

Using growth transform dynamical systems for spatio-temporal data sonification

Oindrila Chatterjee, *Student Member, IEEE*, and Shantanu Chakrabartty, *Senior Member, IEEE*
Washington University in St. Louis, Missouri 63130, USA.

I. ABSTRACT

Sonification, or encoding information in meaningful audio signatures, has several advantages in augmenting or replacing traditional visualization methods for human-in-the-loop decision-making. Standard sonification methods reported in the literature involve either (i) using only a subset of the variables, or (ii) first solving a learning task on the data and then mapping the output to an audio waveform, which is utilized by the end-user to make a decision. This paper presents a novel framework for sonifying high-dimensional data using a complex growth transform dynamical system model where both the learning (or, more generally, optimization) and the sonification processes are integrated together. Our algorithm takes as input the data and optimization parameters underlying the learning or prediction task and combines it with the psychoacoustic parameters defined by the user. As a result, the proposed framework outputs binaural audio signatures that not only encode some statistical properties of the high-dimensional data but also reveal the underlying complexity of the optimization/learning process. Along with extensive experiments using synthetic datasets, we demonstrate the framework on sonifying Electro-encephalogram (EEG) data with the potential for detecting epileptic seizures in pediatric patients.

II. INTRODUCTION

With the multitude of high-dimensional data available today, the search for better techniques for perceptualizing data before applying it to a learning/predictive task has emerged as an important research area. Though visualization remains the modality of choice for most applications, sonification (representation of data using human-recognizable audio signatures [1]) is gaining momentum as a complementary or alternative modality.

Some of the advantages of data sonification, that make it an ideal candidate for augmenting or replacing visualization for analyzing the dataset's properties, are as follows [1]–[3]:

- 1) Temporal resolution for auditory perception is better than that for visual perception. This makes it suitable for perceptualizing time-varying data with complex patterns that visual displays might otherwise miss.
- 2) Audio is orientation-agnostic and has a wider spatial range since the user need not be oriented towards a particular direction. In contrast, for visualization, objects need to be within the field of vision of the user.
- 3) Humans typically respond faster to auditory feedback when compared to visual feedback, making sonification attractive for human-in-the-loop control applications.
- 4) In scenarios where the visual scene is crowded due to multiple displays or when attention is lacking, auditory signals can be used as a means of drawing the user's attention to a particular segment of the visual field.
- 5) Auditory perception provides a natural alternative to shrinking display sizes, especially for monitoring and alerting applications.
- 6) Auditory perception and auditory memory remain resilient to many neurodegenerative diseases, as a result sonification is an attractive choice for rehabilitation technologies.

The most common sonification scheme used in literature involves mapping only a subset of the most important variables in the data directly within the audible range [4], [5]. A second strategy involves using a machine learning algorithm to perform a classification task and then sonifying the output of the classifier [6]. While the first approach does not utilize the entire information available at hand, the second approach lacks a human-in-the-loop component that might be crucial for tasks involving monitoring and feedback. In this paper, we propose a novel sonification

All correspondences regarding this manuscript should be addressed to shantanu@wustl.edu.

This work was supported in part by a research grant from the National Science Foundation (ECCS:1935073).

technique that incorporates both the learning task at hand, as well as psychoacoustic parameters defined by the end-user in the form of constraints placed on the task. The algorithm finally produces an audio signature that can be utilized by the end-user for identifying both spatial and temporal patterns in the dataset. The scope of the paper is illustrated in Fig 1. At the core of the approach is a variant of the complex growth transform dynamical system, proposed in our previous work [7], [8] based on Baum-Eagon growth transforms [9]. Effectively, the high-dimensional space encompassed by the learning task is projected into a network of interacting limit cycle oscillators that form a set of complex basis functions encoding the low-dimensional space of the audio signal. The main model along with its properties is presented in Section IV, while Section IV-A shows the effects of the learning task on the sonified signal. Section IV-B presents the effects of the psychoacoustic parameters and sonification strategies that can be adopted for the sonification module. Section V presents the application of the sonification framework on synthetic data, while Section VI presents the application of the framework on real-world data. Section VI, finally, presents the application of the framework on real-world data for anomaly detection of

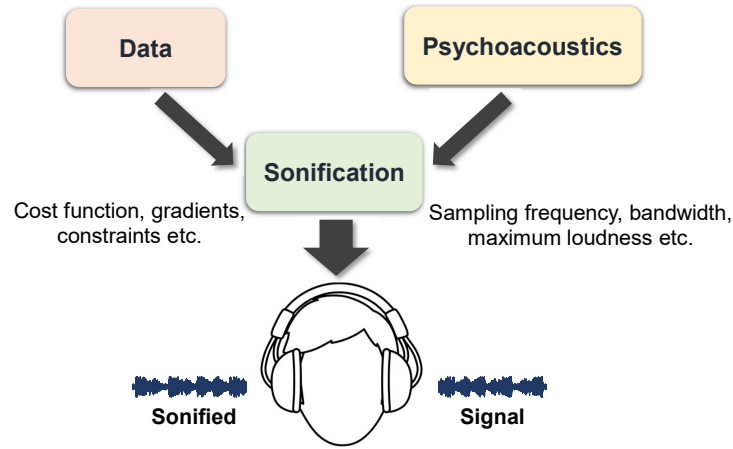


Figure 1: **Proposed approach for data sonification.** The proposed sonification module takes as input (i) a generic optimization/learning problem defined by a cost function, gradients, and constraints on one hand, and (ii) psychoacoustic parameters like sampling frequency etc. on the other. It then outputs a single or dual-channel audio waveform that encodes the underlying optimization problem.

In summary, the key contributions of the paper can be highlighted as follows:

- 1) The proposed method inherently maps high-dimensional data into a lower-dimensional single-channel or dual-channel audio signal. This is particularly suitable when dealing with low throughput systems having limited capacity/ bandwidth.
- 2) The method utilizes the whole spectrum of information available for a task and enables the end-user to arrive at a decision instead of relying on a machine learning model.
- 3) The algorithm automatically combines the learning and sonification stages into a single module. This is unlike the sonification-based decision-making algorithms existing in literature, that typically involve learning the decision parameters and then mapping the same to different sound parameters in subsequent stages.
- 4) The method is versatile and can be tailored to accommodate various learning models, problem dimensionality, and dataset sizes. Additionally, the complex growth transform dynamical system provides a wide range of tunable parameters which can be customized for different applications.

III. BACKGROUND AND RELATED WORK

1) *Visualization of High-Dimensional Data:* Elementary visualization schemes use different visual attributes like color, shape, size, spatial location, etc., for representing information. However, almost all the standard visualization techniques typically use these attributes after mapping the data to a lower-dimensional (usually two or three-dimensional) space. Some of the most commonly used mapping techniques include linear methods like PCA, LDA

and their variants [12] that only preserve the global structure of the data. In contrast, nonlinear methods like Locally Linear Embedding, [13], Isomap [14], Laplacian EigenMap [15], tSNE [16], PHATE [17] etc, have been designed to preserve both global and local information.

However, these techniques are more suitable for time-invariant or static data, and analysis of time-varying, high-dimensional data using these methods would involve frame-wise visualization in a lower-dimensional space. Several visualization methods based on variants of the techniques discussed earlier have been proposed, e.g., m-tSNE [18]. However, they lack interpretability and suffer from a similar limitation of being restricted to only three dimensions for representing the data. Moreover, though other visual attributes like color, size, shape, etc., can be used in conjunction with the three spatial dimensions, these may not be sufficient for complex datasets with very high dimensionality. This can thus potentially lead to information overload of the visual system, thereby pushing the users to the limits of comprehension. Additionally, complex, time-varying patterns in the data are sometimes difficult to capture using only the visual faculties. Hence, employing sonification might help users interpret the data more effectively or draw their attention to the most important aspects of the data that can then be monitored/analyzed by visual inspection [1].

2) *Data Sonification and its Applications:* Sonification, similar to its visual counterpart, provides several attributes that can be used for representing information, like pitch, loudness, timbre, rhythm, duration, harmonic content, etc. Broadly, two different sonification methods exist in the literature depending on how the data is mapped into an audio signal: (i) parameter mapping sonification and (ii) model-based sonification [1]. Of these, the parameter mapping method, where variables can be mapped into different attributes of the sound signal to create a sonified signature, is the most commonly used. Parameter mapping based sonification has been used in a wide range of applications, ranging from sonifying astronomical data like gravitational waves [19] and photons emitted by the Higgs bosons [20], to synthesizing novel protein structures [21], [22] and detecting anomalies in medical data (e.g., CT scans of Alzheimer’s patients [23], EEG [24] and ECG [4] signals, skin cancer detection [25], epileptic seizure detection [5]), detection of anomalous events [26] etc. Sonification based techniques have also been used for intrusion detection in networks [27] and network traffic flow [28], analysis of stock market data [29], and in therapeutic treatment of freezing of gait in Parkinson’s patients [30]. More recently, sonification has been applied to the analysis of RNA sequences in different strains of the Covid19 virus [31]. Unlike the visualization schemes, sonification offers a much wider parameter space for the variables to be mapped into. For example, humans can perceptualize sound signals anywhere from 20 to 20kHz, and frequency differences as low as 3 Hz are easily discernible by the human auditory system [1]. In contrast, the model-based sonification technique involves using a virtual model whose sonic responses are altered according to the data provided. This might involve first passing the data through a machine learning or optimization module which learns the manifold the data resides in, and then mapping the output to different properties of the sound signal. While the parameter mapping based method is restricted in terms of the dimensionality of the data it can handle, the model-based approach uses the entire data available, but does not rely on the end user for the decision-making [1].

IV. SONIFICATION FRAMEWORK USING COMPLEX GROWTH TRANSFORMS

In this section, we introduce a novel framework for sonifying high-dimensional data using a variant of the complex domain dynamical system model proposed in our previous work [8]. Our method utilizes the entire information available at hand to perform a combination of feature extraction and dimensionality reduction. Finally, it allows the human user to arrive at a decision based on the output sound signature. The sonification algorithm has been summarized in Table 2, while the proof has been outlined in VIII-B. The inputs to the sonification process are: (i) the learning parameters and data under consideration; and (ii) the psychoacoustic parameters defined by the user. We then develop a time-evolution operator \mathcal{U} that depends on both the learning task and the psychoacoustics by incorporating the latter as additional constraints on the learning/optimization process. \mathcal{U} is chosen to be a nonlinear unitary transformation that ensures that the total energy of the sonified signal remains conserved over time. This effect mimics an automatic gain control mechanism observed in bioacoustics [32].

Our proposed technique is thus a blend of both the parameter mapping-based and model-based sonification strategies. This is because the entire range of the data is mapped to different parameters of the output sound signal based on a dynamical system model. However, the end-user takes the decision, as illustrated in Fig 1. Within the framework, each variable or data point is mapped to a set of complex growth transform limit cycle oscillators globally coupled together by the conservation constraint on the signal energy. The oscillators can be thought of

as complex basis functions defining the lower-dimensional space of the audio output to which we want to project our high-dimensional data, and are dictated by the psychoacoustic parameters. This is illustrated in Fig 2 for a group of K oscillators represented by complex oscillator variables or waveforms $\psi_1(t), \dots, \psi_K(t)$ having different frequencies and amplitudes. The variables involved are assigned both baseline and relative frequencies as outlined in Appendix VIII-B. As a result of the time evolution of the complex dynamical system, both the amplitudes and frequencies of the oscillators get modulated during the optimization process. For example, if all the variables get mapped to the same constant value of baseline and relative frequencies, each oscillator frequency drifts from this constant value during the transient phase, following which it again returns to the original trajectory. The frequency deviation during the transient phase should ideally encode the complexity of the optimization problem. The output

wa

Visualization

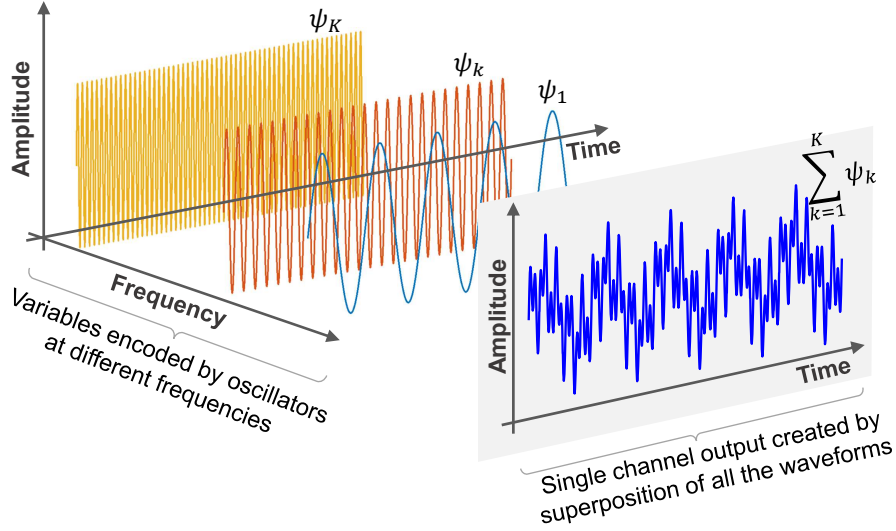


Figure 2: **Illustration of the sonification process.** K oscillator waveforms ψ_1, \dots, ψ_K , representing the set of complex basis functions defined by the user, based on psychoacoustics. Both amplitude and frequency of the basis set oscillators get modulated over time based on the learning task. The sonified output is created by the superimposition of all the oscillator signals, i.e., $\psi_{\text{sum}}(t) = \sum_{k=1}^K \psi_k(t)$.

A. Effect of the learning problem on sonified output

In this section, we will show how the learning problem affects the sonified signal. Any statistical learning task can be expressed in terms of an objective function \mathcal{H} that we are trying to minimize over a constraint space. We will therefore investigate the role of the objective function on the sound signature. Considering default temporal constraints on the total signal energy, the complex growth transform-based sonification technique (please see Table 2) provides a natural framework for optimizing the cost function \mathcal{H} subject to constraints defined by the psychoacoustics. Without loss of generality, in this section, we will focus only on the effect of \mathcal{H} and keep the psychoacoustic parameters constant for all the experiments. In particular, we will consider here that (a) the total energy of the sonified signal remains conserved over time, and (b) all the oscillators or basis waveforms in the network have identical baseline frequencies for better visualization of the network dynamics. The sonified output signal for all cases, as stated earlier, is obtained by the superposition of the wave functions generated by all the oscillators, i.e., $\Psi_{\text{sum},n} = \sum_{k=1}^K \Psi_{k,n}$.

