BEING LOW-RANK IN THE TIME-FREQUENCY PLANE

Valentin Emiya[1], *Ronan Hamon*[1], *Caroline Chaux*[2]

[1] Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France
[2] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France

ABSTRACT

When using optimization methods with matrix variables in signal processing and machine learning, it is customary to assume some low-rank prior on the targeted solution. Nonnegative matrix factorization of spectrograms is a case in point in audio signal processing. However, this low-rank prior is not straightforwardly related to complex matrices obtained from a short-time Fourier – or discrete Gabor – transform (STFT), which is generally defined from and studied based on a modulation operator and a translation operator applied to a so-called window. This paper is a first study of the low-rankness property of time-frequency matrices. We characterize the set of signals with a rank-r (complex) STFT matrix in the case of a unit hop size and frequency step with few assumptions on the transform parameters. We discuss the scope of this result and its implications on low-rank approximations of STFT matrices.

Index Terms— Short-Time Fourier Transform, low-rankness, approximation.

1. INTRODUCTION

The Short Time Fourier Transform (STFT), or more generally the Gabor transform, has been widely used, especially when considering audio signal processing. Indeed, time-frequency representations [1] of such signals may ease their processing [2]. The STFT can be defined as a frame by applying a translation operator $\mathcal{T}$ and a modulation operator $\mathcal{M}$ to a so-called window [3]. Analyzing a signal with such a frame produces a time-frequency matrix, which can be after that used to analyze and process the signal. Much attention has been paid in the literature to such transformations to study both theoretical aspects and implementations [4].

From the perspective of the optimization over matrix variables in signal processing or machine learning problems, one often uses low-rank matrix factorization models to find some structure in the data and/or to add a good regularization that will improve the estimation. Famous examples are given by collaborative filtering problems with applications to recommender systems [5]. In audio, such strategies proved to be successful when applied to spectrograms, i.e., by retaining the modulus of the STFT coefficients, ignoring the phases, using nonnegative matrix factorization (NMF) techniques [6, 7, 8]. It is useful to extract elementary components, such as musical notes or spectral source patterns, alongside their activation patterns in time. These results give credence to the intuition of the low-rank nature of such matrices, easily visible in spectrograms, as in Figure 1, and the limited number of sound elements. While NMF approaches have given good results for audio problems, such as source separation [7], music transcription [6], or spectrogram inpainting [9], two major limitations remain current challenges in research in audio processing. The first limitation is the approximation made by summing the moduli, or squared moduli, instead of summing the original coefficients: only time-frequency components that are not overlapping or that are dominating are preserved. The second limitation is that the phases associated to each coefficient are required to be able to synthesize an audio signal after processing the spectrogram. The problem of phase reconstruction is not only a difficult problem that is still under intensive research, but also requires to split the processing into two separate problems, on the amplitudes and on the phases successively, which is bound to give suboptimal results.

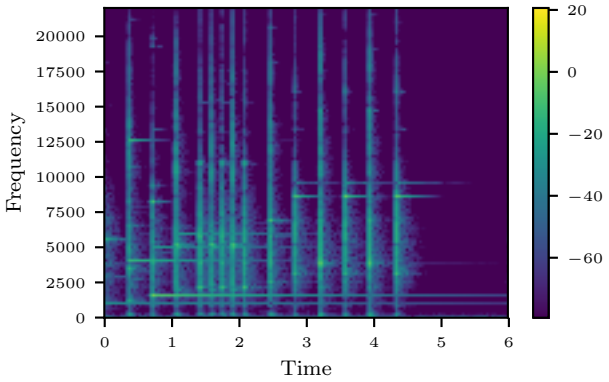

Fig. 1. Spectrogram of the *Glockenspiel*, composed of about 50 spectral peaks distributed on 15 occurrences of 8 notes. Does the approximate rank of the complex-valued STFT matrix equal 8, 15, 50, or another value?

One may wish in having the best of both worlds by modeling the complex-valued STFT matrices using low-rank factorization, allowing true additivity between components, joint modeling of the amplitudes and of the phases of the coefficients, and the use of a wide range of optimization and learning tools. However, the definition of the STFT, based on the modulation and the translation of a window, does not give straightforward evidence about the low-rankness of the resulting time-frequency matrices, nor their capability to be well approximated by low-rank matrices.

