On the relationship between tempo and quantitative metric, melodic, and harmonic information in Chopin’s Prélude Op. 28 No. 4: a statistical analysis of 30 performances

Jan Beran**, Robert Goswitzb, Guerino Mazzolab, and Patrizio Mazzolac

aDepartment of Mathematics and Statistics, University of Konstanz, Konstanz, Germany; 
bSchool of Music, University of Minnesota, Minneapolis, USA; 
cDepartment of Music, Bern University of the Arts, Bern, Switzerland

(Received 14 October 2013; accepted 28 May 2014)

We investigate the relationship between analytical information extracted from the score of Chopin’s Prélude Op. 28 No. 4 and tempo curves obtained from 30 recordings by well-known pianists. The score-related information consists of metric, melodic, and harmonic weight functions. A hierarchic regression approach is used. The statistical analysis yields strong evidence for the approximate validity of an operator defined in terms of hierarchically decomposed weight functions, as originally introduced by G. Mazzola and J. Beran in 1997 (“Rational Composition of Performance.” In: Regulation of Creative Processes in Music, edited by W. Auhagen and R. Kopiez, Staatliches Institut für Musikforschung, Berlin, pp. 37–67) in the context of Schumann’s Träumerei. Further analysis based on these findings reveals some typical features of the performances. Supplemental online material for this article can be accessed at http://dx.doi.org/10.1080/17459737.2014.930192.

Keywords: hierarchic regression; kernel smoothing; tempo curve; metric weights; melodic weights; harmonic weights; performance research; Rubato; Schumann; Chopin

1. Introduction

If there is any truth in Leibniz’s famous claim “musica est exercitium arithmeticae occultum nescientis se numerare animi [music is a hidden arithmetic exercise of the soul, which does not know that it is counting],” performance must play a crucial role as it mediates between rational structures that are hidden in a score and emotional responses evoked by the performative communication to the listener. Performance is a critical pivot when investigating the relation between rational and emotional layers of musical reality and communication. The scientific truth of this thesis is however far from confirmed. One straightforward strategy towards such a proof is to break it down into two partial statements: first, that performance is strongly related to emotional responses, and second, that performance is strongly related to hidden rational score structures.

In a recent paper, music psychologists John A. Sloboda and Andreas C. Lehmann have investigated the first thesis for Chopin’s Prélude Op. 28 No. 4 (Sloboda and Lehmann 2001) – see Figure 1(a). They analysed the emotional responses of listeners to performances by ten pianists

**Corresponding author. Email: jan.beran@uni-konstanz.de

© 2014 Taylor & Francis
by comparing performance parameters of inter-onset intervals (for tempo), dynamics of the melody, and asynchronies between melody and bass notes to emotional responses. Comparisons were made during and after listening to each performance. In their conclusion, the authors state: “This study showed that musical structure had a very important impact on the emotionality ratings made by listeners during performance.” We should emphasize that in their context “musical structure” means structure of the performed music.

The present paper considers the same composition by Chopin, but relates to the second thesis. Understanding musical performance as an expression of analytical rationales (apart from emotional or gestural rationales) is a problem that has gained attention from performance research since the KTH software-implemented model of performance was introduced by Johann Sundberg and his collaborators (Sundberg, Askenfelt, and Frydén 1983). Understanding can be attempted...
by computer-driven models of performance (such as Sundberg’s software or our own approach using the Rubato performance rubette – Mazzola and Zahorka [1994]) or by empirical research based on recorded performances. The latter approach was highlighted by Bruno Repp’s influential paper (Repp 1992) on agogic performance in Schumann’s Träumerei, and we also used Repp’s tempo data to initiate our research on musical performance (Mazzola and Beran 1997; Beran and Mazzola 1999a, 1999b, 2000, 2001). (See also Mazzola [2002, 2011] and references therein.)

Our approach extended Sundberg’s idea by introducing three analytical tools (modules implemented in the software Rubato): the MetroRubette for metric analysis, the MeloRubette for melodic analysis, and the HarmoRubette for harmonic analysis. We thus focused on performance as an expression of metric, melodic, and harmonic curves provided by Rubato. This lead to collaborative interdisciplinary research between a statistician (Jan Beran), a musicologist (Robert Goswitz), a mathematical music theorist (Guerino Mazzola), and a pianist (Patrizio Mazzola).

Let us briefly summarize the main results obtained from Repp’s tempo data of 28 performances and our analytical data extracted from Schumann’s Träumerei. Apart from the general result that there is strong statistical evidence for a correlation between performed agogics and the Rubato-generated analytical data, a more quantitative description of the relationship could be given as follows (cf. Mazzola [2002, equation 44.8]). Using the idea of hierarchical smoothing, analytical functions derived from metric, melodic, and harmonic curves were defined, together with their first two derivatives, and some prima vista data (performance commands written in the score), yielding a 58-dimensional function $X(t)$ of onset time $t$. Denoting the logarithm of tempo by $y(t)$, statistical evidence was found for the approximate validity of the relationship

$$y(t) = \beta_0 + X^T(t) \cdot \beta,$$

where $\beta_0$ is a constant and $\beta = (\beta_1, \ldots, \beta_{58})^T \in \mathbb{R}^{58}$, hereinafter called the shaping vector.

It should be noted that equation (1) was obtained for one single composition only. Though one may conjecture that similar operators may apply to performances of classical music in general, additional data need to be analysed before any general conclusions can be drawn. In the present paper, a new data set consisting of 30 performances of Chopin’s Prélude Op. 28 No. 4 is considered. The following statistical analysis relating metric, melodic, and harmonic weights to the logarithmic tempo curves provides further support for the conjecture. The main qualitative result of this analysis is that there is strong evidence for a relationship between tempo curves and weight functions of all three types, metric, melodic, and harmonic.

The statistical analysis consists of several steps. In a first step, univariate tests provide some statistical evidence for an association between weight functions and tempo curves. The results show however also that the relationship is likely to be nonlinear, and may depend on scale and location. This, together with general theoretical considerations, motivates the approach of hierarchical regression introduced in Mazzola and Beran (1997), and Beran and Mazzola (1999a, 1999b, 2000, 2001). See also Mazzola (2002), Beran (2003), and Mazzola (2011). In comparison with the original definition in the context of Schumann’s Träumerei (Mazzola and Beran 1997) a slightly modified method is used. A minor change is that Nadaraya–Watson kernels are replaced by Priestley–Chao kernels. Since for Chopin’s Prélude onset times are equidistant, this does not make any difference. The second, more significant change is an improved variable selection procedure. Instead of relying on standard stepwise regression based on the assumption of uncorrelated residuals, a backward selection procedure is defined where temporal dependence of residuals is taken into account. This is of particular importance for Chopin’s Prélude, because residuals are likely to contain (almost) periodic components. This dependence needs to be taken into account in order to obtain correct error bounds for estimated regression parameters. Finally, compared to the operator defined in Mazzola (2002), Beran (2003), and Mazzola (2011), a simplified operator is used in that the first two derivatives of the weight functions are used directly,
without decomposing them further. In spite of this simplification, the statistical analysis yields strong evidence for the approximate validity of a suitable version of (1), including derivatives.

Beyond the statistically oriented results, some comments on philosophical and aesthetical aspects of the composition that are supported by our analysis are added (see Section 5), thereby presenting a remarkable congruence of the quantitative approach and a qualitative musical analysis in the realm of the humanities.

The paper is organized as follows. The data are described in Section 2. Univariate tests are discussed in Section 3. The hierarchical regression approach is defined and applied to median as well as individual tempo curves in Section 4. Final comments and a general discussion follow in Section 5.

2. Data

Two types of data are derived from Chopin’s Prélude: (1) analytical data regarding metric, melodic, and harmonic structures; (2) agogical data of tempo curves measured from 30 pianists.

2.1. Analytical data

To obtain analytical data, Chopin’s Prélude, coded in MIDI format, was fed into the Rubato analytical modules (rubettes) MetroRubette (local metric analysis), MeloRubette (topological melodic analysis), and HarmoRubette (Riemannian harmonic analysis). For the theoretical background and the implementation in Rubato we refer to Mazzola (2002, Parts V, VI, X). But let us give a summary of the methods implemented in those three rubettes. The common format of the resulting analytical data consists of three weight curves \( x_{\text{metric}}, x_{\text{melodic}}, x_{\text{harmonic}} : O \to \mathbb{R}_{\geq 0} \), where \( O \) denotes the set of onset times. In other words, we obtain functions that measure the metric, melodic, and harmonic “weight” of notes played at (onset) time \( t \). Denote by \( N \) the set of notes in the composition and by \( p_O : N \to O \) the projection onto \( O \).

