Audio Engineering Society

Convention Paper

Presented at the 129th Convention
2010 November 4–7 San Francisco, CA, USA

Simulating Ensemble Rhythmic Interaction based on Quantifiable Strategy Functions

Nima Darabi [1], U. Peter Svensson [1], Chris Chafe[2]

[1] Centre for Quantifiable Quality of Service in Communication Systems*, NTNU,
Trondheim, O.S. Bragstads plass 2E, NO-7491, Norway
darabi@q2s.ntnu.no
svensson@q2s.ntnu.no

[2] CCRMA, Center for Computer Research in Music and Acoustics
660 Lomita Dr. Stanford, CA 94305
cc@ccrma.stanford.edu

ABSTRACT

This paper studies the strategy taken by a pair of ensemble performers under the influence of delay. A general quantifiable measure of strategy taken by performers in an interactive rhythmic performance is represented in a form of a single-parameter strategy function. This is done by imposing an assumption about a decision-making process for "onset generation" by a participant, with one degree of freedom, to the observed data. We present specific examples of such strategy functions, suitable for different scenarios of rhythmic collaboration. By perpendicular projection of strategy functions of an ensemble performing trail onto Cartesian axis a nominal trial was transformed to a "strategy path" to show how the performers change their strategies during the course of a trial. By mathematical induction it was proven that this transformation from the time domain to a "strategy domain" is conditionally reversible, i.e. time vectors of an ensemble trial can be reconstructed by a domino effect having its time-free strategy path and given an initial state. This algorithm is considered to be a means to simulate the ensemble trials based on the overall strategies leading them.

---

* *Centre for Quantifiable Quality of Service in Communication Systems, Centre of Excellence* appointed by The Research Council of Norway, funded by the Research Council, NTNU, and UNINETT. http://www.ntnu.no/Q2S/

1. INTRODUCING THE IDEA OF TAPPING STRATEGIES

A number of research groups have studied how latency over a communication channel psychophysically influences the two interacting performers [1][2]. There have also been numerous experimental studies on Sensorimotor Synchronization (SMS), i.e. the rhythmic coordination of perception and action, mainly in the form of finger tapping or hand-clapping to a sequence of auditory stimuli (an extensive summary could be find in [3]). Focusing on the objectively measurable properties of the experimental data, the subjective strategies taken by participants involved in ensemble playing or even during tapping to a given pattern of stimuli have still been offered little attention, though.

Studying delayed ensemble performing, remarkable differences between individual rhythm tapping pairs are reported [1][2], i.e. statistical second moments in the data are evident. The dominant approach has been to average the data and study only the properties that are common across clapping trials, discarding information about individual differences. In this paper, we aim at studying tapping strategies that are taken by individuals in a rhythmically collaborating pair. This would reveal information about how trials are different, rather than similar. These differences are of valuable importance for finding mathematical models behind rhythmic aspects of musical collaboration and synchronization.

In continuation of a previous attempt to quantify the strategy taken by a pair of duo-clappers [4], we need to define and quantify possible strategy functions that could describe the individual behavior of human subjects while rhythmically interacting over a delayed channel. This should be done with respect to the experimental setup and given a specific task. We also need to explore ways to show how these mutual strategies, taken by two collaborating partners, interact with each other during the course of a trial.

We thus need to develop an abstract and mathematically precise model, general enough to analyze different scenarios of mutual rhythmic collaboration and applicable to a range of different sounds and acoustical settings. This method should also be applicable to remote jamming sessions just if reduced to two sequences of correspondent clicks (or tones) for two performers.

2. METHODOLOGY

We will develop a mathematical process by which a generally defined single-parameter time-varying strategy function is extracted from a recorded session of rhythmic collaboration, e.g. remote hand-clapping over a network. The applications of the developed method are discussed as well as its consistency; however applying it to real datasets is considered beyond the purpose of the paper, which is for now to develop the theoretical approach. The present study, thus, contains no concrete conclusion about real experiments.