In this paper, we propose to study low-rank time-frequency (complex-valued) matrices, in order to exhibit the nature of these objects and their potential use to model audio data. The following aspects will be particularly of interest:

- from a general viewpoint, how the intuitions of the low-rankness of the spectrograms can be extended to complex-valued time-frequency matrices, and how to validate them or not ?

This work was supported by the Agence Nationale de la Recherche under grant JCJC MAD (ANR-14-CE27-0002).

- what is a rank-one matrix, or more generally a rank-r matrix, in the time-frequency plane? Can the set of rank-r time-frequency matrices be fully characterized?
- do time-frequency matrices of real-world sounds have good low-rank approximations? Which kind of elementary patterns are obtained?

The paper is organized as follows: Section 2 is dedicated to the study of the rank of STFT matrices. Numerical simulations are provided in Section 3, while Section 4 details conclusions on the low-rankness of STFT matrices and consequences for their modeling.

2. CHARACTERIZATION OF LOW-RANK STFT MATRICES

In this first section, we characterize the set of signals with a rank-r STFT matrix, for any $r \in \mathbb{N}_*$, in a context where the signals are discrete and have finite length, and the STFT is circular with maximal redundancy (i.e., unit hop size and a number of frequency bins equal to the length of the signal). These conditions, and the scope of the given results, will be broadly discussed at the end of the section.

2.1. Definitions and properties of STFT representations

We consider complex-valued vectors of length L describing discrete finite signals, denoted by $(\mathbf{s}[m])_{m\in[L]} \in \mathbb{C}^L$, where $[\![L]\!] = \{0,\dots,L-1\}$ is the set of the first L integers. The so-called window is denoted by $(\mathbf{h}[m])_{m\in[\![L]\!]} \in \mathbb{C}^L$. As defined above, a signal $\mathbf{s}$ and a window $\mathbf{h}$ are indexed from 0 to $L-1$. Whenever index m is outside this range, we will consider the L-periodic extension of $\mathbf{s}$ and $\mathbf{h}$ defined by $\forall m \in \mathbb{Z}$, $\mathbf{s}[m] = \mathbf{s}[m\%L]$ and $\mathbf{h}[m] = \mathbf{h}[m\%L]$, where $\%$ denotes the modulo operator.

Definition 2.1 (Fourier matrix)**.** The Fourier matrix $\mathbf{E} \in \mathbb{C}^{L\times L}$ is defined by

$$\mathbf{E} = \left(e^{-2i\pi\frac{kt}{L}}\right)_{k\in[\![L]\!],t\in[\![L]\!]} \tag{1}$$

Definition 2.2 (DFT and IDFT)**.** The discrete Fourier transform (DFT) of $\mathbf{u} \in \mathbb{C}^L$ on L discrete frequencies is $\widehat{\mathbf{u}} = \mathrm{DFT}(\mathbf{u}) = \mathbf{E}\mathbf{u}$. The inverse discrete Fourier transform (IDFT) of $\mathbf{u} \in \mathbb{C}^L$ is $\check{\mathbf{u}} = \mathrm{IDFT}(\mathbf{u}) = \mathbf{E}^{-1}\mathbf{u} = \frac{1}{L}\mathbf{E}^*\mathbf{u}$ where $\mathbf{E}^*$ is the adjoint of $\mathbf{E}$.

There exist different conventions to define the STFT on K discrete frequencies $\{\nu_k\}_{k=0}^{K-1}$ with $\nu_k = \frac{k}{K}$ and N time steps $\{t_n\}_{n=0}^{N-1}$, with $t_n = nh$, where h is an arbitrary hop size. If we denote by $\mathcal{T}_n$ the translation by t_n and by $\mathcal{M}_k$ the modulation by ν_k, these conventions depend on the order of application of operators $\mathcal{T}_n$ and $\mathcal{M}_k$.

