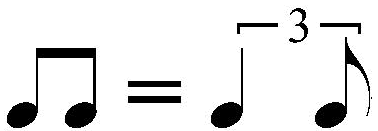
# *Playing it straight: Analysing jazz soloists’ eighth-note distributions with the Weimar Jazz Database*

by Christopher Corcoran and Klaus Frieler

## 1. Introduction

**Swing as a phrasing style**

“Swing” or “swing feel” is a phrasing style applied in jazz. Traditionally, when simply asked to play music, jazz musicians will “swing” it unless explicitly instructed to play it “straight” (Monson, 1996; Sussmann & Abene, 2012; Witmer & Robbins, 1988). Swing melody phrasing is typically played over a near-metronomic tactus beat sequence that emphasises the so-called “backbeats” (beats 2 and 4 in a quadruple-meter context) (Butterfield, 2006; Berliner, 1994; Monson, 1996). Against this steady beat, melodies and solos often feature delayed downbeat attacks (Busse, 2002; Friberg and Sundström, 2002; Wesolowski, 2012), allowing performers to play with a sense of rhythmic elasticity that makes it appear as if they play “laid back” or even “out of time” and yet synchronise with the beat at strategically important moments (Ashley, 2002; Benadon, 2006; Berliner, 1994; Butterfield, 2011; Monson, 1996, 82-83).

Yet, the most prominent aspect of swing phrasing is its characteristic uneven long-short tactus beat subdivisions: the so-called “swing eighths”. While jazz notation typically provides no indication for how uneven these long-short notes should be played, depicting them as even eight notes for quarter-note beats, non-jazz notation often displays them as quarter-eight triplets: . This reflects a wider common understanding of the swing eighth-note ratio as a constant quarter-eight triplet swing. This concept is expressed in empirical literature as a beat-upbeat ratio (from hereon: BUR) of 2:1.

**Swing BURs and tempo**

However, empirical analyses suggest that the steady 2:1 ratio is an oversimplification of what is a more varied and more musician-specific phenomenon. For example, drummers tend to play with much higher swing ratios than 2:1, demonstrating minimal variation within a performance (Collier & Collier, 1996; Friberg & Sundström, 2002; Prögler 1995; Honing & Haas, 2008; see Butterfield 2011, 5-7, for a comparative overview). This likely serves to highlight the downbeat crotchet tactus and so provides a strong and reliable reference grid for other musicians (Butterfield 2011; Monson 1996). Across performances, drummers decrease their BURs with increasing tempo in an approximately linear fashion, starting at around 3 to 3.5 for tempi of 100-150 BPM and levelling out at c. 1.0 at tempo 300 BPM (Friberg and Sundström, 1997 & 2002; Prögler, 1995; Honing, and de Haas, 2008 found a decrease, but no evidence for a linear interpretation). This scaling effect does not seem to hold for other band members, since a similar decrease in BURs with increasing tempo could not be found for melody or solo performances (Busse, 2002; Benadon, 2006; Ellis; 1991, Wesolowski, 2012).

In contrast to drummers, soloists tend to play with a lower swing ratio than 2:1 on average, although ratios vary strongly across different performers and even within individual performances or solos (Benadon, 2006; Busse, 2002; Butterfield, 2011; Ellis, 1991; Wesolowski, 2012). For example, Butterfield (2011) demonstrates that merely a four-bar phrase by Coleman Hawkins includes BURs ranging from 1.31:1 and 2.44:1, while a phrase of similar length by Charlie Parker ranges from 0.97:1 to 1.89:1. In order to nonetheless synchronise with the higher BURs of the drums, soloists tend to play their onbeat eighth note slightly after the drums have already articulated the onbeat note, then link with the drummers’ upbeat eighth note (Friberg and Sundström, 2002; Iyer, 2002).

In summary, it appears that swing is a situation-specific and complex phrasing style that features anything but a straightforward steady 2:1 BUR. The reasons for why the myth of swing as a 2:1 triplet pattern persists so widely are likely rooted in cognitive mechanisms that cause listeners to conceptualise uneven events as easily quantifiable ratios (Butterfield, 2010 & 2011; Iyer, 2002). Some authors have also suggested that certain BURs are associated with particular periods or subgenres in jazz history, with the 2:1-triplet conception possibly harking back to phrasing habits during the classic big band era in the 1930s and 40s (Sussman & Abene, 2012, 58; Benadon, 2006).

**Sampling problems**

However, all the empirical papers referenced here only report results of small case studies, basing their analyses on a limited number of phrases, recordings, performers, or participants (ethnographic studies and personal reports aside, which do not provide quantifiable data). Table 1 shows an overview of the data sources for each paper.

|  |  |
| --- | --- |
| **Paper** | **Sampled data sources** |
| Ashley (2002) | 5 recordings by 4 performers |
| Benadon (2006) | 26 recordings by 24 performers |
| Berliner (1994) | Ethnographic interviews and observations |
| Busse (2002) | 33 recordings by 3 performers |
| Butterfield (2010) | 3 experimental samples assessed by 22, 19, 22 participants |
| Butterfield (2011) | 5 recordings by 5 performers |
| Collier & Collier (1996) | 20 experimental performances by 3 performers each |
| Ellis (1991) | 15 experimental performances by 3 performers each |
| Friberg and Sundström (1997) | 6 recordings by 4 performers |
| Friberg and Sundström (2002) | 6 recordings by 4 ensembles (comparing 4 drummers with 6 soloists) |
| Honing & de Haas (2008) | 18 experimental performances by 3 performers each |
| Iyer (2002) | 2 recordings by 2 performers; personal expertise |
| Monson (1996) | Ethnographic interviews and observations |
| Prögler (1995) | 2-3 experimental performances by 6 performers each (with varying parameters for different performers) |
| Sussmann & Abene (2012) | Personal expertise |
| Wesolowski (2012) | 5 recordings by 1 performer |
| Witmer & Robbins (1988) | Personal expertise |

Table 1: Overview of sampled data sources for the cited papers

Therefore, while offering useful starting points for further research, the results generated by these papers have limited statistical power in elucidating the wider practice. An analysis of a wider corpus of performances is required for more precise and comprehensive data.

A useful recording collection for such purposes is the Weimar Jazz Database (WJD), a corpus of 456 monophonic jazz solos recorded by 78 canonical and lesser-known performers. Devised as part of the Jazzomat Research Project at the Hochschule für Musik Franz Liszt Weimar (University of Music Franz Liszt Weimar), these solos have been coded and segmented for a variety of analyses and have been put online together with customised data extraction tools to facilitate open research (see Pfleiderer et al., 2017 for a full description of features).

In this paper, we describe how we conducted a corpus analysis of the WJD in order to investigate some of the most discussed questions in the swing literature:

1. Do soloists distribute their eighth notes as triplets?
2. Do jazz soloists decrease their eighth-note ratios with tempo?
3. Does jazz style influence a soloists’ beat-upbeat ratios?

By taking into account a large number of recordings from across an extensive time period and performed by a variety of musicians, our analysis provides a more comprehensive investigation than what the literature offers to date. The results contribute with insights relevant to music practice and education in jazz, historical musicological research, and music cognition.