We have here focused on strategies taken by participants in musical interaction. Our approach is to suggest a numerical parametric model with different “strategy assumptions”, and to fit its parameters to observed experimental data. We aim at creating a time-dependent model that can describe how a person’s behavior changes during the extent of a rhythmic session depending on the defined task and the experimental setup.

Since the participants of the session are given a task that will most probably influence the way they perform, by imposing a strategy assumption, dependent on the given task and the experimental setup, to the participants observed behavior (i.e. recorded signals) we will derive a one-dimensional subjective signal called strategy function for each participating person.

We have previously [4] discussed that such a methodology permits a single model parameter only per person: this methodology leads to the solving of a system of linear equations. Each detected event’s timestamp for each person will provide the system with a new equation and one new equation is informative enough to calculate exactly one more unknown variable, which is a single strategy parameter of the tapping participant assigned to that clapping instant.

A discrete definition of such a strategy function, namely the “compensation factor”, has been presented previously by the authors [4]. This parameter was quantified as a discrete function of time based on fitting its definition to the observed timestamps of an ensemble delayed mutual hand-clapping experiment between a pair of participants. In this research we extend this discrete definition of a strategy function that suited that experiment, into a continuous and more general variation. The general definition will then give us other alternatives to define a strategy function.

3. TIME-VARYING STRATEGY FUNCTIONS

Assume that a general rhythmic ensemble performing session, under the influence of a delay, is given numerically by the timestamps of each side's detected events. For example the onset timestamps clapped by persons A and B in an ensemble hand-clapping experiment could be given by vectors $\mathbf{t}_A$ and $\mathbf{t}_B$. These detected events for persons A and B correspond to each other in a one–to-one relationship, and each of them is assumed to represent one musical beat. Notations τ_A and τ_B stand respectively for the channel delays from side A to B and B to A and they are assumed to be constant during a trial.

Given these two data vectors to start with ($\mathbf{t}_A$ and $\mathbf{t}_B$), the block diagram in Fig 1 illustrates the overall procedure of transforming the data of the recorded session into strategy functions.

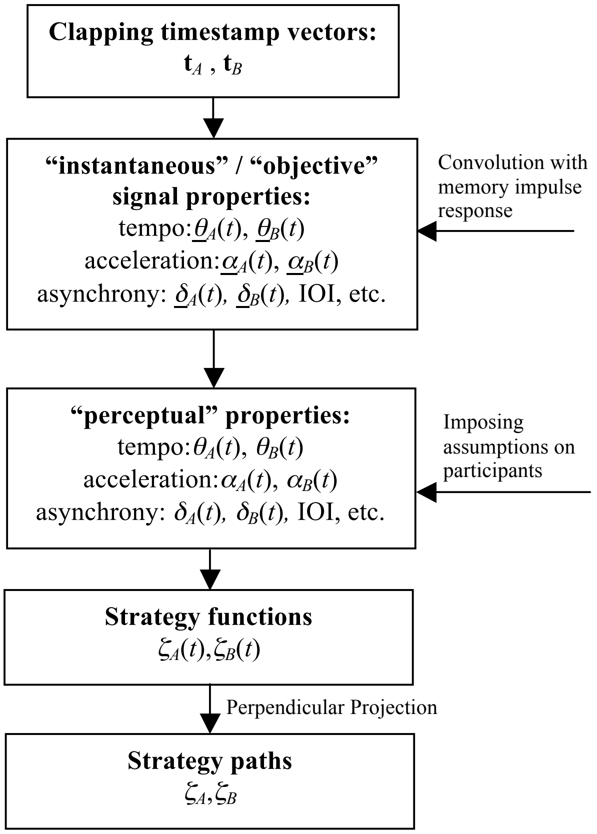


Figure 1 The overall procedure of transforming an objective recorded interactive rhythmic session to its participants' subjective strategies

