

Doctoral Thesis (abridged)
博士論文（要約）

Local and Global Regularity of Musical Time Series

（ 音楽時系列の局所的および大域的規則性 ）

Miwa Fukino
(吹野美和)

Supervisor: Professor Kazuyuki Aihara
(合原一幸 教授)

June, 2016

**Department of Mathematical Informatics
Graduate School of Information Science and Technology
the University of Tokyo**

Abstract

In this thesis, I propose a nonlinear time series method for characterizing two layers of regularity simultaneously. The key of the method is using the recurrence plots hierarchically, which allows us to preserve the underlying regularities behind the original time series. We demonstrate the proposed method with musical data. The proposed method enables us to visualize both the local and the global musical regularities, or two different features at the same time. Furthermore, the determinism scores imply that the proposed method may be useful for analyzing emotional response to the music.

Grasping both macroscopic and microscopic information is very useful, when we aim to understand the underlying dynamics generating a long time series. However, there are few methods of nonlinear time series analysis which can treat these two pieces of information simultaneously. Here, we propose a novel method called "recurrence plot of recurrence plots (RPofRPs)" with which recurrence plots are used hierarchically to represent local and global regularities. At the lower layer, we separate a long time series into many segments of short-term time series, and simply obtain many short-term (unthresholded) recurrence plots. At the higher layer, on the other hand, we calculate distances between every pair of the short-term RPs and obtain a recurrence plot, which is the RPofRPs. We apply this method to music pieces of 44.1kHz 16bit monaural linear PCM. Our results show a continuous periodic regularity, rhythmic pattern change, gradual tempo transition, and phrase similarity between phrases played even in different key scales. Furthermore, we calculated transitions of determinism (DET) scores for both conventional recurrence plots and RPofRPs. We found that DET scores of RPofRPs for Mozart and Beethoven pieces tend to show high scores at the first chorus, then vary largely at the middle, and gradually increase towards the end. Thus, RPofRPs can comprehensively represent both macroscopic and microscopic characteristics of the underlying dynamics behind the data, by only using simple calculations. Therefore, it possibly provides another viewpoint of music based on nonlinear dynamics.

Contents

Chapter 1	Introduction	1
1.1	Motivation	1
1.2	Organization of the Thesis	2
Chapter 2	Overview of Recurrence Plots	5
2.1	Introduction	5
2.2	Definition of Recurrence Plots	7
2.3	Meaning of Graphical Patterns of Recurrence Plots	7
2.4	Recurrence Quantification Analysis	8
2.5	Variations of Recurrence Plots	9
Chapter 3	Hierarchical Recurrence Plots	13
3.1	Introduction	13
3.2	Methods	14
3.2.1	Basic Method	14
3.2.2	Definition of short-term recurrence plots (short-term RP) . .	14
3.2.3	Definition of recurrence plot of recurrence plots (RPofRPs) .	16
3.2.4	Input Data	16
3.2.5	Parameter Values	19
3.2.6	Determinism (DET)	20
Chapter 4	Results	21
4.1	Unthresholded RPofRPs and Short-term RPs	21
4.2	Thresholded RPofRPs	23
4.2.1	Toy Model	23
4.2.2	Bach1	24
4.2.3	Mozart1	24
4.2.4	Beethoven1	24
4.3	Determinism (DET)	27

Chapter 5	Discussion	31
5.1	Extracting a Perturbation	31
5.2	Distance Measure	35
5.3	Jazz and Pops Examples	41
5.4	Future Work	44
Chapter 6	Conclusion	47
	Acknowledgments	49
	Bibliography	51
AppendixA	Thresholded short-term RPs of Bach1	57
AppendixB	Thresholded short-term RPs of Mozart1	67
AppendixC	Thresholded short-term RPs of Beethoven1	97

Chapter 1

Introduction

1.1 Motivation

Music evokes emotion. Many musical factors, such as timbre, melody, harmony, meter, and beat, constitute a musical structure and produce an emotional response in the listener. The relationship between emotion and the physical properties of music has been studied [42, 19, 8, 27, 38, 44, 43], but is not fully understood.

The reciprocal feedback model of musical response [16] explains one aspect of the difficulty. The model describes that the musical response is caused by not only the characteristics of music itself but also listener's property and situational contexts. The states of these three factors change according to the time evolution complexly. Thus, the relationship between emotion and the physical properties cannot be described with a static one-to-one correlation.

Recent researches on the neurobiology have revealed the prediction mechanism [13, 56, 59] and music related brain structures [29, 50, 60, 61]. Based on the idea of prediction or expectation mechanism, several mathematical models are proposed [46, 31, 33, 34, 48, 54, 58].

I raise three key points to analyze the relationship between emotion and physical properties of music.

The first key point is the *regularity*. Our sensory system of cortex is continuously predicting the incoming stimuli and detecting the change of regularities [18, 39, 41, 40]. When some regular pattern starts, we are aware of it consciously. When the regular pattern is continuously presented, we are gradually no longer aware of it. Then we become aware of it consciously, only when our unconscious prediction system finds some irregularity. Unaware information is swiftly processed

2 Chapter 1 Introduction

and thrown away. We feel only emotions without being aware of the detailed information consciously.

The next key point is *hierarchy*. Our prediction mechanism of auditory sensory system processes the auditory properties hierarchically and in parallel [30, 46, 57]. Pitches, pitch patterns, or chords are processed in the lower layers, and bars, phrasings, or musical structures are processed in the higher layers. That is, simple and local information is processed in the lower layers and comprehensive and global information is processed in the higher layers. In addition, there are two pathways of emotion inside the brain [32, 9]: a rough and fast pathway related to the basal ganglia and the limbic system, and a detailed and slow pathway related to the cortex. Two pathways are integrated into our conscious emotions.

The third key point is *scalability*. The ability of both pattern prediction and function of hierarchy vary from person to person. What is more, they change inside the individuals. Therefore, a scalable function should be considered.

In addition, many previous methods use only symbolized properties which we are clearly aware of. However, we should consider the unconscious part of the information, because emotion is essentially related to the unaware information.

Consequently, I focused on the three key points: *regularity*, *hierarchy*, and *scalability* to keep the underlying essential information of music as much as possible.

1.2 Organization of the Thesis

The rest of the thesis is organized as follows.

In Chapter 2, I introduce the overview of the recurrence plots. First, I explain the capability, mathematical definition, and meaning of the graphical patterns of the recurrence plots. Then, I describe the recurrence quantification analysis which is a measure to quantify the patterns. Finally I illustrate the variation of the method.

In Chapter 3, I describe the hierarchical recurrence plots which is our proposal method. First, I describe the basic method and mathematical definition. Next, I

explain input data and parameter values. Then I describe the definition of determinism (*DET*), which is one of measures of recurrence quantification analysis.

In Chapter 4, I show the results of our proposal method described in Chapter 3.

Chapter 5 discusses the results, further extension, and future work.

In Chapter 6, I conclude this thesis.

Chapter 3 and 4 are based on our published paper, "Fukino, M., Hirata, Y., and Aihara, K. (2016). Coarse-graining time series data: Recurrence plot of recurrence plots and its application to music. *Chaos*, 26(2), 023116." Chapter 1, 2, and 5 are partly based on the paper and a news story "Fukino, M., Hirata, Y., and Aihara, K., (2016). Music Visualized by Nonlinear Time Series Analysis, *SIAM News* 49(4), May, 1,4". I extended and deepened them by adding further new analysis.

Chapter 2

Overview of Recurrence Plots

2.1 Introduction

A recurrence plot is based on a concept of recurrence [45], and an important analysis tool for nonlinear time series data [11, 37, 23]. It is a two-dimensional plot that visualizes recurrences of the data. Both horizontal and vertical axes represent the same time axis, preserving the information for the underlying dynamics completely. If a distance between states for two times are close, a dot is plotted at the corresponding place, and if not, a dot is not plotted there.

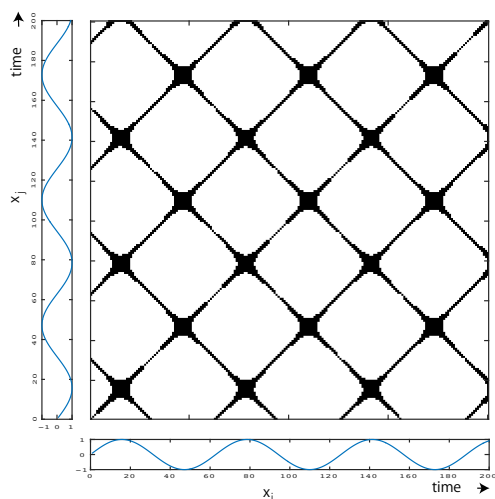


Figure 2.1. Thresholded Recurrence Plots for Sine Wave

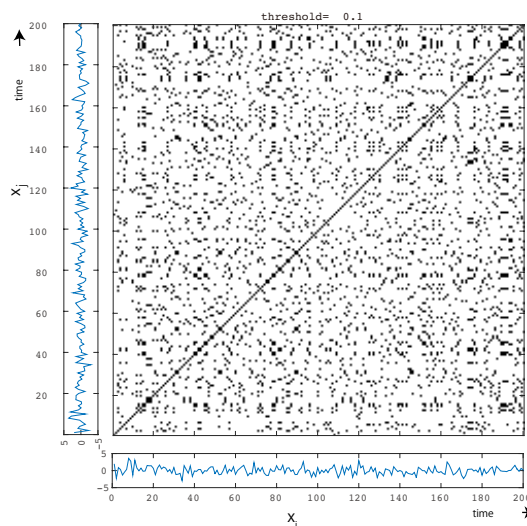


Figure 2.2. Thresholded Recurrence Plots for White Noise

6 Chapter 2 Overview of Recurrence Plots

The recurrence plots can reveal many things about the underlying dynamics of the data. For instance, a periodic time series produces a periodic pattern in the recurrence plot, as shown in Fig. 2.1. A time series of white noise produces a recurrence plot where points are plotted almost randomly (See Fig. 2.2). A time series generated from deterministic chaos creates a recurrence plot containing many short diagonal segments (See Figs. 2.3, 2.4). If two time series generate the same recurrence plot, their underlying dynamical rules are equivalent [47]. By using recurrence plots, one can grasp the underlying dynamics such as serial dependence [22] and consistency with Devaney's mathematical definition of deterministic chaos [21].

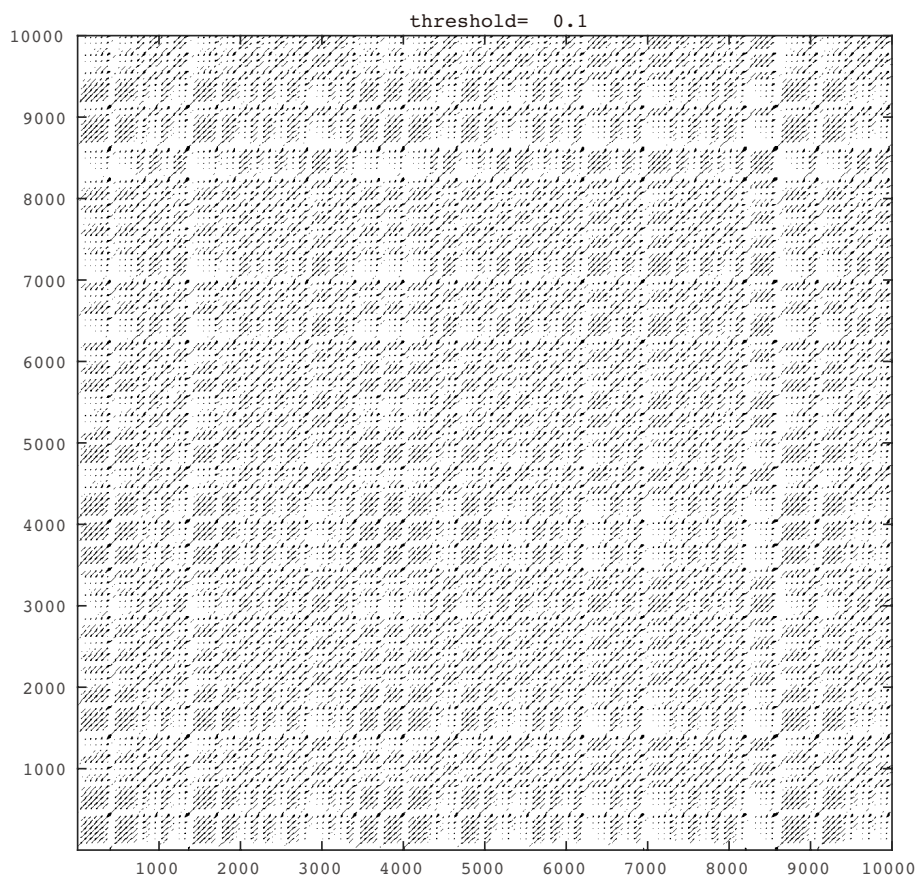


Figure 2.3. Thresholded Recurrence Plots for Lorenz Model. We set $\delta = 10$, $\gamma = 40$, $\beta = 8/3$, and $\varepsilon = 0.1$.

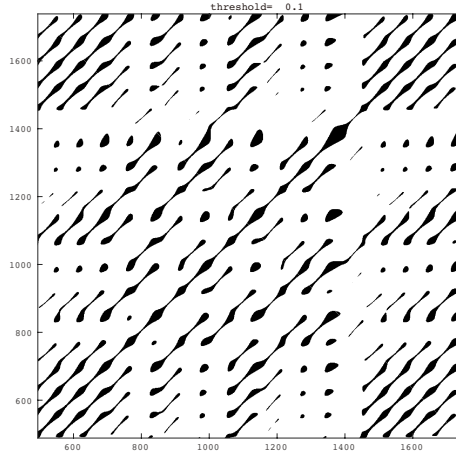


Figure 2.4. Enlarged view of Fig. 2.3. We set $\varepsilon = 0.1$.

2.2 Definition of Recurrence Plots

Let M be a metric space, and d , a distance defined on M . In addition, denote, by $s(i) \in M$, the i th point of a time series s . Then, a recurrence plot is defined as

$$R(i, j) = \begin{cases} 1, & d(s(i), s(j)) < \varepsilon \\ 0, & \text{otherwise.} \end{cases} \quad (2.1)$$

When $R(i, j) = 1$, we plot a dot at (i, j) , and when $R(i, j) = 0$, we do not plot a dot on (i, j) . Here ε is called the threshold. We may use the Euclidean norm, the maximum norm, the minimum norm and others for the distance d . In an unthresholded recurrence plot, we do not use the threshold ε , and express distances as color [25]. Namely the definition for the unthresholded recurrence plot is as

$$R'(i, j) = d(s(i), s(j)). \quad (2.2)$$

2.3 Meaning of Graphical Patterns of Recurrence Plots

Many of nonlinear time series analysis are based on the theory of dynamical system, where the time evolution is defined in a phase space [28]. Thus $s(i)$ and $s(j)$ defined in Section 2.2 are considered to be placed in a phase space, and d is also a distance

measured in the space.

Table 2.1 shows meanings of typical patterns which are found in recurrence plots [37]. As roughly described in Section 2.1, periodic patterns mean periodicity in the process. The time between periodic patterns corresponds to the period of the process. The main diagonal line is called LOI (line of identity). Diagonal lines which are parallel to LOI mean that the state evolution is similar among different epochs. It implies that the process is deterministic. The process could be chaotic when the diagonal lines occur beside single isolated points. Diagonal lines which are orthogonal to the LOI mean that the state evolution is similar at different times when the time axis is reverse. Sometimes, this orthogonal line means an embedding dimension was not sufficient. Vertical and horizontal lines mean that states do not evolve or change slowly. Long bowed lines mean that the state evolution is similar among different epochs in different velocity. This similarity implies that the dynamics of the system is changing.

Table 2.1. Meanings of Patterns in the Recurrence Plots

No.	Pattern	Meaning
1	Periodic patterns	Periodicity in the process
2	Diagonal lines (parallel to the LOI)	The evolution of states is similar at different epochs
3	Diagonal lines (orthogonal to the LOI)	The evolution of states is similar at different times but with reverse time
4	Vertical and horizontal lines	Some states do not change or change slowly for some time
5	Long bowed lines	The evolution of states is similar at different epochs but with different velocity

2.4 Recurrence Quantification Analysis

To quantify the patterns of the recurrence plots, several measures have been proposed. They are called the recurrence quantification analysis (RQA) [37].

Recurrence rate (RR) is the simplest measure that represents the density of points in the recurrence plots defined as

$$RR(\varepsilon) = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j}(\varepsilon), \quad (2.3)$$

where N is the width of the recurrence plots.

The next measures are based on diagonal lines. The histogram $P(\varepsilon, l)$ of diagonal line is

$$P(\varepsilon, l) = \sum_{i,j=1}^N (1 - R_{i-1,j-1}(\varepsilon))(1 - R_{i+l,j+l}(\varepsilon)) \prod_{k=0}^{l-1} R_{i+k,j+k}(\varepsilon), \quad (2.4)$$

where l is the length of diagonal lines.

Using P , the average length of diagonal lines L is defined as

$$L = \frac{\sum_{l=l_{min}}^N lP(\varepsilon, l)}{\sum_{l=l_{min}}^N P(\varepsilon, l)}, \quad (2.5)$$

where l_{min} is the shortest diagonal line to be considered to calculate the histogram P . Then the determinism DET which represents the ratio of recurrence points is

$$DET = \frac{\sum_{l=l_{min}}^N lP(\varepsilon, l)}{\sum_{l=1}^N lP(\varepsilon, l)}. \quad (2.6)$$

2.5 Variations of Recurrence Plots

Since the recurrence plots has a very simple definition, there are many variations and extensions.

Cross Recurrence Plots

Cross recurrence plot is an extension of the recurrence plots to analyze a dependency between two time series [62, 36]. It is defined by

$$CR(i, j) = \begin{cases} 1, & d(s_1(i), s_2(j)) < \varepsilon \\ 0, & \text{otherwise,} \end{cases} \quad (2.7)$$

where $s_1(i)$ is one time series and $s_2(j)$ is another time series (See Fig. 2.5). The phase space between two time series are required to be same. However, the length of time series is not necessarily to be same, so the plot is sometimes not square.

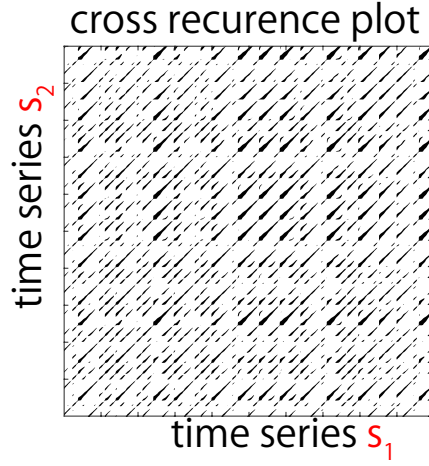


Figure 2.5. Cross Recurrence Plots

Joint Recurrence Plots

While the cross recurrence plots force two time series to be in the same phase space, joint recurrence plots can analyze the two by means of comparing the two recurrence plots calculated individually [49]. It is defined by

$$JR(i, j) = R_1(i, j)R_2(i, j), \quad (2.8)$$

where $R_1(i, j)$ and $R_2(i, j)$ are the first and second recurrence plots to compare, respectively. If both of the recurrences happen at the same time, the rate of plots increases.

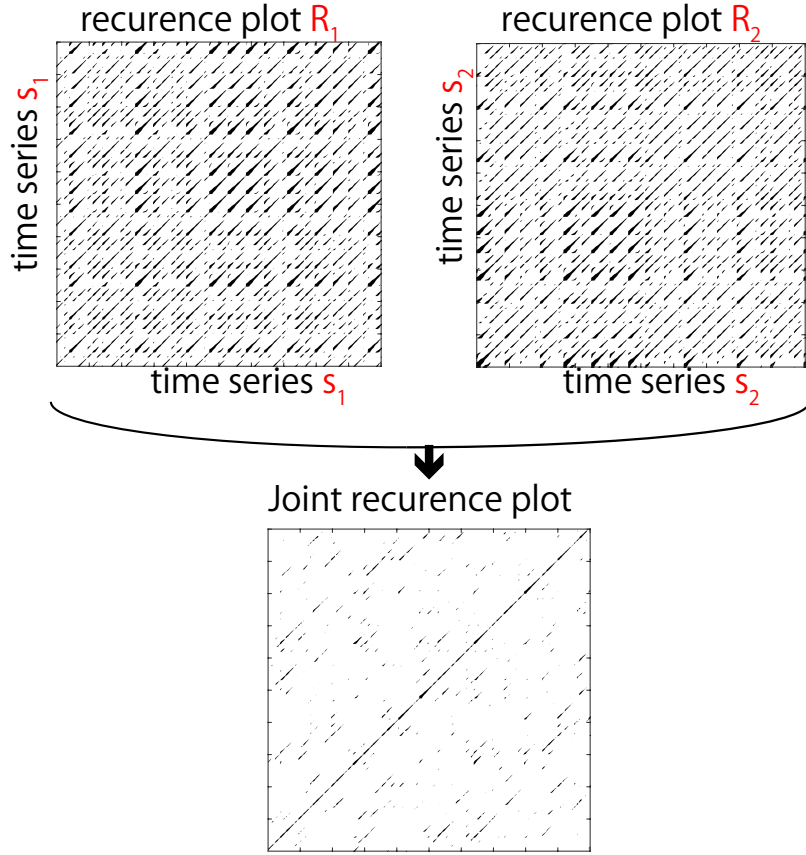


Figure 2.6. Joint Recurrence Plots

Recurrence Plots Using Edit Distance of Marked Point Process

One can apply the edit distance for a marked point process [53, 24] to calculate a distance between every pair of time windows. The marked point process is a series of events that consist of time and additional information, or marks (see Fig. 2.7(a))[24]. We can obtain the edit distance of marked point process by calculating the smallest cost to convert one time window (Fig. 2.7(a)) to another (Fig. 2.7(d)). The cost is calculated by shift (Fig. 2.7(a) to (b)), addition (Fig. 2.7(b) to (c)) and deletion (Fig. 2.7(c) to (d)) for timing information and the value of marks.

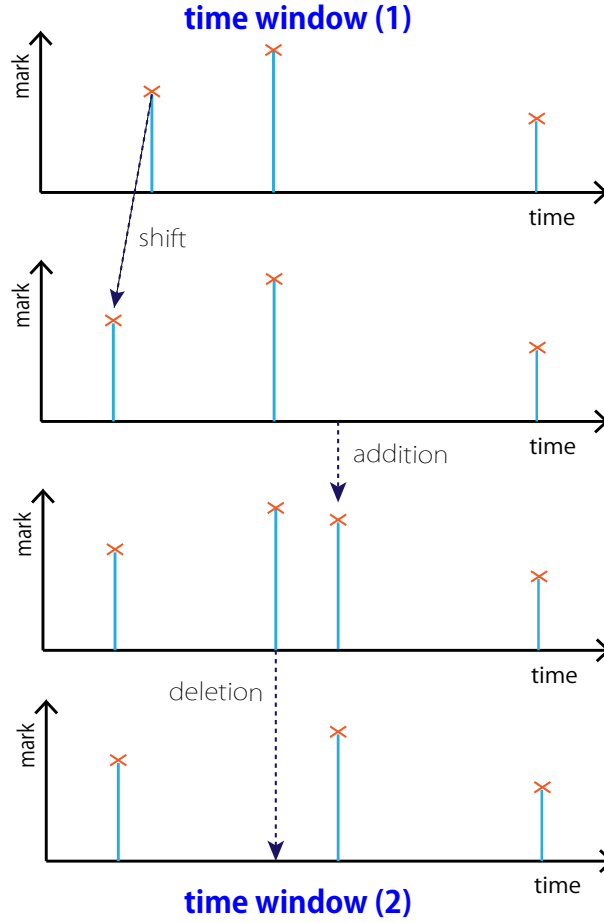


Figure 2.7. Explanation of Edit Distance of the Marked Point Process

Same as the marked point process, by replacing the distance function with another distance function, we can analyze data with various properties; for example, using Canberra distance (for a review of distance measure, see [5]) to focus on the distance near zero, using order pattern instead of the spatial distance in the phase space (order patterns recurrence plots [15], and time series of networks [26]).

Thus, by looking at a recurrence plot directly or extending it, we can grasp the properties of the underlying dynamics.

Chapter 3

Hierarchical Recurrence Plots

In this chapter, I describe our proposal method which uses the recurrence plots hierarchically.

3.1 Introduction

As described in the Chapter 2, the recurrence plots is a very useful tool for analyzing the time series.

Due to the visibility of the plot, we normally apply a recurrence plot to a time series of length up to 10000. Based on this standard, it is impossible to apply a recurrence plot to music time series data directly because it is too huge: such a dataset is usually recorded at 44.1kHz with 16bit linear PCM (Pulse Code Modulation) and thus has 44100 time points per second, meaning that a 5 minutes monaural music piece contains 13,230,000 time points. If we simply subsample such musical data and try to draw its recurrence plots, we cannot grasp its macroscopic structure. The previous works [1, 51] analyzed the structural similarity between pieces by calculating tonality feature vectors and using the vectors as inputs for a recurrence plot. The feature vectors were calculated as a probability distribution of 12, 24 or 36 chromas, and tuned for analyzing which pitch or chord constructs the piece structure. The problem on such analysis is that the underlying rich information except for tonality is lost in such a recurrence plot. Therefore, we focus on describing both microscopic and macroscopic information of music pieces without losing essential information. To do that, we should avoid the linear analysis method. For this sake, we found a hint in our previous work [26], where we divided a long time series, e.g. foreign exchange market data or brain data, into several short-term segments, calculated a recurrence plot for each observable for each segment, represented relations among the observables as a graph structure, and characterized the macroscopic change using the series of graphs. However, we cannot apply this method directly to music pieces

because it assumes that several observables are measured simultaneously, while we here assume that a music piece is given as a scalar time series.

Thus, we propose a novel method [12] which represents a macroscopic regularity, by means of using recurrence plots hierarchically. In this method, first, we divide a long time series into short-term segments, and calculate a short-term recurrence plot for each segment to represent a microscopic regularity. Second, we calculate another recurrence plot using short-term recurrence plots. As a result, we can observe macroscopic properties without losing a microscopic regularity much.

3.2 Methods

3.2.1 Basic Method

We analyze a time series using RP hierarchically. As shown in Fig. 3.1, we divide a long time series into short-term segments at constant interval, and calculate a recurrence plot for each segment. We call this a "short-term recurrence plot (short-term RP)". The window size and overlap with the next window are set arbitrarily according to data. Then we calculate another RP using short-term RPs. We call this as a "recurrence plot of recurrence plots (RPofRPs)". While the short-term RPs represent a regularity of neighbors inside the segment, the RPofRPs represents a regularity among "short-term RPs' regularities".

3.2.2 Definition of short-term recurrence plots (short-term RP)

We here use delay coordinates for reconstructing states for the underlying dynamics. Thus the delay time and the dimension should be considered to calculate short-term RPs. Let $\{s(t) \in R | t = 1, 2, 3 \dots, T\}$ be a time series of length T . When we denote time, delay time, and dimension by t , τ , and D , respectively, delay coordinates $S(t)$ can be written as

$$S(t) = [s(t), s(t + \tau), s(t + 2\tau), \dots, s(t + (D - 1)\tau)], \quad (3.1)$$

where t moves 1 to $T - (D - 1)\tau$. We use the Euclidean distance between $S(t_1)$ and $S(t_2)$ as follows:

$$d_s(S(t_1), S(t_2)) = \sqrt{\sum_{k=0}^{D-1} (s(t_1 + k\tau) - s(t_2 + k\tau))^2}. \quad (3.2)$$

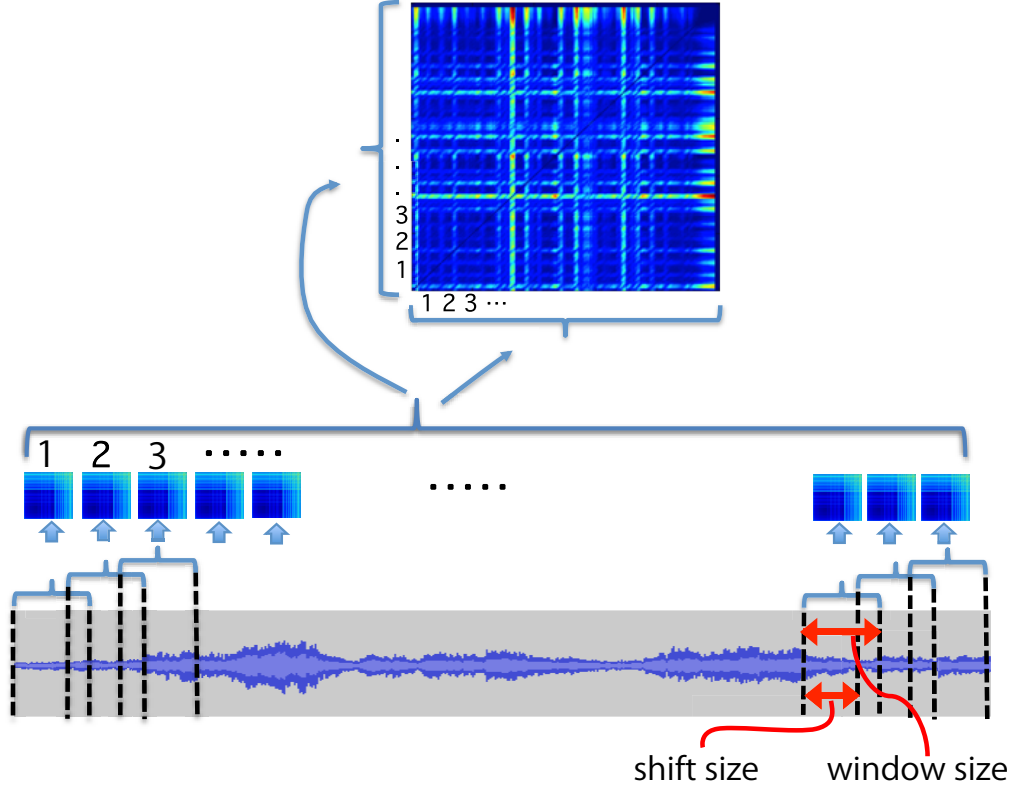


Figure 3.1. Recurrence plot of recurrence plots: we divide a long time series into short-term segments at a constant interval, and calculate multiple short-term RPs. Then using short term RPs, we calculate a RPofRPs.

We select the Euclidean distance as a general measure at this time. Variation of the measures will be discussed in Section 5.2. Using d_s , we define the n th unthresholded short-term RP X'_n as

$$X'_n(i, j) = d_s(S(nh + i\tau + 1), S(nh + j\tau + 1)), \quad (3.3)$$

where h is the shift width. When we let W be a window size, i and j move from 0 to $\lfloor (W/\tau) - 1 \rfloor$, while n moves from 0 to $\lfloor (T - (D - 2)\tau - W - 1)/h \rfloor$. We use the unthresholded short-term RPs X'_n to define a RPofRPs.

Now, we define a thresholded version of short-term RP X_n using X'_n . For each column of the X_n , we plot a dot at the row the distance to which is within top $100 \times \varepsilon_s$ % from the closest of the same column of X'_n . Then, we symmetrize the recurrence plot by plotting a dot at (j, i) if a dot is plotted at (i, j) in the original

recurrence plot.

3.2.3 Definition of recurrence plot of recurrence plots (RPofRPs)

A distance between the n th and m th unthresholded RPs X'_n and X'_m is defined as

$$d_r(X'_n, X'_m) = \sqrt{\sum_{i,j=1}^{N_x} (X'_n(i, j) - X'_m(i, j))^2}, \quad (3.4)$$

where N_X is the largest index for short-term RPs. We select the Euclidean distance as a general measure at this time. Variation of the measures will be discussed in Section 5.2.

