnitudes of the solution $\dot{\mathbf{x}}$ to (P_∞) correspond to $\|\dot{\mathbf{x}}\|_\infty$, whereas only a small fraction of the entries have smaller magnitude; this result has been rediscovered recently [12].

Another line of results that characterize signal representations $\dot{\mathbf{x}}$ with small (but not necessarily minimal) ℓ_∞-norm subject to $\mathbf{y} = \mathbf{D}\dot{\mathbf{x}}$ have been developed in 2010 by Lyubarskii and Vershynin [1]. In particular, [1] proves the existence of matrices $\mathbf{D}$ with arbitrarily small redundancy parameter $\lambda = N/M > 1$ for which every signal vector $\mathbf{y}$ has a representation $\dot{\mathbf{x}}$ satisfying

$$\|\dot{\mathbf{x}}\|_\infty \leq \frac{K}{\sqrt{N}} \|\mathbf{y}\|_2. \tag{1}$$

Here, K is a (preferably small) constant that only depends on the redundancy parameter λ. The existence of such signal representations and matrices can either be shown via results obtained by Kashin [13], Garnaev and Gluskin [14], or by analyzing the outcome of the iterative algorithm proposed in [1], which computes signal representations with small ℓ_∞-norm. The latter (constructive) approach relies on an uncertainty principle (UP) for the matrix $\mathbf{D}$, which establishes a fundamental connection between sensing matrices commonly used in the field of compressive sensing (CS) and sparse signal recovery [15]–[18] and the constant K in (1).

C. Contributions

In this paper, we analyze the fundamental properties of signal representations $\dot{\mathbf{x}}$ obtained from the ℓ_∞-norm minimization problem $(\text{P}_\infty^\varepsilon)$. In particular, we analyze its Lagrange dual problem to derive an improved and more general version of the bound on the ℓ_∞-norm of $\dot{\mathbf{x}}$ proposed in [1]. We furthermore analyze the fundamental PAPR properties of signal representations obtained through $(\text{P}_\infty^\varepsilon)$, which is of particular interest in OFDM-based wireless communication systems. As a byproduct of our analysis, we present the Lagrange duals to a variety of optimization problems, such as ℓ_1-norm minimization frequently used for sparse signal recovery. In addition, we summarize classes of matrices that enable the efficient computation of so-called *democratic representations*, which have both small ℓ_∞-norm and low PAPR. Finally, we show numerical results to support our analysis.

D. Notation

Lowercase boldface letters stand for column vectors and uppercase boldface letters designate matrices. For a matrix $\mathbf{A}$, we denote its conjugate transpose and spectral norm by $\mathbf{A}^H$ and $\|\mathbf{A}\|_{2,2} = \sqrt{\lambda_{\max}(\mathbf{A}^H\mathbf{A})}$, respectively, where $\lambda_{\max}(\mathbf{A}^H\mathbf{A})$

denotes the maximum eigenvalue of $\mathbf{A}^H\mathbf{A}$. The kth entry of a vector $\mathbf{a}$ is designated by $[\mathbf{a}]_k$, and $\Re\{\mathbf{a}\}$ and $\Im\{\mathbf{a}\}$ represent its real and imaginary part. We define the ℓ_p-norm of the vector $\mathbf{a} \in \mathbb{C}^N$ as

$$\|\mathbf{a}\|_p = \begin{cases} \left(\sum_{k=1}^N |[\mathbf{a}]_k|^p \right)^{1/p} & \text{if } 1 \leq p < \infty \\ \max_{k \in \{1,\ldots,N\}} |[\mathbf{a}]_k| & \text{if } p = \infty. \end{cases}$$

Sets are designated by uppercase Greek letters; the cardinality of the set Ω is $|\Omega|$. For $x \in \mathbb{R}$, the operation $\lceil x \rceil$ rounds x to the nearest integer towards infinity.

E. Organization of the Paper

The remainder of the paper is organized as follows. Section II introduces the necessary definitions and our main results are detailed in Section III. Section IV summarizes suitable classes of matrices that satisfy the UP and simulation results are provided in Section V. We conclude in Section VI.

II. SIGNAL REPRESENTATIONS WITH MINIMUM ℓ_∞-NORM

We next introduce the necessary prerequisites and formulate the Lagrange dual problem to $(\text{P}_\infty^\varepsilon)$ which is key in the analysis shown in Section III.

A. Frames

In the remainder of the paper, we often require the overcomplete matrix $\mathbf{D}$ to satisfy the following definition [19].

Definition 1 (Frame): A matrix $\mathbf{D} \in \mathbb{C}^{M \times N}$ with $M \leq N$ is called a *frame* if

$$A\|\mathbf{w}\|_2^2 \leq \|\mathbf{D}^H\mathbf{w}\|_2^2 \leq B\|\mathbf{w}\|_2^2$$

holds for any vector $\mathbf{w} \in \mathbb{C}^M$ with $A \in \mathbb{R}$, $B \in \mathbb{R}$, and $0 < A \leq B < \infty$. The tightest possible constants A and B are called lower and upper *frame bound*, respectively.

In what follows, we refer to $\mathbf{D}$ as a *tight frame* if $A = B$; furthermore, if $A = B = 1$, then $\mathbf{D}$ is called a *Parseval frame*.

B. Democratic Representations

For certain frames, all signal representations $\dot{\mathbf{x}}$ are guaranteed to exhibit certain properties. The class of representations studied in the remainder of the paper is defined next.

Definition 2 (Democratic Representation): Let $\mathbf{D} \in \mathbb{C}^{M \times N}$ be a given frame. If there exists for each signal vector $\mathbf{y} \in \mathbb{C}^M$ a representation $\dot{\mathbf{x}} \in \mathbb{C}^N$ satisfying i) $\|\mathbf{y} - \mathbf{D}\dot{\mathbf{x}}\|_2 \leq \varepsilon$ for a given $\varepsilon \geq 0$, and ii) the following inequalities:

$$\frac{K_1}{\sqrt{N}} (\|\mathbf{y}\|_2 - \varepsilon) \leq \|\dot{\mathbf{x}}\|_\infty \leq \frac{K_{\text{u}}}{\sqrt{N}} (\|\mathbf{y}\|_2 - \varepsilon), \tag{2}$$

with $0 < K_1 \leq K_{\text{u}} < \infty$, then the frame $\mathbf{D}$ enables the computation of *democratic representations* $\dot{\mathbf{x}}$. The constants K_1, $K_{\text{u}} \in \mathbb{R}^+$ are called the lower and upper *Kashin bounds*, which depend only on $\mathbf{D}$ but *not* on the signal vector $\mathbf{y}$ or the democratic representation $\dot{\mathbf{x}}$.

