SONIFICATION OF MARKOV CHAIN MONTE CARLO SIMULATIONS

T. Hermann — Faculty of Technology, Bielefeld University, Germany, thermann@techfak.uni-bielefeld.de

M. H. Hansen — Bell Laboratories, Murray Hill, New Jersey, cocteau@bell-labs.com

H. Ritter — Faculty of Technology, Bielefeld University, Germany, helge@techfak.uni-bielefeld.de

ABSTRACT

Markov chain Monte Carlo (McMC) simulation is a popular computational tool for making inferences from complex, high-dimensional probability densities. Given a particular target density p, the idea behind this technique is to simulate a Markov chain that has p as its stationary distribution. To be successful, the chain needs to be run long enough so that the distribution of the current draw is close to the target density. Unfortunately, very few diagnostic tools exist to monitor characteristics of the chain.

In this paper, we present a new approach to render sonifications of McMC simulations. The proposed method consists of several auditory streams which provide information about the behavior of the Markov chain. In particular, we focus on uncovering modes in the target density function. In addition to monitoring, we have found our sonification to be an effective means for understanding the structure of high-dimensional densities. We have also applied our method to the exploratory analysis of high-dimensional data sets. In this case, we take as our target p a nonparametric density estimate obtained from the data. In this paper, we present a detailed description of our sonification design and illustrate its performance on test cases consisting of both synthetic and real-world data sets. Sound examples are also given.

1. INTRODUCTION

This paper illustrates the use of sonification as a tool for monitoring the performance of Markov chain Monte Carlo (McMC) simulations. Over the last decade, McMC has emerged as one of the most popular computational tools for making inference about complex, high-dimensional density functions. Suppose that we are interested in a target density $p(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d$. Traditional Monte Carlo methods for estimating various features of p use an (independent) sample of points drawn from p. McMC methods are typically employed when it is not possible to sample directly from p.[1] Instead, we generate a sequence of random points $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots\}$ whose distributions converge to p. The samples are drawn sequentially, often with the distribution of $\mathbf{x}_t$ depending only on the value drawn for $\mathbf{x}_{t-1}$; hence the series forms a simple first-order Markov chain. Implicitly, we assume that it is relatively easy to generate samples from this chain. In Section 2 we present the Metropolis algorithm for constructing a Markov chain for a target density p. Considerable research in the statistics literature has extended McMC techniques to very complex modeling situations. In simple cases, it is possible to derive the theoretical convergence rate of the McMC simulation to the stationary distribution p. In most practically important settings, however, determining when the chain has converged is a difficult task. If the chain exhibits poor convergence properties, inference becomes difficult.

In this paper we design several auditory streams that aid in assessing characteristics of an McMC simulation. Our methods can be applied to virtually any chain, regardless of the dimension of p. Our sonification tracks the sequence $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots\}$, so that at time t the audio signal provides information about both the value $p(\mathbf{x}_t)$ as well as the mode structure of p near $\mathbf{x}_t$. In the examples presented in Section 5, we can easily hear how the sampler moves between modes. Higher-level dependence is also audible, as we notice a repeated series of transitions from one mode to the next. The frequency of transitions provides information about the neighborhood between modes. This information is not used explicitly for the sonification but emerges as a consequence of this specific stochastic process. In addition to information about p, we also designed an audio stream that conveys more technical details about how the sequence is constructed. This channel can supply valuable diagnostics about the efficiency of the McMC simulation.

By experimenting with this sonification model, we found that our method also provides considerable insight into the structure of the target density p. For example, we can easily hear the number of modes of p, as well as their size and shape. This kind of insight is extremely helpful in high-dimensional settings. When p is a function of a small number of variables ($d \leq 4$), one can use traditional visualization methods to understand the important features of p. When $d > 4$, however, visual techniques begin to fail, and the resulting plots are more difficult to interpret. The output from a converged McMC simulation provides valuable insights about p that are difficult to capture by purely visual means. Carrying these experiences one step further, we have also applied this sonification scheme to high-dimensional data sets. Here, the target distribution p is taken to be a nonparametric density estimate. In our examples we have used a simple kernel estimator, but any nonparametric technique will work. The auditory streams that track modes in p now provide direct information about clusters in the data.

Our auditory sense is well suited to help orient and guide us toward interesting events in the real world. It is able to process many different information streams and (after minor training) it is excellent in detecting even subtle acoustic patterns. Both of these strengths of our hearing system are exploited in our McMC sonification. Our approach is a hybrid that uses both Model-Based Sonification (see Section 3.1) and Parameter Mapping [2]. The overall design is presented in Section 3. Additional new elements in this sonification are the concept of *Auditory Information Buckets*, which provide a way to *zoom out* acoustically, and nonlinear pitch mapping to facilitate mode distinction and comparisons. These

[1] This situation arises frequently in Bayesian models. An excellent reference is [1].

concepts are discussed in Section 4. Section 5 presents some example data sets and sounds. The paper ends with conclusions and prospects for future work.

2. MCMC SIMULATION

Traditional Monte Carlo techniques use an independent sample of data drawn from a target density p to estimate various features of p. In many practical settings, the density under study is too complex to be used directly and we are forced to rely on asymptotic approximations, numerical integration or McMC methods. The idea behind McMC is that we generate a sequence $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots\}$ that has as its stationary distribution the target density p. In this section, we present a simple McMC scheme known as the Metropolis algorithm [3]. We also illustrate how the output from this simulation can be used to infer properties of p. In the statistics literature, most applications of McMC are associated with so-called Bayesian models. In this case, the variable $\mathbf{x}$ is a vector of parameters in a probability model and p is a *posterior distribution* for $\mathbf{x}$. The characteristics of p relate directly to the uncertainty present in the components of $\mathbf{x}$. McMC can be applied more generally, however, and throughout this section we refer somewhat generically to a density p.

