A Model for Collective Free Improvisation

Clément Canonne, Nicolas Garnier

A Model for Collective Free Improvisation

Clément Canonne[1] and Nicolas Garnier[2]

[1] Université de Lyon, Ecole Normale Supérieure de Lyon
`clement.canonne@ens-lyon.fr`
[2] Université de Lyon, Laboratoire de Physique de l'ENS-Lyon, CNRS UMR 5672
`nicolas.garnier@ens-lyon.fr`

Abstract. This paper presents a model for Collective Free Improvisation (CFI), a form of improvisation that can be defined as *referent-free*. While very simple, it captures some interesting mechanisms of CFI. We use two variables: the *intention* and the *objective*. Both variables are used to describe the production and organization of the improvisers' signals. Using a system of Landau equations, we propose a non-linear dynamics for the intention evolving on a short time-scale while the objective evolves on a long time-scale. In this paper, the model is used to determine if, and within which conditions, a collective structure can emerge from CFI.

Keywords: Free Improvisation, Cognitive Model, Non-linear Dynamic Systems, Emergent Structure.

1 Introduction

Collective Free Improvisation (CFI) is a musical phenomenon produced by at least two persons improvising simultaneously and freely, *i.e.* trying to leave undecided every compositional aspects until the very moment of the performance.

When talking about "free" improvisation, one should carefully distinguish between two time scales. CFI is not deprived of all the automatized behaviors that can generate the improvised musical output on a short-term time scale: Embodied patterns and learned gestures are present as much as in other kinds of improvisation. In this regard, free improvisation is not to be confused with an illusory "pure" improvisation, which would account for instantaneous *ex nihilo* creation.

In return, CFI can be defined as a referent-free improvisation. According to Pressing [1], a referent is an underlying formal scheme or guiding image specific to a given piece, used by the improviser to facilitate the generation and editing of improvised behavior on an intermediate time scale. In CFI, as opposed to referent-based improvisation (like straightforward jazz), there is no founding act (like the common choice of a standard) that confers a given set of musical or extra-musical data the status of *common knowledge* in a group.

In a broader perspective, we are not either denying the importance for CFI of cultural backgrounds or musical knowledge, especially if they are shared among the group. CFI can include idiomatic borrowings: A given CFI can sound, at times, as a *be-bop* piece (with swing articulation, chords, tonal progression) or as a meditation on a *raga* (with a scale and a specific ornamentation) ; but a free improviser is someone that has no *pre-commitment* when the performance begins. His production is of course determined by several self-imposed restrictions, even stylistic restrictions, but he can modify these restrictions at any time.

In CFI, improvisers face two specific problems. First, the generation of improvised musical output on an intermediate time scale is not regulated. The formal unfolding is thus totally undetermined. Second, improvisers' musical coordination is not regulated and free improvisers' simultaneous production is much more difficult to control than in referent-based improvisation.[3] The fact that the way improvisers interact in CFI is not predetermined (roles and places in the ensemble can be redefined by anyone at anytime) makes it even harder.

We propose a model for CFI seen as this set of phenomenons.[4] An important inspiration for this model was the formalization of the improvisation's process proposed by Pressing [2] for a solo improviser. As CFI is a very interesting case of interaction, where shared information and pre-existing structures are almost nonexistent (each improviser can be described as "agnostic" before the interaction begins), it can be seen as paradigmatic. Besides the understanding of basic musical and cognitive processes in CFI, this model can be useful in both understanding social phenomena requiring effective coordination between agents (coordination problems) and in reinforcing the intuitive link between improvisational disposition and an agent's efficiency inside a complex system, following the steps of Borgo [3] that highlighted the numerous links that one can make between free improvisation's understanding and the study of complexity. This model can therefore be of interest in fields absolutely not related to music.