Definition 2 states that for certain frames, one can *always* (i.e., for each signal vector $\mathbf{y}$) find representations $\dot{\mathbf{x}}$ for which the maximum absolute entry is bounded from below and above by the Kashin bounds, the ℓ_2-norm of the signal vector $\mathbf{y}$, and

the accuracy parameter ε. Note that in the case $\varepsilon = 0$, the upper Kashin bound K_u in (2) corresponds K in (1). Hence, the Definition 2 enables us to analyze a generalized setting of the special case studied in [1].

From a practical viewpoint, frames that enable the computation of democratic representations, have, for example, bounded maximum force in kinematically redundant systems. In addition, as it will be shown in Section III-C, the upper Kashin bound K_u can also be used to characterize the PAPR of democratic representations; this property finds use in OFDM-based wireless communication systems, for example.

The key goal in applications relying on ℓ_∞-norm minimization is to find algorithms and frames $\mathbf{D}$ for which one can efficiently compute democratic representations. As it was shown in [1] and is summarized in Section IV, there exist (properly normalized) frames for which K_1 and K_u are close, and where the upper Kashin bound K_u only depends on the redundancy parameter $\lambda = N/M$.

C. Computing Representations with Minimum ℓ_∞-Norm

In order to compute representations $\dot{\mathbf{x}}$ having small (but necessarily minimal) ℓ_∞-norm subject to $\mathbf{y} = \mathbf{D}\dot{\mathbf{x}}$, one can use the iterative algorithm proposed in [1]; this method efficiently computes such representations for real-valued and approximate Parseval frames, i.e., frames $\mathbf{D} \in \mathbb{R}^{M \times N}$ satisfying the UP in [1] with bounds satisfying $A = 1 - \xi$ and $B = 1 + \xi$ for some small $\xi > 0$. The algorithm in [1], however, i) does not solve (P_∞^ε) and is, in general, not guaranteed to find representations having the *smallest* ℓ_∞-norm, ii) was introduced for real-valued systems only, and iii) is only guaranteed to converge for approximate Parseval frames. Moreover, if one is interested in representations $\dot{\mathbf{x}}$ for which $\|\mathbf{y} - \mathbf{D}\dot{\mathbf{x}}\|_2 > 0$ rather than in perfect representations satisfying $\mathbf{y} = \mathbf{D}\dot{\mathbf{x}}$, the algorithm in [1] must be modified accordingly.

In order to overcome the limitations of the iterative algorithm in [1], we propose to directly solve the convex optimization problem (P_∞^ε) instead. Solving this problem enables one to efficiently compute *democratic representations* with *minimum ℓ_∞-norm* for general (i.e., not necessarily tight or Parseval) frames satisfying the UP in [1].[1] Concretely, to efficiently compute the solution to (P_∞^ε), general-purpose solvers for convex optimization problems can be used (see, e.g., [20], [21]). For large-dimensional problems, an efficient first-order method, referred to as fast iterative truncation algorithm (FITRA), was proposed in [4].

D. Lagrange Dual Problem

In order to analyze the properties of signal representations with minimum ℓ_∞-norm obtained from (P_∞^ε), i.e., to derive bounds on the lower and upper Kashin bounds K_l and K_u, as well as to study the associated PAPR properties, our results presented in Section III make use of the following lemma.

[1]Note in the case $\varepsilon \geq \|\mathbf{y}\|_2$, the problem ($P_\infty^\varepsilon$) returns the all-zeros vector and, hence, practically relevant choices of ε are in the range $0 \leq \varepsilon < \|\mathbf{y}\|_2$.

Lemma 1 (Lagrange Dual Problem): Let the ℓ_p-norm primal problem (with $1 \leq p \leq \infty$) be

$$(P_p^\varepsilon) \quad \underset{\tilde{\mathbf{x}} \in \mathbb{C}^N}{\text{minimize}} \; \|\tilde{\mathbf{x}}\|_p \quad \text{subject to} \; \|\mathbf{y} - \mathbf{D}\tilde{\mathbf{x}}\|_2 \leq \varepsilon.$$

Then, the corresponding Lagrange dual problem is given by

$$(D_p^\varepsilon) \quad \begin{cases} \underset{\tilde{\mathbf{z}} \in \mathbb{C}^M}{\text{maximize}} & \Re(\mathbf{y}^H \tilde{\mathbf{z}}) - \varepsilon\|\tilde{\mathbf{z}}\|_2 \\ \text{subject to} & \|\mathbf{D}^H \tilde{\mathbf{z}}\|_d \leq 1 \end{cases}$$

with $1/p + 1/d = 1$; for $p = 1$ we have $q = \infty$ and vice versa. The norm $\|\cdot\|_d$ corresponds to the dual norm of $\|\cdot\|_p$.

Proof: The proof can be found in Appendix A. ∎

Lemma 1 covers not only the Lagrange dual to (P_∞^ε), but also other frequently studied optimization problems. In particular, the Lagrangian dual to (P_1^ε), which is frequently used for CS and sparse signal recovery [15]–[18], is contained in Lemma 1 as a special case. We finally note that the dual problem to (P_∞^ε) with $\varepsilon = 0$ and for real-valued $\mathbf{y}$ and $\mathbf{D}$ was used previously in [6] to characterize the solution $\dot{\mathbf{x}}$ to (P_∞).

III. MAIN RESULTS

We next analyze the fundamental properties of signal representations $\dot{\mathbf{x}}$ obtained from (P_∞^ε). In particular, we start by developing lower and upper Kashin bounds K_l and K_u, and then, we study the associated PAPR behavior.

A. Lower Kashin Bound K_l

We start by the following result, which provides a lower Kashin bound for general (i.e., not necessarily tight) frames.

Lemma 2 (Lower Kashin Bound): Let $\mathbf{D} \in \mathbb{C}^{M \times N}$ be a frame with upper frame bound B. Then, any vector $\mathbf{y} \in \mathbb{C}^M$ admits a signal representation $\dot{\mathbf{x}}$ with lower Kashin bound

$$K_l = 1/\sqrt{B}. \tag{3}$$

Proof: The proof can be found in Appendix B. ∎

We emphasize that representations obtained from (P_∞^ε) are guaranteed to satisfy (3), as the proof for Lemma 2 exploits properties of its solution $\dot{\mathbf{x}}$. It is furthermore interesting to see that the lower Kashin bound in (3) only depends on the upper frame bound B (and on the condition that frames satisfy $A > 0$); this is in stark contrast to the upper Kashin bound derived in Section III-B. We finally note that in the case $\varepsilon = 0$ and for Parseval frames, solving (P_∞^ε) ensures that any vector $\mathbf{y}$ admits a signal representation satisfying

$$\frac{\|\mathbf{y}\|_2}{\sqrt{N}} \leq \|\dot{\mathbf{x}}\|_\infty .$$

This special case of (3) was shown previously in [1, Obs. 2.1b].