Definition 2.3 ($(K\times N)$-STFT, band-pass convention)**.** In the so-called band-pass convention, the $(K\times N)$-STFT of $\mathbf{s} \in \mathbb{C}^L$ is defined on discrete frequency $\nu_k, k \in [\![K]\!]$ and discrete time $t_n, n \in [\![N]\!]$ by

$$\mathbf{S}_{\mathrm{BP}}^{(K\times N)}[k,n] = \langle \mathcal{T}_n\mathcal{M}_k\mathbf{h}, \mathbf{s}\rangle = \sum_m \mathbf{s}[t_n+m]\,\mathbf{h}[m]\,e^{-2i\pi\nu_k m}. \tag{2}$$

Definition 2.4 ($(K\times N)$-STFT, low-pass convention)**.** In the so-called low-pass convention, the $(K\times N)$-STFT of $\mathbf{s} \in \mathbb{C}^L$ is defined on discrete frequency $\nu_k, k \in [\![K]\!]$ and discrete time $t_n, n \in [\![N]\!]$ by

$$\mathbf{S}_{\mathrm{LP}}^{(K\times N)}[k,n] = \langle \mathcal{M}_k\mathcal{T}_n\mathbf{h}, \mathbf{s}\rangle = \sum_m \mathbf{s}[m]\,\mathbf{h}[m-t_n]\,e^{-2i\pi\nu_k m}. \tag{3}$$

Those definitions are linked by the following relation: $\forall k \in [\![K]\!]$, $n \in \mathbb{Z}$, $\mathbf{S}_{\mathrm{LP}}(k,n) = \mathbf{S}_{\mathrm{BP}}(k,n) \times e^{-2i\pi\nu_k m_n}$, that highlights an identical content between the two conventions up to a frequency, as well as a similar resulting spectrogram.

A case of interest is when $K = N = L$, i.e., the redundancy is maximal. The $(L\times L)$-STFT of $\mathbf{s} \in \mathbb{C}^L$ in both conventions, which will be denoted respectively by $\mathbf{S}_{\mathrm{BP}} = \mathbf{S}_{\mathrm{BP}}^{(L\times L)}$ and $\mathbf{S}_{\mathrm{LP}} = \mathbf{S}_{\mathrm{LP}}^{(L\times L)}$, are rewritten

$$\forall k,n, \mathbf{S}_{\mathrm{BP}}[k,n] = \sum_m \mathbf{s}[n+m]\,\mathbf{h}[m]\,e^{-2i\pi\frac{km}{L}} \tag{4}$$

$$\forall k,n, \mathbf{S}_{\mathrm{LP}}[k,n] = \sum_m \mathbf{s}[m]\,\mathbf{h}[m-n]\,e^{-2i\pi\frac{km}{L}}. \tag{5}$$

It is worth noting that any $(K\times N)$-STFT representations can be obtained from the $(L\times L)$-STFT representation by keeping every $\frac{L}{K}$ row and every $\frac{L}{N}$ column.

Proposition 2.1. *Let $K, N \in \mathbb{N}$ be such that $K|L$ and $N|L$. Then for any $k \in [\![K]\!], n \in [\![N]\!]$, we have*

$$\mathbf{S}_{BP}^{(K\times N)}[k,n] = \mathbf{S}_{BP}\left[\frac{kL}{K}, \frac{nL}{K}\right] \tag{6}$$

and

$$\mathbf{S}_{LP}^{(K\times N)}[k,n] = \mathbf{S}_{LP}\left[\frac{kL}{K}, \frac{nL}{K}\right]. \tag{7}$$

Proof. The proof follows from the definition of each matrix. □

2.2. Characterization of rank-r STFT matrices

Lemme 2.2 (Factorization of STFT matrices)**.** *For any signal $\mathbf{s} \in \mathbb{C}^L$ and window $\mathbf{h} \in \mathbb{C}^L$, we have*

$$\mathbf{S}_{BP} = \mathbf{E}\,\mathrm{diag}(\mathbf{h})\,\mathbf{E}^{-1}\,\mathrm{diag}(\widehat{\mathbf{s}})\,\mathbf{E} \tag{8}$$

and

$$\mathbf{S}_{LP} = \mathbf{E}\,\mathrm{diag}(\mathbf{s})\,\mathbf{E}^{-1}\,\mathrm{diag}\left(\widehat{\mathbf{h}}\right)\mathbf{E} \tag{9}$$