3.1. Calculating tempo, IOI and asynchrony

First we describe how different objective properties of the signal will be calculated as time-varying functions. Since the event vectors correspond one by one to each other, we define the "insantanous" *asynchrony* for person A as the time that he/she has generated an event in advance to the other side's corresponding detected event, in milliseconds. Asynchrony vectors are thus given by $\underline{\boldsymbol{\delta}}_A=\mathbf{t}_A-\mathbf{t}_B$ $\underline{\boldsymbol{\delta}}_B=\mathbf{t}_B-\mathbf{t}_A$. The instantaneous *inter-onset interval* (IOI) for each side is the time between two sequential onsets/beats/events. Instantaneous *tempo* for each person at the time of a detected event is another property of interest given by 60/IOI beats per minute, as each element of the vectors $\mathbf{t}_A$ and $\mathbf{t}_B$ represents a beat. The instantaneous tempo is a time-varying measure of each person's action pace. Its time resolution is low, though, as the tempo is irregularly sampled by a varying sampling frequency of number of events per second. In order to get an instant tempo signal for each side with a higher temporal resolution, we need up-sampling.

Up-sampling: To get a continuous model that describes the parameters as a function of their occurring time rather than their corresponding detected event, we up-sample the instant tempo curve and the instant IOI curve with a sampling frequency of F_s. Using a linear interpolation we set values for the in-between samples that do not correspond to any tapped timestamp and that gives us the mentioned property of signal (for person A or B) as a regularly sampled function of time. We denote the instantaneous tempo generated by person A at time t by $\underline{\theta}_A(t)$ and the instantaneous asynchrony by $\underline{\delta}(t)$, which are also calculated by up-sampling the tempo and asynchrony vectors by linear interpolation. Up-sampled versions of the instant tempo curve etc are still discrete, rather than continuous, functions but for our purposes here, we can refer to the up-sampled ones as continuous functions.

Convolving with exponentially decaying memory curve: The calculated curves for signals' "objective" properties usually change rapidly in a short window of time, due to the human jitter or other sources of error. This does not seem to be in accordance to their perceptual values. To map an objectively measured quantity like tempo or asynchrony to the perceptive domain we use an adjustable smoothing function, acting as a first order filter. Considering that auditory short-term memory is decaying exponentially towards the past [5], we define a memory impulse response $g(t)$ adjustable by a parameter (μ). We then convolve the extracted objective properties with the memory IR to get the smooth filtered version of the property that we call here the "memorized" property. Even though this latter property of the signal is calculated based on the objective data, We call "perceptual", as it represents the perceptive match of the measured quantity, and depends on μ that might in practice be different from person to person:

$$g(t) = \ln\frac{1}{\mu}\begin{cases}\mu^{-t}, & t \geq 0 \\ 0, & t < 0\end{cases} \quad (1.)$$

$$\theta(t)=\underline{\theta}(t)*g(t) \quad (2.)$$

$$\delta(t)=\underline{\delta}(t)*g(t) \quad (3.)$$

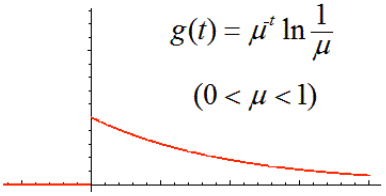
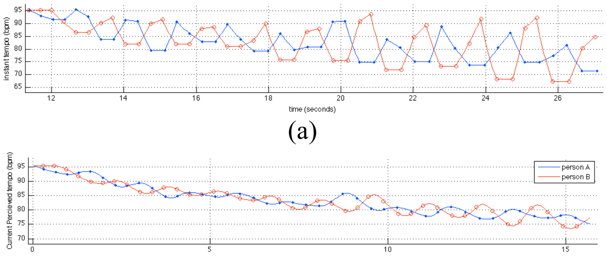


Fig. 2, the exponential IR used using the factor μ as its memory factor decays exponentially by time. Here the positive part of the time axis shows the moments in the past for each of which IR returns how much of that moment is remembered now.