Using d_r , we define an unthresholded RPofRPs as

$$Z'(n, m) = d_r(X'_n, X'_m). \quad (3.5)$$

In addition, we define a thresholded version of RPofRPs Z using Z' . For each column of the Z , we plot a dot at the row the distance to which is within top $100 \times \varepsilon_r \%$ from the closest of the same column of Z' . Then, we symmetrize the RPofRPs by plotting a point on (m, n) if a point is plotted at (n, m) in the original RPofRPs. We first use the unthresholded RPofRPs Z' to grasp a big picture, and then use the thresholded RPofRPs to analyze the details.

3.2.4 Input Data

3.2.4.1 Toy Model Data

First, we tested the performance of our proposed method using toy model data. We used a dataset generated by the following Lorenz model:

$$\frac{dx}{dt} = -\delta(x - y), \quad (3.6)$$

$$\frac{dy}{dt} = -xz + rx - y, \quad (3.7)$$

$$\frac{dz}{dt} = xy - \beta z, \quad (3.8)$$

where t is time, x, y, z are variables, and δ, r, β are constants. We set $\delta = 5, r = 40$,

and $\beta = 8/3$, if not mentioned. The initial values were set to $[x, y, z] = [0.0, 1.0, 1.0]$. We integrated the above equations from $t = 0$ to 2000 at 0.01 intervals, and used generated x, y , and z vector data for the analysis. Figure 3.2 shows the attractor of generated data of first 8000 points.

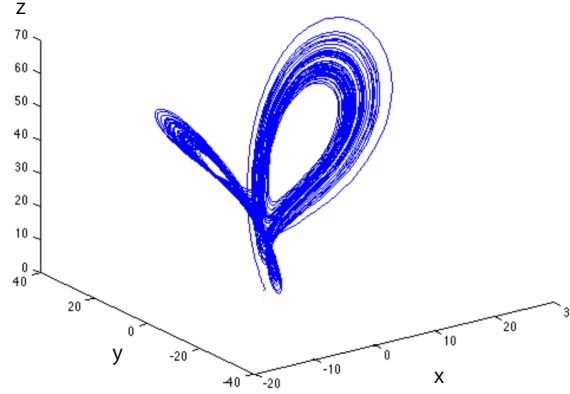


Figure 3.2. Input data (toy model): Attractor of the Lorenz model. We set $\delta = 5$, $r = 40$, and $\beta = 8/3$.

3.2.4.2 Music Data

We used classic piano music pieces shown in Table 3.1 . All pieces are audio recordings of piano only, played by expert pianists. We call these pieces by the piece numbers shown in the right column of Table 3.1. Figures 3.3, 3.4 and 3.5 show the starting part of the musical score and the whole waveform for each title. In Bach1, arpeggio phrases were constantly repeated through whole the piece, chords of arpeggio were gradually changing, and the amplitude was constant. In Mozart1, the theme melody known as "twinkle, twinkle, little star..." was followed by its twelve variations. The variations are for chords, rhythm, tempo, meter, amplitude, or ornament of phrases and notes. The dynamic range of amplitude was wide. Beethoven1 took the Sonata form which consisted of sections: the Introduction, the Exposition, the Development, the Recapitulation, and the Coda. The dynamic range of amplitude was also wide. The musical data were taken from the RWC Music Database (piece number: RWC-MDB-C-2001 No.25 Tr.13; No.27; No.28) [14]. We imported the musical data to PC in the linear PCM 44.1kHz 16bit stereo format, and converted to monaural with audio editing software.

Table 3.1. Musical data

Composer	Title	Piece Number
Johann Sebastian Bach	The Well-Tempered Clavier, Book 1, no.1 in C major BWV.846. Prelude	Bach1
Wolfgang Amadeus Mozart	Variations on " Ah! Vous dirai-je maman " K.265	Mozart1
Ludwig van Beethoven	Piano Sonata no.23 in F minor, op.57 'Appassionata'. 1st mvmt.	Beethoven1

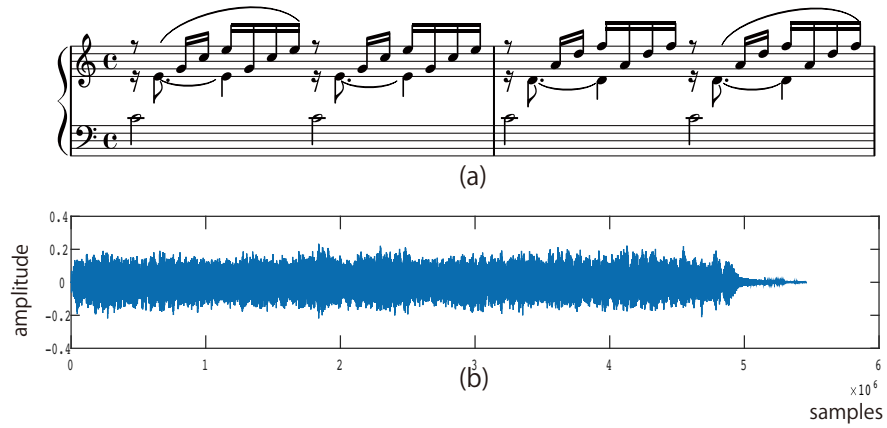


Figure 3.3. Input data (Bach1): (a) Score of the first 2 bars of J. -S. Bach, "The Well-Tempered Clavier, Book 1, no.1 in C major BWV.846. Prelude". (b) Waveform of the whole tune. linear PCM, 44.1kHz 16bit mono.

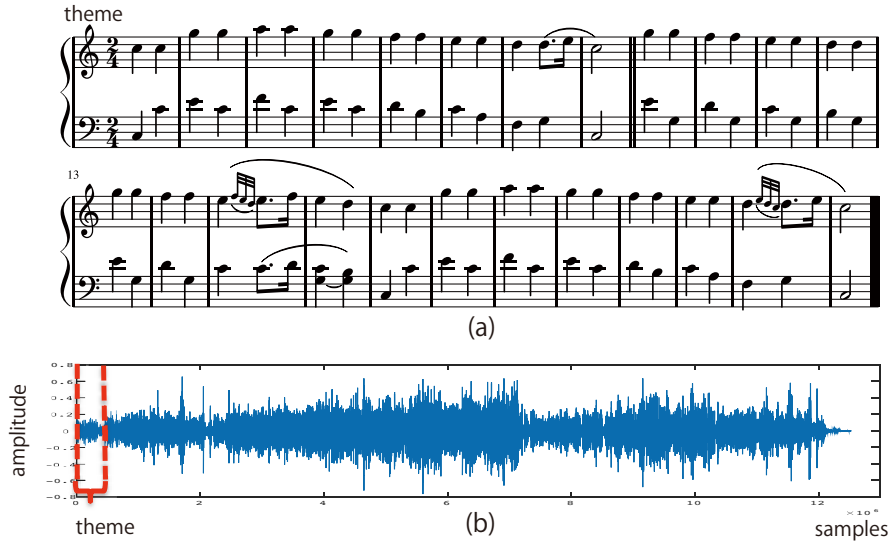


Figure 3.4. Input data (Mozart1): (a) Score of the theme of W. A. Mozart, "Variations on " Ah! Vous dirai-je maman " K.265". (b) Waveform of the whole tune. linear PCM, 44.1kHz 16bit mono.

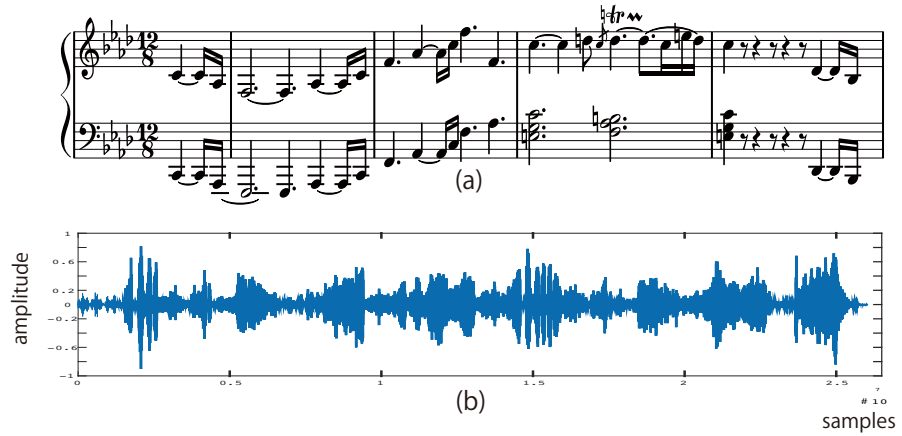


Figure 3.5. Input data (Beethoven1): (a) Score of the starting bars of L. V. Beethoven, " Piano Sonata no.23 in F minor, op.57 'Appassionata'. 1st mvmt.". (b) Waveform of the whole tune. linear PCM, 44.1kHz 16bit mono.

3.2.5 Parameter Values

We set the following parameter values to calculate short-term RPs using input data described above. For the toy model, the window size of short-term RP was set to 4000, and the shift width was 1000 (equivalent to 1/4 of the shift width), while the

delay time was set to 1, and dimension was set to 3 (x, y and z). We set the window size of the musical data to 4 seconds (44.1kHz x 4 seconds = 176,400 samples), and the shift width to 1 second (1/4 of the window size), while we set the delay time to 40, and the dimension to 5. We used mutual information to estimate the delay time, and false nearest neighbors to estimate the embedding dimension. The number of bins for calculating the mutual information was set to 32. We considered more than a half of data points to calculate the false nearest neighbors. Since all of the musical data were on the same recording conditions, we used the resulting values of Beethoven1 as values for all the music pieces. We applied the MATLAB functions (`ami()` and `false_neighbors_kd()`) which are open on the web by the author [6]. We set ε_r to 0.1 for the toy model, Mozart1 and Beethoven1, while we set ε_r to 0.3 for Bach1. That is because the distances were kept constantly close in Bach1 and there were many continuous diagonal lines on the unthresholded RPofRPs of Bach1 (see Fig. 4.3 and the result section).

3.2.6 Determinism (DET)

One of the important measures to quantify a recurrence plot is the determinism (*DET*) [25]. When we denote the length of diagonal lines on RP by l , and the width of RP by N , we can define the histogram for the diagonal lines as

$$P(\varepsilon, l) = \sum_{i,j=1}^N (1 - R_{i-1,j-1}(\varepsilon))(1 - R_{i+l,j+l}(\varepsilon)) \prod_{k=0}^{l-1} R_{i+k,j+k}(\varepsilon), \quad (3.9)$$

where ε is the threshold. Using this P , the *DET* is defined as

$$DET = \frac{\sum_{l=l_{min}}^N lP(\varepsilon, l)}{\sum_{l=1}^N lP(\varepsilon, l)}. \quad (3.10)$$

DET ranges from 0 to 1. If the value of *DET* is larger, the determinism for the underlying dynamics tends to be stronger.

Using *DET*, we quantified the short-term RPs and the RPofRPs. We set the plotting rate per column ε_s of short-term RPs X_n to 0.1, while the threshold ε_r of RPofRPs Z was set to the same value as the one described in the previous section.

Chapter 4

Results

As described in Section 3.2, the short-term RPs represent a regularity of neighbors inside the segment and the RPofRPs represents a regularity among "short-term RPs' regularities". In this chapter, first, we show unthresholded RPofRPs and unthresholded short-term RPs, and describe the overview. Second, we show thresholded RPofRPs, and explain them in detail. Finally, we show transitions of *DET* for the musical data.

4.1 Unthresholded RPofRPs and Short-term RPs

We show unthresholded RPofRPs and unthresholded short-term RPs for toy model data and musical data in Fig. 4.1, where (a), (c), (e) and (g) are unthresholded RPofRPs, and (b), (d), (f) and (h) are multiple short-term RPs. While (a) and (b) are for the toy model, (c) and (d) are for Bach1, (e) and (f) are for Mozart1, and (g) and (h) are for Beethoven1. For the musical data, the unit of the axes (time scale) for RPofRPs is in seconds, which is equivalent to the serial number of short-term RPs, where the unit of axes (time scale) for short-term RPs is in " $\times 10^{-3}$ seconds", which is equivalent to "samples/ τ ". The color shows the relative distance: the dark blue means nearer; the red means farther.

Now we overview the feature of Fig. 4.1. The RPofRPs for the toy model (Fig. 4.1(a)) contains many diagonal lines, whose colors are slightly changing, while in the short-term RP for the toy model shown in Fig. 4.1(b), blue, yellow, and red colors show a clear checkerboard pattern. From this, we can grasp that, in the microscopic view, the distance becomes nearer and farther alternatively. If both time points are at the root of the wings of the Lorenz attractor, the distance is nearer. If two time points are at the tops of the wings, the distance is farther. In the macroscopic

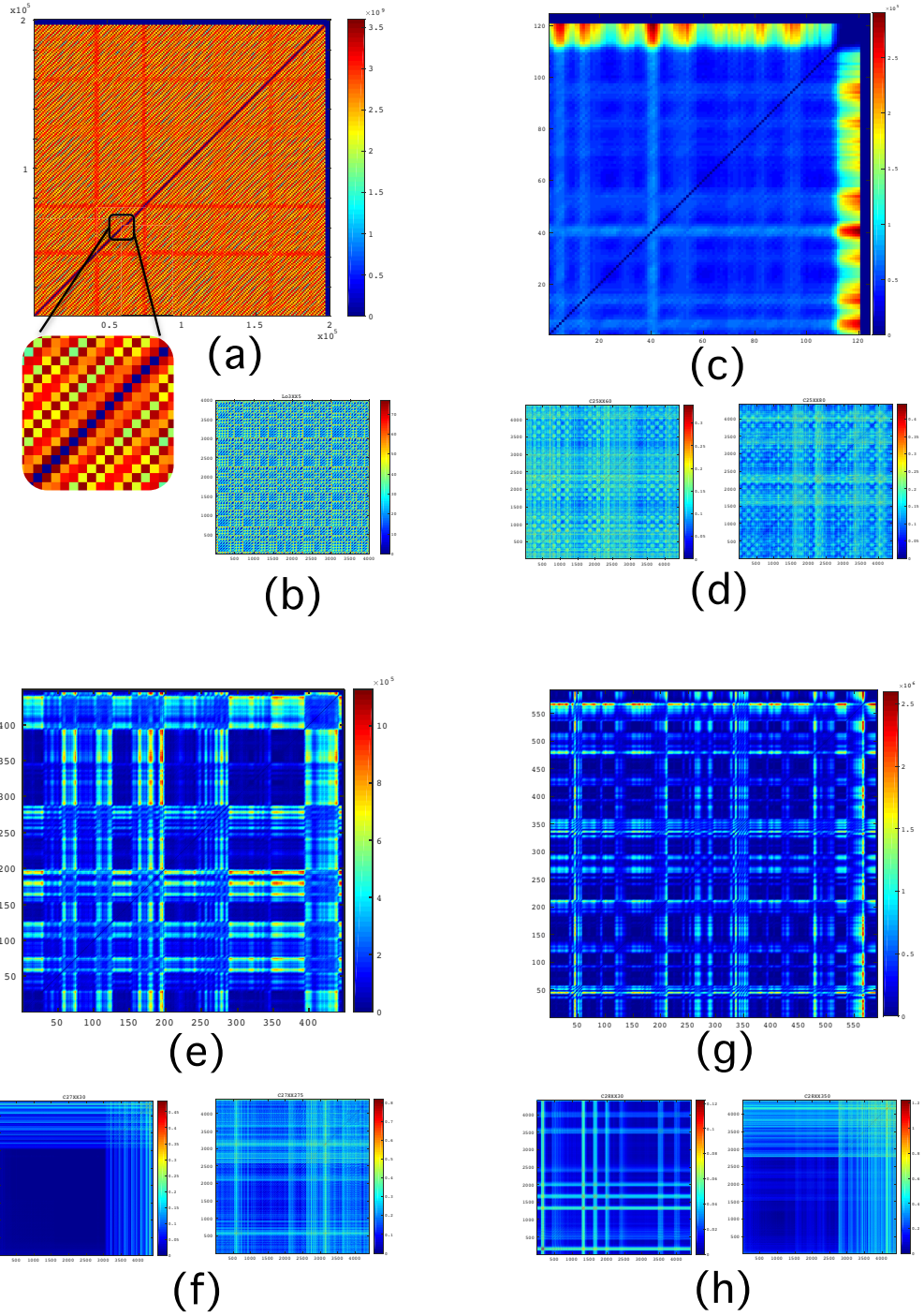


Figure 4.1. Unthresholded RPofRPs and short-term RPs : (a) RPofRPs of the toy model (unthresholded) with enlarged view. (b) short-term RP of the toy model data. (c) RPofRP of Bach1. (d) short-term RP of Bach1 (left:No.60, right:No.80). (e) RPofRPs of Mozart1. (f) short-term RP of Mozart1 (No.30,275). (g) RPofRPs of Beethoven1. (h) short-term RP of Beethoven1 (No.30,350). We set window size = 4000, shift size = 1000, delay time=1, and dimension=3 in (b). In (d), (f), and (h) we set window size = 4 seconds (4410), shift size = 1 second (1100), delay time = 40, and dimension = 5.

view, the shape of the attractor of each segment changes periodically. Whereas, the RPofRPs for Bach1 (Fig. 4.1(c)) represents that the distances are close in almost whole the piece except for the last part, while short-term RPs for Bach1 (Fig. 4.1(d)) represent that the distances become nearer and farther repeatedly. From this, we can understand that the data keep specific similar patterns continuously. RPofRPs for Mozart1 (Fig. 4.1(e)) and Beethoven (Fig. 4.1(g)) represent that there are many sections, for example, where similar patterns were repeated, but the shapes of patterns varied slightly.

4.2 Thresholded RPofRPs

Next, we drew the thresholded RPofRPs.

4.2.1 Toy Model

Figure 4.2 shows the thresholded RPofRPs for the toy model. We can find many diagonal lines clearly. This finding means that the distances between the attractors of short-term segments have a nearly periodic property, as we mentioned above.

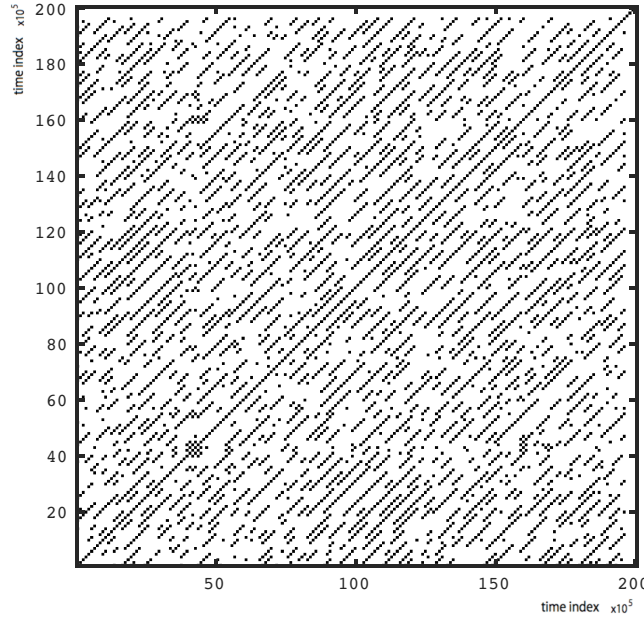


Figure 4.2. Thresholded RPofRPs (the toy model) with $\varepsilon_r = 0.1$

4.2.2 Bach1

Figure 4.3(a) shows the thresholded RPofRPs for Bach1. We can find many diagonal lines on it. An enlarged view shown in Fig. 4.3(b) represents that the width between the two diagonal lines is 3 seconds, that is almost similar to the time to play one bar of the musical score (see Fig. 4.3(c)); in detail, the time within which one bar is played fluctuated according to the performance speed at every bar. In one bar, arpeggio is repeated twice in the same harmony, and harmony change per every bar constantly continues towards the end of the piece. These repeated patterns correspond to many diagonal lines. (For all of the thresholded short-term RPs to calculate the RPosRPs for Bach1, see AppendixA.)

4.2.3 Mozart1

Figure 4.4(a) shows the thresholded RPofRPs for Mozart1. Blue thin lines on the RP are partitions of the movements. We can see that the graphical patterns are different among the movements. There are two kinds of diagonal lines in Fig. 4.4(a): narrow (see Fig. 4.4(b)) and wide (see Fig. 4.4(c)). Figure 4.4(b) was the theme of this piece. The interval of diagonal lines is 3 seconds, which corresponds to about 2 bars. Whereas, Fig. 4.4(c) was the intersection areas between variation VI+VII and X. The interval of these thick lines corresponds to 8 bars. Every variation consists of 24 bars, so the three thick lines in Fig. 4.4(c) demonstrate the similarity between phrases beyond the small difference across the variations. In addition, by focusing on the square of Fig. 4.4(d), we can compare the similarity between the most gorgeous and fast variation XII with another variation. In the same way, by focusing on the square of Fig. 4.4(e), we can compare the similarity between the simplest theme and another variation. Figure 4.4(f) is a table whose numbers indicate the main note length for the rhythm of each variation: 4 means a quarter note, 8 means an eighth note, 16 means a 16th note, and 8tri means eighth triplets. The variations which have many dots in the square of Fig. 4.4(d) correspond to 16 or 8tri (colored cyan), while many dots in the square of Fig. 4.4(e) correspond to 4 or 8 (colored yellow), except for variation XI whose tempo was very slow and the 16th note duration was similar to a quarter note duration of another variations. (For all of the thresholded short-term RPs to calculate the RPosRPs for Mizart1, see AppendixB.)

4.2.4 Beethoven1

Figure 4.5(a) shows the thresholded RPofRPs for Beethoven1. Blue thin lines are partitions of the choruses or verses. We can see that graphical patterns are different

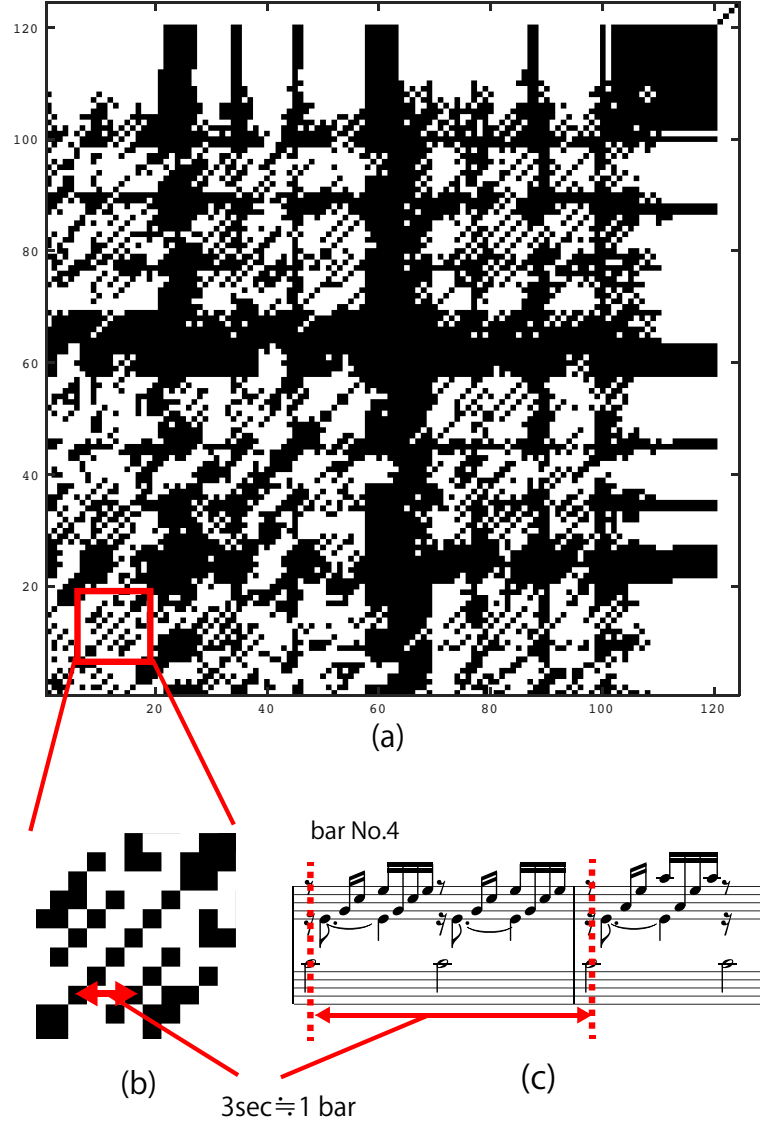


Figure 4.3. Thresholded RPofRPs for Bach1: (a) RPofRPs with $\varepsilon_r = 0.3$. (b) Enlarged view. (c) Corresponding score. The width between the two diagonal lines showed in (b) is 3 seconds, that is almost similar to the time to play one bar of the musical score (c). The unit of the axes is in seconds.

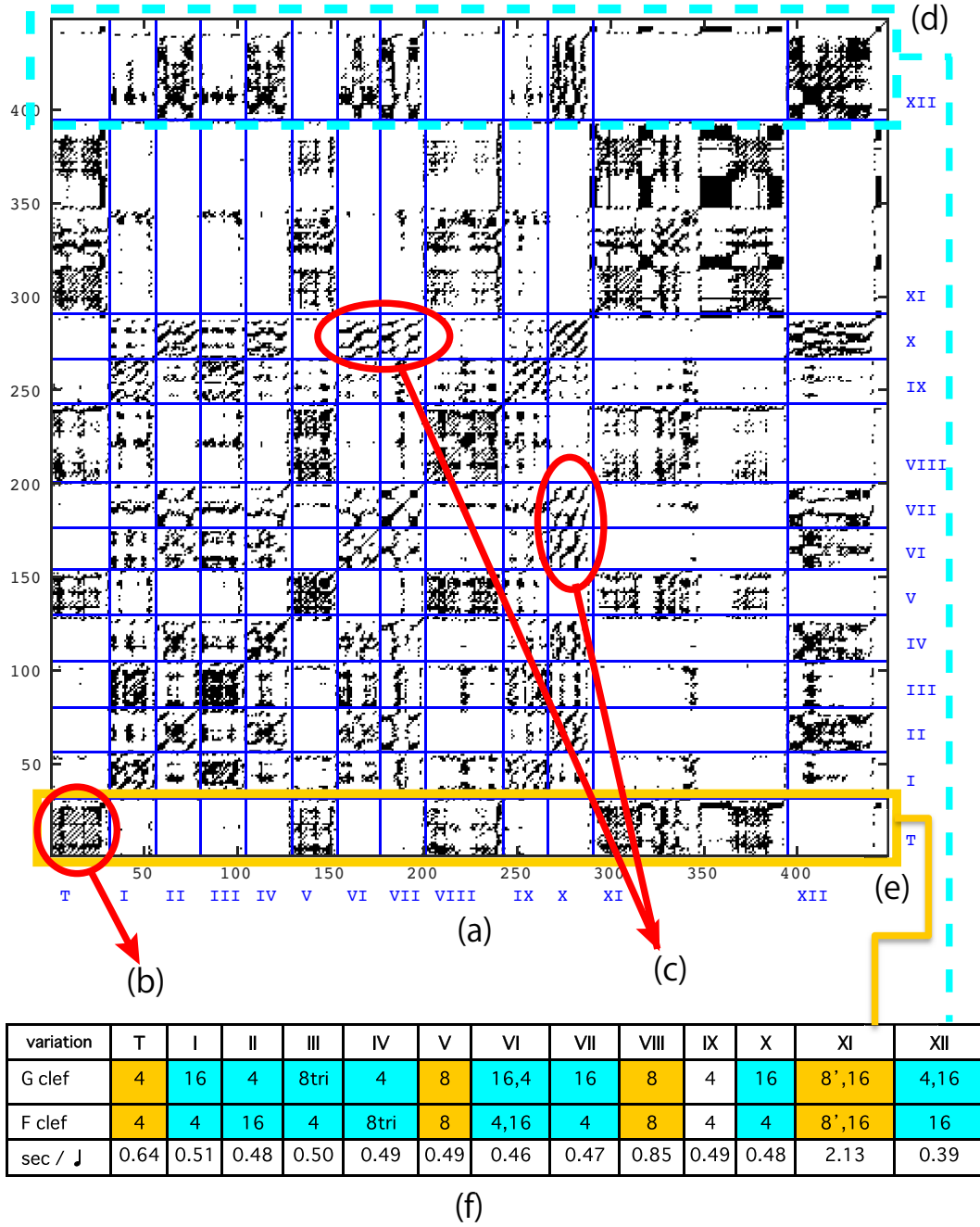


Figure 4.4. Thresholded RPoRPs for Mozart1: (a) RPoRPs with $\varepsilon_r = 0.1$. (b) The theme. (c) Intersection areas between variation VI+VII and X. (d) Intersection areas between variation XII and other variations. (e) Intersection areas between the theme and other variations. (f) The main beat of each variation. The unit of the axes is in seconds.

among the musical chunks, and there are a variety of diagonal lines. Yellow circles with the dotted line (Fig. 4.5(b1)) show the intersections between the part from bar No.23 to No.63 and the part from bar No.162 to No.202. We can see a long thick diagonal line. As shown in Fig. 4.5(b2) and Fig. 4.5(b3), they are similar choruses at the different keys by a major third over 40 bars. Thus, the RPofRPs can represent the similarity for similar melodies played in different keys. Moreover, the circle with the red line corresponds to a chorus where arpeggios and chords were played in fortissimo in succession (see Fig. 4.5(c)). Narrowing down the interval between diagonal lines represents that the tempo becomes fast. Abbreviations used in Fig. 4.5 are: CA: chorus A, CB: chorus B, CC: chorus C, CD: chorus D, VA: verse A, VB: verse B, VC: verse C, VD: verse D, and END: ending. (For all of the thresholded short-term RPs to calculate the RPosRPs for Beethoven1, see AppendixC.)

4.3 Determinism (DET)

Figure 4.6 shows the determinism (DET) of each music piece. Gray lines show DET of the short-term RPs. Red horizontal lines represent DET of the RPofRPs of the whole piece. Blue lines correspond to DET of each musical section (movements, chorus, or verse) of RPofRPs. Note that the gray lines which are the same as the DET of the conventional recurrence plots and the blue lines (the DET by our method) look completely different. In addition, the blue lines tend to show high scores at the first chorus, then vary largely at the middle, and gradually increase towards the end. We calculated the correlation between a linear line going up towards the right end and the DET scores without the first chorus. They had positive correlations in significance level 0.0001. The correlation coefficients were 0.41 for Mozart1, and 0.55 for Beethoven1.