In this paper, the model is used to determine if, and within which conditions, a collective structure can emerge from CFI. This collective structure can be seen as a direct consequence of coordination's effectiveness or, to put it in another way, of group flow, as presented by Sawyer [4] and defined by some, if not all, of the following features: heightened consciousness, clear vision of a common goal, close listening, confidence in each other competency, experience of time dilatation, intuitive understanding of other musicians' intentions, sense of balance, high-enough risky situations to excite each musician's virtuosity and creativity... In the following, we consider that if a collective structure emerge, it is because group flow has happened: as the two notions are strongly correlated, we only focus in this paper on the question of CFI's collective structure.

[3] For example, if the referent is a chord progression, everyone knows at each moment what pitches he can or can't play.

[4] A lot of the music called free improvisation by the musicians, the audience or the record labels does not fall strictly into our category of CFI because it has some minimal form of referent, or because it takes place into an established group, with its own conventions and so on. Maybe Derek Bailey's *Company Week* is the most famous example of CFI in "real life".

2 Definitions and Model

2.1 Time Scales

The improvised production extends over several time scales, that we decompose as:

- The shortest time scale, which is the scale of the musical or acoustical signal, depending on the point of view from which the signal is described: either in a musical way (encoding the signal in "notes" of different durations on a score) or in an acoustical way (following the very evolution of the same timbre through time)... This scale is not explicitly used in our model.
- Short time scale τ_s, of the order of seconds: it is the scale of the "clusters of events" [2].
- Long time scale τ_l, of the order of minutes: it is the scale of the "sequences" [2].
- The scale of the complete improvisation piece. For a real-life improvisation, the duration is not *a priori* established, but in our model, it will be.

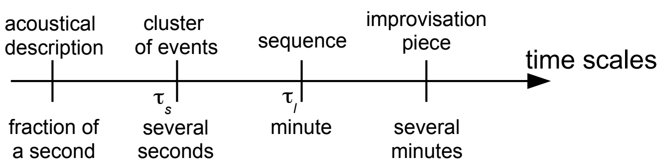

Fig. 1. Separation of the different time scales in the model of CFI.

We detail now the two notions of *clusters of events* and *sequences* which define relevant time scales for our model.

The cluster of events. A cluster is a cognitive chunk that gathers a set of musical, acoustical, cinetical events that were decided at the same point in time; it can be pictured as the subsequent execution of a micro-plan, over its duration of the order τ_s. Pressing [1] noted that the generation of improvised behavior on a short time scale is primarily determined by previous training and embodied patterns, and is not very piece-specific. τ_s is a short time, of the order of seconds. It is short because of the agent's cognitive limitations (one can't decide too much at the same time) and the improvisation's interactive dimension (one does not want to decide too much at the same time, to be able to react quickly to changes in the environment).

The sequence. The sequence is related to the long time scale τ_l. A sequence is defined by a set of processes and/or a number of features (acoustical, cinetical,musical...) holding for a given length. Improvisers try to establish successive identities and stable points in musical's stream. These identities are then developed, played with or eventually negated, until new identities finally emerge. This organization in successive sequences is probably an endogenous feature of CFI [2, 5, 6].

2.2 Signal and Information

What we call "signal" in this paper is not the real musical signal produced by the musician. Its realistic description would require a huge number of variables. On the contrary, our description of the signal is extremely simplified and doesn't contain anything about the acoustic representation. In fact, we consider a real number x, related to the relative complexity of the signal (a cluster on the piano is more complex than a triad; a multiphonic on the clarinet is more complex than a traditionnaly-produced sound; a sub-division in septuplet is more complex than a sub-division in sextuplet; a stretch of music with very quick changes of pitches is more complex than a stretch of music with only one pitch...). As a direct consequence, our model's focus will not be on the signal *per se* but rather on more high-level phenomenas (interaction, long-term intention...). Extension of our model to more realistic signals would imply the consideration of a vector x with as many components as needed (to describe, e.g., pitch, intensity, timbre, duration).

We write N the number of musicians and $x^k(t)$ the temporal signal of musician k ($1 \leq k \leq N$). We write $\boldsymbol{x}(t) = (x^1(t), ..., x^N(t))$ the set of signals produced at a given time t.

