An Analysis of Brazilian Popular Music Based on Quantitative Metrics

Renato Fabbri,[a] Vilson Vieira,[b] Jose Eduardo Fornari Novo Junior,[c] and Luciano da Fontoura Costa[d]
Instituto de Física de São Carlos, Universidade de São Paulo (IFSC/USP)

(Dated: 28 January 2013)

We extend a methodology already used for philosophy and classical music by applying multivariate statistics on representatives, to study Brazilian popular music (MPB). Seven artists were considered in terms of eight main musical features and their historical importance on MPB movements. Grades were assigned to each characteristic and their correlations are analyzed. [DEPRECATED.. reescrever o restante] A bootstrap method is applied to simulate hundreds of artificial artists influenced by the seven representatives chosen. By applying dimensionality reduction, we obtained a multi-dimensional space used to quantify non-numeric relations like dialectics, opposition and innovation. Differences on style and technique were represented as geometrical distances in the planar space, making it possible to quantify, for example, how much X and Y differ from others or how much Z influenced W. In addition, we compared the results with a prior investigation on philosophy and european classical music[1]. The influence of dialectics, strong on philosophy, was not remarkable on music. Instead, supporting an observation already considered by classical music theorists, strong influences were identified between subsequent composers, implying inheritance and suggesting a stronger master-disciple evolution when compared to the philosophy analysis. This article focus on brazilian popular music to further explore this novel approach.

PACS numbers: 89.75.Fb,05.65.+b
Keywords: music, musicology, pattern recognition, statistics

I. INTRODUCTION

[TODO: escrever...]

II. MATHEMATICAL DESCRIPTION

A sequence S of P artists was chosen based on their relevance as representative of each movements of the Brazilian popular music. As done for philosophers and classical composers[1], the set of C measurements define a $C-$dimensional space henceforth referred as the *musical space*. The characteristic vector $\vec{v_i}$ of each artist i defines a respective *artist state* in the musical space.

For the set of P artists, we used the same elements defined for philosophers and classical composers[1]: *average state at time* i, named $\vec{a_i}$; the *opposite state* of a given artist state $\vec{v_i}$, named $\vec{r_i}$; the *opposition vector* of artist state $\vec{v_i}$, named $\vec{D_i}$; and the *opposition amplitude* of that same state, $||\vec{D_i}||$. The dialectics is quantified between a triple of successive artists i, j and k of the given set P.

[TODO: colocar 2 a 3 figuras (ou figura resumo) e tabela de equacoes]

[a] http://www.estudiolivre.org/el-user.php?view_user=gk; Electronic mail: renato.fabbri@gmail.com
[b] http://automata.cc; Electronic mail: vilson@void.cc
[c] http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=B486803; Electronic mail: tutifornari@gmail.com
[d] http://cyvision.ifsc.usp.br/~luciano/; Electronic mail: ldfcosta@gmail.com

III. MUSICAL CHARACTERISTICS

To create the musical space we derived eight variables corresponding to distinct characteristics commonly found in the music considered. The characteristics are related with the basic elements of music. All the eight characteristics are listed below:

Dance - Contemplative (D-C): a music is classified as dancing when it makes listener dance, instead of assuming a contemplative role. Brazilian popular musics are commonly dancing like *samba* and *samba-canção*, although some compositions are just contemplative, common on musics composed during the Brazilian military government.

Short duration - Long duration (S-L): compositions are quantified having short duration when they do not have more than few minutes of execution. Long duration compositions have at least 20 minutes [TODO: 20 minutos eh ainda valido?] of execution or more.

Nationalism - Internationalism (N-I): a nationalist music has a strong identification with the nation, in this case, with the Brazilian history and coustumes. Nationalists reinterpret traditional identities and reforce national characteristics, common on regionalist musics like the south Brazilian nativist music (*música gaúcha*) or the north Brazilian *forró*. Internationalism follows the oposite direction, searching for external relations from different communities, which remembers the cosmopolitanism.

Acoustic - Electronic (A-E): priority use of acoustic instruments, like acoustic guitar, flute and piano commonly used in *bossa-nova* and *choro*. [TODO: melhorar musica eletronica]

Improvisation - Non-improvisation (*I-N*): compositions that encourage improvisation in opposition to musics that inherits the tradition of classical music where the score is interpreted *per se*.

Cheerful - Melancholic (*C-M*): cheerful musics tells joyful histories and even jokes, common on *samba*. Melancholic compositions are common in romantic musics and musics written for criticism.

