# Human Perception of the Soundscape in a Metropolis through the Phenomenology of Neural Networks

**Enrica Santucci, Luca A. Ludovico**
Laboratorio di Informatica Musicale
Department of Computer Science
Università degli Studi di Milano
Via Comelico 39, 20135 Milano, Italy
{enrica.santucci,luca.ludovico}@unimi.it

## ABSTRACT

The aim of this work is providing a quantitative description of the evolution from a hi-fi to a lo-fi soundscape in a metropolis, according to the definition given by R. Murray Schafer. Our theoretical approach has a statistical mechanical foundation and is based on order/disorder transitions. We will show that a system formed by $C$ citizens and by the sound sources related to the $K$ neighborhoods of the metropolis is thermodynamically equivalent to the fully connected network of interacting citizens described by the Hamiltonian of the Hopfield model. The application of statistical-mechanics models to the soundscape will confer an original interpretation to the concepts of ferromagnetic and spin-glass phases: such phases will be associated to different models of society (ordered vs. frustrated system). Transitions will be regulated by the variation of the control parameter, here defined as the ratio between the number of sound sources and the number of citizens.

## 1. INTRODUCTION

Soundscape can be defined as the overall sonic environment of an area. Soundscape does not need to be considered either natural or artificial, since artificial and natural sounds often coexist in the same environment and sometimes become indistinguishable. For example, in a urban context most sound sources come from human activities; anyway, even in a city, sounds produced by non-human species or natural phenomena can be easily found.

Over the centuries, the soundscape has undergone a continuous and rapid evolution. The acoustic universe where contemporary man lives is radically different from any other that preceded it. Some sounds that characterized everyday life in a typical village of the pre-industrial era (ringing bells, craft-work tools, etc.) nowadays have almost disappeared even from rural areas. Investigating the many reasons why this happened go beyond the purpose of this work.

Many scientific works suggest that urban residents have low levels of awareness of soundscape and that the experience of modern urban life involves a high degree of sensory deprivation [1]. Meanwhile, many Countries are subscribing directives and policy intentions to preserve and manage quiet areas [2].

In our society, which is losing the capability to distinguish and preserve sounds, a number of initiatives aim at recording characteristic soundscapes. Many project started building huge databases of sound documents with different goals, ranging from the constitution of stations for permanent listening of artificial as well as natural sounds [3] to the preservation of Chinese classical garden's soundscape [4].

After the phase of recording and organizing digital data, a number of scientific works and initiatives addressed the problem of efficiently and effectively representing them. For example, [5] describes the key specifications of an information system for urban soundscapes. This work refines an earlier proposal of a 2D cartography for the auditory environment concerning urban areas. The visualisation of urban soundscape is realized through dynamic auditory maps enriched by graphic semiology especially dedicated to sounds. An animated version is proposed in [6].

A different approach is discussed in [7], where an XML-based format is employed to provide a multi-layer experience of environmental sounds. The adoption of the IEEE 1599 standard allows to encode within a unique document both visual and audio descriptions of paths and scenarios of the urban life, enabling features such as the comparison of a given soundscape at different hours of the day.

The importance of this matter is reflected also by the attention that architects and urban planners are recently paying to soundscape design. For instance, [8] addresses the problem of sound environment improvement in cities by comparing acousticians', city-users' and planners' categorizations of urban soundscapes. As regards the quality and variability of soundscapes in public and open spaces, it is worth citing - among many others - [9], [10], and [11].

The soundscape we live in can deeply influence the activity of our brain and our emotional response. For instance, the *Positive Soundscape Project* [12] is a large interdisciplinary study that provides a conceptual framework to link the key soundscape components to the response of the

brain, subjected to both affective and neutral soundscapes. Scientific research has been conducted through qualitative fieldwork, lab-based listening tests, and physiological validation. As a result, a number of metrics have been suggested for some perceptual scales.

An interesting experience about audio-visual interaction is described in [13], where the analysis of preferences for combinations of four sounds and eight visual landscapes revealed an interesting result: it is the sound and not the visual component which seems to influence the patterns of preference. This aspect stresses the affective role of sound in combination with the visual setting.