The quantity of information is directly related to the signal. We want that:

- the larger the signal, the larger its information. So we define a static information by analogy with energy in physics $I_s^k = \frac{1}{2}(x^k)^2$,
- the larger the signal varies, the larger its information. Correspondingly we define a dynamical information by analogy with kinetic energy $I_d^k = \frac{1}{2}\tau_i^2 \left(\frac{\mathrm{d}x^k}{\mathrm{d}t}\right)^2$, where τ_i is a normalization time.

Then $I^k = I_s^k + I_d^k$ is the information delivered by player k, whereas the total information seen by any musician of the group is:

$$I = \sum_k I^k = \frac{1}{2}\|\boldsymbol{x}\|^2 + \frac{1}{2}\tau_i^2 \left\|\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}\right\|^2$$

2.3 Signal and Intention

Besides the signal x^k produced by musician k, we define the intention ω^k of this musician, which represents the ideal signal that the musician would like to deliver. The intention is *a priori* more complex than the signal produced, because it contains information that the musician may not be able to actually play, due to, e.g., lack of technicality, lack of time, *etc.* The signal x^k is deduced from the intention ω^k at given discrete time steps by projection, expressing the possible loss of information between the intention and its actualization in the signal:

$$x^k = g(\omega^k)$$

The function g expresses the projection, and for the sake of simplicity in the present paper, we suppose that the musician is "perfect", so that g is identity, *i.e.* $x^k = \omega^k$, when the projection occurs.

The intention ω^k evolves on the short time scale τ_s and we choose a continuous dynamics for it. On the contrary, we impose that x^k is constant and equals x^k_n during a cluster of events, *i.e.* between two projections of the intention separated by the time lag d^k_n which is the duration of the cluster of events. The index n labels the time. Durations d^k_n of clusters are of the order of the short time scale τ_s. To make our model deterministic, we impose

$$d^k_n = \tau_s - a(x^k_n)^2$$

where $a > 0$ is a constant. Clusters of events are shorter when the signal is large, *i.e.* contains more information. Conversely, if the signal is poor in information, the corresponding cluster of events is longer. We choose $a = 0.3$ in the remaining of the paper.

Because the signal x^k is piecewise constant, the information I^k is constant on clusters of events, while there are peaks of dynamic information at the boundaries between clusters (see Fig. 2). For coherence, we choose the normalization time $\tau_i = 2\mathrm{d}t$ where $\mathrm{d}t$ is the time step we use in numerical integration; this way, I^k_d depends on the signal's amplitude variation when there is a change of cluster of events rather than on the time derivative of this amplitude which is infinite.

Intention's dynamics. We propose the following dynamics for ω^k, inspired from dynamical systems' theory:

$$\tau_c \frac{\mathrm{d}\omega^k}{\mathrm{d}t} = \alpha^k x^k + \sum_{l \neq k} \beta^{k,l} x^l - g\|\omega_k\|^2 \omega^k \tag{1}$$

Each of these equations is a Landau equation from phase transition theory. Parameter $g > 0$ is a constant in front of a non-linear that prevents the solutions from diverging; we choose $g = 1$. Note that the evolution of intention ω^k depends on signals $\{x^l\}$, $l \neq k$, from other musicians, not from their intentions which are of course not known by musician k. Solutions of Eq. (1) vary on time scale τ_s and its autocorrelation function decreases to zero over this time scale. Parameter $\alpha^k > 0$ expresses the self-sensitivity of musician k. Parameters $\beta^{k,l}$ express the influence over musician k of signals $\left\{x^l\right\}_{l \neq k}$ from other musicians and therefore quantify the interactions between musicians. $\beta^{k,l}$ are of order 1, they can have any sign or vanish.