Rhythmic Simplicity - Rhythmic Complexity (*R-P*): presence or not of polyrhythms, the use of independent rhythms at the same time. [TODO: melhorar]

Harmonic Stability - Harmonic Variety (*T-M*): rate of tonality change along a piece or its stability. [TODO: melhorar]

IV. RESULTS AND DISCUSSION

Memorable artists were chosen as key representatives of MPB development. This group was chosen purposely to model their influence over contemporaries, creating a concise parallel with Brazilian popular music history. We modeled this group of influenced artists as new artificial samples generated by a bootstrap method, better explained in this section.

The sequence is ordered chronologically and presented on Table I with each artist.

TABLE I. The sequence of Brazilian popular music artists ordered chronologically with the outstanding period their represents.

Artists	Eras
Noel Rosa	Popular songs from early 1900
Tom Jobim	*Bossa-Nova*
Chico Buarque	Festivals of protest songs
Caetano Veloso	*Tropicália*
Roberto Carlos	*Jovem Guarda*
Djavan	Electronic MPB
Lenine	Contemporary

The quantification of the eight musical characteristics was performed jointly by the authors of this article and is shown in Table II. The scores were numerical values between 1 and 9. Values more close of 1 reveals the composer tended to the first element of each characteristic pair, and vice versa.

TABLE II. Quantification of the eight music characteristics for each of the seven artists.

Artists	D-C	S-L	N-I	A-E	I-N	C-M	R-P	T-M
Noel Rosa	1.5	2	1	1	2.5	2.5	2	1.5
Tom Jobim	4	5	5	2	5	8	8	8.5
Chico Buarque	4	7	3.5	2	7.5	5	3	5
Caetano Veloso	3	3.5	3.5	3	7	3.5	7	7.5
Roberto Carlos	7	6	8	6	9	8	1	1
Djavan	7	6	7	8.5	3.5	5	6	6
Lenine	5	7	2	6	2	6.5	9	9

This data set defines an 8-dimensional musical space, each dimension corresponding of one characteristic considering the group of 7 representatives. Such small data set is not adequate to statistical analysis. The analysis of this set would be highly biased by the small sample.

A. Bootstrap method for sampling artificial artists

[TODO: eh necessaria a discussao de bootstrap?]

To simulate a more realistic musical trajectory, we used a bootstrap method for generating *artificial artists* contemporaries of the seven chosen.

The bootstrap routine generated randomized scores $\vec{r}$. The values are not totally random, following a probability distribution that models the original $n = 7$ scores, given by $p(\vec{r}) = \sum_{i=1}^{n} e^{\frac{d_i}{2\sigma^2}}$ where d_i is the distance between a random score $\vec{r}$ and the original scores. For each step a value $p(\vec{r})$ is generated and compared with a random normalized value, characterizing the Monte Carlo[2] method to choose a set of samples. This samples simulates new randomized composers scores – while respecting the historical influence of the main 7 original exponents. Higher the value of $p(\vec{r})$, greater the influence of the original scores over $\vec{r}$. For the analysis we used 500 bootstrap samples obtained by the bootstrap process together with the original scores, considering $\sigma = 1.4$.

The Pearson correlation coefficients between the eight musical characteristics chosen are presented in Table III. The coefficients with absolute value larger than 0.5 are emphasized.

TABLE III. Pearson correlation coefficients between the eight musical characteristics.

-	D-C	S-L	N-I	A-E	I-N	C-M	R-P	T-M
D-C	-	0.35	**0.63**	**0.7**	-0.37	0.21	0.02	-0.08
S-L	-	-	0.13	0.2	0.	0.24	-0.3	-0.19
N-I	-	-	-	**0.55**	-0.23	0.2	0.06	-0.07
A-E	-	-	-	-	**-0.55**	0.04	0.18	0.03
I-N	-	-	-	-	-	-0.17	-0.34	-0.17
C-M	-	-	-	-	-	-	0.16	0.17
R-P	-	-	-	-	-	-	-	**0.76**
T-M	-	-	-	-	-	-	-	-

[TODO: identificar relacoes interessantes e discorrer sobre elas]

PCA was applied to this set of data, yielding the new variances given in Table IV in terms of percentages of total variance. We can note the concentration of variance along the four first PCA axes, a common effect also observed while analyzing philosophers and classical music composers characteristics[1]. This means we could consider just four dimensions.