The soundscape also produces an effect on the sound perception by people, influencing the daily activities and even the social interaction within a community. In fact, the evolution of the soundscape goes hand in hand with the transformation of people's lifestyle. The acoustic environment is an effective indicator of the social conditions that produced it, and it provides information on the development lines of a given society [14].

## 2. HI-FI VS. LO-FI SOUNDSCAPES

The key concept of *acoustic community* has been introduced by B. Truax in [15]. In his opinion, when a sound is clearly heard within a given area and reflects the community, it creates a unifying relationship with the environment. In particular, acoustic cues and signals can help to define an area from a spatial, temporal, social and cultural perspective.

Thus, an acoustic community is linked and defined by those sounds that identify not only daily and seasonal cycles, but also shared activities, rituals and dominant institutions of the area. A community with a good acoustic definition can easily recognize, identify and derive meaning from the soundscape. According to [16], the acoustic features that uniquely characterize a specific soundscape are: sound signals, keynotes, and soundmarks. Through years of listening, these sounds provide strong associations among members of the community and thereby create a continuity with the past.

In the Sixties, R. Murray Schafer caught the essential point of the problem by noting an increasing loss of acoustic identity of human settlements, which consequently reflects on the loss of social identity.

The background noise, predominant almost everywhere, leads to the disappearance of acoustic forms, as well as of the aesthetic plurality and the various depth degrees related to the sounds. Before industrialization the acoustic sphere around the subject could be divided into a foreground zone, an intermediate zone and a background zone. For instance, the sound of a conversation we are engaged in would belong to the foreground, the birds of the garden around us to the intermediate, and the church bells that resonate throughout the city to the background zone. The risk underlined by Schafer is the disappearing in modern cities of heterogeneity and the appearance of an identical and homogeneous sound. At this regard, the key step is the transition from pre-industrial to post-industrial society, that corresponds to the transition from a *hi-fi* to a *lo-fi soundscape*

respectively.

According to Schafer, in a *hi-fi soundscape* we can clearly perceive sounds and their orientation (i.e. location and distance) within the acoustic space. A hi-fi soundscape is an environment where "sounds overlap less frequently; there is perspective - foreground and background". In the atmosphere of a silent hi-fi soundscape, even the most negligible disturbance can communicate important information of vital interest.

On the contrary, in a *lo-fi soundscape* the sonic space is confused, individual sounds lose their identity within a superabundant sonic presence, and masking is common. There is interference in all channels and even the most ordinary sounds must be amplified in order to be heard. As a consequence of amplification, the community becomes noisier and noisier and this adds further difficulty in identifying sound sources.

The lo-fi soundscape comes from congestion. It dates back to the industrial revolution, when new sounds appeared and many natural and archetypal sounds ended up being blacked out. A further development was due to the electric revolution, which introduced new effects and mechanisms able to package sound and transmit in a schizophrenic way through time and space. Today the world is suffering a superabundance of sound and media information in general, so that only a small part can be clearly perceived.

During the transition from pre-industrial to post-industrial society, i.e. from a hi-fi to a low-fi soundscape, the community gets in touch with new sounds and noises having different quality and intensity with respect to the past. Noise was born as a consequence of the invention of machines and nowadays it is present not only in the roaring atmosphere of big cities but also in rural areas.

The main sources of urban noise are due to transport (cars, motorcycles, buses, ambulances, vehicles on rails) and loud work activities (industries, workshops, clubs, construction sites, entertainment venues, concert or outdoor events). Domestic noise is caused by household appliances (vacuum cleaner, hair dryer, washing machine, etc.), technological home devices (television, hi-fi stereo, home entertainment, etc.) and commercial activities close to our houses (shopping centers, restaurants, pubs, etc.).

The evolution of society due to technological discoveries - and the consequent introduction of new sound and noise sources - involves changes in the keynote sounds of the original soundscape. Nowadays, machines have created such sonic variety that pure sound no longer arouses emotion. Overwhelming sounds are so deeply impressed in the soul of people that they influence generalized behavior and lifestyle.