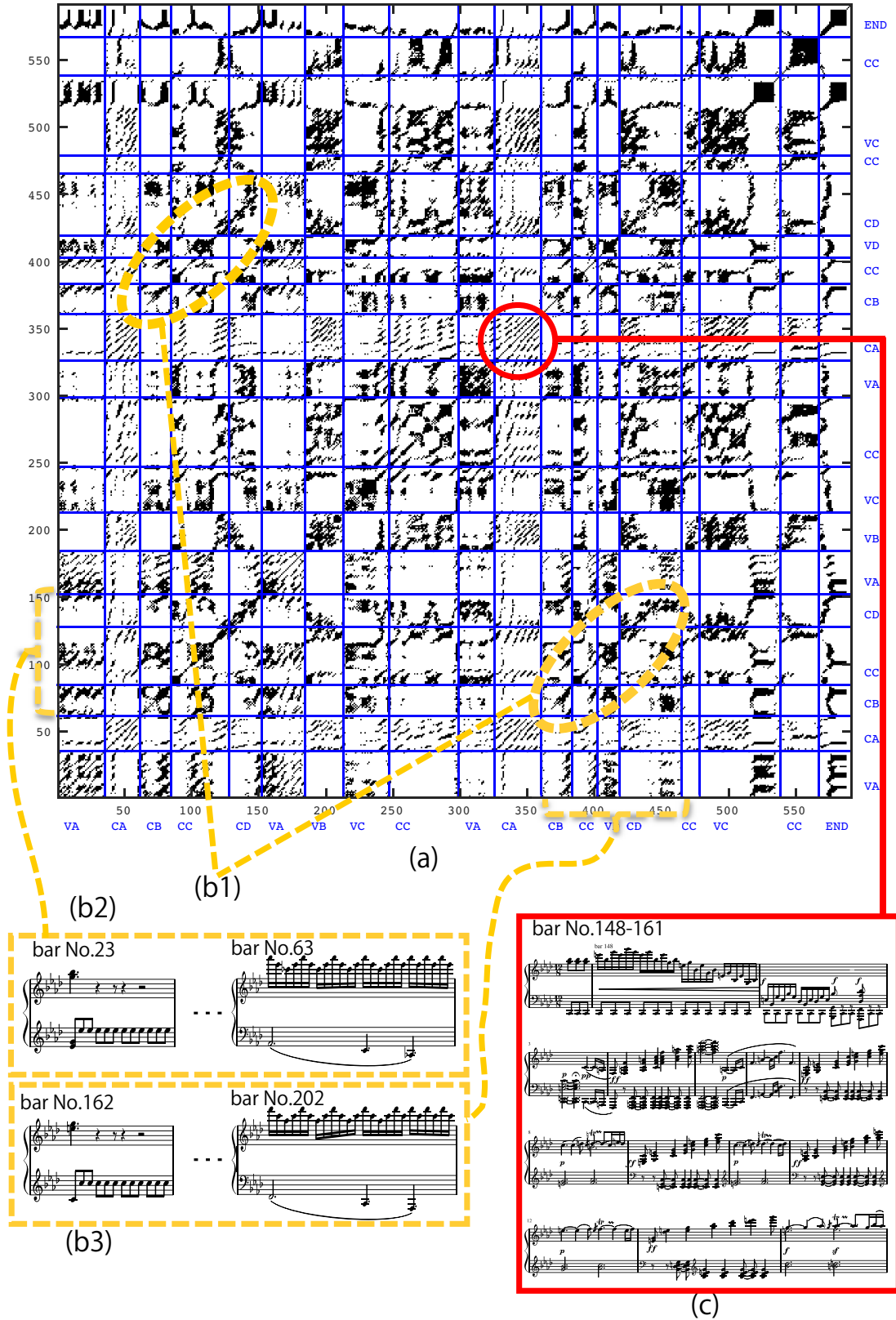


Figure 4.5. Thresholded RPOfRPs for Beethoven1: (a) RPOfRPs with $\varepsilon_r = 0.1$. (b1) Intersections between the part from bar No.23 to No.63 and the part from bar No. 162 to 202. (b2) The score of the part from bar No.23 to No.63. (b3) The score of the part from bar No.162 to No.202. (c) The score of the part from bar No.148 to No.161. The unit of the axes is in seconds.

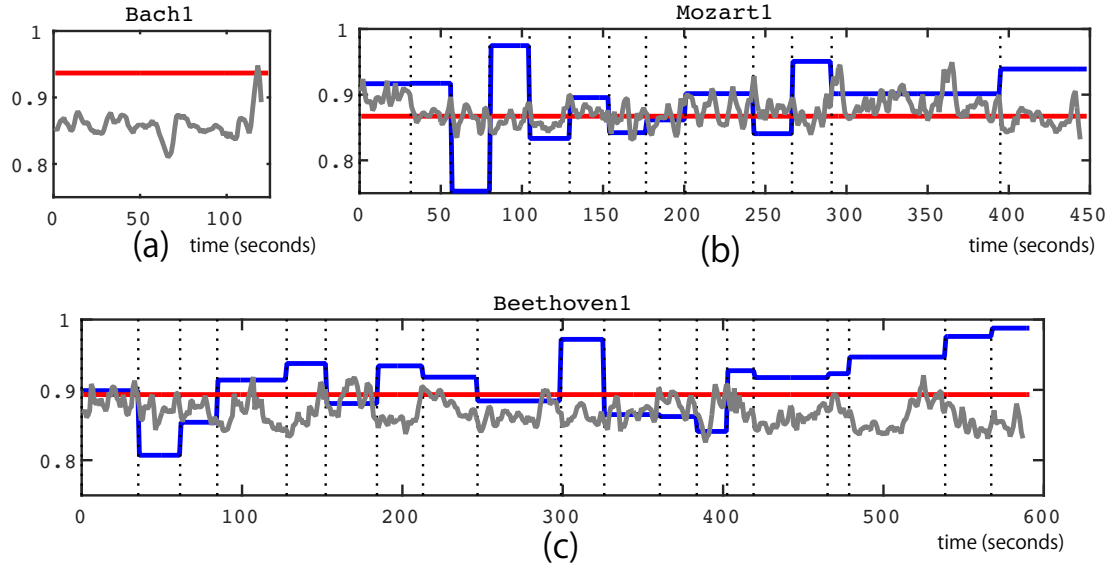


Figure 4.6. Determinism (*DET*): *DET*s of (a)Bach1, (b)Mozart1 and (c) Beethoven1. The gray lines show the short-term RP, the red lines show the whole RPofRPs, and the blue lines show sectioned RPofRPs.

Chapter 5

Discussion

5.1 Extracting a Perturbation

We analyze an additional toy model data to show a characteristic of our method more clearly. The dataset is generated by the Lorenz model (see Eqs. 3.6, 3.7 and 3.8) with the parameter $\sigma = 10$ and $\beta = 8/3$. With these parameters, the trajectory shows the strange attractor during the parameter r exceeds the Hopf bifurcation point $r = 24.748$. We set the parameter r driven by the sine wave as follows:

$$r = 5 \times \sin(t\pi/20) + 29.74, \quad (5.1)$$

where t is the time (See Fig. 5.1).

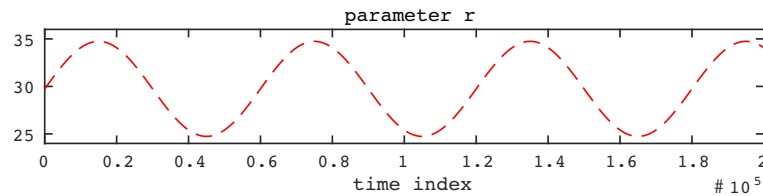


Figure 5.1. Parameter r driven by sine wave

Here the r represents a perturbation for the system. We integrate the above equations from $t = 0$ to 200 at 0.001 intervals. The initial values are set to $[x, y, z] = [0.0, 1.0, 1.0]$. The window size of the short-term RP is set to 4000, and the shift width is 1000, while the delay time is set to 1, and the dimension is set to 3. Figure 5.2 and 5.3 show the time series of x , y , and z of the original and perturbed Lorenz model, respectively. And Fig. 5.4 and 5.5 show attractors of the original and perturbed Lorenz model, respectively.

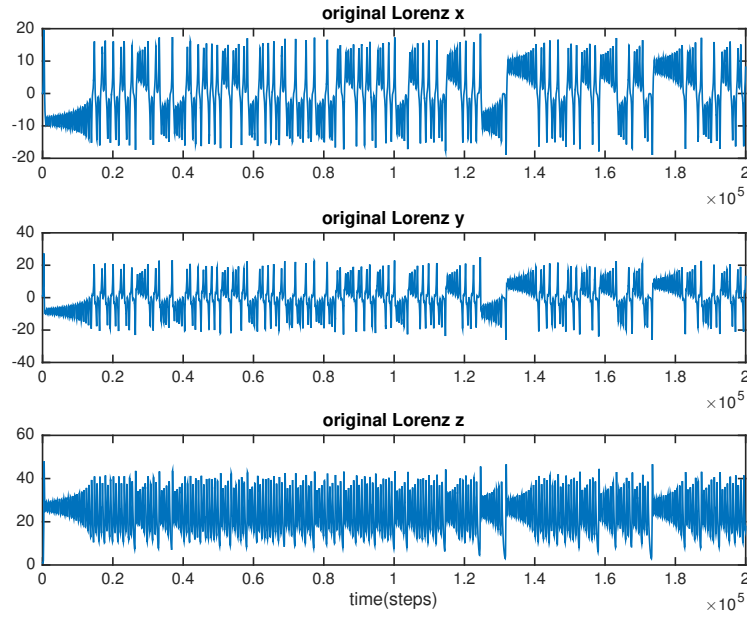


Figure 5.2. Time series of x , y , and z of the original Lorenz model. We set $\delta = 10$, $r = 28$, and $\beta = 8/3$.

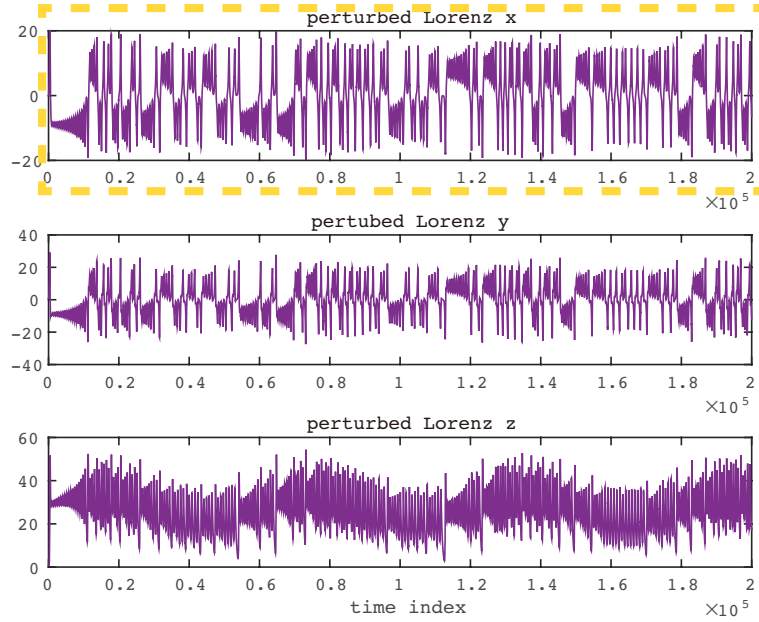


Figure 5.3. Time series of x , y , and z of the perturbed Lorenz model. We set $\delta = 10$, r is perturbed by Eq. 5.1, and $\beta = 8/3$. We analyze the time series x boxed by dotted yellow line.

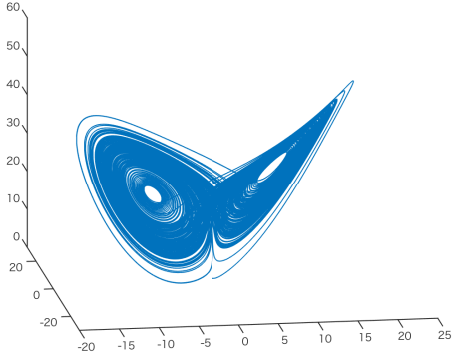


Figure 5.4. Attractor of the original Lorenz model. We set $\delta = 10$, $r = 28$, and $\beta = 8/3$.

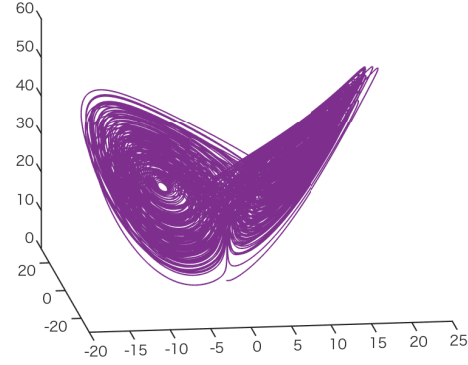


Figure 5.5. Attractor of the perturbed Lorenz model. We set $\delta = 10$, r is perturbed by Eq. 5.1, and $\beta = 8/3$.

To emulate the real world situation, we use only variable x data for the analysis. It is difficult to notice that the time series is influenced by sine wave just by looking at it alone. Figure 5.6 shows the input data of the variable x and its calculated RPofRPs. We obtain a synthesized time series from the RPofRPs by applying the method of reproducing original time series from recurrence plots [20].

Figure 5.7 shows the time series r and the synthesized time series. The dotted red line shows r and the blue line shows the synthesized time series. The mean, the standard deviation and, the orientation of the synthesized time series were adjusted so that these elements match with those of the original r . We can observe an apparent influence from r to the synthesized time series. They have correlation with the correlation coefficient 0.41 at the significance level less than 0.0001.

Thus, RPofRPs can extract slow perturbation from long time series if the slow perturbation exists, and express global regularity due to local regularity, even if we have only a part of the data. While the meta-recurrence plots (MRPs) [4] extract the slow perturbation of the long time series by using sparse information (only 0 and 1) to obtain a global view, our method keeps gradual information in the RPofRPs. Therefore, RPofRPs may be easier to use for further applications. Figure 5.8 is the block diagram of precessing steps described in this section.

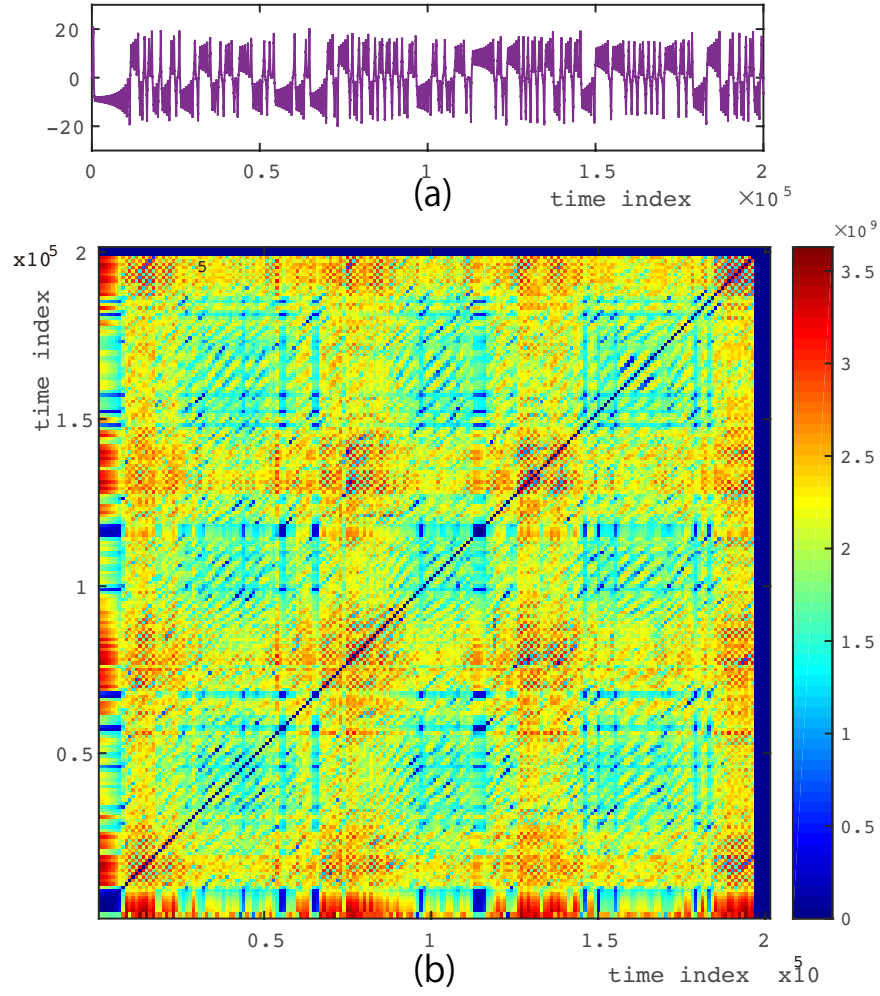


Figure 5.6. RPOFRPs of perturbed Lorenz. (a) Input data (Lorenz variable x , where r is driven by sine wave). (b) RPOFRPs.

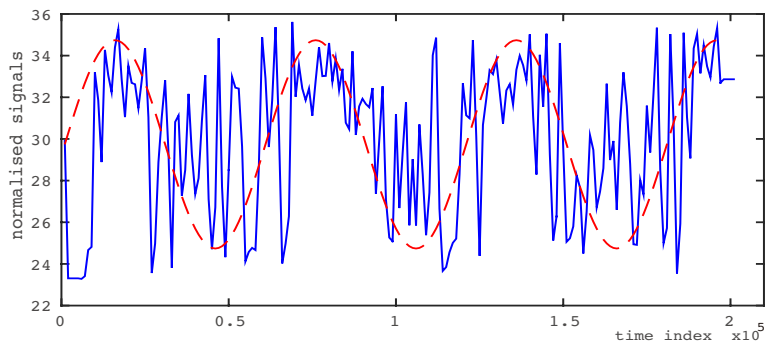


Figure 5.7. Time series r and the synthesized time series. The dotted red line shows r and the blue line shows the synthesized time series.

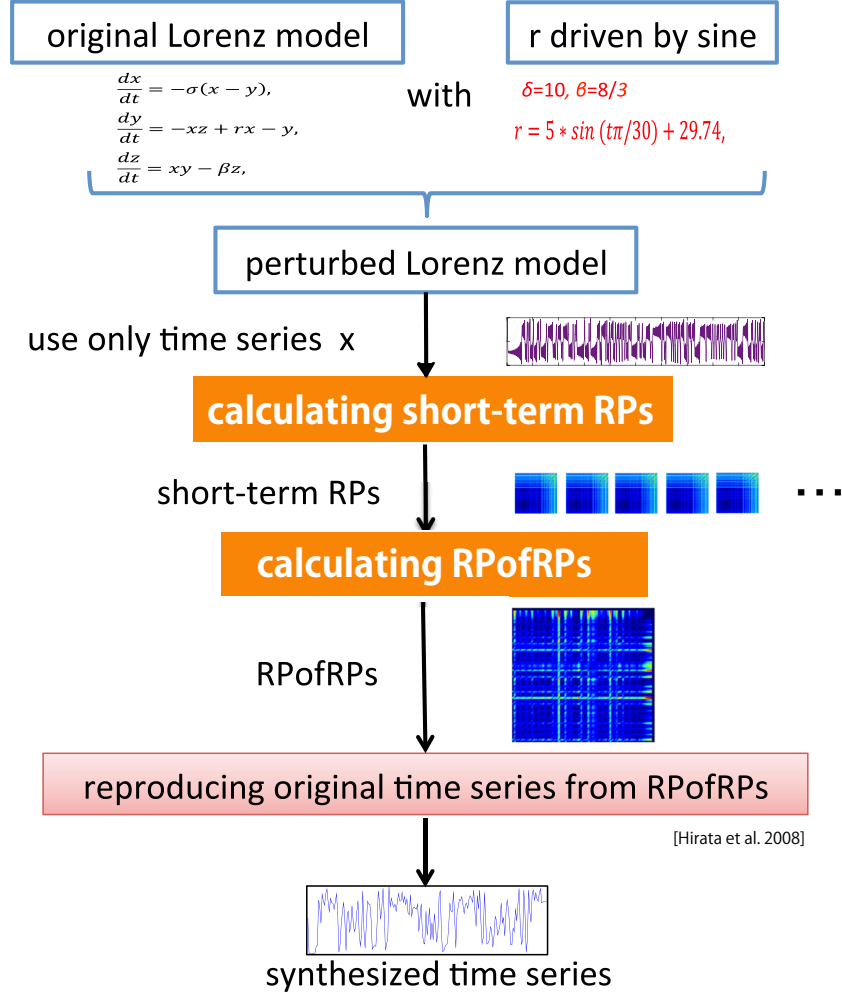


Figure 5.8. Block diagram of processing steps

5.2 Distance Measure

The distance measure to calculate the short-term RP and the RPofRPs might have room for further improvements. Instead of the Euclidean distance which we employed (see Eqs. 3.2 and 3.4), for example, we may use the Canberra metric [5], which is an L1 family and known as very sensitive around zero, for reducing the strong influence of the amplitude for the waveform and/or the unthresholded short-term RPs.

We compare the difference between the distance measures using unthresholded RPofRPs of Beethoven1 with different distance measures. While we use Euclidean distance for short-term RPs d_s (Eq. 3.2) and for RPofRPs d_r (Eq. 3.4), we use the

L1 distance (Eq. 5.2) and the Canberra distance (Eq. 5.3) additionally.

The L1 distance for d_s is defined as

$$d_{s1}(S(t_1), S(t_2)) = \sum_{k=0}^{D-1} |s(t_1 + k\tau) - s(t_2 + k\tau)|, \quad (5.2)$$

and the Canberra distance for d_s is defined as follows:

$$d_{sc}(S(t_1), S(t_2)) = \sum_{k=0}^{D-1} \frac{|s(t_1 + k\tau) - s(t_2 + k\tau)|}{s(t_1 + k\tau) + s(t_2 + k\tau)}. \quad (5.3)$$

Then the L1 distance for d_r is defined as

$$d_{r1}(X'_n, X'_m) = \sum_{i,j=1}^{N_x} |X'_n(i, j) - X'_m(i, j)|, \quad (5.4)$$

and the Canberra distance for d_r is defined as follows:

$$d_{rc}(X'_n, X'_m) = \sum_{i,j=1}^{N_x} \frac{|X'_n(i, j) - X'_m(i, j)|}{X'_n(i, j) + X'_m(i, j)}. \quad (5.5)$$

Figure 5.9 employs d_{s1} and d_{r1} , and Fig. 5.10 employs d_{sc} and d_{rc} . Figure 5.11 employs d_{s1} and d_{rc} .

We can find that Fig. 5.9 and Fig. 4.5(a) are very similar. Figure 5.10 indicates that employing the Canberra distance for both layers is not appropriate. However the Canberra distance is effective when it is used only for RPofRPs (see Fig. 5.11), and seems almost similar to Fig. 4.5 but partially different.

Tables 5.1, 5.2 and 5.3 show contingency tables among the distance measures. Table 5.1 is the Euclidean (Fig. 4.5) versus the L1 distance (Fig. 5.9), Table 5.2 is the Euclidean (Fig. 4.5) versus L1&Canberra distance (Fig. 5.11), and Table 5.3 is the L1 (Fig. 5.9) versus L1&Canberra distance (Fig. 5.11). We calculated the rate of same plots among two recurrence plots as accuracy, that is, sum of both zero and both one are divided by total.

The Euclidean distance and L1 distance are highly similar according to the Table 5.1. We use the L1 distance instead of the Euclidean distance to reduce the computation time in this chapter. The computation time for L1 distance is from one

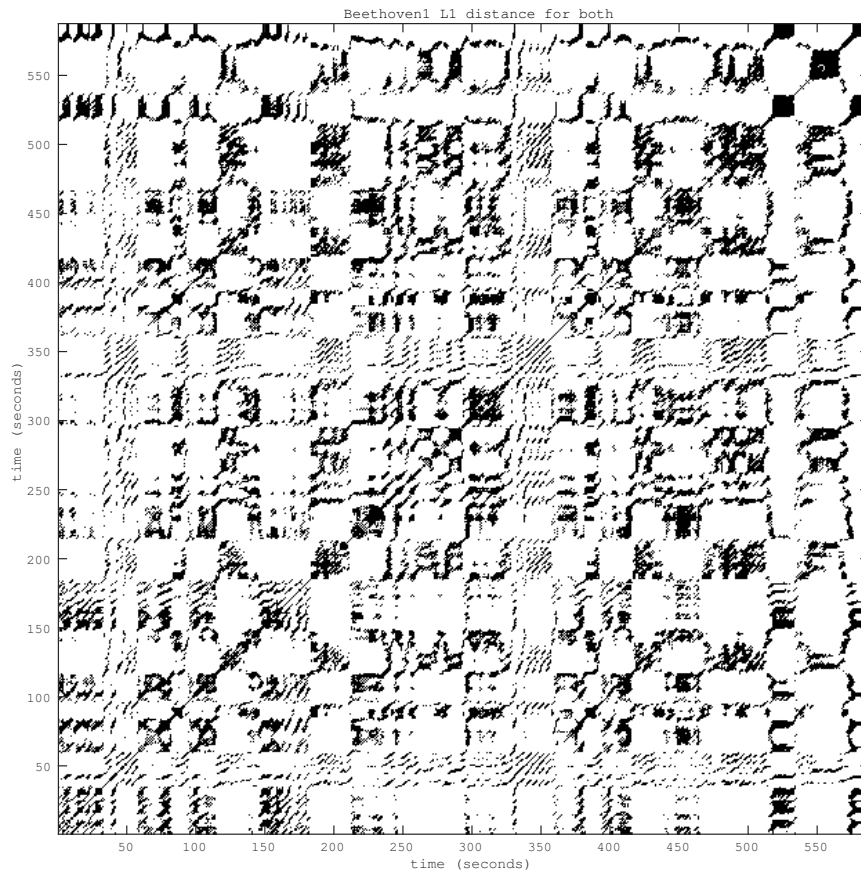


Figure 5.9. Unthresholded RPoFRPs of Beethoven1. We used the L1 distance for both layers.

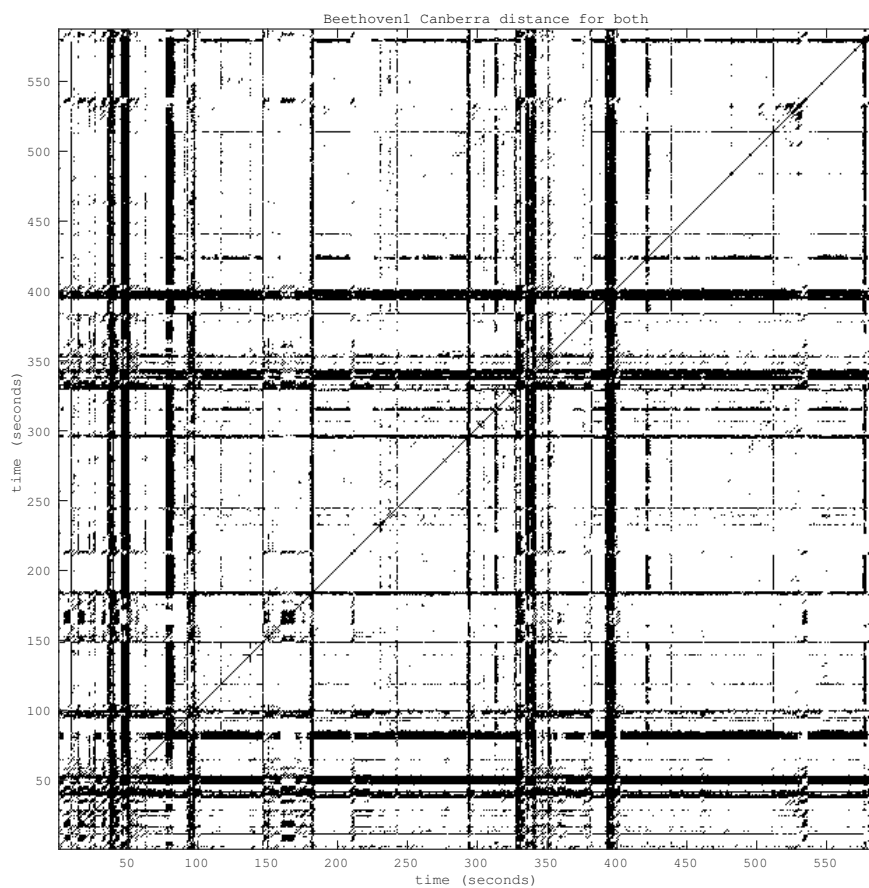


Figure 5.10. Unthresholded RPoFRPs of Beethoven1. We used the Canberra distance for both layers.

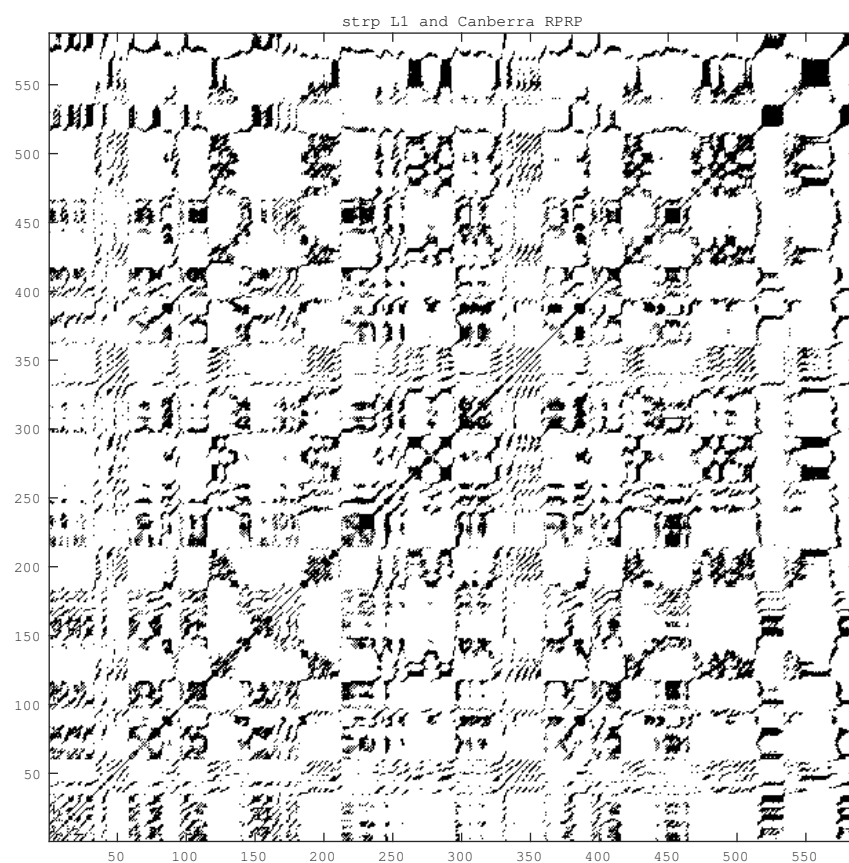


Figure 5.11. Unthresholded RPOfRPs of Beethoven1. We used the L1 distance for short-term RP and the Canberra distance for RPOfRPs.

seventh to one fifteen of the Euclidean on the machine with Xeon E5-2680 (2.70GHz, 8 cores (2 logical cores per physical), 64GB memory).

Table 5.1. Contingency table: Euclidean vs L1 distance

L1 \ Euc	0	1
0	282004 (81.8%)	4942 (1.2%)
1	4238 (1.4%)	53385 (15.5%)
accuracy		97.3%

Table 5.2. Contingency table: Euclidean vs L1&Canberra distance

Mix \ Euc	0	1
0	279098 (81.0%)	7848 (4.3%)
1	14914 (2.3%)	42709 (12.4%)
accuracy		93.4%

Table 5.3. Contingency table: L1 vs L1&Canberra distance

Mix \ L1	0	1
0	276650 (80.0%)	9592 (5.6%)
1	17362 (2.8%)	40965 (11.9%)
accuracy		92.2%

The appropriate distance measure might be different for each property of the target time series. The relationship between music factors and the difference among the distance measures is an issue in the future. From another point of view, we may select special measures that can grasp, for example, the phase, entropy, or geometry, because such measures can represent more complex structural property of the time series naturally.

5.3 Jazz and Pops Examples

Figures 5.12 and 5.13 show the thresholded RPofRPs of an example pop tune. We use the distance measure L1 for both layer (d_{s1} and d_{r1}) for Fig. 5.12, and L1 (d_{s1}) and Canberra (d_{rc}) for Fig. 5.13. Threshold ε is set to 0.2 for both RPofRPs.

It is an amateur live pop tune that was played by drums, a chopper base, an electric guitar, an electric piano, two parts of a synthesizers, and a female vocal (See blue lines and texts in Fig. 5.13 for detail musical structure).

We can find many diagonal lines parallel to the main diagonal line on both of the figures. The lines look thinner and straighter than the classic piano pieces. In addition, the difference of graphical patterns among distance measures is much obvious than the classical piano pieces (see Figs. 4.3, 4.4, and 4.5). The most different pattern between Fig. 5.12 and Fig. 5.13 is chorus part. Red squares in Fig. 5.13 show the intersections of the chorus section. We can find similar graphical patterns on it. On the contrary, we cannot find such patterns on Fig. 5.12.

Next, Figs. 5.14 and 5.15 show the thresholded RPofRPs of an example jazz piece. We use the distance measure L1 for both layer (d_{s1} and d_{r1}) for Fig. 5.14, and L1 (d_{s1}) and Canberra (d_{rc}) for Fig. 5.15. Threshold ε is set to 0.1 for both RPofRPs.