To implement the Metropolis algorithm, we first identify a suitable *jumping distribution*, $J(\mathbf{x}_b|\mathbf{x}_a)$ where $\mathbf{x}_a, \mathbf{x}_b \in \mathbb{R}^d$. We require that J be symmetric; or rather that $J(\mathbf{x}_b|\mathbf{x}_a) = J(\mathbf{x}_a|\mathbf{x}_b)$ for all values of $\mathbf{x}_a$ and $\mathbf{x}_b$.[2] To move the chain from $\mathbf{x}_{t-1}$ to $\mathbf{x}_t$, we first draw a point $\mathbf{x}^*$ from the distribution $J(\mathbf{x}^*|\mathbf{x}_{t-1})$. We then compute the *acceptance ratio*

$$r = \frac{p(\mathbf{x}^*)}{p(\mathbf{x}_{t-1})}. \qquad (1)$$

Finally, with probability $\min(r, 1)$, we set $\mathbf{x}_t = \mathbf{x}^*$; otherwise we remain at $\mathbf{x}_t = \mathbf{x}_{t-1}$. We take as our initial state, $\mathbf{x}_0$, a random point for which $p(\mathbf{x}_0) > 0$. To be practically useful, we need to be able to easily draw samples $\mathbf{x}^*$ from the jumping distribution.

Under this simple scheme, it is not difficult to show that the distributions of the samples $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots\}$ converge to p [3]. The qualitative properties of this Markov chain depend on J. For example, suppose we let J be a Gaussian distribution with variance-covariance matrix $\sigma^2 I_d$, where I_d is the $d \times d$ identity matrix. Then,

$$J(\mathbf{x}_a|\mathbf{x}_b) = \left(\frac{1}{2\pi\sigma^2}\right)^{d/2} \exp\left(-\frac{\|\mathbf{x}_a - \mathbf{x}_b\|^2}{2\sigma^2}\right), \qquad (2)$$

where $\|\cdot\|$ is the standard Euclidean norm on $\mathbb{R}^d$. If σ^2 is small compared to the variance of p, the probability that our chain moves between the different modes of p is small; hence we remain near the same mode for a long time. On the other hand, if σ^2 is very large, the acceptance ratio for each proposed move tends to be small and we rarely leave our current position. Therefore, while convergence is guaranteed at least theoretically for many choices of J, the jumping distribution has considerable influence on the finite-sample properties of the chain. Figure 1 shows the output from several runs of the Metropolis algorithm for a two-dimensional target density. The jumping distributions are Gaussians with σ^2 taking small, medium and large values.

[2] Other technical conditions relating to the support of p must also be satisfied, but these are beyond the scope of this paper. See [4]

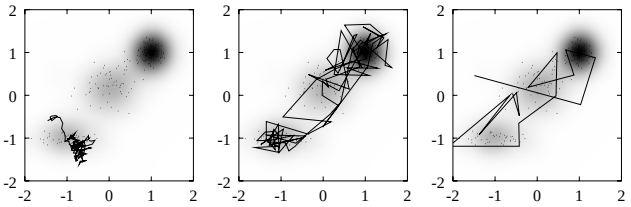


Figure 1: McMC random walk in a 2d distribution with 3 clusters. Grey values represent probability density, data points are plotted with points. 200 McMC steps are shown as line for jumping distribution with variance (a) 10 %, (b) 80 % , (c) 400% of the data set variance. (a) shows low mixing, (c) has only few accepted moves.

The Metropolis algorithm is perhaps the simplest technique for quickly generating a Markov chain with the correct stationary distribution. Many other schemes exist in the statistics literature that extend this approach to much more elaborate modeling contexts. In general, the samples $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots\}$ are used to estimate properties of p like its mean and variance. When p represents a statistical model, understanding its mode structure is an important component in making inferences about the system under study. While this introduction to McMC has been brief, it is sufficient to motivate our sonification model. Our design is sufficiently extensible to provide important information about the behavior of much more complex McMC schemes.

3. MCMC SONIFICATIONS

Mapping numerical data values to attributes of acoustic events, and superimposing a set of these events for all or part of the records in a data set is a popular approach to sonification. This technique is known as Parameter Mapping [2]. Because a large number of sound attributes are available, this idea seems quite promising. Unfortunately, the more variables we map in this way, the more difficult it is to relate sound characteristics back to features in the data, requiring the construction of a code book to draw any conclusions from the sonification. In addition, the number of available sound attributes may be too small for very high-dimensional data so that only a subset of the variables can be used for sonification. These drawbacks are partly overcome by Model-Based Sonification, which was proposed recently in [5]. Model-Based Sonification uses the viewpoint, that sound is better characterized by its sound source and the sound generating processes than by isolated attributes of the sound signal like pitch or envelope shape. A framework for describing sound and hearing based on this perspective can be found in [6, 7, 8].

3.1. Model-Based Sonification

Model-Based Sonification begins with the observation that our auditory system is well-optimized for the interpretation of sounds occurring in our physical environment. The source of these sounds is always a physical material, whose dynamics allows for motions which can be perceived as sound. Most materials are in a state of equilibrium without interaction and thus do not produce sound. Often, humans themselves excite acoustic systems by hitting, striking or shaking a physical system and thus put energy into it. In this case, sound is an important feedback which communicates properties of the material. Arguing that our brain is tuned

to draw information about material properties from such sounds, the goal of Model-Based Sonification is to carry this cycle over to data sonification. A sonification model is a way to build up a kind of virtual *data material* from the data set by identifying dynamical elements, providing a recipe for the dynamics and specifying the types of allowable interactions. The user then might explore the data set by shaking, plucking or hitting the virtual data material. The main advantages of this Model-Based Sonification approach are as follows:

- Sonification models need only a few parameters whose effect on the sound is easily understood and that remain the same independent of the data under consideration.
- Properly designed models can be applied to arbitrary high-dimensional data. As the data is implicitly encoded into sound there is no need for prior feature selection or dimension reduction.
- Knowledge about the model facilitates or even enables an understanding of the sound – no further code book is needed.
- Interaction with a sounding object (in this case the data material) by excitation is something human users are familiar with, so they will not be annoyed by the sound.