## 3. PERCEPTION OF THE SOUNDSCAPE: A THERMODYNAMIC APPROACH

The goal of this paper is giving an interpretation to the main effects of soundscape changes on human perception in order to understand the sociological differences associated with such variations. The novelty of our approach consists in using concepts from statistical mechanics and

the phenomenology of neural networks, considering the system from a macroscopic (i.e. thermodynamic) perspective. The process can be summarized as follows:

1. Describing the interaction between $C$ citizens and the soundscape related to $K$ neighborhoods belonging to a metropolis through a Hamiltonian function. The model is made of one layer of $C$ units (where $c_i$ is the soundscape perception of the $i$-th citizen) and $K$ disconnected layers (each one representing the soundscape of a specific neighborhood), respectively composed of $S_1, \ldots, S_K$ units (where $s_\mu^{(i)}$ is the $\mu$-th sound source listened by the $i$-th citizen);

2. Proving the thermodynamic equivalence, i.e. the equality of the Hamiltonian functions respectively, between the interacting model described above and a Hopfield model where the Hebbian kernel is the sum of the $K$ Hebbian kernels related to the $K$ neighborhoods.

Basic notions about statistical mechanics will be provided in Section 3.1, whereas a brief introduction to Hopfield networks will be presented in Section 3.2.

## 3.1 Fundamentals of Statistical Mechanics and Neural Networks

Statistical mechanics adopts a probabilistic approach in order to study the average and global behavior of a mechanical system made of a large number of particles, focusing the attention on how the relationships and the interactions among parts influence such a collective behavior. The task of describing a system with many degrees of freedom in a deterministic way - namely by studying the motion of each particle - would be very hard. This problem is solved by describing the system from a macroscopic point of view by means of thermodynamic quantities such as free energy, entropy, pressure, and so on. In other words, the goal is finding the phenomenological laws governing the global behavior of a given system, which is not easily deducible from the analysis of the microscopic laws that control each individual constituent.

Among many statistical mechanics models, we are particularly interested in those describing disordered complex systems. In this context, an important concept is the one of *spin glass*, which can be defined [17, 18] as a spin[1] system whose low temperature state appears as a disordered one. A spin glass is a disordered magnet with frustrated interactions, augmented by stochastic positions of the spins, where conflicting interactions are randomly distributed with comparable frequency. Figure 1 provides a schematic representation of the random spin structure of a spin glass (top) and the ordered one of a ferromagnet (bottom).

*Frustration* [19] is a key characteristic of a spin glass system, and it is generally a consequence of randomness. It represents the property for which the Hamiltonian, i.e. the cost function related to a given system, depends not only
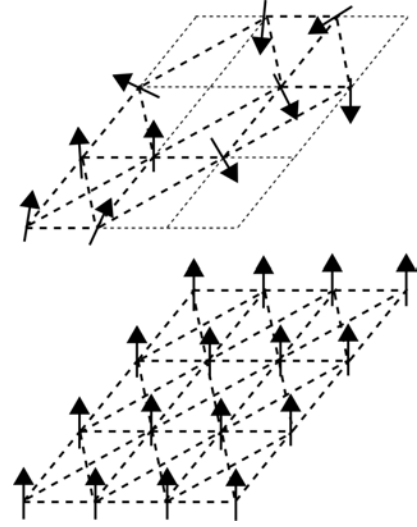
---

[1] Spin is the intrinsic angular momentum of an elementary particle.



**Figure 1**. Schematic representation of a spin glass structure versus a ferromagnet one.

on the configuration of the system and on the strength of the external fields acting on it, but also on random parameters such as the couplings between the spins.

Thanks to particular assumptions, it is possible to find analogies between neural networks theory and the statistical mechanics of frustrated disordered systems.

## 3.2 Hopfield Networks

One of the most popular models describing disordered complex systems is the Hopfield model [20], representing a neural network acting as an *associative memory*. Such a model updates the synaptic weights between neurons according to a specific learning rule, depending on the neural activity driven by a given set of observations. After learning, the network is able to generate a sequence of states whose probabilities match those of the observations.

A Hopfield network has a fully connected structure composed by $N$ neurons (with $N$ large), to which we associate a generic configuration

$$\sigma = (\sigma_1, \ldots, \sigma_N) \quad \in \quad S^N = \{-1, +1\}^N, \quad (1)$$

where $\sigma_i$ for $i = 1, \ldots, N$ are the neuronal binary states (i.e. spin variables) indicating quiescent or spiking neurons.