It is a piano trio played by piano, base, and drums. First, a theme is played by three instruments (0 to 50 seconds), and next, each instrument plays improvisations in solo. Then the piano and drums play in call and response style, and finally the theme is played again by three instruments in the end (after 235 second). The piece was taken from the RWC Music Database (piece number: RWC-MDB-J-2001 No.16 Tr.1) [14].

Same as the pop tune, we can find many thinner and straighter lines than the classic piano pieces, and also the difference of graphical patterns among distance measures is much obvious than the classical piano pieces (see Figs. 4.3, 4.4, and 4.5). This may be because the change of musical factors, for example, timber and dynamic range of amplitude, are not so large as the classical piano pieces. Systematic analysis for many genre remains to be seen.

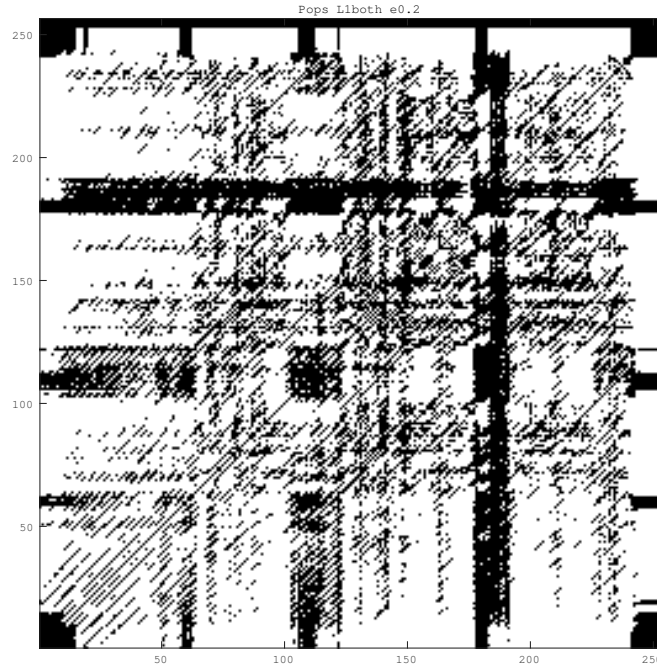


Figure 5.12. Example of pop tune. We used the distance measure L1 (d_{s1} and d_{r1}) and set $\varepsilon = 0.2$.

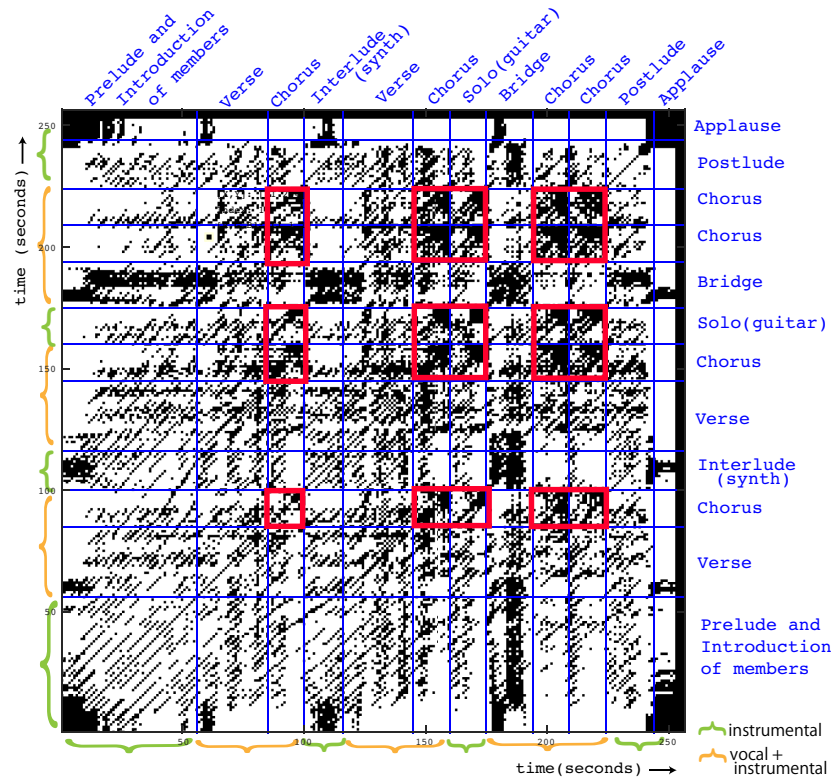


Figure 5.13. Example of pop tune. We used the distance measure L1 (d_{s1}) and Canberra (d_{rc}) and set $\varepsilon = 0.2$.

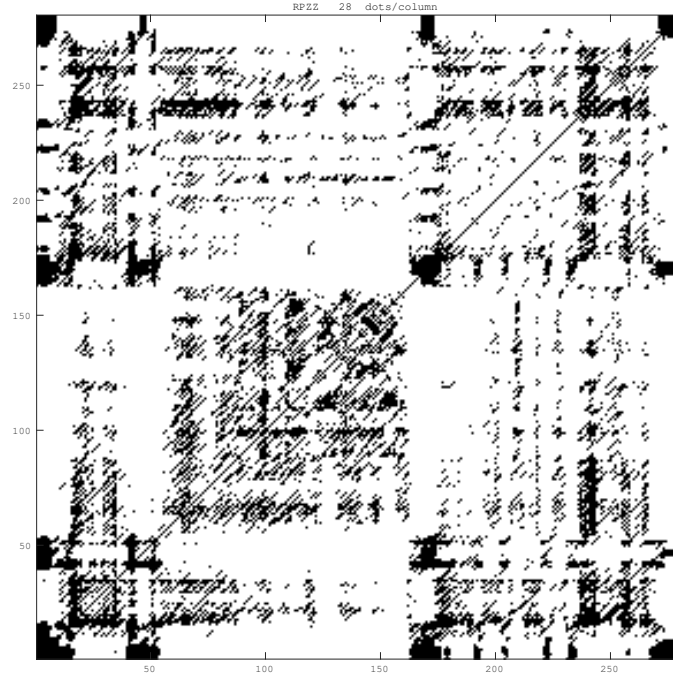


Figure 5.14. Example of jazz tune. We used the distance measure L1 (d_{s1} and d_{r1}) and set $\varepsilon = 0.1$.

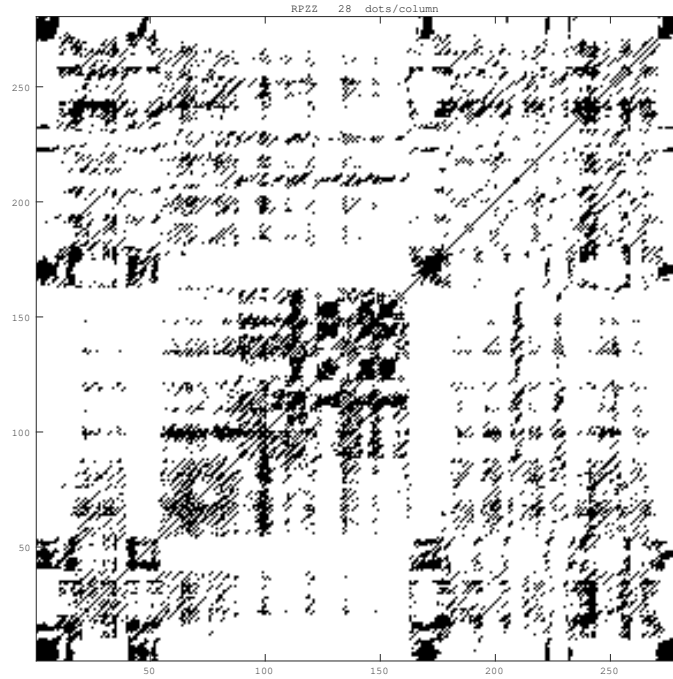


Figure 5.15. Example of jazz tune. We used the distance measure L1(d_{s1}) and Canberra (d_{rc}) and set $\varepsilon = 0.1$.

5.4 Future Work

Window and Shift Size

In this thesis, for the short-term RPs of the musical data, we set the window size to 4 seconds (44100 Hz x 4 seconds = 176,400 points), and the shift size to 1 second (44100 point). 4 seconds is almost equivalent to one bar in a common measure (four-four time), while the length of a quarter note is 1 second. We tried several values, and selected the appropriate values considering the visibility of the P_{Ro}fP_{Rs}. Varying the values may realize different observations of the data. It may be needed to consider the processing time, too. Larger the window size of short-term RPs, longer the total processing time.

Thresholds of the RPs

As mentioned in the Section 3.2, we selected the threshold value considering the visibility. If there are no diagonal lines or other patterns on the recurrence plots, the threshold value is not appropriate. Then, users need to try several values and inspect the corresponding recurrence plots visually.

Window and Shift Size, Threshold, and the DET

The window and shift size and the threshold affect the *DET* scores. For example, when the window size is set to larger and the shift size is set to shorter, the diagonal lines on the RPs tend to be longer, because the overlap between the next windows increases, many of the distance d_s and d_r tend to be shorter. As the result, the *DET* scores tend to be higher. When the threshold value is set to larger, the *DET* scores tend to be higher (e.g. if ε is 1, all of the RPs are plotted and the score is the maximum).

Curving Lines

By the definition of *DET* (see Eqs. 3.9 and 3.10), we counted only straight diagonal lines which are parallel to the central diagonal line, although a R_PofR_Ps shows many curves which are not parallel to the central diagonal line. We need a new measure which can count such non-diagonal but meaningful curves in the future. Besides *DET*, it may be meaningful to use other recurrence quantification analysis (RQA) measures [37, 25] and look at the correlations of the geometric patterns on R_PofR_Ps with various musical features such as tempo and dynamics.

RPofRPs and Network of Networks

If we consider purely theoretically, a recurrence plot can be regarded as a network [20] and thus the proposed RPofRPs might be regarded as the "network of networks" [10]. Therefore, one may view a RPofRPs as a new way of describing a network of networks, while one may use some measures of the "network of networks" for characterizing a RPofRPs. These directions of research contain many open questions.

Chapter 6

Conclusion

I described our proposed nonlinear time series method which uses the recurrence plots hierarchically. The proposed method enabled us to visualize both the local and global regularities at the same time, without corrupting the underlying regularities of the original time series. We applied this method to the music pieces, and showed that the method can represent the global musical features. Furthermore, we calculated the determinism (*DET*) scores of sectioned RPofRPs, and found that *DET*s tend to show high scores at the first chorus, then vary largely at the middle, and gradually increase towards the end. In addition, I showed that the RPofRPs allowed us to observe several aspects of musical time series by changing the distance measures. Moreover, I discussed that the transitions of *DET*s are consistent with the music theory, and might be useful for analyzing the emotional response to the music by selecting the appropriate window size and l_{min} value.

Because music evokes emotion, the structure of musical time series should be influenced by the processing mechanism of the brain consequently. Based on the neurobiological and mathematical studies, I raised the three key points: *regularity*, *hierarchy* and *scalability* to propose the new concept of musical time series analysis. In the proposed method, the *regularity* corresponds to the patterns on the recurrence plots and *DET* scores. The *hierarchy* corresponds to the hierarchical structure of the RPofRPs, of course. And the *scalability* corresponds to the flexibility of the distance scale, the window and shift size, and the l_{min} value.

Thus, the proposed method can comprehensively represent both local and global regularity of the underlying nonlinear dynamics behind the musical time series data—which is influenced by the brain mechanism—by using only simple calculations. I hope that the proposed method will be used for not only the extensive time series analysis that requires microscopic and macroscopic views simultaneously

but also neuroscientific and psychological researches.

Acknowledgments

I would like to express my deepest appreciation to Prof. Kazuyuki Aihara for giving me the very precious opportunity and environment to conduct my research and his invaluable supervision. Also, I am profoundly grateful to Prof. Yoshito Hirata for his insightful advices, fruitful discussions and supports. Moreover, I would like to thank Prof. Yoko Yamaguchi, Prof. Fumiyasu Komaki, and Prof. Tomonari Sei for kindly participating in the thesis committee and providing me helpful comments. I am very grateful to Prof. Yuichi Katori for his helpful discussions from the early stage of my works in the doctoral course. I am also grateful to Dr. Makito Oku for discussions and supports in various aspects. I would like to thank all the members and staffs in the Laboratories for Mathematics, Lifesciences, and Informatics including many research projects of Prof. Aihara.

My sincere thanks go to Prof. Sadahisa Okada (Tokyo Denki University) who was my first supervisor. I learned a spirit of the acoustical system engineering from him and the laboratory members through a study on the acoustic analysis of the "shamisen". Sincere thanks also go to Prof. Kiyoshi Furukawa (Tokyo University of Fine Art), Prof. Kazuo Okanoya (JST ERATO Okanoya Emotional Information Project / The University of Tokyo) and their joint research team of musical emotion, for invaluable discussions on music and emotion at the pre-stage of my doctoral studies. In addition, sincere thanks go to the members of Special Interest Group on Music and Computer (SIGMUS, IPSJ) for discussions on music and computer science for more than 20 years.

I would like to thank the bosses and colleagues of Panasonic for their kind understanding and help. I am deeply grateful to Dr. Tsutomu Muraji, Dr. Hiroshi Kutsumi and Dr. Kazuo Kajimoto for giving me the opportunity and time to study and their generous supports.

Finally, I want to express my gratitude to my family, especially to my partner Shuji Kurokawa. Conversations with him about / via music and technology inspired me very much, and bentos he cooked encouraged me enormously.

Bibliography

- [1] Bello, J. P. (2011). Measuring structural similarity in music. *IEEE Transactions on Audio, Speech and Language Processing*, 19(7), 2013-2025.
- [2] Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, 2(4), 382-7.
- [3] Burunat, I., Alluri, V., Toiviainen, P., Numminen, J. & Brattico, E. (2014). Dynamics of brain activity underlying working memory for music in a naturalistic condition. *Cortex*, 57(M), 254-69.
- [4] Casdagli, M. C. (1997). Recurrence plots revisited. *Physica D*, 108(1-2), 12-44.
- [5] Cha, S. -H. (2007). Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences*, 1(4), 300-307.
- [6] Chelidze, D. (2014). Statistical characterization of nearest neighbors to reliably estimate minimum embedding dimension. *Proceedings of International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, Vol. 8.
- [7] Chew, E. (2000). *Towards a mathematical model of tonality*, unpublished Ph.D. thesis, MIT, Cambridge, MA.
- [8] Cooke, D. (1959). *The Language of Music*, Oxford University Press, London.
- [9] Damasio, A. (2000) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, Mariner Books, New York.
- [10] Donges, J. F., Schultz, H. C. H., Marwan, N., Zou, Y. & Kurths, J. (2011). Investigating the topology of interacting networks: Theory and application to coupled climate subnetworks. *European Physical Journal B*, 84(4), 635-651.

- [11] Eckmann, J.-P., Kamphorst, S. O. & Ruelle, D. (1987). Recurrence Plots of Dynamical Systems. *Europhysics Letters*, 4(9), 973-977.
- [12] Fukino, M., Hirata, Y. & Aihara, K. (2016). Coarse-graining time series data : Recurrence plot of recurrence plots and its application for music. *Chaos* 26(2), 023116 1-12.
- [13] Garrido, M. I., Kilner, J. M., Stephan, K. E. & Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clinical Neurophysiology* 120, 453-63.
- [14] Goto, M. (2004). Development of the RWC music database. *Proceedings of the 18th International Congress on Acoustics*, Vol. I, 553-556.
- [15] Groth, A. (2005). Visualization of coupling in time series by order recurrence plots. *Physical Review E* 72(4), 046220.
- [16] Hargreaves, D. J., MacDonald, R. A. R. & Miell, D. E. (2005). How do people communicate using music? In D. E. Miell, R. A. R. MacDonald & D. J. Hargreaves (Eds.), *Musical Communication* (pp. 1-25). Oxford University Press, UK.
- [17] Helmholtz, H. (1877). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, Dover Publications, New York.
- [18] Herholz, S. C., Boh, B. & Pantev, C. (2011). Musical training modulates encoding of higher-order regularities in the auditory cortex. *The European Journal of Neuroscience*, 34(3), 524-9.
- [19] Hevner, K. (1935). The affective value of the major and minor modes in music. *American Journal of Psychology*, 47, 103-118.
- [20] Hirata, Y., Horai, S. & Aihara, K. (2008). Reproduction of distance matrices and original time series from recurrence plots and their applications. *European Physical Journal : Special Topics*, 164(1), 13-22.
- [21] Hirata, Y. & Aihara, K. (2011). Statistical tests for serial dependence and laminarity on recurrence plots. *International Journal of Bifurcation and Chaos* 21(4), 1077-1084.
- [22] Hirata, Y. & Aihara, K. (2010). Devaney's chaos on recurrence plots. *Physical Review E* 82(3), 036209.
- [23] Hirata, Y., Lang, E. -J. & Aihara, K. (2014). Recurrence plots and the analysis of multiple spike trains. K.Nikola ed., *Springer Handbook of Bio-/Neuroinformatics* Springer, Berlin, pp.735-744.

- [24] Hirata, Y. & Aihara, K. (2015). Edit distance for marked point processes revisited: An implementation by binary integer programming. *Chaos*, 25(12), 123117.
- [25] Iwanski, J. S. & Bradley, E. (1998). Recurrence plots of experimental data: To embed or not to embed? *Chaos*, 8(4), 861-871.
- [26] Iwayama, K., Hirata, Y., Takahashi, K., Watanabe, K., Aihara, K. & Suzuki, H. (2012). Characterizing global evolutions of complex systems via intermediate network representations. *Scientific Reports*, 2, 1-5.
- [27] Juslin, P. N. & Sloboda, J. A. (2001). *Music and Emotion: Theory and Research*, Oxford University Press, Oxford.
- [28] Kantz, H. & Schreiber, T. (1997). *Nonlinear Time Series Analysis*, University Press, Cambridge.
- [29] Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends in Cognitive Sciences*, Vol.14, No.3, pp.131-137.
- [30] Koelsch, S., Rohrmeier, M., Torrecuso, R. & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), 15443-8.
- [31] Large, E. W. & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, 26(1), 1-37.
- [32] LeDoux, J. E. (2000). Emotion Circuits in the Brain. *Annual Review of Neuroscience* 23, 155-184.
- [33] Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J. & Stephan, K. E. (2013). Modelling trial-by-trial changes in the mismatch negativity. *PLoS Computational Biology*, 9(2), e1002911.
- [34] Lieder, F., Stephan, K. E., Daunizeau, J., Garrido, M. I. & Friston, K. J. (2013). A neurocomputational model of the mismatch negativity. *PLoS Computational Biology*, 9(11), e1003288.
- [35] Lorenz, E. N. (1963). Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20(2), 130-141.
- [36] Marwan, N., Thiel, M. & Nowaczyk, N. R. (2002). Cross recurrence plot based synchronization of time series. *Nonlinear Process in Geophysics*, 9, 325-31.
- [37] Marwan, N., Romano, M. C., Thiel, M. & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5-6), 237-329.

- [38] Meyer, L. B. (1956). *Emotion and Meaning in Music*, University of Chicago Press, Chicago.
- [39] Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P. & Winkler, I. (2001). “Primitive intelligence” in the auditory cortex. *Trends in Neuroscience* 24(5), 283-288.
- [40] Näätänen, R., Shiga, T., Asano, S. & Yabe, H. (2015). Mismatch negativity (MMN) deficiency: A break-through biomarker in predicting psychosis onset. *International Journal of Psychophysiology*, 95(3), 338-344.
- [41] Nagai, T., Tada, M., Kirihara, K., Araki, T., Jinde, S. & Kasai, K. (2013). Mismatch negativity as a “translatable” brain marker toward early intervention for psychosis: A review. *Frontiers in Psychiatry*, 4(9), 1-10.
- [42] Narmour, E. (1989). The “genetic code” of melody : Cognitive structures generated by the implication-realization model. *Contemporary Music Review*, 45-63.
- [43] North, A. C. & Hargreaves, D. J. (1995). Subjective Complexity, Familiarity, and Liking for Popular Music. *Psychomusicology*, 14(1-2), 77-93.
- [44] Ohmura, H., Shibayama, Y., Terasawa, H., Hoshi-Shiba, R., Kawakami, A., Fukino, M., Okanoya, K. & Furukawa, K. (2013). Review of musical emotion studies. *The Journal of the Acoustical Society of Japan*, 69(9), 467-478. [In Japanese]
- [45] Poincaré, H. (1890). Sur la probleme des trois corps et les équations de la dynamique. *Acta Mathematica* 13, 1-271.
- [46] Rao, R. P. & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79-87.
- [47] Robinson, G. & Thiel, M. (2009). Recurrences determine the dynamics. *Chaos* 19, 023104.
- [48] Rohrmeier, M. & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, 83, 164-175.
- [49] Romano, M. C., Thiel, M., Kurths, J. & Bloh, W. von (2004). Multivariate recurrence plots. *Physics Letters*, 330, 214-223.

- [50] Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A. & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, 14(2), 257-62.
- [51] Serra, J., Serra, X. & Andrzejak, R. G. (2009). Cross recurrence quantification for cover song identification. *New Journal of Physics*, 11 093017.
- [52] Shepard, R. N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review*, 89(4), 305-333.
- [53] Suzuki, S., Hirata, Y. & Aihara, K. (2010). Definition of distance for marked point process data and its application to recurrence plot-based analysis of exchange tick data of foreign currencies. *International Journal of Bifurcation and Chaos*, 20(11), 3699-3708.
- [54] Temperley, D. (2008). A probabilistic model of melody perception. *Cognitive Science*, 32(2), 418-444.
- [55] Ushakov, Y., Dubkov, a. & Spagnolo, B. (2011). Regularity of spike trains and harmony perception in a model of the auditory system. *Physical Review Letters*, 107(10), 2-5.
- [56] Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C. & Roepstorff, A. (2009). Predictive coding of music - Brain responses to rhythmic incongruity. *Cortex*, 45(1), 80-92.
- [57] Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L. & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 108(51), 20754-9.
- [58] Wacongne, C., Changeux, J. P. & Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *Journal of Neuroscience*, 32(11), 3665-3678.
- [59] Winkler, I. & Czigler, I. (2011). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology*, 83(2), 132-143.
- [60] Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T. & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420-2.

- [61] Zatorre, R. J. & Salimpoor, V. N. (2013). From perception to pleasure: music and its neural substrates. *Proceedings of the National Academy of Sciences of the United States of America*, 110(Suppl.2), 10430-7.
- [62] Zbilut J. P., Giuliani A. & Webber C. L. Jr. (1998). Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Physics Letter A*, 246, 122-8.

Appendix A

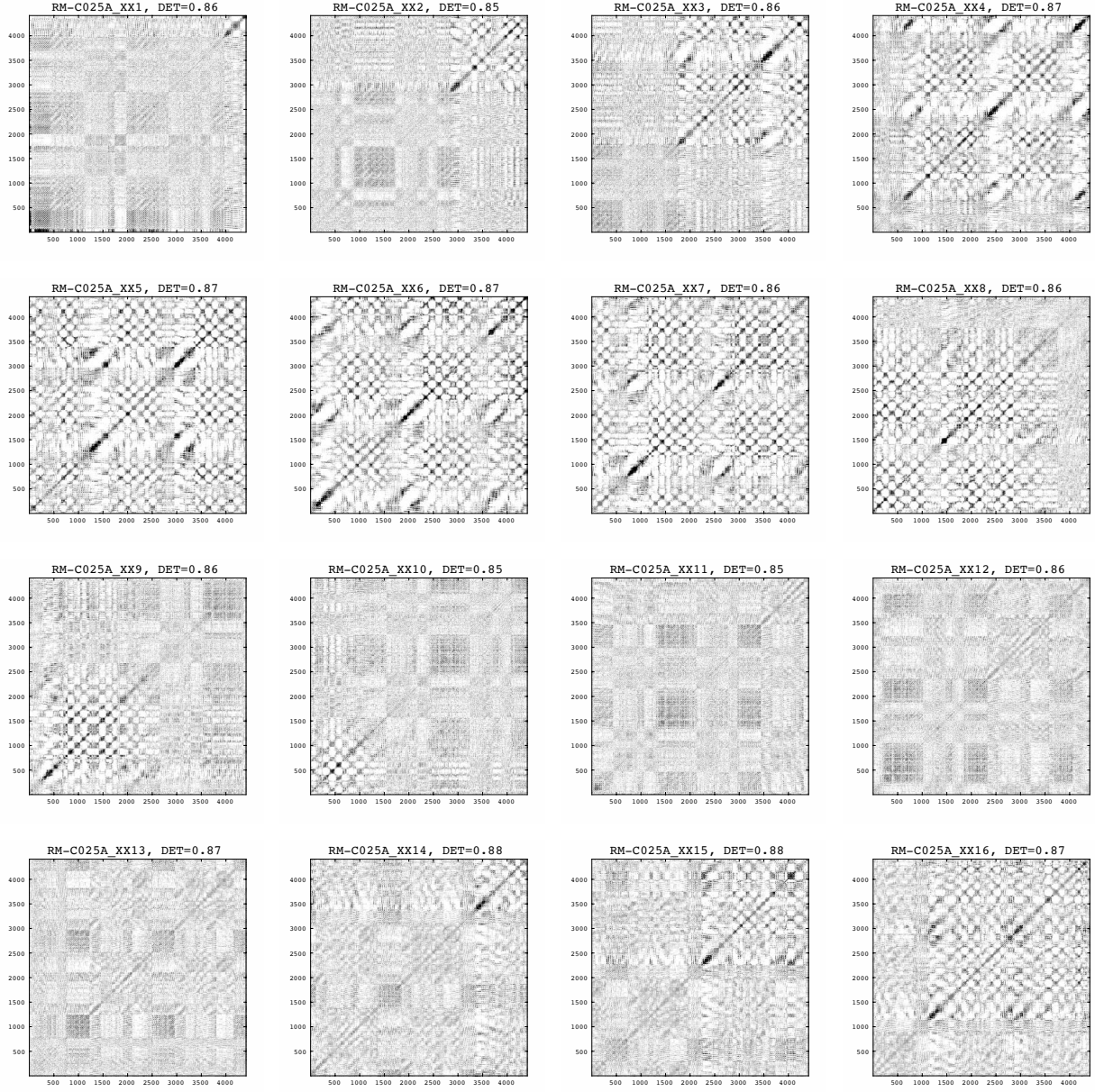
Thresholded short-term RPs of Bach1

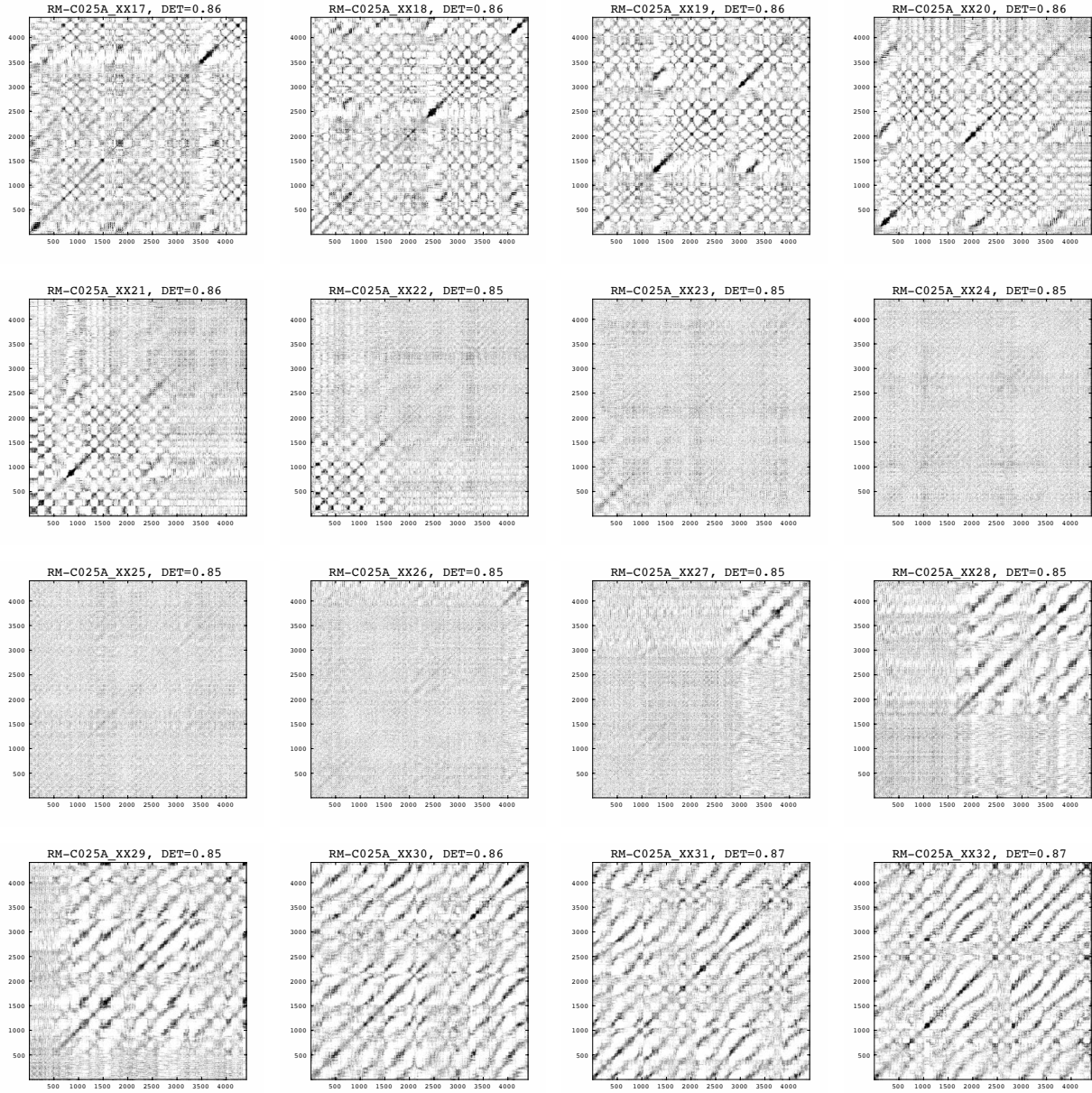
Here I show all thresholded short-term RPs of Bach1 and each corresponding *DET* scores. Parameters are as below.