Examples of Model-Based Sonification include data sonograms, data set spring meshes and particle trajectories in a data potential, all of which are described in [5]. For our McMC sonification, we extend the particle trajectory model and construct a potential function based on a target density p. To motivate this approach, we first review the original particle idea developed in [5].

3.2. Particle Trajectories in a Data Potential

In this sonification model, N data records $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$, are interpreted as fixed masses in a d-dimensional Euclidean space ($\mathbf{y}_i \in \mathbb{R}^d$), contributing to a global potential

$$V(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} \phi_i(\mathbf{x}) \tag{3}$$

with

$$\phi_i(\mathbf{x}) = -\left(\frac{1}{2\pi\sigma^2}\right)^{d/2} \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}_i\|^2}{2\sigma^2}\right) ,$$

Unlike the potential appearing in Newton's law of gravitation, our expression (3) makes use of a flipped Gaussian to avoid singularities. In this sonification model, the data points $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$ are fixed. The dynamical elements are test particles which are thrown into that space at random positions with a given kinetic energy. They are attracted by the data points and thus follow a path according to a given dynamics

$$m\ddot{\mathbf{x}}(t) = -\nabla_{\mathbf{x}} V(\mathbf{x}) - \gamma\dot{\mathbf{x}}(t) , \tag{4}$$

assuming a mass m and friction coefficient γ. The particles' kinetic energy

$$T_{kin}(t) = \frac{1}{2} m\dot{\mathbf{x}}(t)^2 \tag{5}$$

is high-pass filtered and then taken as audio signal for the data sonification. Figure 2 shows a typical trajectory for a data set consisting of data points in 3 clusters. The choice of ϕ determines the

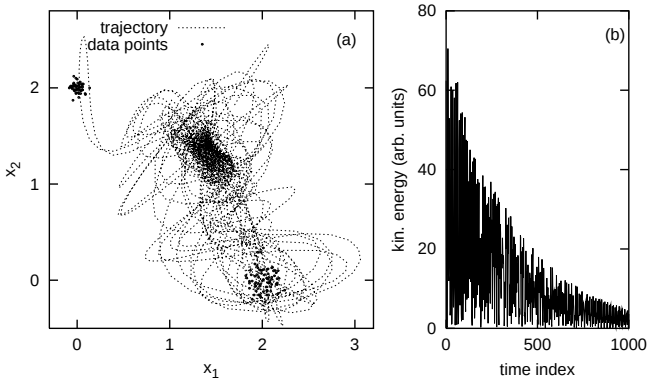


Figure 2: Illustration of particle trajectory audification. Trajectories of this kind are later used to sonify single McMC steps. (a) shows 3 data clusters and 1500 steps of a trajectory, (b) shows the corresponding kinetic particle energy T_{kin}. Obviously, the trajectory converges to the cluster in the middle. A sound example can be found on the Web page [9].

sound: close to the origin, the Gaussian potential can be approximated by a paraboloid which is known to lead to damped harmonic oscillations with constant pitch. Although the sound of an individual particle moving in V depends on its initial state, a qualitative behavior emerges from the sonification which can be perceived when an ensemble of particles is sonified simultaneously: clusters can be discerned as pitch groups. Starting the particles with velocity 0, they will (depending upon their starting position) converge to different troughs of the potential functions. As the pitch is determined by the curvature of the trough at the minimum, clusters of different size will give rise to differently pitched sounds. Besides this, we can make inference about the separation of the clusters based on the dynamical evolution of the particle sounds: particles that start with a high potential energy are able to move around between different clusters. These chaotic trajectories can be perceived as a noisy sound. The better two clusters are separated, the earlier this sound changes into a harmonic pattern. This sonification model was first introduced in [5] and was shown to be an effective tool for assessing the size and separation of clusters in a data set.

3.3. McMC Sonification Model

Our model for McMC sonification is based on the construction of particle trajectories similar to those explained in the previous section. In this case, however, we replace the potential $V(\mathbf{x})$ with a target density $p(\mathbf{x})$. Also, we change the sign of the force term in (4) to $+\nabla_{\mathbf{x}} p(\mathbf{x})$, so that the test particles are accelerated towards local maxima (or modes) of p. Particles then follow paths in the domain of p according to the dynamics

$$m\ddot{\mathbf{x}}(t) = \nabla_{\mathbf{x}} p(\mathbf{x}) - \gamma\dot{\mathbf{x}}(t) , \tag{6}$$

where we again specify the mass m of a particle and the friction coefficient γ. We recover the setup in (4) by taking p to be a kernel estimate of the density of data points $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$, using a Gaussian kernel with bandwidth σ^2.

Similar to the particle model in Section 3.2, the density p is unchanged during the sonification (corresponding to the fixed data

set $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$). The McMC simulation then provides samples in d-dimensional space. These are taken as starting points for a deterministic process which explores the local environment and represents the modal structure acoustically. Different choices have been considered for this step:

Method 1: Injecting a test particle with small kinetic energy and producing an audification of its kinetic energy;

Method 2: Injecting a test particle using high friction coefficient γ and using function values along its aperiodic trajectory to parameterize complex auditory grains[3]; and

Method 3: Running a mode search algorithm to find the nearest local maximum and using parameterized auditory grains to present the results.

The first method leads to particle trajectories which converge to the nearest mode, encoding the mode shape implicitly into its motion and thus into sound as described in Section 3.2. Close to the local maximum, the function has a pure parabolic shape which leads to quasiperiodic oscillations and thus to a pitched audio signal (mixture of discrete lines). The spectrum corresponds to the eigenvalues of the Hessian matrix at the mode center, and thus provides information about the shape of p near the mode. The main advantage of this approach is its conceptual simplicity: only few parameters need to be adjusted and the interpretation of the sound is tightly coupled with the process. Method 1, however, is extremely demanding computationally. For each McMC step, a complete particle trajectory must be computed which requires solving the differential equation (6) numerically. About 10,000 steps of the dynamics are required for a single Markov step.