A stored pattern $\xi = (\xi_1, \ldots, \xi_N)$ is naturally described as a $N$-bit vector (that is an element of the set $S^N$). Each bit, both for the $\sigma$ and for the $\xi$ vector's values, is chosen randomly and independently with equal probability, i.e.

$$P(\rho_i) = \frac{1}{2}\delta(\rho_i - 1) + \frac{1}{2}\delta(\rho_i + 1) \quad (2)$$

with $\rho = \sigma$ and $\rho = \xi$ respectively, and where $\delta$ is the Dirac measure.

When we say that the patterns are *learned* or *stored* by the net, we mean that

$$\sigma_i = \xi_i \quad \forall\, i = 1, \ldots, N. \qquad (3)$$

Equation 3 indicates that each one of the network configurations, for each one of the $P$ patterns $\xi^{(1)} \ldots \xi^{(P)}$, is a fixed point of the dynamics, i.e. an absolute minimum for the free energy.

Taking into account a neural network where synapses fully connect its neurons, we can define the symmetric synaptic matrix $J_{i,j}$, which represents the way a specific neural network stores information, as

$$J_{i,j} = \frac{1}{N} \sum_{\mu=1}^{P} \xi_i^{(\mu)} \xi_j^{(\mu)} \quad \forall i, j = 1, \ldots, N. \qquad (4)$$

Equation 4 represents the Hebbian kernel of a Hopfield network.

Since in a Hopfield network there are no self connections, namely neurons do not connect directly to themselves, we can say that

$$J_{i,i} = 0 \quad \forall i = 1, \ldots, N, \qquad (5)$$

which is compatible with the model of connected citizens that will be detailed in Section 4.

Usually, when we study a generic physical system the main goal is to understand its different behaviors with respect to the variation of a specific parameter. Such different behaviors are the transition phases of the model, which are transformations of a thermodynamic system from a state of matter to another one due to an abrupt change of one or more physical properties [21].

Thanks to rigorous computations, in the Hopfield network case four phases have been discovered [22], depending on the variation of the control parameter defined as

$$\alpha = \frac{P}{N}. \qquad (6)$$

In the Hopfield model, in order to study phase transitions, we assume that the number of patterns $P$ increases linearly with the number of the neurons $N$. Once again, this condition is compatible with our application to the case study of soundscapes.

In the Hopfield model, phases are:

- *Ferromagnetic phase* **F** (for small values of $\alpha$), where the network is able to recall the stored patterns because they are global free energy minima [2];

- *Mixed phase* **M** (for intermediate values of $\alpha$), where pattern recognition is more difficult because of the increase of the $\alpha$ parameter (as a consequence of the increase of the number of patterns $P$); here, the stored patterns are metastable, i.e. the global free energy minima are spin-glass states;

- *Spin glass phase* **SG** (for high values of $\alpha$), where the network is not able to recall the stored patterns because the system is frustrated.

---

[2] In Thermodynamics, free energy is a macroscopic quantity depending on the Hamiltonian function.

- *Paramagnetic phase* **P**, where order - and consequently pattern retrieval - are destroyed by the high temperature $T$, regardless of the $\alpha$ values.

The corresponding phase diagram is shown in Figure 2, where *temperature* $T$ represents the noise level of the network.

In the next section we will contextualize such a model to the soundscape studies, generalizing the model described above to the case of multiple sets of stored patterns.
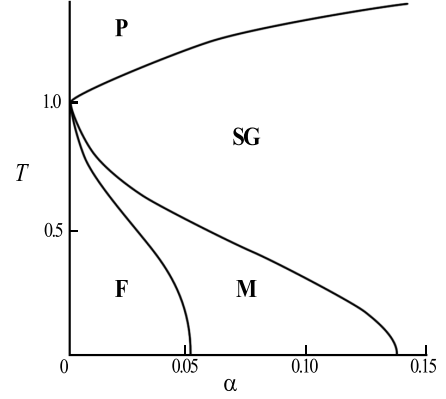