The metrical weight \( x_{\text{metric}}(o) \) of an onset \( o \) is calculated as follows. We consider local metres \( m \subset O \), which are subsets of equally distributed onsets, i.e. \( m = \{a + td : t = 0, \ldots, l\} \), for a time period \( d > 0 \), with \( l = l(m) \) denoting the length of \( m \). We consider all maximal local metres with length \( l(m) \geq 2 \). Then we set \( x_{\text{metric}}(o) = \sum_{o \in m} l(m)^p \), where \( p \) is called the metrical profile. Specifically, the metric weight curve \( x_{\text{metric}} \) considered here was calculated using the following tuning parameters: quantization 1/8, metrical profile 2.

The melodic weight \( x_{\text{melodic}}(o) \) is calculated as follows. We first calculate the melodic weight of every note \( n \in N \). Then we set \( x_{\text{melodic}}(o) = \max_{n \in p_O^{-1}(o)} x_{\text{melodic}}(n) \). To calculate \( x_{\text{melodic}}(n) \), we take the set of all melodic motives from a selection \( \mathcal{M} \) of motives in the given set \( N \) of notes. Our selection is given by the set of motives with span less than or equal to 0.374 (where the “span” is defined as the difference between the onset of the first and the last note of the motive) and maximal motive cardinality 5. This yields a total of 20,022 motives. The melodic weight of a note \( n \) is defined as \( x_{\text{melodic}}(n) = \sum_{n \in p \in \mathcal{M}} \text{mel}(t) \), where \( \text{mel}(t) \) is the melodic weight of motive \( t \), which is defined as follows. We consider all motives \( v \in \mathcal{M} \) containing a sub-motive similar to \( t \) and call their number \( p(t) \), the “presence” of \( t \). Then we consider all motives \( \sigma \in \mathcal{M} \) being similar to a submotive of \( t \) and call their number \( c(t) \), the “content” of \( t \). Then we set \( \text{mel}(t) = p(t) \cdot c(t) \). The meaning of “similarity” is a topological one. Our setting of trivial symmetry group, rigid gestalt paradigm, and motive neighbourhood 2 refers to this concept.

The harmonic weights \( x_{\text{harmonic}}(o) \) of onsets \( o \in O \) are more complex; we give a short description. We first consider the set \( \mathcal{C} \) of all chords in \( N \), one (denoted by \( c(o) \)) for each onset \( o \). This setup is called “duration-included calculation.” For each chord \( c \) we calculate its Riemann matrix whose entries are the harmonic function values of \( c \) with respect to tonality \( \tau \) and Riemann value.
\( \rho \) (in our setup, Riemann values are three major and three minor values: tonic, dominant, and subdominant). A harmonic path \( p \) is a sequence \( \left( p(o) \right)_{o \in O} \) of entries \( p(o) = (\tau(o), \rho(o)) \) of the Riemann matrix at onset \( o \). The values of the Riemann matrix at \( o \) for entries \( p(o) \), together with harmonic transition weights for the movements from an onset \( o \) to its successor \( o' \), define the weight of the path \( p \). We then calculate a best path \( p_{\text{max}} \) with lowest weight, which roughly means that the harmonic transitions are as easy as possible. The weight of a chord \( c(o) \) is defined to be the weight of a best path from the beginning up to onset \( o \). Since such a calculation is too difficult for present computer power, we restrict our calculations to paths that start at causal depth \( cd \) before the onset \( o \) and end at final depth \( fd \) after \( o \). In our calculations we have chosen \( fd = cd = 1 \).

2.2. Tempo data

The tempo data were obtained from 30 CD recordings. Table 1 shows the list of all performances. Each recording was imported from CD to the computer hard drive and then processed by the digital audio workstation Audiosculpt. Each recording was normalized to \(-1\) dBFS. The process of audio normalization has usually been performed in advance by the music producer/engineer in every recording. The goal of a second normalization was to increase the average amplitude of

| Pianist              | \( t \in G \) | \( t \in G^0 \) |
|---------------------|------------|------------|
|                     | \( R^2 \)  | \( p_{\text{total}} \) | \( p_{\text{weights}} \) | \( p_{\text{deriv}} \) | \( R^2 \)  | \( p_{\text{total}} \) | \( p_{\text{weights}} \) | \( p_{\text{deriv}} \) |
| Martha Argerich, 2008 | 0.72       | 21         | 12        | 2         | 0.68       | 16         | 10        | 2         |
| Sheila Arnold, 2009  | 0.73       | 11         | 7         | 2         | 0.51       | 11         | 7         | 1         |
| Claudio Arrau, 1974  | 0.52       | 9          | 6         | 1         | 0.50       | 13         | 9         | 3         |
| Tzimon Barto, 2003   | 0.79       | 15         | 5         | 1         | 0.74       | 21         | 9         | 1         |
| Idil Biret, 1992     | 0.74       | 19         | 10        | 2         | 0.63       | 15         | 6         | 1         |
| Alfred Cortot, 2008  | 0.60       | 15         | 7         | 0         | 0.58       | 18         | 9         | 2         |
| Jeanne-Marie Darre, 1960 | 0.72    | 15         | 7         | 0         | 0.63       | 15         | 7         | 1         |
| Nikolai Demidenko, 2008 | 0.62     | 21         | 12        | 4         | 0.54       | 12         | 7         | 2         |
| Jakob Gimpel, 1976   | 0.72       | 21         | 11        | 2         | 0.56       | 20         | 9         | 1         |
| Leticia Gomez-Tagle, 2010 | 0.75   | 17         | 7         | 2         | 0.50       | 10         | 3         | 0         |
| Rudolf Kerner, 1998  | 0.57       | 9          | 6         | 0         | 0.49       | 8          | 6         | 0         |
| Walter Kien, 2013    | 0.59       | 14         | 9         | 3         | 0.49       | 11         | 6         | 2         |
| Nikolai Lugansky, 2002 | 0.67      | 17         | 4         | 0         | 0.61       | 20         | 7         | 0         |
| Juan Carlos Martins, 1995 | 0.66     | 15         | 11        | 3         | 0.50       | 15         | 7         | 1         |
| Ivan Moravec, 1993   | 0.60       | 17         | 5         | 0         | 0.46       | 16         | 5         | 0         |
| Natalia Nikolai, 2010 | 0.59      | 14         | 8         | 1         | 0.45       | 8          | 3         | 1         |
| Garrick Ohlsson, 2008 | 0.58       | 14         | 6         | 2         | 0.60       | 16         | 6         | 1         |
| Alain Planes, 2001   | 0.71       | 20         | 11        | 1         | 0.68       | 18         | 10        | 1         |
| Maurizio Pollini, 2012 | 0.74     | 14         | 6         | 2         | 0.65       | 22         | 10        | 1         |
| James Rhodes, 2010   | 0.66       | 13         | 1         | 0         | 0.60       | 13         | 4         | 0         |
| Sviatoslav Richter, 2008 | 0.53    | 17         | 8         | 1         | 0.47       | 18         | 7         | 1         |
| Artur Rubenstein, 1991 | 0.41     | 8          | 2         | 0         | 0.34       | 12         | 2         | 2         |
| Russell Sherman, 1986 | 0.65       | 12         | 6         | 1         | 0.46       | 15         | 8         | 1         |
| Grigory Sokolov, 1990 | 0.72       | 16         | 9         | 1         | 0.65       | 16         | 8         | 0         |
| Jean-Yves Thibaudet, 1999 | 0.77     | 8          | 3         | 0         | 0.79       | 12         | 5         | 0         |
| Paris Tsenikoglou, 2012 | 0.58     | 18         | 9         | 1         | 0.52       | 13         | 7         | 1         |
| Hélène Tysman, 2008  | 0.79       | 20         | 10        | 1         | 0.58       | 17         | 9         | 2         |
| Lily Williams, 2006  | 0.84       | 15         | 7         | 1         | 0.63       | 14         | 6         | 0         |
| Irina Zaritskaya, 1988 | 0.77      | 17         | 8         | 2         | 0.63       | 13         | 5         | 1         |
| Houchen Zhang, 2009  | 0.60       | 10         | 4         | 0         | 0.53       | 11         | 5         | 0         |

Note: \( R^2 = ? \); \( p_{\text{total}} \) = total number of selected explanatory variables; \( p_{\text{weights}} \) = number of "weight related" variables (including derivatives); \( p_{\text{deriv}} \) = number of variables defined by derivatives of weight functions.
J. Beran, R. Goswitz, G. Mazzola, and P. Mazzola

This is necessary because, during the mastering process, each recording is viewed in the context of the larger group of recordings that will comprise the album. A good producer or engineer will therefore normalize each recording relative to the others to create a cohesive sound. By performing a second normalization, the average amplitude of each transient was increased beyond the original recordings. The second step was to detect transients. Within Audiosculpt there is a function called “generate markers.” This process involves FFT analysis that detects the frequency of the waveform. The function was set to detect a possible transient every 20 ms, allowing the computer to identify almost every note of the recording and then place a marker at the beginning of waveforms. In a third step we verified every marker and placed markers where a softer transient was ignored by the computer. In the visualization provided by Audiosculpt, waveforms are broken up into different colours to accentuate different pitches and volumes. This increases the accuracy of hand-placed markers because each pitch is separated visually and sonically. If the computer did not detect a soft transient, that section could be highlighted and repeated until the transient was detected. To ensure accuracy further, each highlighted section was systematically shortened until the desired pitch was not audible. It was then far more certain at which millisecond each transient occurred. The final step within Audiosculpt was to export the markers as a text file. The text file contains a list in numerical order of each marker and its corresponding millisecond value.