B. Upper Kashin Bound K_u

1) Uncertainty principle: In order to analytically characterize the upper Kashin bound K_u, we build our results on the uncertainty principle (UP) for frames proposed in [1, Def. 3.4].

Definition 3 (Uncertainty Principle for Frames): A frame $\mathbf{D} \in \mathbb{C}^{M \times N}$ satisfies the UP with parameters η, δ if

$$\|\mathbf{D}\mathbf{x}\|_2 \leq \eta\|\mathbf{x}\|_2$$

holds for $\eta \in \mathbb{R}^+$, $\delta \in (0, 1)$, and for all (sparse) vectors $\mathbf{x} \in \mathbb{C}^N$ satisfying $|\text{supp}(\mathbf{x})| \leq \delta N$.

Verifying the UP for a given frame $\mathbf{D}$ requires, in general, a combinatorial search over all δN-sparse vectors [22]. Hence, in [1] it was shown that certain randomly constructed frames satisfy the UP with high probability. Note that such frames are strongly related to sensing matrices with small restricted isometry constants, which play a central role in CS (see, e.g., [15]–[17]). In Section IV, we briefly summarize suitable classes of frames which satisfy the UP with high probability.

2) Upper Kashin bound: The following theorem establishes an upper Kashin bound K_u, which provides sufficient conditions for which a frame $\mathbf{D}$ enables the computation of democratic representations $\dot{\mathbf{x}}$ for any signal vector $\mathbf{y}$.

Theorem 3 (Upper Kashin Bound): Let $\mathbf{D} \in \mathbb{C}^{M \times N}$ be a frame with frame bounds A, B that satisfies the uncertainty principle (UP) with parameters η, δ. Then, any signal vector $\mathbf{y}$ admits a democratic representation $\dot{\mathbf{x}}$ with upper Kashin bound

$$K_\text{u} = \frac{\eta}{(A - \eta\sqrt{B})\sqrt{\delta}}, \qquad (4)$$

provided $A > \eta\sqrt{B}$. Moreover, democratic representations satisfying both (3) and (4) can be computed using $(\text{P}_\infty^\varepsilon)$.

Proof: The proof can be found in Appendix C. ■

This result shows that if a frame $\mathbf{D}$ satisfies i) $A > \eta\sqrt{B}$ and ii) $\delta > 0$, then one can compute democratic representations for any signal vector $\mathbf{y}$ by solving $(\text{P}_\infty^\varepsilon)$. In addition, the condition $A > \eta\sqrt{B}$ indicates that the use of Parseval frames is beneficial in practice, i.e., leads to representations with smaller ℓ_∞-norm—an observation that was made empirically by Fuchs [23]. Corresponding simulation results are shown in Section V. In order to achieve representations having small ℓ_∞-norm, one is therefore interested in finding frames satisfying the UP with small η and large δ. Both properties can be achieved simultaneously for certain classes of frames (see Section IV for examples).

We emphasize that Theorem 3 improves upon the results in [1], i.e., K_u in (4) is strictly smaller than the Kashin constant obtained in [1, Thms. 3.5 and 3.9]. To see this, consider the case of $\mathbf{D}$ being a Parseval frame and $\varepsilon = 0$, which enables us to establish the following relation between the upper Kashin bound K_u in (4) and the bound K from [1, Thm. 3.5]:

$$K_\text{u} = \frac{\eta}{(1 - \eta)\sqrt{\delta}} < \frac{1}{(1 - \eta)\sqrt{\delta}} = K,$$

The strict inequality follows from the fact that η is required to be smaller than one, which is a consequence of $A > \eta\sqrt{B}$. Hence, by solving $(\text{P}_\infty^\varepsilon)$ rather than using the iterative algorithm proposed in [1], we arrive at an upper bound on the ℓ_∞-norm of the signal representation $\dot{\mathbf{x}}$ that is more tight (i.e., by a factor of η). For frames satisfying $A = 1 - \xi$ and $B = 1 + \xi$ with $0 \leq \xi < 1$ (so-called approximate Parseval frames), the upper Kashin bound in (4) continues to be superior than that in [1, Thm. 3.9]. Moreover, Theorem 3 also encompasses approximate representations, i.e., for which $\|\mathbf{y} - \mathbf{D}\dot{\mathbf{x}}\|_2 > 0$, and the case of complex-valued vectors and frames, which is

in contrast to the results developed in [1].

C. Peak-to-Average Power Ratio (PAPR)

The transmission of signals over frequency-selective channels typically requires sophisticated equalization schemes at the receive side. Orthogonal frequency-division multiplexing (OFDM) [2] is a well-established way of reducing the computational complexity of equalization (compared to conventional schemes). Unfortunaely, OFDM signals are known to suffer from a high PAPR, which requires linear RF components (e.g., mixers, power amplifiers, etc.) to avoid signal distortions and out-of-band radiation. Since linear RF components are, in general, more costly and less power efficient compared than their non-linear counterparts, practical implementations of OFDM usually employ sophisticated PAPR-reduction schemes [24].

1) Limits of ℓ_∞-norm-based PAPR reduction: Prominent approaches for reducing the PAPR exploit either certain reserved OFDM tones [3] or the massive amount of degrees-of-freedom in large-scale multi-antenna wireless systems [4]. For both methods, one can compute signal representations with small PAPR via $(\text{P}_\infty^\varepsilon)$. To this end, we next analyze the fundamental PAPR reduction capabilities of $(\text{P}_\infty^\varepsilon)$.

Definition 4 (Peak-to-Average Power Ratio): Let $\mathbf{x} \in \mathbb{C}^N$ be any nonzero vector. Then, the peak-to-average power ratio (PAPR) or "crest factor" of $\mathbf{x}$ is defined as

$$PAPR(\mathbf{x}) = \frac{\|\mathbf{x}\|_\infty^2}{\|\mathbf{x}\|_2^2 / N}. \qquad (5)$$

Note that for arbitrary vectors $\mathbf{x} \in \mathbb{C}^N$ the PAPR satisfies the following (trivial) inequalities:

$$1 \leq PAPR(\mathbf{x}) \leq N, \qquad (6)$$

which are a consequence of standard norm bounds. The lower bound is achieved for signals having constant modulus, whereas the upper bound is achieved by vectors which have a single nonzero entry. It is, however, important to realize that the PAPR of signal representations obtained through $(\text{P}_\infty^\varepsilon)$ is typically much smaller than the upper bound in (6) suggests. Concretely, the following result bounds the PAPR of signal representations obtained through $(\text{P}_\infty^\varepsilon)$ with the aid of the upper Kashin bound in (4).