## 2. Method

### 2.1 Data

Pfleiderer et al. (2017) describe the Weimar Jazz Database in detail: It is a freely available high-quality database of monophonic jazz solos. It contains 456 solos by 78 different performers, recorded between 1925 and 2009, with a focus on the US-American jazz canon[[1]](#footnote-1). The WJD provides a wide range of metadata for each recording, for example recording year, style, average tempo, or rhythm feel. All solos were manually transcribed and annotated using Sonic Visualiser and the mel-peak spectrogram visualisation, allowing the transcribers to visually inspect and slow down the audio as much as needed. All onsets were cross-checked by a second or sometimes a third expert. As a result, all note events and all beats are marked with precise onsets and offsets, making micro-timing studies possible.

The WJD features several data extraction tools via a user interface called MeloSpyGUI. These tools use the so-called FlexQ algorithm to analyse the WJD’s annotated beats and note onsets (Frieler, 2017). The FlexQ algorithm tries to determine an optimal division of a beat for all events found within a beat interval (i.e. the time span between two annotated beats). It does so based on a loss function that tries to minimize the number of sub-beat divisions necessary to gain a maximally even and simple distribution of tone events within a beat interval. The algorithm takes into account a small forward tolerance window of 30 ms in order to determine which events belong to which beat interval. We used the MeloSpyGUI’s tools (details below) to extract data and analysed it using custom scripts for the statistics programme R v3.6.2.[[2]](#footnote-2)

### 2.2 Procedure

**Event selection**

Measuring and statistically evaluating BURs is a complex undertaking, since it depends on precise identification of note event onsets in the source music. Annotating such event onsets for research is often difficult to do with non-MIDI performances and instruments, and particularly when dealing with historical recordings of varying recording quality, as in the WJD’s case. Any errors in annotating onsets can strongly influence statistical results, since BURs—being ratios of inter-onset intervals (from hereon: IOIs)—are very sensitivity to outliers, particularly in the denominator when expressed as a fraction. Such measurement errors introduce additional variance to any variance already naturally present due to performed music’s inherent microrhythmic variation. Such variation is especially prevalent in music with stylised microrhythmic variation, for example in a wheresoloistsConsequently, when selecting viable rhythmic candidate events for specific research, it is important to take into account how the music was metrically annotated. Below we describe how we have accounted for such issues in our event selection process.

In order to make stable estimates for swing ratios in the WJD, we first used the database’s interface MeloSpyGUI to select all annotated beats that contain exactly two events, i.e. excluding all fully articulated triplet or semiquaver runs. From this sample, we selected all beats for which the sum of both IOIs are less than the local beat duration (plus a small offset of 30 ms, which is the FlexQ’s tolerance window), avoiding any second onsets that are sustained into the next beat. The reason for this is that calculating a BUR requires calculating the length of a downbeat and an upbeat subdivision of a beat. Consequently, one needs three event onsets (what we call an event ‘triple’) for estimating the length of two events: two events of interest subdividing the beat (i.e. a down- and upbeat) as well as a third event at the beginning of the next beat, which indicates when the second event is concluded. If t1, t2 are the onsets for the events of interest in the first beat and t3 is the onset of the third event at the beginning of the next beat, then we can calculate the two IOIs of interest: IOI1 = t2-t1, and IOI2 = t3-t2. We can then estimate the BUR as the ratio of the first to the second IOI: BUR = IOI1/IOI2. However, as fractions are not always easy to handle in statistical analyses, for analysis purposes we additionally converted BURs to binary logarithm of BURs (see below for details).

**Removing artefacts**

A problem resulting from the selection criteria outlined above was that, due to the selection mechanism of the FlexQ algorithm, sometimes the second inter-onset interval of a beat (IOI2) was very small. Such an event, possibly, should have been assigned to the following beat, making it a grace note or a very short attack at the beginning of the next beat. However, probably because of variance in the annotated beats and onsets, the algorithm selected them as belonging to the ongoing beat. In descriptive statistics, these very small second IOI’s contribute to very large BURs, because mathematically they form a BURs denominator. Though such events are likely artefacts, it is difficult to determine when to drop them from the analysis and when to keep them.

We based our selection on musical grounds: A dotted eighth note followed by a sixteenth note () has a BUR of 3:1. The short 16th-note upbeat portion will likely still be perceived as a subdivision of the ongoing beat. The next-higher dotted rhythm in the musical hierarchy (not counting tuplets) is a double-dotted eighth note followed by a 32nd note (), corresponding to a BUR of 7:1. Perceptually, the 32nd note is likely to be conceived of as a slightly early grace note or appoggiatura for the next beat rather than as a subdivision of the ongoing beat. Therefore, we opted to set our cut-off between these two BURs. Hence, we decided to select only those event triples with a BUR up to and including a maximum of 4:1, which corresponds to an upper bound of log2 4:1 = 2. For symmetry reasons, we also selected only short-longs BURs up to and including 1:4 (log2 BUR = –2).

**Error estimation**

Another issue was the estimation of actual measurement errors for onsets. To pre-empt this, we ran simulations with a range of idealized BURs against different tempi using estimated measurement errors of 10, 20, 30 and 40 ms. These showed that for fast tempi, i.e., where small beat durations were more likely, the estimation of BURs from inter-onset interval ratios can be considerably biased if one assumes that onset estimations are normally distributed around true onsets (i.e. with an over-/underestimation in low/high BURs of up to 30 %).

However, the WJD solos contain mostly legato notes, meaning that durations last until close to the onset of the next event. We assumed that it is unlikely for an annotator to place an onset during a sounding tone, meaning there are clear causal constraints on the onset annotation. Due to the careful preparation of the WJD using time-stretch tools and extensive cross-checking by experts (Pfleiderer et al., 2017), we assumed the measurement error of the onsets to be very low—likely in the range of 0–10 ms.[[3]](#footnote-3) We consulted the main expert who carried out the cross-checking and he agreed that this a reasonably estimation of the measurement error. Our simulations showed that a measurement error of 10 ms only has a negligible impact on BURs. Therefore, we decided to accept the original WJD annotations for onsets as suitable for all subsequent calculations.

**Estimating BURs**

In order to examine the overall distribution of the data, we chose not to work with BURs directly, but with the binary logarithm of BURs, which symmetrises and linearizes the distribution. As a result, the conversion followed the principle:

log2(BUR) = log2(IOI1/IOI2) = log2(IOI1) – log2(IOI2).

This way, perfect binary eighths represented by a BUR of 1:1 were mapped to the value 0, since log2(1) = 0. This also transformed any log-BUR of a short-long pattern into the negative value of the log-BUR of the corresponding long-short pattern; for example, a perfect triplet BUR of 2:1 is mapped to 1, whereas the inverse BUR 1:2 is mapped to –1. Arithmetic means of log-BUR values are then equivalent to geometrical means of BUR.

To further facilitate analysis, we grouped BURs into three types:

* The ‘snap’ type (short-long beat subdivisions), defined as log2(BUR) < –0.5 (i.e. BUR < 0.707)
* The ‘even’ type (even beat subdivisions), defined as -0.5 < log2(BUR) < 0.5 (i.e. 0.707 < BUR < 1.41)
* The ‘swung’ type (long-short beat subdivisions), defined as log2(BUR) > 0.5 (i.e. BUR > 1.41:1).

The thresholds for these groups are based on estimations, since they merely serve as illustrative delineations when discussing BUR values and therefore are not crucial for our findings here.