In this formulation $g(t)$ is thought to represent how much a person "remembers" of the moment t, at a "present" moment of zero. The parameter μ for which $0<\mu<1$ is assumed to be a built-in property and then will be kept constant for each participating person. It basically represents what proportion of something from "one second ago" that a person remembers "now". We call it a memory factor as for its 0 value convolving the instantaneous curve with the Dirac delta function returns the same input (no memory scenario). On the other hand, when it is set to 1, all the past history of the signal no matter how distant contributes equally to the memorized property (full memory scenario). In Fig 3 objective tempo and perceived tempo for an example trial of hand-clapping is illustrated.

The perceived tempi $\theta_A(t)$ and $\theta_B(t)$, their time derivative (perceived acceleration) that is denoted by $\alpha_A(t)=d/dt\theta_A(t)$ and $\alpha_B(t)$, and the perceived asynchronies $\delta_A(t)$ and $\delta_B(t)$ will further on be processed to get the strategy functions.



(a)

(b)

Figure 3 "Perceived" tempo (b), of a recorded tapping trial is a smooth version of its "objective", instantaneous tempo (a). The figure illustrates this for an example trial of a pair of duo-clapping subjects A (•) and B (○) with the lead-in tempo of 94 bpm under the delay of 68 ms. The participants were given complementary rhythms (1011 and 1110) and their task was to keep the tempo stable as much as possible. Trial is taken from the experiment done by Farner et. al. described in [1]

3.2. Extracting strategy functions

The idea of extracting strategy functions of a person at any moment of a specific trial works based on comparing an occurred event's time (e.g. timestamp of an observed clap's onset) with the moments in which the subject could have generated (e.g. could have clapped) that event by taking two different extreme rhythmic strategies. In other words, a recorded event marks the tradeoff point within a potential span of possible timestamps that this specific event could have occurred in, bounded by two extreme cases. If the event occurs at the lower boundary of the *strategy span*, the according strategy of the person taken at that time is evaluated as 0 and if that happens at the upper bound, this strategy is quantified as 1.

The strategy span shall be defined so that it could cover the observed detected events, meaning that if an event occurs any time between the two end points, a value between 0 and 1 is linearly assigned to the strategy function of the event-generating person, at that moment of time. As such a mapping rescales the time axis in a way; such a parameter could be negative or may take values greater than 1, for possible events locating outside of the span.

By imposing different assumptions to the on-the-fly perceptive/cognitive decision-making process of a participant during interaction, one can define a strategy span. Focused on each person, the relevant input parameters of such a decision making process can be divided into known and unknown parameters: Known parameters (instantaneous or perceptual) are objectively measurable parameters of the recorded trial. IOI, tempo, acceleration, asynchrony, and even asynchrony change are thought to be relevant. Unknown parameters on the other hand deal with the subject's state of mind and are those subjective properties that influence the process of decision-making.

Presuming different procedures by which the collaborator decides when to generate the next event, such a span can be defined alternatively in different ways. Due to the limited information in the observed data, this definition can't include more than one parameter, or degree-of-freedom, and our approach can't map all unknown subjective properties of the decision-making process to more than that single parameter, i.e. the quantified *strategy function*. Below we will show that actually one is informative enough to regenerate the trial.

3.3. General definition of strategy span

To formulate what is explained above by our notations we discuss how the strategy function for person A at his i^{th} generated event is calculated in our approach: Based on observation event occurs at time t_A^i while with two extreme strategies that person A could take, it could have occurred most probably any time between the lower and upper boundaries of the strategy span (respectively $\overleftarrow{t}_A^i$ and $\overrightarrow{t}_A^i$). f_{lower} and f_{upper} are functions defined in order to give us these two boundaries. Below we will give specific values to them. But, for the most general case they are functions of all the "previously" observed properties of the recorded signals, which are those corresponding to the taps occurring before the i^{th} onset. These arguments of the f functions include tap timestamps of both sides, $(t_A^1, t_A^2, ..., t_A^{i-1}$ and $t_B^1, t_B^2, ..., t_B^{i-1})$ which are denoted shortly by $t_{A,B}^{1,2,...,i-1}$. They also include their calculated tempo/IOI, acceleration, asynchrony, etc. These functions are described in equations 4 and 5. Note that for example $\theta(t_A^i)$, that is person A's perceived tempo at his i^{th} onset, is denoted by θ_A^i and so on:

$$\overleftarrow{t}_A^i = f_{lower}(t_{A,B}^{1,2,...,i-1}, \theta_{A,B}^{1,2,...,i-1}, \alpha_{A,B}^{1,2,...,i-1}, \delta_{A,B}^{1,2,...,i-1}, ...) \quad (4.)$$

$$\overrightarrow{t}_A^i = f_{upper}(t_{A,B}^{1,2,...,i-1}, \theta_{A,B}^{1,2,...,i-1}, \alpha_{A,B}^{1,2,...,i-1}, \delta_{A,B}^{1,2,...,i-1}, ...) \quad (5.)$$

$$\zeta_A^i = \frac{t_A^i - \overleftarrow{t}_A^i}{\overrightarrow{t}_A^i - \overleftarrow{t}_A^i} \qquad (6.)$$

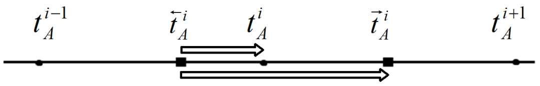


Figure 4 A strategy span is bounded by two timestamps representing two extreme strategies taken by the performer around each detected event. The observed event's time distance to the lower boundary compared to the total span length gives the strategy function at the time of the event.

Equation 6 in accordance to Fig 4 shows how the $\zeta_A(t_A^i)$, or ζ_A^i, that is the strategy taken by person A to decide when he generate the i^{th} onset, is calculated based on these two defined boundaries. $\zeta_A(t)$ would then be obtained by up-sampling $\zeta_A(t_A^i)$ that has calculated values for different t_A^is and is interpolated for the samples in between of two events.

3.4. Examples of strategy span definitions

Any wide enough strategy span defined based on the measured properties of the signal in a recent time window, would lead to a smooth strategy function. Two alternative assumptions are made here to define examples of strategy spans for different tasks:

3.4.1. Listening Span

Listening span tries to quantify to which extent a person pays attention to the co-performer. We define this span for person A:

Lower boundary: At a moment that a tapper does not listen to his/her rhythmic partner and prefers to tap on his own pace, the tap would occur just by keeping his tempo, no matter the properties of the other person:

$$\overleftarrow{t}_A^i = t_A^{i-1} + IOI_A(t_A^{i-1}) = t_A^{i-1} + \frac{\theta_A^{i-1}}{60} \qquad (7.)$$

Upper boundary: In an extreme case that the tapper A has lost the synchronization with person B, he tries to catch up. He thus waits to receive the upcoming tap from the other side (t_B^i) that arrives at $t_B^i + \tau_B$ and then after his reaction time (ρ_A) he generates the i^{th} tap:

$$\vec{t}_A^{\,i} = t_B^{i-1} + \tau_B + \rho_A \qquad (8.)$$

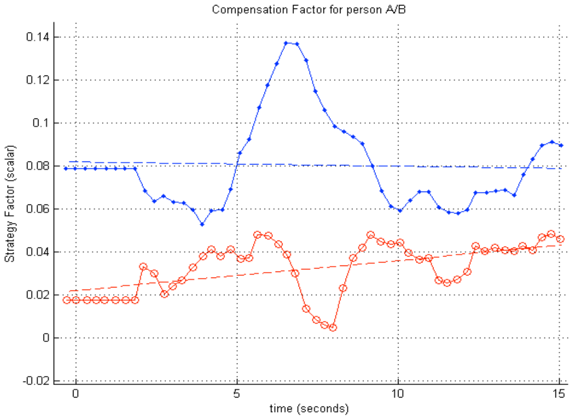


Figure 5 “Compensation” strategy functions of subjects A (•) and B (○) clapping ensemble with an initial tempo of 86 bpm and the delay of 44 ms, according to the definition in the equations (9), (10)