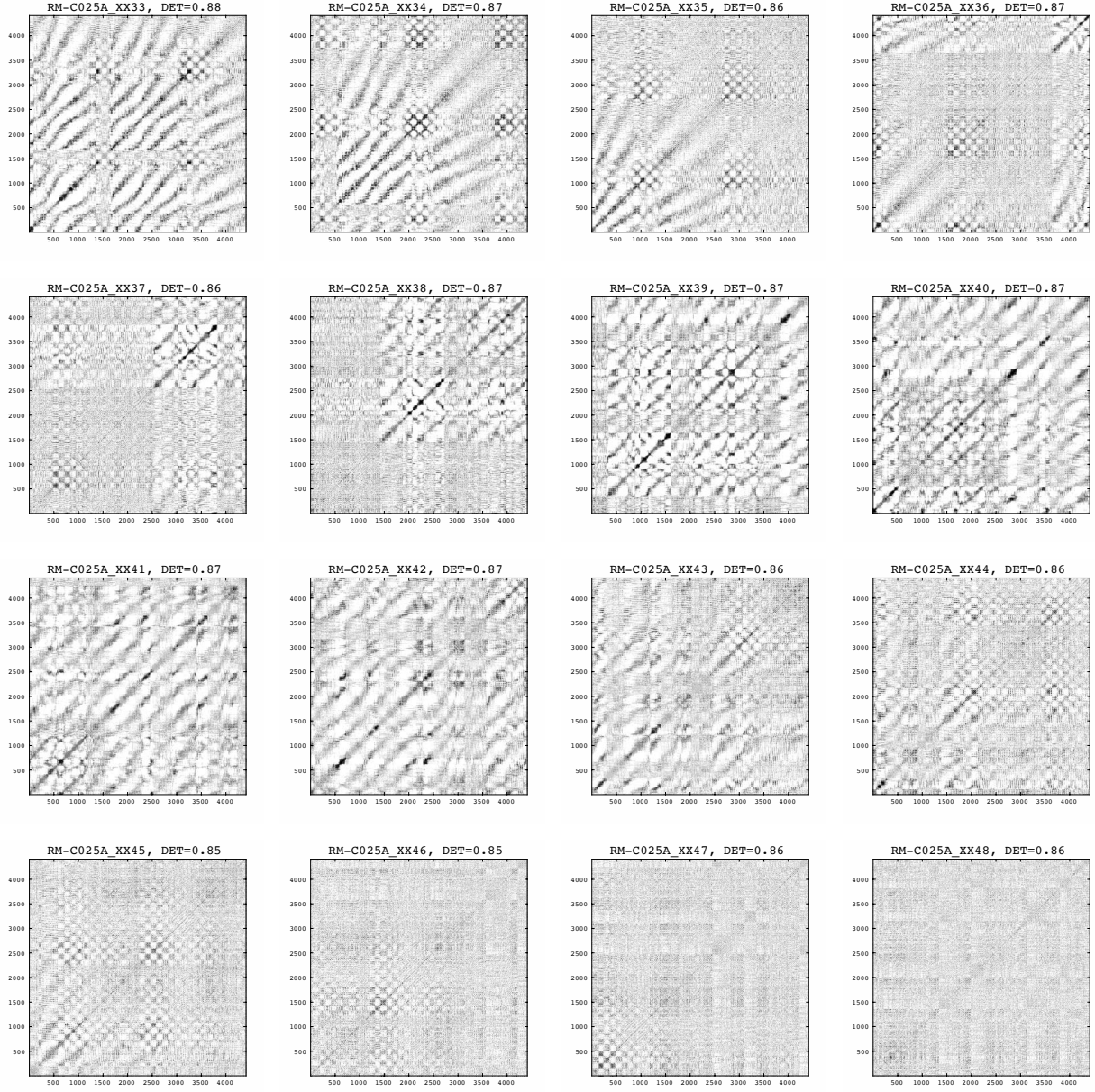
- window size: $W = 4$ seconds (4410)
- shift size: $h = 1$ second (1100)
- delay time: $\tau = 40$
- dimension: $D = 5$
- threshold: $\varepsilon_s = 0.1$
- the shortest diagonal line to be considered to calculate the histogram P : $l_{min} = 2$

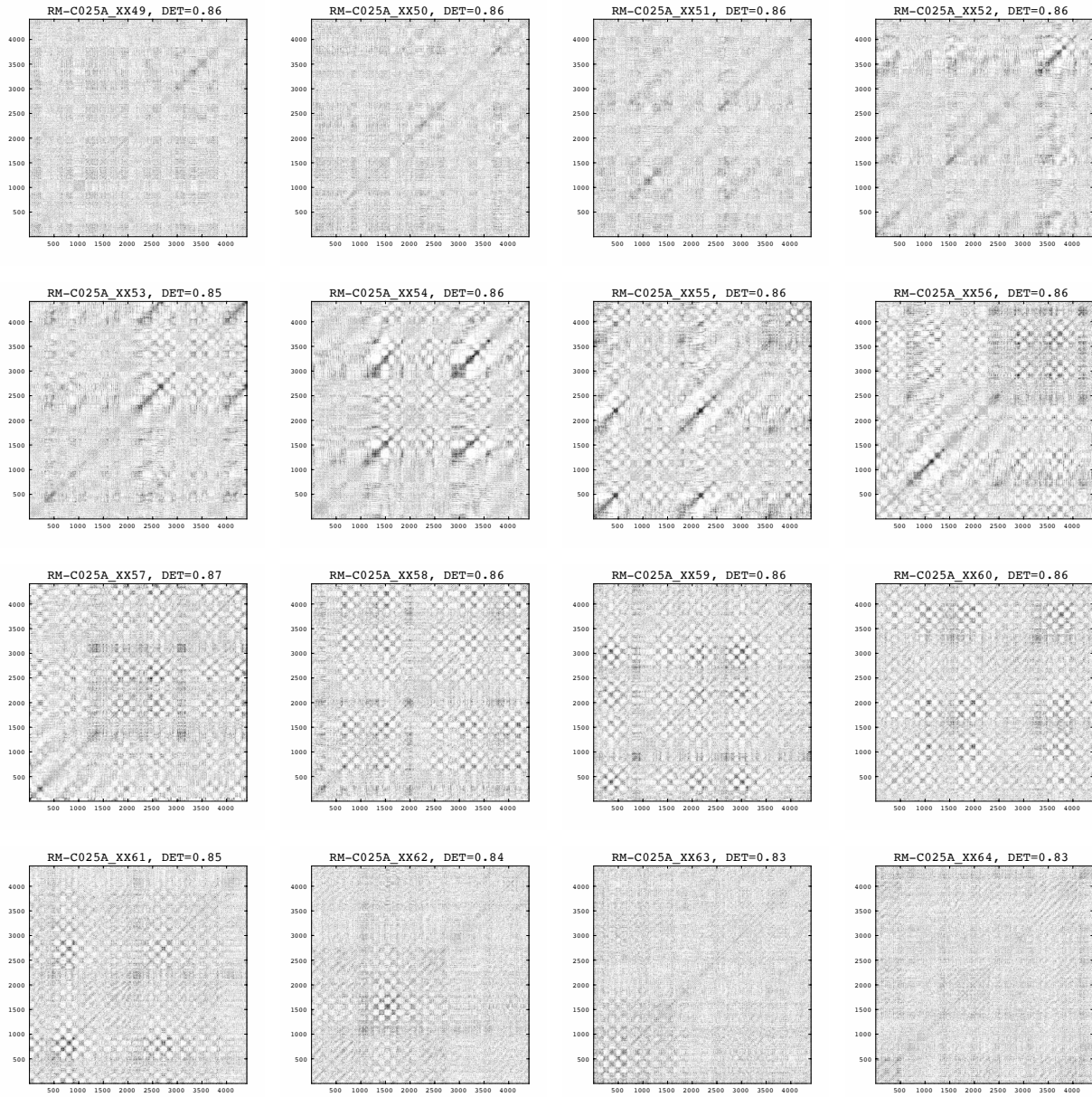
That is, the width of each short-term RP corresponds to 4 seconds, and one short-term RP is overlapped in 3 seconds with the next short-term RP.

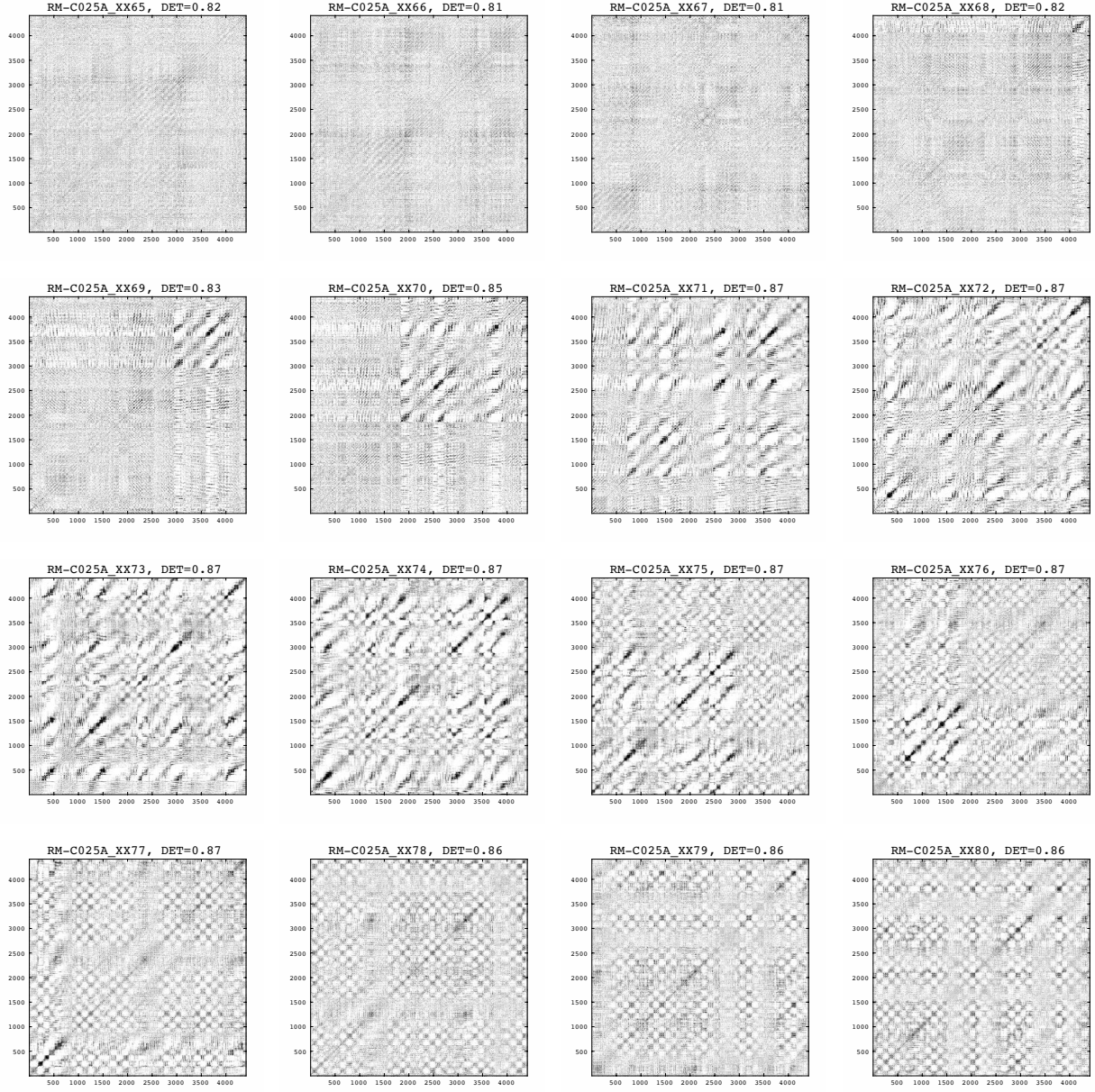
A serial number and *DET* score are shown at the top of each RPs. The number (ex. 1, 2... of RM-C025A_XX1, RM-C025A_XX2...) corresponds to the starting time of the time series for the RP in seconds.

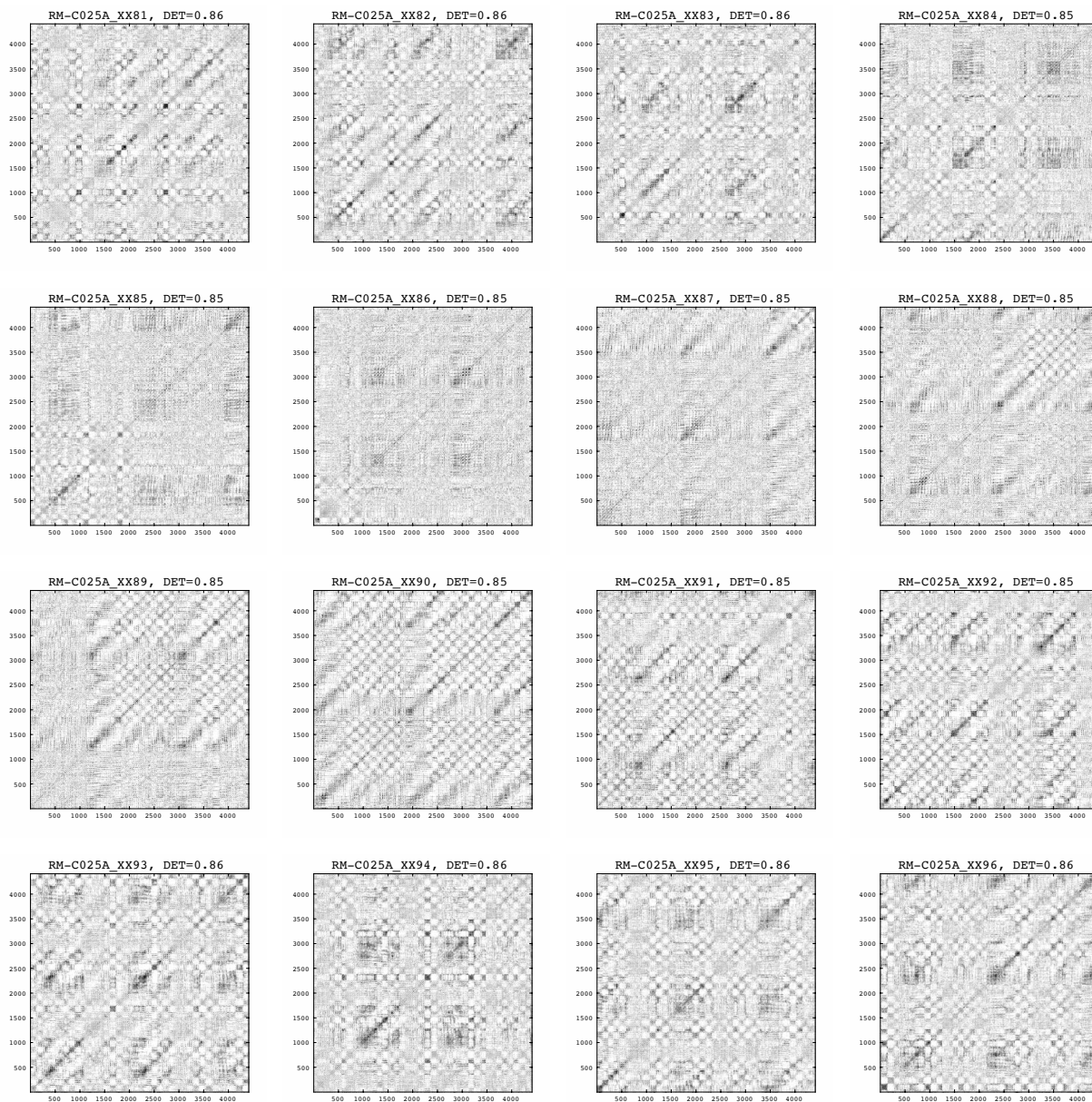




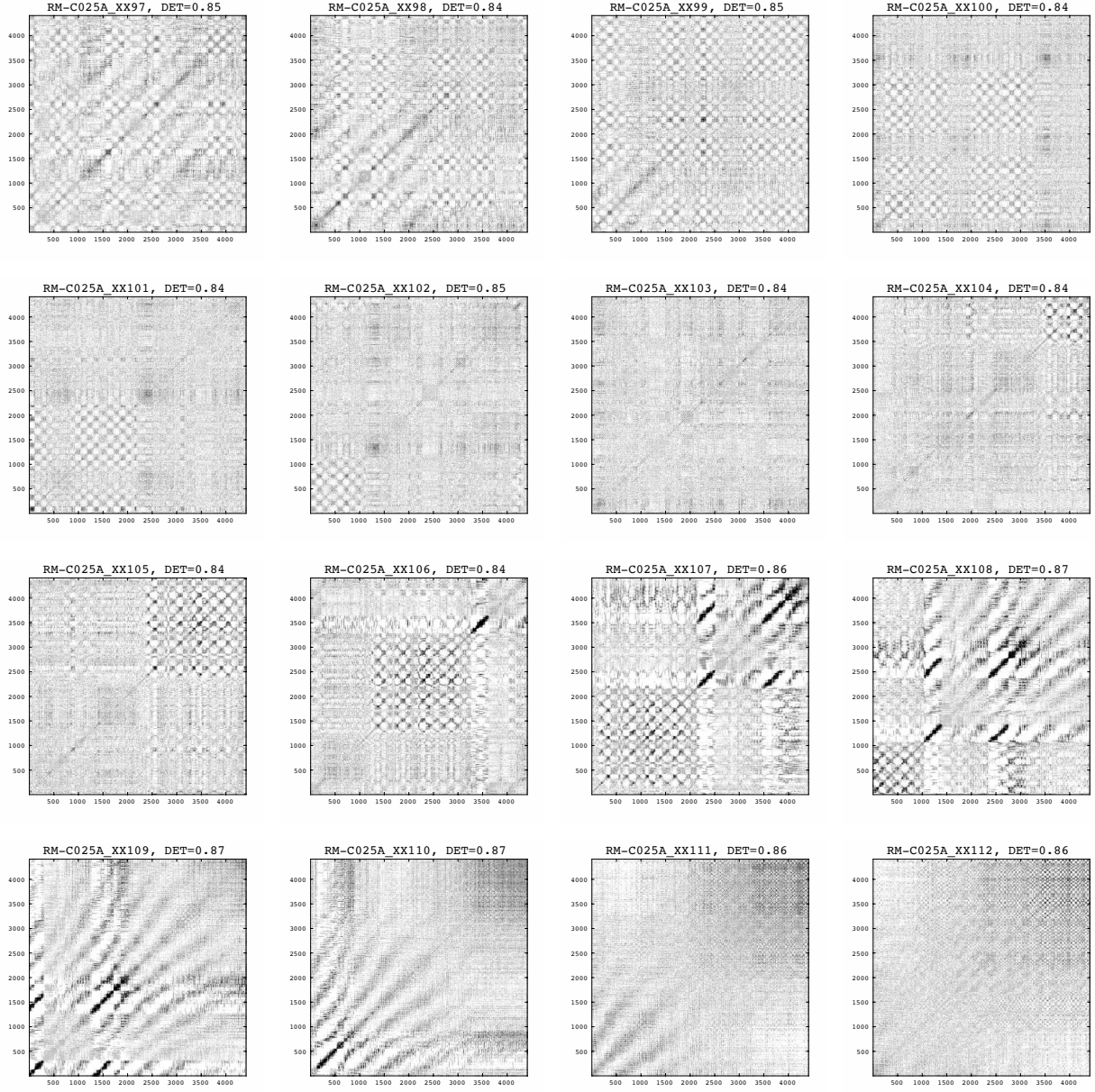


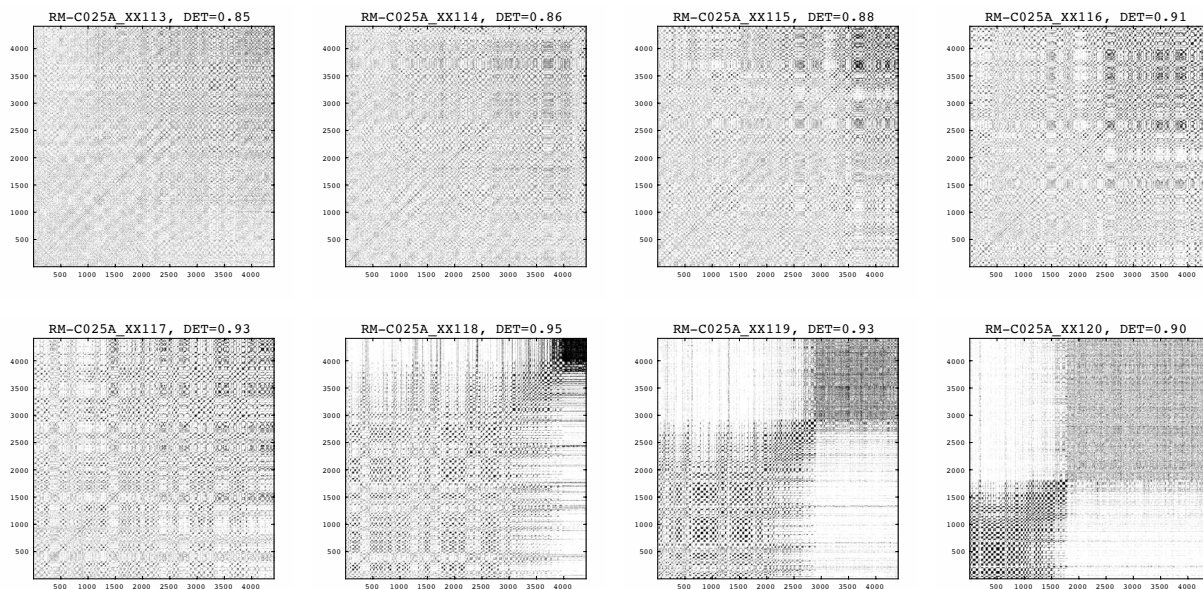






64 AppendixA Thresholded short-term RPs of Bach1





Appendix B

Thresholded short-term RPs of Mozart1

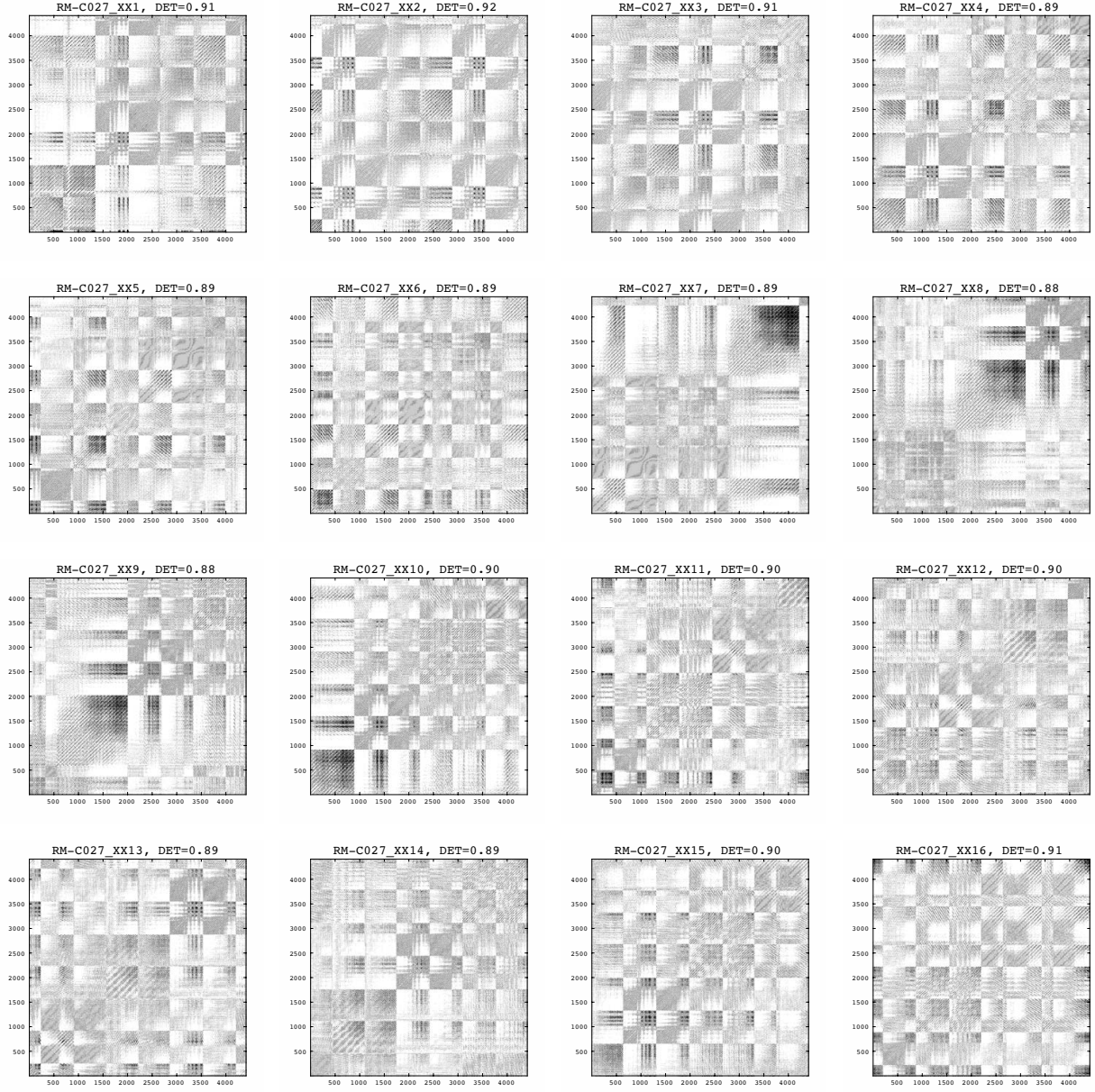
Here I show all thresholded short-term RPs of Mozart1 and each corresponding *DET* scores.

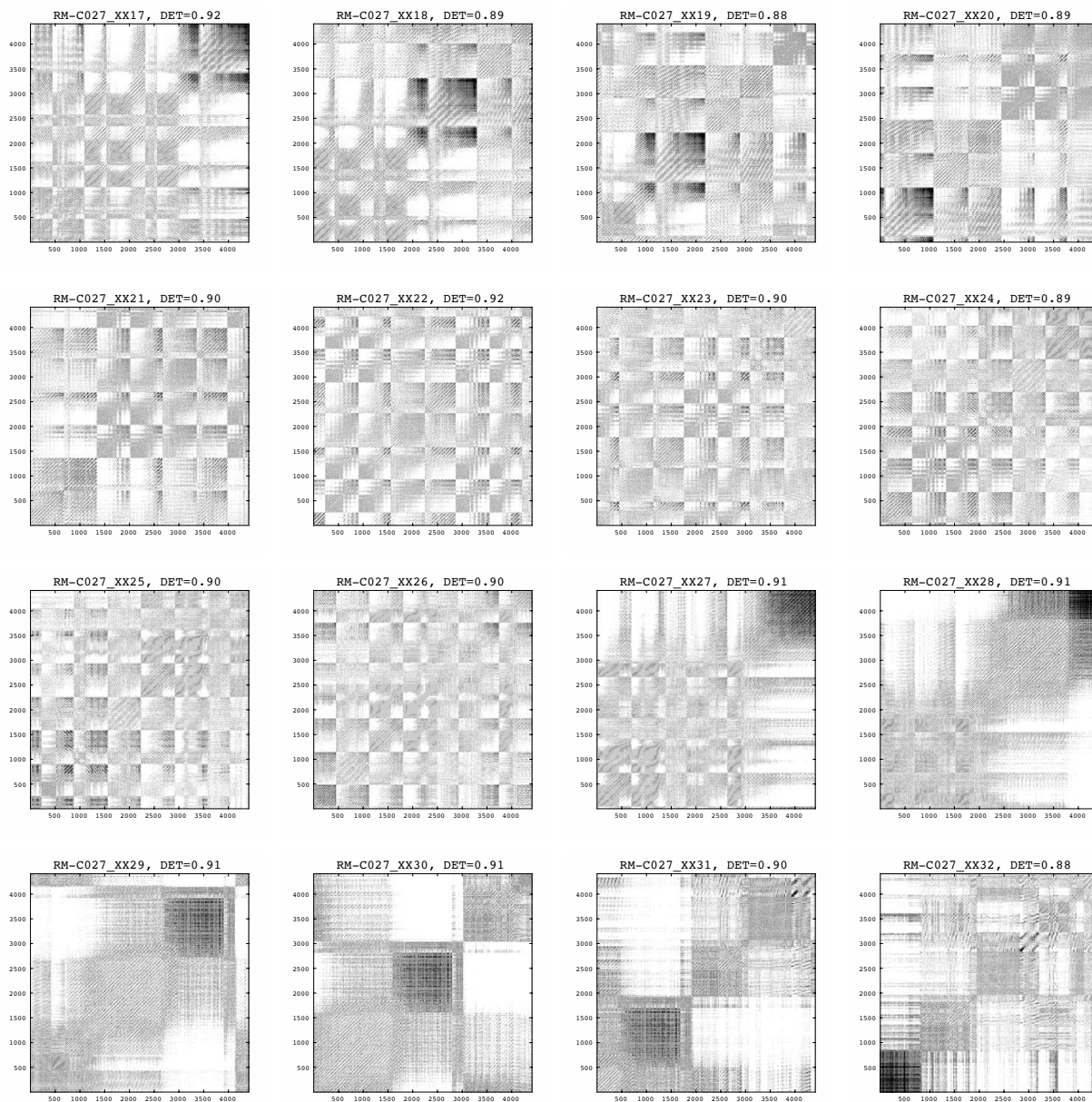
Parameters are as below.