## 3. Results

### Global results

Overall, event triples (i.e., groups of three consecutive eighth notes that start on a beat) are noticeably rare in the WJD solos. We found 14,895 triples that fulfilled our selection criteria, meaning that only about 7.4% of all note events in the WJD form the first event in a swung triple. The number of triples also increases with tempo: Only 1% of notes in a solo in SLOW (<60 beats per minute) and MEDIUM SLOW (60–100 bpm) tempi categories together start a triple; in MEDIUM tempo (100–140 bpm) about 2.3%, in MEDIUM UP (140–180 bpm) 5% and in UP tempo (> 180 bpm) 12% do so. Furthermore, 21 out of 456 solos contained no triple, of which 15 were ballads (i.e. tempo class SLOW). Therefore, eighth note sequences that might be swung are more likely to be found at higher tempi.

The mean log2(BUR) value is 0.3 (BUR = 1.3:1) and the median value is 0.275 (BUR = 1.21:1). Therefore, the data’s distribution is not symmetric around 0. Consequently, we can conclude that jazz soloists play with slightly uneven rather than entirely ‘straight’ eighth notes, i.e., they exhibit a tendency to use long-short patterns.

However, in predominantly swung music, one would also expect a noticeable peak around the log-BUR value of 1 (BUR 2:1) when visually inspecting the data. Such a peak was not clearly visible. To investigate this issue further, we applied Gaussian mixture models[[4]](#footnote-4) (from hereon: GMM, package *mclust* v5.4.6 for R v3.6.2) to the log-BUR distribution, allowing the optimal numbers of clusters to be selected automatically. The 95%-CI intervals were calculated using the MClustBootstrap function with n = 999 bootstrap samples.

This revealed two components: the first component covers 53% of all log-BURs and has a mean value of 0.154 (BUR = 1.11:1, CI95(BUR) = [1.11, 1.12]). The second component is located around 0.51 (BUR = 1.42:1, CI95(BUR) = [1.40, 1.43]) and covers 47%. These are visible in Figure 1, which shows the distribution of all log-BURs over the whole WJD.

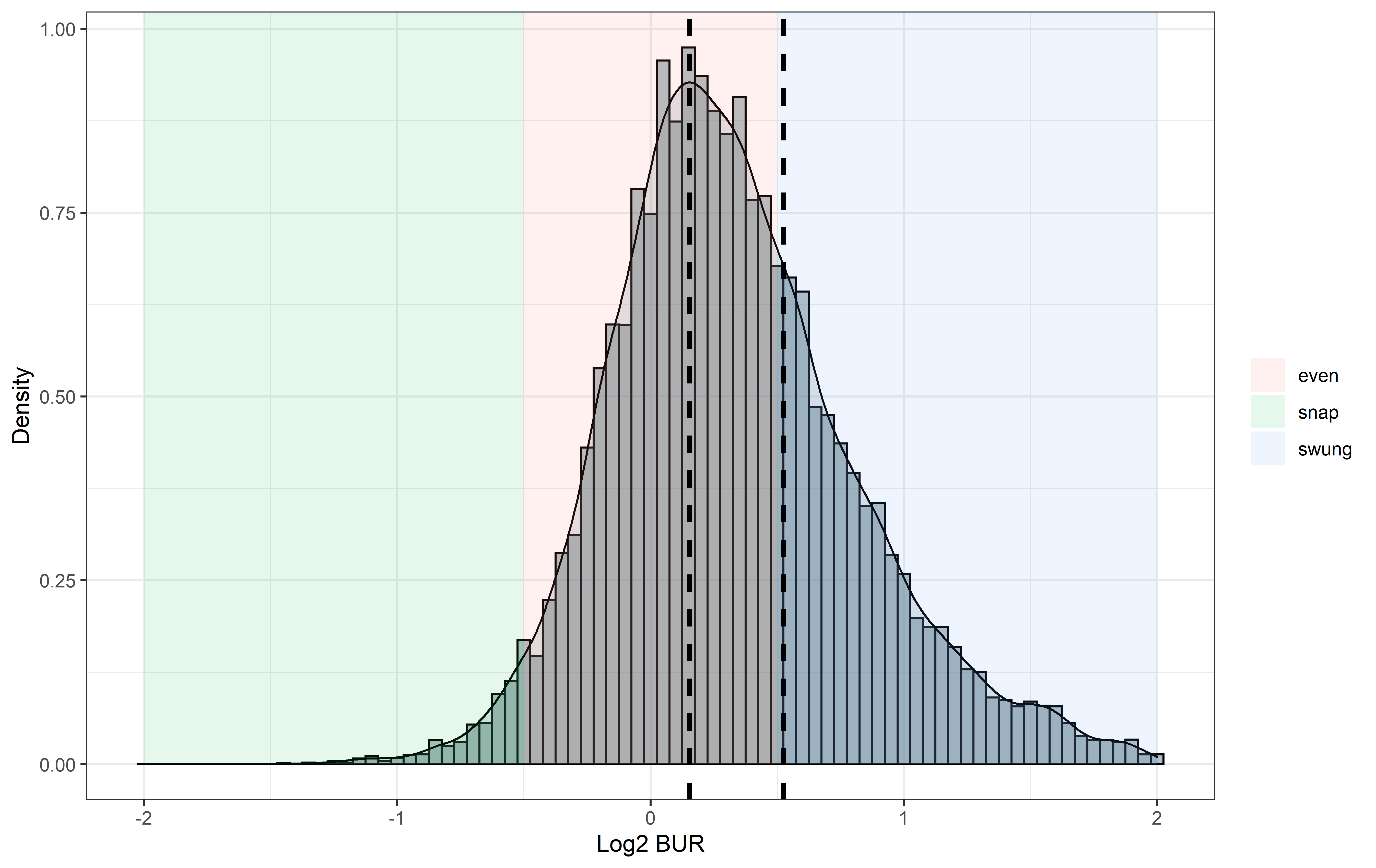


Figure 1: Histogram and density plot of log-BUR across the whole WJD. Vertical lines indicate the two GMM components identified for the optimal numbers of clusters. Shaded areas indicate BUR types: Blue: “swung”, red: “even”, green: “snap”.

As a result, the data suggests that swung eighths do occur in monophonic jazz solo improvisations, but are not the norm. Instead, binary divisions are more common than distinct long-short or short-long patterns. Overall, the distribution is right-skewed, with about twice as many long-short patterns compared to short-long patterns. On the whole, this expresses an overall preference among performers for near-equal, only slightly uneven long-short patterns.

This is, however, only the global view. Since the literature suggests possible differences between jazz styles, performers, as well as tempi, we decided to break the data down further.

### Swing analysis of single solos

The increasing frequency of triples with increasing tempo, reported above, suggested that not all solos feature or focus on eighth-note rhythms. In order to analyse only solos that make consistent use of eighth notes and have enough data to reliably apply a GMM, we calculated optimal GMMs for solos which had more than 15 event triples (*n* = 245). From these, 174 fit an optimal model with only one component (mean BUR: 1.22:1, range: 0.91–1.94, SD = 1.16). About one-fifth (56) exhibited two components (mean BUR of first cluster: 1.08, range: 0.41–1.55, SD = 1.24; mean BUR of second cluster: 2.20, range: 1.04–3.79, SD = 1.28). Only 16 solos exhibited three or more components. In total, 117 solos (48% of the solos in the selection, 25% of solos in the WJD) had a ‘swung’ BUR-type component with a mean BUR greater than 1.41:1.