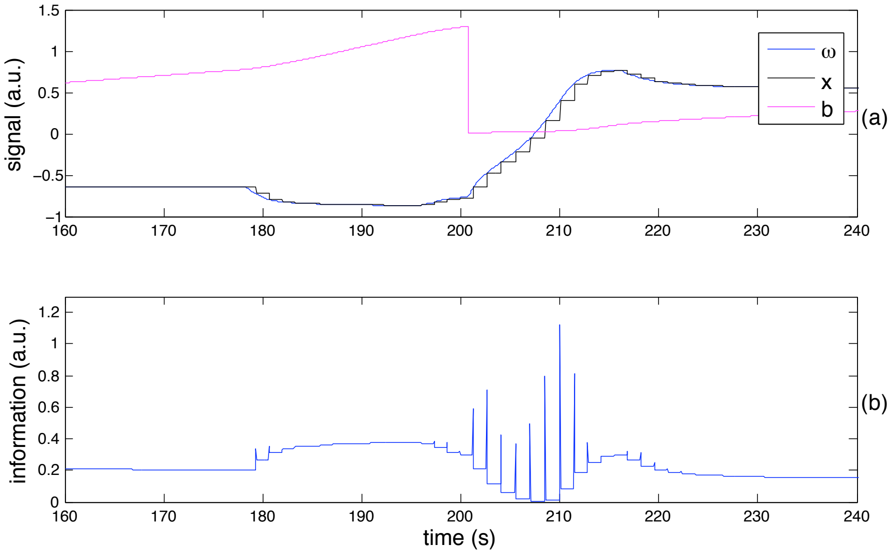

Fig. 2. Example of intention, signal, boreness and information for one musician. (a) the intention ω evolves continuously in time, while the signal x is discrete on short time periods coresponding to cluster of events. Boreness b is reset to zero when a sequence change occurs (here, around time 200 s). (b) Information I^k is constant over a cluster of events, and presents a peak when a new cluster of events begins.

2.4 Objective

We define the objective $\boldsymbol{\Omega}^k$ of the musician k as the set of parameters α^k and $\{\beta^{k,l}\}_{l\neq k}$ that caracterizes the linear part of the equations: $\boldsymbol{\Omega}^k = (\beta^{k,l})$, where we have written $\beta^{k,k} = \alpha^k$ for $l = k$. The objective of musician k is a N-dimensional vector that plays the role of the control parameter in the Landau theory. The set of objectives of all players defines a $N \times N$ matrix. The objective defines the value towards which the intention tends on a time scale larger than τ_s. This is obvious for a solo improvisation ($N = 1$ and all $\beta^{k,l}$ are zero): We then have a Landau equation and its solution ω tends to $\sqrt{\alpha/g}$. But the objective is also a measure of the interaction with the other musicians; in a collective improvisation, any improviser interacts with the others, and the intention tends towards a value given by an appropriate combination of components of the objective. Following Pelz-Sherman [7], we distinguish some paradigmatic cases and define the corresponding coupling of musician k with musician l:

- If $\beta^{k,l} \simeq 1$, then ω^k tends to x^l. Player k is willing to imitate player l: “imitation”.
- If $\beta^{k,l} \simeq -1$, then ω^k tends to -x^l. Player k is willing to have a signal opposite from signal of player l: “contrast”.

– If $\beta^{k,l} \simeq 0$, then player k is not paying any attention to the signal from player l: "independency".

These are limit cases — obvious for $N = 2$ and $\alpha^k = 0$ — and of course any intermediate situation can occur. Note that the couplings are non-symmetrical: $\beta^{k,l} \neq \beta^{l,k}$, *i.e.*, musician k can for example "imitate" musician l while l is "independent" from k.

Contrary to classical Landau theory, we make the objective evolve in time with a specific dynamics, on the long time scale τ_l. We choose a discrete dynamics, and any change in the objective of a musician defines a new sequence for this musician. This dynamics requires the introduction of the *cognitive load* and the *boreness*.