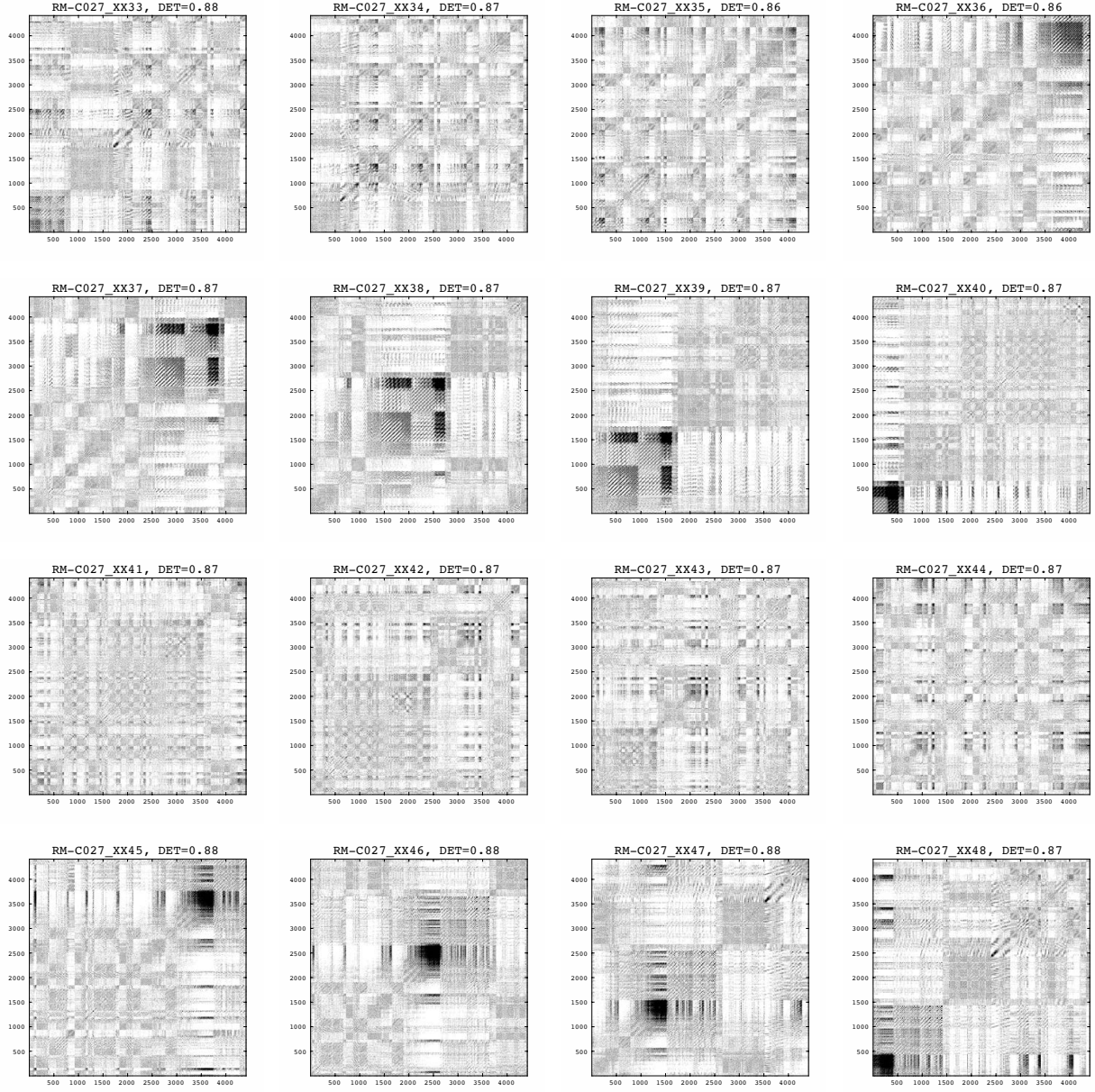
- window size: $W = 4$ seconds (4410)
- shift size: $h = 1$ second (1100)
- delay time: $\tau = 40$
- dimension: $D = 5$
- threshold: $\varepsilon_s = 0.1$
- the shortest diagonal line to be considered to calculate the histogram P : $l_{min} = 2$

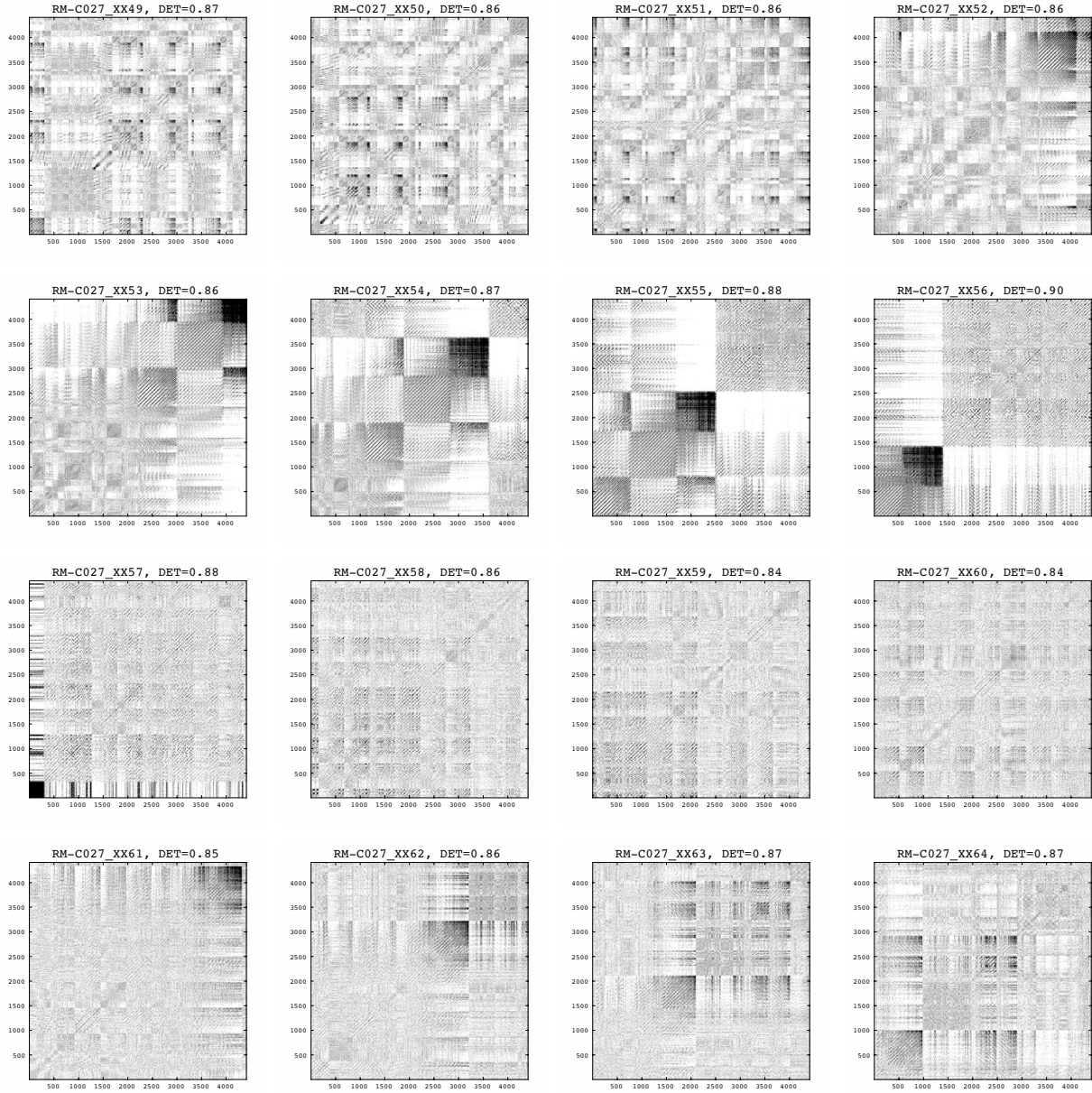
That is, the width of each short-term RP corresponds to 4 seconds, and one short-term RP is overlapped in 3 seconds with the next short-term RP.

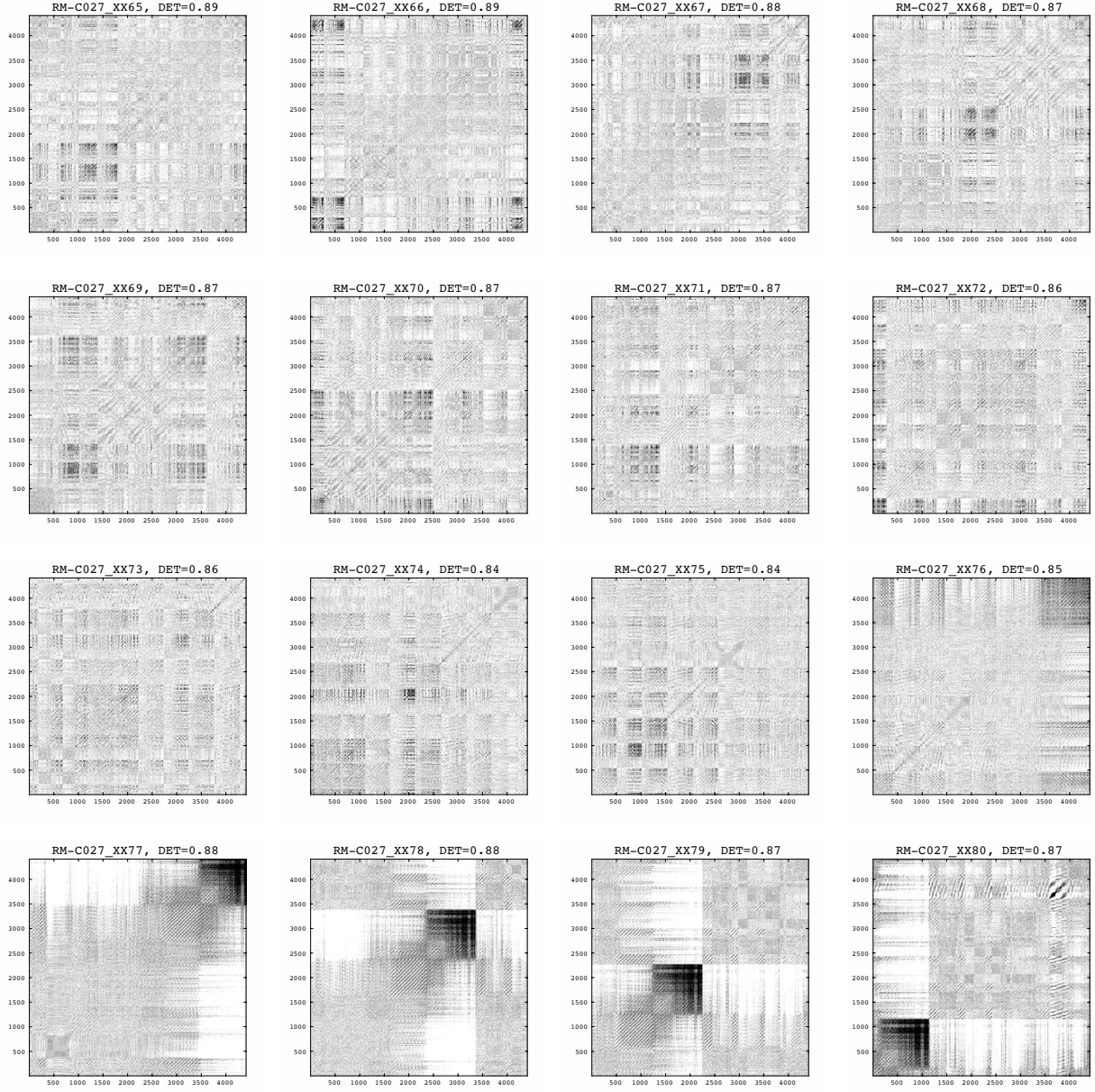
A serial number and a *DET* score are shown at the top of each RPs. The number (ex. 1, 2... of RM-C027_XX1, RM-C027_XX2...) corresponds to the starting time of the time series for the RP in seconds.