Theorem 4 (PAPR Bound): Let $\mathbf{D} \in \mathbb{C}^{M \times N}$ be a frame with the upper Kashin bound K_u as defined in (4). Then, the PAPR of any signal representation $\dot{\mathbf{x}}$ for $\|\mathbf{y}\|_2 \neq 0$ and $\varepsilon < \|\mathbf{y}\|_2$ obtained by solving $(\text{P}_\infty^\varepsilon)$ satisfies

$$PAPR(\dot{\mathbf{x}}) \leq K_\text{u}^2 B. \qquad (7)$$

Proof: The proof can be found in Appendix D. ■

This theorem has immediate implications for practical systems solving $(\text{P}_\infty^\varepsilon)$. In particular, it shows that frames which satisfy the UP and have a small upper Kashin bound K_u are very effective in terms of reducing the PAPR. A practically relevant example of frames satisfying these properties are randomly subsampled discrete Fourier transform (DFT) matrices, which naturally appear in OFDM-based tone-reservation schemes for PAPR reduction (see, e.g., [3] for the details).

We note that a similar (but, in some cases, less tight) PAPR bound as the one in (7) was shown previously in [3] for randomly subsampled DFT matrices. The bound presented in Theorem 4, however, is valid for *general* frames $\mathbf{D}$ and also encompasses the case of approximate signal representations satisfying $\|\mathbf{y} - \mathbf{D}\dot{\mathbf{x}}\|_2 < \varepsilon$ with $\varepsilon \geq 0$.

IV. FRAMES FOR DEMOCRATIC REPRESENTATIONS

In [1] it was shown that random orthogonal matrices, random partial DFT matrices, and random sub-Gaussian matrices satisfy the UP in Definition 3 with high probability. Hence, matrices drawn from such classes are particularly suitable for the computation of *democratic representations* and for PAPR reduction. As an example, we briefly restate the result obtained in [1] for matrices whose entries are i.i.d. sub-Gaussian.

Definition 5 ([1, Def. 4.5]: Sub-Gaussian RV): A random variable X is called sub-Gaussian with parameter β if

$$\Pr\{|X| > u\} \leq \exp(1 - u^2/\beta^2) \quad \text{for all} \quad u > 0.$$

For matrices having i.i.d. sub-Gaussian entries, the following result has been established in [1].

Theorem 5 ([1, Thm. 4.6]: UP for sub-Gaussian Matrices): Let $\mathbf{A}$ be a $M \times N$ matrix whose entries are i.i.d. zero-mean sub-Gaussian RVs with parameter β. Assume that $\lambda = N/M$ for some $\lambda \geq 2$. Then, with probability at least $1 - \lambda^{-M}$, the random matrix $\mathbf{D} = \frac{1}{\sqrt{N}}\mathbf{A}$ satisfies the UP with parameters

$$\eta = C_0 \beta \sqrt{\frac{\log(\lambda)}{\lambda}} \quad \text{and} \quad \delta = \frac{C_1}{\lambda},$$

where $C_0, C_1 > 0$ are absolute constants.

Theorem 5 implies that for random sub-Gaussian matrices the UP with parameters η and δ is satisfied with high probability. Moreover, the UP parameters η, δ only depend on the redundancy $\lambda = N/M$ of $\mathbf{D}$. Since $\mathbf{D} = \frac{1}{\sqrt{N}}\mathbf{A}$ is, in general, not a tight frame, it was furthermore shown in [1, Cor. 4.9] that $\mathbf{D}$ is an approximate Parseval frame with high probability, i.e., $\mathbf{D}$ has frame bounds $A = 1 - \xi$ and $B = 1 + \xi$ for some small $\xi > 0$. Hence, random sub-Gaussian matrices can be used to efficiently compute democratic representations with Kashin bounds K_l and K_u in (3) and (4) by solving $(\mathrm{P}_\infty^\varepsilon)$. We emphasize that results similar to that of Theorem 5 have been established for random orthogonal and random partial DFT matrices. These classes of frames have the key advantage (over sub-Gaussian matrices) of being Parseval frames, which typically yield better Kashin bounds (see (4) and the next section for corresponding simulation results).

V. NUMERICAL RESULTS

To support our analytical results, we next simulate a lower bound on K_u in (4) and evaluate the PAPR behavior of solutions to $(\mathrm{P}_\infty^\varepsilon)$ for i.i.d. Gaussian and randomly subsampled discrete cosine transform (DCT) bases.

A. Behavior of the Upper Kashin Bound

In Fig. 1, we show empirical phase diagrams that characterize the upper Kashin bound K_u for i.i.d. Gaussian and randomly subsampled DCT matrices.
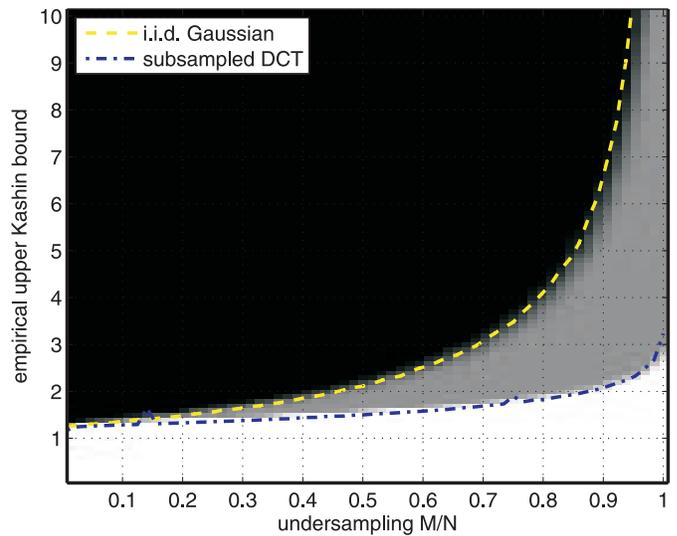

Fig. 1. Empirical phase diagram depending on the undersampling $\lambda = M/N$ for the upper Kashin bound K_u using $(\mathrm{P}_\infty^\varepsilon)$ with $\varepsilon = 0$ and for i.i.d. Gaussian and randomly subsampled DCT matrices (the lines highlight the individual 50% phase-transition boundaries).
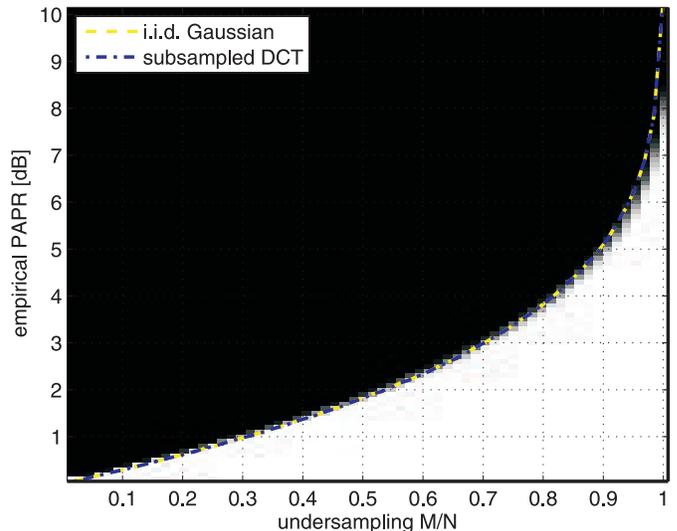

Fig. 2. Empirical phase diagram depending on the undersampling $\lambda = M/N$ for the PAPR using $(\mathrm{P}_\infty^\varepsilon)$ with $\varepsilon = 0$ and for i.i.d. Gaussian and randomly subsampled DCT matrices (the lines overlap for both cases and highlight the sharp 50% phase-transition boundary).