Looking at raw counts of triple types, we found only 4,734 ‘swung’ triples out of 14,895 triples in all of the WJD (32%), whereas 9,726 (65 %) were of the ‘even’ and 435 (3 %) of the ‘snap’ typeThe average number of ‘swung’ triples per solo was 12.1 (SD = 13.9) per solo with a large variation, whereas the average number of ‘even’ triples was 24.7 (SD = 33.7). This corroborates our earlier observation that uneven eights are considerably less common than more even eighths in monophonic jazz solo improvisations, even in solos with comparatively many eighth-note sequences. Moreover, this suggests that uneven eighth notes are an expressive technique that soloists and melody performers only occasionally employ.

Table 1. Basic statistics of triples in the dataset.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Triple type | Count | Percentage | AM count/solo | SD count/solo |
| Even | 9,726 | 65.3 % | 24.7 | 33.7 |
| Swung | 4,734 | 31.8 % | 12.1 | 13.9 |
| Snap | 435 | 2.9 % | 2.8 | 3.0 |

### Effects of tempo on BURs

Next, we investigated relationships between BUR distributions and tempo. Figure 3 shows the mean log-BURs represented in their BUR types (‘even’, ‘snap’, swung’) plotted against the average tempo of the different solos.

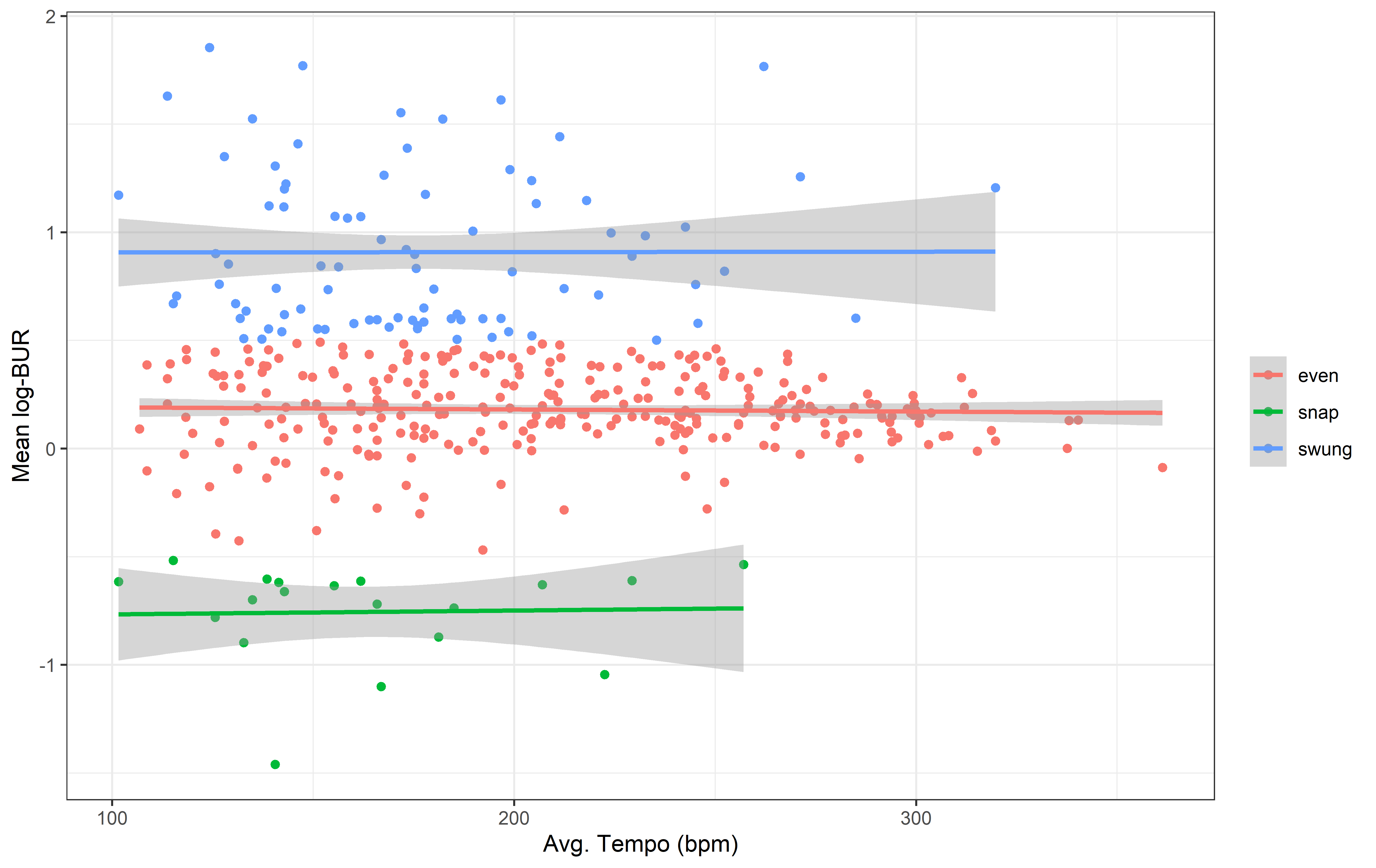


Figure 3: Mean log2-BUR of GMM components plotted against average tempo of solos. Only solos with 15 triples or more are included. Colours indicate BUR types: Blue: “swung”, red: “even”, green: “snap”. Grey shaded areas indicate 95% confidence interval of fitted regressions.

No clear trend is apparent in Figure 3’s graphs. Correspondingly, we found no significant effect when running a linear mixed regression model with log-BUR as dependent variable as well as a second-order polynomial in log-average tempo as fixed effect and BUR-type as random effect (F(2, 393.06) = 0.1083, p = 0.74). However, when plotting all solos’ log-BUR averages together against tempo, a slight compression effect became visible: BURs decrease with increasing tempo. This is shown in Figure 4.