Proof. Let us define $\mathbf{H}[k,n] = \mathbf{h}[k-n]$ for any $k,n \in [\![L]\!]$. Because we consider a L-periodic extension of signals, $\mathbf{H} \in \mathbb{C}^{L\times L}$ is a circulant matrix, and then $\mathbf{H} = \mathbf{E}^{-1}\,\mathrm{diag}\left(\widehat{\mathbf{h}}\right)\mathbf{E}$. In the low-pass convention, let $\mathbf{s}_n \in \mathbb{C}^L$ be the windowed signal in frame $n \in [\![L]\!]$, so that we have, for any $k \in [\![L]\!]$

$$\mathbf{s}_n[k] = \mathbf{s}[k]\,\mathbf{h}[k-n] = \mathbf{s}[k]\,\mathbf{H}[k,n] = (\mathrm{diag}(\mathbf{s})\,\mathbf{H})[k,n], \tag{10}$$

the STFT matrix can be then written for $k,n \in [\![L]\!]$ as

$$\mathbf{S}_{\mathrm{LP}}[k,n] = (\mathbf{E}\mathbf{s}_n)[k] = (\mathbf{E}\,\mathrm{diag}(\mathbf{s})\,\mathbf{H})[k,n] \tag{11}$$

$$= \left(\mathbf{E}\,\mathrm{diag}(\mathbf{s})\,\mathbf{E}^{-1}\,\mathrm{diag}\left(\widehat{\mathbf{h}}\right)\mathbf{E}\right)[k,n], \tag{12}$$

which proves (9). Eq. (8) is obtained in the same way by permuting the role of $\mathbf{s}$ and $\mathbf{h}$. □

Using this new formulation of the STFT, it is possible, under some conditions on the signal or the window, to characterize the rank of the resulting STFT matrices in both conventions.

Theorem 2.3 (Rank-r STFT matrices)**.** *Let $\mathbf{s} \in \mathbb{C}^L$. If $\mathbf{h} \in \mathbb{C}^L$ is a window that does not vanish, i.e., $\forall k \in [\![L]\!], \mathbf{h}[k] \neq 0$, then* $\mathrm{rank}(\mathbf{S}_{BP}) = \|\widehat{\mathbf{s}}\|_0$.

Similarly, if $\mathbf{h} \in \mathbb{C}^L$ is a window such that its DFT does not vanish, i.e., $\forall k \in [\![L]\!], \widehat{\mathbf{h}}[k] \neq 0$, then $\mathrm{rank}(\mathbf{S}_{LP}) = \|\mathbf{s}\|_0$.

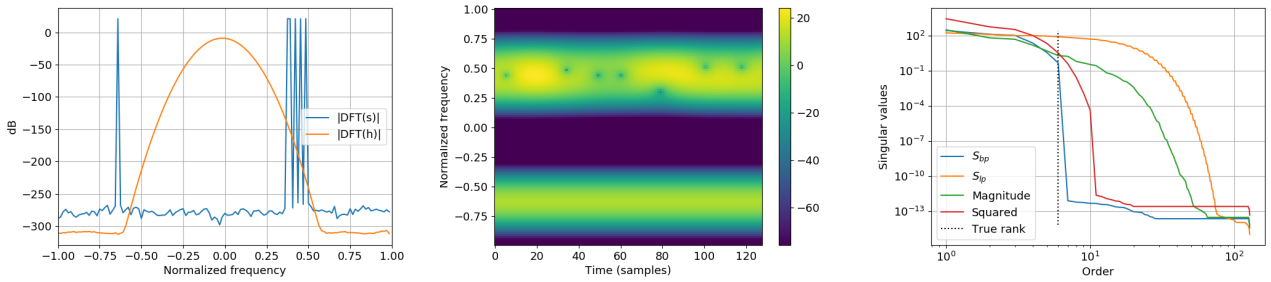

Fig. 2. Analysis of a signal with $N_c = 6$ sinusoids with a Gaussian window: DFT of the signal and of the window (left), spectrogram (middle) and singular values of STFT matrices, magnitude and energy spectrograms (right).

Proof. If the window $\mathbf{h}$ does not vanish, then $\operatorname{rank}(\operatorname{diag}(\mathbf{h})) = L$. According to Proposition 2.2, $\mathbf{S}_{\text{BP}} = \mathbf{E}\operatorname{diag}(\mathbf{h})\,\mathbf{E}^{-1}\operatorname{diag}(\widehat{\mathbf{s}})\,\mathbf{E}$ is the matrix product of $\operatorname{diag}(\widehat{\mathbf{s}})$ with full rank matrices, leading to a resulting matrix with rank $\operatorname{rank}(\mathbf{S}_{\text{BP}}) = \operatorname{rank}(\operatorname{diag}(\widehat{\mathbf{s}})) = \|\widehat{\mathbf{s}}\|_0$.