**Figure 2**. $T$ - $\alpha$ phase diagram.

## 4. THE MODEL

Let us consider a contemporary metropolis made of $C$ citizens and call

$$c_i \in \{-1, +1\} \text{ for } i = 1, \ldots, C$$

a dichotomous variable representing the soundscape perception of the $i$-th citizen during his/her *everyday listening*. In [23] this locution indicates the process through which people extract information about surrounding events by listening the sound they produce. Please note that everyday listening can be influenced not only by physical conditions but also by cultural, educational and social aspects.
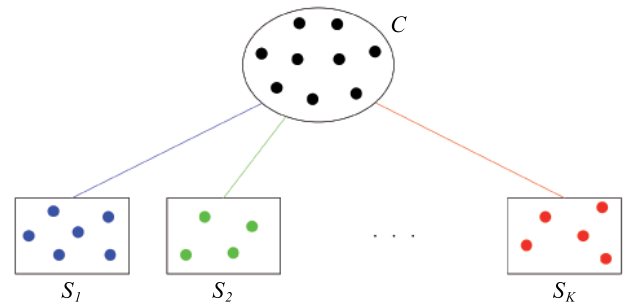


**Figure 3**. Schematic representation of the interaction between the set $C$ of citizens and the soundscape related to $K$ city's neighborhoods. We assume that there is no connections between the layers representing the neighborhoods.

The variable $c_i$ is set to $+1$ if the surrounding soundscape causes a feeling of wellness and to $-1$ if a sense of discomfort is felt. We can assume the occurrence of such values is equiprobable, in accordance to (2).

We stress that sound events in the environment can provide information about hazards or risk situations, as well as reassuring episodes and this obviously has a psychological impact on the individual. In other words, the surrounding soundscape is able to raise emotions in the audience.

Now we suppose that citizens, during their daily life, come in contact with a multitude of sounds and noises when they move across a metropolis to go home, to go to work, to reach entertainment venues, etc.

In order to model this aspect, we introduce $K$ different neighborhoods, each one with its own soundscape reflecting specific socio-economic features. The $K$ neighborhood-related soundscapes are $K$ sets of sound sources

$$\{s^{(1)}\}_{\mu_1} \quad \text{for } \mu_1 = 1, \ldots, S_1$$
$$\vdots$$
$$\{s^{(K)}\}_{\mu_K} \quad \text{for } \mu_K = 1, \ldots, S_K$$

that a priori we can assume to be distributed on the whole territory according to a Normal distribution, i.e.

$$P(s_\mu^{(i)}) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(s_\mu^{(i)})^2}{2}\right]. \qquad (7)$$

Finally, we have to introduce the interactions among people and surrounding sounds $\xi^{(\mu_1)}, \ldots, \xi^{(\mu_K)}$ by representing them as dichotomous variables. As regards their value:

- $\xi_i^{(\mu)} = +1$ when such interaction is positive, in the sense that the $i$-th citizen feels a positive emotion listening to the sound produced by the $\mu$-th source;

- $\xi_i^{(\mu)} = -1$ if the interaction is negative.

The Hamiltonian of our system can be defined as:

$$\mathcal{H}_{C,S}(c, s; \xi) = -\frac{1}{\sqrt{C}}\Big[\sum_{i=1}^{C} c_i \sum_{\mu_1=1}^{S_1} \xi_i^{(\mu_1)} s_{\mu_1}^{(i)} +$$
$$+ \sum_{\mu_2=1}^{S_2} \xi_i^{(\mu_2)} s_{\mu_2}^{(i)} + \ldots + \sum_{\mu_K=1}^{S_K} \xi_i^{(\mu_K)} s_{\mu_K}^{(i)}\Big] \quad (8)$$

where $\beta = 1/T$, namely it is the reciprocal of the network's noise level $T$. Parameter $\beta$ is the strength of the interactions depending on a number of factors: for example, on the distance between citizen and sound sources, on the frequency emitted by a given source, etc.

It is worth underlining that, in accordance with the Hamiltonian defined above, we are adopting a simplified model where all citizens are exposed to all city sounds. In mathematical terms, we are using a *mean field approximation*.

Equation 8 represents the Hamiltonian of a Hybrid Boltzmann Machine [24] containing one layer of visible units $\{c_i\}_{i=1}^{C}$ and $K$ non interacting layers of hidden units. Since we are interested in the perception of the soundscape by the citizens, in our approach the citizens' perceptions constitute the visible layer of the neural network, whereas the soundscape sources are modeled through the hidden layers.