1) Simulation procedure: We fix $N = 512$ and vary M from 1 to 512. For each pair N, M, we perform 100 Monte-Carlo trials, and for each trial, we generate a frame $\mathbf{D}$, either with i.i.d. Gaussian entries or by randomly selecting rows from a DCT basis, and an i.i.d. Gaussian vector $\mathbf{y}$, followed by normalization to $\|\mathbf{y}\|_2 = 1$. We use CVX [21] to compute signal representations $\dot{\mathbf{x}}$ from $(\mathrm{P}_\infty^\varepsilon)$ with $\varepsilon = 0$ for each instance of $\mathbf{D}$ and $\mathbf{y}$. Finally, we compute a lower bound $\widetilde{K}_\mathrm{u}$ on the upper Kashin constant for each trial as

$$\widetilde{K}_\mathrm{u} = \frac{\sqrt{N}\|\dot{\mathbf{x}}\|_\infty}{\|\mathbf{y}\|_2 - \varepsilon} \leq K_\mathrm{u}. \tag{8}$$

We finally generate phase diagrams, which show the empirical probability for which $\widetilde{K}_{\mathrm{u}}$ is larger or smaller than a given value (i.e., the so-called empirical upper Kashin bound).

2) Discussion: The empirical phase diagram shown in Fig. 1 shows a sharp transition between the values of $\widetilde{K}_{\mathrm{u}}$ that have been realized (for a given undersampling ratio $\lambda = M/N$) and the values that were not achieved. Moreover, we see that the subsampled DCT has a smaller (empirical) upper Kashin bound than that of i.i.d. Gaussian matrices. This behavior was predicted by (4) and can mainly be addressed to the fact that subsampled DCT matrices are Parseval frames, whereas i.i.d. Gaussian matrices are, in general, not tight frames (see also Section IV). Hence, the use of Parseval frames for computing democratic representations turns out to be beneficial in practical applications.

B. Behavior of the PAPR

Fig. 2 illustrates the PAPR behavior of signal representations obtained by solving $(\mathrm{P}_\infty^\varepsilon)$.

1) Simulation procedure: We carry out a similar simulation procedure as detailed in Section V-A1, where we compute the *PAPR*$(\dot{\mathbf{x}})$ for each instance of $\mathbf{D}$ and $\mathbf{y}$ instead of $\widetilde{K}_{\mathrm{u}}$ in (8).

2) Discussion: The phase diagram shown in Fig. 2 exhibits a sharp transition between the (empirical) PAPR values achieved in this simulation and the values that were not achieved. It is interesting to see that both 50% phase transitions overlap, which is in stark contrast to the transition behavior of the upper Kashin bound. We, hence, conclude that the choice of the frame has only a small impact for PAPR-reduction (as long as it satisfies the UP).

VI. CONCLUSIONS

In this paper, we have analyzed the fundamental properties of signal representations with minimum ℓ_∞ (or maximum) norm. Specifically, we have developed bounds on such representations using the uncertainty principle (UP) proposed in [1], and we characterized their peak-to-average power (PAPR) properties, which is of particular interest for OFDM-based wireless communication systems. We furthermore demonstrated the existence of matrices for which so-called *democratic representations* with small ℓ_∞-norm and small PAPR exist. To support our analysis, we showed numerical simulation results, which highlight our finding that the use of Parseval frames leads to democratic representations with smaller ℓ_∞-norm compared to general frames.

There are many avenues for follow-on research. An analytical characterization of sharp phase transitions for $(\mathrm{P}_\infty^\varepsilon)$, e.g., as in [9], is interesting open research problem. The development of corresponding computationally efficient algorithms is part of on-going work.

APPENDIX A
PROOF OF LEMMA 1

Let $\|\mathbf{w}\|_p$ and $\|\mathbf{v}\|_d$ denote the primal and dual norm of the vectors $\mathbf{w}$ and $\mathbf{x}$ satisfying

$$\|\mathbf{w}\|_p = \max_{\mathbf{v}} \left\{ \Re\big(\mathbf{v}^H \mathbf{w}\big) : \|\mathbf{v}\|_p \leq 1 \right\}$$

with $1/p + 1/q = 1$. Then, for primal and dual norms, we have the following result [20]:

$$\min_{\mathbf{x}} \left\{ \|\mathbf{x}\|_p - \Re\big(\mathbf{z}^H \mathbf{D}\mathbf{x}\big) \right\} = \begin{cases} 0, & \|\mathbf{D}^H \mathbf{z}\|_d \leq 1 \\ -\infty, & \text{otherwise.} \end{cases} \quad (9)$$

We are now ready to derive the Lagrange dual problem $(\mathrm{D}_p^\varepsilon)$ to the primal problem $(\mathrm{P}_p^\varepsilon)$. To this end, we introduce an auxiliary vector $\mathbf{r} \in \mathbb{C}^M$ to rewrite $(\mathrm{P}_p^\varepsilon)$ as

$$\min_{\mathbf{x}} \left\{ \|\mathbf{x}\|_p : \|\mathbf{D}\mathbf{x} - \mathbf{y}\|_2 \leq \varepsilon \right\}$$
$$= \min_{\mathbf{x},\mathbf{r}} \left\{ \|\mathbf{x}\|_p : \mathbf{D}\mathbf{x} + \mathbf{r} = \mathbf{y}, \|\mathbf{r}\|_2 \leq \varepsilon \right\}.$$

By introducing the Lagrange dual variable $\mathbf{z} \in \mathbb{C}^M$, we obtain

$$\min_{\mathbf{x},\mathbf{r}} \left\{ \|\mathbf{x}\|_p : \mathbf{D}\mathbf{x} + \mathbf{r} = \mathbf{y}, \|\mathbf{r}\|_2 \leq \varepsilon \right\}$$
$$= \min_{\mathbf{x},\mathbf{r}} \max_{\mathbf{z}} \left\{ \|\mathbf{x}\|_p - \Re\big(\mathbf{z}^H(\mathbf{D}\mathbf{x}+\mathbf{r}-\mathbf{y})\big) : \|\mathbf{r}\|_2 \leq \varepsilon \right\}$$
$$= \max_{\mathbf{z}} \min_{\mathbf{x},\mathbf{r}} \left\{ \|\mathbf{x}\|_p - \Re\big(\mathbf{z}^H(\mathbf{D}\mathbf{x}+\mathbf{r}-\mathbf{y})\big) : \|\mathbf{r}\|_2 \leq \varepsilon \right\}. \quad (10)$$