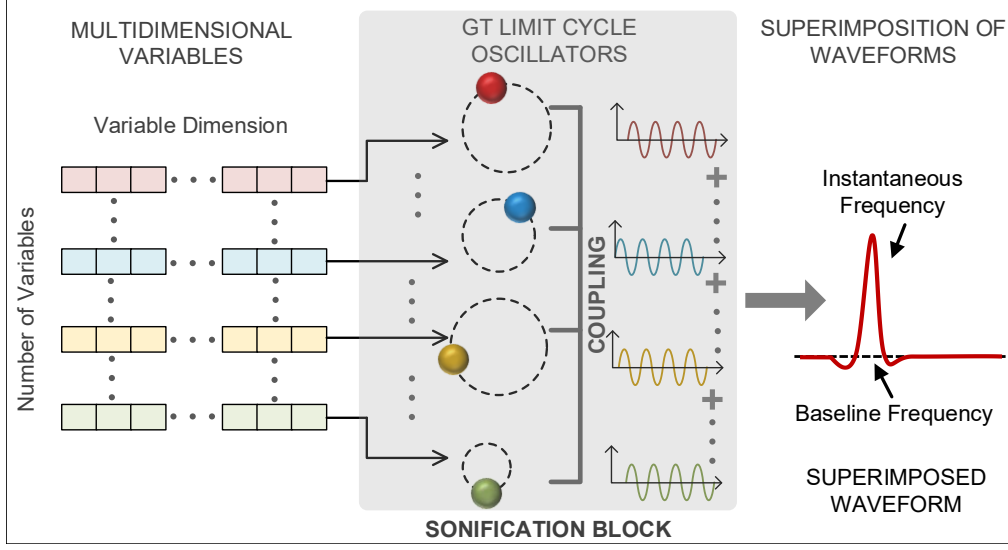


Figure 3: **Illustration of the superimposition of oscillators.** Each multidimensional variable/data vector is mapped into an individual growth transform limit cycle oscillator. Local and global couplings lead to a constant phase difference between each pair of oscillators in a steady-state. This results in a sudden shift of the frequency of the superimposed waveform $\psi_{\text{sum}}(t)$ from the baseline frequency during the transient state, following which the frequency is restored to its baseline value in the steady-state.

Algorithm 2: Sonification using complex growth transform dynamical system (Proof in Appendix VIII-B)

- **Learning task:** Consider a statistical learning task defined by a cost function and gradients.
- **Psychoacoustics:** Consider a set of complex basis functions $\Psi = [\psi_1, \dots, \psi_K] \in \mathbb{C}^K$ defined by the user using psychoacoustic parameters like sampling frequency F_s , bandwidth (f_1, f_2) , frequency mapping strategy etc.
- **Sonification module:** The growth transform-based sonification method produces an output signal of the following form in steady-state:

$$\psi_{\text{sum},n} = \sum_{k=1}^K \psi_{k,n} = \sum_{k=1}^K \psi_{k,n-1} \exp \left(j(\omega_n + \xi_{k,n}) \frac{1}{F_s} \right), \quad (1)$$

where ω_n and $\xi_{k,n}$ are the instantaneous baseline and relative frequencies of the k^{th} oscillator variable $\psi_{k,n} \in \mathbb{C}$ at the n^{th} time step respectively.

- **Update rule:** The corresponding instantaneous time evolution equation for the oscillators is given by the following equation:

$$\Psi_n \leftarrow \mathcal{U}_n(\Psi_{n-1}) \odot \Psi_{n-1}, \quad (2)$$

where the mathematical form of \mathcal{U}_n is given in Appendix VIII-B.

- **Conservation of signal power:** The operator $\mathcal{U}_n : \mathbb{C}^K \rightarrow \mathbb{C}^K$ represents an instantaneous nonlinear unitary transformation that ensures that the total signal energy remains conserved over time and thus imposes a temporal constraint brought about by the psychoacoustics as well, i.e.,

$$\sum_{k=1}^K |\psi_{k,n}|^2 = \gamma \quad \forall n, \quad \gamma > 0. \quad (3)$$

Example 1: Consider the following one-dimensional quadratic optimization problem:

$$\begin{aligned} & \underset{p \in \mathbb{R}}{\text{minimize}} \quad \mathcal{H}'_1(p) = 8p^2 - 2p \\ & \text{s.t.} \quad |p| \leq \gamma \end{aligned} \quad (4)$$

We can map the above to the sonification framework in Table 2 by taking $p = |\psi_1|^2 - |\psi_2|^2$, where $\psi_1, \psi_2 \in \mathbb{C}$ are complex waveforms. (Please see Appendix VIII-B for a detailed outline of the mapping procedure.) The equivalent optimization problem in the complex domain thus becomes:

$$\begin{aligned} \underset{\Psi \in \mathbb{C}^2}{\text{minimize}} \quad & \mathcal{H}_1(\Psi) = 8(|\psi_1|^2 - |\psi_2|^2)^2 - 2(|\psi_1|^2 - |\psi_2|^2)^2 \\ \text{s.t.} \quad & |\psi_1|^2 + |\psi_2|^2 = \gamma, \end{aligned} \quad (5)$$

where $\Psi = [\psi_1, \psi_2]$. For this simple problem, the expression for the unitary operator \mathcal{U} using the procedure outlined in Table 2 (please see Appendix VIII-B for a detailed derivation) is given by:

$$\begin{aligned} \mathcal{U}_n &= \begin{pmatrix} d_{1,n} \\ d_{2,n} \end{pmatrix}, \\ d_{k,n} &= \left[\cos\left(\frac{\omega_n}{F_s}\right) + j \sin\left(\frac{\omega_n}{F_s}\right) \sigma_{k,n-1} \right] \exp\left(j \frac{\xi_{k,n}}{F_s}\right), \\ \sigma_{k,n-1} &= \sqrt{\frac{\left(-\frac{\partial \mathcal{H}}{\partial \psi_{k,n-1}} + \lambda \psi_{k,n-1}^*\right)}{\psi_{k,n-1}^* \sum_{l=1}^{K=2} \psi_{l,n-1} \left(-\frac{\partial \mathcal{H}}{\partial \psi_{l,n-1}} + \lambda \psi_{l,n-1}^*\right)}}, \quad k = 1, 2. \end{aligned} \quad (6)$$

Fig 4 shows the polar coordinate evolution plots of the waveforms ψ_1 and ψ_2 respectively, for two different choices of the parameter γ : (a) $\gamma = 1$ and (b) $\gamma = 0.01$. It can be seen that the oscillators corresponding to ψ_1 and ψ_2 converge to steady limit cycle oscillations for larger values of γ . However, for very small γ values, only the oscillator with the maximum amplitude shows sustained oscillations, while the other converges to the fixed point of zero and stops oscillating. The coupled oscillator network can thus be thought of as some form of a frequency tuner for sufficiently small γ .

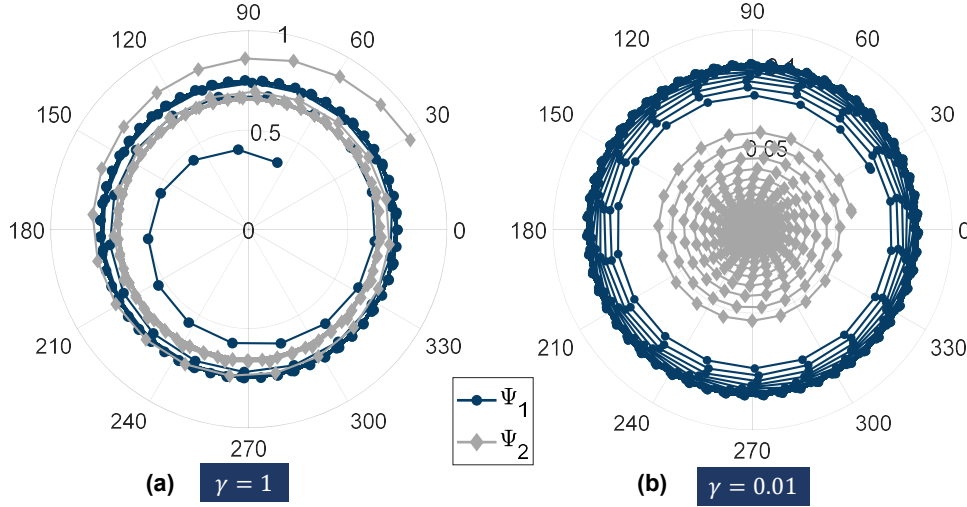


Figure 4: **Effect of γ .** Dynamics of growth transform-based complex dynamical system consisting of two oscillators for two different cases corresponding to problem \mathcal{H}_1 : (a) both oscillators asymptotically reach their respective stable limit cycles in steady-state ($\gamma = 1$) and (b) one reaches a stable limit cycle, while the other goes to a stationary point ($\gamma = 0.01$).

Example 2: Consider a second quadratic optimization problem:

$$\begin{aligned} \underset{\mathbf{p} \in \mathbb{R}^M}{\text{minimize}} \quad & \mathcal{H}'_2(\mathbf{p}) = \frac{1}{2} \mathbf{a} \mathbf{p}^T \mathbf{Q} \mathbf{p} - \mathbf{c}^T \mathbf{p} \\ \text{s.t.} \quad & |p_i| \leq \gamma \quad \forall i = 1, \dots, M, \quad \gamma \in \mathbb{R}_+, \end{aligned} \quad (7)$$

and $\mathbf{c} \in \mathbb{R}^M$, $\mathbf{Q} \in \mathbb{R}^{M \times M}$. $a \in \mathbb{R}_+$ is a parameter that controls the curvature of the optimization surface, and hence controls the rate of convergence to the optimal solution. Taking $p_i = |\psi_{i1}|^2 - |\psi_{i2}|^2$, where $\psi_{i1}, \psi_{i2} \in \mathbb{C}$, we can again follow the mapping procedure in Appendix VIII-C. The equivalent optimization problem in the complex domain is given by:

$$\underset{\Psi_1, \Psi_2 \in \mathbb{C}^M}{\text{minimize}} \quad \mathcal{H}_2(\Psi_1, \Psi_2) = \frac{1}{2}a(|\Psi_1|^2 - |\Psi_2|^2)^T \mathbf{Q}(|\Psi_1|^2 - |\Psi_2|^2) \quad (8)$$

$$- \mathbf{c}^T(|\Psi_1|^2 - |\Psi_2|^2) \\ \text{s.t.} \quad |\psi_{i1}|^2 + |\psi_{i2}|^2 = \gamma \quad \forall i = 1, \dots, M, \quad \gamma \in \mathbb{R}_+, \quad (9)$$

where $\Psi_1 = \{\psi_{11}, \dots, \psi_{i1}, \dots, \psi_{M1}\}$ and $\Psi_2 = \{\psi_{12}, \dots, \psi_{i2}, \dots, \psi_{M2}\}$. Again, defining the unitary operator \mathcal{U} according to Appendix VIII-B, we have:

$$\mathcal{U}_n = \begin{pmatrix} d_{11,n} & d_{12,n} \\ \vdots & \vdots \\ d_{i1,n} & d_{i2,n} \\ \vdots & \vdots \\ d_{M1,n} & d_{M2,n} \end{pmatrix}, \quad \mathcal{U}_n : \mathbb{C}^{M \times 2} \mapsto \mathbb{C}^{M \times 2} \\ d_{ik,n} = \left[\cos\left(\frac{\omega_{i,n}}{F_s}\right) + j \sin\left(\frac{\omega_{i,n}}{F_s}\right) \sigma_{ik,n-1} \right] \exp\left(j \frac{\xi_{ik,n}}{F_s}\right), \\ \sigma_{ik,n-1} = \sqrt{\frac{\left(-\frac{\partial \mathcal{H}}{\partial \psi_{ik,n-1}} + \lambda \psi_{ik,n-1}^*\right)}{\psi_{ik,n-1}^* \sum_{l=1}^{K=2} \psi_{il,n-1} \left(-\frac{\partial \mathcal{H}}{\partial \psi_{il,n-1}} + \lambda \psi_{il,n-1}^*\right)}}, \quad i = 1, \dots, M; k = 1, 2. \quad (10)$$

The final sonified output is then given by $\psi_{\text{sum},n} = \sum_{i=1}^M [\psi_{i1,n} + \psi_{i2,n}]$.

Time evolution and phase portrait of the sonified signal: Fig 5 shows the time evolutions and phase portrait for a one-dimensional problem ($M = 1$), with $a = 1$, $\mathbf{Q} = \mathbf{I}$, $\mathbf{c} = 0.8$ and $\gamma = 2$. The sampling and natural/baseline frequencies were considered to be $F_s = 22\text{kHz}$ and $\omega_n = 600\text{Hz}$ respectively, while the relative frequency was considered to be $\xi_{ik,n} = \xi_k = 0$ for all the oscillators. In the sonification framework, $t = n\Delta t$, where t seconds is the time duration corresponding to the n -th time step, and each time step $\Delta t = (t + \Delta t) - \Delta t = (n+1)\Delta t - n\Delta t = (1/F_s)\text{s}$. The simulation duration was 0.1s, with the optimization onset being at 0.03s for same initial amplitudes of ψ_1 and ψ_2 , and zero initial phase difference between the waveforms. Fig 5(a) shows the time evolutions for the waveforms $\psi_{11} = \psi_1$, $\psi_{12} = \psi_2$ and ψ_{sum} , along with their zoomed-in views during the initial (I), transient (T) and steady-state (S) stages respectively. Fig 5(b) shows the phase portrait between the real parts of ψ_1 and ψ_2 . It can be seen that the two waveforms evolve over time from an initial state (I) of same amplitude and zero phase difference, through a transient phase (T) of varying amplitudes and phase difference, to a final steady state where ψ_1 and ψ_2 have constant values of amplitudes, and a constant phase difference between them. This indicates in a non-zero phase shift of the final trajectory of $\psi_{\text{sum}}(t)$ from the initial trajectory.