Please note that we are considering a restricted version of a Hybrid Boltzmann Machine (rHBM) as we suppose that each neighborhood has its own soundscape which does not influence the other ones, as shown in Figure 3.
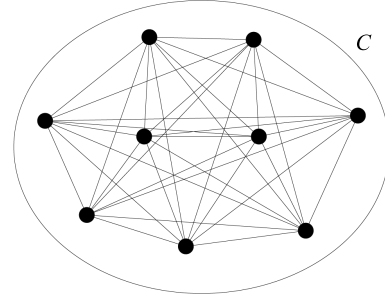


**Figure 4**. Schematic representation of the equivalent Hopfield neural network built upon the visible units $c_i$ only, with an internal fully connected structure.

## 5. THERMODYNAMIC EQUIVALENCE

This section will demonstrate that a rHBM made of one layer of visible units and $K$ disconnected layers of hidden units (see Section 4) is thermodynamically equivalent to a Hopfield model, namely a fully connected network of only visible units.

First, we introduce the *partition function* related to the Hamiltonian 8, which is

$$\mathcal{Z}_{C,S}(\beta; \xi) = \sum_{\{c\}} \int \prod_{\mu_1=1}^{S_1} ds_{\mu_1}^{(1)} \ldots$$
$$\ldots \int \prod_{\mu_K=1}^{S_K} ds_{\mu_K}^{(K)} \exp[-\sqrt{\beta}\mathcal{H}_{C,S}(c, s; \xi)]. \quad (9)$$

Now, by applying the Gaussian integral represented by

$$\int_{-\infty}^{+\infty} dz \exp\left[-\beta\left(\frac{z^2}{2} - az\right)\right] = \sqrt{2\pi} \exp\left(\frac{a^2}{2}\right) \tag{10}$$

we obtain the partition function

$$\mathcal{Z}_C(\beta, \xi) \propto \sum_c \exp(-\beta\mathcal{H}_C(c; \xi)), \qquad (11)$$

where

$$\mathcal{H}_C(c; \xi) = -\frac{1}{C}\sum_{i<j}^{C} J_{i,j} c_i c_j \qquad (12)$$

and

$$J_{i,j} = \sum_{\mu_1=1}^{S_1} \xi_i^{(\mu_1)}\xi_j^{(\mu_1)} + \sum_{\mu_2=1}^{S_2} \xi_i^{(\mu_2)}\xi_j^{(\mu_2)} + \ldots +$$
$$+ \sum_{\mu_K=1}^{S_K} \xi_i^{(\mu_K)}\xi_j^{(\mu_K)}. \quad (13)$$

Equation 12 represents the Hamiltonian of a Hopfield model where the Hebbian kernel is the sum of the $K$ Hebbian kernels related to $K$ different sets of patterns, where

$$J_{i,i} = 0, \ \forall i = 1, \ldots, C \quad (14)$$

because, as we briefly mentioned in the previous section, we assume that citizens do not communicate with themselves.

We can infer that the system described by the Hamiltonian in Equation 8 can be mapped into that one described by the Hamiltonian in Equation 12, where $\xi^{(\mu_1)}, \ldots, \xi^{(\mu_K)}$, i.e. the interactions among citizens and the sounds composing the soundscape, are in the latter case the patterns stored by the citizens network. In this case, patterns correspond to the emotions felt by citizens.

As a consequence, we can discover the main phase transitions of the aforementioned Hopfield model and associate to them different global behaviors of the network from a social and psychological point of view.

First, we define $K$ control parameters denoting the ratio between the number of sound sources $S_1, \ldots, S_K$ in each neighborhood and number of the citizens $C$ respectively:

$$\alpha_1 = \frac{S_1}{C}, \ \alpha_2 = \frac{S_2}{C}, \ \ldots \ \alpha_K = \frac{S_K}{C}. \quad (15)$$

The thermodynamic properties of the Hopfield model, in terms of memory, depend on how the numbers $S_1, \ldots, S_K$ of stored patterns scale with the size $C$ of the system, i.e. on the choice of the parameters $\alpha_1, \ldots, \alpha_K$. We can assume that $S_1, \ldots, S_K$, i.e. the number of sound and noise sources, increases linearly with the number of citizens. Consequently, adopting the ratio between such quantities as a control parameter is plausible.