Cognitive load. In CFI, a musician's attention is shared between two tasks: generating his own signal and monitoring other musicians' signals. We introduce the cognitive load to account for finiteness of a musician's attention. Cognitive load bounds possible values of the objective components. We write a first part of the cognitive load as the part devoted to monitoring the signals:

$$C^k_{\text{monitor}} = \frac{1}{2}\left(\alpha^k x^k\right)^2 + \sum_{l \neq k} \frac{1}{2}\left(\beta^{k,l} x^l\right)^2 = \frac{1}{2}\left\|\boldsymbol{\Omega}^k.\boldsymbol{x}\right\|^2$$

Another component of the cognitive load is related to the production of the signal. We suppose that difficulty for musician k to produce a signal is proportional, with a proportionality coefficient $(a^k)^2$, to the quantity of static information in his signal:

$$C^k_{\text{prod}} = \frac{1}{2}(a^k)^2 I^k_s = \frac{1}{2}(a^k x^k)^2$$

The total cognitive load of the musician k will be noted $C^k = C^k_{\text{monitor}} + C^k_{\text{prod}}$, and we require that this variable be bounded from above by a constant $C^k_{\max}$ representing the maximal cognitive capacity of the musician k.

Boreness. When a sequence is lasting too long, the musician gets bored and ultimately breaks it. We define the boreness $b^k(t)$ of the musician k to quantify this effect. Boreness grows in time, until it reaches a limit $b^k_{\max}$; then a change of objective occurs (see Fig. 2), which is also a change of sequence. We simply choose:

$$\frac{\mathrm{d}b^k}{\mathrm{d}t} = C^k$$

with initial condition $b^k = 0$ at $t = 0$. When the objective is changed, at the end of a sequence, boreness is reset to 0. Maximal boreness $b^k_{\max}$ is related to maximal cognitive charge; we choose:

$$b^k_{\max} = \frac{\tau_l}{3\tau_c} C^k_{\max}$$

Objective's dynamic. We choose that the objective remains constant as long as $b^k(t) < b^k_{\max}$. On the contrary, when the boreness b^k becomes larger than the maximal value $b^k_{\max}$, we choose a new objective such that the cognitive charge remains bounded from above. All possible choices of α^k and $\beta^{k,l}$ are therefore not possible. For the sake of simplicity, the new components of the objective are choosen randomly, between -1 and $+1$ for $\beta^{k,l}$, and between 0 and 1 for α^k; If the new cognitive charge C^k resulting from these new values is larger than $C^k_{\max}$, we apply the factor $C^k_{\max}/C^k$ to all components of $\boldsymbol{\Omega}^k$. We also decide to project ω^k into x^k at the very same time such a change of sequence occurs.

3 Results and Discussion

3.1 Collective sequences and their articulation

To analyse the complete production of the group, we define collective sequences, not to be confused with individual sequences simply defined above by a given objective. This is a way to probe and quantify coordination efficiency in the group. We call collective sequence a time frame during which each improviser maintain a relative musical identity (*i.e.* his intention stays more or less constant). If we find a lot of collective sequences, and if collective sequences are long enough, we will say that coordination amongst musicians in the group is good. One of the main interest of this model is to show the existence of collective sequences.

If the position of each improviser in the group stays more or less constant, *i.e.* all objectives are contant, then we expect a collective sequence. Nevertheless, a constant objective is not a sufficient condition for the occurence of a collective sequence. Also, a collective sequence can be composed of a series of individual sequences; if after a change of individual sequence, each musican's position into the group is not strongly altered, then the same collective sequence will continue.

This collective structuring in successive sequences is one of CFI's great challenges. Fig. 3 shows that it is not always possible to detect collective sequences in CFI. One can clearly see two types of local structure in our model of CFI:

- A stable solution which can be seen as a "collective sequence" (labelled 1,2,3 in Fig. 3); this corresponds to a fixed point in the phase space of the system.
- An oscillating solution which can be seen as a phase of discoordination among the musicians (labelled B in Fig. 3); this corresponds to a limit cycle, and it is obtained when some of the eigenvalues of the matrix of objectives are not real numbers but complex conjugates.

3.2 Contributing factors to CFI's structuring in collective sequences

The model is now used to see within which conditions the emergence of collective sequences is facilitated.