Effect of \mathcal{H} and γ on the sonified signal: Next, we investigate the performance of the model for different values of curvature of the optimization surface as shown in Fig 6, as well as different levels of total energy available to the system at any instant of time. Fig 7 demonstrates the model behavior for a two-dimensional variant of Example 2 (i.e., $M = 2$). As before, we consider all the four oscillators in this case, namely, $\psi_{11}, \psi_{12}, \psi_{21}, \psi_{22}$ to have the same value of natural frequency $\omega = 600\text{Hz}$ and relative frequency $\xi_{ik} = \xi_k = 0$. Simulations were carried out with $\mathbf{Q} = \mathbf{I}_2$ and $\mathbf{c} = 0.8\mathbf{1}_2$. Figs 7(a)-(c) show the spectrograms for the waveform $\Psi_{\text{sum}}(t)$, with $a = 0.5, 1$ and 2 respectively, and $\gamma = 1$. In all cases, the spectrogram was computed using a 1024-point Short Time Fourier Transform (STFT) with a sliding Kaiser window of size 1024 and steepness parameter 5, with an overlap size of 1023. Figs 7(d)-(f) show the zoomed-in versions of the instantaneous frequency shift computed using Hilbert transform. Figs 7(g)-(i) show similar results for the same set of values for a and a higher value of $\gamma = 2$. The simulation duration is 1.0s, and the optimization process starts 0.1s after the start of the simulation. From the figures, we can conclude the following: (i) the duration of the transient phase as well as the magnitude of frequency shift

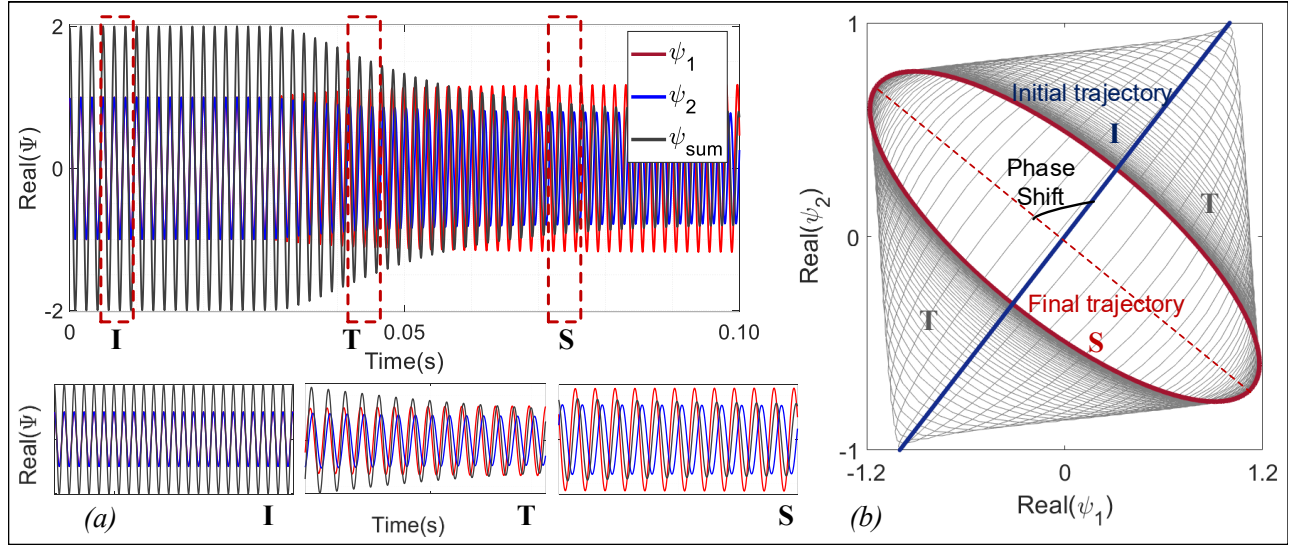


Figure 5: **Time evolution and phase portraits.** Time evolution and phase portrait for ψ_1, ψ_2 and Ψ_{sum} corresponding to a 1D variant of the optimization problem \mathcal{H}_2 : (a) Time evolution of the waveforms, with zoomed-in versions of the initial(I), transient(T) and steady-state(S) regions respectively; (b) Phase portrait of ψ_1 vs. ψ_2 . The oscillators go from an initial state(I) of the same amplitude and no phase difference, through a transient state(T) of varying amplitudes and varying phase difference, to a final steady-state(S) where they have constant amplitudes and a constant phase difference as well.

during the same decreases with an increase in the value of a , (i.e., a steeper curve implies a faster optimization process and hence smaller values of frequency deviation); (ii) higher value of the total energy γ leads to shorter transient phases with smaller frequency drifts. Thus, the complexity of the optimization problem is encoded in the final phase shift of the output signal from its initial phase.

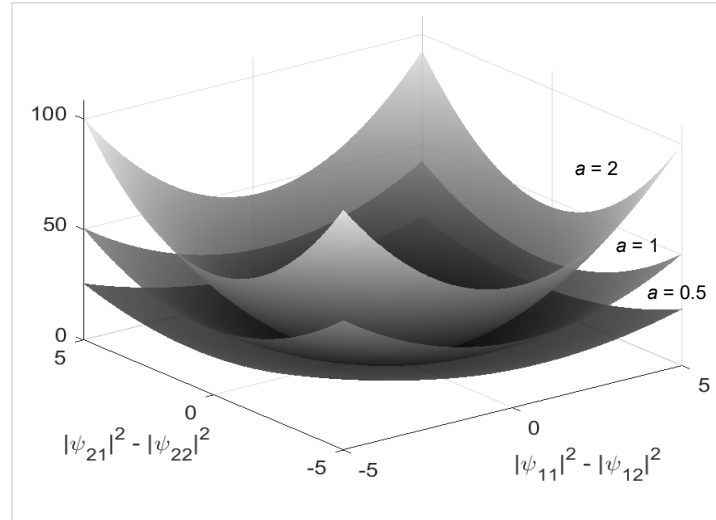


Figure 6: **Quadratic optimization surface.** Optimization surface for the 2D variant of the problem \mathcal{H}_2 for different values of the parameter a .

Effect of cost function-induced frequency perturbations on the sonified signal: Next, we examine the effect of variations in the relative frequencies on the system's performance. Fig 8 shows the spectrogram and instantaneous frequency plots for different choices of time evolution of the relative frequency $\xi_k(t) \forall k$, for the optimization problem in Example 2 and same parameter settings as used in the simulations in Fig 7. The baseline frequency was chosen to be $\omega_n = 100$ Hz for all the oscillators. Additionally, the total simulation duration and the onset of

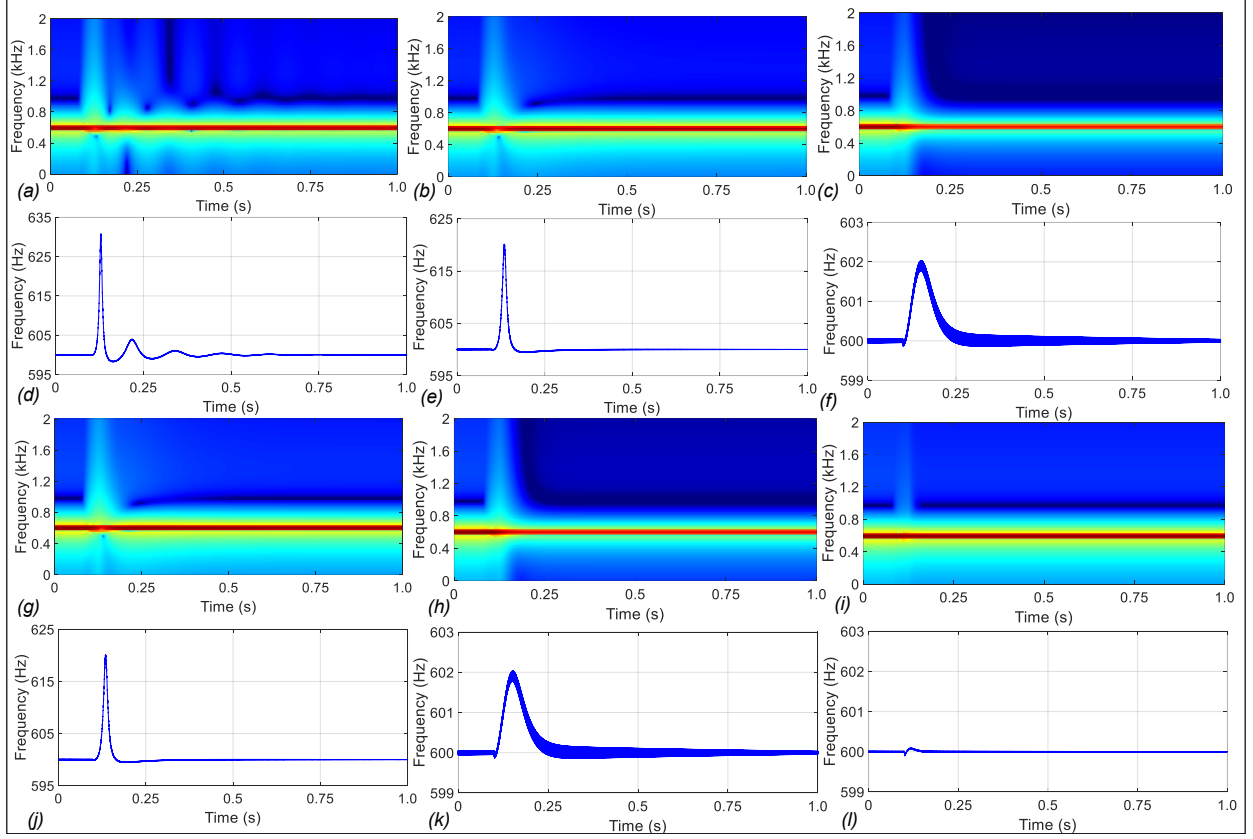


Figure 7: **Effect of \mathcal{H} and γ .** Variation of $\Psi_{\text{sum}}(t)$ with different values of curvature (a) and total energy available (γ) for a 2D variant of the optimization problem \mathcal{H}_2 : (a)-(c) show the spectrograms; (d)-(f) show the zoomed-in versions of the instantaneous frequency shift computed using Hilbert transform, for $a = 0.5, 1$ and 2 respectively, and $\gamma = 1$; (g)-(l) show similar plots for $\gamma = 2$.

optimization were 1s and 0.2s, respectively, in all of the following experiments.

Case 1 (Constant relative frequency): For the first set of simulations, the relative frequencies are chosen to be $\xi_k = \xi = 500$ Hz for all the oscillators, and no amplitude modulations or frequency modulations were added to the relative frequency trajectories. The plot of $\text{Re}(\exp(j\xi_k t))$, which is a unit amplitude sinusoid with a constant frequency for all the oscillators in this case, is shown in Fig 8(a). Figs 8(b) and (c) show the corresponding spectrogram and instantaneous frequency shift (computed using Hilbert transform) of the superimposed waveform $\psi_{\text{sum}}(t)$.

Case 2 (Relative frequency modulation): Figs 8(d)-(f) show similar results for the case when each relative frequency trajectory ξ_k is modulated by the corresponding value of σ_k , where $\sigma_k \rightarrow 1$ in steady state, i.e., $\xi_{k,n} = \xi \sigma_{k,n}$. This type of frequency modulation thus incorporates information about the optimization process in the output signal, and leads to faster convergence and larger frequency deviation compared to the output in Case 1.

Case 3 (Amplitude decay and constant relative frequency): In this case, a varying amplitude term constant for each oscillator was added to the relative frequency term, and no frequency modulation was added. The multiplicative term in the update equation is thus replaced by $\exp(j\xi_{k,n}\Delta t) \leftarrow \exp((- \rho + j\xi)\Delta t)$, where $\rho = 1$ was chosen for the experiment. Figs 8(g)-(i) show the relative frequency modulations, spectrogram, and frequency deviation respectively for this case.

Case 4 (Amplitude decay and frequency modulation): In the last set of experiments, both the amplitude and frequency of the relative frequency trajectories were modulated for all the oscillators, i.e., $\exp(j\xi_{k,n}\Delta t) \leftarrow \exp((- \rho + j\xi \sigma_{k,n-1})\Delta t)$. Figs 8(j)-(l) show the relative frequency modulations, spectrogram, and frequency deviation respectively for this case. In both Cases 3 and 4, the relative frequencies were applied only during the transient phase of the simulation, and set to zero both during the initial stage as well as after convergence to a steady-state.

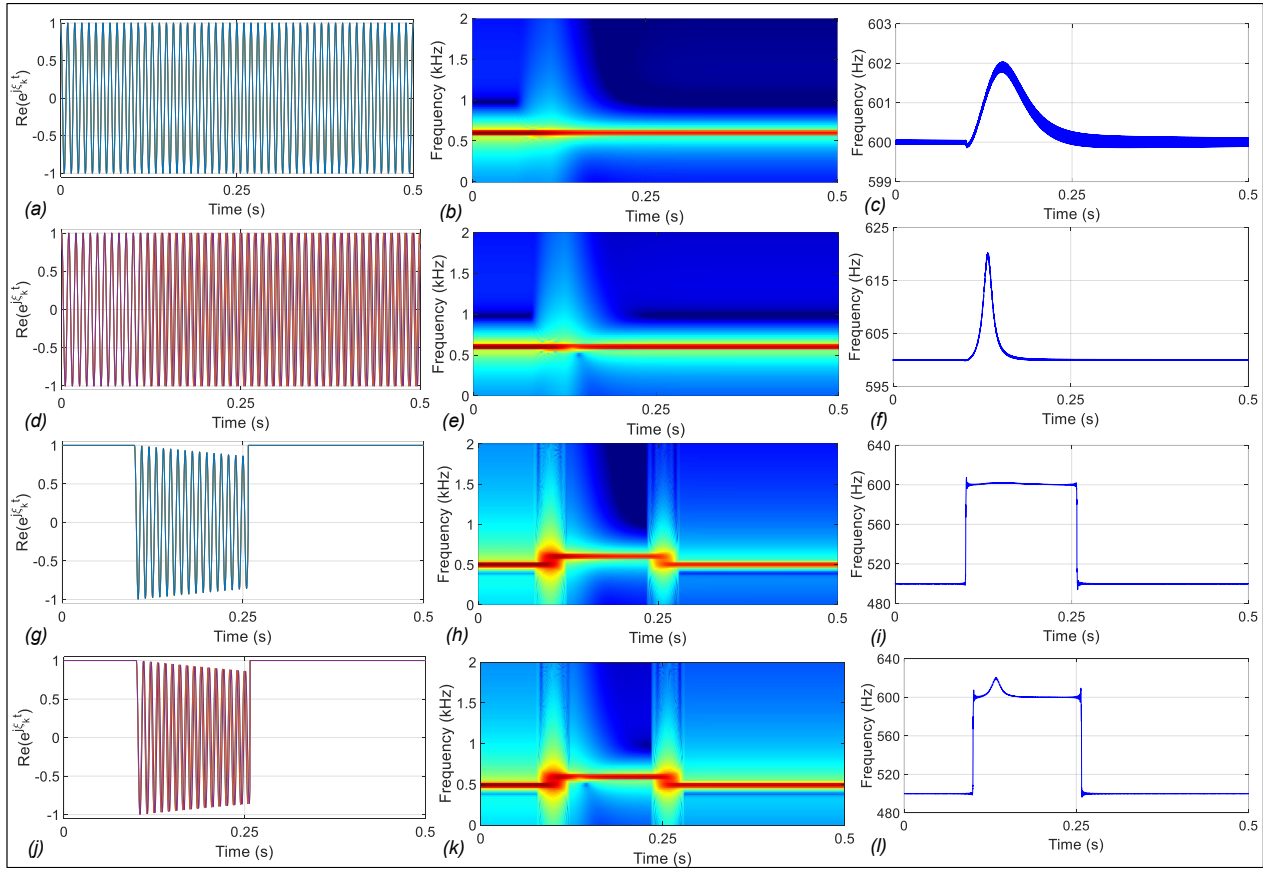


Figure 8: **Effect of frequency perturbations induced by \mathcal{H} .** Variation of $\Psi_{\text{total}}(t)$ for a 2D variant of the optimization problem \mathcal{H}_2 , for different choices of the relative frequency $\xi_k(t)$: (i) **Case 1:**(a) ξ_k is constant $\forall k, t$, (b) spectrogram of the output waveform, (c) instantaneous frequency shift using Hilbert transform ; (ii) **Case 2:** (d) value of $\xi_k(t)$ depends on the convergence rate of each variable $\psi_k(t)$, (e) spectrogram of the output waveform, (f) instantaneous frequency shift using Hilbert transform; (iii) **Case 3:** (g) $\xi_k(t)$ exists only during the optimization stage and is an exponentially decaying function which is the same for all ψ_k , (h) spectrogram of the output waveform, (i) instantaneous frequency shift ; (iv) **Case 4:** (g) $\xi_k(t)$ exists only during the optimization stage and is an exponentially decaying function which depends on the convergence rate of each variable ψ_k , (h) spectrogram of the output waveform, (i) instantaneous frequency shift.