Needless to say, an increasing number of citizens implies only a greater number of artificial sounds (due to transport, technologies, industries etc.), whereas the number of natural sounds remain fixed. However, the latter quantity is negligible with respect to the former one. Moreover, in a metropolis soundscape, also the relevance in terms of loudness of natural sounds with respect to artificial ones becomes negligible as well.

The aforementioned process leads to detect the main transition phases of the model (see Section 3.2). At this purpose, the parameter $T$ plays an important role: in fact, we recall that $T$ (the reciprocal of the $\beta$ parameter), represents the temperature of the system, to which we associate the communication level of the citizens' network.

Besides, let us define

$$\alpha = \alpha_1 + \ldots + \alpha_K, \quad (16)$$

which comes from Equation 13.

As a result:

- For high level of noise $T$ no retrieval is possible;

- For $T$ at an intermediate value of noise retrieval is possible, provided that the number of sound sources is not too large, i.e. $\alpha \leq 0.05$;

- For $T$ low, the Hopfield network can retrieve a number of patterns near its maximum, namely $\alpha = 0.14$.

## 6. CONCLUSIONS

As explained in Section 3.2, four phases can be observed in a Hopfield network. In our opinion, each phase can be associated with a different type of society.

The *ferromagnetic phase* corresponds to an *ordered society*, where citizens can recall the memorized patterns, namely the emotions related to the sounds heard during their everyday life. The reason is that sound sources are few and distinguishable: keynote sounds, sound signals, and soundmarks are clearly preserved. In the theoretical works by Schafer, this situation - typical of pre-industrial society - is defined as a *hi-fi soundscape*.

On the contrary, the *spin glass phase* corresponds to a *chaotic society*, where the community is no more able to distinguish sounds because of the dominant presence of noise. In Schafer's works, this situation is defined as a *lo-fi soundscape*, typical of post-industrial society. This kind of community loses its own reference points with respect to the surrounding soundscape because of masking, distraction, disorientation, etc. The social and cultural identity of the community within the specific environment is lost. In such a context, individuals tend to isolate themselves, even when they try to recover the mentioned reference points. In this sense, a clue is the habit of listening to music through headphones. The transition from a hi-fi to a lo-fi soundscape can be described in the Hopfield model through the transition from the ferromagnetic phase to the spin glass phase.

The model includes two other phases. The *mixed phase* is halfway between the ferromagnetic phase and the spin glass one. Even if Schafer's theories do not explicitly define intermediate states, from a social and sociological point of view this "nuance in meaning" can represent many real-world situations: e.g., the transformation of a pre-industrial into a post-industrial society, a process that took several decades during which natural and artificial sounds slowly lost their identity and became more and more indistinguishable. Another example is provided by low-technology societies and communities, where the transformation of a hi-fi soundscape into a lo-fi one is not complete.

Finally, we can identify the *paramagnetic phase* with the case in which community is so noisy that, regardless of the number of surrounding sound sources, the retrieval is not possible because of the confusion created by the citizens' network itself.

As shown in our work, these considerations about soundscapes can be modeled in terms of statistical mechanics. When the sound and noise sources increase in number, it is no longer possible to simultaneously minimize all the

Hamiltonian components. As a consequence, we get a frustrated system representing the inability to recall objects to mind as well as the inability to perceive distinct sounds having a specific meaning for the citizen.

Moreover, we stress that the network noise $T$ grows together with the increase of sound sources since the confusion of the surrounding soundscape leads people to communicate louder. As a consequence, today community is much more noisy than in the past. This fact contributes to produce even a more congested soundscape for a contemporary city.

We deduce from Equation 13 that, if hidden layers are disconnected (i.e. $K$ neighborhoods have their own soundscapes which do not interfere with the other ones), the corresponding patterns (i.e. the emotions stored by the citizens) contribute linearly to the capacity of the Hopfield network. The linear additivity property for the Hebbian kernel holds in the case of uncorrelated experiences (i.e. uncorrelated patterns). As a consequence, the sum of the control parameters defined in Equation 15 and referred to different neighborhoods implies the impossibility to distinguish among patterns related to neighborhoods. The control parameter is unique, as shown in Equation 16. In this homogeneous environment, citizens can still recall sounds and consequently feel emotions, but they are no more able to associate them to specific places: now the sound of the metropolis is indiscernible.