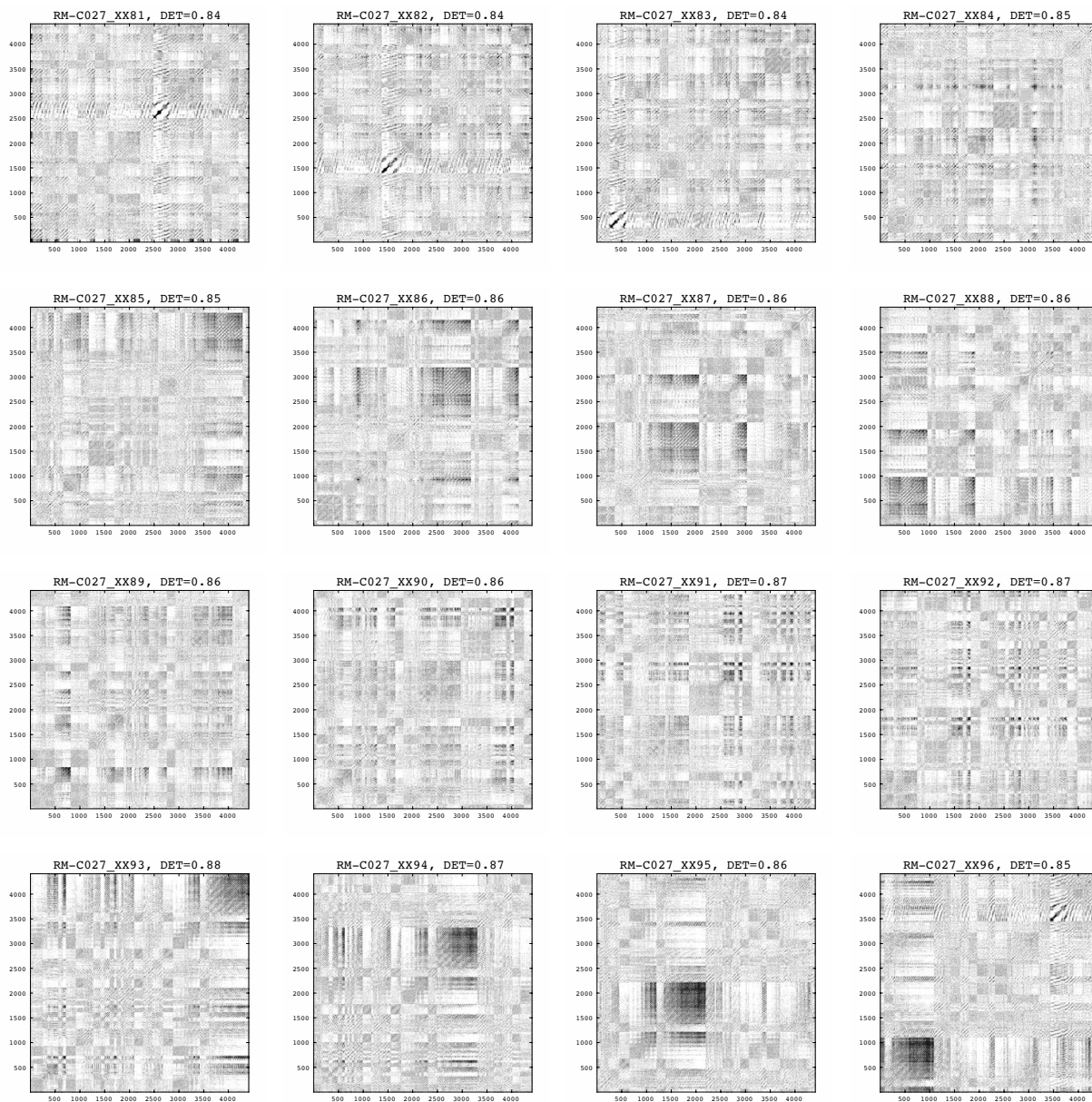


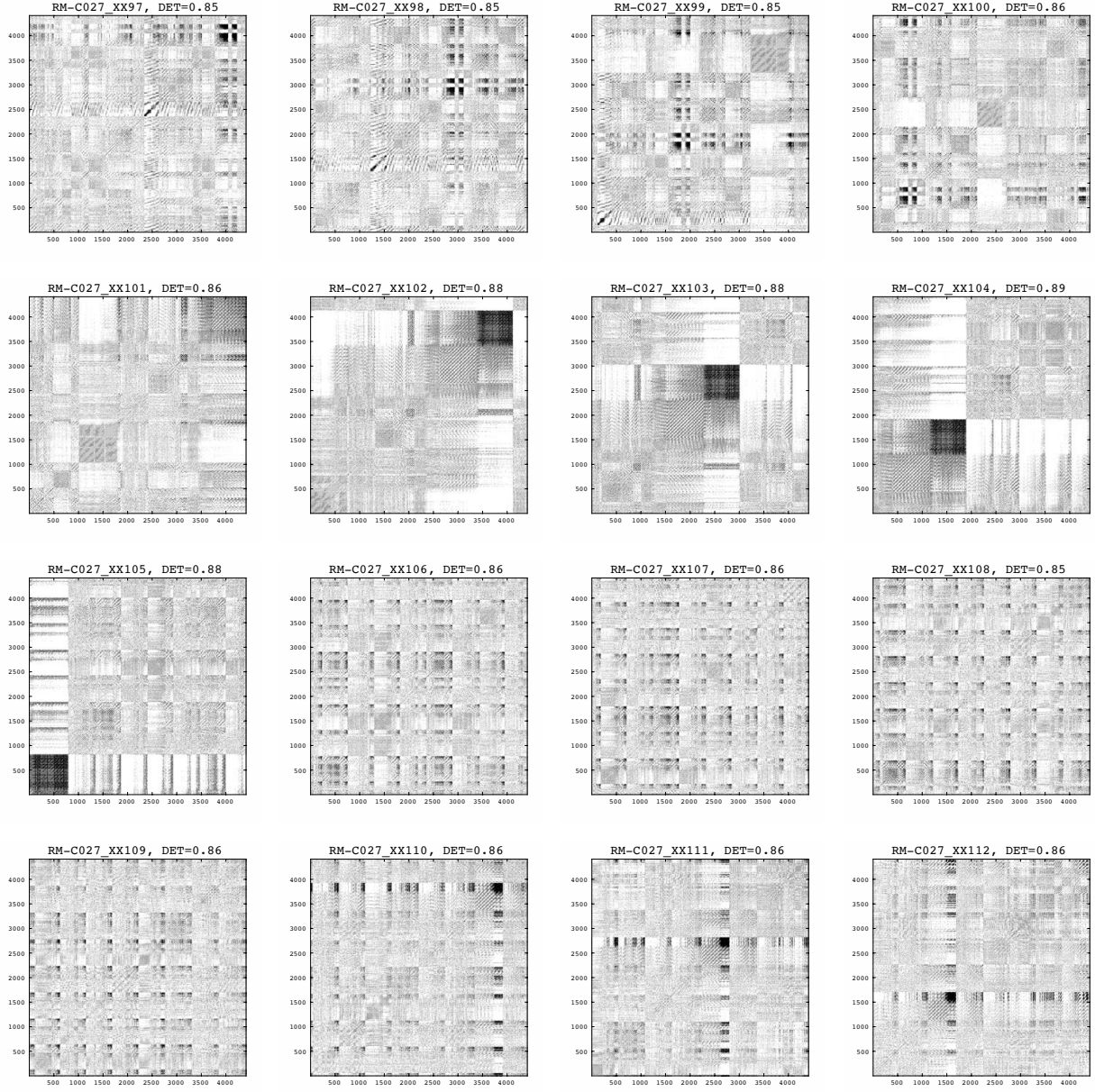


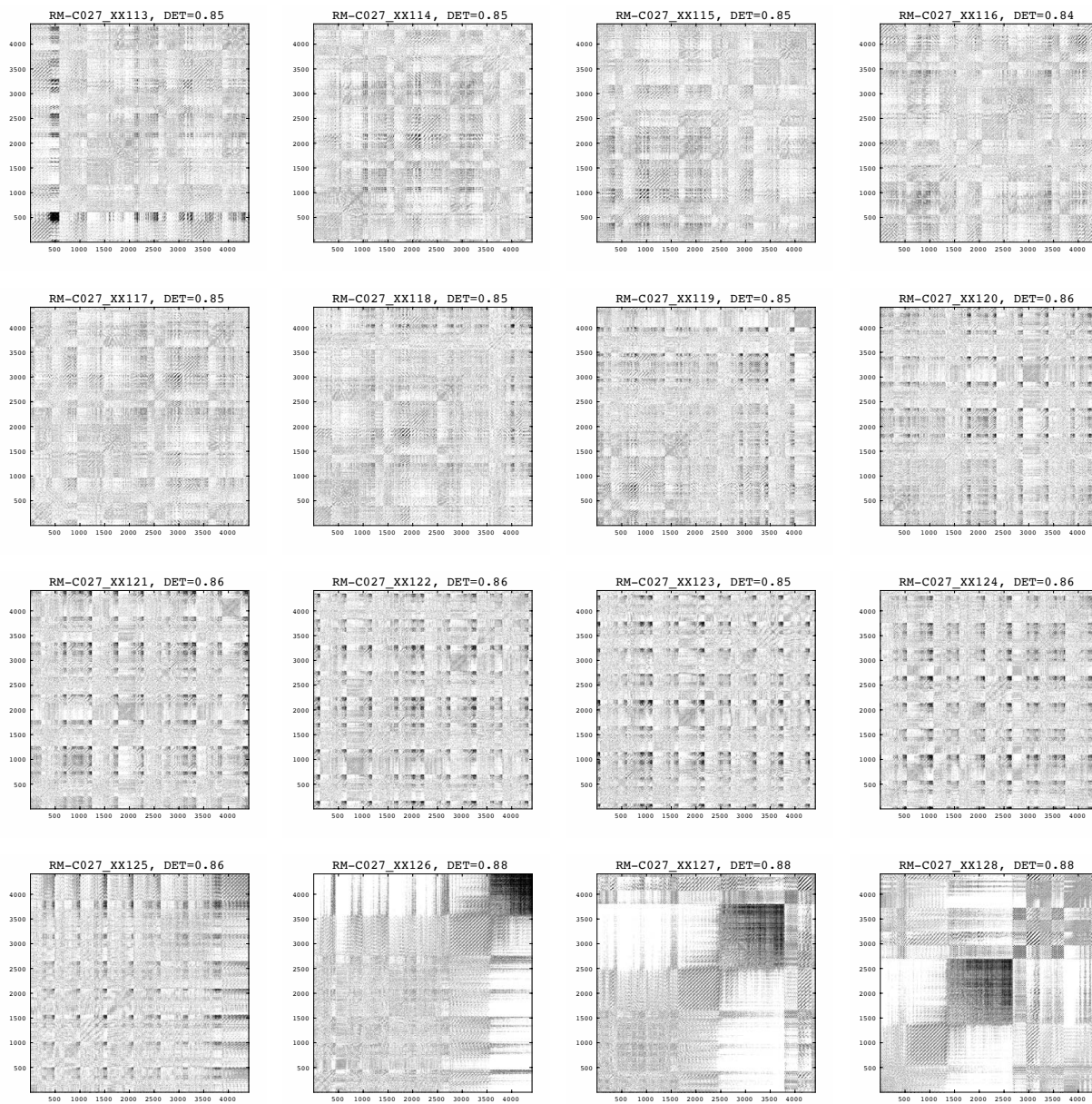


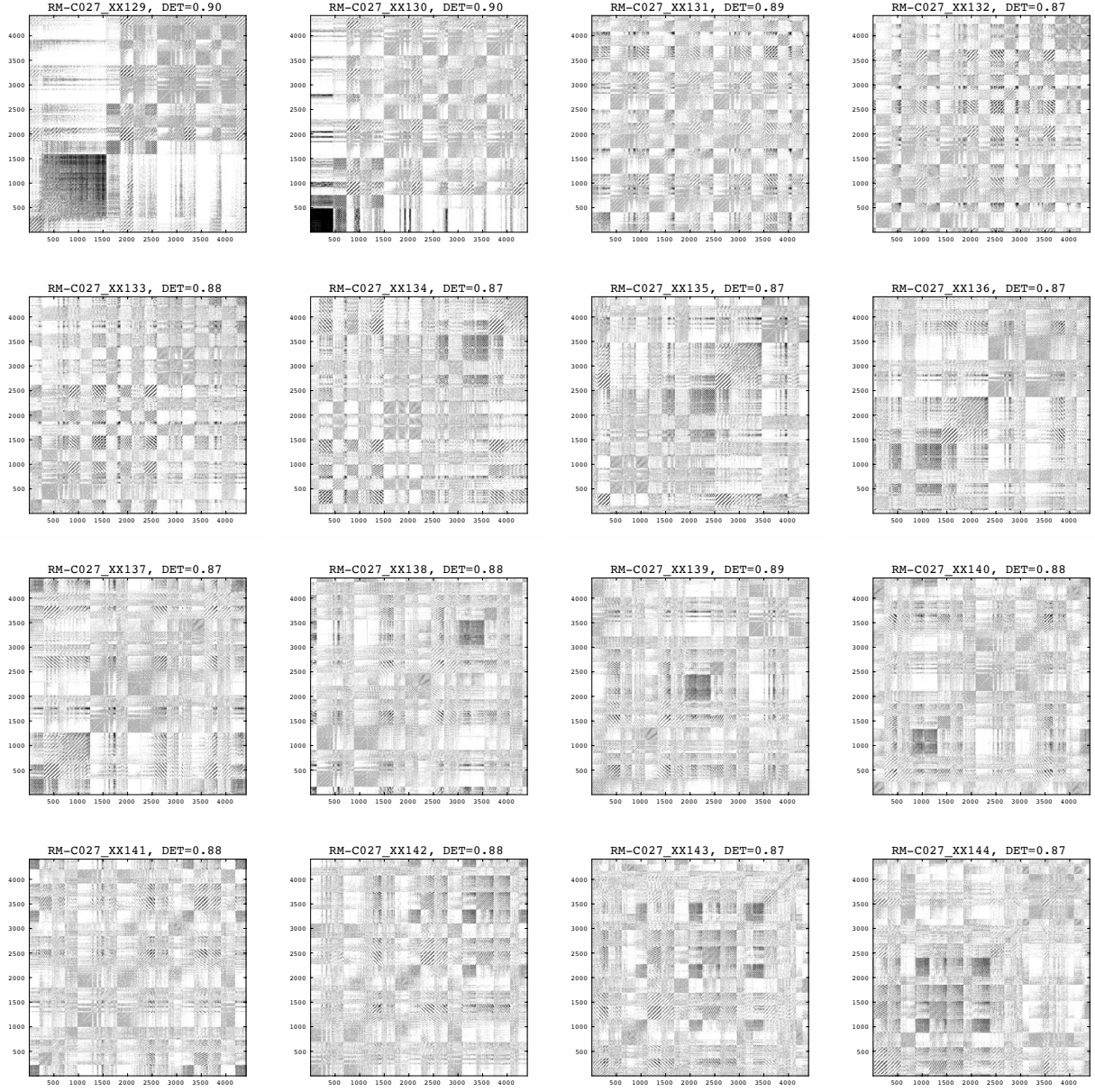


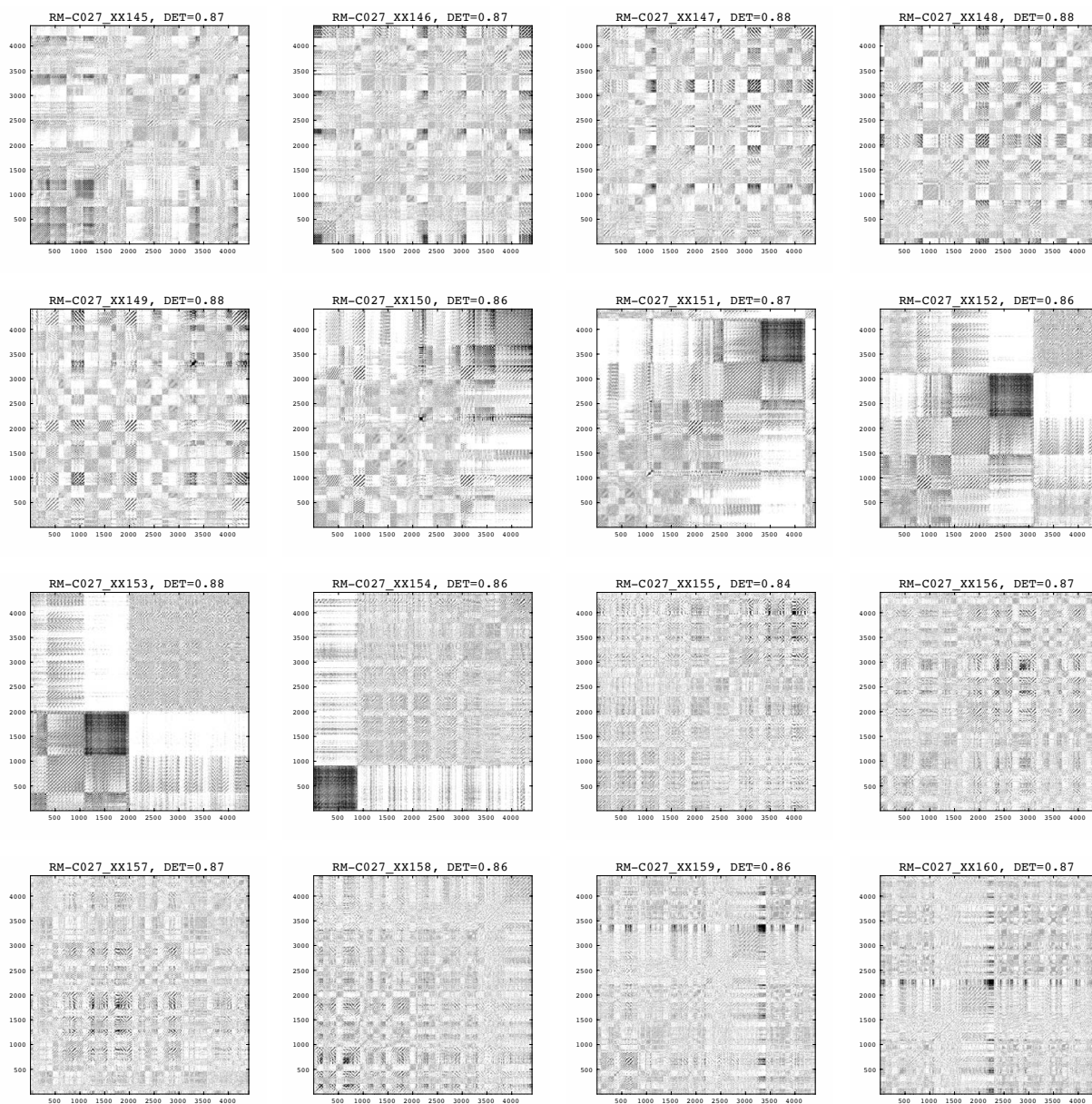


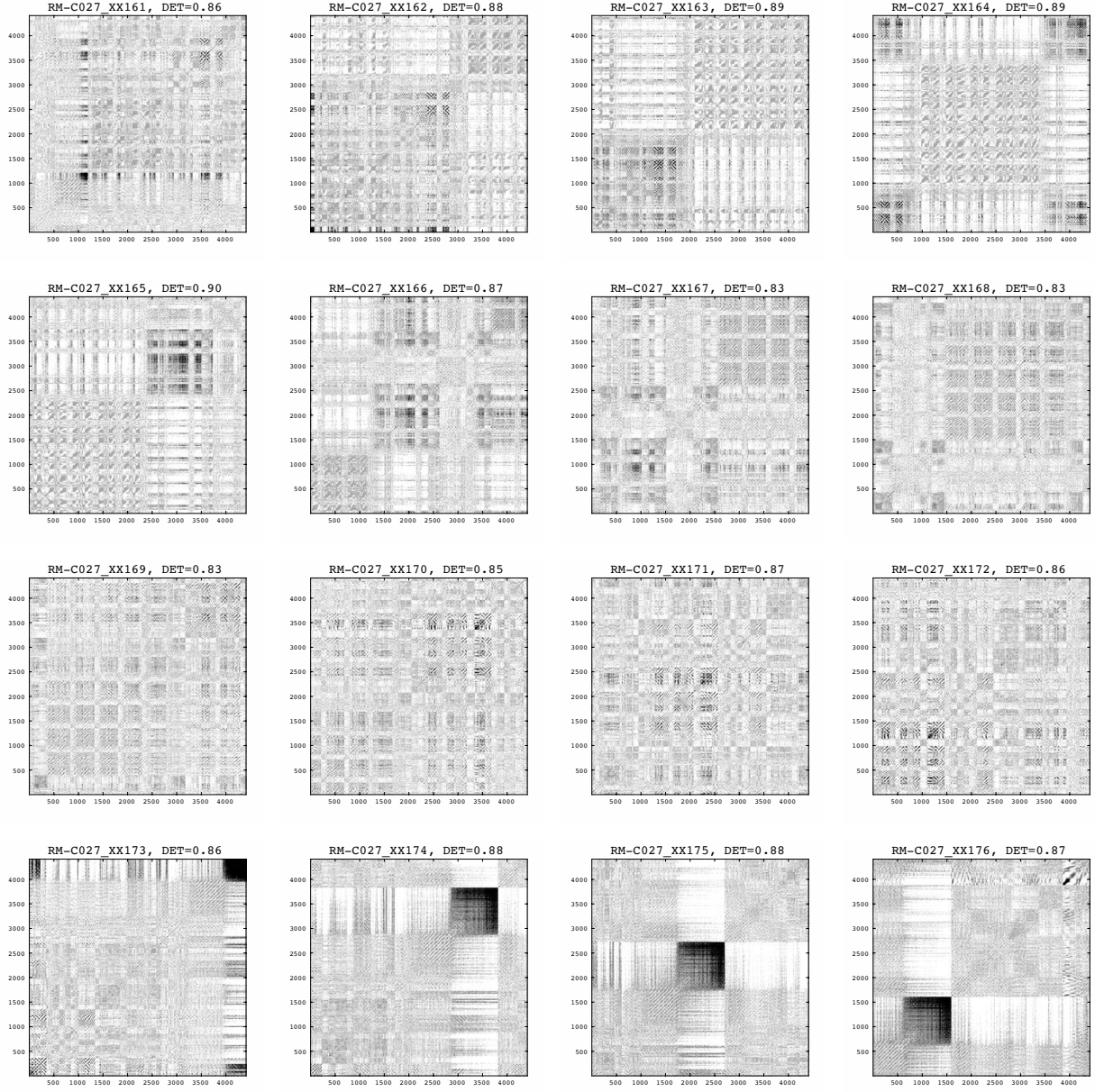


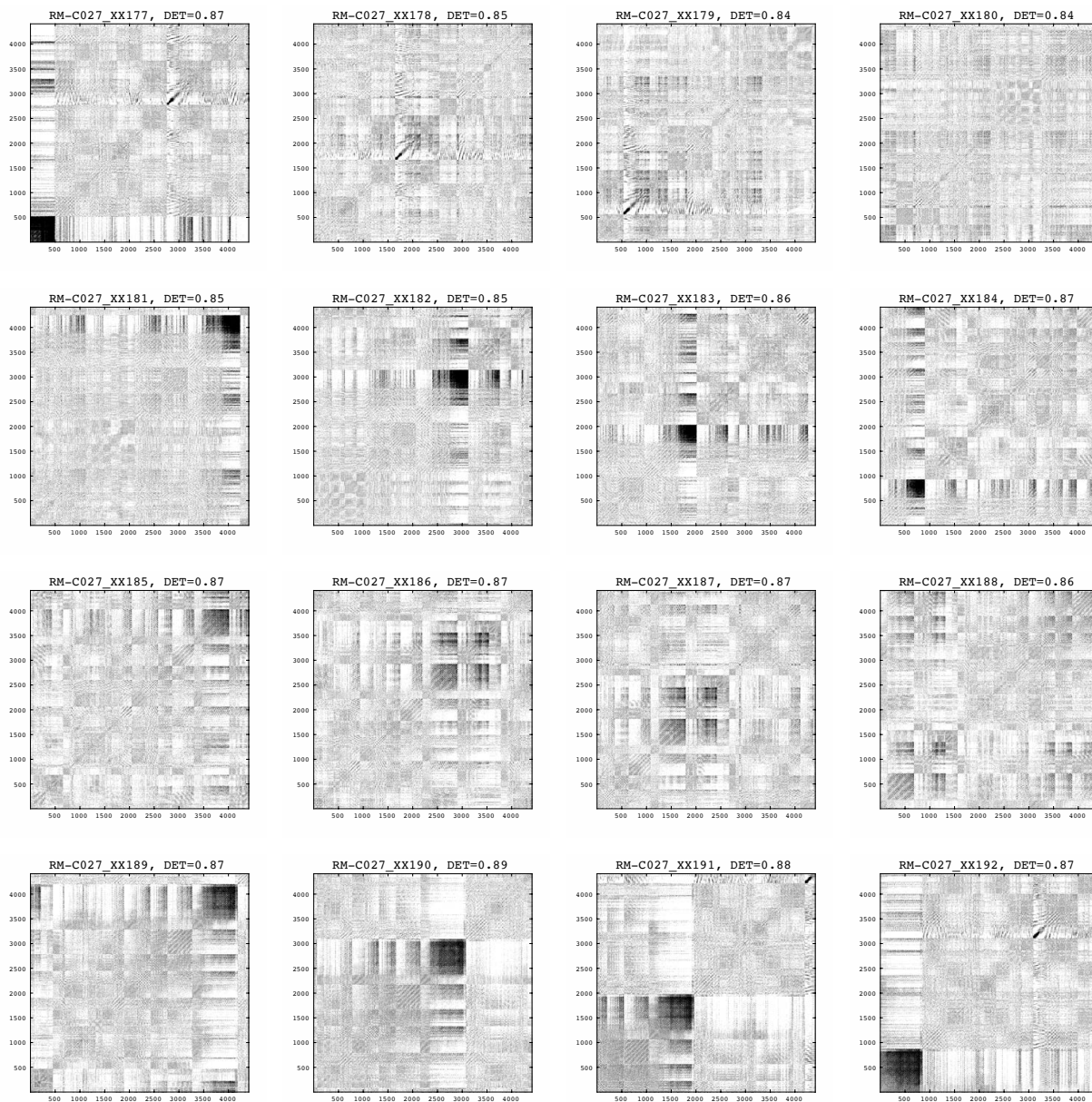


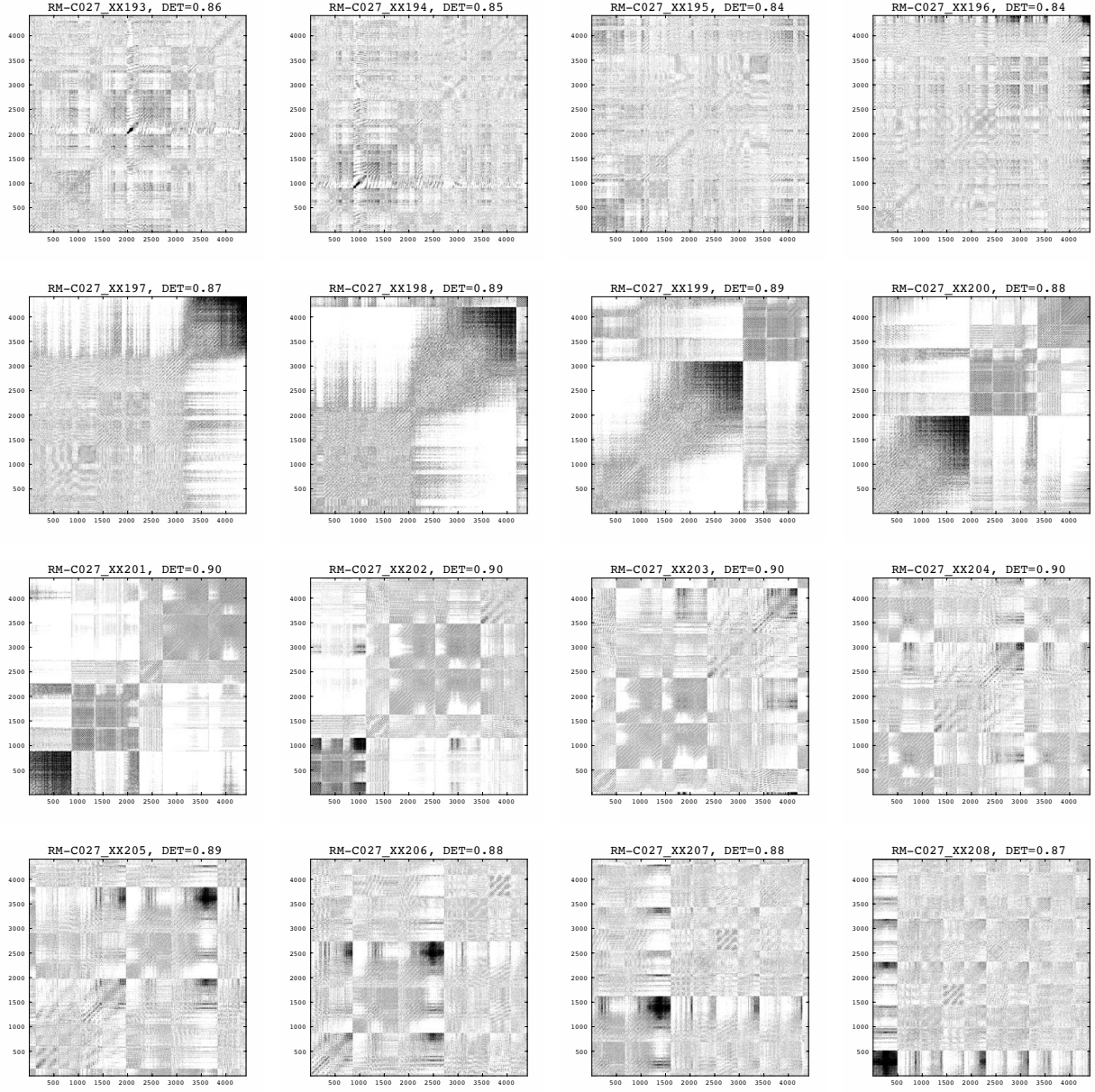


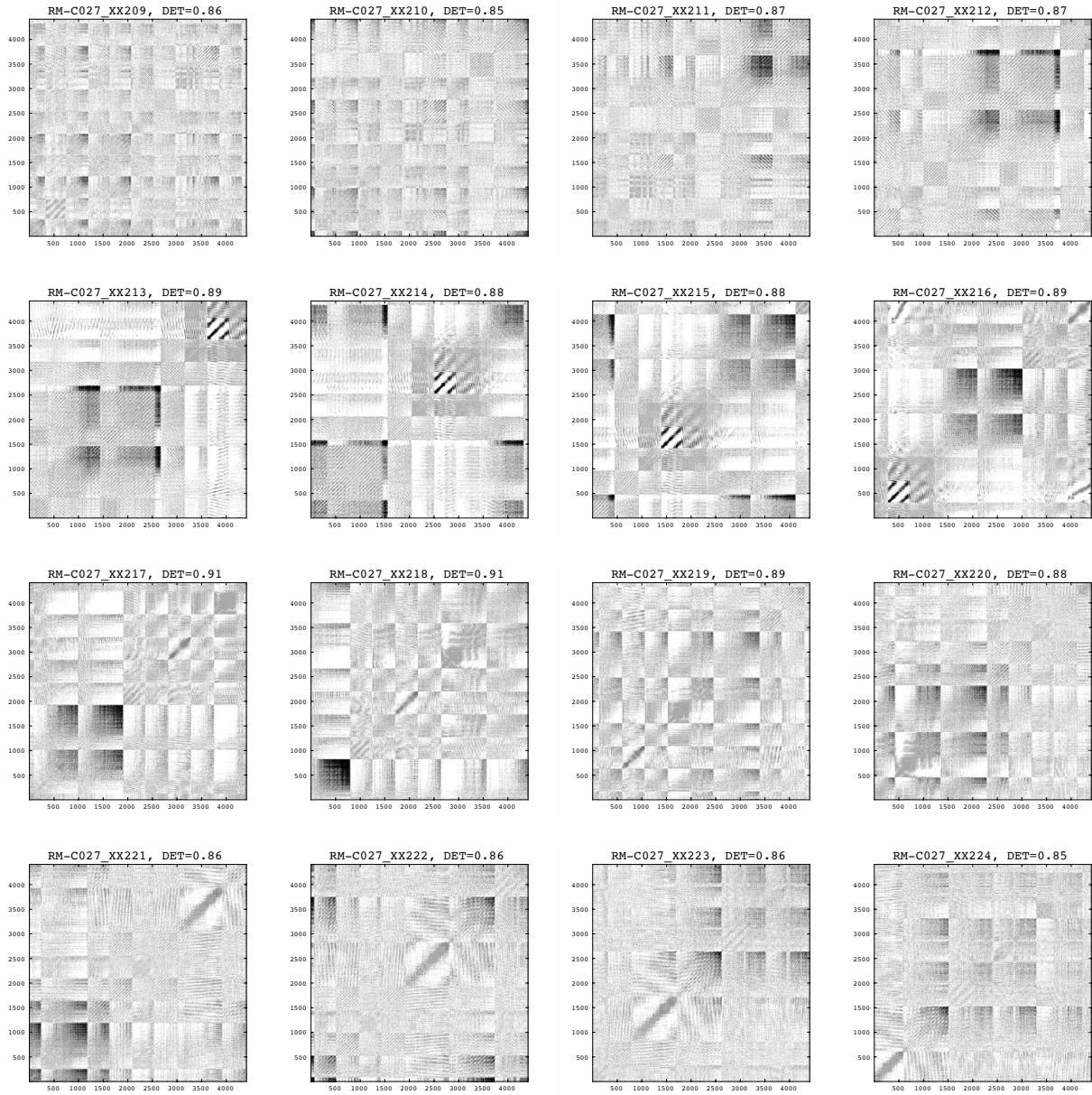


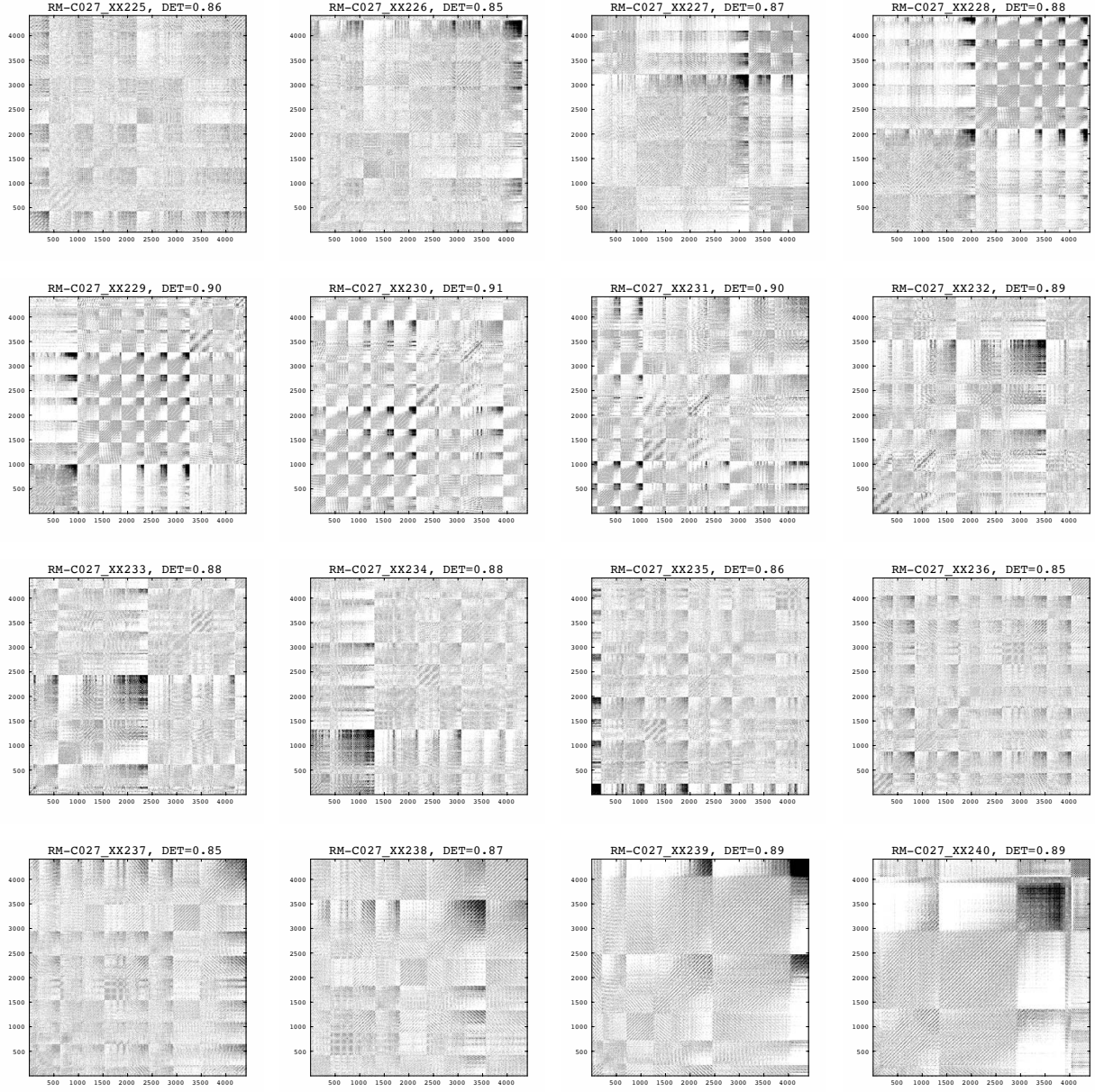


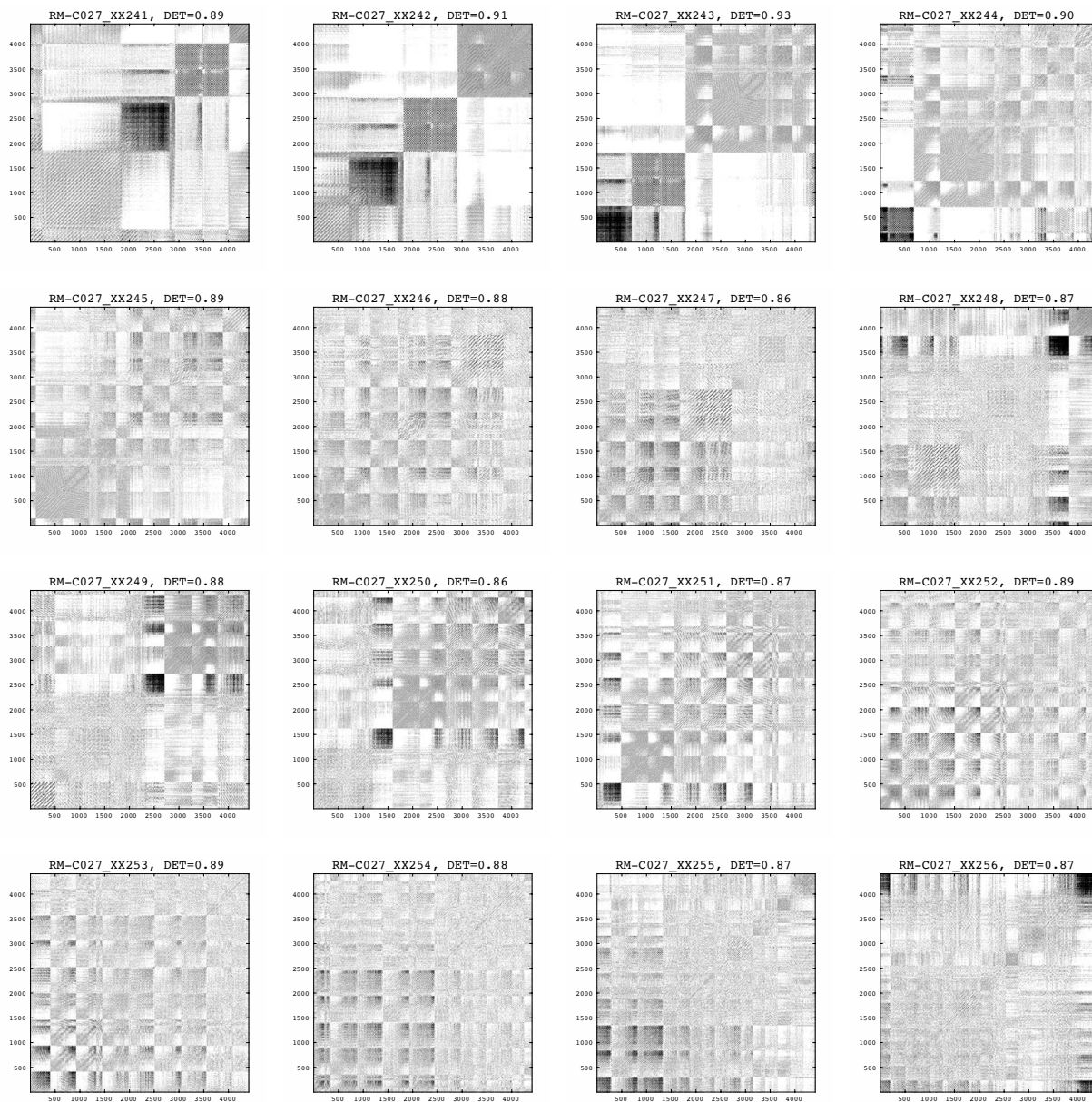


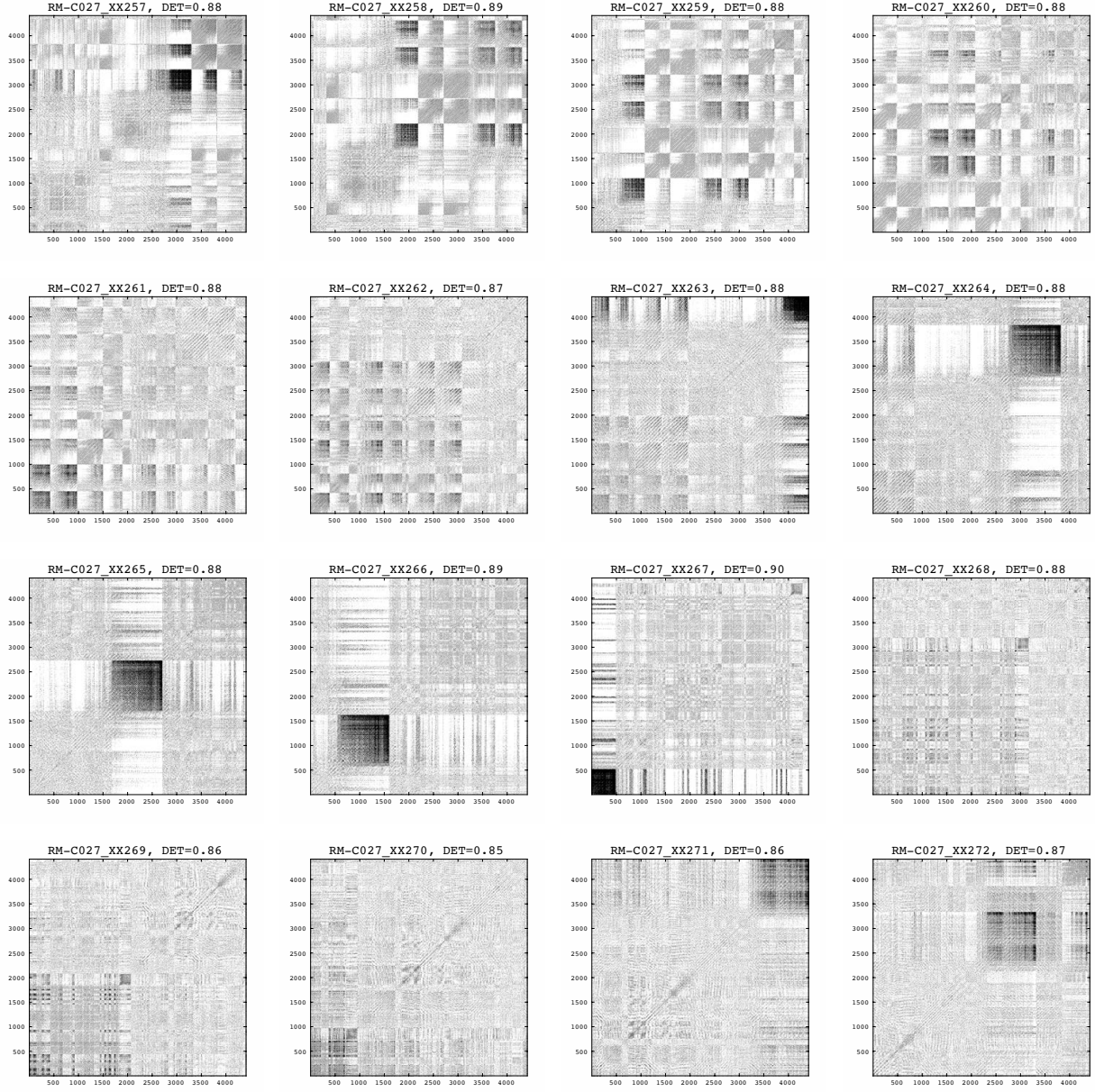


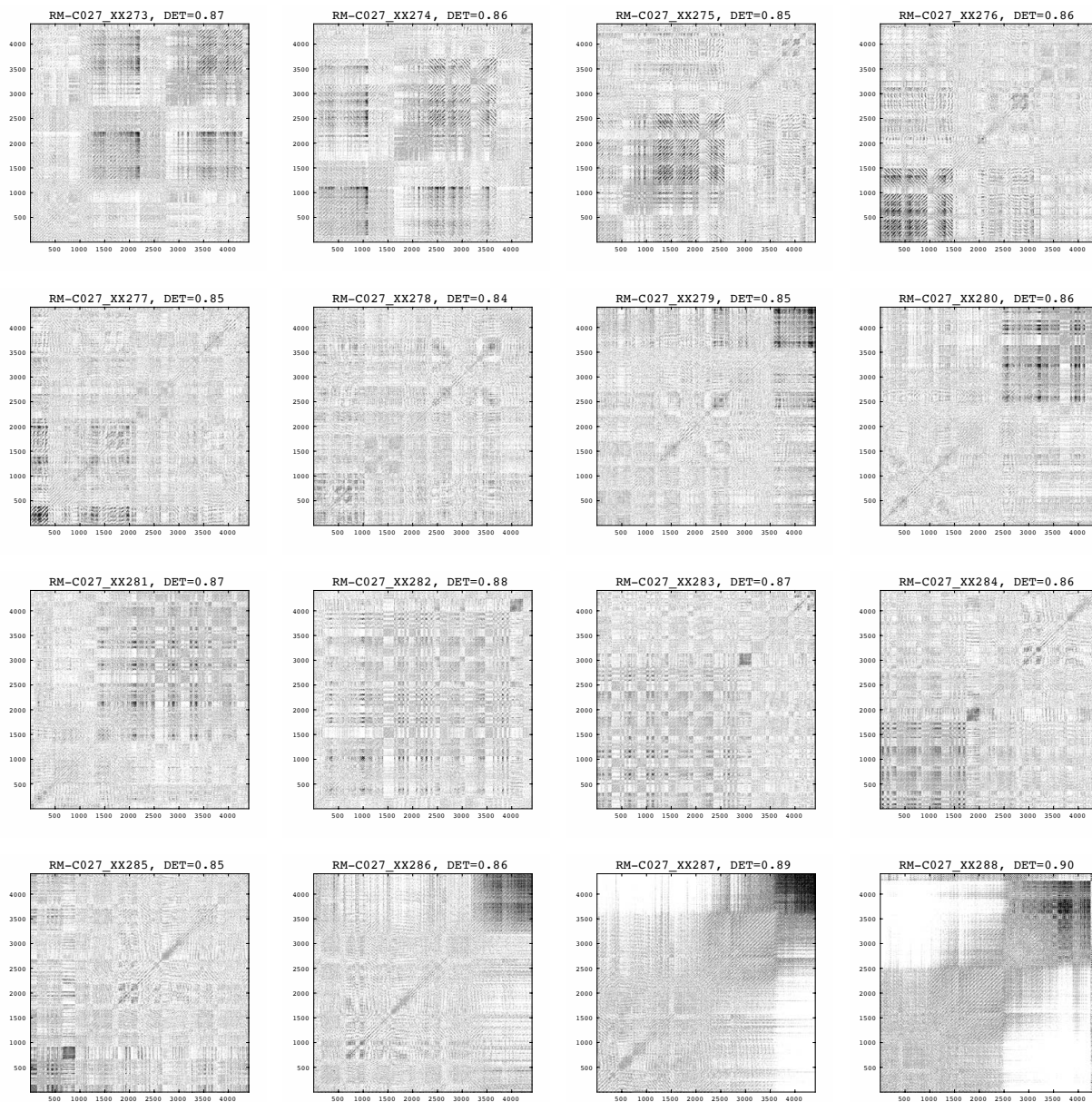


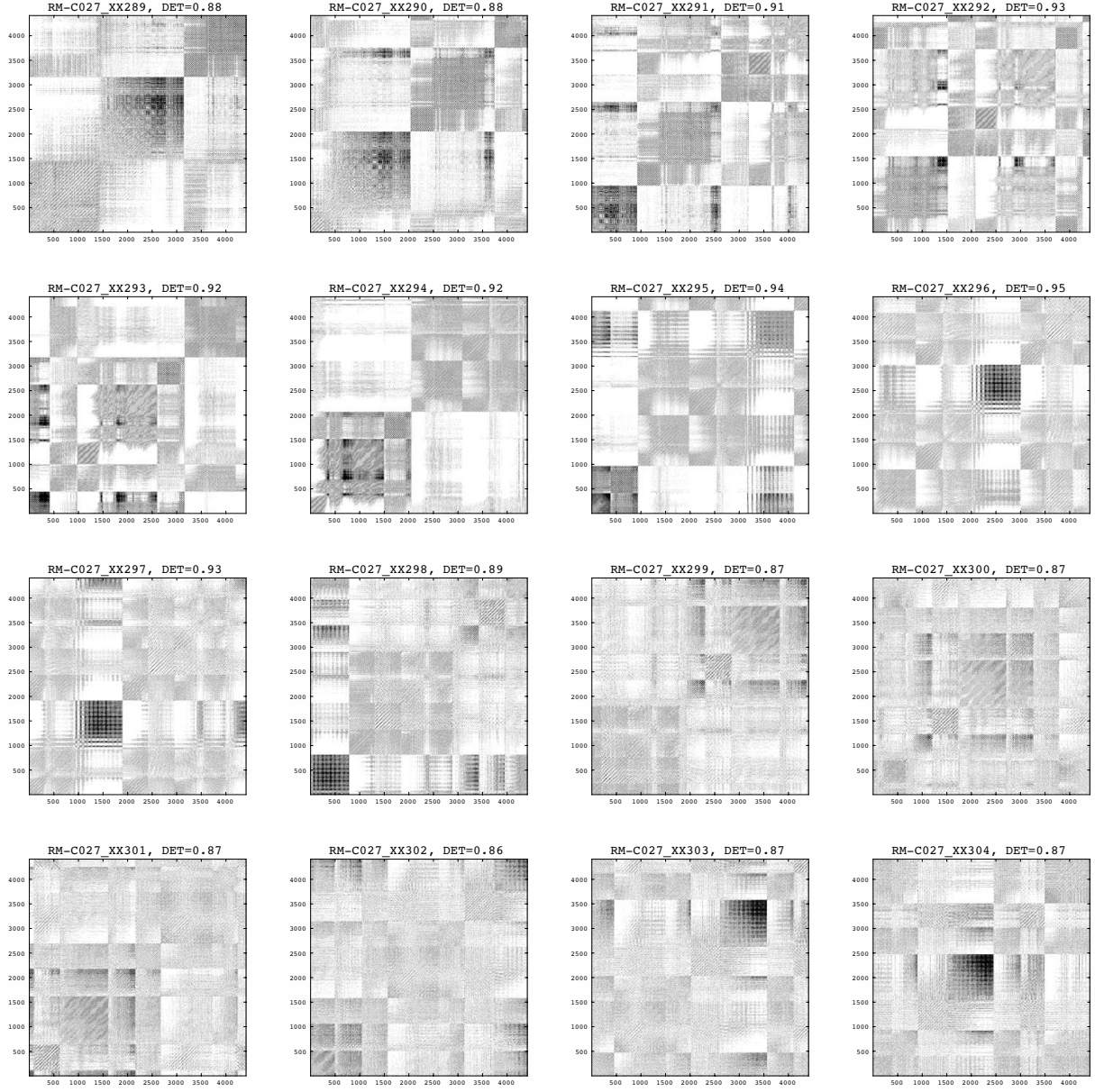


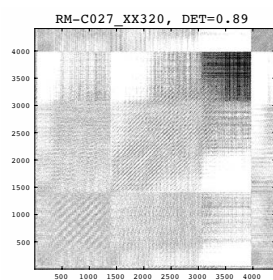
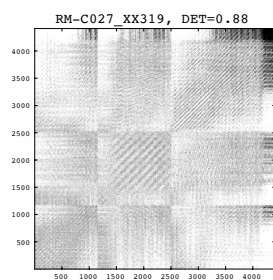
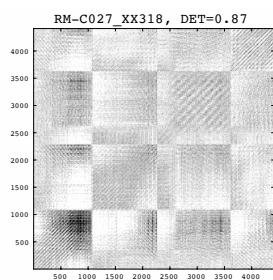
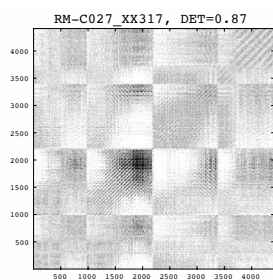
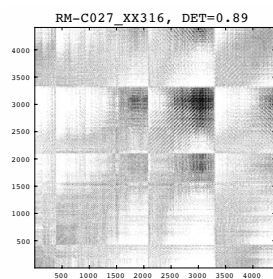
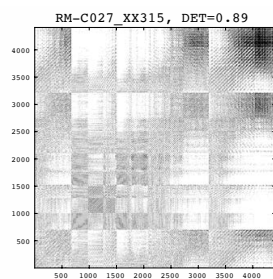
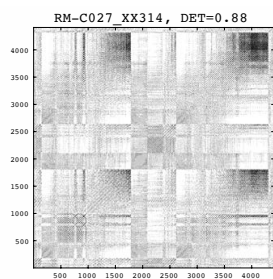
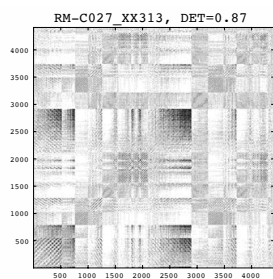
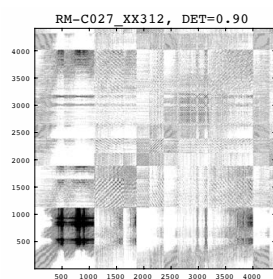
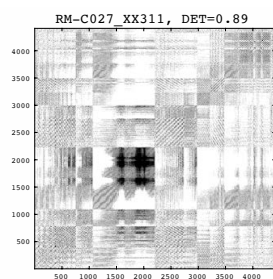
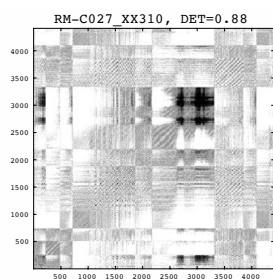
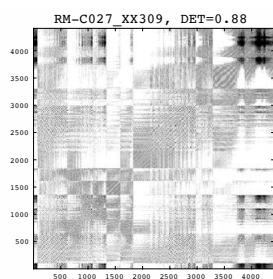
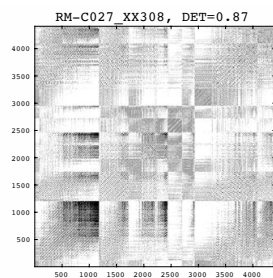
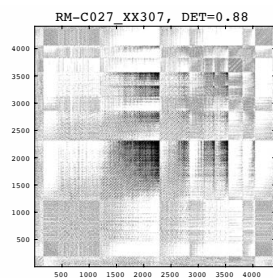
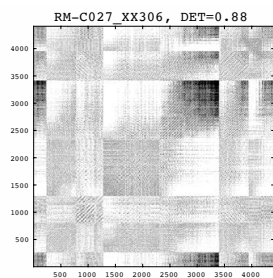
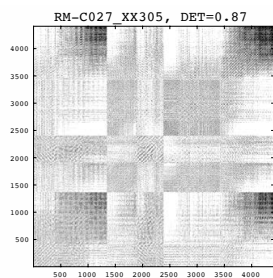