It can be seen that the frequency shift during the transient phase in Case 4 is significantly higher than all the other stages.

Based on the above experiments, we can infer the following:

- The framework can encode information about the optimization problem's complexity and the total energy available to the network.
- Different encoding strategies can be employed by exploiting the relative frequencies of the oscillators.
- Frequency selectors/tuners can also be implemented by assigning sufficiently lower levels of energy to the oscillator network, such that only the oscillator with the maximum energy content is sustained.

B. Effect of Psychoacoustics on Sonified Output

On the psychoacoustics side, we can incorporate both temporal and spatial attributes, as well as different choices of the sampling frequency, frequency mapping schemes, bandwidth, etc. Table 1 presents a summary of the various psychoacoustic parameters at our disposal. A temporal constraint is imposed on the sonified output by default because of the inherent power normalization provided by the growth transform framework. The temporal constraint thus imparts an automatic gain control feature [32] on the sonified signal. Since the output signal is inherently

complex, spatial constraints can be imposed on the model using the left and right channels of the audio for the real and imaginary parts respectively, or vice versa. Finally, different sonification strategies can be adopted by considering different mapping schemes for both the baseline frequencies ω_i 's, and the relative frequencies ξ_k 's. In a sense, the sonification process can be thought of as projecting high-dimensional data into a low-dimensional basis space created by predetermined frequency trajectories. We will present here three different sonification strategies: (a) using frequencies equally spaced on the Bark-scale, (b) creating chords based on a musical scale of choice, and (c) extracting dominant frequency trajectories from a chosen musical piece and mapping the basis set of frequencies to these trajectories. For generating human recognizable auditory signatures, all the baseline and relative frequencies should lie within the range of human perception, i.e., 20 Hz-20 kHz. Furthermore, the largest frequency assigned to a variable should be less than the Nyquist frequency to avoid aliasing.

Table 1: Psychoacoustic Parameters

Constraint	Parameters
Temporal	Power normalization
Spatial	Binaural hearing
Frequency	Sampling frequency, frequency mapping, bandwidth

Some of the desirable characteristics of a candidate sonification strategy are as follows:

- 1) Different data distributions can be encoded by different sound signatures, depending on the underlying optimization task.
- 2) The complexity of the dataset or the underlying optimization problem affects the output sound signature.
- 3) For time-varying data, drift in the data distribution over time would lead to a drift in the sonified signal as well.

For example, if we consider clustering as the underlying optimization problem, where the number of allowable clusters (K) is fixed apriori, then we would ideally want the sonified output to give an indication of (a) the instantaneous cluster densities, (b) the orientation of the clusters and (c) the time it takes for the optimization problem to converge to the optimal cluster assignments.

1) *Bark scale-based sonification*: This method involves mapping the relative/baseline frequencies to equally spaced frequencies on the Bark scale. Any number of frequency trajectories can be selected if masking is acceptable in the end application. However, if we want to eliminate masking effects from the output sonified signal, the Bark scale frequencies should be chosen in a way such that the critical bands around these frequencies do not overlap with each other. Since the Bark scale has 24 critical bands, this implies that the number of frequencies in the basis set is limited to a maximum of 24 if we want to avoid masking. Additionally, the sonification module should be designed so that each frequency trajectory remains within its critical band on the Bark scale, even during the transient phase. This method's advantage is that changes in each frequency trajectory (i.e., each sonified variable) can be discerned unambiguously since there is no mixing of trajectories throughout the duration of sonification.

2) *Musical chord-based sonification*: This approach is similar to the Bark scale-based approach, with the relative (or baseline) frequencies being mapped to a predetermined musical scale, e.g., the equally-tempered Western scale. Depending on the user requirements, we may create a chord by choosing notes in a single octave as the basis set. We can also select notes over multiple octaves, associating a different timbre to each, giving the impression of multiple different instruments being played. For example, we can map the frequencies to every other note in a diatonic scale such that they form a triad (e.g., the notes C, E, and G form the C major triad in the equally-tempered scale around $A_4 = 440$ Hz). Similarly, four or more random notes from the same octave create a generic chord. The method also allows for the creation of arpeggios (or a broken chord being played multiple times in succession) by suitably defining periodic repetitions of the set of chords. Overlapping of the frequency trajectories corresponding to different sonification variables may or may not occur, depending on (a) the pairwise distances between the frequency trajectories and (b) the maximum extent of frequency perturbation caused by the sonification module. Additionally, since humans are more adept at recognizing time-varying audio signatures than static tones, a small, slowly varying sinusoidal variation may be added on top of the original frequency trajectories. The amplitudes and frequencies of

these sinusoidal variations may be either (a) kept constant, or made to vary based on (b) the convergence properties of the optimization problem, or on (c) the instantaneous statistical properties of the sonification variables.

3) *Sonification using an existing musical piece:* This method of sonification involves extracting a certain number of dominant frequency trajectories (depending on the desired size K of the basis set) from an existing musical piece. Additionally, the following steps need to be taken for extracting the dominant frequency trajectories from the chosen musical piece:

- 1) Typically, a musical piece may be much longer in duration than the simulation duration and has a much higher sampling rate (44 kHz). Thus we would need to extract a segment of the entire composition and compress its time scale to ensure a proper mapping.
- 2) Next, we analyze the spectrogram of the time-scaled musical sample and extract K dominant frequency components in each time window of the spectrogram, based on the power content of each in that window.
- 3) Finally, we carry out upsampling and interpolation to form continuous frequency trajectories that represent K most dominant components of the musical composition.

The sonification variables are then mapped to these frequency trajectories. The trajectories undergo perturbations in the transient phase during the sonification process, depending on the dataset complexity and the optimization problem being solved. The baseline frequency perturbations may remain unmodified or can be made a function of the convergence properties only. They can also be made to depend on certain statistical properties of the dataset or the optimization variables. Finally, for better interpretability, the sonification module's output can be treated as a "noise" signal and superimposed on the original musical composition. The degree of deviation of this superimposed signal from the original composition thus encodes the dataset's complexity and the optimization process. Fig 9 shows an overview of the musical composition-based sonification technique for a sample optimization problem, considering a basis set of size 3.

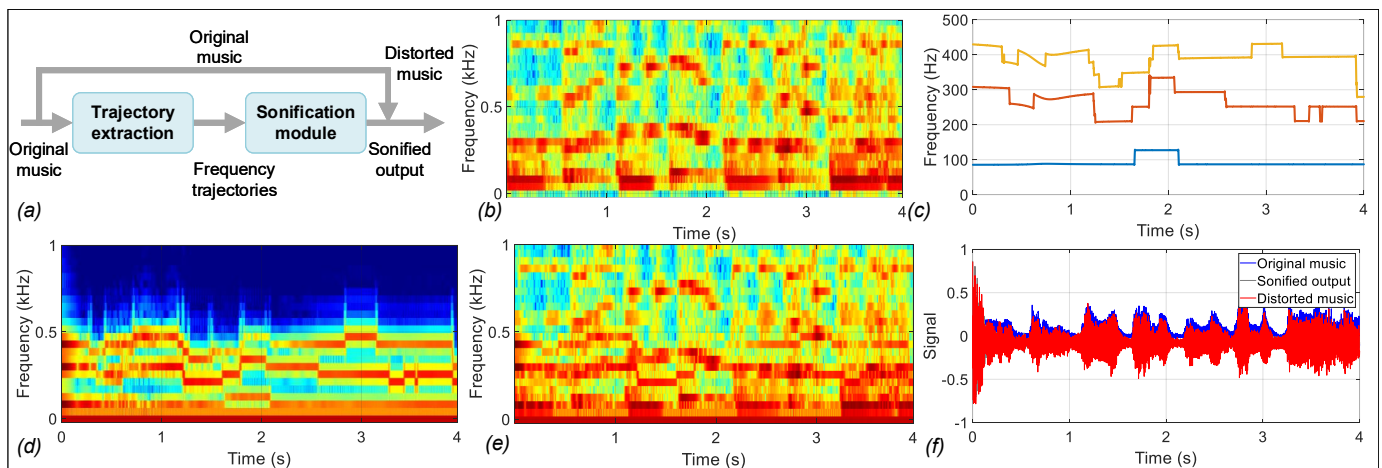


Figure 9: **Musical composition based sonification technique.** (a) Illustration of the sonification process in this technique; (b) Spectrogram of an original musical piece; (d) Raw frequency trajectories extracted from the music; (d) Spectrogram of the sonified output; (e) Spectrogram of the distorted music, obtained by superimposing the sonified output signal on the original music signal; and (f) Time evolution of the original musical composition, the sonified output, and the distorted music signal.

V. EXPERIMENTS ON SYNTHETIC DATASETS

As a case study, we will consider a data clustering problem. The goal is to sonify the data by mapping each cluster to a particular tone or frequency trajectory, with the amplitude of each trajectory being proportional to the cluster density. Since the focus of the paper is on advocating a new sonification framework and not improving the efficiency of the particular clustering algorithm chosen, for the sake of simplicity, we will consider a similarity-based probabilistic clustering approach proposed in [33]. This involves solving a non-negative matrix factorization problem that minimizes the distance between a similarity matrix computed pairwise between the data points, and the actual likelihoods of the different data points to be clustered together in space. Consider a D -dimensional

dataset $\mathbf{X} \in \mathbb{R}^{M \times D}$ and a similarity matrix $\mathbf{W} \in \mathbb{R}_+^{M \times M}$ computed using a pairwise distance metric between the data points. Then the following optimization problem assigns to the i -th data point a K -dimensional vector \mathbf{p}_i consists of probability values of the data point of belonging to each of K possible clusters:

$$\begin{aligned} & \underset{\mathbf{P} \in \mathbb{R}_+^{M \times K}, \beta \in \mathbb{R}_+}{\text{minimize}} && \mathcal{H}(\mathbf{P}, \beta) = \|\mathbf{W} - \beta \mathbf{P}^T \mathbf{P}\|_2^2 \\ & \text{s.t.} && \sum_{k=1}^K p_{ik} = 1 \quad \forall i = 1, \dots, M \end{aligned} \quad (11)$$

Thus, $p_{ik} \in \mathbb{R}_+$ denotes the probability of the i th data point of belonging to the k th cluster, and β denotes a scaling factor such that $\beta \mathbf{p}_i^T \mathbf{p}_j$ represents the true likelihood of the i th and j -th data points to be clustered together. In this approach, each w_{ij} is assumed to be normally distributed about its corresponding true likelihood $\beta \mathbf{p}_i^T \mathbf{p}_j$ with a constant variance. Here, the similarity matrix \mathbf{W} is chosen to be the RBF kernel computed pairwise between the data points by mapping them to a high-dimensional space. We can use the procedure outlined in Table 2 by applying the mapping $p_{ik} = |\psi_{ik}|^2$, $\psi_{ik} \in \mathbb{C}$. Sonification of the clustering problem can then be achieved in the following manner:

- We assign the same baseline frequency to all the subgroups (individual data points in this case), i.e., $\omega_i = \omega \quad \forall i$.
- Each cluster is assigned to a particular relative frequency trajectory according to the sonification strategy chosen, i.e., $\xi_{ik} = \xi_k \quad \forall i$. Depending on the sonification strategy, the unperturbed version of the relative frequencies $\xi_k(t)$'s may be chosen as follows:

(a) *Bark scale based*: Each of the K clusters is mapped to a distinct frequency on the Bark scale (with or without masking effects, depending on the frequency spacing). A slow sinusoidal variation may be added about each original frequency trajectory depending on the instantaneous cluster density, as well as the convergence characteristics of the optimization problem. The relative frequencies are thus modulated over time governed by the following equation:

$$\begin{aligned} \xi_k(t) &\leftarrow \xi_k(t)[1 + a_k \sin(2\pi b_k \Delta f_k(t)t)] s_k, \\ \text{where } \Delta f_k(t) &= K \frac{c_k(t) - c_k(0)}{c_k(0)}, \quad s_k = \frac{1}{N} \sum_{i=1}^M \sigma_{ik}. \end{aligned} \quad (12)$$

$a_k, b_k \in \mathbb{R}_+$ are constants, $c_k(0) = M/K$ is the initial cluster density (assuming the data points to be uniformly distributed among the clusters), and $c_k(t) = \sum_{i=1}^M |\psi_{ik}|^2$ is the instantaneous cluster density of the k -th cluster.

(b) *Musical chord-based*: Each of the K clusters is mapped to a distinct frequency on a chosen musical scale. In this case, too, a slow sinusoidal variation may be added about each original frequency trajectory depending on the instantaneous cluster density and the convergence characteristics of the optimization problem. The evolution equations for the $\xi_k(t)$'s are similar to those used for the Bark scale-based method.

(c) *Using an existing musical composition*: Each of the K clusters is mapped to a distinct frequency trajectory extracted from a musical composition. The instantaneous statistical properties of each cluster and the convergence properties of the optimization variables may be used as a scaling factor, to enhance the perturbations caused by the growth transform optimization framework. In this case, the evolution equations may have a sinusoidal variation as in the previous two approaches, or maybe of the following form:

$$\xi_k(t) \leftarrow \xi_k(t)[1 + a_k \Delta f_k(t)] s_k, \quad (13)$$

where $a_k, \Delta f_k(t)$ and s_k have the same definitions as before. This is because the original frequency trajectories vary over time, and hence an additional sinusoidal perturbation for recognizing the audio signature in the steady state might not be necessary.

- The update equation for β is given by:

$$\beta \leftarrow \frac{\text{Tr}(\mathbf{W} \mathbf{P}^T \mathbf{P})}{\|\mathbf{P}^T \mathbf{P}\|^2}, \quad (14)$$

while those of the complex waveforms ψ_{ik} 's are obtained using the complex growth transform updates in Table 2 (please see Appendix VIII-B for details). The time evolution of $\xi_k(t) \forall k$ occurs according to the update rules in (12) or (13), depending on the chosen sonification strategy.

- The final output of the sonification module is obtained by superimposing the waveforms of all the oscillators, i.e., $\psi_{\text{sum}}(t) = \sum_{i=1}^M \sum_{k=1}^K \psi_{ik}(t)$.