Given this substantial overhead, we looked for modifications to the sonification model that might provide similar local information but with less effort. For Methods 2 and 3 we define a computationally simpler link between the stochastic McMC process and the rendered sound. However, these approaches require additional parameters that need to be set, making the auditory representation slightly more indirect. Method 2 reduces computation by taking a high friction coefficient for the particles. This scheme requires only about 10% of the time used by Method 1 for each McMC step, because the high friction quickly slows the particles to a stop. As each particle sweeps out its trajectory, the shape of the neighborhood close to the nearest local maximum can be perceived as a pitch variation pattern. Modes can be identified by their pitch which corresponds to the value of p. (Therefore, only modes with distinct density values can be discerned; we will address this problem in Section 3.5.) By using large friction forces, Method 2 is almost the same as a mode search by gradient ascent. For that reason, we considered Method 3, which finds the nearest mode more efficiently.

The available information during the McMC process can be divided into three groups: local data, which is valid only for a single step, global data about the McMC process, and mode-specific data which we compile in a mode database after each iteration. The mode database provides a summary of our current knowledge about the modal structure of p. Table 1 summarizes these three categories.

The trajectory audification as used in Method 1 provides information about the modes in an analogic manner, according to the taxonomy of Kramer [11]. However, we can also identify symbolic information by making explicit the attracting mode of the particle started at the current McMC step. This is done in Methods 2 and 3, where we create a list of mode structures that contains the coordinates of each mode and the of height of p, the number of McMC steps for which test particles were attracted to each mode, the average distance attracted particles traveled to reach the mode, and the covariance matrix of all contributing McMC steps. Most of this information is not displayed in the sonification at each McMC step, but instead is summarized through *Auditory Information Buckets* explained below.

Summarizing, the sonification model is a continuously running stochastic process, where interleaved deterministic steps are used to explore local attributes of $p(\mathbf{x})$. The sonification superimposes relevant information about this process in the domain of p. Using audification of particle trajectories like in Method 1, the sound can be very complex and hard to interpret. If auditory grains are used like in Methods 2 and 3, the ear seems better able to pro-

[3]The acoustic presentation for McMC step i is given in the sonification at a time $t_i = T \cdot i$ in form of a short audio signal of about 20-100 ms. This short signal is the basic element for granular synthesis [10] which we refer to as an "auditory grain".

Local Data	
$\mathbf{x}_i$	McMC step coordinates
$p(\mathbf{x}_i)$	McMC step function value
$\nabla p(\mathbf{x}_i)$	gradient
d_i	distance to last McMC step
r_i	acceptance ratio
$\mathbf{m}_i$	coordinates of nearest mode
m_i	index of nearest mode in mode database
$p(\mathbf{m}_i)$	mode density at nearest mode m_i
d_i^m	distance between McMC step and nearest mode
A_i	acceptance flag
η_i	nr. of steps until convergence in mode search
Global Data	
c_i^a, c_i^r	counter for accepted and rejected steps
N_m	number of modes in mode database
p_{max}	$= \max_j p(\mathbf{m}_j)$
Mode Database — for all $j < N_m$	
$p(\mathbf{m}_j)$	mode coordinates and density value
c_{ma}^j, c_{mr}^j	counter for accepted/rejected steps
δ_j	average distance of all attracted steps to mode
$\mathbf{v}_m^j$	mean of all attracted steps
Σ_j	covariance matrix of all attracted steps

Table 1: This table summarizes the available information on the McMC process.

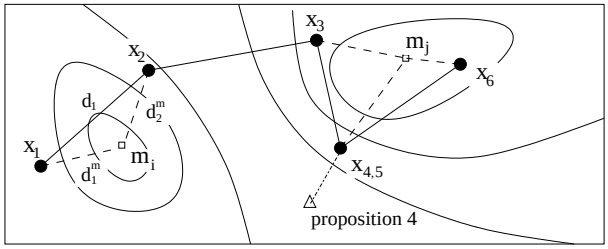


Figure 3: Illustration of 6 McMC steps in a bimodal distribution. At each McMC step, furthermore the density values $p(\mathbf{x}_i)$ and the gradient $\nabla p(\mathbf{x}_i)$ is available. At step 4, the proposition is rejected, so the McMC step remains at $\mathbf{x}_5 = \mathbf{x}_4$.

cess the sound. Therefore we enriched the model in this case with two additional streams.

Using a kind of parameter mapping technique to drive the sound attributes of the grains, several problems arise. Sound attributes normally show perceptual couplings, so that the perceived loudness of tones with identical amplitude strongly depends on the frequency of the tone. The implications of such perceptual interactions for the design of sonifications are investigated e.g. in [12]. To mitigate such effects, the selections for assignments from data attributes to sound attributes are mainly motivated by drawing analogies with the particle trajectory audification.

3.4. Auditory Information Buckets (AIB)

Information buckets provide a way to choose the granularity of information which is presented acoustically. A bucket can be thought of as a place where information is collected about certain time steps. Both a counter for the number of items in the bucket and the value(s) of the data are stored. Furthermore, a threshold is defined which limits the bucket size. Exceeding the threshold triggers a flushing of the bucket which results in the synthesis of a sonification for the bucket content. Thus the rate and complexity of these bucket sonifications can be easily chosen by adjusting the threshold. Within the McMC sonification model, buckets are used to summarize the characteristics of the modes. Each McMC step contributes to its respective mode bucket, where its starting position $\mathbf{x}_i$ is stored. On a flushing event for mode j, the data covariance matrix Σ_j of this sample is computed and a sound is rendered using the eigenvalues of Σ_j as described in Section 4.3. From the bucket sonification, we infer the shape and intrinsic dimension of the mode.