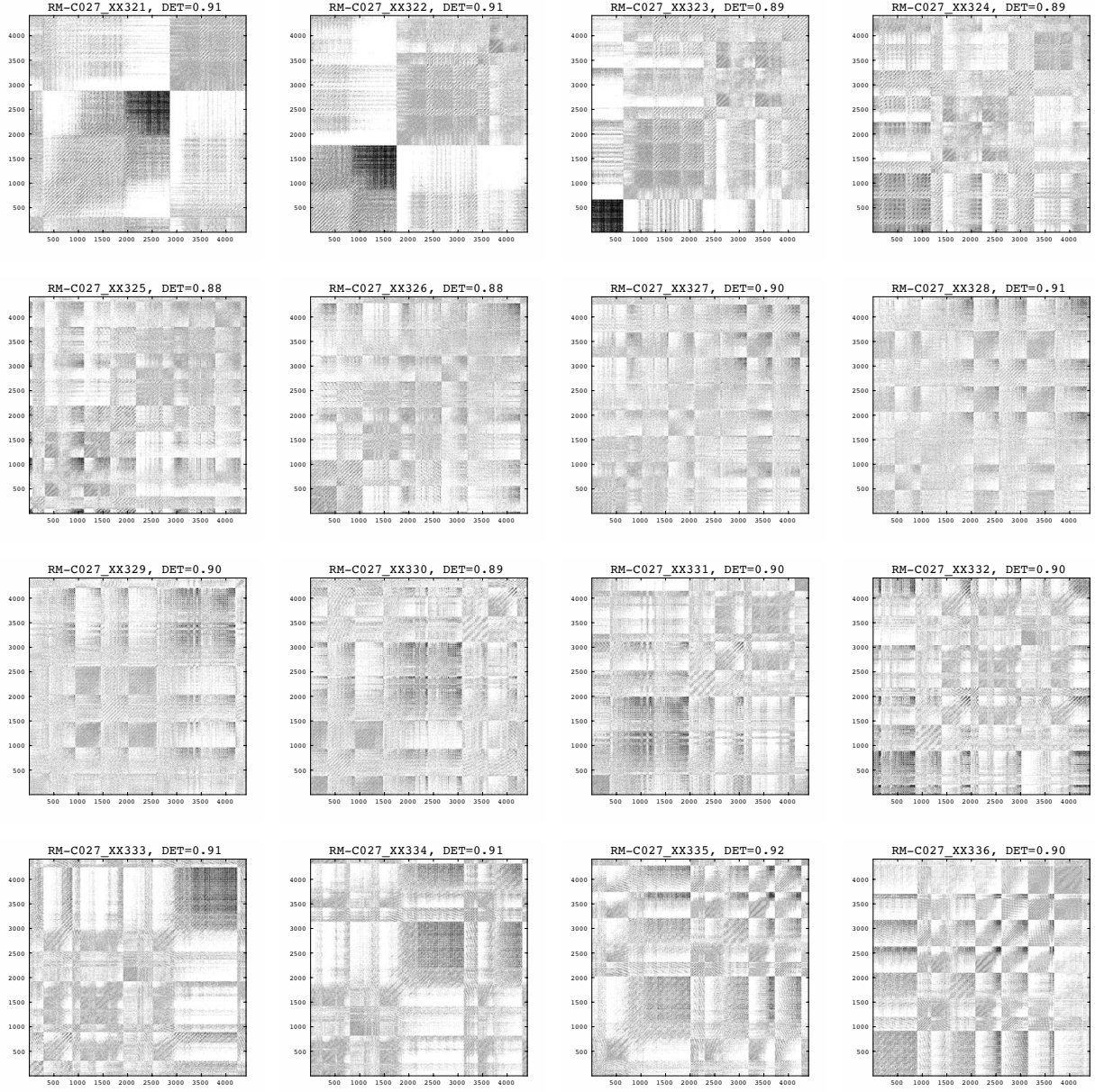


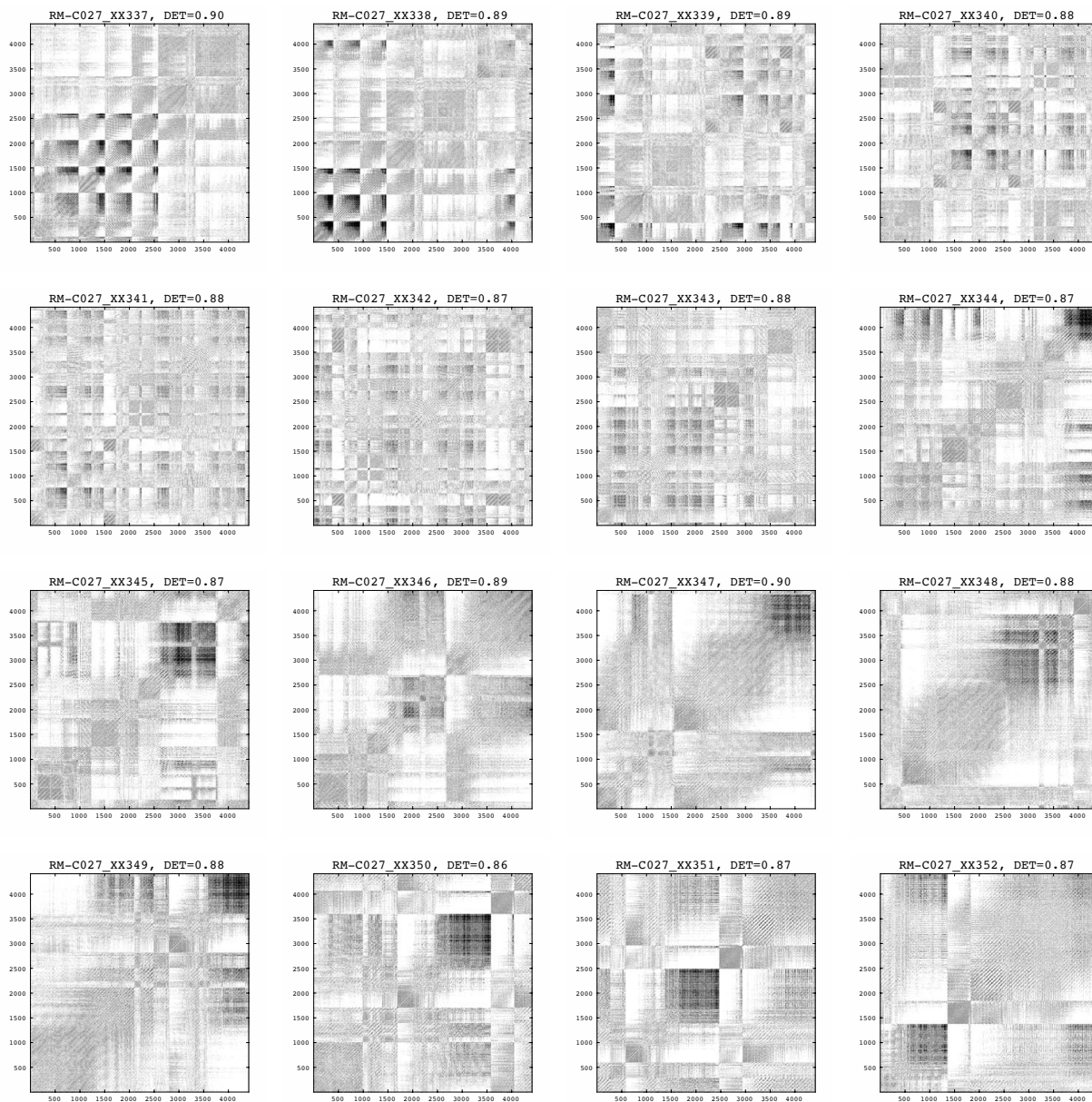


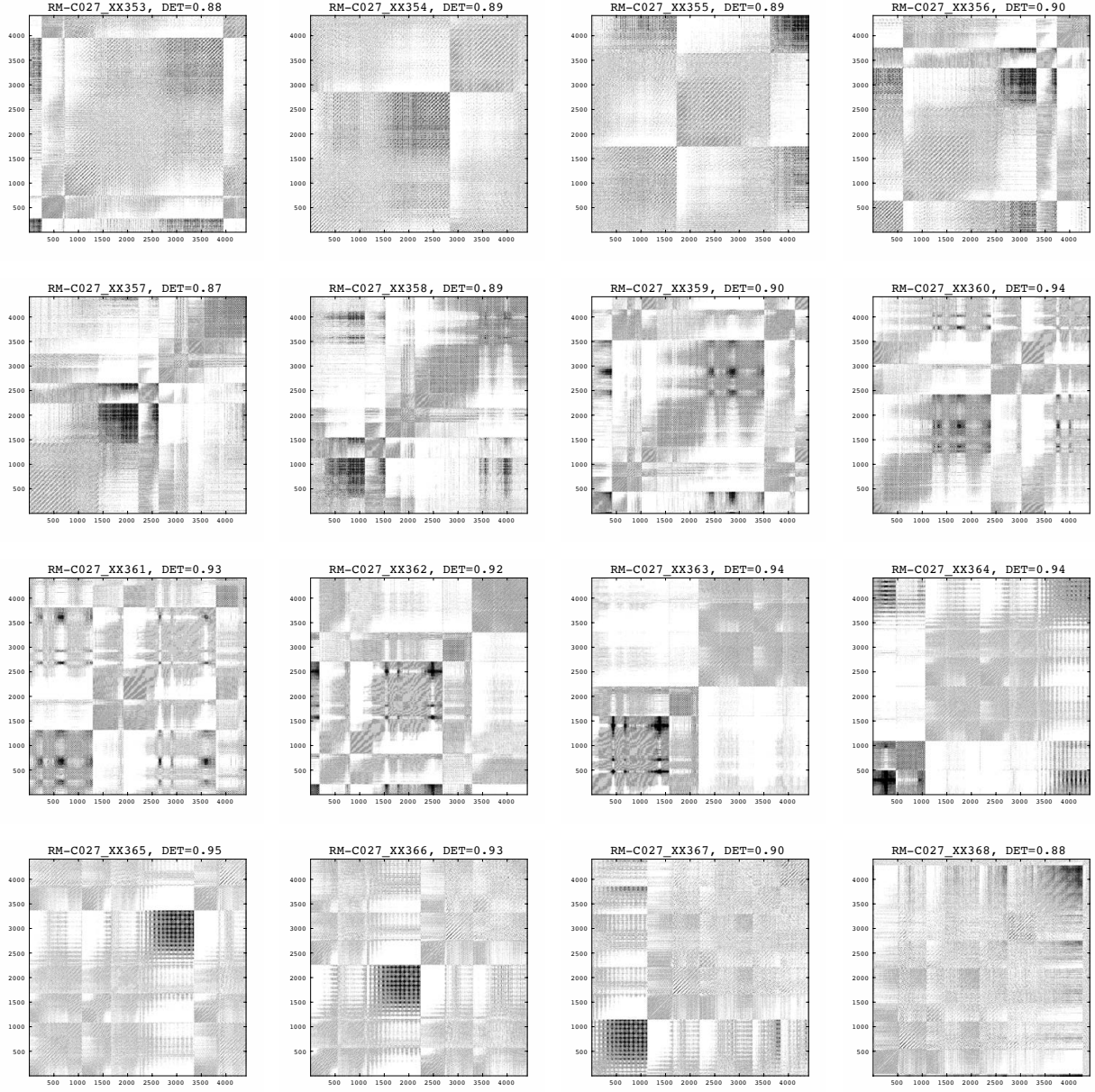


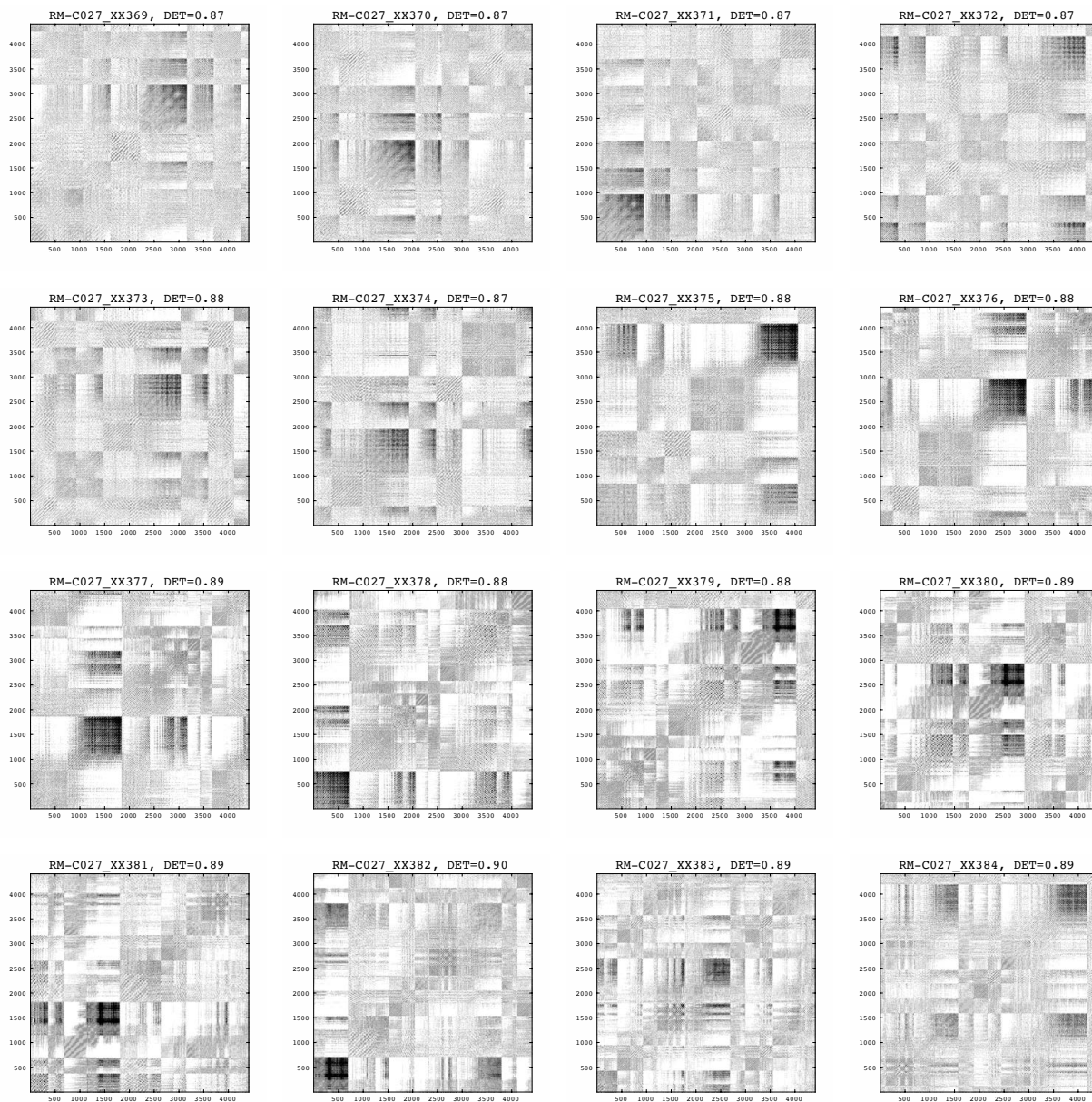


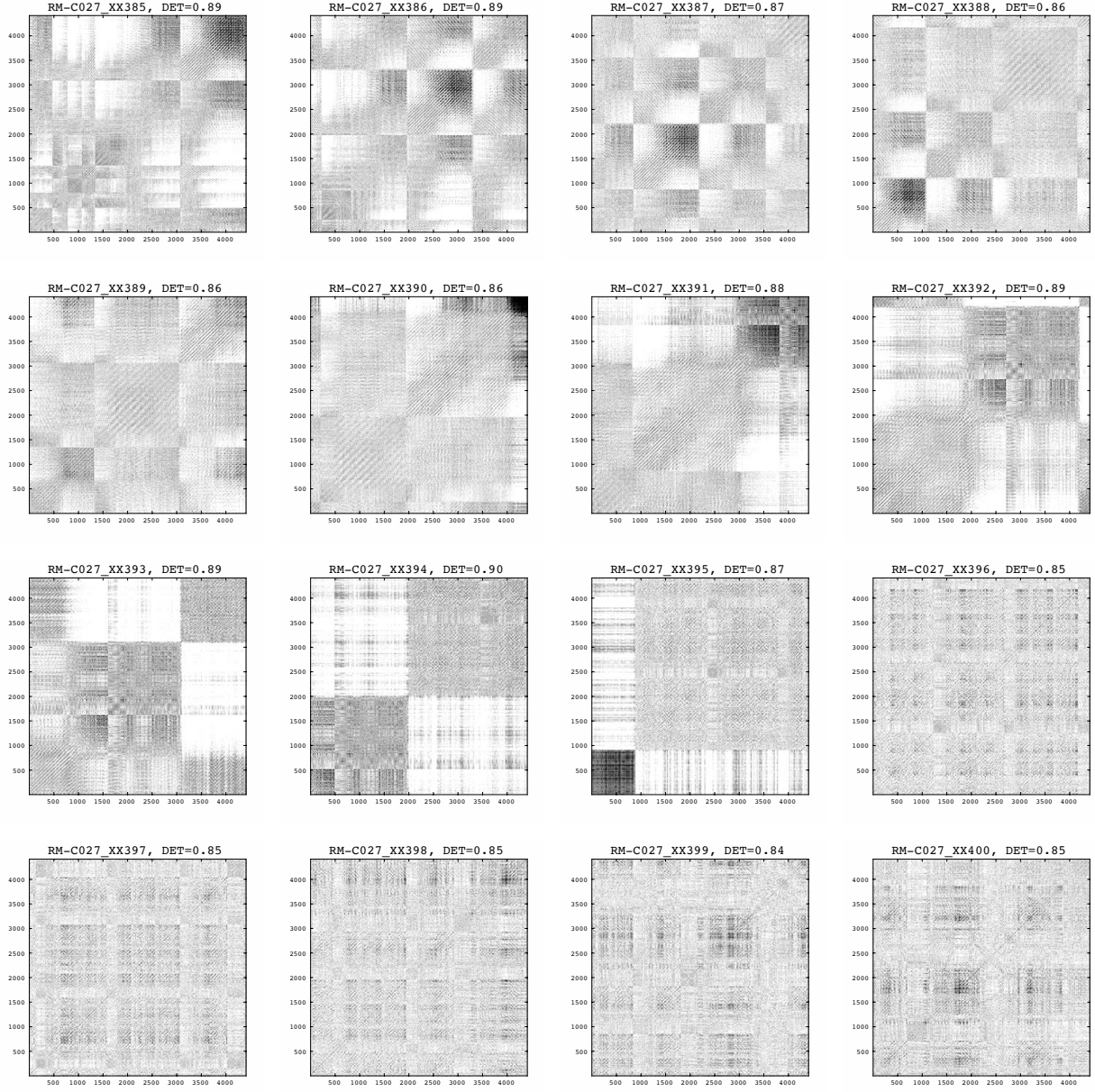


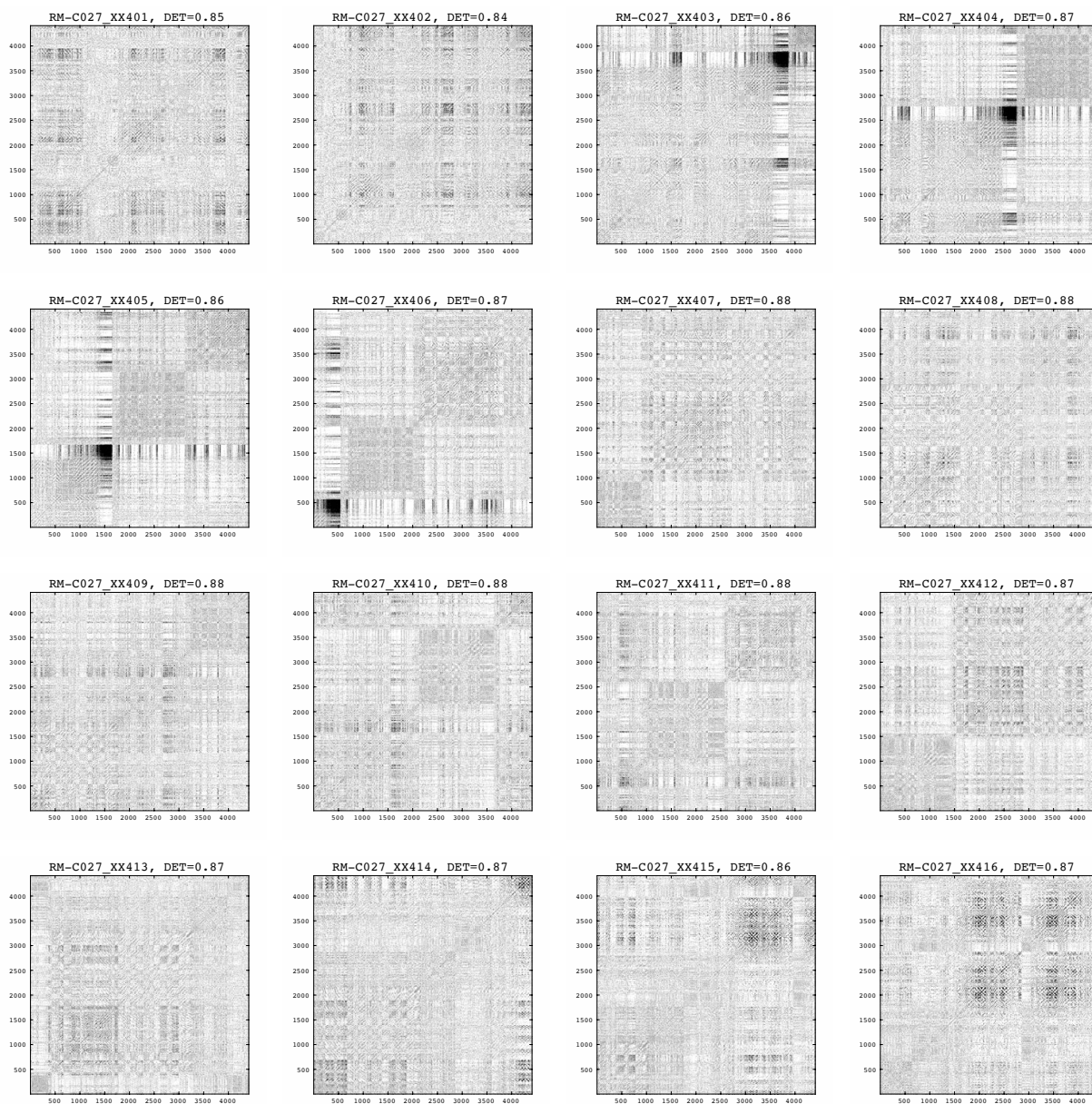


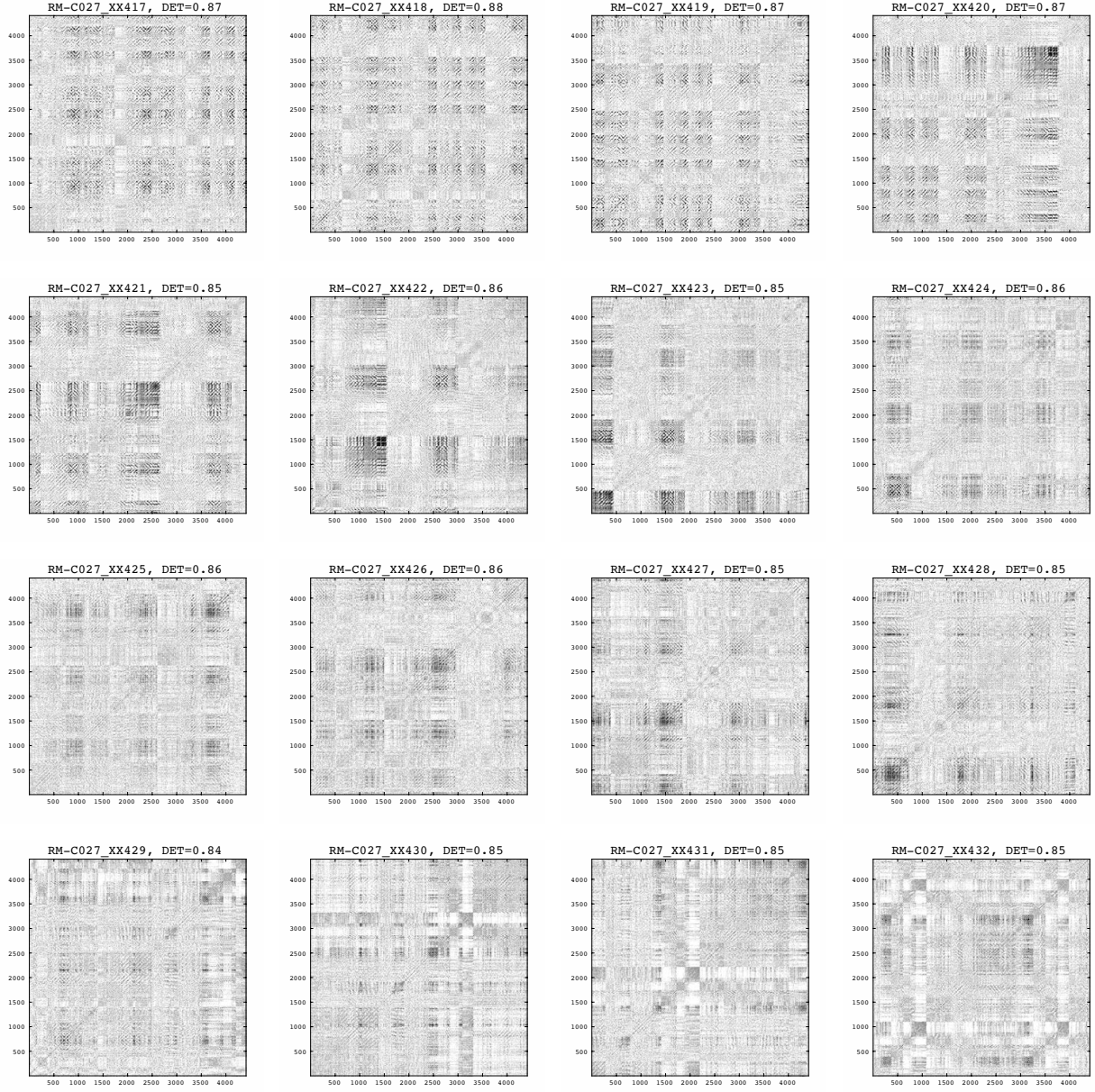


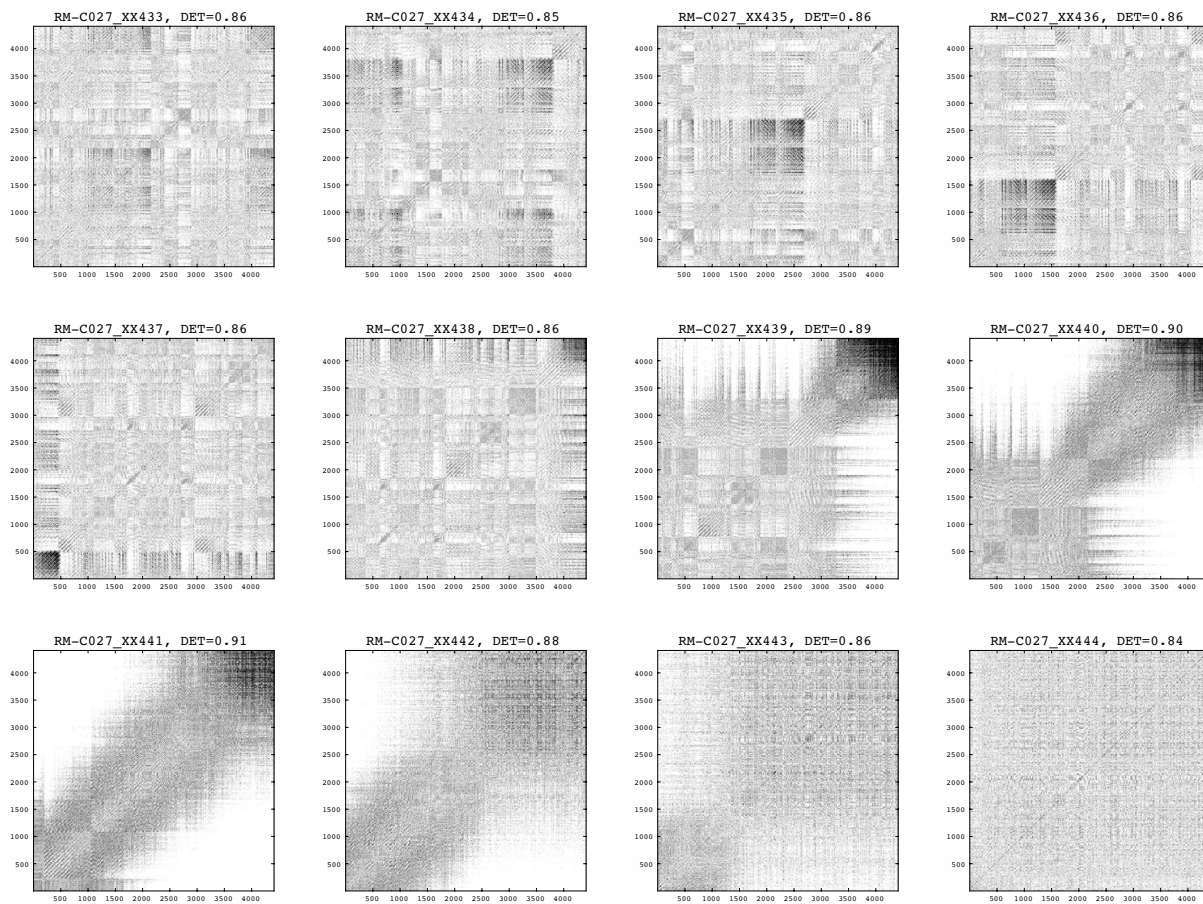












Appendix C

Thresholded short-term RPs of Beethoven1

Here I show all thresholded short-term RPs of Beethoven1 and each corresponding *DET* scores.

Parameters are as below.

- window size: $W = 4$ seconds (4410)
- shift size: $h = 1$ second (1100)
- delay time: $\tau = 40$
- dimension: $D = 5$
- threshold: $\varepsilon_s = 0.1$
- the shortest diagonal line to be considered to calculate the histogram P : $l_{min} = 2$

That is, the width of each short-term RP corresponds to 4 seconds, and one short-term RP is overlapped in 3 seconds with the next short-term RP.

A serial number and a *DET* score are shown at the top of each RPs. The number (ex. 1, 2... of RM-C028_XX1, RM-C028_XX2...) corresponds to the starting time of the time series for the RP in seconds.