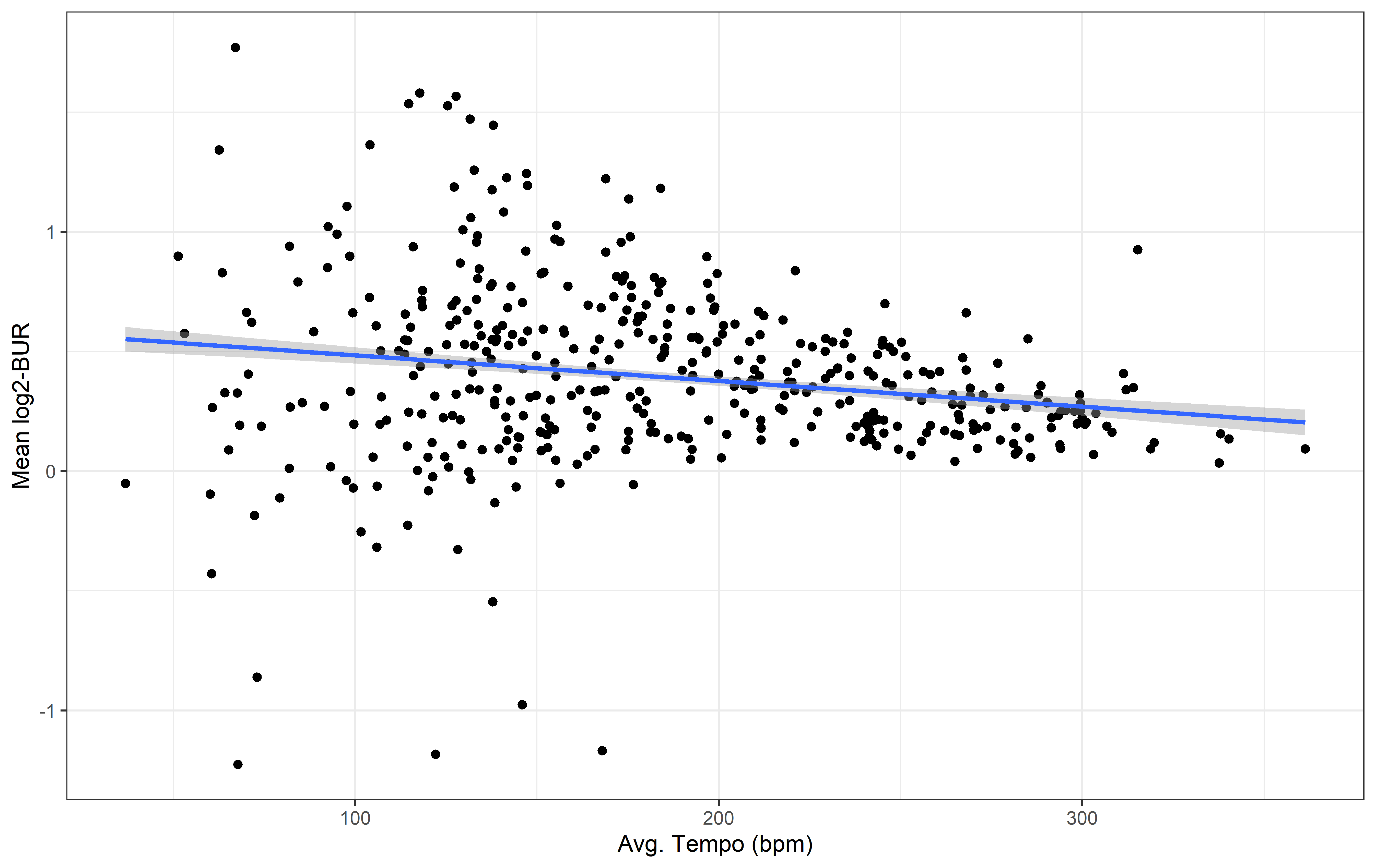


Figure 4: Mean log2-BUR of solos plotted against the average tempo. Linear trend in blue. Grey shaded area indicates 95% confidence interval of the fitted regression.

In order to investigate further, we ran a second-order polynomic linear mixed regression of all log-BURs against average tempo as fixed effect and solo as random effect. This yielded a significant but weak negative contribution of both polynomial coefficients (F(2, 330.97) = 26.34, p < .001, ω2= .11, 95% CI[.05, .17]). This suggests that performers on average are likely to play with lower BURs at higher tempi.

### Differences between jazz styles and rhythmic feels

We conducted several GMM analyses based on the WJD's metadata in order to explore BURs across different musical contexts. First, we used the so-called “rhythm feel” annotations. These were clustered into two groups: “swing feel” (containing feels “swing” and “twobeat”) and “non-swing feel” (containing “latin”, “funk”, and songs featuring more than one rhythm feel). Results are shown in Table 2.

Table 2. GMM components for different rhythmic feels.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Rhythmic Feel** | **Cluster Type** | **N Triples** | **Mean BUR** | **Proportion (%)** | **CI95(BUR)** |
| Swing Feel | Even | 13,884 | 1.11 | 52 | [1.11, 1.13] |
|  | Swung |  | 1.43 | 48 | [1.41, 1.45] |
| Non-Swing | Even | 1,011 | 1.07 | 57 | [1.05, 1.10] |
|  | Swung |  | 1.34 | 43 | [1.28, 1.43] |

As Table 2 shows, “swing feel” exhibited two components, one each from the even (mean BUR = 1.11:1) and swung type (1.43:1). “Non-swing feel” exhibited two even components (1.07:1 and 1.34:1). Surprisingly, the data indicates that rhythmic feels that are traditionally considered non-swing feels still contain a large proportion of swung triples—just not as much as swing feels. Consequently soloists are likely to swing over backing music even when this backing music is not traditionally considered swinging. Therefore, it is likely that solo swinging and backing swinging are two separate processes.

In order to further investigate in which jazz styles swing eighths are most used, we pooled all triples and then divided the data by the WJD’s style annotations (“TRADITIONAL”, “SWING”, “BEBOP”, “HARDBOP”, “FREE”, and “FUSION”, “POSTBOP”, see Pfleiderer (2017) for details). Then we calculated GMM for each style. Table 3 shows the results for each style.

As Table 3 shows, only solos in the categories BEBOP, HARDBOP, COOL and POSTBOP featured a BUR component that falls into our “swing” cluster type. However, the “even” cluster types in TRADITIONAL, SWING and FUSION still features BURS noticeably above 1.0 and according to their 95% confidence intervals could also fall into our “swing” cluster type. This means that all the mentioned styles feature a notably uneven BUR cluster that may induce a sense of swing, even though none of their clusters approach a BUR of 2:1. The only two styles to exhibit a BUR cluster of 2:1 or higher are HARDBOP and POSTBOP, though here the swung cluster covers less than a sixth of all examined triples (14% and 11% respectively). Therefore, in line with our already presented data, these results suggest that BURs of 2:1 are relatively rare among jazz soloists, even when taking stylistic differences into account. Furthermore, ‘swung’ clusters seem to come in two distinct types, one with a mean closer to 4:3 or 3:2 (TRADITIONAL, SWING, BEBOP, COOL) and one closer to 2:1 (in HARDBOP and POSTBOP), which might be called ‘soft’ and ‘hard’ swinging, respectively. Furthermore, the mean BURs of these swung clusters are monotonically increasing with time, from 1:37:1 in TRADTIONAL to 2.14:1 in POSTBOP.

Table 3. GMM components for different jazz styles and rhythmic feels. Note that the category FREE contains only solos by Ornette Coleman.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Style** | **Cluster Type** | **N Triples** | **Mean BUR** | **Proportion (%)** | **CI95(BUR)** |
| TRADITIONAL | Even | 287 | 1.37 | 100 | [1.33, 1.42] |
| SWING | Even | 1,064 | 1.41 | 100 | [1.38, 1.44] |
| BEBOP | Even | 1,510 | 1.17 | 49 | [1.13, 1.20] |
|  | Swung |  | 1.44 | 51 | [1.38, 1.49] |
| COOL | Even | 2,303 | 1.13 | 50 | [1.11, 1.15] |
|  | Swung |  | 1.47 | 50 | [1.42, 1.54] |
| HARDBOP | Even | 2,825 | 1.12 | 86 | [1.11, 1.14] |
|  | Swung |  | 2.02 | 14 | [1.95, 2.15] |
| FREE | Even | 208 | 1.29 | 100 | [1.23, 1.35] |
| FUSION | Even (low) | 200 | 1.02 | 46 | [0.98, 1.05] |
|  | Even (high) |  | 1.31 | 54 | [1.19, 1.49] |
| POSTBOP | Even | 6,498 | 1.13 | 89 | [1.12, 1.14] |
|  | Swung |  | 2.14 | 11 | [2.09, 2.23] |

However, there is the possibility that individual performers consistently swing across different jazz styles while others do not. Swing may be a result of individual performers’ behavioural patterns rather than of jazz subgenre. Consequently we decided to investigate individual performers’ swing.