TABLE IV. New variances after PCA, in percentages for scores on III.

Eigenvalue	Value
λ_1	34 %
λ_2	25 %
λ_3	14 %
λ_4	9 %
λ_5	8 %
λ_6	4 %
λ_7	3 %
λ_8	3 %

TABLE V. Average and standard deviation of the deviations for each composer and for the first 4 eigenvalues.

Artists	μ_Δ	σ_Δ
Noel Rosa	5.66	0.93
Tom Jobim	3.61	0.79
Chico Buarque	3.49	0.81
Caetano Veloso	3.44	0.79
Roberto Carlos	5.53	0.96
Djavan	3.85	0.88
Lenine	4.28	0.88

Eigenvalues	μ_Δ	σ_Δ
λ_1	-0.1965	0.0056
λ_2	-0.1135	0.0038
λ_3	-0.0073	0.0027
λ_4	0.0357	0.0025

B. Perturbations robustness of the original scores

As done for philosophers and classical composers analysis, we performed 1000 perturbations of the original scores by adding the values -2, -1, 0, 1 and 2 with uniform probability. In other words, we wanted to test if scoring errors could be sufficient to cause relevant effects on the PCA projections. Interestingly, the values of average and standard deviation for both original and perturbed positions listed in Table V show relatively small changes. It is therefore reasonable to say small errors in the values assigned as scores of composers characteristics did not affected too much its quantification.

C. Results

Table VI shows the normalized weights of the contributions of each original property on the four new main axes. Most of the characteristics contribute almost equally in defining the first two main axes.

Figure 1 presents a 2-dimensional space considering the first two main axes. The arrows follows the time sequence along with the seven artists. Each of these arrows corresponds to a musical move from one artist state to another – for clarity, just the lines of the arrows are preserved. The bootstrap samples define clusters around the original artists.

[TODO: DEPRECATED... identificar as principais caracteristicas da projecao e relata-las]

TABLE VI. Percentages of the contributions from each musical characteristic on the four new main axes.

Musical Characteristics	C_1	C_2	C_3	C_4
D-C	20.22	8.67	1.64	5.07
S-L	7.34	16.78	24.01	18.82
N-I	17.45	6.15	4.89	30.03
A-E	20.33	2.15	13.79	6.92
I-N	15.61	8.64	8.31	27.75
C-M	8.47	3.17	36.81	7.18
R-P	7.14	27.53	1.29	2.78
T-M	3.40	26.86	9.22	1.42

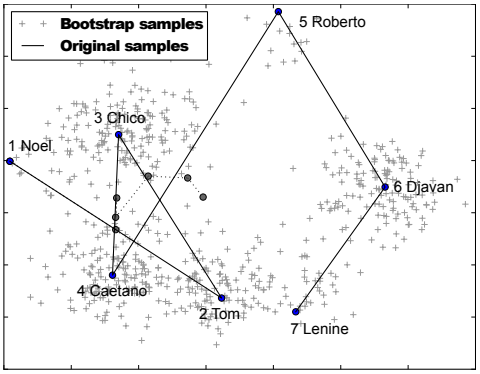

FIG. 1. *2-dimensional projected musical space.*

Bach is positioned far from the rest of composers, which suggests his key role admitted by other great composers like Beethoven and Webern[3]: "In fact Bach composed everything, concerned himself with everything that gives food for thought!". The greatest subsequent change takes place from Bach to Mozart, reflecting a substantial difference in style. We can identify a strong relationship between Beethoven and Brahms, supporting the belief by the *virtuosi* Hans von Bülow[4] when he stated the 1^{st} Symphony of Brahms as, in reality, being the 10^{th} *Symphony of Beethoven*, clamming Brahms as the true successor of Beethoven. Stravinsky is near of Beethoven and Brahms, presumably due to his heterogeneity[5,6]. Beethoven is also near of Mozart who deeply influenced Beethoven, mainly in his early works. For Webern, Beethoven was the unique classicist who really came close to the coherence found in the pieces of the Burgundian School: "Not even in Haydn and Mozart do we see these two forms as clearly as in Beethoven. The period and the eight-bar sentence are at their purest in Beethoven; in his predecessors we find only traces of them"[3]. It could explain the proximity of Beethoven to the Renaissance Monteverdi. Stockhausen is a deviating point when compared with the others and it could