Next, we demonstrate the properties of the sonification technique when applied to the clustering problem for different sonification strategies, the actual number of clusters in the dataset, dataset complexity (i.e., cluster alignments), and the number of clusters assigned apriori (i.e., the size of the basis set of frequencies K).

Effect of different frequency mapping strategies: Fig 10 shows the results on the 'Iris' dataset ($M = 150$), considering a basis of size $K = 3$, for all three types of sonification strategies discussed in Section IV-B. Fig 10(a) shows a PCA plot of the data for reference, where the data points have been colored according to their cluster assignments by the complex growth transform-based clustering algorithm. Fig 10(b) shows the frequency evolution plots of the relative frequency trajectories are chosen according to the Bark scale, with added sinusoidal variations, as discussed before. Fig 10(c) and (d) show the corresponding time evolution plot and the spectrogram respectively of the sonified output signal $\psi_{\text{sum}}(t)$. Figs 10(e)-(h) show similar results for the musical scale-based technique, while Figs 10(i)-(l) show those corresponding to the musical composition-based sonification.

Effect basis set size and number of clusters: Fig 11 shows the results of the sonification approach on synthetic Datasets I, II, and III, each containing $M = 500$ data points, but the number of underlying clusters being 2, 3 and 5 respectively. The synthetic datasets are generated using a Gaussian mixture model consisting of 2, 3 and 5 different clusters respectively, with a fixed cluster mean and variance associated with each cluster. We consider the basis set to consist of 3 frequency trajectories (i.e., $K = 3$), and the musical composition-based sonification strategy for our experiments. Fig 11(a) shows a scatter plot of Dataset I, with the points color-coded according to their cluster assignment, Fig 11(b) shows the frequency evolutions, while Figs 11(c) and (d) show the time evolution and spectrogram of the sonified output, respectively. Fig 11(e)-(h) show the corresponding plots for Dataset II, while Fig 11(i)-(l) show those for Dataset III. Fig 12 shows similar results on Datasets I, II, and III, for a basis set of frequencies of size $K = 5$.

Effect of different cluster alignments: Fig 13 shows the results of the experiments on Datasets III and IV, where both have the same number of data points ($M = 500$) and 5 underlying clusters, but differ in the cluster alignments and geometries (which represents different data complexities).

VI. EXPERIMENTS ON REAL DATASETS: SONIFICATION OF EEG RECORDINGS FOR EPILEPTIC SEIZURE DETECTION

In this section, we demonstrate how our framework can be applied for sonifying non-invasive electro-encephalogram (EEG) recordings for the detection of epileptic seizures. This is particularly useful in scenarios where trained physicians may not be readily available for analyzing the patterns in the EEG waveforms, as in the case of sub-clinical seizure onset. In instances of refractory epilepsy in particular, where patients are non-responsive to anti-epileptic drugs, neurostimulation or surgery is an option if the epileptogenic focus/foci can be identified [34], [35]. Sonification can be useful in such scenarios since it usually implies faster feedback than visualization, leading to faster injection of the radiotracer for effective localization of the epileptogenic foci.

Traditional approaches for seizure detection from EEG recordings include a sliding window-based feature extraction stage. This involves inferring network connectivities or frequency components of the EEG channels in each window [36]. The feature extraction stage is usually followed by a classification stage formed by support vector machine or neural network-based classifiers, and more recently, deeper architectures like CNNs [6]. This is illustrated in Fig 14(c). On the other hand, sonification-based approaches for seizure detection usually adopt a parameter mapping approach where the voltage values of a single channel of the EEG signal are mapped to the auditory domain [5]. There exists a second approach for sonification of EEG signal. This involves mapping the output of the binary classifier into two distinct audio waveforms representing the presence or absence of seizure. While the first approach discards valuable information in terms of the network interactions during a seizure event, the second approach conveys the decision of the classifier using sound [37]. It thus does not employ the end-user (clinician or caregiver) in the decision-making process.

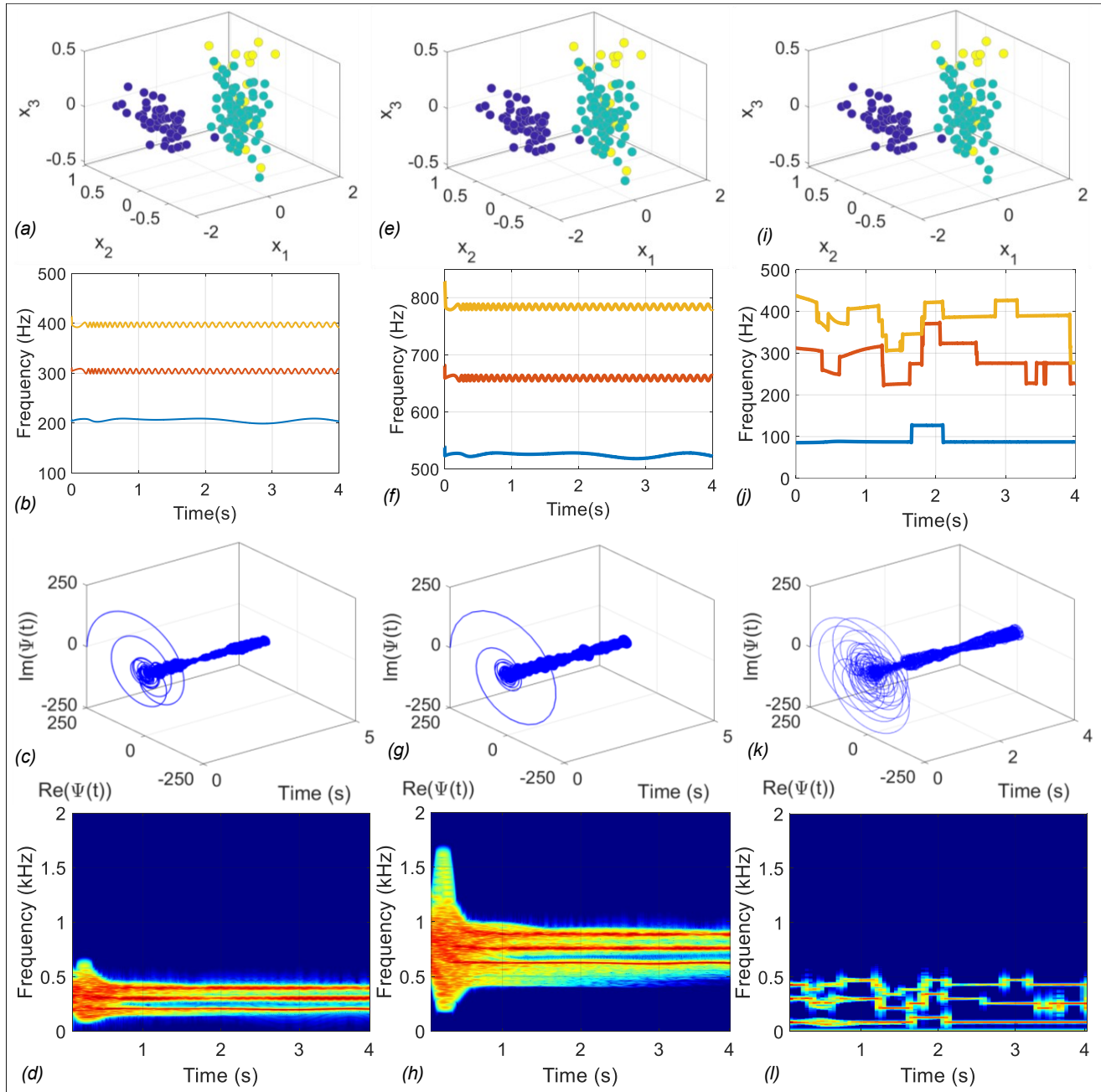


Figure 10: **Different sonification strategies for clustering the ‘Iris’ dataset.** (a)-(d) show the PCA (labeled according to cluster assignments), frequency trajectories, time evolution, and spectrogram of the sonified output for the bark scale-based sonification; (e)-(h) show similar plots for the musical scale based sonification; and (i)-(l) show the results for the musical composition based sonification.

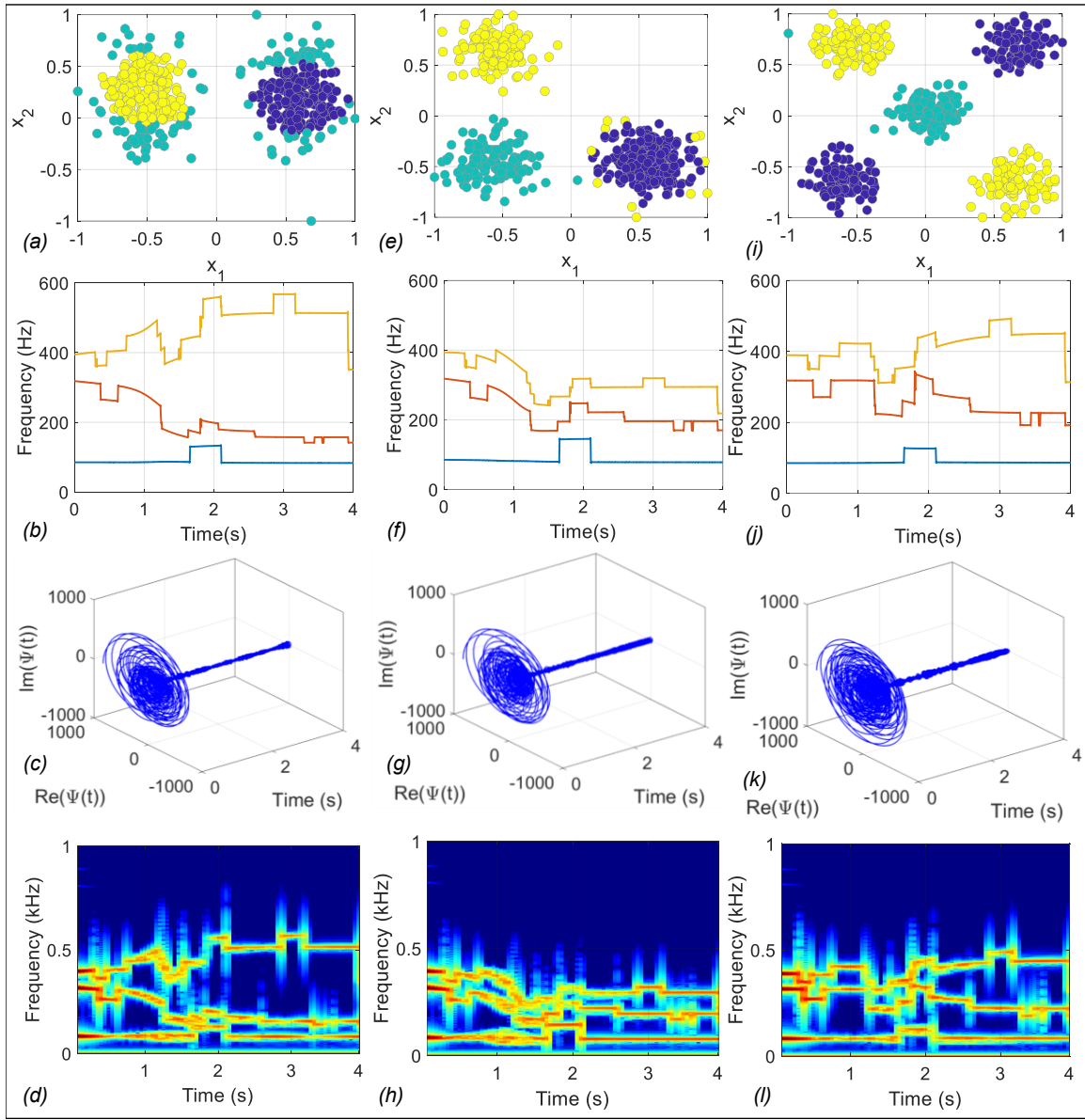


Figure 11: **Musical composition based sonification on synthetic datasets with same number of points ($M = 500$) but different number of clusters, with a basis set size of 3.** (a)-(d) show the scatter plot (colored according to the cluster assignments), frequency trajectories, time evolution, and spectrogram of the sonified output respectively for Dataset I; (e)-(h) show similar results for Dataset II; and (i)-(l) show the results for Dataset III.

In this section, we propose a novel sonification method that considers the interactions between all EEG channels to detect a seizure event, and creates a sonified output that enables the listener to make a decision. The proposed sonification module thus acts as a feature extractor in this case, and the task of detecting a seizure event still lies with the human-in-the-loop. We applied the sonification strategy to scalp EEG recordings of pediatric patients provided by the CHB-MIT database collected at the Children’s Hospital Boston [10], [11]. The recordings were collected from 23 channels placed according to the International 10-20 system of electrode positions, and all signals were sampled at 256 Hz. The original dataset contains EEG recordings from 23 patients (5 males aged 3-22 years, 17 females aged 1.5-19 years). Only the recordings that had one or more instances of the seizure (as marked by domain experts in the dataset) were used for our purpose.