### Differences between performers

When pooling GMM component by individual performers, we found noticeable differences. Figure 5 shows the data for performers who are represented with at least five solos in the WJD.

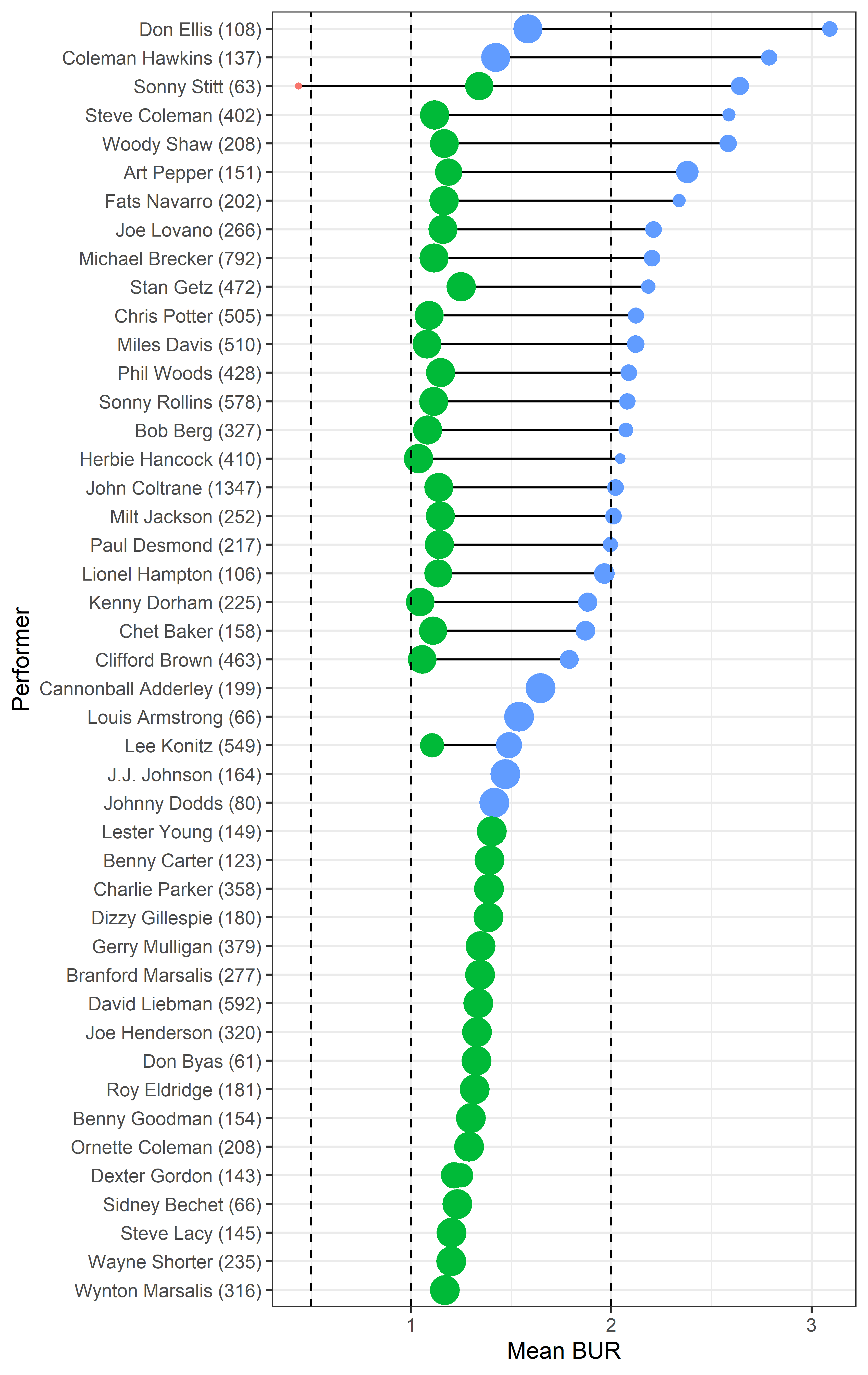


Figure 5. Pooled GMM components for performers with at least 5 solos in the WJD. The number in brackets following a performer’s name indicates the number of examined swing triples. Point colours indicate BUR types: Blue: “swung”, green: “even”, red: “snap”. Dashed lines represent BURs of 1:2, 1:1, 2:1, 3:1. Point size is proportional to probability mass of the components.

As Figure 5 shows, for many performers we only found an ‘even’-type component, with only few performers exhibiting a low ‘swung’-type component instead of an ‘even’ one. However, for a number of performers we also found a secondary ‘swung’ type component in addition to a primary ‘even’ type. Noticeably, in the majority of cases where we found such a secondary ‘swing’ type component, its BUR approached or exceeded 2:1. This suggests that those among selected performers who tend to divide their eighth-note phrasing into ‘even’ and ‘swung’ types, tend to play the swung eighths particularly uneven—possibly to enhance the contrast with even rhythms. For these musicians, the stereotypical 2:1 conception of a BUR seems to hold. For some it may even be an underestimation, since we found a small number of musicians play with BURs approaching or even surpassing a BUR of 3:1 (Woody Shaw, Steve Coleman, Sonny Stitt, Coleman Hawkins, Don Ellis). However, on the whole the majority of selected performers tend to play primarily slightly uneven, nearly straight eighths. This again supports our earlier observation that a BUR of 2:1 is not the norm among jazz soloists.

## 4. Discussion

In this research paper, we presented results from a corpus analysis using the Weimar Jazz Database to investigate swing rhythms in monophonic jazz solos. Overall, our results show that jazz soloists tend to play with nearly even long-short beat divisions and that the famed 2:1 BUR is only employed by comparably few performers. This confirms the observation of earlier case studies (Benadon, 2006; Butterfield, 2011; Ellis, 1991) that jazz soloists tend to play considerably more even notes than the traditionally assumed 2:1 BUR for swing.

**High BURs as a stylistic tool**

Even in jazz styles traditionally associated with swing, BURs tended to be slightly uneven (between 4:3 and 3:2), indicating that high BURs are used sparingly—also in styles that are traditionally considered to have a “swing feel”. This indicates that swing feel and uneven beat subdivisions in a solo context are not necessarily associated. While most performers simply play with near-even BURs with only slightly longer first eighths, only a select few soloists also frequently play with very uneven BURs. The literature indicates that jazz players likely use high BURs to temporarily heighten the sense of forward momentum associated with swing (Butterfield, 2011), using it as an additional parameter to note choice, articulation, and dynamics by which to control the expressive properties of a solo (Butterfield, 2011; Sussman & Abene, 2012, 58; Williams, 1993). Consequently, our data here suggests that high BURs seem to be an expressive tool reserved only for specific occurrences, and only few performers choose to use this tool less sparingly.