3.5. Nonlinear Pitch Mapping

A very important property of the auditory grains is their pitch. We design the auditory grains so that pitch correlates with the density value of the nearest mode. Doing this, however, two modes with very similar densities cannot be distinguished acoustically. This problem is solved by a nonlinear mapping function $z = g(\pi)$ from density values π to pitch z which allocates a higher portion of the available pitch range $[p_{min}, p_{max}]$ in density regions containing many values. At some fixed point in time, let $\mathbf{m}_1, \ldots, \mathbf{m}_{N_m}$ be the locations of the N_m modes discovered by the McMC sonification process. A resolution requirement is given by the density of modes along the π-axis which can be estimated from the sample of mode densities $\{\pi_1, \ldots, \pi_{N_m}\}$ with $\pi_j = p(\mathbf{m}_j)$. We then construct a pitch map via

$$g(\pi) = c(\pi) \cdot (p_{max} - p_{min}) + p_{min} \tag{7}$$

where

$$c(\pi) = \frac{1}{N_m} \sum_{j=1}^{N_m} \Phi\left(\frac{(\pi - \pi_j)^2}{\tau^2}\right), \tag{8}$$

is the cumulative distribution function (CDF) derived from a kernel density estimate of the mode distribution using a Gaussian kernel with bandwidth τ^2. In this expression, $\Phi(\cdot)$ is the standard normal CDF. A reasonable choice for τ^2 is the average distance between the π_j. Figure 4 shows the transfer function for a simple example. As a special case of this general approach, we might use the empirical (step function) CDF of the sample $\{\pi_1, \ldots, \pi_{N_m}\}$. If g is only evaluated at the π_j, this is equivalent maintaining an ordered list of modes and assigning pitch with a table lookup. Figure 4, (b) shows the corresponding pitch values for this approach. The effect of the nonlinear pitch mapping is to amplify pitch differences while maintaining the ordering.

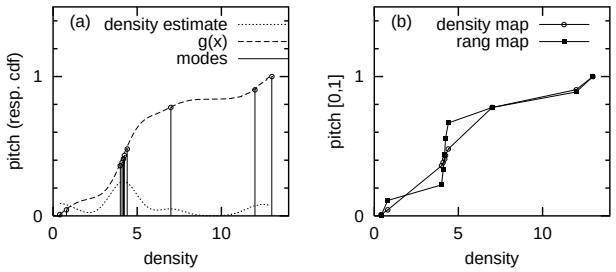


Figure 4: Nonlinear pitch mapping for a given sample of modes. (a) shows the kernel density estimate, the CDF and the modes and their assigned function values $g(\lambda)$. It can be seen how the nonlinear mapping increased the resolution at densities around 4. (b) shows the pitch assignment using the index of the density ordered mode list. Obviously, pitch differences are even stronger amplified

4. SONIFICATION DESIGN

The sonification currently uses 3 auditory streams which can be independently switched on or off. The first stream contains auditory grains using granular synthesis to present the Markov chain random walk through data space. The second stream provides information about rejected propositions of the Markov process, and thus informs the listener about the borders of the modes. The third stream contains auditory information buckets which summarize information about the modes in the sense of a *zoom out*.

4.1. McMC process monitoring stream

The basic element of the McMC sonification is the process monitoring stream, which provides transient information about the running McMC process. On each McMC step, one auditory grain is added to this stream. As the McMC process is a serial evolution in time, the step index i has the natural meaning of a time axis and thus "process time" is a linear mapping $t(i) = T \cdot i$, where T is a user-specified scaling factor. Depending upon the focus of attention, different scales are useful for inspection of the McMC process. With $T \approx 0.1\text{sec/step}$, individual steps can be resolved, whereas $T \approx 0.002\text{sec/step}$ yields an overall impression about different modes of p and their relation to each other. However, arranging a series of grains on a regular time grid can lead to either a monotonic rhythmical pattern or the perception of pitch at frequency $1/T$. To avoid these effects, we add a random time jitter of $T/4$ for the onset of the grains.

For the particle trajectory approach in Method 1, there is no need for further sonification design, as the sound is canonically derived from the particle's kinetic energy. The only parameters are the mass of the particle, the friction coefficient, and the sampling rate of the trajectory computation which can be adjusted by the system user. However, given currently available computation power, this audification approach is unsuited to inspect larger

datasets, so we will shift our discussion to the proposed alternatives which use auditory grains. This kind of sonification requires a number of design decisions which we now describe.

The auditory grain signals for Methods 2 and 3 are nonharmonic periodic functions which are multiplied with an envelope function having smooth attack and decay. The parameters of the grain synthesis are pitch, duration, volume and timbre (spectral composition, which is a multidimensional parameter). The choice of mapping is inspired by the acoustic properties of the audification, in order to maintain the meaning of the sound related to an imaginary process. In particle trajectory audification, the duration is dependent upon its initial energy $E_i^0 = p(\mathbf{m}_i) - p(\mathbf{x}_i)$. We truncate this value via $\max(0, \min(E_i^0, p_{max}))$ and then apply the linear map that carries $[0, p_{max}]$ to $[0.2T, 3T]$. In functional notation, the transformation is given by

$$\text{duration} = \text{map}(E_i^0, [0, p_{max}], [0.2T, 3T])\,. \tag{9}$$

For Method 1, pitch is varying while a particle converges to the nearest mode, $\mathbf{m}_j$, say. At some point, the motion of the particle stabilizes and the audification can be used to identify the mode by its pitch. This is maintained in Method 2 and 3 by driving the pitch of the the grains by $p(\mathbf{m}_j)$. Because this pitch depends on the density value $p(\mathbf{m}_j)$, it can be difficult to distinguish between modes of similar height. Therefore, we apply the nonlinear pitch mapping described in Section 3.5,

$$\log_2(\text{frequency}) = \text{map}(c(p(\mathbf{m}_i)), [0, 1], [8, 10]) \tag{10}$$

with $c(\cdot)$ given in equation 8. As $c(\cdot)$ is a monotone function, different modes can be compared with respect to their densities.