Throughout this paper, the logarithm of the original tempo measurements will be considered. Thus, “tempo” will always refer to the logarithmic values. Figure 1(b) shows all 30 tempo curves. The mean of all 30 curves is shown in Figure 1(c). Figure 1(d) displays the standard deviation curve, obtained by calculating the standard deviation of the tempo for each onset time separately. This curve indicates where the main differences between the curves occur.

The onset times for the tempo measurements correspond to the regular grid \( G \) given in the score, i.e. \( G = \{ t_j = \frac{3}{4} + (j - 1) \frac{1}{8} \mid j = 1, \ldots, 202 \} \), where \( n = 202 \). The original metric, melodic, and harmonic weights were calculated for irregularly spaced grids that differed partially from \( G \). In a first step, the original weights were replaced by weights on the grid \( G \), with values obtained by linear interpolation. The resulting curves \( x_{\text{metric}}(t) \), \( x_{\text{melodic}}(t) \), \( x_{\text{harmonic}}(t) \) \((t \in G)\) are shown in Figure 2(a).

### 3. Exploratory analysis and univariate tests

In a first step, we consider some univariate comparisons that should indicate whether there may be any relationship between the observed tempo curves and the metric, melodic, and harmonic weights. Since here the shape of tempo curves, and not the absolute values nor the scale, are of interest, the measurements are standardized (to mean zero and variance one) for each of the 30 performances.

Let \( x(t) \) be one of the three weight functions, and \( y_k(t) \) \((k = 1, 2, \ldots, N)\) the \( N = 30 \) tempo curves. Denote by \( q(\alpha) \) the \( \alpha \)-quantile of \( x(t) \) where \( \alpha \in (0, \frac{1}{2}) \), and define

\[
G_1(\alpha; G) = \{ t \in G \mid x(t) \leq q(\alpha) \},
\]

\[
G_2(\alpha; G) = \{ t \in G \mid x(t) \geq q(1 - \alpha) \}.
\]

Recall that, for a univariate random variable \( X \), an \( \alpha \)-quantile is defined as the smallest value \( q \in \mathbb{R} \) such that \( P(X \leq q) \geq \alpha \). Thus, \( G_1 \) and \( G_2 \) are onset times where \( x(t) \) is below and above (or equal to) its \( \alpha \)- and \( (1 - \alpha) \)-quantile, respectively. If \( x(t) \) has any influence on the tempo

---

1 A transient refers to the very first compression of air from a given sound, the very beginning of a waveform.
Figure 2. (a) metric, melodic, and harmonic weights; (b), (c), and (d) show the components $x_1(t), \ldots, x_6(t)$ for the metric, melodic, and harmonic weights respectively.

curves, then this is likely to be visible in a comparison of $y_k(t) (t \in G_1(\alpha))$ with $y_k(t) (t \in G_2(\alpha))$, provided that $\alpha$ is not too close to $\frac{1}{2}$. For each $k \in \{1, 2, \ldots, N\}$, we therefore define

$$\bar{y}_{k,1} = \frac{1}{n_1} \sum_{t_j \in G_1(\alpha)} y_k(t_j),$$

$$\bar{y}_{k,2} = \frac{1}{n_2} \sum_{t_j \in G_2(\alpha)} y_k(t_j)$$

and

$$\Delta_k = \bar{y}_{k,1} - \bar{y}_{k,2},$$

where $n_1$ and $n_2$ denote the number of onset times in $G_1$ and $G_2$, respectively. The statistics $\bar{y}_{k,l}$ are sampled from two populations with expected values $\mu_l (l = 1, 2)$. Since the different tempo curves are independent from each other, the differences $\Delta_k (k = 1, 2, \ldots, N)$ are independent.
identically distributed (iid) realizations of a random variable with expected value $\mu_1 - \mu_2$. The null hypothesis

$$H_0 : \mu_1 = \mu_2$$

against the alternative $H_1 : \mu_1 \neq \mu_2$ can therefore be tested using a standard test. Here, the Wilcoxon signed-rank test is applied to $\Delta_k (k = 1, 2, \ldots, N)$ based on $x_{\text{metric}}(t)$, $x_{\text{melodic}}(t)$, and $x_{\text{harmonic}}(t)$, respectively, and $\alpha = 0.25$. In all three cases, all values of $\Delta_k$ are negative so that the $p$-values of the two-sided tests are zero. Thus, there is very strong evidence that, when considering the whole composition, tempo in regions with relatively low values of the metric, melodic, or harmonic weights tend, respectively, to be lower than in regions with high values of the weights. This is illustrated by boxplots of $\Delta_{k,1}$ (metric), $\Delta_{k,2}$ (melodic), and $\Delta_{k,3}$ (harmonic) in Figure 3(a). However, it should be noted that, as shown in Figures 3(b), 3(c), and 3(d), each

![Figure 3](image-url)

Figure 3. (a) Boxplots of $\Delta_{k,i}$ ($i = 1, 2, 3$) based on $t \in G$; (b), (c), and (d) show points that correspond to $G_{i,1}$ and $G_{i,2}$ ($i = 1, 2, 3$).
of the sets $G_{i,1}$ ($i = 1, 2, 3$) consists predominantly of onset times at the end of the piece, as well as onset times at the very beginning. A ritardando towards the end may stem from a general principle of musical performance, and may not necessarily be linked directly to the structure characterized by the weight functions used here. Therefore, to avoid the dominance of “border” effects, the same test is applied to $y_{i}(t)$ ($t \in G_{i,1}(\alpha; G^{0})$) versus $y_{i}(t)$ ($t \in G_{i,2}(\alpha; G^{0})$), where

$$G^{0} = G \cap [1, 22.75].$$

This means that time points at the border, $t \in [0.75, 1) \cup (22.875, 25.875]$, are excluded. In this case, all $p$-values are again below 5% (0.04, 0.00002, 0.00002); however, for $x_{\text{metric}}$ the opposite effect is observed. This is illustrated by boxplots in Figure 4(a). Figures 4(b), 4(c), and 4(d) show that this time the sets $G_{i,1}$ ($i = 1, 2, 3$) are quite different for the three weight functions. In particular, $G_{2,1}$ (melodic weights) contains most time points in the region around bars 12 and
13. The reason is that the melodic line in bar 12 deviates considerably from most of the other sections of the piece.

A third comparison is made considering bars 1 to 8 only, i.e. \( t \in G_{i,1}(\alpha; \tilde{G}_0) \) and \( G_{i,2}(\alpha; \tilde{G}_0) \) with \( \tilde{G}_0 = G \cap [1, 8.875] \). This corresponds to the most “homogeneous” part of the composition – see Figure 1(a). Again, using the two-sided Wilcoxon signed-rank test, significant results at the 5% level are found (with \( p \)-values of 0.003 and 0) for the melodic and harmonic weights, not however for the metric function. Moreover, this time the effect for \( x_{\text{melodic}} \) is reversed – see Figure 5(a). The locations of the respective time points are shown in Figures 5(b), 5(c), and 5(d).

For comparison, the remaining bars 9 to 22, i.e. \( t \in G_{i,1}(\alpha; \tilde{G}_0) \) and \( G_{i,2}(\alpha; \tilde{G}_0) \) with \( \tilde{G}_0 = G \cap [9, 22.75] \), are considered in Figures 6(a)–6(d). Here, significant results are found (with
Figure 6. (a) Boxplots of $\Delta_{k,i}$ ($i = 1, 2, 3$) based on $t \in \tilde{G}^0$; (b), (c), and (d) show points that correspond to the two extreme sets.