present even more detachment if we had considered vanguard characteristics – e.g. timbre exploration by using electronic devices[6] – not shared by his precursors. In general, the musical movements had minor opposition and, remembering the beginning of this work, it reflects the master-apprentice tradition present in music: the composers tend to build their own works confirming their precursors legacy. This reveals a crucial difference considering the *memory treatment* along the development of philosophy and music: while a philosopher was influenced by the opposition of ideas from his two predecessors, composers were commonly influenced by their direct predecessor. Therefore, we can argue that philosophy presents a *memory-2* state, while music presents *memory-1*, considering *memory-N* being the number N of past generations those influenced a philosopher or composer. Considering the linearity of musical movements we can identify the abscissa as a "time axis" representing the development of music along the history, with some composers like Beethoven returning to Monteverdi and others advancing to the modern age like Stravinsky and Stockhausen.

[TODO: DEPRECATED... rever baseando-se na nova analise]

To complement the analysis, Table VII gives the opposition and skewness indices for each of the six musical moves, showing the movements are driven by rather small opposition and strong skewness. In other words, most musical moves do not benefit from opposition as far as innovation is concerned. Dialectics is also analyzed on Table VIII where we identified an alternation of values along the pairs of subsequent musical movements: the first value of counter-dialectics is greater than the second, that is lesser than the third and so on. There is no strong dialectics, but a continuous variation.

TABLE VII. Opposition and skewness indices for each of the six musical moves.

Musical Move	$W_{i,j}$	$s_{i,j}$
Noel Rosa → Tom Jobim	1.00	0.00
Tom Jobim → Chico Buarque	0.52	1.92
Chico Buarque → Caetano Veloso	0.64	1.66
Caetano Veloso → Roberto Carlos	1.15	2.97
Roberto Carlos → Djavan	0.43	3.04
Djavan → Lenine	0.34	3.34

TABLE VIII. Counter-dialectics index for each of the five subsequent pairs of musical moves.

Musical Triple	$d_{i\to k}$
Noel Rosa → Tom Jobim → Chico Buarque	0.04
Tom Jobim → Chico Buarque → Caetano Veloso	0.21
Chico Buarque → Caetano Veloso → Roberto Carlos	3.23
Caetano Veloso → Roberto Carlos → Djavan	0.07
Roberto Carlos → Djavan → Lenine	4.00

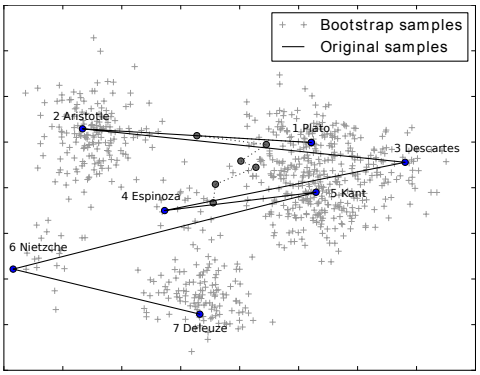

FIG. 2. *2-dimensional projected philosophical space.*

V. COMPARISONS WITH PRIOR ANALYSIS

[TODO: DEPRECATED... refazer comparacao]

The results of Brazilian popular musicians analysis when compared with philosophers and classical composers[1] reveals surprising results. It is important to note we preserved the number of characteristics and performed the same bootstrap method to generate a larger set of samples, making possible this comparison. The variances after PCA (Table IX) concentrates in the four first new axis, similar to the variances for composers shown at Table IV. If we compare the discussed musical space with the philosophical one in Figure 2 we identify opposite movements along all the philosophy history in contrast to music. This reveals a notorious characteristic of the way philosophers seem to have evolved their ideas, driven by opposition, while composers tend to be more influenced by their predecessors.

TABLE IX. New variances after PCA for philosophers scores in percentages.

Eigenvalue	Value
λ_1	?? %
λ_2	?? %
λ_3	?? %
λ_4	?? %
λ_5	?? %
λ_6	?? %
λ_7	?? %
λ_8	?? %

The opposition and skewness indices for philosophers listed in Table X endorses the minor role of opposition in composers. We can observe strong opposition and rather small skewness in philosophical moves contrasted to small opposition and strong skewness in musical movements. Also, the oscillating dialectics of both are out of phase, indicating transfer latency.

TABLE X. Opposition and skewness indices for each of the six philosophical moves.