For a given $\mathbf{z}$, the inner minimization problem of (10) is separable in the unknown vectors $\mathbf{x}$ and $\mathbf{r}$. The optimal auxiliary vector $\mathbf{r}$ is given by

$$\mathbf{r} = \begin{cases} \varepsilon \mathbf{z}/\|\mathbf{z}\|_2, & \mathbf{z} \neq \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1}, & \text{otherwise,} \end{cases}$$

and in either case, we have $\Re\big(\mathbf{z}^H \mathbf{r}\big) = \varepsilon\|\mathbf{z}\|_2$. Together with (9), we find that (10) is equal to

$$\max_{\mathbf{z}} \left\{ \Re\big(\mathbf{y}^H \mathbf{z}\big) - \varepsilon\|\mathbf{z}\|_2 : \|\mathbf{D}^H \mathbf{z}\|_d \leq 1 \right\},$$

which corresponds to the Lagrange dual problem $(\mathrm{D}_p^\varepsilon)$. Note that since in the derivation of $(\mathrm{D}_p^\varepsilon)$ all intermediate steps hold with equality, there is no duality gap.

APPENDIX B
PROOF OF LEMMA 2

The proof follows from a lower bound on the value of the dual problem $(\mathrm{D}_\infty^\varepsilon)$. Specifically, we have

$$\|\dot{\mathbf{x}}\|_\infty = \max_{\mathbf{z}} \left\{ \Re\big(\mathbf{y}^H \mathbf{z}\big) - \varepsilon\|\mathbf{z}\|_2 : \|\mathbf{D}^H \mathbf{z}\|_1 \leq 1 \right\}, \quad (11)$$

which we bound from below by replacing the optimal solution $\dot{\mathbf{z}}$ by the estimate

$$\hat{\mathbf{z}} = \frac{\mathbf{y}}{\|\mathbf{D}^H \mathbf{y}\|_1}, \quad (12)$$

which satisfies the constraint $\|\mathbf{D}^H \hat{\mathbf{z}}\|_1 \leq 1$. Hence, inserting (12) in the right-hand side (RHS) of (11) leads to the following lower bound:

$$\|\dot{\mathbf{x}}\|_\infty \geq \frac{\|\mathbf{y}\|_2^2 - \varepsilon\|\mathbf{y}\|_2}{\|\mathbf{D}^H \mathbf{y}\|_1}. \quad (13)$$

To further bound the RHS of (13) from below, we use standard norm bounds and the upper frame bound B of $\mathbf{D}$ to compute an upper bound to $\|\mathbf{D}^H \mathbf{y}\|_1$ as follows:

$$\|\mathbf{D}^H \mathbf{y}\|_1 \leq \sqrt{N}\|\mathbf{D}^H \mathbf{y}\|_2 \leq \sqrt{NB}\|\mathbf{y}\|_2. \quad (14)$$

Combining (14) with (13) finally yields

$$\|\dot{\mathbf{x}}\|_\infty \geq \frac{\|\mathbf{y}\|_2 - \varepsilon}{\sqrt{NB}}. \tag{15}$$

Note that in (12) we assumed that $\left\|\mathbf{D}^H\mathbf{y}\right\|_1 > 0$. Since $\mathbf{D}$ is a frame with lower frame bound $A > 0$, we have

$$\left\|\mathbf{D}^H\mathbf{y}\right\|_1 \geq \left\|\mathbf{D}^H\mathbf{y}\right\|_2 \geq \sqrt{A}\|\mathbf{y}\|_2 > 0,$$

which is satisfied whenever $\|\mathbf{y}\|_2 > 0$. In the case $\|\mathbf{y}\|_2 = 0$ the bound (15) continues to hold, which concludes the proof.

APPENDIX C
PROOF OF THEOREM 3

The proof proceeds in two stages. First, we separate the objective function of the Lagrange dual problem $(\mathrm{D}_\infty^\varepsilon)$ into two independent terms and then, we derive an upper bound on the ℓ_2-norm of the solution $\dot{\mathbf{z}}$ to $(\mathrm{D}_\infty^\varepsilon)$.

A. Separating the Result of the Lagrange Dual Problem

From the Lagrange dual problem $(\mathrm{D}_\infty^\varepsilon)$, we have

$$\|\dot{\mathbf{x}}\|_\infty = \Re(\mathbf{y}^H\dot{\mathbf{z}}) - \varepsilon\|\dot{\mathbf{z}}\|_2 \leq \left|\mathbf{y}^H\dot{\mathbf{z}}\right| - \varepsilon\|\dot{\mathbf{z}}\|_2$$
$$\leq \|\dot{\mathbf{z}}\|_2(\|\mathbf{y}\|_2 - \varepsilon), \tag{16}$$

as an immediate consequence of the Cauchy-Schwarz inequality.[2] In the remaining steps of the proof, we derive an upper bound on $\|\dot{\mathbf{z}}\|_2$ in (16). To this end, we expand

$$\|\dot{\mathbf{z}}\|_2 = \left\|(\mathbf{D}\mathbf{D}^H)^{-1}\mathbf{D}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 \tag{17}$$

where $\mathbf{D}\mathbf{D}^H$ is invertible since $\mathbf{D}$ is a frame with lower frame bound satisfying $A > 0$. Application of the Rayleigh-Ritz theorem [25, Thm. 4.2.2] to the right-hand side (RHS) of (17) leads to the following upper bound:

$$\|\dot{\mathbf{z}}\|_2 \leq \left\|(\mathbf{D}\mathbf{D}^H)^{-1}\right\|_{2,2}\left\|\mathbf{D}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 \leq \frac{1}{A}\left\|\mathbf{D}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2, \tag{18}$$

where the second inequality is a result of

$$\left\|(\mathbf{D}\mathbf{D}^H)^{-1}\right\|_{2,2} = \frac{1}{\|(\mathbf{D}\mathbf{D}^H)^{-1}\|_{2,2}} \leq \frac{1}{A}$$

and the assumption that $\mathbf{D}$ is a frame with lower frame bound $A > 0$. We next derive an upper bound on $\left\|\mathbf{D}\mathbf{D}^H\mathbf{z}\right\|_2$ using (18).