Similarly, $\mathbf{S}_{\text{LP}} = \mathbf{E}\operatorname{diag}(\mathbf{s})\,\mathbf{E}^{-1}\operatorname{diag}\left(\widehat{\mathbf{h}}\right)\mathbf{E}$ is the matrix product of $\operatorname{diag}(\mathbf{s})$ with full rank matrices if $\widehat{\mathbf{h}}$ does not vanish, leading to a resulting matrix product with rank $\operatorname{rank}(\mathbf{S}_{\text{LP}}) = \operatorname{rank}(\operatorname{diag}(\mathbf{s})) = \|\mathbf{s}\|_0$.

□

This theorem states that under weak conditions that are commented below, the set of rank-r STFT matrices in the band-pass convention is composed of signals that are the sum of r pure complex exponentials at Fourier frequencies, while the set of rank-r STFT matrices in the low-pass convention is composed of signals that are the sum of r diracs at integer times. Due to the duality between the two conventions and between dirac and sinusoids, we will mainly comment on the case of $\mathbf{S}_{\text{BP}}$.

Rank-one elements. Theorem 2.3 states that the elementary rank-one components in low-rank $\mathbf{S}_{\text{BP}}$ matrices are pure sines at Fourier frequencies. This is a very restricted set compared to the intuition we can have on spectrograms, where an exact or approximate rank-one matrix may be composed by *any fixed spectral pattern* – e.g., a sum of sinusoids or a wide-band noise – possibly *modulated in amplitude* in consecutive frames. No such degrees of freedom are allowed in complex-valued rank-one STFT matrices.

Finite signals, discrete frequencies and circularity. The fact that only discrete Fourier frequencies are allowed in Theorem 2.3 can be viewed as a boundary effect. Indeed, since signals are finite and circular, only sinusoids with discrete Fourier frequencies have an integral number of periods and do not have discontinuities. Hence, this constraint on discrete frequencies is not very restrictive: one may obtain rank-r matrices with a finer frequency discretization by increasing the length of the signal, until having a continuous set of frequencies when the signal is supported on $\mathbb{Z}$.

Condition on the window. The condition on the absence of zeros in window $\mathbf{h}$ or in its DFT can be satisfied easily, e.g., by choosing a (truncated) Gaussian window, as illustrated in Figure 2. Theorem 2.3 does not cover the use of windows with a more compact support that results in zeros in the time and Fourier domains – e.g., Hamming, Hann, rectangular, and so on. We conjecture that in those cases, some additional signals may have a low-rank STFT matrix, with a very specific content that depends on the STFT parameters, without altering the main analysis and conclusions of this paper. A dedicated study of those signals would extend the scope of Theorem 2.3. Another remark is that in practice when the signal length L is large, even if the condition on the window is satisfied, the amplitude of a window or its DFT may decrease down to the numerical precision: this may cause some changes in the actual rank of the STFT matrix, as illustrated in Figure 3 below.

Extension to rank $K \times N$-STFT matrices. Under the same conditions as in Theorem 2.3, the rank of any $K \times N$-STFT matrices is upper-bounded by the rank of the $N \times N$-STFT given by Theorem 2.3, thanks to eq. (6) and (7). An exact characterization of the rank of $\mathbf{S}_{\text{BP}}^{(K\times N)}$ and $\mathbf{S}_{\text{LP}}^{(K\times N)}$ would be an interesting extension of Theorem 2.3.

3. NUMERICAL EXPERIMENTS

All the experiments in this section have been realized using the Python version[1] of the LTFAT toolbox [4]. For reproducibility purposes, the code will be made available with the paper[2].