Philosophical Move	$W_{i,j}$	$s_{i,j}$
Plato → Aristotle	1.0	0.0
Aristotle → Descartes	0.8922	0.0145
Descartes → Espinoza	0.7838	1.1218
Espinoza → Kant	0.6548	0.7059
Kant → Nietzsche	1.5124	1.4479
Nietzsche → Deleuze	0.3235	1.9040

TABLE XI. Counter-dialectics index for each of the five subsequent pairs of philosophical moves.

Philosophical Triple	$d_{i\to k}$
Plato → Aristotle → Descartes	0.968
Aristotle → Descartes → Espinoza	0.287
Descartes → Espinoza → Kant	0.138
Espinoza → Kant → Nietzsche	1.247
Kant → Nietzsche → Deleuze	0.054

When comparing dialectics, other curious facts arise: the dialectics indices in Table XI are considerably stronger philosophical moves than for composers. Both indices are also shown in Figure 3 where we can see a constantly decrease of counter-dialectics, contrasting the continuously variation of the indices when considering the composers. This makes possible to argue that dialectics is stronger in philosophy than in music where a constantly return to the origins are clearly visible on some composers. This reveals the nature of the musical development, based on the search for a unity. Using the words of Webern, the search for the "comprehensibility" but always influenced by their old masters.

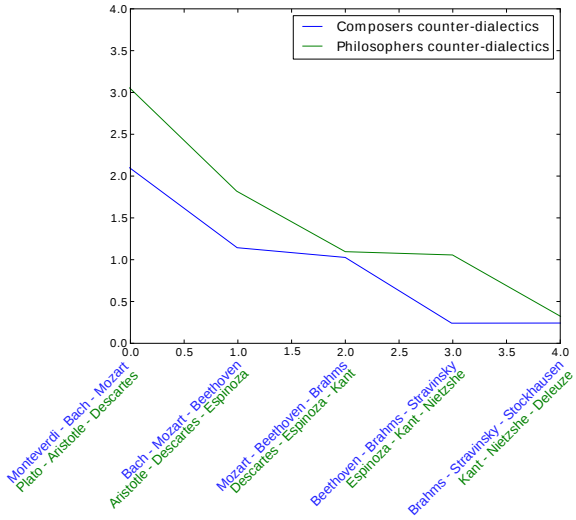

FIG. 3. *Comparison between composers and philosophers counter-dialectics indices.* *[TODO: DEPRECATED... fazer nova analise]*

VI. CONCLUDING REMARKS

[TODO: DEPRECATED... escrever]

Motivated by the understanding of how innovation evolves in music history, we extended a quantitative method recently applied to the study of philosophical characteristics[1] and compared the results. Statistical methods have been commonly applied to the study of music features and composers productivity, but analysis of composers characteristics along the music history has been less explored. The method differs on the aspect of how the characteristics concerning composers are treated: scores are assigned to each feature common in musical works. These scores reveal not the exact profile of composers, but a tendency of how their techniques relate one another. To make the simulation more realistic, we considered not just the small number of 7 composers, but derived other 500 new "artificial composers" through a bootstrap method. A larger data set made possible the statistical analysis, considering not just the original scored composers, but other samples those respect the historical tendency of the formers, modeled by a probabilistic distribution. In order to investigate the relationship between this scoring we applied Pearson correlation analysis. The results demonstrated a strong correlation between some characteristics, which allows us to group this values, creating a reduced number of features that summarizes the most important characteristics. PCA was also applied to these components, reducing the complex space to a planar graph where the most interesting properties were identified.

Historical landmarks in music are well-defined in the planar space, like the isolation of Bach, Mozart and Stockhausen, the proximity between Beethoven and Brahms and the distance from Bach and Mozart, the heterogeneity of Stravinsky and the vanguard of contemporary composers like Stockhausen. Even not so visible relations, like the trend to return to the maximum domain of polyphony – present on Renaissance – by Beethoven could also be clearly observable, demonstrating the time nature of the space.

The dichotomy between master-apprentice tradition on music and the quest for innovation that opened this discussion could be visualized quantitatively. Each composer demonstrated his own style, differing considerably from his predecessor – clearly shown when analyzing pairs of subsequent composers like Bach and Mozart, Mozart and Beethoven or Stravinsky and Stockhausen. Otherwise, the inheritance of predecessors styles is also present when analyzing the direct relations between Mozart and Beethoven or Beethoven and Brahms, or indirect ones between Bach and Beethoven or Beethoven and Monteverdi. The entire scenario presented a "continual pattern" between composers – motivated by the influence of theirs predecessors – but also showed a force repelling both of them: the innovation, or in the words of William Lovelock[6], the "experimentation" that makes progress possible.