Note that one can straightforwardly arrive at an upper bound on $\|\dot{\mathbf{x}}\|_\infty$ as follows:

$$\left\|\mathbf{D}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 \leq \|\mathbf{D}\|_{2,2}\left\|\mathbf{D}^H\dot{\mathbf{z}}\right\|_1 \leq \|\mathbf{D}\|_{2,2} \tag{19}$$

using $\left\|\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 \leq \left\|\mathbf{D}^H\dot{\mathbf{z}}\right\|_1$ and the constraint $\left\|\mathbf{D}^H\dot{\mathbf{z}}\right\|_1 \leq 1$ of the dual problem $(\mathrm{D}_\infty^\varepsilon)$. Hence, by combining (16), (18), and (19) one would arrive at the following result:

$$\|\dot{\mathbf{x}}\|_\infty \leq \frac{\|\mathbf{D}\|_{2,2}}{A}(\|\mathbf{y}\|_2 - \varepsilon). \tag{20}$$

[2]Note that the bound (16) appears to be tight for $\varepsilon = 0$, i.e., we were able to construct signal and frame instances for which we have $\|\dot{\mathbf{x}}\|_\infty = \|\mathbf{y}\|_2\|\dot{\mathbf{z}}\|_2$ up to machine precision. A systematic characterization of such signal and frame instances is, however, left for future work.

This bound is, however, overly pessimistic and does not exploit additional properties of the frame $\mathbf{D}$. Note that for a Parseval frame, the result (20) leads to the bound $\|\dot{\mathbf{x}}\|_\infty \leq \|\mathbf{y}\|_2 - \varepsilon$.

B. Refined Upper Bound

In order to arrive at a refined bound on $\left\|\mathbf{D}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2$, we define an N-dimensional vector $\mathbf{v} = \mathbf{D}^H\dot{\mathbf{z}}$ and divide its coefficients into $S = \lceil 1/\delta \rceil$ disjoint support sets, each[3] of cardinality δN such that

$$\Omega_1 \cup \cdots \cup \Omega_\ell \in \{1, \ldots, N\}.$$

Moreover, the magnitudes of the entries in $\mathbf{v}$ associated to set Ω_ℓ are no smaller than the magnitudes associated with the sets Ω_k, $k > \ell$. In other words, Ω_1 contains the indices associated to the largest δN entries in $\mathbf{v}$, Ω_2 the δN coefficients associated to the second largest entries, etc. This partitioning scheme now allows us to rewrite $\left\|\mathbf{D}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2$ as

$$\left\|\mathbf{D}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 = \left\|\mathbf{D}\sum_{i=1}^S \mathbf{P}_{\Omega_i}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2,$$

where the matrix $\mathbf{P}_{\Omega_i}$ realizes a projection onto the set Ω_i. Application of the triangle inequality, followed by applying the UP with parameters η, δ leads to the following bounds:

$$\left\|\mathbf{D}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 \leq \sum_{i=1}^S\left\|\mathbf{D}\mathbf{P}_{\Omega_i}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 \leq \sum_{i=1}^S \eta\left\|\mathbf{P}_{\Omega_i}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2$$
$$= \eta\left\|\mathbf{P}_{\Omega_1}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 + \sum_{i=2}^S \eta\left\|\mathbf{P}_{\Omega_i}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2. \tag{21}$$

Since the sets Ω_i order the entries of $\mathbf{v} = \mathbf{D}^H\dot{\mathbf{z}}$ according to their magnitudes, we can use a technique developed in [26], which states that for $i \in \{2, \ldots, S\}$ we have

$$\left\|\mathbf{P}_{\Omega_i}\mathbf{v}\right\|_2 \leq \sqrt{\delta N}\left\|\mathbf{P}_{\Omega_i}\mathbf{v}\right\|_\infty \leq \frac{1}{\sqrt{\delta N}}\left\|\mathbf{P}_{\Omega_{i-1}}\mathbf{v}\right\|_1.$$

This result in combination with the RHS of (21) leads to

$$\left\|\mathbf{D}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 \leq \eta\left\|\mathbf{P}_{\Omega_1}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 + \sum_{i=1}^S \frac{\eta}{\sqrt{\delta N}}\left\|\mathbf{P}_{\Omega_i}\mathbf{D}^H\dot{\mathbf{z}}\right\|_1$$
$$= \eta\left\|\mathbf{P}_{\Omega_1}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 + \frac{\eta}{\sqrt{\delta N}}\left\|\mathbf{D}^H\dot{\mathbf{z}}\right\|_1$$
$$= \eta\left\|\mathbf{P}_{\Omega_1}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 + \frac{\eta}{\sqrt{\delta N}}, \tag{22}$$

where the first equality follows from the fact that $\left\|\mathbf{D}^H\dot{\mathbf{z}}\right\|_1 \leq 1$ for any solution $\dot{\mathbf{z}}$ to the dual problem $(\mathrm{D}_\infty^\varepsilon)$.

We can now bound the first RHS term in (22) as

$$\left\|\mathbf{P}_{\Omega_1}\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 \leq \left\|\mathbf{D}^H\dot{\mathbf{z}}\right\|_2 \leq \sqrt{B}\|\dot{\mathbf{z}}\|_2 \tag{23}$$

using the facts that i) $\mathbf{P}_{\Omega_1}$ is a projector and ii) $\mathbf{D}$ is a frame with (upper) frame bound B. By combining (18), (22),

[3]Note that the last support set Ω_ℓ can have a cardinality that is smaller than δN; such cases, however, leave the proof unaffected.

and (23) we arrive at

$$\|\dot{\mathbf{z}}\|_2 \leq \frac{1}{A}\left(\eta\sqrt{B}\|\dot{\mathbf{z}}\|_2 + \frac{\eta}{\sqrt{\delta N}}\right),$$

which can be rewritten as

$$\|\dot{\mathbf{z}}\|_2 \leq \frac{\eta}{(A - \eta\sqrt{B})\sqrt{\delta N}} \tag{24}$$

provided that $A > \eta\sqrt{B}$ holds. Combining (16) with (24) finally yields

$$\|\dot{\mathbf{x}}\|_\infty \leq \frac{\eta}{(A - \eta\sqrt{B})\sqrt{\delta N}}\left(\|\mathbf{y}\|_2 - \varepsilon\right), \tag{25}$$

which concludes the proof. We finally note that (25) is able to scale in $1/\sqrt{N}\|\mathbf{y}\|_2$ for certain frames (see Section IV).