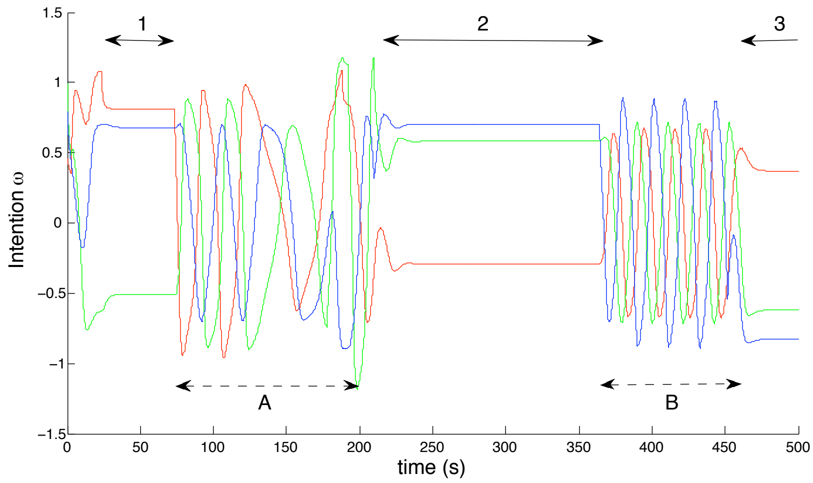

Fig. 3. Signals from a team of 3 musicians, all with $a^k = 0$, as produced by our model. We can discriminate 3 collective sequences (labelled 1, 2, 3). At the begining, a short transient period is observed before collective sequence 1. Between collective sequences 1 and 2, resp. 2 and 3, we observe a chaotic behavior A, resp. periodic behavior B.

Effects of improvisers' features. We consider here two specific features:

- *Virtuosi* produce high-information signals at a lower cognitive cost. This is represented by a low value of a^k.
- *Leaders* have a superior cognitive capacity. As a direct consequence, they tend to get bored more slowly (*i.e.* from a musical perspective, they try to "work out" the different ideas and situations).

Fig. 4, where $a^k = 0.4$ for all 3 musicians, has to be contrasted with Fig. 3, where we selected $a^k = 0$ for all 3 musicians (highly virtuoso improvisers). It clearly shows the impact of this feature on improvisers' coordination.

Fig. 5 shows that the existence of leaders enhance the organization of CFI in collective sequences.

Number of improvisers. As one could have expected, the fewer the musicians, the easier the collective organization, as shown on Fig. 6 where we increase the number of musicians to 5, and see the difficulty to define clear collective sequences.

Emergence of sub-teams. To adress the issue of obtaining collective sequences in large groups of musicians, we allow our improvisers to seek for the creation of sub-teams in CFI. This can be done in two different ways:

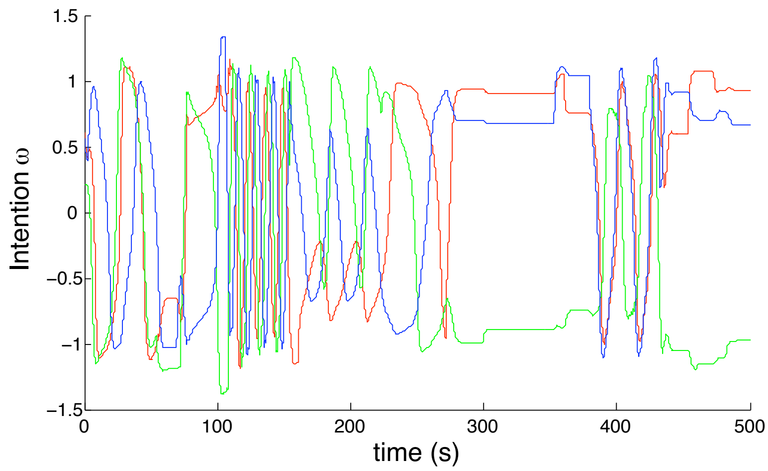

Fig. 4. Signals from a team of 3 musicians, all with $a^k = 0.4$. Although collective sequences still exist, they occur less often.