The literature offers several indications for when high BURs might be used. Case studies suggest that soloists increase their BURs to lock in with the underlying drum pattern at structurally important moments, such as phrase endings (Ashley, 2002). This would explain the relative rarity of swing BURs found here and why performers were shown to play more ‘straight’ for the most part. Since increasing BURs in this way may highlight perceptual awareness of the ongoing beat sequence (Benadon, 2006) and increase a sense of forward momentum (Butterfield, 2011), this technique is likely to increase in listeners the overall sensation of ‘groove’—defined by Janata, Tonic, and Haberman (2012) as the quality that causes a pleasurable urge to move along to an ongoing beat sequence. As a result, performers may use increased BURs strategically to engage their audiences’ locomotive senses and increase their physical arousal in order to gain their attention when phrase units chance—much like the way voice modulation is used at the end of sentences in spoken language. Confirming this would require future studies on swing syntax.

**The effect of tempo**

Perceptual factors may also be involved in why soloists and drummers reduce their BURs at high tempi. While this effect is likely primarily due to physical limitations (i.e., increasing motoric difficulties in articulating short upbeat eights as tempo increases), there may be an additional psychological component: At high tempi, BUR variations may be more difficult to perceive than at slow tempos (Benadon, 2006). More concretely, we propose that the upbeat portion of two very uneven swing eights is likely to fall below the perceptual limit for rhythmization, which London sets at c. 100 ms (2012, 27). Consequently, high BURs at high tempi are likely less efficient at facilitating pulse perception than equal tactus beat subdivisions. This dovetails with Friberg and Sundström (1997) findings showing that drummer’s upbeats to stabilise at a length no shorter than 100 ms above c. 150 BPM (although it should be noted that Honing & de Haas (2008) found no such floor limit).

Therefore, playing more even BURs may allow performers to keep upbeats long enough to stay within human limits for both rhythm perception and synchronisation at fast tempi. Since beat subdivisions at slow and medium tempi lie more comfortably between human perceptual limits, they possibly afford listeners and performers both more capacity (and indeed a preference for) greater rhythmic variety in swing. This may be associated with findings showing that for listeners syncopation and groove form a relationship approximating an inverted U-shape, meaning that too little and too much syncopation is both detrimental to groove perception in the long run (Witek et al., 2014). Maintaining optimal groove within stylistic limits is therefore another likely reason why jazz soloists keep their eighth notes even and only play high BURs occasionally.

**The perceptual myth of a 2:1 BUR**

What is clear form our data is that a constant 2:1 BUR in swing rhythms among melody instruments is a myth. However, if varying BURs are more likely to arouse audiences than even ones, where does the idea of a stable 2:1 BUR stem from and why is it so persistent? Butterfield applies findings by Povel, which suggest that listeners have a tendency to reduce durationally uneven ratios between 1:4 and 4:5 to a simpler 1:2. Butterfield suggests that 'subjective interpretation of swing ratios tends toward the triplet model simply because any succession of unequal durations tends to sound to us like triplets’ (2011, 3-4). This corresponds with London’s (2012, 35-36) suggestion that two evenly spaced notes which are moved further apart will gradually produce a sense of triplet, causing listeners to divide the two notes naturally into simplistic long and short categories. Consequently, the assumed 2:1 BUR, being the simplest rational long-short ratio, is likely a reductionist effect of entrainment that guides pattern recognition and corresponding expectations towards salient beat occurrences.

This perceptual effect may arise from various musical factors. As our data showed, while soloists' overall BURs reduce slightly with increasing tempo, their eighth-note sequences increase. Research has shown that drummers also reduce their BURs as tempo increases, though to a much stronger degree (Collier & Collier, 1996; Friberg & Sundström, 2002; Honing & de Haas, 2008). This suggests that soloists adjust their beat divisions according to the underlying beat subdivision articulated by the drums. Butterfield (2010) showed that the difference in attacks between drums and bass in jazz can only be perceived subliminally, since spacing them too widely will make them feel like performance errors. This indicates that the first instrument to attack defines the beginning of an event duration, even if another instrument only attacks shortly after. Therefore, we propose that when solo performers play a delayed onbeat eighth after the quasi-metronomic drum attack, listeners associate this event as a continuation of an ongoing event initiated by the drums' earlier attack. If the drums are perceived as initiating a soloist’s first eighth note, even though the soloist attacks later, a sense of 2:1 may still emerge from the solo. Since our data showed that solos in styles associated traditionally with a swing feel (SWING and TRADTIONAL) feature lower swing BURs on average than the historically later and less swing-focussed styles BEBOP, COOL, HARDBOP, or POSTBOP, it is possible that a 2:1 BUR swing feel is more strongly associated with drumming than with melodic rhythm patterns. However, in this study, we were unable to take into account how melody lines sit in relation to the backing band, and consequently these thoughts remain speculative.

It is interesting that swung BURs in all styles but POSTBOP are closer to 3:2 or 4:3 than to 2:1, given that recent studies suggest that a preference for small integer ratios in the production of rhythms may be universal (Jacoby & McDermott, 2017; Roeske et al., 2020). However, consistent deviation from small integer ratios is regularly exhibited in jazz as well as in other music cultures, e.g., in Mande drumming in Mali (Polak & London, 2014) or in Urugayan Candombe (Fuentes et al., 2019; Jure & Rocamora, 2016). These findings pose a challenge to theories of regular meter theories (e.g., Lerdahl & Jackendoff, 1983) and to oscillator models of rhythms (Bose et al., 2019; Large et al., 2015).

One last point concerns the dimension of intensity in assessing microrhythmic deviation from idealised norms. Typically, psychological and musicological timing research operates with idealized and measured sharp onset points of events. However, it also well known that the perception of an event onset (perceptual onset time) is related to the perceptual attack time (PAT) of the actual sounding event (Gordon, 1987). Whereas in experiments involving percussive events (drums, tapping), which have rather steep rise times, this approximation seems to be justified, in our case of monophonic (mostly wind) instruments there is a potential divide between the ‘production time’ of the performer and the perceived onset time in a listener. In addition, the intensity and other spectral characteristics of the tone events might play a role, as these can the influence the PAT and therefore the (perceived) BUR. For instance, in a study on the PAT of saxophone tones, Bechtold & Senn (2018) found that louder tones have an earlier PAT whereas the PAT of soft tones was influenced by the articulation (with/without tongue attack). Therefore, future research into jazz microrhythms should also take into account the influence of intensity and articulation on the perception and production of BURs. First steps in this direction were already pursued in a study by Abeßer et al. (2014), which, however, found no consistent effects for intensity differences between first to second eighths notes.

## 5. Conclusion

In conclusion, our corpus analysis indicates that jazz soloists tend not to play with uneven swing BURs and that the famed 2:1 BUR for swing eighths is the exception rather than the norm in solos. Instead, it is likely that soloists use varying BURs as an expressive technique to play with audiences’ perceptual mechanisms. This in turn is likely one of swing phrasing’s engaging qualities. The corpus analysis also strongly indicated that jazz soloists very slightly decrease their BURs with tempo, which they possibly do to maximise groove perception among audiences. This may suggest that perceptual limits are involved in how much microrhythmic deviation is deemed stylistically permissible at different tempi. Therefore, our results are relevant to research in microrhythmic behaviour, jazz performance and history, and groove perception.

## Bibliography

Abeßer, J., Cano, E., Frieler, K., & Pfleiderer, M. (2014). Dynamics in Jazz Improvisation—Score-informed Estimation and Contextual Analysis of Tone Intensities in Trumpet and Saxophone Solos. Proceedings of the 9th Conference on Interdisciplinary Musicology (CIM).