In the particle sonification of Method 1, volume is tightly coupled with E_i^0. For Methods 2 and 3, however, we are free to assign amplitude in a different but intuitive way: we use the amplitude of the grains to communicate the relevance of the mode, using loud grains for modes which are visited rarely. This is achieved by mapping $S = c_i^{m_i}/i$ to the grains amplitude

$$10\log_{10}(\text{amplitude}) = \text{const} + \text{map}(S^{-1}, [0, N_m], [0, 30]) \tag{11}$$

Thus amplitude corresponds to the "interestingness" of the event. New modes are introduced with loud grains, while the relative variation of amplitude heard as the McMC random walk moves around a mode provides information about how far we are from equilibrium. Figure 5 illustrates this stabilization process.

4.2. McMC details stream

This stream aims to provide more insight into the McMC process. Currently, a decaying white noise of 2 msec duration is added for all rejected propositions. From that, the listener gets a coarse impression about the efficiency of the McMC process.

Furthermore, for all accepted McMC steps some information is given about the distance to their nearest mode relative to the average distance of all steps which where attracted by the mode. This provides clues about whether the McMC random walk is exploring the center or the edge of a mode. Let i denote the index of the accepted step and let j be the index of the nearest mode. We use d_i^j to denote the distance between the position of the chain at i and the location of the mode $\mathbf{m}_j$. Finally, let δ_j be the average

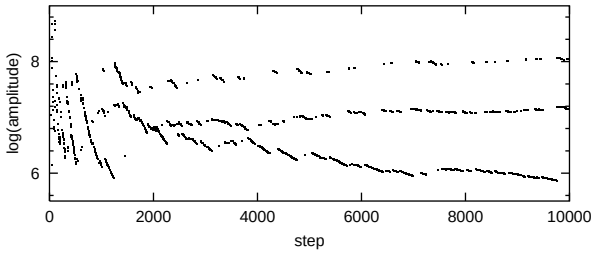


Figure 5: log(amplitudes) of the auditory grains in the McMC process monitoring stream for a density p with 3 modes of different probability mass. Whenever the McMC process moves around in one of the modes, the grain amplitude decreases. These volume differences get smaller if convergence is reached.

distance traveled by particles finding mode j. We map d_i^j/δ_j to the amplitude of a grain using

$$10\log_{10}(\text{amplitude}) = \text{map}(d_i^j/\delta_j, [1, 2], [0, 30]) \tag{12}$$

and use grains whose pitch is the same as in the McMC process stream but shifted 2 octaves higher. In so doing, we modify the perceived timbre of the grains in the McMC process stream: McMC steps from the tails of a mode are perceived as brighter sound.

4.3. AIB stream

The AIB stream provides a summary of what is known about the modes during the run of the McMC simulation. On each McMC step, a value of 1 is thrown into the respective AIB. If the bucket's content exceeds a threshold T_b, an AIB event is generated. We choose T_b to be a constant, so that the average number of bucket events per minute is so small that overlays are rare. The AIB events have a duration of about 0.5 sec, starting with a pitched tone whose pitch is a 2-octave downshifted version of the mode pitch. The duration of this mode marker is a mapping of the average distance, relative to the square root of the trace of the variance-covariance matrix associated with p, Σ. We can estimate this matrix using the McMC samples themselves. (When our target density p is a kernel smooth of a set of data $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$, we can estimate this matrix directly.) Thus small clusters or narrow modes are introduced with short tones, using

$$\text{duration} = \text{map}(\delta_j, [0, \sqrt{\text{trace}(\Sigma)}], [0.3, 1])\,. \tag{13}$$

After 100 msec, an uprising chain of percussive tones is started. These represent the ordered eigenvalues λ_k, $1 < k < d$ of Σ_j, the sample variance-covariance matrix of all McMC observations having mode j as their nearest mode. The number of tones played within this chain is given by

$$n = \min_l \left(\sum_{k=1}^{l} \lambda_k < 0.9\ \text{trace}(\Sigma_j) \right) \tag{14}$$

Thus, the number of tones provides insight into the intrinsic dimension of the mode. The tones within the arpeggio are located on a 50 msec time grid. The pitch of the kth tone is given by

$$\log_2(\text{frequency}) \approx \text{pitch} = \text{map}(k/d, [0, 1], [12, 13]) \tag{15}$$

where d is the dimension of the domain of p.

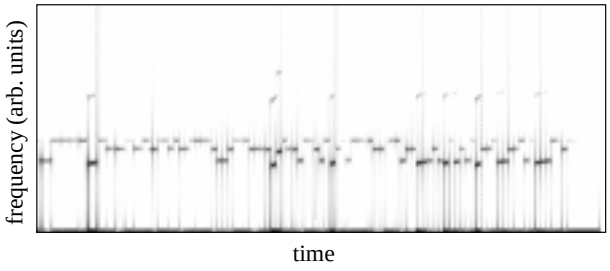


Figure 6: Short Time Fourier Transform of a sound signal obtained by sonifying 100 McMC steps in a distribution p which is a mixture of 3 Gaussians with different variances. The three different pitches correspond to the curvature of the modes as described in section 3.2.

5. EXAMPLES

In this section, we present examples of our sonification method applied to both synthetic and real world problems. Sound files for these examples can be found at the Web site [9]. We begin with a short demonstration of sound examples using an audification of particle trajectories as described in Method 1. All further examples are generated using auditory grains as in Methods 2 and 3. For the purpose of learning the sonifications, we split the sound examples in different pieces so that it gets easy to follow the different streams.

5.1. McMC Sonification using Trajectory Audification

This example (E1) presents a sonification using direct audification of the kinetic particle energy as described in Method 1. We demonstrate this sonification with a density p which is a mixture of 3 Gaussians with different variances and mixing proportions. For the example, every 10th McMC step is audified, so that the samples are almost uncorrelated. Figure 6 shows a spectrogram of the resulting sonification. The particle audifications are perceived as percussive sounds with rich spectra during the attack phase due to the nonharmonic shape of the attracting mode of p.