APPENDIX D
PROOF OF THEOREM 4

The proof follows from separately bounding the numerator and denominator of the PAPR defined in (5). We first bound $N\|\dot{\mathbf{x}}\|_\infty^2$ using (4) to arrive at

$$N\|\dot{\mathbf{x}}\|_\infty^2 \leq K_\mathrm{u}^2(\|\mathbf{y}\|_2 - \varepsilon)^2. \tag{26}$$

The second part of the proof bounds $\|\dot{\mathbf{x}}\|_2^2$ from below. To this end, it is important to realize that

$$\|\dot{\mathbf{x}}\|_2 \geq \min_{\mathbf{x}}\left\{\|\mathbf{x}\|_2 : \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \varepsilon\right\} \tag{27}$$

because $\dot{\mathbf{x}}$ satisfies $\|\mathbf{y} - \mathbf{D}\dot{\mathbf{x}}\|_2 \leq \varepsilon$ and the RHS is the minimizer for all vectors $\mathbf{x} \in \mathbb{C}^N$ satisfying $\|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \varepsilon$. We next compute a lower bound on the RHS of (27). From Lemma 1 with $p = 2$ and $q = 2$, we have

$$\min_{\mathbf{x}}\left\{\|\mathbf{x}\|_2 : \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \varepsilon\right\}$$
$$= \max_{\mathbf{z}}\left\{\Re\left(\mathbf{y}^H\mathbf{z}\right) - \varepsilon\|\mathbf{z}\|_2 : \left\|\mathbf{D}^H\mathbf{z}\right\|_2 \leq 1\right\}. \tag{28}$$

Using a similar strategy as in Appendix B, we replace the optimal solution $\dot{\mathbf{z}}$ of the dual problem in (28) by the estimate

$$\dot{\mathbf{z}} = \frac{\mathbf{y}}{\|\mathbf{D}^H\mathbf{y}\|_2}, \tag{29}$$

which satisfies the constraint $\left\|\mathbf{D}^H\mathbf{y}\right\|_2 \leq 1$. Hence, inserting the estimate (29) into the RHS of (28) leads to the following lower bound:

$$\max_{\mathbf{z}}\left\{\Re\left(\mathbf{y}^H\mathbf{z}\right) - \varepsilon\|\mathbf{z}\|_2 : \left\|\mathbf{D}^H\mathbf{z}\right\|_2 \leq 1\right\}$$
$$\geq \frac{\|\mathbf{y}\|_2^2 - \varepsilon\|\mathbf{y}\|_2}{\|\mathbf{D}^H\mathbf{y}\|_2}. \tag{30}$$

The upper frame bound $\left\|\mathbf{D}^H\mathbf{y}\right\|_2 \leq \sqrt{B}\|\mathbf{y}\|_2$ enables us to further bound the RHS of (30) from below as

$$\frac{\|\mathbf{y}\|_2^2 - \varepsilon\|\mathbf{y}\|_2}{\|\mathbf{D}^H\mathbf{y}\|_2} \geq \frac{\|\mathbf{y}\|_2 - \varepsilon}{\sqrt{B}}. \tag{31}$$

By combining (27), (28), (30), and (31), we finally obtain

$$\|\dot{\mathbf{x}}\|_2^2 \geq \frac{\left(\|\mathbf{y}\|_2 - \varepsilon\right)^2}{B}. \tag{32}$$

Consequently, if $\varepsilon < \|\mathbf{y}\|_2$, then we can bound the PAPR of the democratic representation $\dot{\mathbf{x}}$ obtained from (P_∞^ε) using (26) and (32) as $PAPR(\dot{\mathbf{x}}) \leq K_\mathrm{u}^2 B$. Note that in (29) we assumed that $\dot{\mathbf{x}} \neq 0$, i.e., we require $A > 0$ and $\|\mathbf{y}\|_2 \neq 0$, which concludes the proof.

REFERENCES

[1] Y. Lyubarskii and R. Vershynin, "Uncertainty principles and vector quantization," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3491–3501, Jul. 2010.

[2] R. van Nee and R. Prasad, *OFDM for wireless multimedia communications*. Artech House Publ., 2000.

[3] J. Illic and T. Strohmer, "PAPR reduction in OFDM using Kashin's representation," in *Proc. IEEE 10th Workshop on Sig. Proc. Advances in Wireless Comm. (SPAWC)*, Perugia, Italy, June 2009, pp. 444–448.

[4] C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO-OFDM downlink," *IEEE J. Sel. Areas Comm.*, to appear.

[5] H. Jégou, T. Furon, and J.-J. Fuchs, "Anti-sparse coding for approximate nearest neighbor search," *arXiv:1110.3767v2*, Oct. 2011.

[6] J. A. Cadzow, "Algorithm for the minimum-effort problem," *IEEE Trans. Autom. Control.*, vol. 16, no. 1, pp. 60–63, Feb. 1971.

[7] A. S. Deo and I. D. Walker, "Minimum effort inverse kinematics for redundant manipulators," *IEEE Trans. Robotics and Automation*, vol. 13, no. 5, pp. 767–775, Oct. 1997.

[8] Y. Zhang, J. Wang, and Y. Xu, "A dual neural network for bi-criteria kinematic control of redundant manipulators," *IEEE Trans. Robotics and Automation*, vol. 18, no. 6, pp. 923–931, Dec. 2002.

[9] D. L. Donoho and J. Tanner, "Precise undersampling theorems," *Proc. of the IEEE*, vol. 98, no. 6, pp. 913–924, June 2010.

[10] O. Mangasarian and B. Recht, "Probability of unique integer solution to a system of linear equations," *Europ. J. of Operational Research*, vol. 214, no. 1, pp. 27–30, 2011.

[11] L. W. Neustadt, "Minimum effort control systems," *J. Soc. Indus. and Appl. Math. Ser*, vol. 1, no. 1, pp. 16–31, Apr. 1962.

[12] J.-J. Fuchs, "Spread represenations," in *Proc. 45th Asilomar Conf. on Signals, Systems, and Comput.*, Pacific Grove, CA, USA, 2011.

[13] B. Kashin, "Sections of some finite dimensional sets and classes of smooth functions," *Izv. Acad. Nauk SSSR*, vol. 41, no. 2, pp. 334–351, 1977.

[14] A. Y. Garnaev and E. D. Gluskin, "On widths of the Euclidean ball," *Soviet Math. Dokl.*, vol. 30, no. 1, pp. 200–204, 1984.

[15] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[16] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections and universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, pp. 5406–5425, 2006.

[17] R. G. Baraniuk, M. Davenport, R. A. DeVore, and M. B. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, pp. 253–263, Jan. 2008.

[18] C. Studer and R. G. Baraniuk, "Stable restoration and separation of approximately sparse signals," *submitted to Apppl. Comput. Harm. Anal.*, July.

[19] V. I. Morgenshtern and H. Bölcskei, *A short course on frame theory*, E. Serpedin, T. Chen, and D. Rajan, Ed. Chapter in Mathematical Foundations for Signal Processing, Communications, and Networking: CRC Press, 2012.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[21] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," http://cvxr.com/cvx/, Apr. 2011.

[22] M. E. Pfetsch and A. M. Tillmann, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *arXiv:1205.2081v2*, May 2012.

[23] J.-J. Fuchs, Personal communication, 2011.

[24] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Comm.*, vol. 12, no. 2, pp. 1536–1284, Apr. 2005.

[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, NY: Cambridge Press, 1985.

[26] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *C. R. Acad. Sci. Paris, Ser. I*, vol. 346, pp. 589–592, 2008.