$p$-values of 0.006 and 0.00003) for $x_{\text{metric}}$ and $x_{\text{melodic}}$, whereas no significant difference is seen for $x_{\text{harmonic}}$. Moreover, for the metric weights the effect is reversed, i.e. tempo is higher for low values of $x_{\text{metric}}$, see Figure 6(a). Figures 6(b), 6(c), and 6(d) show the corresponding locations. Note that a large portion of $\tilde{G}^0$ consists of points in bars 12 and 13, see Figure 6(c). Finally, bars 9 to 22 are considered with $\alpha$ changed to 0.05. In this case, all three comparisons point in the direction of reduced tempo for lower values of $x(t)$, with a significant two-sided $p$-value ($p = 0.03$) for $x_{\text{harmonic}}$ only. The results are illustrated in Figures 7(a)–7(d).

In summary, there is reasonably strong evidence for an association between metric, melodic, and harmonic weights and tempo curves.

The functional relationship however seems to be quite complicated, and possibly nonlinear. This also explains why a standard cross-correlation analysis does not detect these dependencies. The most consistent result is found for harmonic weights, with tempo values always significantly lower in regions with low weights. For $x_{\text{metric}}$ and $x_{\text{melodic}}$ the relationship is more complex, and
appears to depend on scale and localization. It should be noted however that univariate comparisons do not necessarily give a full picture when a response variable depends simultaneously on several explanatory variables. This comment applies in particular when explanatory variables are highly correlated. This is indeed the case here, with sample correlations of \( \text{corr}(x_{\text{metric}}, x_{\text{melodic}}) = 0.62, \text{corr}(x_{\text{metric}}, x_{\text{harmonic}}) = 0.46, \text{and corr}(x_{\text{melodic}}, x_{\text{melodic}}) = 0.52. \)

4. **Hierarchical smoothing and regression models**

4.1. **Motivation**

Based on the findings above, relating tempo measurements to the three weight functions directly by a standard linear regression model does not seem to be promising. This can also be seen
empirically by regressing each of the 30 curves, $y_k(t)$ ($t \in G$), on the three explanatory variables $x_{\text{metric}}(t)$, $x_{\text{melodic}}(t)$, and $x_{\text{harmonic}}(t)$. The resulting values of $R^2$ vary between 0.019 and 0.52 with a median value of 0.23. Thus, the maximal explanatory power does not go far beyond 50%, and is virtually non-existent for many of the performances. A possible reason could be seen already in the previous section: the relationship between weights and tempo appears to be more complex, possibly nonlinear and dependent on scale, location, or both.

To obtain a better model, we adopt the approach of hierarchical smoothing introduced in Beran and Mazzola (1999a, 1999b, 2000, 2001). See also Mazzola and Beran (1997), Beran (2003), and Mazzola (2002, 2011). The idea is that the structure of a composition consists of a hierarchy of partial structures at different scales and locations. The weight functions that are supposed to characterize at least parts of the inherent structure can therefore be understood best when decomposed into components at different scales. For Chopin’s Prélude No. 4, natural scales are even multiples of 1/8. Specifically, for each of the three curves, $x_{\text{metric}}(t)$, $x_{\text{melodic}}(t)$, and $x_{\text{harmonic}}(t)$, the following decomposition was applied. Let

$$K(u) = \frac{1}{2} 1(|u| \leq 1).$$

where 1("statement") is 1 if the statement is true, and is 0 if the statement is false. Denote by $n = |G|$ the number of onset times, and define, for any $b > 0$, the matrix

$$K_b = \left[ K\left( \frac{t_i - t_j}{b} \right) \right]_{i,j=1,...,n}.$$

Then the decomposition of a function $x(t)$ ($t \in G$), or equivalently the vector $x = [x(t_1), \ldots, x(t_n)]^T$, is defined by

$$x = \sum_{s=1}^{6} x_s$$

with

$$x_1 = K_{b_1} \cdot x,$$

$$x_s = K_{b_s} \cdot \left( x - \sum_{r=1}^{s-1} x_r \right), \quad (2 \leq s \leq 6),$$

and bandwidths

$$b_s = \frac{2^{6-s}}{8} \quad (1 \leq s \leq 5), \quad b_6 = 0$$

(with the convention $0/0 := 0$ for $b = 0$). The components $x_1(t), \ldots, x_6(t)$ ($t \in G$) obtained for the metric, melodic, and harmonic weights are displayed in Figures 2(b), 2(c), and 2(d), respectively. The corresponding vectors will be denoted by $x_{\text{metric, } j}$, $x_{\text{melodic, } j}$ and $x_{\text{harmonic, } j}$ ($j = 1, 2, \ldots, 6$), respectively.

### 4.2. Explicit explanatory variables

Before defining a model based on these decompositions, some general remarks should be made. The hierarchical components of the three weight functions may be interpreted as intrinsic metric, melodic, and harmonic information encoded in the onset-pitch specifications defined by the score. This may not be the only information contained in the score. For instance, obvious explicit (or prima vista) information such as ritardando, accelerando, stretta, slurs, etc. are likely to have an influence on a performance, even if the performer is a less skilled musician. The interesting
question is therefore whether highly skilled pianists incorporate additional information in their performance, other than the obvious instructions in the score. The conjecture is that additional, intrinsic information, is, at least partially, included in the weight functions and their hierarchical components. The difficulty is of course that some of the instructions in a score may coincide with the information in the weight functions, since that may be the very reason why the composer wrote them down. This means however also that, if in spite of this confounding effect one is able to detect a significant additional effect of the weight functions, then the result is a strong indication for a genuine connection between performance and weight functions, as defined here.

Specifically for Chopin’s Prélude, the following explicit information will be used.

- There is an obvious periodic structure. Natural choices for a basic period are for instance $T = 16/8 = 2$ bars or $T = 8/8 = 1$ bar. To keep it simple, we include four sine-components

  $$x_{\text{sin}, j}(t) = \sin \left( \frac{2\pi j}{16/8} t \right) \quad (j = 1, 2, 3, 4)$$

  (7)

only. Note that these are the first four basis functions in the Fourier expansion of a periodic function with period $T = 16/8$. More generally, one would need an infinite Fourier expansion. However, since our data set is relatively small, only the first four terms will be used here. We will use the notation $x_{\text{sin}, j} = \left[ x_{\text{sin}, j}(t_1), \ldots, x_{\text{sin}, j}(t_n) \right]^T$ for the vector of values at onset times $t_k \in G$.  

- To capture (in a simple manner) extreme tempo variations due to the unique melodic line in bar 12 we include the variable

  $$x_{\text{break}}(t) = |t - 13| \textbf{1}(12.625 \leq t \leq 13.125).$$

  (8)

  We will use the notation $x_{\text{break}} = \left[ x_{\text{break}}(t_1), \ldots, x_{\text{break}}(t_n) \right]^T$ for the vector of values at onset times $t_k \in G$.  

- Slurs with starting point $t_0$ and end point $t_1$ are captured by quadratic functions of the form

  $$x_{\text{slur}}(t) = \left( t - \frac{t_0 + t_1}{2} \right)^2 \textbf{1}(t_0 \leq t \leq t_1).$$

  (9)

The following slurs are included:

- $x_{\text{slur, right}, 1}(t) : t_0 = 1, t_1 = 10$,
- $x_{\text{slur, right}, 2}(t) : t_0 = 11.75, t_1 = 12.875$,
- $x_{\text{slur, right}, 3}(t) : t_0 = 13, t_1 = 19$,
- $x_{\text{slur, right}, 4}(t) : t_0 = 19.75, t_1 = 23$,
- $x_{\text{slur, left}, 1}(t) : t_0 = 1, t_1 = 12$,
- $x_{\text{slur, left}, 2}(t) : t_0 = 13, t_1 = 23$.

We will use the notation $x_{\text{slur}, j} = \left[ x_{\text{slur}, j}(t_1), \ldots, x_{\text{slur}, j}(t_n) \right]^T$ for the vector of values at onset times $t_k \in G$.  

• The *stretta* passage in bars 16 to 18, and possible "aftershocks" in bar 19 are captured by variables of the form

\[ x_{\text{stretta}, j}(t) = \left( t - \frac{t_{0,j} + t_{1,j}}{2} \right)^2 \mathbf{1}(t_{0,j} \leq t \leq t_{1,j}) \]  

(10)

with

\[ t_{0,j} = 16 + 0.5 (j - 1), \quad t_{1,j} = 16.5 + 0.5 (j - 1) \quad (j = 1, 2, \ldots, 8). \]  

(11)

We will use the notation \( x_{\text{stretta}, j} = [x_{\text{stretta}, j}(t_1), \ldots, x_{\text{stretta}, j}(t_n)]^T \) for the vector of values at onset times \( t_k \in G \).