3.4.2. Delay compensating span

We have previously presented this definition [4]. When a person detects that the other one has a falling tempo (e.g. the incoming tap comes later than the anticipated time instant due to a channel latency), then the person decides how early he should generate the upcoming event. This parameter tends to describe to what degree a person preempts to compensate the delay:

Lower boundary: Represents the lazy strategy by which person A tries to slow down to make his generated event synchronized with what he/she anticipates from B:

$$\vec{t}_A^{\,i} = t_B^{i-1} + \tau_B + IOI(t_B^{\,i-1}) = t_B^{i-1} + \tau_B + \frac{\theta_B^{\,i-1}}{60} \qquad (9.)$$

Upper boundary: Stands the extreme sharp strategy for which person A tries to preempt and keep his tempo stable. This gives us the same lower boundary of the listening span, equation (7):

$$\vec{t}_A^{\,i} = t_A^{i-1} + \frac{\theta_A^{\,i-1}}{60} \qquad (10.)$$

Examples of derived strategy functions from a recorded ensemble trial are illustrated in Fig 6.Two other strategy functions (“Leading 1” and “Leading 2”) are illustrated too in the figure but not formulated in this paper.

4. STRATEGY PATH

The strategy path for a trial is achieved by perpendicularly projecting the strategy functions of its pair of performers, onto the Cartesian axis. In other words, by projecting the strategy function quantified for person A during the course of a specific trial on axis X, and the same strategy for the person B on axis Y, we get a two dimensional parametric path.

A strategy path for a recorded trial then represents how two ensemble musicians change their mutual strategy during the course of a jamming session. This path would involve noise, that is, stochastic uncertainty, as well as systematic trends, both individual and co-varying ones.

Strategy path is “time-free”, meaning that the time information is removed by the projection. It doesn’t directly tell us at what timestamps a node has occurred. What we know is the order in which the nodes are connected but not the time they correspond to.

Emerging clear patterns in the strategy path would reveal how that strategy of the performers changes over time for a specific trial or a group of trials sharing a common feature, e.g. the paired persons, defined subjective task, acoustical conditions, latencies or other experimental settings. Refer to Fig. 5 and comment what we see, for instance if the X-value is always higher than the Y-value (red curve in the upper left graph), then person A is “working harder” towards keeping the tempo up.

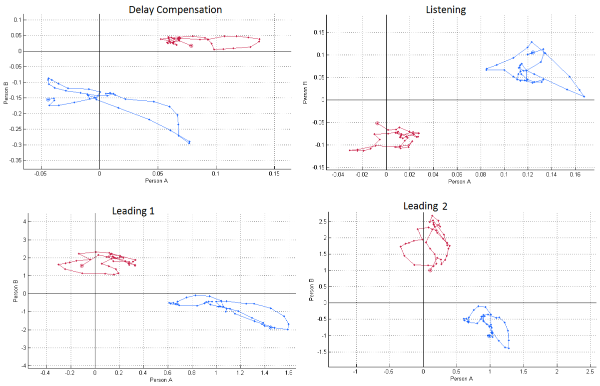


Figure 6 Four different strategy paths of the same two recorded hand-clapping trial between two different pairs both with the starting tempo of 94 bpm and the delay of 78 ms. The strategy paths are plotted based on different alternative definitions of “strategy span”. Trials are from the experiment done by Chafe. et. al [2]

5. TRIAL SIMULATION

In Fig 1 we described the process of transforming the timestamp vectors of a duo-tapping trial, into its two-dimensional strategy path.

Are the steps of this process reversible?