Currently, the nature of the present work is exquisitely theoretical and it lacks in experimentation. On one hand, it would be very difficult to recreate a pre-industrial soundscape in order to either verify or confute our theory. On the other hand, some tests can be conducted in populations where technology level is quite low, and future work will focus on this aspect.

## 7. REFERENCES

[1] J. D. Porteous and J. F. Mastin, "Soundscape," *Journal of Architectural and Planning Research*, 1985.

[2] B. De Coensel and D. Botteldooren, "The quiet rural soundscape and how to characterize it," *Acta Acustica united with Acustica*, vol. 92, no. 6, pp. 887–897, 2006.

[3] J. M. Berenguer, "Sounds in cause: Soundscape and evolution," *The Journal of the Acoustical Society of America*, vol. 130, no. 4, pp. 2531–2531, 2011.

[4] Y. Xiao-mei and H. Kun-peng, "Sound and video recording for chinese classical gardens soundscape database," *Journal of Temporal Design in Architecture and the Environment*, vol. 9, no. 1, pp. 104–107, 2009.

[5] S. Servigne, R. Laurini, M.-A. Kang, and K.-J. Li, "First specifications of an information system for urban soundscape," in *Multimedia Computing and Systems, 1999. IEEE International Conference on*, vol. 2. IEEE, 1999, pp. 262–266.

[6] M.-A. Kang and S. Servign, "Animated cartography for urban soundscape information," in *Proceedings of the 7th ACM international symposium on Advances in geographic information systems*. ACM, 1999, pp. 116–121.

[7] L. A. Ludovico and D. A. Mauro, "Sound and the city: Multi-layer representation and navigation of audio scenarios," in *Proceedings of the 6th Sound and Music Computing Conference. Oporto: SMC*, 2009.

[8] M. Raimbault and D. Dubois, "Urban soundscapes: Experiences and knowledge," *Cities*, vol. 22, no. 5, pp. 339–350, 2005.

[9] M. E. Nilsson and B. Berglund, "Soundscape quality in suburban green areas and city parks," *Acta Acustica united with Acustica*, vol. 92, no. 6, pp. 903–911, 2006.

[10] J. Kang and W. Yang, "Soundscape in urban open public spaces," *World Architecture*, vol. 144, pp. 76–79, 2002.

[11] B. Krause, S. H. Gage, and W. Joo, "Measuring and interpreting the temporal variability in the soundscape at four places in Sequoia National Park," *Landscape ecology*, vol. 26, no. 9, pp. 1247–1256, 2011.

[12] W. J. Davies, M. D. Adams, N. Bruce, M. Marselle, R. Cain, P. Jennings, J. Poxon, A. Carlyle, P. Cusack, D. A. Hall *et al.*, "The positive soundscape project: A synthesis of results from many disciplines," 2009.

[13] J. Carles, F. Bernáldez, and J. de Lucio, "Audio-visual interactions and soundscape preferences," *Landscape Research*, vol. 17, no. 2, pp. 52–56, 1992.

[14] R. M. Schafer, "The soundscape. our sonic environments and the tuning of the world," *Rochester, Vermont: Destiny Books*, 1994.

[15] B. Truax, *Acoustic communication*. Greenwood Publishing Group, 2001, vol. 1.

[16] R. M. Schafer, "The tuning of the world," 1977.

[17] K. H. Fischer and J. A. Hertz, *Spin glasses*. Cambridge University Press, 1993, vol. 1.

[18] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond*. World scientific Singapore, 1987, vol. 9.

[19] G. Toulouse, "Theory of the frustration effect in spin glasses: I," *Commun. Phys*, vol. 2, no. 4, pp. 115–119, 1977.

[20] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the national academy of sciences*, vol. 79, no. 8, pp. 2554–2558, 1982.

[21] J. M. Yeomans, *Statistical mechanics of phase transitions*. Oxford University Press, 1992.

[22] D. J. Amit, *Modeling brain function: The world of attractor neural networks*. Cambridge University Press, 1992.

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

[24] A. Barra, A. Bernacchia, E. Santucci, and P. Contucci, "On the equivalence of Hopfield networks and Boltzmann machines," *Neural Networks*, vol. 34, pp. 1–9, 2012.