- Possible reinforced fluctuations in the final ritardando part are captured by

\[ x_{\text{rit}, j}(t) = \left( t - \frac{t_{0,j} + t_{1,j}}{2} \right)^2 \mathbf{1}(t_{0,j} \leq t \leq t_{1,j}) \]  

(12)

with

\[ t_{0,j} = 19.5 + (j - 1), \quad t_{1,j} = 20 + (j - 1) \quad (j = 1, 2, 3, 4). \]

We will use the notation \( x_{\text{rit}, j} = [x_{\text{rit}, j}(t_1), \ldots, x_{\text{rit}, j}(t_n)]^T \) for the vector of values at onset times \( t_k \in G \).

4.3. Derivatives of weight functions

As demonstrated in Mazzola and Beran (1997) and Beran and Mazzola (1999a, 1999b, 2000, 2001), tempo curves may also depend on derivatives of weight functions. This idea is incorporated here in a simple manner by defining, for each of the three weight functions, first and second differences \( Dx, D^2x \) where \( D \) is the differencing operator characterized by

\[ Dx(t_k) = x(t_k) - x(t_{k-1}) \quad (t_k \in G). \]

The corresponding vectors with values at all \( t_k \in G \) will be denoted by \( Dx_{\text{metric}}, D^2x_{\text{metric}}, Dx_{\text{melodic}}, D^2x_{\text{melodic}}, Dx_{\text{harmonic}}, \) and \( D^2x_{\text{harmonic}} \), respectively.

4.4. The design matrix

Next, we define the design matrix used in the regression analysis below. Let

\[ X_{\text{metric}} = \begin{pmatrix} x_{\text{metric},1}, x_{\text{metric},2}, x_{\text{metric},3}, x_{\text{metric},4}, x_{\text{metric},5}, x_{\text{metric},6} \end{pmatrix}, \]

\[ X_{\text{melodic}} = \begin{pmatrix} x_{\text{melodic},1}, x_{\text{melodic},2}, x_{\text{melodic},3}, x_{\text{melodic},4}, x_{\text{melodic},5}, x_{\text{melodic},6} \end{pmatrix}, \]

\[ X_{\text{harmonic}} = \begin{pmatrix} x_{\text{harmonic},1}, x_{\text{harmonic},2}, x_{\text{harmonic},3}, x_{\text{harmonic},4}, x_{\text{harmonic},5}, x_{\text{harmonic},6} \end{pmatrix}, \]

be the three \( n \times 6 \)-matrices obtained by joining the corresponding six components in equation (3), define by

\[ X_{\text{explicit}} = \begin{pmatrix} x_{\text{sin},1}, \ldots, x_{\text{sin},3}, x_{\text{slur},\text{right},1}, \ldots, x_{\text{slur},\text{left},2}, x_{\text{stretta},1}, \ldots, x_{\text{stretta},8}, x_{\text{rit},1}, \ldots, x_{\text{rit},4} \end{pmatrix} \]

the \( n \times 23 \)-matrix with explicit variables as specified above, and by

\[ X_{\text{deriv}} = \begin{pmatrix} Dx_{\text{metric}}, D^2x_{\text{metric}}, Dx_{\text{melodic}}, D^2x_{\text{melodic}}, Dx_{\text{harmonic}}, D^2x_{\text{harmonic}} \end{pmatrix} \]

the \( n \times 6 \)-matrix with the first and second differences of the weight functions. The initial design matrix is defined by

\[ X = X_{n \times 47} = \begin{pmatrix} X_{\text{metric}}, X_{\text{melodic}}, X_{\text{harmonic}}, X_{\text{explicit}}, X_{\text{deriv}} \end{pmatrix}. \]  

(13)
To simplify readability, we will also use the notation

\[ x_{\text{metric}, 32/8} := x_{\text{metric}, 1}, \quad x_{\text{metric}, 16/8} := x_{\text{metric}, 2}, \quad \ldots \]
\[ x_{\text{melodic}, 32/8} := x_{\text{melodic}, 1}, \quad x_{\text{melodic}, 16/8} := x_{\text{melodic}, 2}, \quad \ldots \]
\[ x_{\text{harmonic}, 32/8} := x_{\text{harmonic}, 1}, \quad x_{\text{harmonic}, 16/8} := x_{\text{harmonic}, 2}, \quad \ldots \]

i.e. we write instead of the index \( s \) the bandwidth \( b_s \) as defined in equation (6), and

\[ x_{\text{stretta}, 16a} := x_{\text{stretta}, 1}, \quad x_{\text{stretta}, 16b} := x_{\text{stretta}, 2}, \quad x_{\text{stretta}, 17a} := x_{\text{stretta}, 3} \quad \ldots \]

i.e. we write the “location” instead of the index, see equation (11).

4.5. Dependence adjusted variable selection

Compared to the number of observations, the number of explanatory variables (i.e. columns of \( X \)) is quite large. Variable selection is therefore an essential part of the analysis. Standard simple variable selection procedures in linear regression include stepwise regression, likelihood based criteria such as the AIC or BIC (Akaike 1973; Schwarz 1978), or more recent methods such as LASSO (Tibshirani 1996). It should be noted however that when regressing \( y(t) (t \in G) \) on \( X \) or a subset of its columns, residuals \( \hat{\varepsilon}(t) = y(t) - \hat{y}(t) \) obtained after subtracting the regression fit \( \hat{y}(t) = X\hat{\beta} \) are likely to exhibit temporal dependence. This is not taken into account in most procedures available in standard statistics packages. Therefore, an alternative approach is used here. It essentially consists of a modified backward stepwise selection with an adjustment for possible dependence in the residual process. First, a critical level \( \alpha \in (0, 1) \) is specified. Starting with the full design matrix, a linear least squares regression is carried out and an autoregressive process of order \( p \) (denoted as AR\((p)\)-process) is fitted to the residuals series \( \hat{\varepsilon}(t) (t \in G) \), with \( p \) being chosen by minimizing the AIC. Based on the estimated AR\((p)\)-parameters, standard errors for the estimated coefficients \( \hat{\beta}_i \) and individual \( p \)-values \( p_i \) (for testing \( H_0 : \beta_i = 0 \) versus \( H_1 : \beta_i \neq 0 \)) are calculated based on these standard errors. If \( p_{\text{max}} = \max p_i > \alpha \), then the variable with \( p_i = p_{\text{max}} \) is excluded, i.e. a new design matrix \( X_1 \) is defined by removing the corresponding column. The procedure is repeated until either \( p_{\text{max}} \leq \alpha \) or no explanatory variables are left.

4.6. Regression for median tempi

In a first step, we consider the median tempo curve for \( t \in G \) and \( t \in G^0 \), respectively. For each \( t \in G \), let

\[ F_{n,t}(u) = \frac{1}{N} \sum_{k=1}^{N} 1(y_k(t) \leq u) \]

be the marginal empirical distribution function (at time \( t \)) computed from \( N \) replicates (in our case \( N = 30 \) performances). The median (tempo) curve \( y_{\text{median}}(t) (t \in G) \) is defined by

\[ y_{\text{median}}(t) = F_{n,t}^{-1} \left( \frac{1}{2} \right) = \inf \left\{ u \in \mathbb{R} \big| F_{n,t}(u) \geq \frac{1}{2} \right\}. \tag{14} \]

First we consider \( y_{\text{median}}(t) \) for all onset times \( t \in G \). Applying the dependence adjusted variable selection algorithm as defined above leads to an \( R^2 \)-value of 0.81. The excellent fit is shown
in Figure 8(a). The final selected weight variables are

\text{metric} : x_{\text{metric}, 4/8}
\text{melodic} : x_{\text{melodic}, 32/8, 16/8, 2/8, 0}.