The first step in the sonification process involves (i) removing the baseline, (ii) re-referencing the electrodes to the average potential, and finally (iii) applying a bandwidth filter with a passband of 0.5-50 Hz for noise reduction. All of these operations were performed using the publicly available EEGLAB toolbox [38]. The EEG waveforms

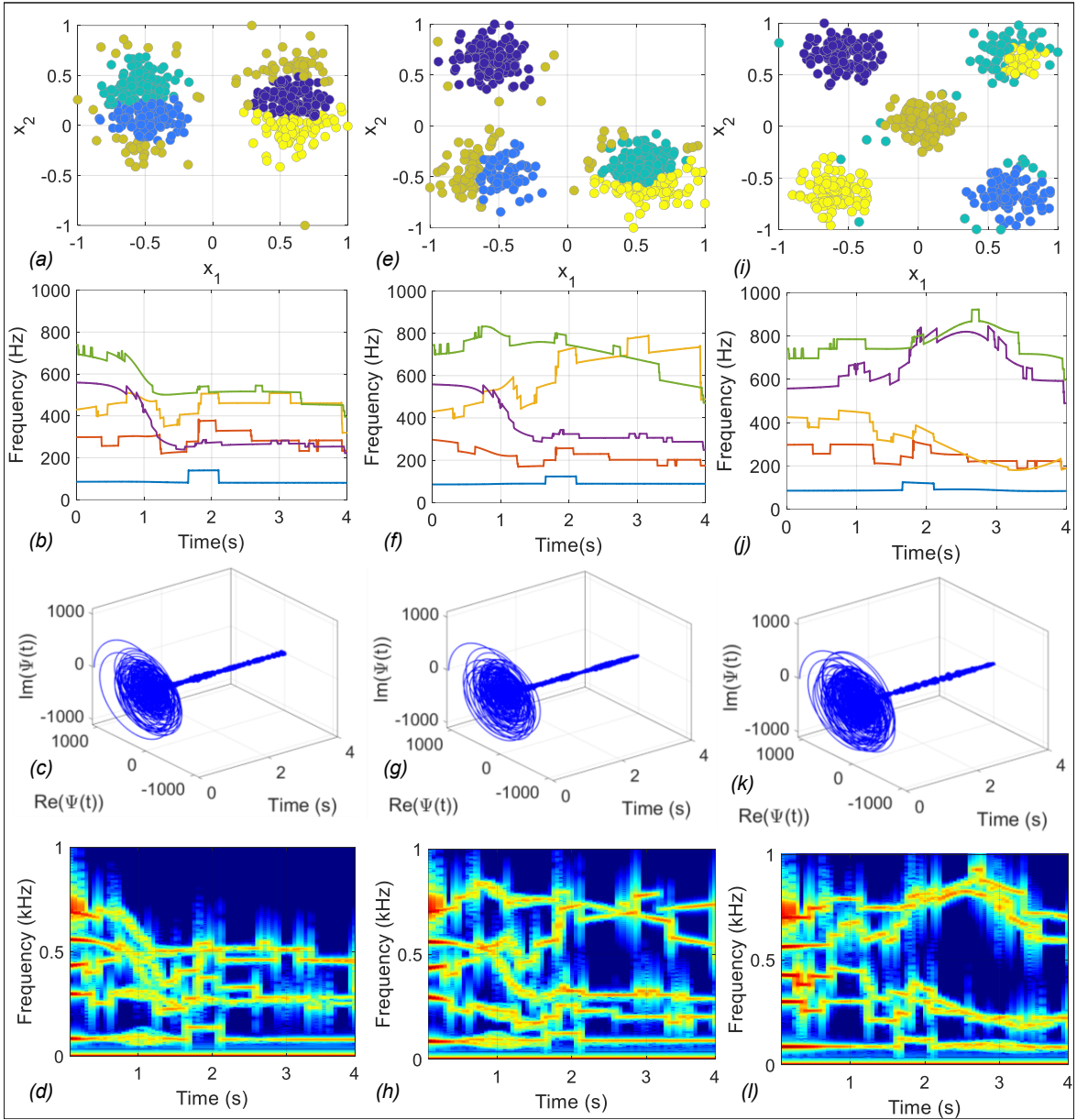


Figure 12: **Musical composition based sonification on synthetic datasets with same number of points ($M = 500$) but different number of clusters, with a basis set size of 5.** (a)-(d) show the scatter plot (colored according to the cluster assignments), frequency trajectories, time evolution, and spectrogram of the sonified output respectively for Dataset I; (e)-(h) show similar results for Dataset II; and (i)-(l) show the results for Dataset III.

were then normalized channel-wise in the range $[-1, +1]$. Fig 14(a) shows a sample EEG recording with a seizure event lasting from 1467-1494s. Fig 14(b) shows a zoomed-in version of the seizure event. The EEG channel signals were then analyzed using a non-overlapping sliding window of 2s duration (denoted by the matrix \mathbf{E}). Note that 2s sliding windows have been used extensively in literature for epileptic seizure detection from EEG signals since they are short enough for retaining relevant information about the dynamics of the signal [10]. Fig 14(c) shows the traditional supervised approach where the output of a classifier trained on features extracted from each of the 2s windows is converted to an audio waveform. In contrast, our proposed unsupervised approach (shown in Fig 14(d)) involves passing each such window through the sonification module to produce an audio signal. The output signal is finally used by the user to decide whether the window belongs to a seizure or a non-seizure event. In this framework, for each 2s epoch, we solve a quadratic optimization problem of the following form for inferring the

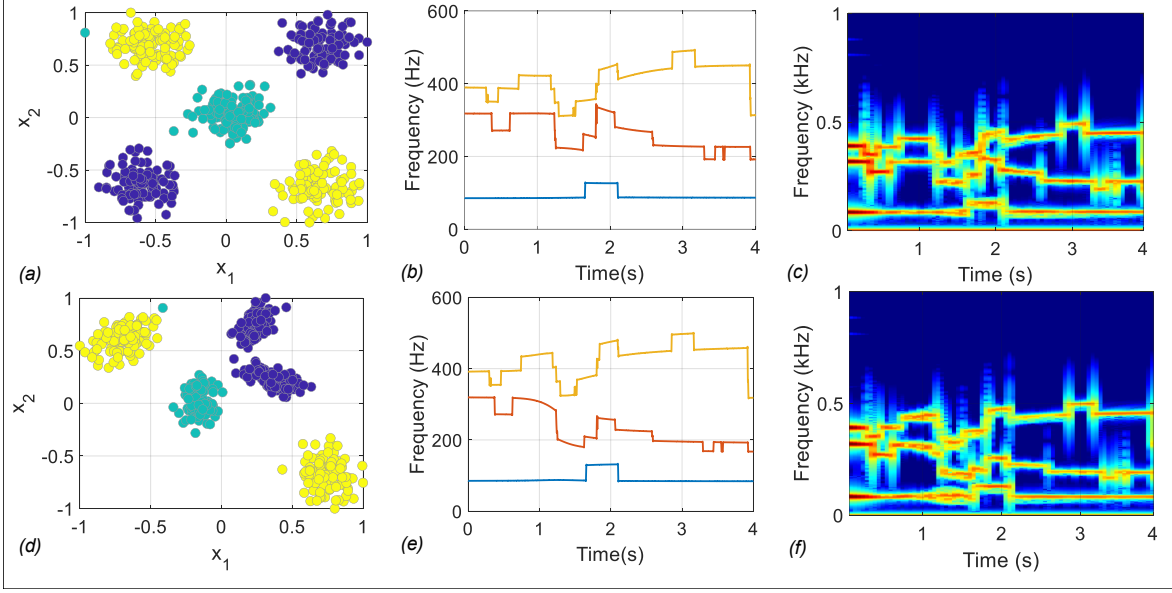


Figure 13: **Musical composition based sonification considering synthetic datasets with same number of points ($M = 500$) and same number of clusters ($K = 5$), but with different cluster geometries, considering a basis set of size 3.** (a)-(c) show the scatter plot (colored according to the cluster assignments), frequency trajectories, and spectrogram of the sonified signal respectively for Dataset III; and (d)-(f) show similar results for Dataset IV.

relationship between different EEG channels:

$$\underset{\mathbf{p} \in D \subset \mathbb{R}_+^{n_c}}{\text{minimize}} \quad \mathcal{H}(\mathbf{p}) = \mathbf{p}^T \mathbf{Q} \mathbf{p} \quad (15)$$

$$\text{s.t.} \quad \sum_{k=1}^{n_c} p_k = 1, \quad p_k \geq 0 \quad \forall k = 1, \dots, n_c, \quad (16)$$

where $n_c = 23$ is the number of EEG channels under consideration. $\mathbf{Q} \in \mathbb{R}^{n_c \times n_c}$ represents a measure of the pairwise connectivity between the EEG channels, and is computed as $\mathbf{Q} = f(\mathbf{E}\mathbf{E}')$. Here, $f(\cdot) : \mathbb{R}^{n_c \times n_c} \mapsto \mathbb{R}^{n_c \times n_c}$ represents an element-wise square-root operator, and \mathbf{E} represents the EEG data over the 2s window. Fig 14(e) illustrates the proposed approach. We can apply the complex growth transforms to arrive at the optimal solution of Eq (15)-(16). For the simulations presented in this section, the baseline frequency of all the oscillators is chosen to be $\omega_n = 300$ Hz. The relative frequencies $\xi_{k,n}$'s of the oscillator variables ψ_k 's (where $p_k = \psi_k \psi_k^*$, $\psi_k \in \mathbb{C}$) are chosen according to the relative placements of the electrodes from which the corresponding channels are recorded. For example, oscillators corresponding to channels recorded from the frontal, temporal, parietal, and occipital lobes of the brain are can be mapped to the notes G, E, C, and A based on the equally-tempered Western musical scale around A = 440 Hz. This mapping scheme based on the electrode locations is illustrated in Fig 15. Similar mappings can also be achieved using the Bark scale-based and musical composition-based approaches. Note that higher relative frequencies are assigned to the channels from the frontal and parietal lobes, as these brain regions are affected more by epileptic seizure events in pediatric patients than in the other two regions [39]. Thus allocating higher frequencies to the oscillators/channels mapping to these lobes would result in greater perturbation in the corresponding frequency bands. This would potentially lead to better discrimination between seizure and non-seizure events. The framework thus allows for incorporating spatial information about the electrode locations because of the proposed mapping strategy, in addition to the temporal information captured by the sonification process. Spatial information can also be embedded in the sonification process by utilizing a dual-channel audio representation instead of the single-channel audio representation discussed until now in the paper. Since the output sonified signal is inherently complex, we can use the left and right channels for outputting the real and imaginary part respectively of the signal, or vice versa.

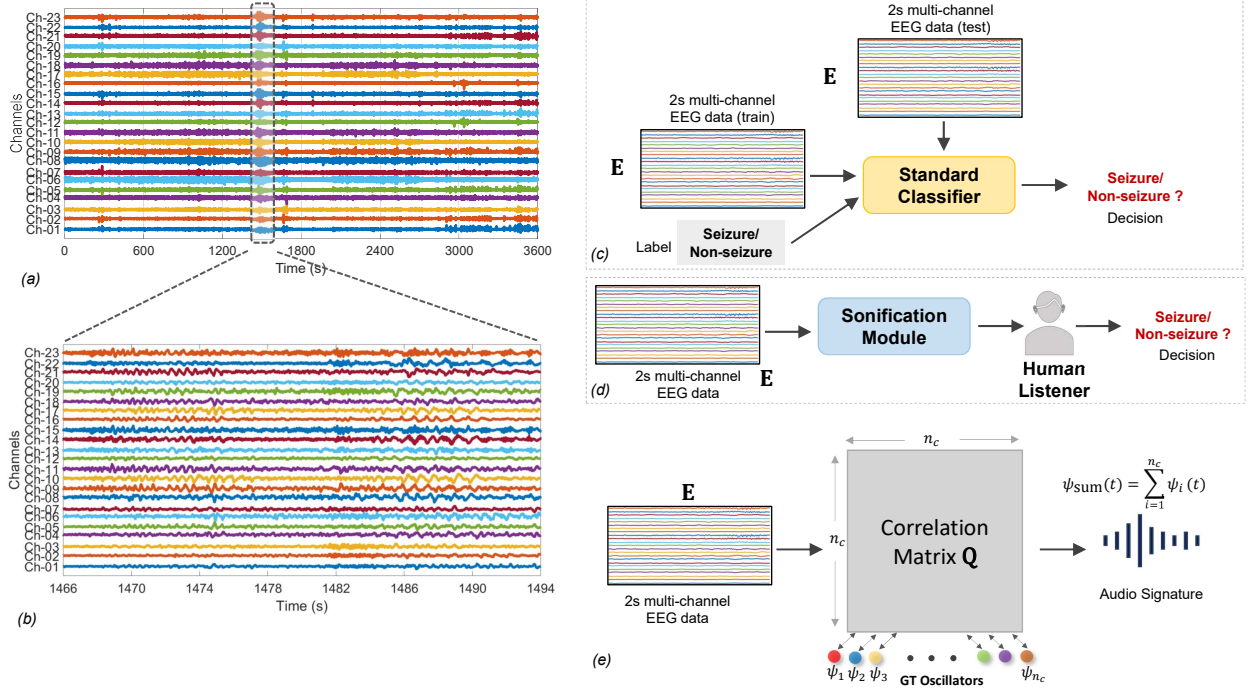


Figure 14: **Sonification of EEG recordings.** (a) 1 hour EEG recording with a single seizure event between 1467-1494s; (b) zoomed-in version of the seizure event; (c) An automatic seizure classification module which involves (i) a supervised learning phase using 2s windows of EEG recordings (represented by the matrix \mathbf{E}) along with the corresponding labels, and (ii) a testing phase where the presence of seizure is determined by using the trained model on a previously unseen 2s window; (d) Proposed approach, where features extracted from each 2s EEG window are passed directly through a sonification module and presented to a human listener, who takes the decision; and (e) Sonification module involves computing the pairwise connectivity matrix \mathbf{Q} between all the n_c EEG channels in the window. \mathbf{Q} thus interconnects n_c globally coupled growth transform oscillators, and the final sonified output is obtained by the superposition of all the oscillator outputs.

Fig 16 shows the results of applying our sonification technique to five different instances of seizure recordings (recording #03, #04, #15, #16, and #18 respectively) for the same patient (Patient #01). Figs 16(a)-(e) show spectrograms for 2s windows corresponding to non-seizure windows in each of the five recordings. Figs 16(f)-(j) show similar figures for 2s windows corresponding to seizure events in the same five recordings. Based on the above results, we can conclude that the spectrograms for the sonified windows corresponding to seizure events seem to have a higher level of entropy than those corresponding to non-seizure events.

Fig 17, on the other hand, presents a comparison of the results of the sonification process between seizure recordings across ten different patients (Patient #01, #03 #05, #08, #11, #14, #19, #20, #21 and #22 respectively). Figs 17(a)-(j) show the spectrograms for 2s windows corresponding to non-seizure events. Figs 17(k)-(t) show similar results for 2s windows corresponding to seizure events across the ten patients. The sonified signals corresponding to seizure vs non-seizure windows thus show consistent patterns not only across different recordings for the same patient, but also across patients. Our proposed sonification framework is thus capable of identifying both spatial and temporal patterns in the time-varying EEG signal, which can be potentially used for real-time detection of epileptic seizure events.

VII. DISCUSSION AND CONCLUSIONS

This paper presented a novel technique for the sonification of high-dimensional data that incorporates both learning and sonification stages into the same module. The end-user then uses the sonified audio output for making a decision. This is in contrast to existing sonification techniques that either involve (a) using a learning algorithm upstream that inherently controls the decision-making process and maps the outcome to an audio signature, or (b) directly

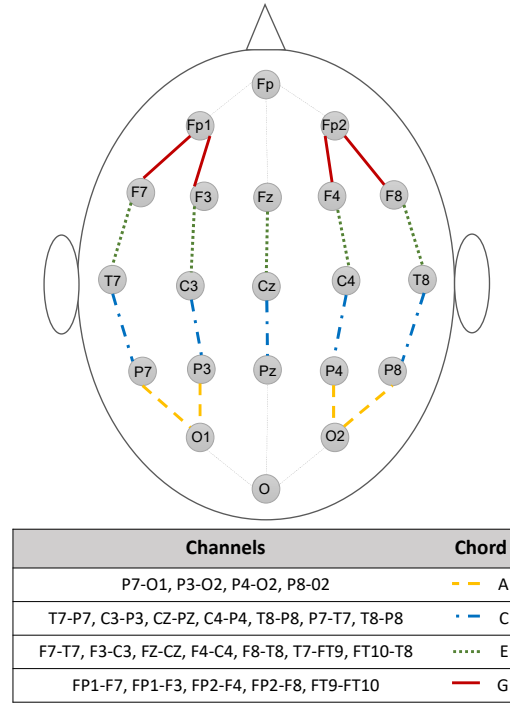


Figure 15: **Illustration of relative frequency assignment to oscillators.** Relative frequencies of oscillators were assigned to notes A, C, E, or G of the equally-tempered musical scale around A=440 Hz. Frequencies are assigned based on the electrode placement of the corresponding channel.