The answer is no. Some signal information is lost from step to step: As discussed in the previous section, we have removed the time information during the perpendicular projection. In fact a 2D strategy path does not directly tell us at what timestamps the nodes have occurred. Besides, deconvolution to get the measured instantaneous properties back from their convolved curves is non-trivial. But still:

Could we derive the tapping timestamps of a recorded trial just by having its time-free strategy path?

We will prove, by mathematical induction that the answer to this question is yes. This is not only a mathematical argument to prove that an object (here the time vectors of the original trial) exists, but also a constructive proof that give us the mathematically existing object itself.

5.1. Theorem

Knowing the time-free strategy path for a trial of a duo-tapping over a given delayed channel and given an initial state, the objective properties of the signals in time domain (including time vectors, tempo curves, IOIs, asynchronies, acceleration) can be reconstructed.

Proof by mathematical induction: To reconstruct the vectors $\mathbf{t}_A$ and $\mathbf{t}_B$ we start with a given time-free two-dimensional strategy path (ζ_A^i, ζ_B^i). The fixed properties of the channel (delays τ_A and τ_B, time signature, etc.) as well as the definition of the strategy factor (f_{lower} and f_{upper}) are also known.

Induction basis: An initial state including all the objective properties for a "zero" state is given: t_A^0, t_B^0, θ_A^0, θ_B^0, and in case that they are needed α_A^0, α_B^0, δ_A^0, δ_B^0, etc. are known.

The induction step: knowing the mentioned properties for a timestamp i-1 we can calculate them for the time stamp i:

Given t_A^{i-1}, t_B^{i-1}, θ_A^{i-1}, θ_B^{i-1} (and α_A^{i-1}, α_B^{i-1}, δ_A^{i-1}, δ_B^{i-1} in case that f_{lower} and f_{upper} have asynchrony and/or acceleration as their arguments), equations 4 and 5 provide us with $\overleftarrow{t}_A^i$ and $\overrightarrow{t}_A^i$. Since we know ζ_A^i equation 6 gives us t_A^i. The IOI of the tap correspond to the starting node is now t_A^i-t_A^{i-1} and its instantaneous tempo ($\underline{\theta}_A^i$) is then $60b/(t_A^i - t_A^{i-1})$. Having instantaneous tempos for the i^{th} tap and the previous ones we construct an up-sampled signal (again by linear interpolation) called $\underline{\theta'}(t)$ that equals the original instantaneous tempo curve $\underline{\theta}(t)$ for t≤t_A^i while is yet zero for the future moments (t>t_A^i):

$$\underline{\theta}'(t) = \begin{Bmatrix} 0 & ;t > t_A^i \\ \underline{\theta}(t) & ;t \le t_A^i \end{Bmatrix} \quad (11.)$$

Convolving $\underline{\theta'}(t)$ with $g(t)$ we get $\theta'(t)$ that happens[*] to be equal to the original smoothed (we previously called it perceived) tempo curve $\theta(t)$ for the passed times (t≤t_A^i) and equal to zero for the future moments:

$$\theta'(t) = \underline{\theta}'(t) * g(t) = \begin{Bmatrix} 0 & ;t > t_A^i \\ \theta(t) & ;t \le t_A^i \end{Bmatrix} \quad (12.)$$

$\theta_A^i = \theta(t_A^i) = \theta'(t_A^i)$ is thus known as well as t_A^i. For more complicated strategy functions in which f_{lower} and f_{upper} includes acceleration, derivatives of $\theta'(t)$ would also determine the original perceived acceleration $\alpha(t)$ for t≤t_A^i that give us α_A^i. In a similar way by knowing ζ_B^1 the objective properties for person B ($t_B^i, \theta_B^i, \alpha_B^i$) comes out too. Asynchronies then come from and $\delta_A^i = t_A^i - t_B^i$ and $\delta_B^i = t_B^i - t_A^i$ and the theorem is proved.