In addition, the explicit variables

\begin{align*}
&x_{\text{sin}, 1}, x_{\text{sin}, 2}, x_{\text{sin}, 4}, x_{\text{break}}, x_{\text{slur, right}, 3}, x_{\text{slur, right}, 4}, x_{\text{stretta}, 17}, x_{\text{stretta}, 18}
\end{align*}

and the derivative

\[ D^2 x_{\text{melodic}} \]

were included. Table 2 (columns 2 through 4) gives an overview with estimated coefficients, standard deviations and \( p \)-values. Note in particular that all \( p \)-values are very low. Also, the coefficients that match are often similar in numerical value, sign and significance.

| Variable | \( t \in G \) | \( \hat{\beta}_j \) | s.d. | \( p \)-value | \( t \in G^0 \) | \( \hat{\beta}_j \) | s.d. | \( p \)-value |
|----------|---------------|-----------------|------|-------------|---------------|-----------------|------|-------------|
| \( x_{\text{metric}, 32/8} \) | \(-\) | \(-\) | \(-\) | \(-1.061 \) | \( 0.126 \) | \( 0 \) |
| \( x_{\text{metric}, 16/8} \) | \(-\) | \(-\) | \(-\) | \(-0.731 \) | \( 0.156 \) | \( 0 \) |
| \( x_{\text{metric}, 8/8} \) | \(-\) | \(-\) | \(-\) | \( 0.172 \) | \( 0.086 \) | \( 0.05 \) |
| \( x_{\text{metric}, 4/8} \) | \( 0.119 \) | \( 0.026 \) | \( 0 \) | \(-\) | \(-\) | \(-\) |
| \( x_{\text{melodic}, 32/8} \) | \( 0.427 \) | \( 0.039 \) | \( 0 \) | \( 0.281 \) | \( 0.068 \) | \( 0 \) |
| \( x_{\text{melodic}, 16/8} \) | \( 0.111 \) | \( 0.040 \) | \( 0 \) | \( 0.218 \) | \( 0.054 \) | \( 0 \) |
| \( x_{\text{melodic}, 2/8} \) | \( 0.132 \) | \( 0.034 \) | \( 0 \) | \( 0.177 \) | \( 0.039 \) | \( 0 \) |
| \( x_{\text{melodic}, 0} \) | \( 0.063 \) | \( 0.024 \) | \( 0.01 \) | \( 0.068 \) | \( 0.026 \) | \( 0.01 \) |
| \( x_{\text{harmonic}, 16/8} \) | \(-\) | \(-\) | \(-\) | \( 0.111 \) | \( 0.039 \) | \( 0 \) |
| \( x_{\text{sin}, 1} \) | \(-0.141 \) | \( 0.036 \) | \( 0 \) | \(-0.154 \) | \( 0.041 \) | \( 0 \) |
| \( x_{\text{sin}, 2} \) | \(-0.402 \) | \( 0.053 \) | \( 0 \) | \(-0.516 \) | \( 0.044 \) | \( 0 \) |
| \( x_{\text{sin}, 4} \) | \( 0.130 \) | \( 0.044 \) | \( 0 \) | \( 0.182 \) | \( 0.065 \) | \( 0 \) |
| \( x_{\text{break}} \) | \( 7.279 \) | \( 0.806 \) | \( 0 \) | \( 9.093 \) | \( 0.886 \) | \( 0 \) |
| \( x_{\text{slur, right}, 3} \) | \( 0.056 \) | \( 0.016 \) | \( 0 \) | \( 0.067 \) | \( 0.017 \) | \( 0 \) |
| \( x_{\text{slur, right}, 4} \) | \(-0.277 \) | \( 0.064 \) | \( 0 \) | \(-0.034 \) | \( 0.096 \) | \( 0 \) |
| \( x_{\text{slur, left}, 1} \) | \(-\) | \(-\) | \(-\) | \(-0.014 \) | \( 0.003 \) | \( 0 \) |
| \( x_{\text{stretta}, 17} \) | \( 9.129 \) | \( 3.889 \) | \( 0.02 \) | \( 10.458 \) | \( 3.944 \) | \( 0.01 \) |
| \( x_{\text{stretta}, 18} \) | \(-12.426 \) | \( 4.141 \) | \( 0 \) | \(-10.389 \) | \( 4.151 \) | \( 0.01 \) |

The individual contributions of the explanatory variables can be seen in Figure 8(b). The main contribution of the metric component seems to be due to the two borders, which coincide with regions of an initial accelerando and final ritardando respectively. An obvious question is therefore whether the weight variables are selected even if \( t \) is restricted to \( G^0 = G \cap [1, 22.75] \), as defined in equation (2). The resulting fit yields an \( R^2 \) of 0.75 and is shown in Figure 8(c). Individual contributions of the explanatory variables are displayed in Figure 18(d). Numerical results are given in Table 2 (columns 4 through 6). The selected variables are almost the same as before so that their inclusion in the model is not just due to “boundary effects.”

4.7. Individual regressions for \( t \in G \)

The regression results for the median curve provide some evidence for the conjecture of a relationship between tempo curves and (hierarchically decomposed) metric, melodic, and harmonic
weight functions. A possible objection is that a series consisting of \( n = 202 \) (for \( t \in G \)) or \( n = 175 \) (for \( t \in G^0 \)) observations is used to select among a relatively large number (namely \( p = 47 \)) of explanatory variables, and the \( p \)-values given in Table 2 are obtained after the selection procedure has been carried out. It is therefore important to corroborate the results by additional data. This can be done by applying the same analysis to each of the \( N = 30 \) individual tempo curves. First we consider \( y_k(t) \) \( (k = 1, 2, \ldots , N) \) for all onset times \( t \in G \). Figures 9(a)–9(e) show the individual fits. The average of all series and fitted curves respectively are displayed.

Thus, there appears to be strong evidence for an influence of the hierarchically decomposed weight functions on the tempo curves.

A possible objection to this testing approach may be that multiple tests are performed so that the level of significance actually differs from the nominal one. As a conservative correction we can apply Bonferroni’s rule. This means that, to achieve an actual level of \( \alpha \), a corrected significance level of \( \alpha^* = \alpha/K \) is used, where \( K \) is the number of tests. Thus, for \( \alpha = 0.05 \),
and $K = 47$ tests, we would use $\alpha^* = 0.05/47 \approx 0.001$. The variables with $p$-values below this threshold are

$x_{\text{metric}, 16/8}, x_{\text{melodic}, 32/8}, x_{\text{melodic}, 2/8}, x_{\text{sin}, 2}, x_{\text{sin}, 4}, x_{\text{break}}$. 

Thus, even with a very conservative statistical test there is evidence for a relationship with metric and melodic weights. The significant effect of the periodic components and $x_{\text{break}}$ is less surprising. Note also that these variables are among those selected in Table 2. (As a cautionary remark, it should be noted here that, since the weight functions are correlated, the less prominent occurrence of harmonic weights does not necessarily imply that the harmonic structure is not important.)

### 4.8. Individual regressions for $t \in G^0$

The same analysis as above is carried out for “interior points” $t \in G^0$. Figures 12(a)–12(e) show the individual fits, the average of all series and fitted curves is displayed in Figure 13. A summary of individual $R^2$-values and the numbers of selected variables is given in Table 1 (columns 6 to 8).

Applying the Wilcoxon signed-rank test to $\hat{\beta}_{1,j}, \ldots, \hat{\beta}_{p,j}$ ($p = 47$) yields the results illustrated in Figure 14. The null hypothesis that $E[\beta_j] = 0$ is rejected at $\alpha = 0.01$ for the weight variables

- metric: $x_{\text{metric}, 32/8}, x_{\text{metric}, 16/8}, x_{\text{metric}, 8/8}, x_{\text{metric}, 4/8}$
- melodic: $x_{\text{melodic}, 32/8}, x_{\text{melodic}, 16/8}, x_{\text{melodic}, 4/8}, x_{\text{melodic}, 2/8}, x_{\text{melodic}, 0}$
- harmonic: $x_{\text{harmonic}, 16/8}$

the explicit variables

$x_{\text{sin}, 1}, x_{\text{sin}, 2}, x_{\text{sin}, 4}, x_{\text{break}}, x_{\text{slur, right}}, 3, x_{\text{slur, left}}, 1$

and the derivative

$D^2 x_{\text{melodic}}$.

If $\alpha = 0.05$ is used, then another five variables are added, namely:

$x_{\text{metric, 0}}, x_{\text{stretta, 17b}}, x_{\text{stretta, 19b}}, Dx_{\text{metric}}, D^2x_{\text{metric}}$.

Finally, applying a Bonferroni correction with $\alpha^* = 0.05/47$, $H_0$ is rejected for the following variables:

$x_{\text{metric}, 16/8}, x_{\text{melodic}, 32/8}, x_{\text{melodic}, 2/8}, x_{\text{sin}, 2}, x_{\text{sin}, 4}, x_{\text{break}}$.

Thus, again there is strong evidence for a relationship with the weight functions.

### 4.9. Cluster analysis

The regression results above can be used to identify clusters of performances where similarity is measured in terms of estimated coefficients. Figure 15 shows the result of a hierarchical cluster analysis (R-function hclust). In a first step, the estimated coefficients (for $t \in G^0$) were standardized to zero mean and variance one. The reason is that results should not depend on the scaling of explanatory variables. Only those coefficients were included for which the null hypothesis $H_0 : E[\beta_j] = 0$ was rejected by the Wilcoxon signed-rank test at $\alpha = 0.05$. The cluster analysis
was then based on Euclidean distances between the reduced parameter estimates. The following two main clusters can be identified.

- Cluster 1: Alfred Cortot, 2008; Sviatoslav Richter, 2008; Hélène Tysman, 2008; Claudio Arrau, 1974; Natalia Nikolai, 2010; Grigory Sokolov, 1990; Jeanne-Marie Darre, 1960; Rudolf Kerer, 1998; Juan Carlos Martins, 1995; Idil Biret, 1992; Llŷr Williams, 2006; Irina Zaritzkaya, 1988; Walter Klien, 2013; Nikolai Demidenko, 2008.