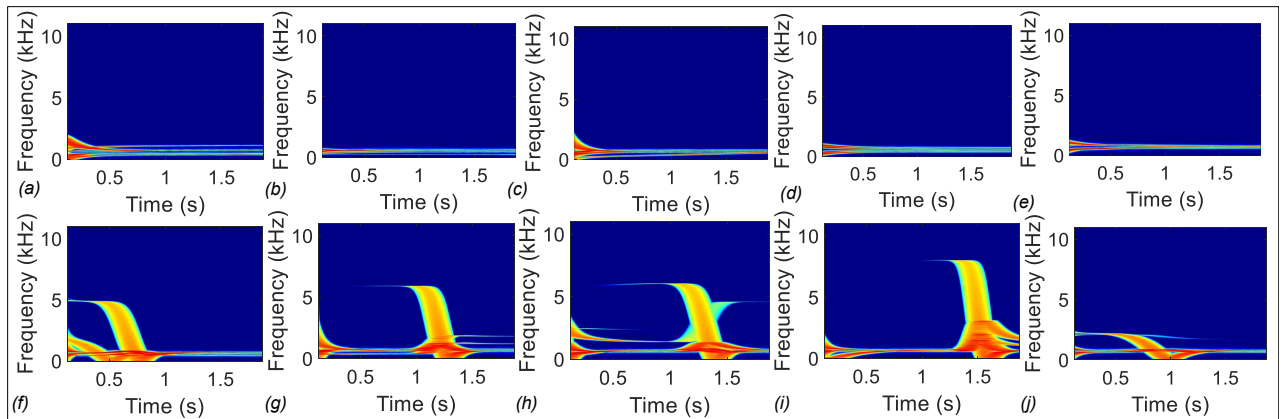


Figure 16: **Sonification results on arbitrary 2s windows for 5 different recordings of Patient #01** (#03, #04, #15, #16 and #18 respectively) . (i) **Non-seizure events:** (a)-(e): Spectrograms for arbitrary non-seizure windows in each of the five recordings; (ii) **Seizure events:** (f)-(j): similar results for arbitrary seizure windows in each of the five recordings.

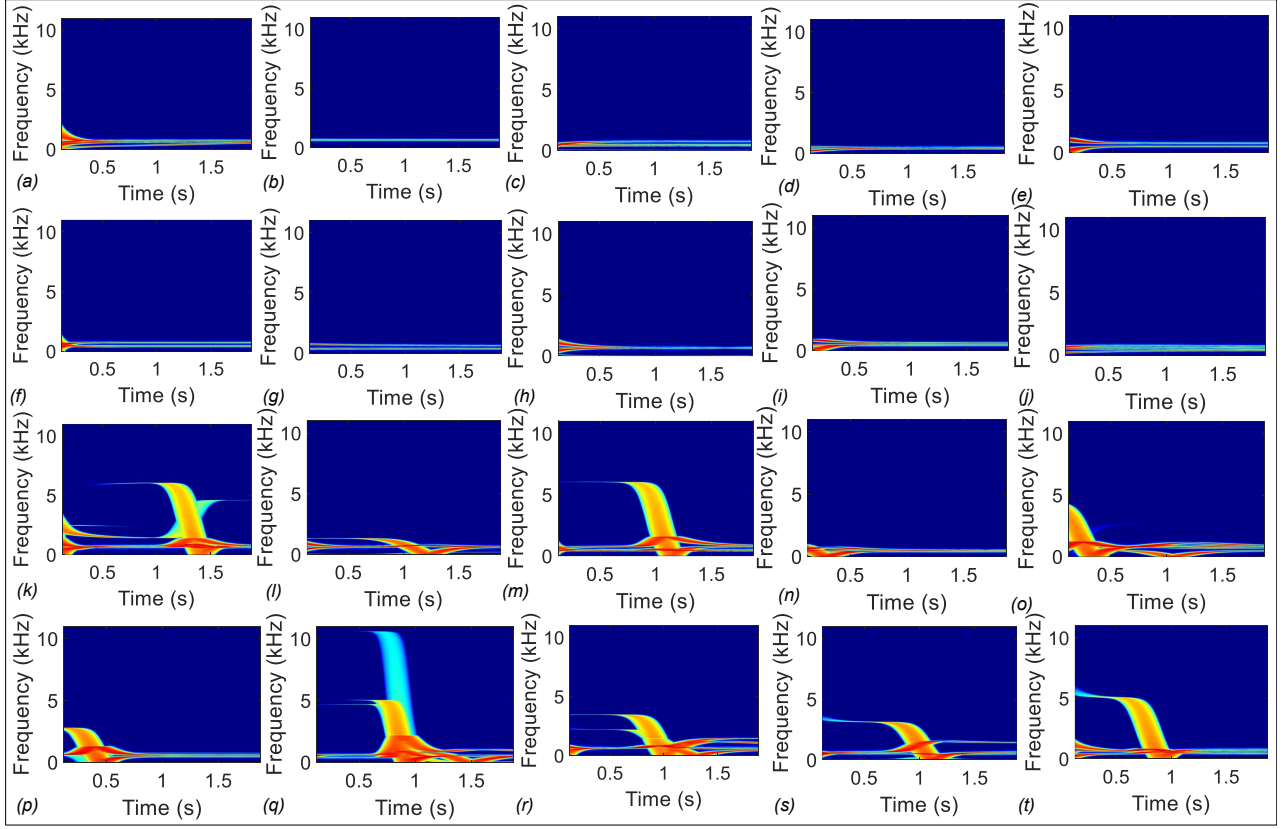


Figure 17: **Sonification results on arbitrary 2s windows for a single recording of each of 10 different patients** (Patient #01, #03 #05, #08, #11, #14, #19, #20, #21 and #22 respectively). (i) **Non-seizure events**: (a)-(j): Spectrograms for arbitrary non-seizure windows in each of the ten patients; (ii) **Seizure events** : (k)-(t): similar results for arbitrary seizure windows in each of the ten patients.

mapping the underlying variables to different parameters of the sound wave, without accounting for the correlations and patterns in the data. At the core of the framework lies the complex growth transform dynamical system, which simultaneously utilizes the learning variables and psychoacoustic parameters defined by the user. Thus, our proposed sonification module outputs a binaural audio signature that can be used for human-in-the-loop decision-making. The output sonified signal encodes the high-dimensional space data, which might be particularly useful for low throughput systems. Additionally, the method can be utilized for solving a range of learning problems of varying dimensionality and provides several tunable parameters that can be customized to adapt different sonification strategies. Experiments on synthetic and real data show encouraging results, proving that the method can be used for real-time applications involving high-dimensional, temporally varying data. Future directions for this work involve explore using sonification in multimodal perceptualization tasks [40], [41] with potential applications in visual rehabilitation [42], neuroprosthesis [43], and for cohesive perception in autonomous robots [44].

VIII. APPENDIX

A. Nomenclature

Following are the notational conventions used in this paper:

B. Proof of Main Result

Theorem 1: Considering $\Psi \in D^C = \{\Psi \in \mathbb{C}^{M \times K} : \sum_{k=1}^K |\psi_{ik}|^2 = 1, \forall i = 1, \dots, M\} \subset \mathbb{C}^{M \times K}$, a time evolution of the form given below converges to limit cycles corresponding to the optimal point of a Lipschitz

Table 2: Notations

Variable	Definition
\mathbb{R}_+	One-dimensional positive real vector space
\mathbb{R}^M	M -dimensional real vector space
\mathbb{C}^M	M -dimensional complex vector space
\mathbf{P}	a matrix
\mathbf{p}	a vector
$ z $	magnitude of a complex variable z
$\text{Re}(z)$	real part of a complex variable z
$\text{Im}(z)$	imaginary part of a complex variable z
z^*	complex conjugate of a complex variable z
$z(t)$	a continuous-time complex variable at time t
z_n	a discrete-time complex variable at the n^{th} time step
\odot	Hadamard product of two matrices

continuous objective function $\mathcal{H}(\Psi)$:

$$\psi_{ik,n} \leftarrow \psi_{ik,n-1} [\cos(\theta_{i,n}) + j \sin(\theta_{i,n}) \sigma_{ik,n-1}(\Psi_{n-1})], \quad (17)$$

where $\sigma_{ik} \rightarrow 1 \quad \forall i = 1, \dots, N, k = 1, \dots, M$, in steady state.

Proof: Consider a constrained optimization problem of the following form :

$$\underset{\mathbf{P} \in D \subset \mathbb{R}_+^{M \times K}}{\text{minimize}} \quad \mathcal{H}'(\mathbf{P}) \quad (18)$$

$$\text{s.t.} \quad \sum_{k=1}^K p_{ik} = 1, \quad \forall i = 1, \dots, M, \quad p_{ik} \geq 0 \quad \forall i, k \quad (19)$$

In [7], we used the Baum-Eagon inequality [9] to show that the optimal point of the optimization problem represented by Eqs (18)-(19) for a generic Lipschitz continuous cost function $\mathcal{H}'(\mathbf{P})$ ($\mathbf{P} \in D \subset \mathbb{R}^{M \times K}$) corresponds to the steady-state solution of a multiplicative update-based discrete-time growth transform dynamical system model given by:

$$p_{ik,n} \leftarrow (1 - \alpha_{i,n}) p_{ik,n-1} + \alpha_{i,n} p_{ik,n-1} g_{ik,n-1}(\mathbf{P}_{n-1}), \quad (20)$$

where $g_{ik,n-1}(\mathbf{P}_{n-1}) = \frac{(-\frac{\partial \mathcal{H}'}{\partial p_{ik,n-1}} + \lambda)}{\sum_{l=1}^K p_{il,n-1} (-\frac{\partial \mathcal{H}'}{\partial p_{il,n-1}} + \lambda)}$, $\forall i = 1, \dots, M$, and $0 \leq \alpha_{i,n} \leq 1 \quad \forall i$. The constant

$\lambda \in \mathbb{R}_+$ is chosen to ensure that $|\frac{\partial \mathcal{H}'(\mathbf{P})}{\partial p_{ik,n-1}} + \lambda| > 0, \forall i, k$. Note that convergence to the optimal solution is guaranteed even though $\alpha_{i,n}$'s are time-varying, since this would still ensure the invariance of the manifold D .

Taking $g_{ik,n-1}(\mathbf{P}_{n-1}) = \sigma_{ik,n-1}^2(\mathbf{P}_{n-1}) \quad \forall i, k$ and $\alpha_{i,n} = \sin^2(\theta_{i,n})$, we get:

$$p_{ik,n} \leftarrow [\cos^2(\theta_{i,n}) p_{ik,n-1} + \sin^2(\theta_{i,n}) p_{ik,n-1} \sigma_{ik,n-1}^2(\mathbf{P}_{n-1})] \quad (21)$$

Representing $p_{ik,n} = \psi_{ik,n}\psi_{ik,n}^*$, $\psi_{ik,n} \in \mathbb{C}$, the update equations become:

$$\begin{aligned} \psi_{ik,n}\psi_{ik,n}^* &\leftarrow \psi_{ik,n-1}[\cos(\theta_{i,n}) + j\sin(\theta_{i,n})\sigma_{ik,n-1}(\mathbf{P}_{n-1})] \\ &\times \psi_{ik,n-1}^*[\cos(\theta_{i,n}) + j\sin(\theta_{i,n})\sigma_{ik,n-1}(\mathbf{P}_{n-1})]^* \end{aligned} \quad (22)$$

Considering $\mathcal{H}'(\mathbf{P}) = \mathcal{H}(\mathbf{\Psi})$ to be analytic in D^C , since by Wirtinger's calculus,

$$\frac{\partial \mathcal{H}}{\partial \psi_{ik,n-1}} = \frac{\partial \mathcal{H}}{\partial \psi_{ik,n-1}\psi_{ik,n-1}^*} \cdot \left(\frac{\partial \psi_{ik,n-1}\psi_{ik,n-1}^*}{\partial \psi_{ik,n-1}} \right) = \frac{\partial \mathcal{H}}{\partial \psi_{ik,n-1}\psi_{ik,n-1}^*} \cdot \psi_{ik,n-1}^*, \quad (23)$$

we have

$$\sigma_{ik,n-1}(\mathbf{\Psi}_{n-1}) = \sqrt{\frac{\left(-\frac{\partial \mathcal{H}}{\partial \psi_{ik,n-1}} + \lambda \psi_{ik,n-1}^* \right)}{\psi_{ik,n-1}^* \sum_{l=1}^K \psi_{il,n-1} \left(-\frac{\partial \mathcal{H}}{\partial \psi_{il,n-1}} + \lambda \psi_{il,n-1}^* \right)}} \quad (24)$$

The discrete time update equations for $\psi_{ik,n}$ is thus given by:

$$\psi_{ik,n} \leftarrow \psi_{ik,n-1}[\cos(\theta_{i,n}) + j\sin(\theta_{i,n})\sigma_{ik,n-1}(\mathbf{\Psi}_{n-1})], \quad (25)$$

Now, let us define $\theta_{i,n} = \omega_{i,n}\Delta t$, where $\omega_{i,n}$ is the baseline frequency for the i -th group of oscillator variables at the n -th time step, and Δt is a unit time step. Eq (25) can then be rewritten as:

$$\psi_{ik,n} \leftarrow \psi_{ik,n-1}[\cos(\omega_{i,n}\Delta t) + j\sin(\omega_{i,n}\Delta t)\sigma_{ik,n-1}(\mathbf{\Psi}_{n-1})], \quad (26)$$

■

Theorem 2: *Different oscillation frequencies can be assigned to each element in the dynamical system represented by Eq (17) by adding an instantaneous relative phase term, i.e., the following dynamical system*

$$\psi_{ik,n} \leftarrow \psi_{ik,n-1}[\cos(\omega_{i,n}\Delta t) + j\sin(\omega_{i,n}\Delta t)\sigma_{ik,n-1}] \exp(j\xi_{ik,n}\Delta t), \quad (27)$$

converges to the same solution as that attained by the one described by Eq (17).