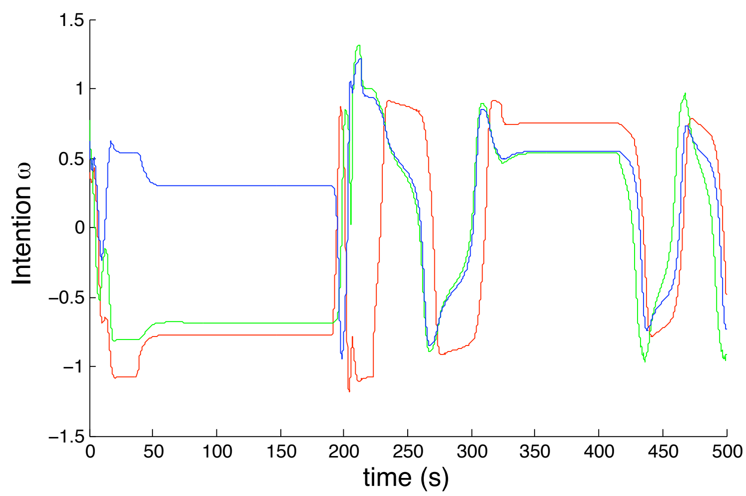

Fig. 5. Signals from a team of 3 virtuoso musicians, ($a^k = 0$). One of the musicians is a leader, with $C^k_{\max}$ twice larger than for regular improvisers.

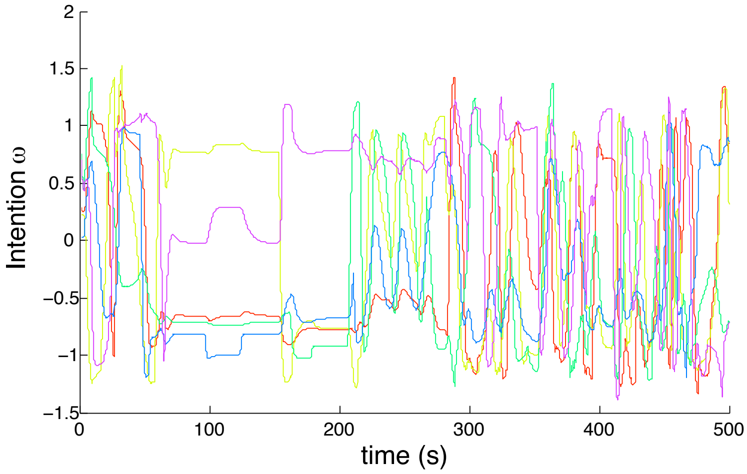

Fig. 6. Signals from a team of 5 virtuoso musicians ($a^k = 0$). Collective sequences are difficult to produce and/or discriminate.

- Improvisers are looking for symmetrical interactions: if A is in imitation with B, B will also try to imitate A, thus unifying a sub-team by introducing *reciprocal actions*: $\beta^{k,l}$ has the same sign as $\beta^{l,k}$ and is of the same order of magnitude. Results are illustrated in Fig. 7.
- In large groups, improvisers do not interact with every other musicians. On the contrary, they focus on one or two specific musicians, so that they interact only with them. This is obtained by imposing a maximum of two non vanishing $\beta^{k,l}$ for every musician k. A typical case is depicted in Fig. 8.

As shown in Fig.7 and Fig. 8, this "sub-team" kind of reasoning is efficient to organize CFI in collective sequences.

3.3 Future plans

The model's first results presented above are encouraging. In particular, the model is successful in showing the possibility of self-organization in CFI, despite the absence of *a priori* structures. But this self-organization depends on several features: the musicians' virtuosity, their leadership quality, their team and sub-team reasoning... and probably other features still to discover. Next step is to quantify the effects of improvisers and sub-teams' features on CFI and its organization; for this purpose, a statistical approach will be used. Some assumptions we have made can be relaxed. For example, we have tried a different projection of intention into signal, by adding some noise, thus making the musician imperfect: This does not affect the observed behavior of the model. Future plans also