3.1. Analysis of low-rank STFT matrices

A direct illustration of Theorem 2.3 is represented in Figure 2, that shows the analysis of a signal with length $L = 128$ and composed by a sum of $N_c = 6$ complex sinusoids at exact Fourier frequencies. A group of sinusoids have closed frequencies $\frac{25}{128}, \frac{26}{128}, \frac{28}{128}, \frac{30}{128}, \frac{32}{128}$ while the other sinusoid with frequency $\frac{-40}{128}$ is isolated, in terms of the frequency resolution imposed by the Gaussian window in use. This can be observed in the DFT of the signal and of the window in the left plot of Figure 2. As a result on the spectrogram, the group of sinusoids generates a wide, modulated strip, while the isolated component appears as a time-invariant pattern. When looking at the singular values in decreasing order on the right plot of Figure 2, one can see that the rank of $\mathbf{S}_{\text{BP}}$ exactly equals the number of sinusoids, as predicted by Theorem 2.3. Matrix $\mathbf{S}_{\text{LP}}$, in the other STFT convention, has many high singular values, which prevents any good low-rank approximation. This illustrates the importance of the choice of the STFT convention. The dual counterpart, with a signal composed of diracs at discrete times 25, 26, 28, 30, 32, and 88, would result in a permutation between the $\mathbf{S}_{\text{BP}}$ and $\mathbf{S}_{\text{LP}}$ curves. The rank of the energy spectrogram is larger than the number of components, while the magnitude spectrogram has singular values that decrease more slowly.

Another illustration is given in Figure 3, that displays the rank of the STFT representations and the corresponding spectrograms with

[1] `ltfatpy` module, `http://pythonhosted.org/ltfatpy/`.
[2] `https://mad.lis-lab.fr/`

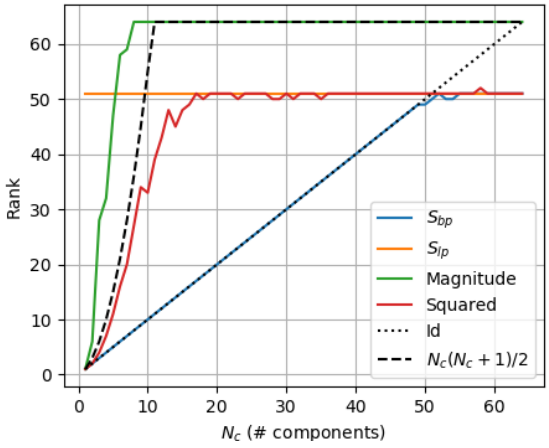

Fig. 3. Rank of the STFT matrix with respect to the number of sinusoids in the signal, for several types of time-frequency matrices.

respect to the number of components N_c of the signal of length $L = 64$. In this example, the frequencies of components are drawn randomly at exact Fourier frequencies. The rank of $\mathbf{S}_{\text{BP}}$ perfectly equals the number of components, as predicted by Theorem 2.3, except for high values where a plateau is reached, which may be due to numerical precision. As expected, the rank of $\mathbf{S}_{\text{LP}}$ is high. The rank of the energy spectrogram seems to follow a trend in $\frac{N_c(N_c+1)}{2}$, which is an upper bound that can be found by the fact that for any matrix $\mathbf{M}$ with rank r, the rank of $|\mathbf{M}|^2$ is upper bounded by $\frac{r(r+1)}{2}$ (the proof is easy and omitted by lack of space). This is due to the interference between close components, which are perfectly modeled as a low-rank part in the $\mathbf{S}_{\text{BP}}$ matrix and not in the spectrogram matrix. Together with the magnitude spectrogram, which has a higher rank, this shows that in this context, an STFT matrix may be of lower rank than the related spectrograms.

3.2. Low-rank approximation of STFT matrices of audio signals

While the previous section demonstrates that low-rank STFT matrices are in a very narrow set, we illustrate in this section how this impacts the low-rank approximation of STFT matrices of real sounds. The (classical) problem considered here consists in finding the best approximation $\widetilde{\mathbf{X}} \in \mathbb{C}^{K\times N}$ of rank r of a matrix $\mathbf{X} \in \mathbb{C}^{K\times N}$, that is to say to solve the following problem:

$$\widetilde{\mathbf{X}} = \underset{\mathbf{Y}\in\mathbb{C}^{K\times N}, \text{rg}(\mathbf{Y})\leq r}{\arg\min} \|\mathbf{X}-\mathbf{Y}\|_F^2 . \tag{13}$$

The Eckart-Young theorem [10] states that the solution is $\widetilde{\mathbf{X}} = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^*$, where $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ is the singular value decomposition (SVD) of $\mathbf{X}$ and $\widetilde{\mathbf{U}}$ (resp. $\widetilde{\mathbf{\Sigma}}$, $\widetilde{\mathbf{V}}$) is the truncated version of $\mathbf{U}$ (resp. $\mathbf{\Sigma}$, $\mathbf{V}$) related to the r largest singular values of $\mathbf{X}$. In addition, the minimum satisfies $\left\|\mathbf{X}-\widetilde{\mathbf{X}}\right\|_F^2 = \sum_{k=r+1}^{\min(K,N)} \sigma_k^2$. We solve this problem for $\mathbf{X} = \mathbf{S}_{\text{BP}}$ and $\mathbf{X} = |\mathbf{S}_{\text{BP}}|$ for different values of r.