- Cluster 2: Martha Argerich, 2008; Houchen Zhang, 2009; Nikolai Lugansky, 2002; Jean-Yves Thibaudet, 1999; Alain Planes, 2001; Leticia Gomez-Tagle, 2010; Garrick Ohlsson, 2008; James Rhodes, 2010; Artur Rubenstein, 1991; Ivan Moravec, 1993; Jakob Gimpel, 1976; Sheila Arnold, 2009; Paris Tsenikoglou, 2012; Tzimon Barto, 2003; Maurizio Pollini, 2012; Russell Sherman, 1986.

To what extent this is in correspondence with the subjective experience when listening to the performances is an interesting question worth pursuing in future research. Here one can observe for instance that Thibaudet being present in the same cluster as Rubenstein is plausible, since Rubenstein was one of the pianists who had impressed Thibaudet most in his youth. Also, for instance, Jakob Gimpel and Russell Sherman had shared Eduard Steuermann as their teacher and are very close in the cluster hierarchy. Further observations along this line are omitted due to lack of space.

The tempo curves, together with corresponding cross-sectional means and mean fits, are displayed for the two clusters separately in Figure 16. The curves $y_k(t)$ seem to have cluster-specific properties. This can be made more visible by considering second derivatives. For each tempo curve, second derivatives are calculated numerically after first extrapolating the curves by cubic (interpolation) splines to a finer grid of onset times. Figure 17 shows $(y''_k(t))^2$ for the two clusters and the first eight bars (i.e., $t < 10$). Also shown are cross-sectional means for each cluster (red line). There is an obvious difference between the clusters. In particular, the (absolute value of) the curvature is lower in the second cluster. (Note that $y''_k$ corresponds to the curvature of $y_k$.) This is also illustrated in Figure 18(b) with boxplots of the integrated squared curvatures

$$\kappa = \int_1^{10} [y''_k(t)]^2 \, dt.$$

A similar pattern can be seen for integrated squared derivatives, $\int [y'_k(t)]^2 \, dt$ in Figure 18(a).

Alternative clusters can be obtained by other criteria. For instance, we may ask the question how many of the first derivatives $D_{x\text{metric}}, D_{x\text{melodic}}, D_{x\text{harmonic}}$ have been selected. As it turns out, for any of the 30 tempo curves the answer is either 0 or 1. This yields the following two clusters.

- Cluster 1 (no first derivative included): Rudolf Kerer, 1998; Idil Biret, 1992; Llŷr Williams, 2006; Leticia Gomez-Tagle, 2010; Irina Zaritzkaya, 1988; Juan Carlos Martins, 1995; Sheila Arnold, 2009; Paris Tsenikoglou, 2012; Grigory Sokolov, 1990; Garrick Ohlsson, 2008; Ivan Moravec, 1993; Houchen Zhang, 2009; James Rhodes, 2010; Nikolai Lugansky, 2002; Martha Argerich, 2008; Alfred Cortot, 2008; Tzimon Barto, 2003; Maurizio Pollini, 2012; Sviatoslav Richter, 2008; Jean-Yves Thibaudet, 1999; Alain Planes, 2001; Russell Sherman, 1986.

- Cluster 2 (first derivative included): Walter Klien, 2013; Jeanne-Marie Darre, 1960; Jakob Gimpel, 1976; Nikolai Demidenko, 2008; Natalia Nikolai, 2010; Hélène Tysman, 2008; Artur Rubenstein, 1991; Claudio Arrau, 1974.

It is interesting to see that these two clusters also differ distinctly with respect to curvature. The boxplots in Figure 21 show that $\kappa$ based on the first eight bars is generally higher in cluster 2. This is confirmed by a two-sample Wilcoxon test with a two-sided $p$-value of 0.027. Also, the
non-parametric estimate of the probability density function of $\kappa$ displayed in Figure 22 indicates a bimodal distribution. The circles and triangles in Figure 22 represent points from clusters 1 and 2, respectively.

4.10. Additional remarks

In the analysis considered here, the residuals in the regression model are assumed to be independent between the 30 performances. This assumption is realistic, if all relevant explanatory variables have been included. Non-parametric estimates of the coherence functions between the residual series yielded mostly small values, thus providing only weak evidence against the hypothesis of zero cross-correlation. Estimated values of the coherences between the residuals of the different performances are very small (mostly not beyond 0.2) so that the effect, if any, on the analysis discussed above can be expected to be small. Finding additional explanatory variables may improve the results further. This is the subject of current research.

A further comment should be made regarding the Wilcoxon signed-rank test for regression coefficients that are set equal to zero when not selected by the iterative procedure. To check where possible ties may influence the results we also carried out an adjusted Wilcoxon signed-rank test that allows for ties. The results turned out to be almost the same. Finally note that the common approach of a longitudinal data analysis was not suitable here, because the stochastic structure of the residuals differs substantially among the 30 performances.

5. Philosophical comments relating to the results

A general philosophical interpretation of Chopin’s Prélude that is, at first sight, completely different from the quantitative analysis described above and also somewhat speculative, can be given as follows.

In this short composition by Chopin, we can experience in realtime how an initially floating and vaguely searching voice unfolds its consciousness in a growing melodic autonomy. This illustrates to our ears Arthur Schopenhauer’s idea of a human consciousness as reified in the melody. To Schopenhauer, the harmonic series with its integer multiples of interval relations represents the building block of the evolutionary universe which in its limitation to harmony does not allow for self-awareness yet. The latter only appears through melody (in soprano), which represents the peak of the thinking mind. But it is more than mere representation, it is the very substance, “thinking music” so to speak. This exemplifies Schopenhauer’s idea that in music, the essence of will is immediately present, more than in other arts, where only the will’s shadow appears. Johann Wolfgang von Goethe supported these considerations with his idea that the arts do not only map nature but also act similarly to nature, see also Staiger (1956). This general philosophy immediately points to the Sanskrit nada brahma (the world is sound), where the world’s beginning and ending are rooted in sound. Schopenhauer even points out the extra-temporal nature of musical sound, whose existence is given beyond the world’s substance.

In Chopin’s Prélude we witness the unfolding of self-awareness “in nuce”: in the first eight measures, the soul floats without real melody “over the waters” in search of its (re)incarnation. What initiates a movement towards a comprehensive self-awareness in measures 9–12 densifies towards a self-awareness of the ego as a subjective peak (principio individuationis) with a short

---

2 Exposed in his chief opus *Die Welt als Wille und Vorstellung*, Das Objekt der Kunst, Section 52, von Löhneysen (1986).

3 This coincides with Gottfried Wilhelm Hegel’s thought that God is not an accomplished being but the worldspirit’s (“Weltgeist”) awakening that defines itself through the evolution of intelligent human subjects.
and dramatic climax after a non-melodic recapitulation in measures 13–15. It now represents a fusion of the self (ego) with its object (nature). After that (measures 16–20) the soul has evolved into a thinking subject before it dissolves into eternity in measure 21. Concluding this short philosophical excursus, we find ourselves confronted with the big mystery, similar to Goethe’s Gretchen, who remained untouchable for Faust and Mephistopheles. Chopin’s Prélude is at the same time a musical and “divine” revelation in its reconciliation of ratio and emotio in their mediation by the performative action. This again harmonizes with Leibniz’s statement cited in our Introduction.