5.2. "Domino" Algorithm

Following the procedure described in the theorem above we have developed an algorithm called "Domino" that given the strategy path of a pair (or some general

---

[*] It apparently happens because of the nature of the memory curve $g(t)$ that is set to zero for it positive (future) inputs. This theorem is no more held for a definition of $g(t)$ that "remembers" the future. In a deterministic model in predicting "present" given the "past" we don't need to know about "future".

features of such a path) reconstructs the timing vectors of a rhythmic trial by a domino effect. The initial state (including the starting tempo) and the fixed properties of the network (e.g. latencies) are also arguments given to the algorithm and Domino then generates auditory events for both sides that could be also in real-time.

Since the strategy path in an interactive rhythmic session is also decided on-the-fly and is not known in advance, the algorithm has routines to generate random strategy paths by some initial parameters.

6. CONCLUSION

A general quantifiable measure of strategy taken by performers in an interactive rhythmic performance was presented in a form of a single-parameter strategy function. Examples of specific definitions of such functions, suitable for different scenarios of rhythmic collaboration were given.

By perpendicular projection of strategy functions onto the Cartesian axes, we transformed an ensemble performed time series to a "strategy path", to be able to describe how the performers change their behavior during the course of a trial.

Using mathematical induction it was proven that this transformation from a time domain to a strategy domain is conditionally reversible, i.e. time vectors of an ensemble trial can be reconstructed by a domino effect having its time-free strategy path and given an initial state. This algorithm provides a tool to simulate ensemble trials based on the overall strategies presumed to be driving them.

7. FUTURE WORK

Applying this deterministic method on experimental data is a way of classifying the trials in a pool of experimental data. In future studies we will learn whether strategy paths are pair-dependant and if they vary systematically from a pair of ensemble performers to another. This can be shown by applying statistical analysis on different features of strategy paths, e.g. size, pattern, direction, and center of mass.

The aspect of noisiness/uncertainty should be considered in a model aiming at describing how people behave. By recognizing recurring patterns in the strategy paths from the noise/randomness, common features of strategy paths could be revealed to explain how strategies taken by two collaborators change over time and how these functions interact with each other.

A model for strategy dynamics has practical applications: the idea of "compressing" a trial to its general strategy features and an initial state, enables us to simulate an ensemble trial even without having its exact strategy path: instead of feeding the domino algorithm with a detailed noisy path, we can provide it with a randomly generated path elaborated by those general features. Transforming this approximated strategy path to the time domain, would then construct the simulated trial. This can be applied to design a musical computer co-performer that generates auditory events adjustably collaborating with a human performer.

8. ACKNOWLEDGEMENT

We acknowledge Snorre Farner, author of [1], for early discussions on defining different strategies taking by ensemble players over a delayed network and Ola Vikholt Brunborg for important comments on the topic.

9. REFERENCES

[1] S. Farner, A. Solvang, A Sæbø, U. P. Svensson, "Ensemble hand-clapping experiments under the influence of delay and various acoustic environments," *J. Audio Eng. Soc.*, vol. 57, no. 12, pp. 1028-1041 (2009).

[2] C. Chafe, J. P. Càceres, M. Gurevich, "Effect of temporal separation on synchronization in rhythmic performance," *Perception*, vol. 39, pp. 982 – 992 (2010).

[3] B. H. Repp, "Sensorimotor Synchronization: A review of the tapping literature," *Psychonomic Bulletin & Review*, vol. 12 (6), pp. 969-992 (2005).

[4] N. Darabi, P. Svensson, S. Farner, "Quantifying the strategy taken by a pair of ensemble hand-clappers under the influence of delay," pres. at the 125th AES convention, San Francisco, CA, USA, Oct. 2008, Preprint no. 7567 (2008).

[5] W. A. Wickelgren, "Acoustic similarity and retroactive interference in short-term memory," *J. of verbal learning and verbal behavior*, Vol. 4, pp. 53-61 (1965).