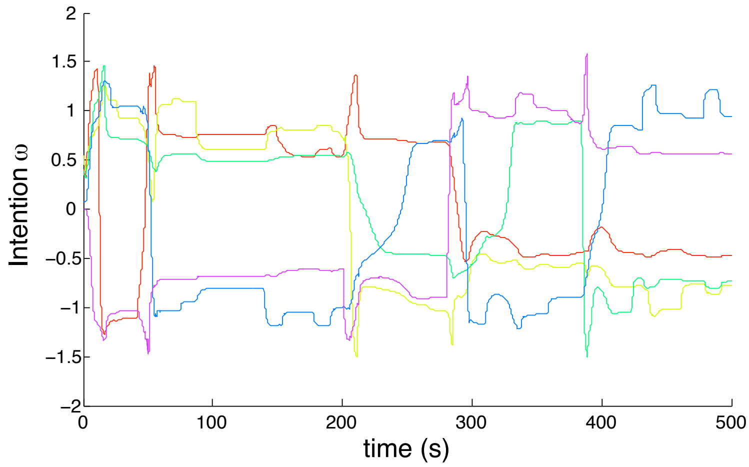

Fig. 7. Signals from a team of 5 musicians ($a^k = 0.4$). Here, musicians tend to have symetrical interactions.

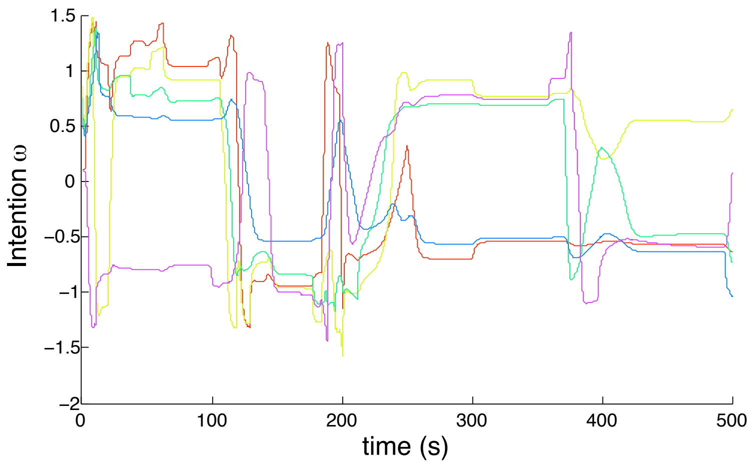

Fig. 8. Signals from a team of 5 virtuoso musicians ($a^k = 0$) with one leader. Here, all improvisers are interacting with at most 2 other musicians.

include the study of more realistic dynamics for the objective, and possible inclusion of a long-term memory: in this regard, Dubnov's works [8] could be useful. This should provide a finer understanding of the way a collective structure can emerge from CFI. Confrontations of our model with laboratory human production is also under consideration, which may suggest in return some non-trivial modifications of our signal's conception.

Acknowledgements

This work has been supported by the Fonds Recherche of ENS-Lyon.

References

1. Pressing, J.: Cognitive Processes in Improvisation. In: Crozier, W. R., Chapman, A. (eds.) Cognitive Processes in the Perception of Art, pp. 345-363. Elsevier, Amsterdam (1984)
2. Pressing, J.: Improvisation: Methods and Models. In: Sloboda, J. (ed.) Generative Processes in Music, pp. 129-178. Clarendon, Oxford (1988)
3. Borgo, D.: Sync or Swarm: Improvising Music in a Complex Age. Continuum, New York (2005)
4. Sawyer, R. K.: Group Creativity: Music, Theater, Collaboration. Routledge, London (2003)
5. Nunn, T.: Wisdom of the Impulse: On the Nature of Musical Improvisation, http://www20.brinkster.com/improarchive/tn.htm(1998)
6. Canonne, C.: L'improvisation Collective Libre: De l'Exigence de Coordination âĂą la Recherche de Points Focaux. Thèse de Doctorat en Musicologie de l'UniversitÃĹ de Saint-Etienne (2010)
7. Pelz-Sherman, M.: A Framework for the Analysis of Performer Interactions in Western Improvised Music, http://pelz-sherman.net/mpsdiss.pdf (1998)
8. Dubnov, S.: Unified View of Prediction and Repetition Structure in Audio Signals with Application to Interest Point Detection. IEEE Transactions on Audio, Speech and Language Processing. 16 (2), 327–337 (2008)