Analogies to the philosophical, and at first sight perhaps somewhat speculative, description can be found in the quantitative analysis above. For instance, consider the regression for the median tempo function, denoted here by \( y(t) \) \((t \in G)\). The explanatory variables chosen using a significance level of \( \alpha = 0.05 \) were \( x_{\text{metric},4/8}, x_{\text{melodic},32/8}, x_{\text{melodic},16/8}, x_{\text{melodic},2/8}, x_{\text{melodic},0}, x_{\text{sin},1}, x_{\text{sin},2}, x_{\text{sin},4}, x_{\text{break}}, x_{\text{slur,right},3}, x_{\text{slur,right},4}, x_{\text{stretta},17b}, x_{\text{stretta},18a}, \text{and } D^2x_{\text{melodic}}\). Denote by \( X_{\text{med}} \) the \( n \times 14 \)-matrix (with \( n = 202 = |G| \)) with columns given by these 14 functions (in the sequence given here) evaluated at onset times \( t_k \in G \), and by \( \hat{\beta}_{\text{med}} = (\hat{\beta}_{\text{med},1}, \ldots, \hat{\beta}_{\text{med},14})^T \), the corresponding vector of estimated coefficients. Furthermore, denote by \( \hat{y}(t) \) \((t \in G)\) the fitted regression function. The contribution of the melodic variables is given by

\[
\hat{y}_{\text{melod}}(t) = \hat{y}_{\text{melod},1}(t) + \hat{y}_{\text{melod},2}(t) + \hat{y}_{\text{melod},3}(t) + \hat{y}_{\text{melod},4}(t) + \hat{y}_{\text{melod},5}(t)
\]

with

\[
\hat{y}_{\text{melod},1}(t) = \hat{\beta}_{\text{med},2}x_{\text{melodic},32/8}, \quad \hat{y}_{\text{melod},2}(t) = \hat{\beta}_{\text{med},3}x_{\text{melodic},16/8},
\]

\[
\hat{y}_{\text{melod},3}(t) = \hat{\beta}_{\text{med},4}x_{\text{melodic},2/8}, \quad \hat{y}_{\text{melod},4}(t) = \hat{\beta}_{\text{med},5}x_{\text{melodic},0},
\]

\[
\hat{y}_{\text{melod},5}(t) = \hat{\beta}_{\text{med},14}D^2x_{\text{melodic}}.
\]

The contribution of the other (non-melodic) variables will be denoted by \( \hat{y}_{\text{other}}(t) := \hat{y}(t) - \hat{y}_{\text{melod}} \). Figure 23 shows \( y_{\text{med}}(t) := y(t) - \hat{y}_{\text{other}}(t) \) and \( \hat{y}_{\text{melod}}(t) = \hat{y}(t) - \hat{y}_{\text{other}} \) (the two upper curves, in black and red, respectively), together with \( \hat{y}_{\text{melod},1}(t) + \hat{y}_{\text{melod},2}(t) \) (the second black curve from the top), and \( \sum_{j=1}^{5} \hat{y}_{\text{melod},j}(t) = \hat{y}_{\text{melod}}(t) \) (the lowest curve, in red). The curves are shifted vertically for better visibility. The tempo curve exhibits a dip around measure 13 which coincides with the non-melodic recapitulation mentioned above and is well explained by \( x_{\text{melodic},32/8} \) and \( x_{\text{melodic},16/8} \). This is followed by a climax culminating at the beginning of measure 16 and slowly dissolving “into eternity” in measure 21. Here also, the five melodic components provide an amazingly good “explanation.” Also remarkable is that adding the more detailed melodic curves \( x_{\text{melodic},2/8}, x_{\text{melodic},0}, \text{and } D^2x_{\text{melodic}} \) helps significantly in matching more local maxima and dips of the tempo curve. In summary, after adjusting for “non-melodic” components, the melodic components extracted analytically from the score are in very good agreement with the (equally adjusted) median tempo curve, and with the philosophical interpretation given above.

**Acknowledgements**

We thank Thomas Fiore for his editorial work, both in managing the referee process and in editing the typography. We would like to thank two referees for their insightful and constructive comments and suggestions.

**Supplemental online material**

Supplemental online material for this article can be accessed at http://dx.doi.org/10.1080/17459737.2014.930192. For convenience, the captions of the figures available in the online supplement are given after the references section herein.
References

Akaike, H. 1973. “Information Theory and an Extension of the Maximum Likelihood Principle.” In: Proceedings of the Second International Symposium on Information Theory, edited by F. Csáki and B. N. Petrov, 267–281. Budapest: Akadémiai.

Beran, J. 2003. Statistics in Musicology. New York: Chapman & Hall/CRC Press.

Beran, J., and G. Mazzola. 1999a. “Analyzing Musical Structure and Performance – a Statistical Approach.” Statistical Science 14 (1): 47–79.

Beran, J., and G. Mazzola. 1999b. “Visualizing the Relationship Between Two Time Series By Hierarchical Smoothing.” Journal of Computational and Graphical Statistics 8 (2): 213–238.

Beran, J., and G. Mazzola. 2000. “Timing Microstructure in Schumann’s Träumerei as an Expression of Harmony, Rhythm, and Motivic Structure in Music Performance.” Computers & Mathematics with Applications 39 (5-6): 99–130.

Beran, J., and G. Mazzola. 2001. “Musical Composition and Performance – Statistical Decomposition and Interpretation.” Student 4 (1): 13–42.

Mazzola, G. 2002. The Topos of Music: Geometric Logic of Concepts, Theory, and Performance. Basel, Switzerland: Birkhäuser.

Mazzola, G. 2011. Musical Performance: A Comprehensive Approach: Theory, Analytical Tools, and Case Studies. New York: Springer.

Mazzola, G., and J. Beran. 1997. “Rational Composition of Performance.” In: Regulation of Creative Processes in Music, edited by W. Auhagen and R. Kopiez, Staatliches Institut für Musikforschung, Berlin, 37–67. Frankfurt: Lang Verlag.

Mazzola, G., and O. Zahorka. 1994. “The RUBATO Performance Workstation on NeXTSTEP.” In: Proceedings of the International Computer Music Conference: The Human Touch (ICMC 94), Musikhuset Aarhus, Denmark, 12–17 September 1994. San Francisco, CA: ICMA.

Repp, B. 1992. “Diversity and Commonality in Music Performance: An Analysis of Timing Microstructure in Schumann’s ‘Träumerei’.” Journal of the Acoustic Society of America 92 (5): 2546–2568.

Schwarz, G. 1978. “Estimating the Dimension of a Model.” Annals of Statistics 2 (6): 461–464.

Sloboda, A. S., and A. C. Lehmann. 2001. “Tracking Performance Correlates of Changes in Perceived Intensity of Emotion During Different Interpretations of a Chopin Piano Prelude.” Music Perception, 19 (1): 87–120.

Staiger, E. 1956. Goethe, Vol. II. Zürich: Atlantis.

Sundberg, J., A. Askenfelt, and L. Frydén. 1983. “Musical Performance: A Synthesis-by-Rule Approach.” Computer Music Journal 7 (1): 37–43.

Tibshirani, R. 1996. “Regression Shrinkage and Selection via the Lasso.” Journal of the Royal Statistical Society, Series B 58 (1): 267–288.

von Löhnysen, Wolfgang Freiherr, Hrsg. 1986. Arthur Schopenhauers sämtliche Werke, Band I. und II. Frankfurt am Main: Suhrkamp Verlag.

Appendix A. Captions of additional figures available online as supplemental material

Figure 8. Hierarchical regression fit for median tempo: (a), (b) $t \in G$; (c), (d) $t \in G^0$.

Figure 9. Individual hierarchical regression fits (black) and observations (blue) for $t \in G$.

Figure 10. Mean of individual hierarchical regression fits (black) and observations (blue) for $t \in G$.

Figure 11. $p$-Values for testing $H_0 : \mu_j = E[\beta_j] = 0$ against $H_1 : \mu_j \neq 0$, based on $t \in G$.

Figure 12. Individual hierarchical regression fits (black) and observations (blue) for $t \in G^0$.

Figure 13. Mean of individual hierarchical regression fits (black) and observations (blue) for $t \in G^0$.

Figure 14. $p$-Values for testing $H_0 : \mu_j = E[\beta_j] = 0$ against $H_1 : \mu_j \neq 0$, based on $t \in G^0$. 
Figure 15. Cluster analysis based on coefficients that are significant using the Bonferroni correction.

Figure 16. Tempo curves and average fits for the two main clusters in Figure 15.

Figure 17. Second derivatives of tempo curves for the two main clusters in Figure 15, and the first eight bars.

Figure 18. Boxplots of integrated squared first (a) and second (b) derivatives of tempo curves, for the two main clusters in Figure 15 and the first eight bars.

Figure 19. Second derivatives of tempo curves for the two main clusters in Figure 15, and bars 16 to 18.

Figure 20. Boxplots of integrated squared first (a) and second (b) derivatives of tempo curves, for the two main clusters in Figure 15 and bars 16 to 18.

Figure 21. Boxplots of integrated squared first and second derivatives of tempo curves for the first eight bars, and two clusters defined by non-inclusion and inclusion, respectively, of derivatives in the hierarchical regression model.

Figure 22. Estimated distribution (probability density) of $\kappa$ (based on the first eight bars). Circles mark performances where no derivatives were included in the model, triangles mark the other performances.

Figure 23. The two upper curves, in black and red, respectively: $y_{\text{melod}}(t) := y(t) - \hat{y}_{\text{other}}(t)$, $\hat{y}_{\text{melod}}(t) = \hat{y}(t) - \hat{y}_{\text{other}}$. Second black curve from the top: $\hat{y}_{\text{melod}, 1}(t) + \hat{y}_{\text{melod}, 2}(t)$. The lowest curve, in red: $\sum_{j=1}^{5} \hat{y}_{\text{melod}, j}(t) = \hat{y}_{\text{melod}}(t)$. 