Proof: Considering $\varphi_{ik,n}$ to be the instantaneous relative phase of the ik -th element with respect to an absolute reference at the n -th time instance, we have:

$$\psi_{ik,n} \leftarrow \psi_{ik,n-1}[\cos(\omega_{i,n}\Delta t) + j\sin(\omega_{i,n}\Delta t)\sigma_{ik,n-1}] \exp(j\varphi_{ik,n}), \quad (28)$$

Let $\varphi_{ik,n}$ be the instantaneous relative phase of the ik -th oscillator variable. Then we can define $\varphi_{ik,n} = \xi_{ik,n}\Delta t$, such that $\xi_{ik,n}$ is the corresponding relative frequency. The addition of the relative frequency term only alters the oscillator dynamics without altering the instantaneous solution of the optimization problem. Additionally, we can incorporate information about the instantaneous oscillator dynamics in the relative phase by using $\varphi_{ik,n} = c_{ik}\sigma_{ik,n}\xi_{ik,n}\Delta t$, where c_{ik} 's are arbitrary scaling coefficients. Note that as before, this mapping only perturbs the dynamics, without impacting the true solution. Eq (28) can be rewritten as

$$\begin{aligned} \psi_{ik,n} &\leftarrow \psi_{ik,n-1} \sqrt{\cos^2(\omega_{i,n}\Delta t) + \sin^2(\omega_{i,n}\Delta t)\sigma_{ik,n-1}^2} \\ &\times \exp \left(j \left(\arctan(\sigma_{ik,n-1} \tan(\omega_{i,n}\Delta t)) + \xi_{ik,n}\Delta t \right) \right), \end{aligned} \quad (29)$$

Finally, in the steady state, since $\sigma_{ik,n-1} \xrightarrow{n \rightarrow \infty} 1 \quad \forall i, k$, and taking $\Delta t = \frac{1}{F_s}$ (F_s is the sampling frequency), we have:

$$\psi_{ik,n} \leftarrow \psi_{ik,n-1} \exp \left(j \left(\omega_{i,n} + \xi_{ik,n} \right) \frac{1}{F_s} \right). \quad (30)$$

The output sonified signal in the steady-state is then obtained by the superposition of all the oscillator waveforms, i.e.,

$$\psi_{\text{sum},n} = \sum_{i=1}^M \sum_{k=1}^K \psi_{ik,n} = \sum_{i=1}^M \sum_{k=1}^K \psi_{ik,n-1} \exp \left(j(\omega_{i,n} + \xi_{ik,n}) \frac{1}{Fs} \right) \quad (31)$$

In particular, when there is a single global constraint involving all the variables, i.e., when $M = 1$, then $\psi_{ik,n} = \psi_{k,n}$, $\omega_{i,n} = \omega_n$ and $\xi_{ik,n} = \xi_{k,n}$. The complex growth transform dynamical system for this special case can be written as:

$$\psi_{k,n} \leftarrow \psi_{k,n-1} \exp \left(j(\omega_n + \xi_{k,n}) \frac{1}{Fs} \right) \quad (32)$$

$$\psi_{\text{sum},n} = \sum_{k=1}^K \psi_{k,n-1} \exp \left(j(\omega_n + \xi_{k,n}) \frac{1}{Fs} \right). \quad (33)$$

■

Theorem 3: *The complex growth transform dynamical system follows a nonlinear unitary transformation.* Defining

$$\mathcal{U}_n = \begin{pmatrix} d_{11,n-1} & & & \\ & \ddots & & \\ & & d_{ik,n-1} & \\ & & & \ddots \\ & & & & d_{MK,n-1} \end{pmatrix},$$

$$\text{where } d_{ik,n} = [\cos(\omega_{i,n}\Delta t) + j \sin(\omega_{i,n}\Delta t)\sigma_{ik,n-1}] \exp(j\xi_{ik,n}\Delta t), \quad (34)$$

and using Eq (28), we have:

$$\Psi_n \leftarrow \mathcal{U}_n(\Psi_{n-1}) \odot \Psi_{n-1}. \quad (35)$$

Additionally, since $\sum_{i=1}^M \sum_{k=1}^K |\psi_{ik,n}|^2 = \sum_{i=1}^M \sum_{k=1}^K |\psi_{ik,n-1}|^2$, $\mathcal{U}_n : \mathbb{C}^{M \times K} \mapsto \mathbb{C}^{M \times K}$ can be thought of as an instantaneous nonlinear unitary operator [45] that ensures that the signal energy is conserved over time.

For the special case of a single global constraint ($M = 1$), we have

$$\mathcal{U}_n = \begin{pmatrix} d_{1,n} \\ \vdots \\ d_{k,n} \\ \vdots \\ d_{K,n} \end{pmatrix}, \quad \mathcal{U}_n : \mathbb{C}^K \mapsto \mathbb{C}^K,$$

$$d_{k,n} = [\cos(\omega_n\Delta t) + j \sin(\omega_n\Delta t)\sigma_{k,n-1}] \exp(j\xi_{k,n}\Delta t), \quad \forall k = 1, \dots, K \quad (36)$$

$$\text{and } \sum_{k=1}^K |\psi_{k,n}|^2 = \sum_{k=1}^K |\psi_{k,n-1}|^2. \quad (37)$$

■

C. Mapping procedure

Let us consider an optimization problem of the following generic form:

$$\begin{aligned} & \min_{\{p_i \in \mathbb{R}_+\}} \mathcal{H}'(\{p_i\}) \\ & \text{s.t. } |p_i| \leq \gamma, \quad \gamma \in \mathbb{R}_+ \quad \forall i. \end{aligned} \quad (38)$$

Considering $p_i = p_i^+ - p_i^- \forall i = 1, \dots, N$, where both $p_i^+, p_i^- \geq 0$. Since by triangle inequality, $|p_i| \leq |p_i^+| + |p_i^-|$, enforcing $p_i^+ + p_i^- = \gamma \forall i$ would automatically ensure $|p_i| \leq 1 \forall i$. Thus we have,

$$\begin{aligned} \underset{\{p_i\}}{\operatorname{argmin}} \quad \mathcal{H}'(\{p_i\}) &\equiv \underset{\{p_i^+, p_i^-\}}{\operatorname{argmin}} \quad \mathcal{H}'(\{p_i^+, p_i^-\}) \\ \text{s.t.} \quad |p_i| &\leq \gamma, \quad p_i \in \mathbb{R} \quad \text{s.t.} \quad p_i^+ + p_i^- = \gamma, \quad p_i^+, p_i^- \geq 0 \end{aligned} \quad (39)$$

Finally, we replace $p_i^+ \leftarrow \frac{p_i^+}{\gamma}$, $p_i^- \leftarrow \frac{p_i^-}{\gamma}$ to arrive at the following equivalent optimization problem over a probabilistic domain:

$$\begin{aligned} \min_{\{p_i^+, p_i^-\}} \quad &\mathcal{H}'(\{p_i^+, p_i^-\}) \\ \text{s.t.} \quad &p_i^+ + p_i^- = 1, \quad \forall i. \end{aligned} \quad (40)$$

The above problem, in turn, can be mapped to the complex growth transform-based sonification framework by considering $p_i^+ = |\psi_{i1}|^2$, $p_i^- = |\psi_{i2}|^2$, and following the procedure in VIII-B.

REFERENCES

- [1] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*. Logos Verlag Berlin, Germany, 2011.
- [2] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. H. Flowers, N. Miner, and J. Neuhoff, "Sonification report: Status of the field and research agenda," 2010.
- [3] G. Dubus and R. Bresin, "A systematic review of mapping strategies for the sonification of physical quantities," *PloS one*, vol. 8, no. 12, p. e82491, 2013.
- [4] J. N. Kather, T. Hermann, Y. Bukschat, T. Kramer, L. R. Schad, and F. G. Zöllner, "Polyphonic sonification of electrocardiography signals for diagnosis of cardiac pathologies," *Scientific reports*, vol. 7, no. 1, pp. 1–6, 2017.
- [5] P. Loui, M. Koplin-Green, M. Frick, and M. Massone, "Rapidly learned identification of epileptic seizures from sonified eeg," *Frontiers in human neuroscience*, vol. 8, p. 820, 2014.
- [6] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, and J. Xiang, "Epileptic seizure detection based on eeg signals and cnn," *Frontiers in neuroinformatics*, vol. 12, p. 95, 2018.
- [7] O. Chatterjee and S. Chakrabartty, "Decentralized global optimization based on a growth transform dynamical system model," *IEEE Transactions on Neural Networks and Learning Systems*, 2018.
- [8] O. Chatterjee and S. Chakrabartty, "Resonant machine learning based on complex growth transform dynamical systems," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2020.
- [9] L. E. Baum and G. Sell, "Growth transformations for functions on manifolds," *Pacific Journal of Mathematics*, vol. 27, no. 2, pp. 211–227, 1968.
- [10] A. H. Shoen and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 975–982.
- [11] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [12] A. M. Martinez and A. C. Kak, "Pca versus lda," *IEEE transactions on pattern analysis and machine intelligence*, vol. 23, no. 2, pp. 228–233, 2001.
- [13] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [14] M. Balasubramanian, E. L. Schwartz, J. B. Tenenbaum, V. de Silva, and J. C. Langford, "The isomap algorithm and topological stability," *Science*, vol. 295, no. 5552, pp. 7–7, 2002.
- [15] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.
- [16] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.
- [17] K. R. Moon, D. van Dijk, Z. Wang, S. Gigante, D. B. Burkhardt, W. S. Chen, K. Yim, A. van den Elzen, M. J. Hirn, R. R. Coifman *et al.*, "Visualizing structure and transitions in high-dimensional biological data," *Nature Biotechnology*, vol. 37, no. 12, pp. 1482–1492, 2019.
- [18] M. Nguyen, S. Purushotham, H. To, and C. Shahabi, "m-tsne: A framework for visualizing high-dimensional multivariate time series," *arXiv preprint arXiv:1708.07942*, 2017.
- [19] J. St George, S. Y. Kim, and H.-P. Bischof, "Sonification of simulated black hole merger data," in *Proceedings of the International Conference on Modeling, Simulation and Visualization Methods (MSV)*. The Steering Committee of The World Congress in Computer Science, Computer ..., 2018, pp. 3–9.
- [20] A. Supper, "Sublime frequencies: The construction of sublime listening experiences in the sonification of scientific data," *Social Studies of Science*, vol. 44, no. 1, pp. 34–58, 2014.
- [21] J. Dunn and M. A. Clark, "Life music: the sonification of proteins," *Leonardo*, vol. 32, no. 1, pp. 25–32, 1999.

- [22] C.-H. Yu and M. J. Buehler, "Sonification based de novo protein design using artificial intelligence, structure prediction, and analysis using molecular modeling," *APL bioengineering*, vol. 4, no. 1, p. 016108, 2020.
- [23] L. Gionfrida and A. Roginska, "A novel sonification approach to support the diagnosis of alzheimer's dementia," *Frontiers in neurology*, vol. 8, p. 647, 2017.
- [24] A. Våljamäe, T. Steffert, S. Holland, X. Marimon, R. Benitez, S. Mealla, A. Oliveira, and S. Jordà, "A review of real-time eeg sonification research," 2013.
- [25] A. Dascalu and E. David, "Skin cancer detection by deep learning and sound analysis algorithms: A prospective clinical study of an elementary dermoscope," *EBioMedicine*, vol. 43, pp. 107–113, 2019.
- [26] M. Ballora, R. J. Cole, H. Kruesi, H. Greene, G. Monahan, and D. L. Hall, "Use of sonification in the detection of anomalous events," in *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2012*, vol. 8407. International Society for Optics and Photonics, 2012, p. 84070S.
- [27] L. Axon, J. R. Nurse, M. Goldsmith, and S. Creese, "A formalised approach to designing sonification systems for network-security monitoring," *International Journal on Advances in Security*, vol. 10, no. 1-2, 2017.
- [28] M. Debashi and P. Vickers, "Sonification of network traffic flow for monitoring and situational awareness," *PloS one*, vol. 13, no. 4, p. e0195948, 2018.
- [29] P. Janata and E. Childs, "Marketbuzz: Sonification of real-time financial dataa." Georgia Institute of Technology, 2004.
- [30] S. Mezzarobba, M. Grassi, L. Pellegrini, M. Catalan, B. Kruger, G. Furlanis, P. Manganotti, and P. Bernardis, "Action observation plus sonification. a novel therapeutic protocol for parkinson's patient with freezing of gait," *Frontiers in neurology*, vol. 8, p. 723, 2018.
- [31] M. J. Buehler, "Nanomechanical sonification of the 2019-ncov coronavirus spike protein through a materiomusical approach," *arXiv preprint arXiv:2003.14258*, 2020.
- [32] R. F. Lyon, "Automatic gain control in cochlear mechanics," in *The mechanics and biophysics of hearing*. Springer, 1990, pp. 395–402.
- [33] S. R. Buló and M. Pelillo, "Probabilistic clustering using the baum-eagon inequality," in *2010 20th International Conference on Pattern Recognition*. IEEE, 2010, pp. 1429–1432.
- [34] J. Parvizi, K. Gururangan, B. Razavi, and C. Chafe, "Detecting silent seizures by their sound," *Epilepsia*, vol. 59, no. 4, pp. 877–884, 2018.
- [35] B. H. Scheid, A. Ashourvan, J. Stiso, K. A. Davis, F. Mikhail, F. Pasqualetti, B. Litt, and D. S. Bassett, "Time-evolving controllability of effective connectivity networks during seizure progression," *Proceedings of the National Academy of Sciences*, vol. 118, no. 5, 2021.
- [36] W. Bomela, S. Wang, C.-A. Chou, and J.-S. Li, "Real-time inference and detection of disruptive eeg networks for epileptic seizures," *Scientific Reports*, vol. 10, no. 1, pp. 1–10, 2020.
- [37] J.-W. Lin, W. Chen, C.-P. Shen, M.-J. Chiu, Y.-H. Kao, F. Lai, Q. Zhao, and A. Cichocki, "Visualization and sonification of long-term epilepsy electroencephalogram monitoring," *Journal of Medical and Biological Engineering*, vol. 38, no. 6, pp. 943–952, 2018.
- [38] A. Delorme and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [39] "Childhood epilepsy: The brain," <https://www.massgeneral.org/children/epilepsy/education/the-brain>.
- [40] L. Lemus, A. Hernández, R. Luna, A. Zainos, and R. Romo, "Do sensory cortices process more than one sensory modality during perceptual judgments?" *Neuron*, vol. 67, no. 2, pp. 335–348, 2010.
- [41] G. H. Recanzone, "Interactions of auditory and visual stimuli in space and time," *Hearing research*, vol. 258, no. 1-2, pp. 89–99, 2009.
- [42] H. Jiang, B. E. Stein, and J. G. McHaffie, "Multisensory training reverses midbrain lesion-induced changes and ameliorates haemianopia," *Nature communications*, vol. 6, no. 1, pp. 1–11, 2015.
- [43] N. R. Stiles, V. R. Patel, and J. D. Weiland, "Multisensory perception in argus ii retinal prosthesis patients: Leveraging auditory-visual mappings to enhance prosthesis outcomes," *Vision Research*, vol. 182, pp. 58–68, 2021.
- [44] S.-R. Shiang, A. Gershman, and J. H. Oh, "A generalized model for multimodal perception," in *Proceedings of AAAI '17 Fall Symposium*, November 2017.
- [45] C. Schwartz, "Nonlinear operators. ii," *Journal of Mathematical Physics*, vol. 38, no. 7, pp. 3841–3862, 1997.