5.2. Cluster-Analysis

In this example, McMC sonification is used to explore clusters in a 6-dimensional data set. The data were drawn from a mixture of some spherical Gaussian distributions in 6d space. The mixing proportions, the covariance matrices, the cluster centers and the number of clusters are chosen randomly. In this simple example, the clustering can be easily depicted from a 2d-plot, as shown in Figure 7. Our target density p is a kernel density estimate derived from the sampled data using a Gaussian kernel with covariance matrix $0.2V$, where V is the sample 6×6 covariance matrix estimated from the data. The jumping distribution J is a simple Gaussian as in (2) with σ^2 taken to be 80% of the data set variance. This example is given to present the different streams and the information they provide. The first data set contains 3 clusters, one of them in a 3 dimensional submanifold, the others in a 5 dimensional subspace. All 3 clusters have a different probability mass. The time per step is $T = 0.01$ sec/step. Figure 8 shows a signal plot of the different streams.

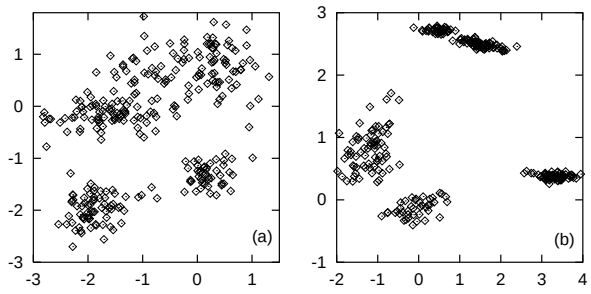


Figure 7: Clustered data in 6d space. The plots show 5 clusters, projected onto (a) axis x_0 and x_1, (b) first two principal axis. Plots depend strongly upon the selection of the axis. McMC sonification is independent upon rotations of the coordinate system.

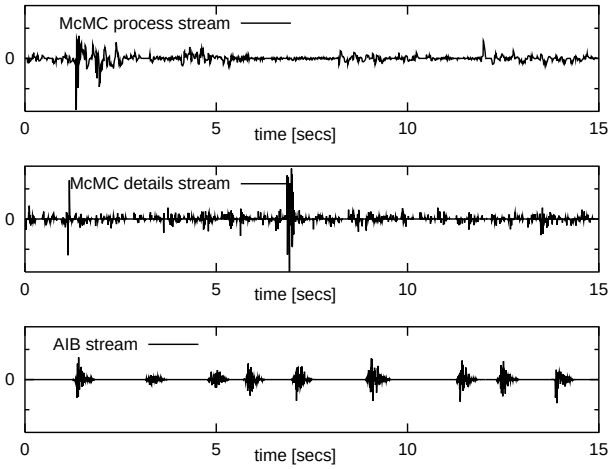


Figure 8: The signal plot of Example E2.a gives a course orientation about the content of the different streams and may be a help for the novice listener to distinguish parts in the mixed sound track.

On the Web page you find separate sound files for this sonification. Listen to first to example E2.a. In stream 1, we present the McMC process monitor. You can hear 3 different pitches which alternate. These are the pitches related to the modes as described above. You hear how the amplitude changes while the McMC steps move around in one mode. At the beginning you notice some high pitched percussive sounds: we add this marker whenever a new mode is found. From the absence of such sounds you know that all modes were found after the first 100 steps. Listening to the McMC details stream you observe that the noisy sound keeps rather constant. This means, that the rejection rate is independent of where the McMC step is. This is due to the large jumping variance. Reducing the variance of J, the noisy ticks get more infrequent. Also in this stream, you can hear pitched auditory grains, whose pitch corresponds to the mode. They are played when the sampler moves around in the tails of a mode. Together with the first stream this results in the brighter sound of the grain. The AIB stream contains the buckets. Here you can hear 3 different pitches. The middle one is the most frequent, and thus this mode has the highest probability mass (mixing proportion). From the eigenvalue arpeggio you can conclude that this cluster also exhibits variability

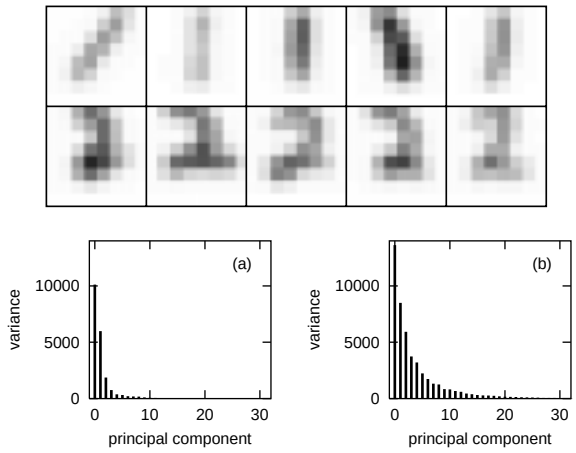


Figure 9: 5 instances of the digit classes '1' and '2'. Each record consists of the 64 dimensional vector of intensity values. The eigenvalue spectrum of the covariance matrix of all instances in class '1' and '2' are shown in figures (a) and (b). Obviously the instances in class '2' show a higher complexity.

along more than 3 directions. Example E2.b sonifies the same data set, this time with $T = 0.002$. We hear 6000 McMC steps in 12 secs. You can hear that transitions between the middle pitched and the higher pitched clusters are very frequent. Indeed, the cluster centers have a smaller distance. Example E2.c sonifies a different data set where 6 clusters are found in data space. Using $T = 0.008$ secs/step and an appropriate variance for J we can discern all 6 clusters.