Ashley, R. (2002). Don't Change a Hair for Me: The Art of Jazz Rubato. *Music Perception*, *19*(3), 311-332. doi: 10.1525/mp.2002.19.3.311

Bechtold, T. A., & Senn, O. (2018). Articulation and Dynamics Influence the Perceptual Attack Time of Saxophone Sounds. Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.01692

Benadon, F. (2006). Slicing the Beat: Jazz Eighth-Notes as Expressive Microrhythm. *Ethnomusicology,* *50*(1), 73-98. ISSN: 00141836

Berliner, P. (1994). *Thinking in Jazz: The Infinite Art of Improvisation.* University of Chicago Press.

Bose, A., Byrne, Á., & Rinzel, J. (2019). A neuromechanistic model for rhythmic beat generation. PLOS Computational Biology, 15(5), e1006450. <https://doi.org/10.1371/journal.pcbi.1006450>

Busse, W. G. (2002). Toward Objective Measurement and Evaluation of Jazz Piano Performance Via MIDI-Based Groove Quantize Templates. *Music Perception: An Interdisciplinary Journal*, *19*(3), 443-461. doi: 10.1525/mp.2002.19.3.443

Butterfield, M W. (2006). The Power of Anacrusis: Engendered Feeling in Groove-Based Musics. *Music Theory Online,* *12*(4). doi: http://dx.doi.org/10.30535/mto.12.4.2

Butterfield, M. (2010). Participatory Discrepancies and the Perception of Beats in Jazz. *Music Perception: An Interdisciplinary Journal*, *27*(3), 157-176. doi: 10.1525/mp.2010.27.3.157

Butterfield, M. W. (2011). Why Do Jazz Musicians Swing Their Eighth Notes?. *Music Theory Spectrum*, *33*(1), 3-26. doi: 10.1525/mts.2011.33.1.3

Collier, G. L., & Collier, J. L. (1996). The Swing Rhythm in Jazz. In B. Pennycook & E. Costa-Giomi (Eds.), *Proceedings of the International Conference of Music Perception and Cognition* (pp. 477-80). Montreal: McGill University.

Ellis, M.C. (1991). An analysis of “swing” subdivision and asynchronization in three jazz saxophonist. *Perceptual and Motor Skills*, 73, 707-713. Retrieved from https://doi.org/10.2466/pms.1991.73.3.707

Friberg A., & Sundström, A. (1997). Preferred swing ratio in jazz as a function of tempo. *Tmh-Qpsr (Speech, Music, and Hearing: Quarterly Progress and Status Report)*, *38*(4), 19–27. Retrieved from <http://kth.diva-portal.org/smash/get/diva2:1246291/DATASET01.pdf>

Friberg, A., & Sundström, A. (2002). Swing Ratios and Ensemble Timing in Jazz Performance: Evidence for a Common Rhythmic Pattern. *Music Perception,* *19*(3), 333–349. doi: 10.1525/mp.2002.19.3.333

Frieler, K- (2017). [The FlexQ algorithm](http://www.mu-on.org/frieler/docs/frieler_flexq_2017.pdf). In: M. Pfleiderer, K. Frieler, J. Abeßer, W.-G. Zaddach, & B. Burkhard (Eds.): *Inside the Jazzomat. New Perspectives for Jazz Research* (pp. 319-322). Mainz: Schott Campus.

Gordon, J. W. (1987). The perceptual attack time of musical tones. The Journal of the Acoustical Society of America, 82(1), 88–105. <https://doi.org/10.1121/1.395441>

Honing, H., De Haas, W.B. (2008). Swing once more: Relating timing and tempo in expert jazz drumming. *Music Perception*, 25(5), 471-476. doi: 10.1525/mp.2008.25.5.471

Iyer, V. (2002). Embodied Mind, Situated Cognition, and Expressive Microtiming in African-American Music. *Music Perception: An Interdisciplinary Journal, 19*(3), 387-414. doi: 10.1525/mp.2002.19.3.387

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General, 141*(1), 54-75. doi: 10.1037/a0024208

Jacoby, N., & McDermott, J. H. (2017). Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction. Current Biology, 27(3), 359–370. https://doi.org/10.1016/j.cub.2016.12.031

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural Networks for Beat Perception in Musical Rhythm. Frontiers in Systems Neuroscience, 9. <https://doi.org/10.3389/fnsys.2015.00159>

Lerdahl, F., & Jackendoff, R. S. (1983). A generative theory of tonal music. MIT Press.

London, J. (2012). *Hearing in Time: Psychological Aspects of Musical Meter.* Oxford University Press.

Monson, I. (1996). *Saying something: Jazz improvisation and interaction.* University of Chicago Press

Pfleiderer, M. (2017.) The Weimar Jazz Database.In M. Pfleiderer, K. Frieler, J. Abeßer, W.-G. Zaddach, & B. Burkhart (Eds.), *Inside the Jazzomat: New Perspectives for Jazz Research* (pp. 19-40). Schott Music GmbH & Co. KG.

Polak, R., & London, J. (2014). Timing and Meter in Mande Drumming from Mali. *Music Theory Online*, *20*(1). <https://mtosmt.org/issues/mto.14.20.1/mto.14.20.1.polak-london.html>

Prögler, J.A. (1995). Searching for Swing: Participatory Discrepancies in the Jazz Rhythm Section. *Ethnomusicology,* *39*(1), 21-54. doi: 10.2307/852199

Ravignani, A., Thompson, B., Lumaca, M., & Grube, M. (2018). Why Do Durations in Musical Rhythms Conform to Small Integer Ratios? Frontiers in Computational Neuroscience, 12. <https://doi.org/10.3389/fncom.2018.00086>

Roeske, T. C., Tchernichovski, O., Poeppel, D., & Jacoby, N. (2020). Categorical Rhythms Are Shared between Songbirds and Humans. Current Biology. https://doi.org/10.1016/j.cub.2020.06.072

Sussman, R., & Abene, M. (2012). *Jazz Composing and Arranging in the Digital Age.* Oxford University Press.

Wesolowski, B. C. (2012). *Testing a Model of Jazz Rhythm: Validating a Microstructural Swing Paradigm* (Doctoral thesis). Retrieved from <https://scholarlyrepository.miami.edu/cgi/viewcontent.cgi?article=1747&context=oa_dissertations>

Williams, M. (1993). *The Jazz Tradition.* Oxford University Press.

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, Body-Movement and Pleasure in Groove Music. *PLoS ONE*, 9. Retrieved from <http://dx.doi.org/10.1371/journal.pone.0094446>

Witmer, R., & Robbins, J. (1988). A Historical and Critical Survey of Recent Pedagogical Materials for the Teaching and Learning of Jazz. In *Bulletin of the Council for Research in Music Education,* *96,* 7-29.

1. A complete list of solos and links to solo vignettes with basic metadata and stats can be found here <https://jazzomat.hfm-weimar.de/dbformat/dbcontent.html>. [↑](#footnote-ref-1)
2. Data and analysis scripts can be accessed at <https://github.com/klausfrieler/swing_ratios>. [↑](#footnote-ref-2)
3. BURs are not affected by constant annotation errors in onset time, such as for example perceptual onset times (PAT), since BURs are only based on onset differences. [↑](#footnote-ref-3)
4. GMMs try to model a probability (frequency) distribution as a weighted sum of single Gaussian distributions with distinct mean and standard deviation value. The weights can then interpreted as the probability for a value to be ‘produced’ by a single component. [↑](#footnote-ref-4)