Figure 4 displays the normalized approximation error of the STFT representation, obtained after performing a low-rank approximation of the STFT and of the corresponding spectrogram matrices, for the Glockenspiel sound, whose spectrogram is displayed in Figure 1. The sound is sampled at 44.1 KHz, and the STFT is computed using a Hann window of size 2048 with 75% overlapping. The approximation is better for the spectrograms than for the STFT, confirming that low-rank models are too constraining for STFT matrices. Yet, the approximation error being directly related to the singular values as explained above, one can see that the STFT matrix is somehow low-rank approximable.

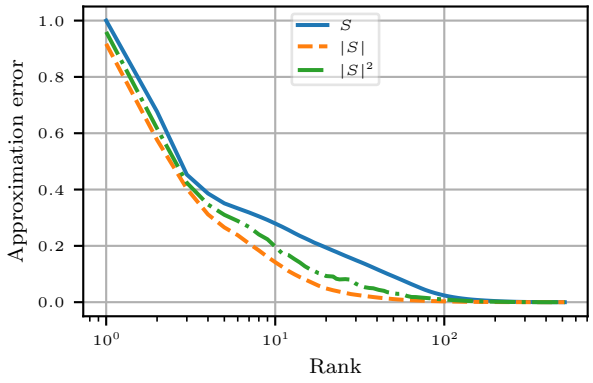

Fig. 4. *Glockenspiel* sound: normalized approximation error of the STFT representation when considering a low-rank decomposition of the STFT matrix S_{BP} (blue), the magnitude spectrogram (green) and the energy spectrogram (yellow).

4. CONCLUSIONS

In this paper, we have characterized exactly the set of low-rank matrices in a general context that allows to draw conclusions for low-rank models on this kind of data. First, this set of low-rank matrices is very narrow, which highly limits the modeling capacity. Second, it appears that the phase convention used to define the STFT is critical, and that the STFT of a mixture of sinusoids and dirac cannot be jointly described by a low-rank model, which is another major limitation. One may conclude that using a low-rank prior on complex-valued STFT matrices is not a good strategy. For instance, one may better design optimization problems with such matrices, and use a low-rank constraint on its magnitude or square magnitude.

However, we have shown that sum of sinusoids or diracs gives exact low-rank STFT matrices: it would be then possible to extend these results by designing low-rank models for local parts of the time-frequency plane, like patches in images, instead of using a low-rank prior on the full matrix. These study, despite the negative results when applied on real audio signals, provides nevertheless a better understanding of the low-rankness property of STFT matrices, and, to our opinion, brings real hope for the development of innovative structured models considering low-rank constraints in the time-frequency plane.

5. REFERENCES

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

[2] P. Balazs, M. Doerfler, M. Kowalski, and B. Torresani, "Adapted and adaptive linear time-frequency representations:

A synthesis point of view," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 20–31, Nov 2013.

[3] T. Strohmer, "Computational frameworks for discrete gabor analysis," in *In Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations VI.*, San Diego, 1997.

[4] P. L. Søndergaard, B. Torrésani, and P. Balazs, "The linear time frequency analysis toolbox," *Int. J. Wavelets Multidim. Inform. Proc.*, vol. 10, no. 4, 2012.

[5] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug 2009.

[6] P. Smaragdis and J.C. Brown, "Non-negative matrix factorization for polyphonic music transcription," in *Proc. WASPAA*, New Paltz, NY, USA, Oct. 2003, pp. 177–180.

[7] A. Ozerov and C. Févotte, "Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation," *IEEE Trans. Audio Speech Lang. Process.*, vol. 18, no. 3, pp. 550–563, Mar. 2010.

[8] C. Févotte and J. Idier, "Algorithms for nonnegative matrix factorization with the β-divergence," *Neural Comput.*, vol. 23, no. 9, pp. 2421–2456, 2011.

[9] P. Smaragdis, B. Raj, and M. Shashanka, "Missing data imputation for spectral audio signals," in *Proc. of MLSP*, Grenoble, France, Sept. 2009.

[10] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, Sept. 1936.