5.3. Sonification of Handwritten Digits

In this example, the high-dimensional function is a kernel density estimate for a data set of bitmaps showing hand-written digits. Samples are given in Figure 9. We used the MNIST database [13] which offers a large data set of images of size 28×28 which we sampled down to 8×8. Each image can thus be represented as a point in a 64-dimensional data space, using the pixel intensities as features. The data set contains 1135 instances of class '1' and 1032 instances of class '2'. Using kernel density estimation, a 64-dimensional function p is generated which is sonified in Example E3.a. The sonification of the digits showing an '1' show less complexity than the distribution of the '2'. Listening to the AIB stream, the intrinsic dimension seems to be smaller as well. This is in agreement with the eigenvalue spectrum of the sample covariance matrices of these data sets (shown in Figure 9).

5.4. Convergence of McMC Simulations

With this example, we demonstrate how McMC sonification can be used to assess convergence of the Markov chain. We consider a Bayesian model and use McMC to simulate from the posterior distribution. The simplicity of the Metropolis algorithm introduced in Section 2 is ideal for motivating Markov chain methods, but is not sufficiently rich to cover most modern statistical applications. As the purpose of this paper is to highlight the audio representation, we will begin with audio streams derived from several runs of McMC. The underlying problem involves a Bayesian formulation of Poisson regression. The goal of the original analysis was to ascertain the effects of pollution on daily mortality counts in Philadelphia, Pennsylvania between the years 1974 and 1988. Because this is a Bayesian model, our target density is a posterior distribution on the model parameters. In this case, we have 186 parameters, leading to a Markov chain in 186-dimensional space. Our sound examples illustrate how changes in the form of the model, estimates for the hyperparameters as well as the proposal distribution (similar in spirit to the jumping distribution introduced in Section 2) effect the audio display. We also illustrate how the audible dependence in a chain effects variance estimates based on the samples $\{x_0, x_1, x_2, \dots\}$. See Example E4 on the Web site for the sounds and further description.

5.5. Sonification of the Iris Dataset

This example presents sonifications for the 5-dimensional iris dataset, which contains 3 classes of 50 instances each, where each class refers to a type of iris plant. Removing the class labels, 4-dimensional data vectors remain. Two clusters can be easily discerned. The presented McMC sonification also finds this structure.

This example discusses the influence of σ^2 on the sonification. As explained above, p is derived from the data set by kernel density estimation using a Gaussian kernel of bandwidth σ^2. Thus σ^2 can be adjusted to select a specific resolution while inspecting the data. Taking large values for σ^2, p shows only one mode, as can be heard in sound Example E5.a and E5.b. Reducing σ^2, p is slowly transformed to a bimodal distribution. This can be clearly heard in Example E5.c to E5.f. Obviously, the bimodal structure is quite stable under variation of σ^2. However, decreasing σ^2 further, a modal substructure of the main clusters gets apparent, which can be heard in Examples E5.g and E5.h.

6. CONCLUSION

We have presented a new tool for monitoring the convergence of McMC simulations. Markov chain methods are extremely popular for making inferences about complex, high-dimensional densities. Unfortunately, it can be difficult to assess important characteristics of a simulated chain with visual displays. Through a multi-stream sonification model, we are able to directly infer rather complex dependencies evident in a chain. After experimenting with this approach, we found that the sonification also provides information about the target density itself. Furthermore, by taking the target to be a nonparametric density estimate, this sonification tool can be applied to explore the local features of data sets. Using a Model-Based Sonification approach, the McMC sonification can be used for arbitrary densities/data and must only be learned once. It has only a few control parameters, which can easily be understood. While our current auditory display is already quite informative and gives a lot of interesting information about the function at hand, there are lots of possibilities to enrich it even further. For example, we are considering the use of spatialization, speech annotations and additional auditory streams. Another sensible enhancement involves pairing our sonification with a visual display. A projection of the data may be used as a map where certain regions may be highlighted aiming to amplify the amplitude of all grains related to McMC steps within the selected region. These and other extensions are the subject of our ongoing research and will be reported elsewhere.

7. REFERENCES

[1] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman & Hall, 1996.

[2] C. Scaletti, "Sound synthesis algorithms for auditory data representations," in *Auditory Display*, G. Kramer, Ed. 1994, Addison-Wesley.

[3] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, 1995.

[4] L. Tierney, "Markov chains for exploring posterior distributions (with discussion)," *the Annals of Statistics*, vol. 22, no. 2, pp. 1701–1727, 1994.

[5] T. Hermann and H. Ritter, "Listen to your Data: Model-Based Sonification for Data Analysis," in *Advances in intelligent computing and mulimedia systems*, M. R. Syed, Ed. 1999, Int. Inst. for Advanced Studies in System Research and Cybernetics.

[6] W. W. Gaver, "How do we hear in the world? Explorations in ecological acoustics," *Ecological Psychology*, vol. 5, no. 4, pp. 285–313, 1993.

[7] W. W. Gaver, "What in the world do we hear? An ecological approach to auditory source perception," *Ecological Psychology*, vol. 5, no. 1, pp. 1–29, 1993.

[8] W. W. Gaver, "Using and creating auditory icons," in *Auditory Display*, G. Kramer, Ed. 1994, Addison-Wesley.

[9] T. Hermann and M. H. Hansen, "Demonstration of sonification of McMC simulations," `http://www.techfak.uni-bielefeld.de/~thermann/projects/index.html`, 2001.

[10] C. Roads, "Automated granular synthesis of sound," *Computer Music Journal*, vol. 2 (2), pp. 61, 1978.

[11] G. Kramer, "An introduction to auditory display," in *Auditory Display*, G. Kramer, Ed. 1994, Addison-Wesley.

[12] J. G. Neuhoff, G. Kramer, and J. Wayand, "Sonification and the interaction of perceptual dimensions: Can the data get lost in the map," in *Proceedings of the Int. Conf. on Auditory Display*, P. R. Cook, Ed. 2000, ICAD.

[13] Y. LeCun, "The MNIST Database of Handwritten Digits," `http://www.research.att.com/~yann/ocr/mnist/index.html`, 2001.