Along the analysis we noticed interesting differences when comparing composers with philosophers. While on philosophy the innovation is notably marked by opposition of each philosophers ideas, it is less present for music composers. The lack of strong opposition movements in musical space indicates the music innovation is driven by a constant heritage of each composer from his predecessor. We represented this characteristic referring to a *memory state* where philosophers shows *memory-2* – each philosopher was influenced by the opposite ideas of its two predecessors – while composers shows *memory-1* – inheriting the style of their direct predecessor. The analysis of both dialectics values also shown surprising results: while on philosophy the dialectics indices are arranged on a increasing series – showing a strong influence of dialectics to philosophy development – the same dialectics indices on music exhibits a constantly variation. This behavior presumably indicates a constantly quest for coherence by the composers, a fact notably observed by the studies of Anton Webern[3].

The quantitative methodology initially applied to the analysis of philosophy[1] proved to be extensible to other fields of knowledge like music, reflecting with considerable efficiency, specific details concerning each field.

Computational analysis of music scores could be applied to automate the quantification of composers characteristics, like identification of melodic and harmonic patterns or the presence or not of polyrhythms, motivic and harmonic stability[7]. More composers could be inserted in the set for the analysis of a wider time-line, possibly including more representatives of each music periods.

While taking the first steps on the direction of a quantitative approach to arts and philosophy we believe that an understanding of the creative process could also be eventually quantified. We want to end this work going back to Webern, who early envisioned these relations: "It is clear that where relatedness and unity are omnipresent, comprehensibility is also guaranteed. And all the rest is dilettantism, nothing else, for all time, and always has been. That's so not only in music but everywhere."

ACKNOWLEDGMENTS

Luciano da F. Costa thanks CNPq (308231/03-1) and FAPESP (05/00587-5) for sponsorship. Vilson Vieira and Renato Fabbri is grateful to CAPES and the Postgrad Committee of the IFSC.

[1] Vilson Vieira, Renato Fabbri, Gonzalo Travieso, Osvaldo N Oliveira Jr, and Luciano da Fontoura Costa, "A quantitative approach to evolution of music and philosophy," Journal of Statistical Mechanics: Theory and Experiment **2012**, P08010 (2012), `http://stacks.iop.org/1742-5468/2012/i=08/a=P08010`.

[2] Christian P. Robert, "Simulation in statistics," in *Proceedings of the 2011 Winter Simulation Conference* (2011) arXiv:1105.4823.

[3] Anton Webern, *The Path To The New Music* (Theodore Presser Company, 1963).

[4] Alan Walker, *Hans von Bülow: a life and times* (Oxford University Press, 2010).

[5] Roy Bennett, *History of Music* (Cambridge University Press, 1982).

[6] William Lovelock, *A Concise History of Music* (Hammond Textbooks, 1962).

[7] Debora C Correa, Jose H Saito, and Luciano da F Costa, "Musical genres: beating to the rhythms of different drums," New Journal of Physics **12** (2010).

[8] G. Deleuze, *Difference and Repetition* (Continuum, 1968).

[9] Luciano da Fontoura Costa Renato Fabbri, Osvaldo N. Oliveira Jr., "Philosometrics," (October 2010), arXiv:1010.1880v2.

[10] Aaron Kozbelt, "Performance time productivity and versatility estimates for 102 classical composers," Psychology of Music **37**, 25–46 (2009).

[11] Aaron Kozbelt, "A quantitative analysis of Beethoven as self-critic: implications for psychological theories of musical creativity," Psychology of Music **35**, 144–168 (2007).

[12] Dean Keith Simonton, "Emergence and realization of genius: The lives and works of 120 classical composers," Journal of Personality and Social Psychology **61**, 829 – 840 (1991), ISSN 0022-3514.

[13] Dean K. Simonton, "Creative productivity, age, and stress: A biographical time-series analysis of 10 classical composers," Journal of Personality and Social Psychology **35**, 791 – 804 (1977), ISSN 0022-3514.

[14] Peter van Kranenburg, "Musical style recognition – a quantitative approach," in *Proceedings of the Conference on Interdisciplinary Musicology (CIM04)* (2004).

[15] Peter van Kranenburg, "On measuring musical style – the case of some disputed organ fugues in the J.S. Bach (BWV) catalogue," Computing In Musicology **15** (2007-8).